Hydrologic model forcing over the Himalaya

Bikas Chandra Bhattarai

Dissertation submitted for the degree of Philosophiae Doctor (PhD)

Department of Geosciences  
Faculty of Mathematics and Natural Sciences  
University of Oslo  

Oslo, Norway  
June 2020
Abstract

The Himalayas are the source for ten of Asia’s biggest rivers, which bring drinking water, power and irrigation directly to millions of people living downstream. The Himalayan river system contribute to the economy of the region. Hydropower plays a critical role in improving the economy and quality of life for the people living in the Himalayan countries like Nepal and Bhutan. Recognizing the importance of the Himalaya, there is a great demand for up-to-date information on various climate data to support planning for hydropower production, water management, and to forecast extreme weather events.

Despite the present level of technological advancement, hydro-meteorological observation in the Himalayan regions are minimal because of the adverse geographical conditions; hence the region is characterized as data scarce. To overcome these challenges, remote sensing and climate models offer a regional and global perspective on many atmospheric climatic variables. However, there are some limitations in remote sensing to measure some atmospheric properties such as aerosol optical depth (AOD) over the brighter surface. In previous works, although some hydrological modelling studies were conducted using remote sensing and reanalysis data, it has been given little attention to assess and improve the data quality for hydrologic simulation. I aim to assess and improve the available hydrologic model, forcing data set for the hydrologic simulation over the Himalayan region.

The thesis focuses on three aspects which are presented in three papers. In the first paper, I have proposed a novel approach for estimating AOD over the cryosphere portion of the Himalaya. To date, no such a model can be used to predict aerosol optical properties over the Himalayan region, particularly over the brighter surface such as snow cover area. In the research, empirical model was developed using multiple regression method. Model prediction was carried out using the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-interim) derived variables (i.e. relative humidity (RH) and U and V wind components (zonal and meridional wind components)). Model simulation validation with AErosol RObotic NETwork (AERONET) and the Moderate
Resolution Imaging Spectroradiometer (MODIS) observation shows the adequacy of the used model. The result presented in the paper-I supports that the use of an empirical model yields good estimation for daily average AOD in Nepal and could be used in other mountain regions for climate research.

In the second paper, I have evaluated different available hydrological forcing data sets for the hydropower inflow simulation in Nepal. The hydrological simulation was performed using the available hydrological modelling stack from the Statkraft’s hydrological modelling toolbox (Shyft). Standard model forcing data for the used model were temperature (T), precipitation (P), relative humidity (RH), wind speed (WS), and incoming shortwave radiation (S). Two global reanalysis data sets, i.e. ERA-Interim and the WATCH Forcing Data methodology applied to ERA-Interim data (WFDEI), and climate model data (CORDEX (Coordinated Regional Downscaling Experiment)) was used for the model simulation. Station observed temperature and precipitation data were available over the study region, thus, a hybrid data set was formed by combining observation with WFDEI forcing data sets (i.e. observed T, P and WFDEI RH, WS, and S). Reanalysis forcings are ERA-I and WFDEI, while climate model data were taken from CORDEX. We found significant variations in the precipitation from different forcing data sets; hence, the precipitation correction factor as a calibrating parameter was used during the hydrological model calibration.

Results show that used hydrological model (i.e. PT_GS_K) with all forcing data sets performed very well for the hydrological simulation in Nepal. Model performance in terms of Nash-Sutcliffe efficiency (NSE) during the model calibration and validation were found higher for the hybrid datasets (Observed+WFDEI). Similarly, the second-highest performance was observed for the WFDEI. Comparatively model performance for CORDEX forcing datasets was found lower, but the simulation was able to capture observed peak. This study shows that the global reanalysis and climate model datasets hold great potential for understanding the hydrology of the data-sparse Himalayan region. However, for the hydrological simulation in the ungauged catchments, CORDEX and ERA-I datasets should be bias-corrected.

In the third paper, I have evaluated three spatial catchment discretization methods, i.e. hypsography (HYP), regular square grid (SqGrid), and triangulated irregular network (TIN) were evaluated in a steep and glacierized Marshyangdi-2 river catchment, central
Abstract

Himalaya, Nepal. To assess the impact of radiation on the model response, we translated shortwave radiation using two approaches, one with the measured solar radiation assuming a horizontal surface, and another with a translation to slopes. The results show that the catchment discretization has a significant impact on simulation results. Evaluation of the simulated streamflow value using NSE and log-transformed NSE (LnNSE) shows that highest model performance was obtained when using TIN followed by HYP (during the high flow condition) and SqGrid (during the low flow condition). A similar order of precedence in relative model performance was obtained both during the calibration and validation periods. A notable difference between snow simulation from two radiation approaches was observed. However, the impact of translated radiation into hydrological model efficiency was not found significant.

The results presented in the thesis provide valuable new regional knowledge and contribution in the Himalayan region. My thesis opens new possibilities for the application of newly developed aerosol model and global data sets in water resources studies in the Himalayan region. With including values of TIN and and the bias corrected WFDEI dataset are recommended as the best performing methods for future water resource-related model application in the Himalayan region. Further implementation of newly developed AOD model in the Shyft, which is a community modelling framework within hydrology, will again improve our current understanding of the Himalayan hydrology.
Acknowledgments

I have great pleasure in expressing my deep sense of gratitude to my supervisor John F. Burkhart for giving me the opportunity to be a part of his research group. Thanks for all your support, for being available for discussion, and teaching me how to write scientific papers. Your guidance helped me in all the time of research and writing of this thesis.

I express my sincere thanks to my co-supervisor Frode Stordal for his excellent guidance and support during my PhD.

My sincere thanks are to Profs. Chong-Yu Xu and Lena Merete Tallaksen for always listening and giving me words of encouragement and insightful comments.

For IT support, I would like to thank Hans Peter Verne and Anne Fouillous. A big thanks also goes to Olag Silantyeva and Felix Matt for assisting me with the setting up models.

I extend my gratitude to the Department of Hydrology and Meteorology (DHM), Government of Nepal (GoN) for their cooperation.

I am extremely grateful to my parents for their love, caring and sacrifices for educating and preparing me for my future. I am very much thankful to my wife and my son for their love, understanding, prayers and continuing support to complete PhD. Also, I express my thanks to my brother and sisters for their support and encouragement.

At last but not least gratitude got to all of my colleagues at the Department of Geosciences who directly or indirectly helped me to complete my PhD.

Thanking You
Bikas Chandra Bhattarai
Contents

Abstract i

Acknowledgments iii

I Overview 1

1 Introduction 2

1.1 Motivation 4

1.2 Goal and Objectives 6

1.3 Scope 7

1.4 Thesis outline 7

2 Background 9

2.1 Modelling geophysical processes 9

2.2 Aerosol and climate 10

2.2.1 Aerosol: properties, observation, and modelling 12

2.3 Climate of the Himalayan region 13

2.4 Hydrological modelling in the Himalayas 17

3 Hydrological model and study area 21

3.1 Hydrological model 21

3.1.1 Catchment response function: Kirchner (K) 22

3.1.2 Potential evapotranspiration: Priestley-Taylor (PT) 23

3.1.3 Snow distribution and melt: Gamma snow (GS) 24

3.2 Study area 26

3.2.1 Geographical location, demography, and climate 26

3.3 Model forcings and validation datasets 30

3.3.1 Model forcings 30

v
# Contents

4 **Overview of the papers**

4.1 Summary of paper 1 ........................................ 41

4.2 Summary of Paper 2 ........................................ 43

4.3 Summary of Paper 3 ........................................ 45

5 **Discussion** ........................................ 47

5.1 General discussion ........................................ 47

5.1.1 Aerosol optical depth (AOD) estimation over the cryospheric portion of Nepalese Himalaya ........................................ 47

5.2 Challenges and opportunities for the use of remote-sensing and global forcing data over the Himalayan region ........................................ 49

5.3 Catchment discretization, model forcing, and hydrological response ........................................ 51

6 **Conclusion and recommendations for further works** ........................................ 53

6.1 Conclusion ........................................ 53

6.2 Recommendations for further works ........................................ 55

References ........................................ 59

II **Journal Publications** ........................................ 75

Paper I: Aerosol Optical Depth Over the Nepalese Cryosphere Derived From an Empirical Model ........................................ 77

Paper II: Evaluation of global forcing data sets for hydropower inflow simulation in Nepal ........................................ 95

Paper III: Implementation of Triangulated Irregular Network (TIN) on shyft and its effects on model response: A case study from Central Himalayan catchment ........................................ 121

III **Appendices** ........................................ 151

A **Observed Station** ........................................ 152

A.1 Temperature observation stations ........................................ 152

A.2 Precipitation observation stations ........................................ 153
Part I

Overview
1 Introduction

The Himalayan region is the primary source for freshwater supply to the millions of people living downstream (Immerzeel et al., 2015), and their livelihoods are directly linked to the Himalayas originated rivers. The supply and quality of water in this region is under extreme threat, both from the natural and anthropogenic processes, and climate change related variations (Ashraf, 2013). A rapid expansion of agriculture and increase in the region’s population are massively pressuring regional water resources, affecting them in terms of water-use patterns and management practices (Gupta et al., 2019). Change in the climatic condition in the Himalayan region may lead to a change in runoff volume, and in the seasonality of flow, and is expected to affect regional water demands. Water resource management and its development have been important measures in the Himalayan regions in such forms as hydropower development, irrigation project, domestic and industrial water supply and river navigation (Shekhar et al., 2010).

A large area in the Himalayan region is covered by seasonal snow and glaciers and change in its extent can influence the availability of water in the Himalayan Rivers (Kulkarni et al., 2008; Immerzeel et al., 2010, 2012). The seasonal snow constitutes a major portion of the water budget, contributing up to 50%, and even more, to the annual discharge (Bookhagen and Burbank, 2010). Therefore, a good representation of the snow cover area in the hydrological models is important for accurate hydrological simulation. To measure the snow cover area and properties, remote sensing offers a new and valuable tool for getting snow related data, however it is limited to determining only the snow cover area. Numerical modelling of snow could be an efficient tool to quantify physical parameters such as snow water equivalent, and snow depth (Brun et al., 1994). However, there are large uncertainties in many variables defining the physical parameters of the snowpack, and the snow albedo is one of the most significant of those for its direct effect on the energy input to the snow from solar radiation (Anderson, 1976).
Aerosol (suspended solid or liquid particles in the atmosphere) such as black carbon (BC) is a major anthropogenic pollutant originating from biomass burning and fossil fuel combustion (Yasunari et al., 2010). When BC is deposited to snow/glacier, it may accelerate melt through perturbations to snow albedo and their associated radiative forcing (Matt and Burkhart, 2018). Moreover, when BC is in the atmosphere, it absorbs solar radiation, thus, resulting in a warming of the atmosphere (positive radiative forcing) (Cappa et al., 2012). However, the strength of these effects depends on the concentrations of these particles in the atmosphere or at the surface. Previously, a wide range of global and regional studies have been focusing on aerosol-radiation interactions via observations, laboratory experiment, and model simulation (Ramanathan, 1998; Ramanathan et al., 2001; Levy et al., 2007; Mei et al., 2013; Sič et al., 2015). These studies showed that aerosol-radiation interaction significantly contributes to the atmospheric energy balance.

To represent complex hydrological processes, hydrologists have developed the different numerical models also called hydrological models, which are appropriate for different situations and purposes. These models are globally used to understand the water resource availability and its dynamics in the region. Understanding these changes in the Himalayan region is complicated due to intricate monsoon dynamics, and the varying dependence on snow and glacier melt (Smadja et al., 2015). The accurate hydrological simulation requires consistent forcing data for a long period. There is a lack of forcing data with high quality and consistency between variables in the Himalayan region (Pellicciotti et al., 2012). Regional climate model and reanalysis data fulfill the spatial and temporal consistency but suffer from bias that limits their use for accurate hydrological simulations. Commonly used regional and climate model forcing data for hydrological simulation in the Himalayan region are from ERA-Interim (Dee et al., 2011) e.g. used by Sapkota (2016) in the centre Himalayan catchment, Water and Global Change (WATCH) Forcing Data (Weedon et al., 2014) e.g. used by Hegdahl et al. (2016) in Upper Beas catchment, India, and Coordinated Regional Climate Downscaling Experiment (CORDEX) (Ozturk et al., 2012) e.g. used by Ghimire et al. (2018) over the Himalayan region.

In terms of formulation, three different types of hydrological models can be considered. In order of complexity, these are empirical, conceptual and physically based hydrological models. Empirical models are often developed from an investigation of simple data sets and provide few information about the physical processes. An opposite to en empirical models, physically based models attempt to represent all the relevant processes in
1 Introduction

the hydrological system. Their complexity, long computation time and enormous data requirement lead them to have limited practicality in most contexts. One of the most complicated, physically based models developed is the SHE (System Hydrologique Européen) model (Abbott et al., 1986). Conceptual models, on the other hand, make use of laws derived from environmental data and are lumped or distributed. These models are widely used for water resource management. An example is Statkraft’s Hydrological Forecasting Toolbox (Shyft) (Burkhart et al., 2020) which is commonly used for operational hydrology and also used in this thesis.

However, hydrological modelling outputs from these conceptual models are subject to the uncertainty resulting from different sources. The hydrological model uncertainty depends on model structure, theories of runoff generation, routing, calibrated parameters, and the quality of model forcing data sets (Yi et al., 2018). Dehotin and Braud (2008) mentioned that the uncertainty in the hydrological simulation mainly depends on the choice of a proper level of spatial catchment discretization to handle the land use heterogeneity. Commonly used catchment discretization methods in hydrological applications are grid, e.g., Tomassetti et al. (2005); Burkhart et al. (2016); Hegdahl et al. (2016); Ragettli et al. (2016), and hypsography based, e.g., Martinec et al. (2008). Recently, triangulated irregular networks (TIN) based discretized hydrological models are also used for discharge simulation (Singh and Fiorentino, 1996; De Wulf et al., 2012) but to date, TIN based distributed hydrological models are not tested in the Himalayan catchment.

1.1 Motivation

Atmospheric aerosols arising from the variety of natural and anthropogenic emission sources are known to affect the global and regional climate through the modification to the radiative budget of the atmosphere and the surface below (Ramanathan et al., 2001; Esteve et al., 2014) through the scattering and absorption of short-wave (SW) and long-wave (LW) radiation (the aerosol direct effect). There are several types of research focus on studying aerosols over the Himalayan region due to its impact on the regional climate and hydrology (Ramanathan et al., 2001; Menon et al., 2002; Lau and Kim, 2006; Matt and Burkhart, 2018; Matt et al., 2018). The feedback mechanisms between aerosols over the Himalaya and regional climate are also found to affect the dynamics and hydrology of the regions (Gautam et al., 2009; Yasunari et al., 2010). This would have implications
1.1 Motivation

on the hydrological cycle, water resources, hydropower generation, and agriculture in the downstream regions, possibly affecting the living conditions of a population of more than a billion people (Immerzeel et al., 2010).

Despite their significant contribution towards radiative and hydrologic effects, spatio-temporal distribution of aerosols and their climatic impacts are poorly observed and understood (IPCC, 2007), particularly over the Himalayan region (Srivastava et al., 2015) due to the limited aerosol observation stations. High-altitude harsh topography and inaccessibility generally limit the installation and maintenance of the stations in the Himalayan region. In this regard, satellites offer a global perspective on many atmospheric variables, including aerosols and their properties. Remotely sensed data from satellites has the potential to account for the highly variable aerosol properties on regional as well as on global scales and to provide repeated observations over long periods.

A well-known example is the MODIS (Moderate resolution Imaging Spectroradiometer) instrument, which can provide daily aerosol optical properties with nearly global coverage at the resolution of 3-10 km (Remer et al., 2013). However, Govaerts et al. (2009) and Mei et al. (2012) identify the snow cover area as a great challenge for aerosol properties retrieval from remote sensing. The high surface reflectance makes it challenging to separate radiation at the top of the atmosphere due to reflection from the snow and atmospheric scattering by aerosol particles. As a result, information on aerosol properties over the cryospheric portion is mainly missing. To fully understand the effect of aerosols over the cryospheric portion of the Himalayas, detailed knowledge regarding the spatio-temporal distributions of aerosols, and their seasonal variability in the atmosphere are required (Bonasoni et al., 2012). To date, there is no model that can be used to estimate aerosol or its properties such as aerosol optical depth (AOD) over the Himalayan region. Therefore, we were motivated to develop a model that can be used to estimate AOD over the Himalayan region.

Hydrologic simulation of discharge and other water balance components (evapotranspiration, snow and groundwater storage) is used for the analysis of available water resources in the Himalayan region, e.g. (Ragettli et al., 2016; Immerzeel et al., 2010; ICIMOD, 2015; Immerzeel et al., 2009), but the quality of discharge simulation remains a challenge (Pellicciotti et al., 2012). The choice of a suitable hydrological model and appropriate forcing data is critical for any analysis and significantly affects the outcome (Immerzeel et al., 2010).
Currently, no standardized or community modeling framework within hydrology exist, so prior studies of the water resources of the Himalayan region have been conducted using different types of hydrologic models. Recent studies have addressed the performance of global and regional forcing data sets, including WFDEI, CORDEX, and ERA-I, for discharge simulation in various regions across the world, but to date, the use of these global forcing data sets to predict discharge in a Himalayan catchment is limited. Thereby, we want to address this research gap by evaluating the impact of different forcing datasets particularly; WFDEI, CORDEX, and ERA-I using community based hydrological modeling framework (i.e. Shyft) in the data-scarce Himalayan region.

Distributed hydrological models are commonly used for the hydrological simulation in the Himalayan region. These models hold the promise of representing the spatial variability of hydrological processes in the region. Distributed hydrological modeling approaches commonly involve the discretization of a catchment into several modeling elements. A very common approach in practice is to use elevation, regularly spaced grid and TIN based catchment discretization. Although the use of grid and elevation based discretized models are more prevalent because of their simplicity and ease of implementation (Freer et al., 1997; Woods et al. 1997), TIN based discretized models have a better potential of accurately representing topographic features than grids and elevation (Singh and Fiorentino, 1996).

Different catchment discretization methods are available and tested globally, but limited studies have been focused on the selection of appropriate discretization methods for hydrologic simulation in the Himalayan catchment. It is, therefore, our motivation to find an appropriate model according to different catchment discretization methods.

1.2 Goal and Objectives

Having shortly described the context and the main research gaps in the Himalayan region, the goal of this thesis is now introduced. The main goal of this thesis is to investigate the relative impact of hydrological model forcing and spatial discretization methods on hydrological response. This goal is achieved through the following objectives:

1. To develop an empirical model by using multiple regression, to increase the present understanding of spatio-temporal variability of aerosol optical properties over the cryospheric portion of Nepal.
2. To evaluate global and regional forcing data sets for hydrological simulations in the Himalayan catchment.

3. To investigate the influence of catchment discretization on the hydrological response of the Himalayan catchment.

1.3 Scope

This thesis work is dedicated to a continuous process, which has received most of the attention by the academic research. The present work is focused on an assessment of hydrological model forcings and catchment discretization for hydrologic predictions. This thesis aims at contributing to hydrological modelling in the Himalayan catchment by developing an empirical model for estimating aerosols optical depth (AOD) and evaluating different global forcing data sets and spatial catchment discretization methods.

1.4 Thesis outline

This thesis consists of the two parts. **Part I** provide an overview of the presented research and yet unpublished work, and is consist of six chapters. Chapter 1 describes the introduction of subject topic including the motivation, objectives, and scope of the study. Chapter 2 provides general background on the state of knowledge of the aerosol observation and hydrological modeling over the Himalayan region. Chapter 3 contains details on the study area, and methodologies applied in the analysis. Different hydrological model forcing datasets are also described in the chapter. Chapter 4 summaries the main findings of three research papers. Discussion on the topics are presented in Chapter 5. Finally, the conclusions are summarized in Chapter 6 alongside with some recommendations.

**Part II** consists of three scientific journal articles forming the basis of this thesis. Paper I and II are peer-reviewed and open-access, published in internationally recognized scientific journals. The manuscript of the Paper III is submitted to the Hydrology and Earth System Sciences journal of Copernicus Publications.
2 Background

2.1 Modelling geophysical processes

Different types of models are developed to understand and better explain the natural environment. Selection of an appropriate model depends on a different situation and purposes. No one type of model can be considered being better than any other; each has strength and weaknesses, and the choice of model type depends greatly on the system to be modelled and the problem to be addressed (Wilby 1998). Generally, there are three different types of models based on their formulation. In order of complexity, these geophysical models such as hydrological and atmospheric models are empirical, conceptual and physically based models.

Empirical models

An empirical model describes the mathematical relationship that supports a quantifiable analysis of the system parameters. These models are concerned only with describing how the system behaves, with little attempts to explain the underlying physical principles. Having said that these models are an indispensable tool used by the operational and atmospheric research communities for data analysis, initialization of detailed physics-based models, and instrument design. Empirical models can reflect the recent state of the atmosphere and be able to forecast by continually adding current data to the models and then modifying their parameter sets.

Conceptual models

The conceptual model is a descriptive representation of the system that incorporates the modeller’s understanding of the relevant physical, chemical, and hydrologic conditions. They attempt at describing the main processes occurring in the system following a systemic approach. These models help us to investigate correlations and processes in natural and in complex Earth system models. Sometimes conceptual models are also called grey-box models. Most of the hydrological models are the conceptual models. These
models are an excellent strategy for getting to the bottom of the research issue. For example, in hydrology; conceptual precipitation-runoff models, build based on observed or assumed empirical relationship among different variables. Conceptual rainfall-runoff models are designed to predict the magnitude of stream flow by conceptualizing rainfall-to-runoff generating processes and simulating internal variables, such as soil moisture, by various types of the response function. Similar to hydrology, there is also some conceptual atmospheric model such as bulk model [Naumann et al. 2017].

**Physically based models**

Physics-based models are built from fundamental conservation laws describing the physical laws of mass, energy, and momentum conversion. The formulation of the parametrizations is based on the physical knowledge of smaller-scale processes along with some dynamical-statistical up-scaling assumptions and therefore require a significant amount of information [Jinkang et al. 2007]. Depending upon the computing machines and available model forcing data, physical models are 2D-var, 3D-var and 4D-var. Generally, the 4D variational data (4D-var) used for assimilation for global scale air quality in atmospheric models [Daescu and Carmichael 2002].

### 2.2 Aerosol and climate

The geographic distribution of the energy is an important driver of the climate system and is determined by the incoming short-wave radiation [Trenberth and Stepaniak 2004]. The shortwave radiation received at a site varies in time: between day-night and between seasons. At a given time it also varies with the concentration of atmospheric constituents such as aerosols. Aerosols are the suspension of liquid and solid particles (size: $10^{-4}$ to $10 \, \mu m$) in the atmosphere, excluding clouds and precipitation. Aerosol particles reach the atmosphere from a wide variety of natural and anthropogenic sources such as biomass burning, incomplete combustion of fossil fuel, volcanic eruptions, and biological materials such as plant fragments, microorganisms, pollen [Srivastava et al. 2015]. They act as ice nuclei and cloud condensation nuclei, therefore, affect regional cloud properties and may affect precipitation amounts.

Atmospheric aerosols arising from the variety of natural and anthropogenic emission sources are known to affect the global and regional climate through the modification to the radiative budget of the atmosphere and the surface below [Ramanathan et al. 2001].
through the scattering and absorption of short-wave (SW) and long-wave (LW) radiation (the aerosol direct effect). Depending upon the sources, they fall under the following categories: sulphate, black carbon, organic carbon, dust, and sea salt. These particles scatter and absorb solar and terrestrial radiation in the atmosphere and can affect the radiative balance of the climate system.

These effects are usually quantified in terms of aerosol radiative forcing and are defined as the effect of the aerosols on the radiative fluxes at the top of the atmosphere (TOA) and at the surface of the Earth. Solar radiation absorbed by aerosols warms the air directly instead of allowing sunlight to be absorbed by the surface of the Earth, leading to the enhanced warming of the atmosphere \cite{Ramanathan1998}. Moreover, variations in the surface radiation balance cause changes in evapotranspiration as well as precipitation, and thereby govern the intensity of the hydrological cycle. Furthermore, when aerosols like black carbon (BC) are deposited on snow/glacier surface, it can significantly affect regional climate by modifying snow/ice reflectance and thus altering the snowmelt rate and cryosphere spatial coverage \cite{Matt et al. 2018}. Even though its significant contribution towards radiative and climatic effects, the impact of aerosol radiative forcing are poorly understood \cite{IPCC2007}, particularly over the Himalayan region \cite{Srivastava et al. 2015}, partly because of insufficient ground measurements data and inadequate information about the aerosols optical and chemical characteristics on spatial and temporal scales.

There has been attention and interest of studying aerosol properties and its radiative forcing over the Himalayan region due to its unique topography, geographic location weather driven by monsoon circulation and seasonal transportation of aerosols. The Himalayan region received aerosols derived from natural as well as anthropogenic emission sources \cite{Chatterjee et al. 2012, Li et al. 2016a}. \cite{Li et al. 2016a} reported that the Himalayan mountains and the southern parts of the Tibetan Plateau receive much of their aerosols, particularly BC from emissions in the Indo-Gangetic Plain (IGP). The feedback mechanisms with closed links between the aerosols over the Himalayas and regional climate are found to affect the dynamics and hydrology of the regions \cite{Gautam et al. 2013}. From an economic perspective, changes in the hydrological cycle can impose great pressures and damages on a variety of industrial sectors, such as water management, hydropower development, agricultural production and tourism. Despite their obvious environmental and societal importance, our understanding of the aerosol and its distribution over the Himalayan re-
2 Background

gion is limited.

2.2.1 Aerosol: properties, observation, and modelling

Aerosol optical depth (AOD) is a fundamental parameter in the aerosol optical properties. It is the measure of aerosols distributed within a column of air between the instrument and the top of the atmosphere. It quantifies the degree to which aerosol particles absorb or scatter solar radiation. AOD values for clean atmospheric conditions is close to zero values within ± 0.05 (Remer et al., 2005). There are different methods available to monitor atmospheric aerosols and their properties. Ground-based observation from AErosol Robotic NETwork (AERONET) (see: Fig. 2.1) and observation from remotely sensed satellite images are well-established methods to monitor aerosols and its properties. Ground-based observations provide fairly accurate AOD measurements and better temporal information (≃15 min), but they lack spatial coverage. Satellite-based observation gives more comprehensive spatio-temporal information but aerosol retrievals are challenged over the bright surface (e.g., desert and snow cover area) and have significant retrieval uncertainties.

Figure 2.1: AERONET station at Jomsom, Nepal.

Therefore, using only one method with one or a few instruments seems not to be suitable for describing the spatio-temporal distribution of aerosols. The combination of ground-based measurements, remotely sensed data, and model simulations could be a suitable
approach to overcome this difficulty. There are different types of atmospheric models that can be used for AOD estimation. The classification of models depends on how the processes are described in the structure of the model. Generally, atmospheric models are classified as empirical or physical models.

**Current aerosols modelling approach**

Physically based models are generally used for estimating atmospheric aerosols and its properties. For example, ECHAM-HAM \cite{Zhang2012}, The Lagrangian particle dispersion model FLEXPART \cite{Brioude2013}, MOCAGE \cite{Sic2015}, NorESM \cite{Schwinger2016}, Integrated Forecasting System of European Centre for Medium Range Weather Forecasts (ECMWF-IFS) \cite{Roberts2018}. All these models have their own limitations and strengths. For all physically based model, the accurate prediction of aerosols and its properties often requires the observed or remotely sensed forcing data. For example, the aerosol fields of the ECMWF-IFS model are constrained by assimilating AOD retrievals product by the MODIS instruments \cite{Bozzo2020}. AOD information from MODIS is lacking over the Himalayas and is remain a source of uncertainty for aerosol predictions. Although these models provide a global perspective on aerosol, the spatial resolutions of these models are usually too coarse for the local level study.

Empirical models are also used for the study of atmospheric aerosol and its properties. For example, Chaloulakou et al. \cite{Chaloulakou2012} used a neural network model and regression model to estimate daily average PM10. Yang and Zong \cite{Yang2014} used an empirical model for estimating the stratospheric aerosol extinction profile, and comparison with measurement found significant. Using empirical models for different studies shows that it seems a very promising tool for atmospheric and climate research. Although, there are some empirical models that are already used to estimate aerosols and its properties, none of the models has been used in the Himalayan region, particularly over the cryospheric portion.

### 2.3 Climate of the Himalayan region

Climate change is expected to have a strong impact on the Himalayas, where the rise in temperature is higher than the global average \cite{Smadja2015}. The impacts of climate change on river flows, ground water recharge, natural hazards, could be significant.
2 Background

Given the present understanding about climate change, determining diversity of impacts is a challenge for researchers, risk assessment is needed to guide future action (Oreskes, 2004). IPCC (2007) and Takemura and Suzuki (2019) reported an unprecedented warming trend during the 20th century, mainly because of anthropogenic greenhouse gases and aerosols (particularly black carbon) concentrations in the atmosphere. Temperatures of the last half-century are unusual in comparison with those of previous 1300 years (IPCC, 2013). The current average global surface temperature of 15 °C is nearly 0.6 °C higher than it was 100 years before and is 0.74 °C (0.56 °C to 0.92 °C) higher than the past 50 years (1906-2005) which show that there is a rapid warming of global surface temperature (IPCC, 2008).

Several studies e.g. Lau et al. (2010); Singh et al. (2010); Rangwala and Miller (2012); Wiltshire (2014) indicate that the Himalayas are warming significantly faster than the global average. For example, studies carried out by Kuang and Jiao (2016) and Yao et al. (2006) reported that the Tibetan Plateau has experienced warming at the rate of 0.02 °C to 0.03 °C per year over the last 50 years, which is much greater than the global average. Ren et al. (2017) found the surface mean air temperatures over 1901-2017 show a significant increase of 0.104 °C per decade in the entire Himalayan range. Similarly, the mean air temperature in Nepal has increased at a rate of 0.04 °C per year for the years 1975 to 2005 (Baidya et al., 2008) which is much higher than the mean global rate of 0.74 °C for the years 1906-2005 (IPCC, 2007) while the analysis of annual mean minimum and maximum temperature for the years 1976 to 2005 by Marahatta et al. (2009) showed higher increase in maximum temperature (0.05 °C per year) than minimum temperature (0.03 °C per year). Moreover, not only in the regional scale studies, but also some of the catchment scale studies, such as Gautam et al. (2013); Krishnan et al. (2019); Scott et al. (2019); Chand et al. (2019) have also shown the significant increasing temperature trends in the Himalayan river catchment.

Besides the historical temperature trend analysis, some of the future temperature change analysis in the Himalayan region were also carried out. For example, a study carried out by ICIMOD (2015) demonstrated that by 2050, temperatures across the Himalayan river catchment are projected to increase by about 1-2 °C on average, with winters seeing greater warming than summers. The study suggested that the mountainous and high altitude areas are particularly affected, with warming reaching 4-5 °C in some places. Similarly, Kulkarni et al. (2013) used high-resolution regional climate model
data to project climate change over the Hindu-Kush-Himalayan (HKH) region. In their study, the annual average temperature was projected to increase by 4 to 5 °C toward the end of the 21st century.

Similar to temperature change, global precipitation change over the past few decades were also reported in the different research works e.g. Jiang and Yang (2012); Nguyen et al. (2018); Fujita et al. (2019). Annual mean and pixel-based trends of global precipitation from 1983 to 2015 from the PERSIANN-CDER datasets were analysed by Nguyen et al. (2018). In their study, global precipitation (60 °N to 60 °S) was found to show the increasing trend (2.36%), which is statistically significant. According to IPCC (2008), precipitation over land generally increased over the 20th century between 30 °N and 85 °N, but the notable decrease has occurred in the past 30 to 40 years from 10 °S to 30 °N. However, there is a pattern of negative and positive trends across the globe with decreases over some middle latitude regions and an increases over tropical oceans. But in Nepal, there is an increase in precipitation (analysis from 1978 to 2001) i.e. 0.6% annually (Chaulagain, 2003) though it lies between 26 to 30 °N. The precipitation fluctuation in Nepal is not the same as the all-India precipitation trend, but it is well related with rainfall variations over northern India (Kumar et al., 2006; Pokharel and Hallett, 2015).

Future global and regional precipitation projection have been analysed in several studies e.g. Kulkarni et al. (2013); Giorgi et al. (2019); Bhowmick et al. (2019); Kusunoki et al. (2019). Their studies show that the future temporal and spatial variation characteristics of precipitation are different in different regions. The CGCM (Canadian Global Climate Model) suggests a modest increase in global precipitation and an El-Nino like change in precipitation pattern under warmer climates. By about 2050, precipitation changes in this projection are generally relatively small across medium to high latitudes but show large increases for the tropical Pacific and west American coastline, and large decreases over the south-eastern United States and southern Europe. Rajbhandari et al. (2018) used projected precipitation from high resolution climate models (PRESIS: Jones et al. (2004)) over the central Himalayan catchment and found an increase in precipitation in the near future (2011-2040) and a progressive increase towards the end of the century (2071-2098). The projected change in rainfall was also non-uniform, with an increase over the southern plains and the middle mountains and decreased over the trans-Himalayan region.
The hydrological cycle of the Himalayan region is complicated by the Asian monsoon (Gain et al., 2011), but there is little doubt that snowmelt and icemelt are to be important regulators of the seasonal discharges of the Himalayan rivers. Seasonal distribution of snow and ice melt contributions to the river discharge are heavily dependent on snow/ice storage and air temperature. Temperature and precipitation change along with black carbon aerosols deposited in the Himalayan region is expected to affect the glacier area and ice volume adversely (Immerzeel et al., 2012). It will affect the regional hydrological cycle and water budget. The fourth assessment report of the Inter Governmental Panel on Climate Change (IPCC) compiled existing knowledge on climate change, including the key indicators, based on previous research (IPCC, 2007). The rapid shrinkage of these glaciers due to climate change is likely to seriously threaten water availability in the region, particularly during lean flow seasons when melt water contribution is crucial to sustaining the river flow which supports human activities and ecosystem services in these areas and downstream (IPCC, 2008). One of the significant impacts of climate change in the Himalayan region is the retreat of glaciers, which are the reliable sources of freshwater to many people living downstream to meet their needs for water supply, irrigation, hydropower and navigation.

Increasing temperature and precipitation trends are expected to drive consistent increases in the total streamflow of the Indus, the Ganges, and Brahmaputra rivers (Immerzeel et al., 2013; Scott et al., 2019; Lutz et al., 2019). Hydrological modelling of the Upper Indus catchment using SRM by Immerzeel et al. (2009) found that regional warming is affecting the discharge in the basin because of the accelerated melting of glaciers. Similarly, Immerzeel et al. (2012) made a study about the impact of climate change on the hydrology of a glacierized catchment in central Nepal using the PCRaster environment for numerical modelling (Karssenberg et al., 2002). The study simulated the glacier evolution (including location and permanent snow) using an empirical model and estimated the impact of future climate change on glacier hydrology. The study results suggested that both glacial area and glacial ice volume will substantially decrease in the future due to increasing temperature scenarios. Similarly, Li et al. (2016b) made a study about the water resources under climate change in two Himalayan basins (i.e., Chamkhar Chhu basin Bhutan and Beas basin, India). The study simulated future climate data using two Regional Climate Models (RCMs) and three Representative Concentration Pathways (i.e., Rcp2.6, Rcp4.5, and Rcp8.5). The study showed that glaciers in the Chamkhar Chhu basin are predicted to disappear or reduce to small size before the 2050s, whereas the
glaciers in the Beas basin are expected to lose mass before the 2060s. There is a prediction of an increase in annual river discharge until around 2030 and then decrease because of the rapid melting of snow and glacier in the beginning, and then significant decrease of available snow and glacier mass thereafter [IPCC, 2007]. Although there are many uncertainties in different researches, it is crucial to consider the impact of climate change on glacier retreat, and on regional water resources of the Himalayan river catchment.

The impacts on river flows, ground water recharge, natural hazards, could be dramatic, although not the same in terms of rate, intensity or direction in all parts of the region. Existing knowledge about climate change, determining diversity of impacts is a challenge for researchers, risk assessment is needed to guide future action.

2.4 Hydrological modelling in the Himalayas

Hydrology of mountainous catchments has significant implications on regional climate and socio-economic aspects. Their water supplying role is particularly relevant for millions of people living downstream of the major mountain range. Hydrological modelling can give an improved understanding of the Mountain hydrology with important insights into the possible present and future changes to the hydrological dynamics of the region. Application of hydrological models is necessary both for prediction of the hydrological responses of the catchment over the large area and for simulations of the catchment future responses [Pellicciotti et al., 2012]. To understand the past and current state of water resources and to predict future hydrological characteristics in the Himalayas, hydrological models are the primary tools today. Different types of hydrological models are used for different purposes in the Himalayan region. In terms of their formulation, the hydrological model used in the Himalayan region are three types. In order of complexity, these are empirical, conceptual and physically based models.

Most of the empirical models used in Himalayan catchment are for glacio-hydrological modelling. For example, an empirical model used by [Ageta and Higuchi, 1984] to calculate glacier melt, a simple empirical model for estimating glacier melt under debris layer used by [Nakawo and Takasashi, 1982], and energy balance modeling or glacier mass balance on the glacier AX010 in Nepal used by [Kayastha et al., 1999]. These model provides no explanations of the physical processes of glacier melt. Mostly daily average temperature is used to derive the model. Despite their conceptual naivety, empirical models can have
significant applications for the region with a lack of observations. These models can be applied quickly and easily for designing flood protection in catchments with little or no flow gauging history. For example, a Clark unit hydrograph (Clark, 1945) was previously used by the Department of hydrology and meteorology, Government of Nepal for flood forecasting.

Direct field observations are very difficult to carry out in the Himalayas because of the rugged and remote mountain terrain. So, the model and method to predict streamflow should be simple with a minimum field data requirements. Conceptual hydrological models only approximate the general hydrologic processes based on limited representations of the processes occurring in the hydrological system. Following this concept of simplicity, different conceptual hydrological models were used for hydrological modeling in the Himalayan region. Because of simplicity and reasonably good result, conceptual hydrological models have been used by many authors. Some of the conceptual hydrological models extensively used in the Himalayan catchments are Snowmelt Runoff Model (SRM) (Martinec et al., 2008), used by Bhattarai and Regmi (2015); Rasouli et al. (2015); Khadka et al. (2016); Hayat et al. (2019), Hydrologiska Byrans Vatenbalansavdelning (HBV) (Bergström, 1976), used by Normand et al. (2010); Li et al. (2014); Parajuli et al. (2015); Ciupak et al. (2019), Hydrologic Engineering Centre’s River Analysis System (HEC-RAS) (Horritt and Bates, 2002), used by Parhi (2013); Jha and Khare (2016).

All these studies found that data scarcity is the main limitation for the application of more advanced physically based hydrological models. Rana et al. (1997) used a conceptual runoff model called HYCYMODEL in Langtang River catchment, Nepal, using the method developed by Ageta and Higuchi (1984). They used satellite derived surface-temperature data to calculate glacier ablation rate. Similarly, Li et al. (2019) used WASMOD and Matt and Burkhart (2018) used Shyft modelling framework (Burkhart et al., 2016) to assess the impact of black carbon aerosols on the hydrological response in the Beas catchment, India. Furthermore, Hasson et al. (2019) used the energy-balance hydrological model of the University of British Columbia (UBC) to assess water availability in Pakistan from HKH watersheds. Some of the conceptual hydrological models also give opportunities to combine with remotely sensed data, for example, Immerzeel et al. (2010) used the Normalized Melt Index (NMI) method to quantify the importance of meltwater from the upstream area on the Himalayan basins. Studies show that climate change will
affect water availability and food security in the regions.

Based on the development in computational power and availability of distributed data sets, various physically based hydrological models are used for the hydrological simulation in the Himalayan regions. The main characteristics of a physically based models are their parameters are physically realistic and can therefore be obtained independently of the model. TOPKAPI (TOPographic Kinematic APproximation and Integration) model is a fully distributed physically based hydrological model used by Peng et al. (2008); Ragettli et al. (2016); Atif et al. (2019) in various Himalayan catchments. This model provides high resolution information on the hydrological state of the studied catchments. Similarly, Badar et al. (2013) used the Generalized Watershed Loading Function (GWLF) model to simulate the hydrological process under the impact of changing land use conditions in Dal lake catchment of Kashmir Himalayas. This model was able to simulate discharge under landuse conditions. Kayastha and Kayastha (2019) used physically based glacio-hydrological degree-day model (GDM) in the Himalayan river catchment and suggested that this model can be used as a promising tool to study hydrological system dynamics, and potential impacts of climate change on the Himalayan river catchment.
3 Hydrological model and study area

3.1 Hydrological model

Statkraft’s Hydrological Forecasting Toolbox (Shyft; https://gitlab.com/shyft-os) is used for the hydrological simulation in the thesis. Shyft is an open-source computing framework developed by Statkraft AS, Norway. It provides an optimized platform for the implementation of many well known hydrological models from conceptual to physically based distributed hydrological models. High-performance generic time-series framework in the Shyft allows for rapid calculations of hydrologic response at the regional to a grid cell scale. Different governing equations for calculating hydrologic response, evapotranspiration, and snow distribution and melt in the models forms different hydrological models within the Shyft and represented by different stacks (Fig. 3.1).

![Figure 3.1: Hydrological models and modelling stacks in the Shyft.](image-url)
In this study, PT_GS_K and R_PT_GS_K model stacks (see Fig. 3.1) were used for the discharge simulation. Both modelling stacks are conceptually distributed hydrological models. Standard model input variables are temperature (T), precipitation (P), wind speed (WS), relative humidity (RH), and shortwave radiation (S). In PT_GS_K and R_PT_GS_K modelling stacks, potential evapotranspiration was estimated according to Priestley-Taylor (PT) approach (Priestley and Taylor, 1972), snow distribution and melt are calculated using Gamma snow (GS) method (Sec. 3.1.3), and the catchment response function (K) was based on the storage-discharge relationship concept described in Kirchner (2009).

3.1.1 Catchment response function: Kirchner (K)

The main core of the Shyft (except HBV model) is based on the rainfall-runoff modelling approach presented by Kirchner (2009). Kirchner’s model is based on the law of conservation-of-mass, and stated as the change in the volume of water stored in a catchment (S) is equals to the incoming precipitation (P) minus evapotranspiration (E) and runoff (Q), and mathematically expressed as:

\[
\frac{dS}{dt} = (P - E - Q) \tag{3.1}
\]

Where, P, Q, E, and S are understood to be a function of time and measured in units of depth (e.g. mm of water per hour).

In the analysis, discharge in the stream (Q) is assumed to solely depend on the amount of water stored in the catchment (S), and can be defined by the storage-discharge function \( f(S) \) such that:

\[
Q = f(S) \tag{3.2}
\]

Equations 3.1 and 3.2 form a first-order dynamical system, where 'Q' be a non-linear function of 'S'. Catchment sensitivity to changes in storage with respect to time is determined by differentiating Equation 3.2 and substituting to Equation 3.1 and given by Equation 3.3

\[
\frac{dQ}{dt} = \frac{dQ}{dS} (P - E - Q) \tag{3.3}
\]
3.1 Hydrological model

Where \( \frac{dQ}{dS} \) represents the sensitivity of discharge to changes in storage. It can be expressed as a function of discharge called the 'sensitivity function' and defined as \( g(Q) \). Replacing \( \frac{dQ}{dS} \) by \( g(Q) \) and rearranging Equation 3.3 becomes Equation 3.4.

\[
g(Q) = \frac{dQ/dt}{(P - E - Q)}
\] (3.4)

Under the conditions of very low precipitation and evapotranspiration as compared to discharge (i.e. \( P \ll Q \) and \( E \ll Q \)), sensitivity function \( g(Q) \), can be estimated by fitting the quadratic function to the observed time series discharge \( Q \) (Kirchner (2009), Eqs. 7, and 9).

\[
g(Q) = e^{c_1 + c_2 \ln(Q) + c_3 (\ln(Q))^2}
\] (3.5)

Where \( c_1, c_2, \) and \( c_3 \) are the catchment specific outlet parameters and obtained during model calibration.

Now from the Equations. 3.4 and 3.5, catchment response to precipitation, evapotranspiration is expressed by Equation. 3.6.

\[
\frac{dQ}{dt} = e^{c_1 + c_2 \ln(Q) + c_3 (\ln(Q))^2} \cdot (P - E - Q)
\] (3.6)

3.1.2 Potential evapotranspiration: Priestley-Taylor (PT)

In the PT_GS_K and R_PT_GS_K hydrological modelling stacks, potential evapotranspiration is calculated by using Priestley and Taylor (1972) approach and is given by Equation 3.7. It is a modification of the Penman-Monteith equation (Beven, 1979) (PM) through the empirical approximations to remove the dependence on observations data. For Priestley-Taylor, only radiation (irradiance) observation is required to estimate potential evapotranspiration.

\[
E_{pot} = \frac{\alpha}{\lambda} \frac{s(T_a)}{s(T_a) + \gamma} K_n
\] (3.7)

Where \( \alpha = 1.26 \) being a dimensionless empirical multiplier, \( \gamma \) the psychometric constant, \( s(T_a) \) is the slope of the relationship between temperature \( (T_a) \), and the saturation vapour pressure, latent heat of vaporization is denoted by \( \lambda \), and \( K_n \) is the net radiation. These parameters are explicitly described in Westergren (2016).
3 Hydrological model and study area

3.1.3 Snow distribution and melt: Gamma snow (GS)

In gamma snow (GS), snow distribution and ablation in each modelling element (or cell) is described by two different approaches. Snow Depletion Curve (SDC) concept of Kolberg et al. (2006) is used to estimate the snow state within each grid cell, and an energy balance approach is used to calculate snow ablation.

Snow ablation

The energy balance concept based on DeWalle and Rango (2008) approach is used to calculate snow net energy flux $\Delta E$ available for snow ablation in each grid cell and given by Equation 3.8.

$$\Delta E = S(1 - \alpha) + L_{in} + L_{out} + H_{SE} + H_L + E_G$$  \hspace{1cm} (3.8)

Where $S$ is the incoming shortwave radiation moderated for albedo ($\alpha$), $L_{in}$ and $L_{out}$ are the incoming and outgoing longwave radiation, $H_{SE}$ and $H_L$ represent sensible and latent heat fluxes respectively and are obtained by using vapour pressure from Bosen (1960) approach, wind speed, and temperature gradient measurements. $E_G$ is the subsurface energy flux. All terms are given in $W/m^2$. For a given time step (t), snow albedo ($\alpha$) at each grid cell depends on the minimum ($\alpha_{min}$) and maximum albedo ($\alpha_{max}$) values as well as on decay rate, temperature and snowfall as described in Hegdahl et al. (2016) and given as:

$$\alpha_t = \begin{cases} 
\alpha_{min} + (\alpha_{t-1} - \alpha_{min})(1 - \frac{1}{FDR}) & \text{if } T_{\alpha} > 0^\circ C \\
\alpha_{t-1} + (\alpha_{max} - \alpha_{min})(1 - \frac{1}{SDR}) & \text{if } T_{\alpha} \leq 0^\circ C 
\end{cases}$$  \hspace{1cm} (3.9)

In Equation 3.9, FDR and SDR denote fast and slow snow decay rates. FDR and SDR parameters can be understood as the time it takes for to decrease 95% of the value defined by the difference between the $\alpha_{max}$ and $\alpha_{min}$.

Finally, the simulated water equivalent ablation rate ($Melt_{we}$) at each grid cell is determined by Equation 3.10.

$$Melt_{we} = \frac{\Delta E}{\rho_w L_f} - \frac{H_L}{\rho_w L_v}$$  \hspace{1cm} (3.10)

Where $\rho_w$ is the density of water, $L_f$ is the latent heat of fusion, and $L_v$ is the latent heat of vaporization. Only positive values of $\Delta E$ are considered in the ablation rate calculation, but both negative and positive values of $H_L$ are considered.
3.1 Hydrological model

**Snow distribution**

For the sub-grid distribution of snow in the model, snow depletion curve (SDC) is used. The SDC, shown in Figure 3.2, is based on a three parameter gamma distribution i.e. average snow storage (m), the snow coefficient of variation (cv), and an initial bare ground fraction (1-A0). SDC describes the relationship between the Snow Cover Area (SCA) and the mass balance of the snowpack. In the model, SDC describes how SCA reduces gradually through the melt season and, moreover, relates the SCA to its respective Snow Water Equivalent (SWE), which is a model derived response variable.

![Figure 3.2: Snow Depletion Curve parametrized by average snow storage (m), the snow coefficient of variation (cv), and the initial bare ground fraction (1 − A0) (after Kolberg et al. (2006)).](image)

At each time step, fractional snow cover area at each grid cell is calculated by using Equation 3.11.

\[
A(t) = A_0 \times \{1 - F[\lambda(t)]\} \tag{3.11}
\]
3 Hydrological model and study area

In Eq. 3.11 $F[\lambda(t)]$ is the cumulative distribution function of SCA at a time 't' and is determined by Eq. 3.12

$$F[\lambda(t)] = \int_0^{\lambda(t)} p(m,cv)dx = \gamma \left(\frac{1}{cv^2}, \frac{\lambda(t)}{m \times cv^2}\right)$$ (3.12)

Where $p(m,cv)$ is the probability density function, and $\gamma$ is the cumulative gamma function with shape and scale arguments.

3.2 Study area

3.2.1 Geographical location, demography, and climate

The Himalayas is a mountain range in the Asia (Fig. 3.3), extends in an arc about 2400 kilometres and pass through the nations of Pakistan, China, India, Nepal, and Bhutan. North of the Himalayas is the Tibetan Plateau, and the south is the Indo-Gangetic Plain (IGP). North-South range varies in width from approx 350 kilometres (Kashmir) in the west to 150 kilometres (Arunachal Pradesh). The Hindu-Kush range in Afghanistan and Hkakabo Razi in Myanmar are not included in the Himalayan range, but they are both parts of the grater Hindu-Kush-Himalayan (HKH) river system. The main rivers sourced in the Himalayas are the Ganges, Indus, Yarlung, Yangtze, Yellow, Mekong, and Nujiang. The Himalayas are the largest deposit of ice and snow outside the polar regions, and are often referred to as the world’s ‘Third Pole’.

In the last 50 years (1961-2011), The Himalayas population has grown by 250%, from 19.9 to 52.7 million (Apollo, 2017). Rural to urban migration is one of the most widespread global demographic trends, and this is also predominant in the Himalayan region. The most populous cities include Kabul (Afghanistan), Kathmandu (Nepal), Srinagar (India), Peshawar (Pakistan), Quetta (Pakistan), Xinning (China), and Dehra Dun (India). Large cities such as Kathmandu, Lhasa are growing at rates that double the population of these cities every ten to fifteen years or so (ICIMOD, 2012). If the population keeps growing at the same rate, this can be a big challenge for Himalayan regions due to poor infrastructure or simply lacking enough water to meet demand (Rasul, 2014).
3.2 Study area

Himalayan weather and climate are governed by the summer and winter monsoon systems of Asia and is caused by the differential response of land and sea to incoming solar radiation. These are often pictured as giant land and sea breezes which blow across the subcontinent once a year with remarkable regularity. Winter monsoon is brought by the cyclonic storms associated with the mid-latitude subtropical westerly jet, also referred to as ‘western disturbances’ and the summer monsoon brought by the summer monsoon winds (Huang and Sun, 1992). The Himalayan climate is mostly alpine but varies significantly with location, elevation, and topography. The average summer and winter temperatures are about 30 and 18 °C in the southern foothills. The middle Himalayan valleys experience mean summer temperatures between 15 and 25 °C (Krishnan et al., 2019). In the high elevation region (> 4800 m asl), the temperature is mostly below the freezing point.
In this thesis, an empirical model to estimate AOD was developed and tested over the South Asian Himalayan country, Nepal (Figs. 3.3 and 3.4). It is located between latitudes 26.366°N to 30.45°N and 80.066°E - 88.2°E, and extends approx 885 kilometres along the east-west with an average width of 193 kilometres along the north-south. Within this relatively small latitudinal extent, the altitude rises from 57 m asl to the world’s highest peak. The country can be divided into three broad ecological zones, i.e., the Terai plains (18%), Churia hills (58%), and High mountains (Middle and Greater Himalayas, 24%).

The hydrological simulations were conducted in the Narayani River catchment and its sub-catchment (i.e., Marsyangdi-2 river catchment) located in central Nepal. The Narayani River catchment has a catchment area of 31692 km². It is a trans-boundary catchment, with 13% of the total area in Tibet, China (Fig. 3.4). Tributaries of the Narayani River are monsoon generated (originated in the middle and high mountain region) and snow-fed (originated in the higher Himalaya), and it finally contributes to the Ganges in India.
3.2 Study area

Figure 3.4: Map of Nepal with the location of study catchments.
3 Model forcings and validation datasets

3.3 Model forcings

Reanalysis and regional climate model data

**ERA-Interim (ERA-I):** ERA-I is a global atmospheric reanalysis dataset produced by the European Centre for Medium-Range Weather Forecast (ECMWF). It has sixty vertical model levels from the surface to 0.1 hPa, 27 of which are below 100 hPa, and covers a period from 1979 to the present day. ERA-I temperature results from the assimilated surface temperature (Essou et al., 2016), while precipitation data are based on meteorological model using a reanalysis of precipitation fields (Dee et al., 2011; Berrisford et al., 2011). Thusly obtained precipitation data are not rescaled using observation data. ERA-I is updated in near real-time, using data from the operational ECMWF forecast system. Data have a horizontal resolution of 0.75 ° x 0.75 ° and a 3-h temporal resolution for surface parameters, and 6-h for upper-air parameters. ERA-I dataset is freely available from http://apps.ecmwf.int/datasets/.

As per the requirement for Shyft, units for each forcing datasets are converted to a standard input unit. The temperature in ERA-I is in Kelvin (K) scale hence converted into degree Celsius. ERA-I precipitation is an accumulated field, and its unit is in Kg m\(^{-2}\) s\(^{-1}\). As per the requirement for Shyft, accumulated precipitation is converted to mm/hr. Wind speed (m/s) is not readily available in ERA-I, hence derived from the wind components \(U\) (m/s) and \(V\) (m/s) using Eq. 3.14. Incoming short wave radiation is in J/m\(^2\), hence converted to Wt/m\(^2\). Similarly, relative humidity is also derived from air and dew point temperature using Eq. 3.13 given by (Dingmann, 2015).

\[
RH = \frac{exp\left[\frac{17.625\times td}{243.04+td}\right]}{exp\left[\frac{17.625\times tas}{243.04+tas}\right]} \tag{3.13}
\]

Where \(tas\) and \(td\) are the two meter air, and dew point temperatures, respectively.

\[
WS = \sqrt{U^2 + V^2} \tag{3.14}
\]

Where \(WS\) (m/s) is the wind speed, and \(U\) (m/s) and \(V\) (m/s) are wind components.

**Water and Global Change (WATCH) Forcing Data ERA-Interim (WFDEI):** WFDEI is a global meteorological forcing dataset at 0.5° x 0.5° horizontal resolution obtained by
3.3 Model forcings and validation datasets

bias-correcting ERA-I data (Weedon et al., 2014). It covers the period 1979-2016 and contains eight metrological variables at a 3-h time step. WFDEI then corrects ERA-Interim for precipitation biases using data from the Climatic Research Unit (CRU) or the Global Precipitation Climatology Center (GPCC). In this study, GPCC products were preferred to CRU because of their higher resolution and data quality (Weedon et al., 2014). The WFDEI dataset is freely available online from ftp.iiasa.ac.at.

Similar to ERA-I, the temperature is in Kelvin (K) scale, hence converted to degree Celsius. Wind speed (m/s) and incoming short wave radiation (w/m$^2$) is readily available in WFDEI. Precipitation is also an accumulated (Kg m-2 s-1) hence converted into mm/hr. Since pressure and specific humidity is provided by WFDEI, relative humidity is calculated using Eq. 3.15 given by Salby (1996); Wallace and Hobbs (2006).

$$RH = 0.263 \times p \times q \left[ \exp \left( \frac{17.67(tas - 273.16)}{tas - 29.65} \right) \right]^{-1} \quad (3.15)$$

Where $p$ is pressure (pascal), and $q$ is specific humidity (dimensionless).

COordinated Regional Climate Downscaling EXperiment (CORDEX): CORDEX is a program sponsored by the World Climate Research Programme (WCRP) to produce an improved generation of regional climate change projections (Giorgi et al., 2009). CORDEX has two datasets as 'evaluation' and 'historical'. Evaluation is a run within reanalysis and is the one to use to 'mimic' observation (i.e., represent real weather), whereas historical is run within a climate model, and results can only be used in a climatological sense. Daily evaluation datasets available from CORDEX experiments 'The Abdus Salam International Centre for Theoretical Physics (ICTP) Regional Climatic Model version 4 (RegCM4)' with contributing institute 'Indian Institute of Tropical Meteorology (IITM), India', and Rossy Centre regional atmospheric model version 4 (RCA4) as a regional model with contribution from Swedish Meteorological and Hydrological Institute (SMHI), Sweden are evaluated before used in this study. Since the studied region is dominated by monsoons and evaluation datasets from RCA4 as CORDEX experiment and SHMI as contributing institute (hereafter CORDEX-SMHI), captures the monsoon better, hence selected for further study. The horizontal resolution of CORDEX-SMHI data is (0.44 ° x 0.44°) and freely downloaded on request from http://cccr.tropmet.res.in/home/index.jsp.
3 Hydrological model and study area

Similar to ERA-I and WFDEI, in CORDEX, temperature is measured in Kelvin (K) and, the precipitation is measured in Kg m-2 s-1; therefore, converted to degree Celsius and mm/hr. Wind speed is derived from the wind components ($U$ and $V$) using Eq. 3.14. The Relative humidity is derived using Eq. 3.15. Since incoming shortwave radiation is already in w/m$^2$ so no need to convert for the application.

Observed data

Nepal has a very short history of hydro-meteorological data collection. Since 1969 nationwide meteorological and hydrological networks were established. The Department of Hydrology and Meteorology (DHM), Government of Nepal (GoN), is solely responsible for collecting and disseminating hydrological and meteorological information for water resources, agriculture, energy, and other development activities. Most of the measured data are based on conventional manual observations. Measurements from the Automatic weather Station (AWS) are available only after 2000 AD (Karki, 2010). In total, 282 meteorological and 51 hydrological stations are currently operated by DHM, GoN. Geographic location and installation date for each station are available from http://dhm.gov.np/meteorological-station/. In this study, daily temperature and precipitation observations from the Narayani River catchment was used.
3.3 Model forcings and validation datasets

Figure 3.5: (A) station altitude (temperature and precipitation) plotted against the hypsographic curve for the Narayani river catchment, and (B) station average daily mean temperature and annual mean precipitation (for the period 2000-2009) plotted against altitude. Station name and number are provided in Appendix A.

Observed data obtained from the DHM over the Narayani catchment show that all observation stations are located below 4000 m asl covering only 60% of the total catchment area (Fig. 3.5A). The temperature and precipitation observations at higher elevation are missing, which might add some uncertainties to the current study. To get the overview of mean temperature and precipitation distribution over the basin, the mean daily temperature and mean annual precipitation for each station are plotted and shown in Figure 3.5B. Similar to the previous study on the Narayani river catchment (i.e. Panthi et al. (2015); Chand et al. (2019)), the temperature is decreasing with altitude. Unlike temperature, precipitation up to 2000 m asl is somewhat increasing but not in a distinct way, but after 2000 m asl precipitation decreases significantly. This significant decrease in precipitation is likely due to the orthographic effect of high mountains, as mention by Nayava (1974).

Climatic characteristic of the Narayani River basin

Monthly average forcing data sets over the Narayani river catchment were plotted and shown in Figure 3.6. From the plots, significant differences between different forcing datasets are seen. The average temperature from the CORDEX-SMHI, ERA-I, and WFDEI forcing datasets followed similar trends, but observed average temperature was
higher all over the year. The relatively higher observed average temperature was due to a lack of the temperature observation stations at higher altitudes (Fig. 3.5).

Precipitation was notably different between the forcing data sets. A significant deviation from the observation in precipitation was observed during the months of July-September. Notably, higher precipitation was observed for ERA-I and CORDEX-SMHI as compared with WFDEI and the observed precipitation data. Monthly averaged wind speeds and incoming short-wave radiation from WFDEI and ERA-I were fairly similar but different for CORDEX-SMHI.

Figure 3.6: Narayani river catchment average (2000-2009) monthly mean metrological variables for the periods 2000-2009) for each forcing data sets in Narayani River catchment.
3.3 Model forcings and validation datasets

Relative humidity from WFDEI followed the similar seasonality with precipitation, but from CORDEX-SMHI lower relative humidity is seen during August as compared to WFDEI and ERA-I. Although we did not have the station observed relative humidity, wind speed, and shortwave incoming global radiation; studies carried out by Iizumi et al. (2014) and Weedon et al. (2014) indicate that the WFDEI represents observation better.

Validation data sets

Streamflow: Discharge series of the Himalayan river stations (Narayani river gauge station, and Marsyangdi-2 river gauge station) are available from the Department of Hydrology and Meteorology (DHM), Government of Nepal (GoN). In Figure 3.7, monthly average discharge observed at Narayani river gauge from 1963 to 2015 was plotted. The Narayani river gauge is the main outlet of the Narayani River catchment. The mean annual discharge at Narayani river gauge station was found to be 1553 m$^3$/s. The timing of discharge coincides closely with seasonal maxima and minima of precipitation in the catchment. Discharge maxima generally occurred in August, coinciding with the peak of the monsoon. About 75% of the annual volume of water leaves the catchment during the monsoon season (June-September). Minimum values occurred during February-March. This observed discharge data was validated with the simulated discharge from the hydrological models.

![Figure 3.7: Monthly average discharge (1963-2015) observed at Narayani river gauge station.](image)
Moderate Resolution Imaging Spectroradiometer (MODIS) data: For the validation of model AOD from an empirical model, AOD observed from the MODIS data set was used. Generally, there are two MODIS aerosol data products provided by Terra (MOD04_L2) and Aqua (MYD04_L2) satellites. In this study, daily aerosols from Terra satellite, Level_2 (Collection 6) with spatial resolution of 10 km x 10 km (at nadir) were used. It provides global AOD from the deep blue (DB), and dark target (DT) algorithms. In this thesis, AOD retrieved with DB algorithm was used for the validation of AOD estimated from the model and the observation. The seasonal average AOD (2002-2015) over Nepal is plotted in Figure 3.8. From the figure, AOD during the witter and pre-monsoon was found higher than Monsoon and post-monsoon season. It is important to note that from the MODIS, it is not able to measure AOD over the Himalayan region of Nepal.

Figure 3.8: Seasonal averaged MODIS AOD (2002-2015) over Nepal.

Similarly, the simulated snow cover area from the hydrological model was validated with the eight-day MODIS snow product. MODIS generally provides two snow products from two different satellites, i.e., Terra (MOD10A1) and Aqua (MYD10A1). Both products contain snow cover, snow albedo, fractional snow cover, and quality assessment (QA)
3.3 Model forcings and validation datasets

data in a compressed HDF-EOS format along with corresponding metadata. Aqua and Terra products are with 463.3 m spatial resolution in sinusoidal projection. Only snow cover information is used for the current validation purposes. To minimize the effect of clouds on satellite derived snow images, composite data were formed from the Aqua and Terra satellites. A composite data were formed by following the methods described by Muhammad and Thapa [2019]. Finally, validation of model-simulated snow was validated with a composite snow product. Figure 3.9 shows the snow for four different dates representing four different seasons. From the plot, a higher snow cover area was observed in winter and pre monsoons seasons.

Figure 3.9: A composite MODIS snow cover area over Nepal. In figures green colour represents snow.
AERONET AOD: Estimated AOD from the model was also validated with the observation from the AErosol Robotic NETwork (AERONET) available in Nepal. AERONET is a federation of ground-based sun photometers which derive total column AOD and other aerosol characteristics based on a radiative transfer inversion algorithm. It provides measurements, with every 15 min, on AOD with an accuracy of ±0.01-0.02. There are eight AERONET stations in Nepal (Fig. 3.10). AERONET station at Pokhara (2010-2016) and the EVK2CNE (2006-2015) has longest time series AOD observation, while AOD data from other stations are for relatively short periods.

Figure 3.10: AERONET measured AOD over the different sites in Nepal

Catchment characteristic data

In this study, a digital elevation model (DEM) of 90 meters spatial resolution from NASA’s Shuttle Radar Topography Mission (NASA-SRTM) is used. The NASA-SRTM DEM is freely available for download from https://eros.usgs.gov/. Two land cover maps providing forest, lake, glacier, and reservoir cover of the Narayani river catchment were extracted from the MODIS land cover data and Global land Cover products (GlobCover). The
3.3 Model forcings and validation datasets

Spatial resolution of MODIS land cover data is 500m x 500m and available free online at http://glcf.umd.edu/data/lc/. Similarly, the resolution of the GlobCover product is 300m x 300m and freely available for research from http://due.esrin.esa.int/page_project68.php.
4 Overview of the papers

4.1 Summary of paper 1

Title: Aerosol Optical Depth over the Nepalese cryosphere derived from an empirical model

Atmospheric aerosols arising from a variety of natural and anthropogenic emission sources are known to affect the hydrological cycle and climate through a modification to the radiative budget of the atmosphere and the surface below through scattering and absorption of short-wave (SW) and long-wave (LW) radiation (the aerosol’s direct effect). Aerosols received much attention in the Himalayan region because of their effect on the regional as well as the local climate. Recent studies have pointed out that aerosols, particularly black carbon transport and deposition, can significantly affect the Himalayan cryosphere by modifying snow/ice albedo and thus altering the snowmelt rate, with implications for the regional hydrological regimes and the availability of freshwater. Alterations in water availability would have an impact on water resource management, hydropower generation, and agriculture in the downstream regions, possibly affecting the living conditions of more than a billion people.

Despite its substantial contribution towards radiative and climatic effects, spatio temporal distributions of aerosols in the Himalayan regions are poorly understood and are mainly due to the lack of observation. AErosol Robotic Network (AERONET) provides highly accurate, ground-truth measurements of the aerosol optical depth (AOD). The installation of AERONET station all over the Himalayan region is complicated, unpractical, and expensive. Alternatively, remotely sensed data from satellites has the potential to provide AOD on global as well as on a regional scale. However, observations from the satellite are also limited in the Himalayan region, mainly due to the high surface reflectance of the snow. To fill the data gaps in the cryospheric portion of the Himalaya, we have conducted a multivariate regression analysis to estimate the AOD over the Himalayan region of Nepalese Himalaya. Regression model was developed using three meteorolog-
4 Overview of the papers

cal variables from ERA-Interim: relative humidity (RH), wind velocity components (U10 and V10) independent variables, and the longest time series AOD observation at Pokhara station is used as a dependent variable.

The developed regression model’s coefficients were found significant at the 95 percent level with 0.53 coefficients of determination for daily values. Simulated AOD from the model is correlated and validated with the AERONET and MODerate resolution Imaging Spectrophotometer (MODIS) observation data. The correlation coefficients between model output and AERONET observations were found higher than 0.68 for all validation stations. However, the model overestimate AOD at EVK2CNR and Jomsom AERONET stations while slightly underestimates in KTM-UN, KTM-BO and Lumbini station, respectively. Similarly, simulated AOD over Nepal was also correlated with corresponding MODIS AOD observations. The correlation coefficients for most of the grids (Fig. 4.1B) were found above 0.5, with some exceptions. It was observed that the correlation coefficients were higher at the lower elevation region of Nepal. As the model was developed by using AERONET data, the systematic error between MODIS observation and AERONET observation also leads to a lower correlation coefficient. Seasonal AOD analysis shows that from both model output and MODIS observation, the highest AOD was observed during winter and pre-monsoon season, while we observed the lowest AOD during the monsoon and post-monsoon seasons.

![Figure 4.1: (A) Spatial average dry AOD (2000-2015) from daily model simulation over whole Nepal. (B) Correlation coefficient between M1 output and AOD from MODIS satellite for each grid cell over Nepal.](image)

The result presented in the paper supports the use of a multiple regression model yields good estimation for daily average AOD over the Himalayan region. Results showed that
the presented model could simulate AOD over the cryospheric region of Nepal, showing the presented model could be an efficient AOD prediction tool for the Himalayan region. The model is especially useful for the complex terrain with minimal observed data and will help to enhance the understanding of aerosol and its impacts on the Himalayan hydrology. We have only tested the model to simulate AOD over Nepal, but it can also be used for other cryospheric regions.

4.2 Summary of Paper 2

Title: Evaluation of global forcing data sets for hydropower inflow simulation in Nepal

Nepal has been endowed with immense hydropower generation potential due to glaciers in the Himalayas, regular monsoon rain, and local topography. However, climate change and consequent changes in the hydrological regime pose a great challenge for the future of sustainable hydropower development in the region. Effective assessment and implementation of hydropower projects or other water resource management projects mainly depend on hydrological simulation. Realistic hydrological simulation cannot be produced without near real-time or historical data. With a highly complex topography of Himalayas, the region has a great challenge for accurate measurement of hydro-meteorological variables; hence the region is characterized as data scarce and is a major impediment to harnessing the full potential of available resources. Several global and regionally distributed datasets hold great promise for hydrologic prediction by providing a more comprehensive spatial and temporal coverage. Most of the available global and regionally distributed data sets for hydrologic simulation are not bias-corrected (for example, Fig. 4.2 for precipitation) and can be challenging for the modeler. Despite data scarcity in the Himalayan region, use of the reanalysis and the regional model datasets for hydrological modelling, particularly in Nepal, are limited.

The main purpose of this research was to assess the quality of hydrological simulation, particularly for the place that is typically difficult to model based on different forcing data sets. Furthermore, the specific objective of this research is to evaluate different forcing datasets for inflow hydrologic modelling. In this study, discharge over the Narayani River catchment of Nepal is simulated using Shyft (Statkraft’s Hydrologic Forecasting Toolbox) forced with observations and three global forcing data sets: i) ERA-Interim (ERA-I), ii) Water and Global Change (WATCH) Forcing Data ERA-interim (WFDEI), and iii) CO-
4 Overview of the papers

ordinated Regional Climate Downscaling EXperiment with contributing institute Rossy Centre, Swedish Meteorological and Hydrological Institute (CORDEX-SMHI).

![Average precipitation from different forcing data sets.](image)

Figure 4.2: Average precipitation from different forcing data sets. In the figure inverted bar diagrams are for daily average seasonal rainfall (2000-2009). Ten year average accumulated rainfall from each product is plotted as dash line plot while shadings are for daily deviation from the mean of ten year daily precipitation data.

Results show the successful application of global forcing data sets in regional catchment-scale modelling. Results are also promising for capturing the inter-annual variability in discharge. This study showed that significant biases in precipitation could be reduced by applying a precipitation correction factor as a calibrating parameter in a model. However, the best result was obtained using bias-corrected forcing data as input to Shyft, i.e., the WFDEI performed better than the others. Accordingly, the WFDEI forcing dataset holds great potential for predicting and improving understanding of the hydrology of data-sparse Himalayan regions, whereas the use of CORDEX-SMHI and ERA-I derived data requires further validation, bias correction, particularly over the high mountain regions.

Therefore, based on this current analysis, we conclude that in the data-poor Himalayan catchment, WFDEI forcing dataset could be the best alternative to the observed forcing datasets for the water resource planning and hydropower estimation. If the model is not constrained with the application of precipitation correction factor, then the ERA-Interim forcing data set could be an alternative for hydrological simulation at daily time steps.
CORDEX-SMHI products relatively capture higher peaks so that it could be better for peak-flow analysis. Moreover, this study will be useful in validating the use of reanalysis and regional model data sets in the data-scarce catchment. Further investigation by implementing Shyft in the different Himalayan catchments for different forcing data sets could be the next potential step to check the validity of the current results, which is currently out of our scope.

### 4.3 Summary of Paper 3

**Title: Impact of catchment discretization and imputed radiation on model response: a case study from central Himalayan catchment**

Accurate runoff prediction is one of the fundamental challenges for effective and sustainable water resource management in the Himalaya region. Hydrological models are among the available tools used for a better understanding of the hydrological processes and their interaction. They describe processes and allow predicting the results of making management decisions, and are an important water resource management tool, especially for the Himalayan catchment. However, the application of hydrological models (mainly distributed models) fundamentally requires watershed partitioning or discretization. Partitioning the watershed into modelling elements represent an abstraction of the actual watershed surface and its relevant hydrological properties. Distributed and semi-distributed hydrological modelling approaches commonly involve the discretization of a catchment into several modelling elements. However, a crucial issue that must be addressed by any user of these models prior to their application in the definition of an acceptable level and type of watershed discretization. In previous studies, although some simulations were conducted using triangulated irregular networks (TIN), a little attention has been given to assess the impact of TIN as compared to the standard catchment discretization techniques.

In this study, we examine how different catchment discretization approaches and radiation forcings influence hydrological simulation results. Three catchment discretization methods, i.e., elevation zones (HYP), regular square grid (SqGrid), and TIN, were evaluated in central Himalaya catchment (Marshyangdi-2), Nepal. To evaluate the impact of radiation on the hydrological model response incoming short-wave radiation was imputed using two approaches, one with the measured incoming solar radiation assuming a horizontal plain surface and another with a translation to sloped surface. The results
presented in the study show that the catchment discretization has a significant impact on simulation results. Evaluation of the simulated streamflow value using Nash-Sutcliffe efficiency (NSE) and log-transformed Nash-Sutcliffe efficiency (LnNSE) shows that the highest model performance was obtained when using TIN followed by HYP (during the high flow condition) and SqGrid (during the low flow condition). A similar order of precedence in relative model performance was obtained both during the calibration and validation periods. Better representation of glacier area (Fig. 4.3) into each modelling elements could be one of the reasons for better performance obtained from TIN based discretization.

Snow simulated with TIN based discretized models were validated with MODIS snow products. Critical Success Index (CSI) between MODIS snow and model with and without using translated radiation was found satisfactory. Higher CSI were found during winter and pre-monsoon season, while lower CSI were observed during the monsoon season. No significant differences in the NSE and LnNSE were observed for the model with and without using translated radiation. However, notable difference in snow cover area were observed from the two radiation approaches.

Figure 4.3: Cell area distribution with respect to aspect.
5 Discussion

5.1 General discussion

In the remote Himalayan region, there is a high demand for high-quality hydro-meteorological data sets for the water resource management problem. Accurate and reliable hydrological data are crucial not only for the study of hydrological variability but also for the management of water resources and hydrological forecasting. On-site observation produces an accurate measurement, but are not readily available because of remote, high relief, and inaccessible mountains in the Himalaya. Various large-scale climate models and remote sensing dataset are the alternative for the observations but these datasets have several limitation, such as coarse resolution, significant biases, and unable to measure on some specific area. Hence the selection of climate model and remote-sensing datasets with an appropriate complexity and resolution are most required to achieve reliable hydrological simulation over the Himalayan region. In this section, I will provide a general discussion on the main outcomes of current research presented in this thesis.

5.1.1 Aerosol optical depth (AOD) estimation over the cryospheric portion of Nepalese Himalaya

The hydrological cycle is driven primarily by the energy from the sun. The location and atmospheric conditions determine the amount of sunlight received on the Earth’s surface. Globally, aerosols are spread all over the world; it has nonetheless a strong regional imbalance, which generally results in modification of climate through their direct and indirect effects on radiative forces and condensation particles. Accordingly, the geographical location of Nepal in south-east Asia makes aerosol pollution a familiar phenomenon in which its concentrations exhibit substantial seasonal variability, mainly driven by seasonally changed air mass patterns, monsoon, and winter seasons. Considering the significant impact of aerosols in the regional hydrology, it’s accurate prediction is most important for realistic hydrological simulation [Matt et al., 2018]. Although some previous works
5 Discussion

e.g. Nair et al. (2013); Lau et al. (2010) were conducted over the Himalayan region at the regional scale, little attention has been given to estimating the impact of aerosol on snow and glacier melt. This is primarily due to the limited aerosol observation data from the Himalaya region. However, satellite remote sensing provides data on aerosol optical depth (AOD), the snow and glacier-covered surfaces are identified as a main challenge for aerosol retrieval from remote sensing.

To overcome the data limitation problem over the Himalayan region, we have presented a novel approach to estimate AOD. The rationale for developing a model for the AOD over Nepal is the failure of satellites to provide accurate measurements, in particular over bright surfaces such as snow and ice. As mention by many researcher e.g. Ramanathan (1998); Govaerts et al. (2009); Remer et al. (2005), AOD is one of the main parameters to determine the aerosol load on the atmosphere. The results presented in the first paper suggested that the model-based simulation of AOD has merits for a better understanding of aerosols both on local and regional scales. However there are some uncertainty in estimation. For e.g. while developing model, observed AOD was converted in to dry-form, and is likely to introduce uncertainty into the model because the hygroscopic growth factor depends on the chemical composition of the aerosol particles which is widely vary over the study area (e.g. westerly winds transporting dust whereas southerly winds transport polluted aerosol, in addition to local sources from e.g. combustion processes). As mention by (Remer et al., 2005) identification of aerosol source and its inclusion into the model is important for the development of aerosol retrieval model.

As this study focus on local scale, greater impact of aerosol on local level were already reported by many scholars e.g. Ramanathan (1998); Safarpour et al. (2014); Sić et al. (2015). Aerosol not only has an impact on the extinction of solar radiation but also affects atmospheric long-wave radiation, which is a key driver of the local energy balance of a place. Changes in the local or regional energy balance cause change in the regional water budget. The model developed for AOD estimation in the study was not aimed at representing the best physically-based model. Although the presented results were good, uncertainties remain. The model is aimed at providing AOD data over the Nepalese Himalayan region, which will help to address existing uncertainties in climate research due to data scarcity. Further inclusion of AOD estimated for the snow and glacier region into the hydrological region could improve hydrological simulation and forecasting (Yoshioka et al. 2019). Again, this present study has contributed to the research knowledge on the
5.2 Challenges and opportunities for the use of remote-sensing and global forcing data over the Himalayan region

Hydrological modelling in the Himalayan region suffers from multiple issues that affect our ability to represent the hydrological dynamics in the region. Due to a lack of observation station, especially in the high-altitude areas, the spatial distribution of metrological variables are unknown. Temperature, precipitation, humidity, wind, and radiation provides the drivers and input for the land surface and subsurface water fluxes. Study by Fan et al. (2019) show the importance of topography for hydrological simulation, where topography will have impact on incoming radiation. An accurate representation of the spatio-temporal variability of these forcing variables including incoming short-wave radiation is thus essential for achieving good simulation over the region. As we know, global incoming radiation received at the surface not only affected by the slope but also by the atmospheric constituent present in the atmosphere. Hence its accurate estimation is necessary. Remote sensing and global climate forcing datasets provide an opportunities for the hydrological modelling in the data-scarce region across the Himalaya, but the quality of model forcing remains a challenge to the researcher. Similarly, in terms of aerosol and its properties estimation from the remote sensing, there are again some problems. To overcome this limitation and problems, the AOD estimation model presented in the thesis could be a valuable tools for the hydro-metrological simulation over the Himalayan region.

Hydrological simulation over the centre Himalayan catchment results presented in our second paper show that the selected hydrological model (i.e., PT_GS_K from Shyft) appears sensitive to the choice of the global forcing data leading to different levels of model efficiency in terms of discharge simulations. Similar level of sensitivity were also presented by Sapkota (2016) in his research. Determining the factors explaining these sensitivity within the forcing data sets is not easy as different forcing datasets have different characteristics such as the resolution, the data assimilation method, the number
5 Discussion

of stations used to correct predicted global data. However, as mention by De Wulf et al. (2012), a good representation of topography by forcing variables leads to better model efficiencies. Precipitation is notably different among the forcing data sets, most probably because of the difficulty of capturing the spatial distribution of precipitation compared to continuous variables like temperature and short wave incoming radiation (Raimonet et al., 2017). Studies carried out by Larsen et al. (2016) also revealed that precipitation has a major role in simulating discharge from distributed hydrological modelling.

In the presented study, the performance of the selected model (i.e. PT_GS_K) during the model calibration and validation showed that all forcing datasets (WFDEI, CORDEX-SMHI, and ERA-I) perform well, although significant differences are observed between precipitation datasets (Fig. 5.1). Despite differences in precipitation, good agreement between simulations and observations during the model calibrations and validations were because of the application of a precipitation correction factor. Similar results were also reported by Bisselink et al. (2016); Naseer et al. (2019) etc, where the performance of the bias corrected data performed better then the rest. However, by applying a multiplicative correction factor (i.e. precipitation correction factor) to the precipitation, a significant amount of bias in discharge estimation could be removed. Precipitation correction factor for WFDEI is near to 1, likely due to the fact that precipitation in WFDEI is bias-corrected. But the precipitation correction factor for observed datasets was 1.34, indicating insufficient observation station causing under-catch precipitation amounts. Not only the fewer observation stations in the studied catchment cause under-catch, wind, type of rain-gauge, measurement methods are also responsible for precipitation under-catch (Mekonnen et al., 2015). Similar to this study, better performance in discharge simulation by the application of scale factor to the precipitation is also shown by previous studies e.g. Lawrence et al. (2009); Lakew et al. (2017). Although we achieved suitable error statistics (higher NSE and KGE with lower Dv) from all forcing data sets, still all model forcing simulations were not able to capture the peaks as seen in the observed discharge data. However, CORDEX-SMHI captures relatively higher peaks as compared to other forcing datasets but produces a higher positive bias overall in the simulation.

Overall from the thesis, I demonstrates the successful application of global forcing datasets in the regional catchment scale for the hydrological modelling over the Himalayan region. This study explores the opportunities for using global forcing data sets for hydrological simulation in the data-scarce Himalayan region. WFDEI forcing data sets may be
5.3 Catchment discretization, model forcing, and hydrological response

Several previous studies e.g. De Wulf et al. (2012); Kolberg et al. (2006) highlighted the importance of spatial discretization on catchment response. As mention by the Fan et al. (2019), in the hydrological models, the topography is assumed to be a dominant control of the hydrological process. There is a link between hydrological model efficiency and catchment characteristics such as catchment size, mean catchment elevation, and slope. Singh and Fiorentino (1996) show that the catchment discretization types and size of modelling elements has significant impacts on model efficiency. In this regard, we have evaluated three spatial catchment discretization methods based on hypsography (HYP), regular square grid (SqGrid), and triangular irregular networks (TIN). In previous studies such as Matt et al. (2018); Matt and Burkhart (2018); Hegdahl et al. (2016); Bhattarai and Regmi (2015), hypsography and square grid-based catchment discretization methods were widely used for the hydrological simulation in the Himalayan region. To date, TIN based spatial discretization method has not been tested in Nepalese Himalayan catchments. There are different ways to create TIN, but the Delaunay triangulation is a widely appreciated and investigated mathematical model to represent the topography and
is highly efficient for building TINs. Vivoni et al. (2004) mentioned that in the regions of high terrain variability, the catchment discretized with TINs could be modelled more precisely due to the higher flexibility of the mesh. TINs also allows a more continuous description of stream paths and networks in conjunction with the topography.

Theoretically, the ability to account for spatial-temporal variability of hydro-meteorological forcing and physical features within a catchment should lead to better simulations. Results presented in our third paper shows that all discretization methods provide an excellent representation of the general flow pattern and the overall water balance components while maintaining the significant inter-annual variability. The highest performance efficiencies in terms of NSE and LnNSE were obtained for TIN based catchment discretization method. Similar to the prior studies by Reed et al. (2004) and Smith et al. (2004), we found that grid-based distributed modelling approaches do not always provide improved discharge simulation as compared to hypsography discretized conceptual models. The poor performance of the grid-based models was partly attributed to the complex topography of the region and might be because of hydro-meteorological data interpolation scheme in to the selected model.

Most of the distributed hydrological models use radiation measured on the horizontal plane surface for the hydrological simulation. But as Fan et al. (2019) shows that the radiation received on the surface depends on aspect and slope and may have impact on discharge simulation. Currently several approaches were presented but the in this thesis, approach by Allen et al. (2006) was used to translate radiation on the slope surface. After implementation, as expected, we found that the more radiations are distributed to the surface facing toward the south-east. Finding presented in our third paper demonstrates that the application of translated radiation into the model have a significant impact on the snow simulation; although the model efficiency in terms of discharge simulation was not significantly different. Since the model is a heavily parametrized calibrated model, the impact of one parameter could be easily compensated by others. That could be the main reason for getting insignificant differences in model response after implementing translated radiation into the model.
6 Conclusion and recommendations for further works

6.1 Conclusion

The research work presented in this thesis was focused on estimation, analysis, and evaluation of hydrological model forcing data sets over the Himalaya. Three paper presented in the thesis are representing the three aspect of the thesis. Including aerosol properties and calculating its impacts on the hydrological simulation was the primary goal. But with limited time, integrating AOD part in the model is still missing and could be a potential topic for further study. This section presents the conclusions drawn in light of the three main objectives of the work.

The main outcome of our first objective is the development of an empirical model for the estimation of AOD over the cryospheric portion of Nepal. Results showed that the presented model is able to simulate AOD over the different regions of Nepal, indicating model adequacy and establishing the model as an efficient AOD prediction model for the Himalayan region. In the presented model, emission flux and sources were not considered while developing the model. Identifying the source of the aerosol is beyond the scope of this study. The results presented in our first paper confirm that the use of even the simplest linear regression model yield very good estimation results for daily average dry AOD data in Nepal. These simple models are indispensable prediction tools for scientists requiring AOD information in the data-scarce Himalayan region. It is evident that the model from multivariate regression analysis has universality in statistics, and it can be able to predict most of the spatial AOD variability in the Himalayan regions. It is especially useful for a situation where we do not have explicit knowledge about AOD. This includes modelling in very complex terrain where minimal observed data sources exist. Our model only uses ASTER DEM and meteorological data from ECMWF. Therefore, these models could be easily applied to other regions with the mountain environment and
in related climate research in the mountain region.

To attain the second objective, we have evaluated the quality of hydrologic simulation based on different forcing data sets for runoff simulation on a daily scale over the Narayani River basin located in central Nepal, from January 2000 to December 2009. Forcing datasets represented by WFDEI, CORDEX-SMHI, ERA-I, and ground-based observation. Amongst forcing, only WFDEI data sets are bias-corrected. In the study, PTGSK from Shyft is used as a simulation tool. Prior to the hydrological simulation, significant differences in precipitation were observed between different forcing data sets. To deal with the bias present in precipitation, a precipitation correction factor was used in the model as a calibrating parameter. Relatively higher model performance in terms of NSE, KGE with lower biases are found for the Observed+WFDEI. The second-best performance was observed for the model with WFDEI forcing data sets.

Based on current results, we have concluded that, in the data-poor Himalayan catchment, WFDEI forcing dataset could be the best alternative to the observed forcing data sets for the water resource planning and hydropower estimation. If the model is not constrained with the application of precipitation correction factor than, ERA-Interim forcing data set could be another option for hydrological simulation at daily time steps in the Himalayan river basin. We observed that the simulation from CORDEX-SMHI forcing data sets relatively captures higher peaks as compared with other simulations. Therefore CORDEX-SMHI forcing data sets could be a better option for peak-flow analysis in the Himalayan catchment. But It should be noted that if the catchment is un-gauged than ERA-I and CORDEX-SMHI should be bias-corrected before using it into the model as a model forcing. Since this study only evaluated forcing data sets in one catchment located in centre Nepal so, further analysis in different Himalayan catchments could be the next potential step to check the validity of the current results, which is currently out of or scope.

Similarly, to attain the third objective, we introduced three catchment discretization approaches; (1) hypsography (HYP) approach, (2) square grid approach (SqGrid), and (3) triangulated irregular network (TIN) approach, for hydrological simulation in the Marshyangdi-2 river catchment. It is a sub-catchment of the Narayani river catchment. To evaluate the impact of radiation on the model response, shortwave radiation was translated using two approaches, one with the measured solar radiation assuming a horizontal surface and another with a translation to slopes. Results presented in our third paper
shows the model performance according to the Nash-Sutcliffe efficiency (NSE) and the log-transformed Nash-Sutcliffe efficiency (LnNSE) of the outlet discharge, with the rank of discretized models efficiency in the descending order being TIN, HYP, and sqGrid both in the calibration and validation mode. Critical success index (CSI) based on MODIS snow and model-simulated snow results show the validity of snow simulation for the model. A significant difference between the snow simulation was also observed with and without using translated radiation into the model.

From our analysis, we concluded that TIN based discretized models are a priori the preferred option for hydrological simulation in the Himalayan catchment. The selection of the catchment discretization methods depends upon the types of the study area and required level of model efficiency. As regular SqGrid and HYP based discretized, models are less flexible, but offer higher speed, lower memory requirements, and easier implementation algorithms as most important assets, making them be preferred when the studied area is flatter.

### 6.2 Recommendations for further works

Results presented in the thesis are focus on the Himalayan region, where finding in the thesis provide valuable new regional knowledge and contribution. Based on the current results, we have some recommendations for further works:

1. The bias-corrected WFDEI dataset and the triangulation grid discretization methods are recommended as the best performing methods for future water resources-related model application in the Himalayan region.

2. The model presented in our first paper included in the thesis is the first application of an empirical model to estimate AOD over the cryospheric portion of Nepal. Model validation is limited to Nepal. It is recommended to validate the model in other parts of the Himalayan region.

3. We have evaluated the hydrological simulation only from the reanalysis and CORDEX-SMHI forcing data sets. It would be nice to compare the simulation results from other globally available forcing data sets.
6 Conclusion and recommendations for further works

4. Spatial catchment discretization methods were evaluated in the Himalayan catchments with steep descent/ascent terrain. It would be interesting to follow the same experiment and compare the results from the more flat catchment.

5. Integrating aerosol properties and its impact on hydrological modelling is still missing and will be the potential future work.
References


Atif I, Iqbal J and Su Lj (2019) Modeling hydrological response to climate change in a data-scarce glacierized high mountain astore basin using a fully distributed topkapi model. Climate, 7(11), 127


References

Beven K (1979) A sensitivity analysis of the Penman-Monteith actual evapotranspiration estimates. *Journal of Hydrology*, 44(3-4), 169–190, ISSN 00221694


Brun E, Durand Y, Martin E and Braun L (1994) Snow modelling as an efficient tool to simulate snow cover evolution at different spatial scales. *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences*, 223, 163–176


of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137**(656), 553–597, ISSN 00359009


Freer J, McDonnell J, Beven K, Brammer D, Burns D, Hooper RP and Kendal C (1997) Topographic controls on subsurface storm flow at the hillslope scale for two hydrologically distinct small catchments. *Hydrological Processes (United Kingdom)*


Ghimire S, Choudhary A and Dimri A (2018) Assessment of the performance of cordex-south asia experiments for monsoonal precipitation over the himalayan region during present climate: part i. Climate dynamics, 50(7-8), 2311–2334


63


Immerzeel WW, van Beek LP and Bierkens MFP (2010) Climate change will affect the Asian water towers. Science (New York, N.Y.), 328(5984), 1382–5, ISSN 1095-9203


IPCC (2007) Intergovernmental panel on climate change: Climate change 2007. IPCC Secretariat Geneva


Jinkang D, Shuping X, Youpeng X, Xu CY and Singh VP (2007) Development and
testing of a simple physically-based distributed rainfall-runoff model for storm runoff

Precis. Office, 40

Temperature and Precipitation Data Between Automatic Weather Station and Manual
Observation. Technical report, IOM-104, Kathmandu, Nepal

and Course Materials for Teaching Numerical Modelling in the Environmental Sciences.
Transactions in GIS, 5(2), 99–110, ISSN 13611682 (doi: 10.1111/1467-9671.00070)

Useful for the Himalayan River Basins. In Himalayan Weather and Climate and their
978-3-030-29684-1_19)

Kayastha RB, Ohata T and Ageta Y (1999) Application of a mass-balance model to
a Himalayan glacier. Journal of Glaciology, 45(151), 559–567, ISSN 0022-1430 (doi:
10.3189/s002214300000143x)

Hydrology of Koshi River Basin. Journal of Hydrology and Meteorology, 9(1), 28–44,
ISSN 1818-2518 (doi: 10.3126/jhm.v9i1.15580)

Kirchner JW (2009) Catchments as simple dynamical systems: Catchment characteri-
45(2)

scheme for snow coverage observations in a gridded snow model. Hydrology and earth
system sciences, 10(3), 369–381

R, Xu Y, You Q and Ren Y (2019) Unravelling Climate Change in the Hindu Kush
Himalaya: Rapid Warming in the Mountains and Increasing Extremes. In The Hindu
Kush Himalaya Assessment, 57–97, Springer International Publishing (doi: 10.1007/
978-3-319-92288-4_3)

Kuang X and Jiao JJ (2016) Review on climate change on the Tibetan Plateau during the
last half century. Journal of Geophysical Research: Atmospheres, 121(8), 3979–4007,
ISSN 2169897X (doi: 10.1002/2015JD024728)


Larsen MA, Christensen JH, Drews M, Butts MB and Refsgaard JC (2016) Local control on precipitation in a fully coupled climate-hydrology model. Scientific Reports, 6, ISSN 20452322 (doi: 10.1038/srep22927)


References


Nair VS, Babu SS, Moorthy KK, Sharma AK, Marinoni A and Ajai (2013) Black carbon aerosols over the himalayas: direct and surface albedo forcing. Tellus B: Chemical and Physical Meteorology, 65(1), 19738


Rana B, Nakawo M, Fukushima Y and Ageta Y (1997) Application of a conceptual precipitation-runoff model (HYCYMODEL) in a debris-covered glacierized basin in the
References


References


Part II

Journal Publications
Paper I: Aerosol Optical Depth Over the Nepalese Cryosphere Derived From an Empirical Model
Aerosol Optical Depth Over the Nepalese Cryosphere Derived From an Empirical Model

Bikas Chandra Bhattarai*, John Faulkner Burkhart, Frode Stordal and Chong-Yu Xu

Department of Geosciences, University of Oslo, Oslo, Norway

In the Himalayan region, aerosols received much attention because they affect the regional as well as local climate. Aerosol Optical Depth (AOD) observation from satellite are limited in the Himalayan region mainly due to high surface reflectance. To overcome this limitation, we have conducted a multivariate regression analysis to predict the AOD over the cryospheric portion of Nepalese Himalaya. Prediction using three meteorological variables from ERA-Interim: relative humidity, wind velocity components (U10 and V10) were taken into account for model development as independent variables, while the longest time series AOD observation at Pokhara station is used as dependent variable. Model coefficients were found significant at 95 percent level with 0.53 coefficients of determination for daily values. Correlation coefficients between model output and AERONET observations were found to be 0.68, 0.73, 0.75, 0.83, and 0.82 at Lumbini, Kathmandu Bode (KTM-BO), Kathmandu University (KTM-UN), Jomsom, and Pyramid laboratory/observatory (EVK2CNR) AERONET stations, respectively. Model overestimate AOD at Jomsom, and EVK2CNR AERONET stations while slightly underestimates AOD in Lumbini, KTM-UN, and KTM-BO AERONET station, respectively. Both model output and MODIS observation showed that the highest AOD over Nepal is observed during winter and pre-monsoon season. While lowest AOD is observed during monsoon, and post-monsoon season. The result of this research supports that the use of linear regression model yields good estimation for daily average AOD in Nepal. The model that we have presented could possibly be used in other mountain regions for climate research.

Keywords: Himalaya, MODIS aerosol optical depth, AERONET aerosol optical depth, empirical model, cryosphere

1. INTRODUCTION

Aerosols are a focal point of climate research due to their role, and significant uncertainty, in atmospheric processes. Atmospheric aerosol particles scatter, reflect, and absorb incoming solar radiation (as a direct effect) (Chylek and Wong, 1995; Solomon et al., 2007), and modify cloud properties (as an indirect effect) (Charlson et al., 1992; Kim et al., 2014). The uncertainty associated with these processes thusly is considered as one of the huge gaps in current climate prediction capabilities (Parry et al., 2007; Istomina et al., 2011; Alexandrov et al., 2016). Considering the significant role of aerosol in climate processes in the Himalaya (Ramanathan, 2001; Meehl et al., 2008; Nair et al., 2013), different studies have evaluated this region (Ramanathan and Ramana, 2005; Ramanathan et al., 2007; Srivastava et al., 2012) focusing on aerosol emissions,
optical–physical properties, and its climatic implications (Tripathi et al., 2007; Srivastava et al., 2012; Lau, 2014; Soni, 2015; Paliwal et al., 2016; Zhang et al., 2017) as well as impacts for regional hydrology (Matt et al., 2018). These research showed that the aerosols over the Himalayan region are in increasing trend, which is mainly detected during the winter and post-monsoon seasons and are forced by the high anthropogenic emissions, composed of bio and fossil fuel combustions (Acharya and Sreekesh, 2013). Ramanathan and Carmichael (2008) state that aerosols (particularly black carbon) in the high Himalayas likely play significant role in the snow and glacier melt by increasing solar heating. Li et al. (2016) claim the Himalayan region should be considered as the most vulnerable due to the impact of black carbon. Aerosol deposition and its transport over the Himalaya is attracting more attention due to its impact on the transformation of hydrological processes, and regional energy balance, affecting billions of people living downstream (Nepal et al., 2014).

Satellites offer a global perspective on many atmospheric variables, including AOD (Kaufman et al., 2002). Remotely sensed data from satellites has potential to account the highly variable black carbon aerosol properties on global as well as on regional scales and to provide repeated observations over long periods. A well-known example is the MODerate Resolution Imaging Spectroradiometer (MODIS) instrument which can provide daily aerosol and its different properties with nearly global coverage at the resolution of 10 and 3 km (Remer et al., 2013). Several works provide an overview of MODIS aerosol retrieval algorithms and products (Kaufman et al., 1997; Chu et al., 2002; Remer et al., 2005; Martonchik et al., 2009; de Leeuw et al., 2011). The basis of MODIS AOD retrievals is that two independent algorithms are used to derive aerosol, one over ocean, and a second to derive over land. The land algorithm is mainly based on the dark target approach (Kaufman et al., 1997; Remer et al., 2005). However, there are some limitations over brighter surface. In both Govaerts et al. (2010) and Mei et al. (2012) the snow and glacier covered surfaces are identified as a great challenge for aerosol retrieval from remote sensing due to the fact the high surface reflectance makes it difficult to separate radiation at the top of atmosphere due to reflection from the snow and from atmospheric scattering by aerosol particles. As Mei et al. (2012), indicate that the crucial issue with using satellite for AOD retrieval over brighter surface is due to very high spectral albedo of the brighter surface like snow at wavelengths in the visible region.

To fully understand the effect of aerosols over the Himalayan region, detailed knowledge regarding the spatio-temporal distributions of aerosols, and their seasonal variability in the atmosphere are required (Bonasoni et al., 2012). Several methods have been used to retrieve AOD over pure snow (Istomina et al., 2009; Mei et al., 2012, 2013), but all these algorithms are restricted to the Arctic region in order to meet the requirement of having a sufficient snow BRDF model (Mei et al., 2013). To date, no algorithm exists to retrieve AOD products over Himalayan cryospheric region (snow and ice surfaces). A more detailed understanding of spatial, and temporal variations of aerosols is required in order to quantify the dynamic influence on the regional climatic conditions.

The objective of this research is to develop an empirical proxy model by using multiple regression, to increase the present understanding of spatio-temporal variability of AOD over the cryospheric portion of Nepal. Three meteorological variables from ERA-interim reanalysis dataset: relative humidity, wind velocity components (U10 and V10) (describe in section 3.3) and observed AERONET AOD from Pokhara AERONET station are used to develop our proxy en empirical model. Our study region is presented in section 2, while the dataset used in this research, and the proposed methods to retrieve AOD is explained in section 3. Results and discussion are presented in section 4, and finally conclusions are presented in section 5.

2. STUDY AREA

The domain of our analysis is the country of Nepal (see Figure 1), with our results applicable to the cryospheric portion of the country. Nepal is between India and China, and extends 885 km east-west and 145–248 km north-south. Within this small geographical range, the altitude varies from ∼60 m above sea level (m asl.) in the southern plain, tropical Terai, to the highest peak on the earth in the northeast. Along a south-north transect, Nepal is divided into three ecological belts: Mountain in the northern range, the mid range is called Hill, and the low elevated southern range called Terai (CBS, 2014). Area of the country is 147,181 km², out of which about 15% is comprised of high Himalaya, 68% covers by mid hill regions, and the remaining 17% flat valley floor Terai. Around 50% of the total population lives in Terai region, 43% of country population lives in the Hill region, and 7% in the Himalayan region (CBS, 2012). Predominant economic (and aerosol producing) activities (farming, industrial establishment) are conducted in Terai region.

Rapid changes in elevation within a short north-south distance creates a wide range of climatic conditions, from subtropical to alpine/arctic within a span of <200 km. The temperature variation in Nepal is mainly related with the seasons. Within a season temperature varies with topographic variations along north to south direction. Eighty percent of the total precipitation in Nepal occurs during the monsoon (June to September) season (Nayava, 1974; Shrestha et al., 2000) with winter (5%) (December to February) rains more common in the western hills (Ichiyanagi et al., 2007). Pre-monsoon (March-May) season receives about 10% of rainfall while 5% of rainfall occurs during Post-Monsoon (October to November) (Nayava, 1974).

3. DATASETS AND METHODS

3.1. AERONET Data

The Aerosol Robotic NETwork (AERONET) is a federation of ground-based sun photometers which derive total column AOD and other aerosol characteristics based on a radiative transfer inversion algorithm. The network requires standardized instruments, calibration, and processing (Holben et al., 1998). AERONET stations provide measurements every 15 min from a 2.1° field of view, and eight solar
Spectral bands from 340 to 1020 nm are used to calculate, for each wavelength, the AOD, with an accuracy of ±0.01–0.02 (Eck et al., 1999). There are eight AERONET stations in Nepal. The longest time series data come from Pokhara (2010–2016) and the EVK2CNR (2006–2015) station while data from other stations are less comprehensive and for relatively short periods from different years.

The mean AOD at 550 nm from the AERONET stations at different locations is presented as boxplots in Figure 2. While we note an elevation dependence of AOD, it must be recognized some stations have limited data. Figure 2 shows that the Lumbini station located in lowest elevation has highest AOD mean followed by the Hetuda, while the lowest AOD mean value is observed at EVK2CNR at the highest altitude. Systematically the mean are greater than the median indicating that the positive skewed distributions are characteristic of naturally occurring phenomena as indicated by Sriram et al. (2004). Since the number of observed data points are limited in Hetuda and Langtang (up to 2016), further analysis does not include the data from these stations. AERONET datasets were screened for outliers, and these values were removed using a mean ± 3 standard deviation conventional approach (Miller, 1991; Leys et al., 2013). As in prior studies, to provide an effective comparison and analysis, AERONET data are interpolated from 500 to 550 nm using the following computation (Kaskaoutis et al., 2007; Prasad and Singh, 2007; Alam et al., 2014).

$$AOD_{550nm} = AOD_{500nm} \left( \frac{550}{500} \right)^{-\alpha} \quad (1)$$

where \(AOD_{500nm}\) in Equation (1) is the AOD measured in 500 nm wavelength. Here \(\alpha\) is the Ångstrom exponent from the wavelength of 440–870 nm (Sayer et al., 2013):

$$\alpha = -\frac{\ln (\tau_1/\tau_2)}{\ln (\lambda_1/\lambda_2)} \quad (2)$$

where \(\tau_1\) and \(\tau_2\) are the AOD at wavelengths \(\lambda_1\) and \(\lambda_2\).

As explained below, in section 3.5, our empirical proxy aerosol model performs better with dry than wet aerosols. To derive dry AOD from wet AOD we use the approach of Zhang et al. (2017) to account for hygroscopic growth.

$$AOD_{dry} = \frac{AOD}{f(RH)} \quad (3)$$

where, RH is relative humidity, \(AOD_{dry}\) represents the AOD with a dehydration adjustment, \(f(RH)\), the hygroscopic growth factor, denotes the ratio of the aerosol scattering coefficient in ambient with a certain relative humidity to that in the dry air condition (Li et al., 2014; Zheng et al., 2017). \(f(RH)\) can be expressed as:

$$f(RH) = \frac{1}{(1 - RH/100)} \quad (4)$$

We have tested the different hygroscopic factors given by Li et al. (2014) to convert observed AOD in to a dry state and Equation (4) performed best to yield a higher correlation of \(AOD_{dry}\) with meteorological parameters. Hereinafter, all references to AERONET AOD denote \(AOD_{dry}\) at 550 nm from AERONET, unless otherwise indicated, and wavelength subscripts are not assigned for conciseness.

### 3.2. MODIS Data

In this study we use 16 years (2000–2015) of AOD data obtained from the MODIS instrument on-board, the NASA EOS satellites.
The MODIS products provide three processing levels of data: Level 1 (geolocated radiance, and brightness temperature), Level 2 (retrieved geophysical data products), and Level 3 (gridded averages of geophysical retrievals) data. There are two MODIS Aerosol data product files: MOD04_L2, containing data collected from the Terra platform; and MYD04_L2, containing data collected from the Aqua platform. Daily Level 2 (Collection 6) data produced at the spatial resolution of a 10 km x 10 km (at nadir) from the Terra platform MOD04_L2 is used in this study. The MOD04_L2 product provides global AOD from the dark target (DT) (Kaufman et al., 1997; Levy et al., 2013), and deep blue (DB) algorithms (Huss, 2013). The DT algorithm is applied over the ocean and dark land surfaces, while the DB algorithm is used for brighter surfaces. The data analysis that follows uses AOD at 550nm to be consistent with the wavelength used by many climate transport and chemistry models (Kinne et al., 2013) and prior MODIS validation studies (Levy et al., 2007, 2010; Safarpour et al., 2014). MODIS scientific datasets (SDS) in Level 2, collection 06 is used to retrieve AOD for this study.

Table 1 provides the names for the relevant scientific datasets within the MODIS Level 2 aerosol products. In order to select the optimal MODIS scientific dataset for our purposes, we conducted a brief validation and evaluation of different scientific dataset performance in relation to the Pokhara AERONET observations. Validation of MODIS AOD with AERONET observed AOD (dry AOD from both observation) is carried out to find the best fit MODIS scientific dataset with AERONET datasets. For the validation, we followed the procedure described by Ichoku et al. (2002) and Li et al. (2009). Spatial and temporal variability of AOD distributions were taken into account. MODIS retrieval at 10 km x 10 km and AERONET measurement within ±30 min of MODIS overpass time and at least 3 out of 9 MODIS retrieval in a square box of 30 × 30 km centered over AERONET site were used. After that mean values of co-located spatial and temporal values were used for calculating error statistics (i.e., RMSE, correlation coefficient). Scatter plots between mean AERONET and MODIS AOD for different scientific dataset are shown in Supplementary Figure 1. Although different datasets have different expected error (EE), ±(0.05 + 0.15AOD_{AERONET}) is used for direct comparison (Remer et al., 2005). Validation results shows that the lowest RMSE (0.13), with highest correlation coefficient (0.75) and EE (71%) is found for aerosol optical depth estimated from deep blue algorithm. The correlation coefficient for the best estimate is higher (0.75), but MODIS has large underestimation. Large AOD underestimation by MODIS in the studied site does not affect the current study because MODIS AOD is only used in a relative sense to determine gradient along the mountain slope.

Using classical regression model evaluation statistics, we also calculated at the Pearson Correlation coefficient (Adler and Parmryd, 2010) for the different MODIS scientific dataset in two forms (i.e., dry (converted by using Equation 3) and normal) against the observations. The correlation coefficient between AERONET AOD and MODIS AOD in normal (0.75) as well as in dry form (0.87) are also highest for aerosol optical depth estimated from deep blue algorithm (Table 1). We selected this scientific datasets and hereinafter this is refers to as MODIS AOD.

### 3.3. ERA-Interim Data

We use the daily average ERA-Interim global atmospheric reanalysis dataset (Berrisford et al., 2011) to obtain the meteorologic parameters over Nepal. The data of this reanalysis are available from ECMWF website http:apps.ecmwf.int/ datasets/. Nine variables are included in the analysis: albedo, 10m wind velocity components (U10 & V10), total columnar water vapor, total columnar water, 2m dew temperature, 2m surface temperature, sea level pressure, and surface level pressure. Obtained datasets are linearly interpolated to the resolution of 10 × 10 km. Additionally, we include relative humidity in our analysis, which is calculated using the equation given by Dingmann (2015):

\[
Relative \, humidity(\%) = 100 \times \frac{exp\left[\frac{17.625 \times TD}{243.04 + TD}\right]}{exp\left[\frac{17.625 \times T}{243.04 + T}\right]}
\]  

(5)
TABLE 1 | Pearson correlation coefficients between daily average AERONET AOD with MODIS AOD in dehydrated (dry) and normal (without dehydration) form.

<table>
<thead>
<tr>
<th>S.N.</th>
<th>MODIS scientific datasets</th>
<th>Normal AOD</th>
<th>Dry AOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Opital_Depth_Land_And_Ocean</td>
<td>0.68</td>
<td>0.85</td>
</tr>
<tr>
<td>2</td>
<td>AOD_550_Dark_Target.Deep_BlueCombined</td>
<td>0.68</td>
<td>0.85</td>
</tr>
<tr>
<td>3</td>
<td>Corrected_Optical_Depth_Land</td>
<td>0.57</td>
<td>0.76</td>
</tr>
<tr>
<td>4</td>
<td>Deep_Blue_Aerosol_Optical_Depth_Dark_Land</td>
<td>0.75</td>
<td>0.87</td>
</tr>
</tbody>
</table>

where, $T$, and $TD$ in Equation (5) are the 2 m air temperature, and dew point temperature in °C, respectively.

### 3.4. Digital Elevation Model (DEM)

A detailed map of land surface elevation was obtained from the Advanced Spaceborne Thermal Emission, and Reflection Radiometer (ASTER) (Fujisada et al., 2005) with 30 m resolution. The DEM is re-sampled to the ground resolution of 10×10 km using a nearestneighbor interpolation technique available in the re-sampling tools of spatial analyst in ESRI ArcGIS.

### 3.5. Regression Analysis

Multiple linear regression methods with Ordinary Least Square (OLS) assumptions (i.e., the model is linear in the parameter and error terms are identically and independently distributed) (Kleiber, 2001) are applied to develop an empirical model for dry AOD based on ERA-Interim predictors. Daily average AERONET AOD in dry form at Pokhara station (see Figures 1, 2) is selected as the dependent variable, as the statistical significance was greatly improved compared to the use of the wet aerosols. We have selected Pokhara to develop our model as it has the longest time series (2010–2016) with consistent data quality. The nine aforementioned meteorological variables from ERA-Interim are initially used as independent variables and we conducted a step-wise multiple linear regression analysis (Bendel and Afifi, 1977).

All the data were linearly de-trended (Tanabe et al., 2002) and normalized ($\frac{X - \bar{X}}{S}$) before doing the statistical analysis. Scatter plots between the dependent and independent variables are presented in Figure 3 with the resulting correlation coefficients ($r$).

To maximize the estimation power of the model using the minimum number of independent variables, forward step-wise regression process (Bendel and Afifi, 1977; Khatibi Bardsiri et al., 2014; Silhavy et al., 2017) is used to identify the independent variables. The selected regression model from the step-wise regression process, based on Pokhara AERONET data as a dependent variables, and average daily ERA-interim as independent variables is given by:

$$ AOD_{sim} = -0.030 - 0.6035 \times RH + 0.2140 \times V10 + 0.3140 \times U10 + \epsilon $$  \hspace{1cm} (6)

where, $\epsilon \sim N(0, \sigma^2)$ is an error term that is associated with the model prediction. $AOD_{sim}$ is the simulated AOD in dry form. We found that the coefficient of determination ($R^2$) of the model is 0.53 for daily values. Observed $P$-value for all the regression coefficients are <0.05 indicating that the model is statistically significant at 95% confidence level. In the developed empirical model given by Equation (6) simulated dry AOD ($AOD_{sim}$) is determinedly the relative humidity, and wind components (U10 and V10). Higher relative humidity is associated with removal of aerosols, reducing the AOD (loss term, negative coefficient), and aerosol sources are predominantly to the south, and west of Nepal, so southerly (V10 > 0), and westerly (U10 > 0) winds should bring higher aerosol content (source term, positive coefficients). Although Equation (6) is a proxy model only (as we have not considered emission flux and sources), this interpretation shows a relation to processes controlling the AOD.

### 3.6. Model Correction

Possible biases in simulated AOD due to the altitudinal gradients (see section 4.1) are corrected by three methods using:

- Model-1: average regression slope from MODIS,
- Model-2: monthly regression slope from MODIS data,
- Model-3: monthly regression slope calculated from the AERONET station data,

abbreviated as M1, M2, and M3, respectively. Now the correction equation for all models becomes Equation (7), except that the value of regression slope will change according to the model.

$$ AOD_{corr} = AOD_{sim} + AOD_{sim} \times \text{Exp}(E_i - EB) \times \text{Slope} \hspace{1cm} (7) $$

where, $E_i$ is the elevation of respective station, $EB$ is the elevation of base station i.e., elevation of AERONET station at Pokhara.

### 3.7. Model Evaluation Statistics

We briefly describe three standard statistical metrics which were used throughout the analysis to evaluate the performance model results: Root Mean Square Error (RMSE), Pearson correlation coefficient ($r$), and percent bias (PBIAS). These are described below.

- Root Mean Square Error (RMSE): is the standard deviation of the residual (prediction errors). It is the distance, on average, of the fitted line for a data point to the observed value. Hence it is consistent in terms of measurement units and provides a metric that is easy to interpret: the smaller an RMSE value, the
FIGURE 3 | Scatter plots between aerosol optical depth with meteorological variables. “r” represents Pearson correlation coefficient between aerosol optical depth and meteorological variables.

closer predicted values are to the observed. Mathematically it is describe by Chai and Draxler (2014)

$$\text{RMSE} = \sqrt{\frac{(AOD_{\text{sim}} - AOD_{\text{aeronet}})^2}{n}}$$  \hspace{1cm} (8)

where $n$ is the total number of observations, $AOD_{\text{sim}}$ is the simulated AOD from the model, and $AOD_{\text{aeronet}}$ is the AERONET AOD, respectively.

- Pearson correlation coefficient ($r$): $r$ (Pearson, 1896) measures the strength and the direction of a linear relationship between our observed and simulated AOD. The mathematical formula for computing $r$ given in Pearson (1895) is used. The numerical value of the correlation coefficient can vary numerically between $-1$, and $1$. The closer the correlation is to $1.0$, the stronger the relationship between the two variables, whereas a negative value defines an anti-correlation.

- Percent bias (PBIAS): measures the average tendency of the simulated AOD to be larger or smaller than their observed counterparts (Moriasi et al., 2007). The optimal value of PBIAS is 0.0, with a low-magnitude value indicating accurate model performance. Positive values indicate model underestimation bias while negative values indicate overestimation bias.

$$\text{PBIAS} = \left[ \frac{(AOD_{\text{aeronet}} - AOD_{\text{sim}}) \times 100}{\sum(AOD_{\text{aeronet}})} \right]$$  \hspace{1cm} (9)

4. RESULTS AND DISCUSSION

4.1. Altitudinal Distribution of AOD

Shown in Figure 1 is the spatial distribution of average observed AOD from MODIS over Nepal for the period of 2000–2015. It is clear that the spatial gradients of AOD are inversely related to the topography, i.e., higher AOD over the southern low land and lower AOD over the mountainous regions of Nepal. Scatter plots (Figure 4) show that the AOD non-linearly decreases with the altitude (steep negative gradient) up to $\sim$500 m asl. and thereafter varying more smoothly with the altitude. We found that the altitudinal distribution follows a semi-logarithmic form with a slope and coefficient of determination of $-0.135$ and $0.899$, respectively (blue points in Figure 4 is the linear transfer of elevation by taking natural log). Since the aerosol load over the region varies greatly with the seasons (Acharya and Sreekesh, 2013), slope coefficient of the regression equation on the monthly average MODIS and AERONET observations are calculated over Nepal. In Figure 5 the blue points represent an average AOD per each grid cell with respective elevation in semi-logarithmic scale. The red triangles represent the monthly average AOD from the AERONET stations. As compare to AERONET AOD, MODIS underestimates AOD in all months. Monthly gradients from MODIS and AERONET observations are different. But from both observations, the strongest elevation gradients are observed during the winter season, while the weaker gradients are observed during the monsoon season.

4.2. Seasonal Variability and Spatial Distribution of MODIS AOD

We calculated a spatial average (2000–2015) of the MODIS AOD in the dry forms for each season to evaluate the seasonal variability. These are shown in Figure 6. The maps show average AOD retrievals over Nepal are concentrated toward lower elevations (Terai) where a high density of data pixels exists, and the retrievals are limited in higher elevation regions due to high surface reflectance. However, the seasonal spatial distribution of AOD over Nepal clearly shows typical cycle of AOD with high AOD in winter and pre-monsoon, and low in monsoon and post-monsoon seasons. We present a further discussion of the seasonal dynamics of AOD over Nepal in section 4.5.3.
4.3. Validation of MODIS AOD With AERONET

Because of enormous altitude variation within a short south to north distance, Nepal has remarkable climatic variability (tropical to Arctic) (Li et al., 2017) which poses a challenge for satellite remote sensing of aerosol. Validation of remotely sensed AOD with ground based instruments (AERONET) is worthwhile in a region where such studies have not yet been completed. We followed the general way to validate MODIS AOD against AERONET AOD and the detailed procedure that we followed is explained in section 3.2. Scatter plots between mean AOD from MODIS and AERONET for different stations are plotted and shown in Supplementary Figure 2. The AERONET sites are located in different elevations ranging from 110 to 5,050 (m asl.) (Figure 2), but the validation is done in four AERONET stations (i.e., Lumbini, Pokhara, KTM-BO, and KTM-UN) with MODIS observations. The comparison is performed using co-located points. Seasonal analysis of these co-located AOD data shows that, most of these data points are from winter, and pre-monsoon season in all stations. In Pokhara, KTM-BO, and KTM-UN AERONET stations about 30, 55, and 41% of the co-located data are from winter season, while 55, 29, and 25% of the co-located data are from pre-monsoon season, respectively. The linear correlation coefficient, with RMSE, and regression slopes between AERONET, and MODIS observation over four station are given in Table 2. The regression slope, and intercept from Table 2 shows that AOD observations from AERONET in all station are higher than from MODIS. Chu et al. (2002) and Li et al. (2009) showed that the intercept of linear regression different from zero represents the errors in surface reflectance estimates, and the regression slope differing from 1.0 represents a systematic bias of MODIS AOD retrievals. We also tried to do similar comparison in our study. Table 2 shows good correlation (0.75–0.91) between MODIS derived and AERONET observed average daily dry AOD with intercept values between 0.29 and 0.44, and a regression slope between 0.62 and 1.65. About 43, 74, and 69% of the observations fall inside expected error (EE) from KTM-BO, KTM-UN and Lumbini AERONET station. Lower percentage inside EE from KTM-BO might be due to fewer observation compared to KTM-UN and Pokhara (Figure 2). However, in comparison with global validation results (Levy et al., 2007), with intercept of 0.029, and slope of 1.009, MODIS C005 AOD retrieval has higher errors in Nepal. The relatively high positive offset of MODIS AOD in Nepal (e.g., intercept 0.29–0.44) is indicative of the poor estimates in surface reflectance (Li et al., 2009). Moreover, daily MODIS, and AERONET derived AOD are not concurrent in time. Therefore, the sampling time inconsistency for AERONET and MODIS AODs is also a source of uncertainty.

4.4. Average MODIS AOD Distribution With Inter Annual Variability, 2000–2015

Analysis of seasonal average dry AOD for Nepal shows the nature of the dynamics of aerosol concentration during the study period. Figure 7 shows inter annual variability and 16 years AOD trends for each season from MODIS. It shows that the AOD exhibited an increasing trend in all seasons except in winter, but only the trend in monsoon season is statistically significant at 95% significant level (faint dashed lines in Figure 7). The highest average AOD throughout the period was in the pre monsoon, whereas the monsoon season showed an increasing trend after 2012. Between 2002–2004, and 2007–2010, the post-monsoon curves showed significant lowering of average AOD. The high AOD in pre-monsoon during 2001–2005 may be attributed to low rainfall over most parts of Nepal (Department of Hydrology and Meteorology Government of Nepal, 2017), which caused an increase in aerosol loading in the atmosphere. The winter decline in AOD (0.36–0.16) from 2000 to 2003 (0.046–0.01), and 2007 to 2010 (0.03–0.02) was possibly due to a reduction of aerosol load through higher precipitation observed over Nepal. Overall, significant increasing trend with slope of 0.0006, and p-value of 0.042 is observed for average AOD over Nepal. This weak but significant increasing trend may be attributed to annual decreasing trend in precipitation in Nepal (1.3 mm per year), although this precipitation decreasing trend is not significant (Department of Hydrology and Meteorology Government of Nepal, 2017).

In order to see the average inter-annual variability of AOD in Nepal, monthly average dry AOD from both MODIS and AERONET data are plotted in Figure 8. Standard deviation about the mean from corresponding nine cell MODIS observation is plotted as shaded part in the plot. From the seasonal AOD analysis, AERONET AOD is found higher than most of AOD observation from the MODIS. Although the seasonal AOD comparison between MODIS and multi wavelength radiometer (MWR) by Galeria et al. (2012), and between MODIS, and Multiangle Imaging Spectro-Radiometer(MISR) by Prasad and Singh (2007) over Indian subcontinent shows that MODIS is overestimating during summer, and underestimating during winter. But we found that in four stations MODIS AOD is lower
than AERONET AOD (Figure 8). From MODIS and AERONET observation, the highest AOD is observed during the period of March-May when Nepal experiences heavy spring dust from north-west a feature that might be attributed to the fact that we use dry aerosol in this analysis. The scavenging effect of the rain can be seen from both observations as the lowest AOD is observed during the period of July-August when the monsoon season initiates. The highest AOD observed is at Lumbini during November and December, and likely results from biomass burning for heating in the Terai region (Wang et al., 2013).

4.5. Model Result Analysis
The empirical model obtained from Equation (6) resulting from the multiple linear regression is used to simulate daily dry AOD at different AERONET stations in Nepal. Simulated daily dry AOD values with observed data are plotted in time in Figure 9. The time series can be effectively used to understand the predictability of the model. It is observed that the simulated dry AOD values are similar to the measured values except in Jomsom, and EVK2CNR station. The average bias (difference between observed mean with simulated mean) between simulated and observed value are found to be 0.0002, 0.034, 0.009, 0.013, −0.045, −0.043 for Pokhara, Lumbini, KTM-BO, KTM-UN, Jomsom, and EVK2CNR station, respectively. Results showed that the model underestimated when is used to simulate AOD in lower elevation region than Pokhara station, and overestimated AOD in Jomsom and EVK2CNR, which are located in higher altitude. The correlation coefficient between observed and modeled dry AOD (presented inside each plot) are found to be higher at stations in higher altitude indicating their higher similarity between observed and simulated values.

We found that the disagreement (i.e., highest PBIAS and RMSE) between the simulated, and observed AOD is due to the altitudinal dependencies (Figure 5), as it follows the topography

---

FIGURE 5 | Monthly average AOD at 550 nm from MODIS (blue) and AERONET (red) observation with log elevation. Blue and red lines show the best fit linear lines for each datasets. Numbers in each subplots show the slope for each best fit lines, respectively.
with higher AOD values over the low land than over the mountain area (Figure 1). Roux et al. (2008) showed that the lower aerosol load in French mountains are because of high wet deposition due to orographic effects, but in the case of Nepalese mountainous region, lower AOD observed from MODIS and AERONET might be lower anthropogenic activity in the region in addition. As we discussed in sections 4.2 and 4.4 that the aerosol distribution in Himalayan region are not only dependent upon altitude, but also dependent upon seasons, because the Hindu-Kush-Himalayan region is strongly influenced by large-scale atmospheric circulation, which alternates between the wet summer monsoon, and dry season. As the distribution of aerosols over time and space is determined by its type, size, and source (Cristofanelli et al., 2014), aerosols transported to Himalayas mainly from Indo-Gangetic plain during the pre-monsoon season (Ramanathan et al., 2007; Kopacz et al., 2011) are deposited differently over the space. Dhungel et al. (2016) mentioned that the different sources of AOD for the Himalayan region (especially to Nepal) are from biomass burning (mainly in mountain region) and fossil fuels combustion (in low land or Southern parts of Nepal).

To overcome biases between the model and observed AOD, monthly average regression slopes obtained from the linear regression between log elevation and AOD (Figures 4, 5) are used to correct the simulated AOD values from the model over the different stations. Results from three different methods (see section 3.6) are discussed in section 4.5.1.

### 4.5.1. Model Performance Evaluation After Using Different Regression Slope Values

The three different slope correction models M1, M2, and M3 are compared with the daily average dry AOD values observed from respective AERONET stations. Model evaluation statistics were calculated, and presented in Figure 10. AERONET stations in Figure 10 are arranged in such a way that the station elevations (Figure 2) are in increasing order from Lumbini to EVK2CNR in anticlockwise direction.

From the calculated error statistics (Figure 10), the models for the stations at the lower elevations are performing at a very close PBIAS, if we compare the PBIAS among all. With increasing station elevation, performance of the model decrease for M2, and M3. Performance difference can be clearly seen at EVK2CNR and Jomsom stations, where correlation coefficient of 0.1, 0.5, 0.82, RMSE of 0.018, 0.01, 0.002 with PBIAS of −62.8, 4.9 are observed for the models M3, M2, and M1, respectively at EVK2CNR. Although a higher correlation

---

**TABLE 2 | Different statistics calculated from daily average AERONET AOD and daily MODIS AOD for each station.**

<table>
<thead>
<tr>
<th>Station</th>
<th>RMSE</th>
<th>Correlation coefficient</th>
<th>Slope</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lumbini</td>
<td>0.33</td>
<td>0.91</td>
<td>0.62</td>
<td>0.44</td>
</tr>
<tr>
<td>Pokhara</td>
<td>0.13</td>
<td>0.75</td>
<td>1.34</td>
<td>0.31</td>
</tr>
<tr>
<td>KTM-BO</td>
<td>0.39</td>
<td>0.75</td>
<td>1.65</td>
<td>0.29</td>
</tr>
<tr>
<td>KTM-UN</td>
<td>0.39</td>
<td>0.84</td>
<td>1.29</td>
<td>0.29</td>
</tr>
</tbody>
</table>

---

**Figure 6 | Seasonal average MODIS AOD (2000–2015) over Nepal.**
A coefficient of 0.85 is observed in Jomsom station from the model M2 than model M1 (0.83), but the PBIAS and RMSE are better for model M1 (-68.2% and 0.014) compared to M2 (91% and 0.018). M3 performs similar to M1 and M2, however, it also tends to underestimate AOD with increasing elevation. Underestimation of AOD at EVK2CNR station from the model M3, is mainly due to the steeper regression slope obtained from the available monthly average AERONET data compared to the monthly regression slopes from the monthly MODIS AOD. Referring to Figure 5, monthly regression slopes mainly in Oct, Nov, Dec, April, May, and June are steeper when compared with monthly MODIS slope as well as average slope. These steeper slopes force the model M3 to predict relatively lower AOD values resulting overall underestimation and low predictability of the model M3.

Overall, our conclusion is that M1 performs the best of the three. In addition to the improved performance the correction is most simple, using an average regression slope from MODIS, and therefore M1 is recommended for the extrapolation of AOD in the Nepalese Himalaya region. But the selection of model might be different for the different purpose of study. In a study about seasonal patterns, model correction by monthly slope...
might be appropriate, although the difference between these two models (M1 and M2) is not big. In Figure 11, observed and predicted AOD from M1 is presented. The correlation coefficient between the simulated AOD from model M1 and observed AOD from AERONET stations, Pokhara, Lumbini, KTM-BO, and KTM-UN are 0.7, 0.52, 0.72, and 0.67, respectively. We found that PBIAS in Pokhara, Lumbini, KTM-BO, and KTM-UN are 0.5, 34.4, 36.1, and 41.9%, respectively, indicating a general underestimation from the model. Highest RMSE (0.06) is calculated at Lumbini while the lowest is observed at Pokhara (0.03).

To better understand the seasonal predictability of the selected model, the longest time series of daily observed and simulated dry AOD data at Pokhara (2010–2016) and EVK2CNR AERONET stations (2006-2015) are taken into account. During the monsoon season, the simulation gives results that agree with measured dry AOD at Pokhara station where correlation coefficient of 0.89, with 0.019 RMSE, and lowest PBIAS of $-0.6$ were observed, and these values are better than in other seasons. In EVK2CNR, higher correlation coefficient between AERONET and simulated AOD were observed in winter (0.92), and in post-monsoon (0.89) season with RMSE of 0.002 and 0.0015, respectively. Low model performance (i.e., low correlation coefficient, large PBIAS and high RMSE) during the winter season at Pokhara is due to local air pollution. Since 87% of the energy requirement of the county...
is fulfilled by traditional sources like firewood, animal dung, and some paper residue leaves of trees to warm houses and as a kitchen fuels (Ranabhat et al., 2015), producing human-made aerosols in the form of smoke (Panday and Prinn, 2009). Higher model performance during the post-monsoon and winter season at EVK2CNR is due to the absence of local sources of pollution as compared to Pokhara station. Observed higher performance during the post-monsoon and winter season at EVK2VNR is also due to the number of sufficient observations during that season as the percentage of data point observations in post-monsoon and winter seasons are 35.63 and 26.24%, respectively.

4.5.2. AOD Simulation Compared With MODIS AOD

Since the simulation results from M1 is found better than the other models, it is further used to simulate dry AOD over all Nepal, and results are compared with the corresponding MODIS AOD observations. Comparison is made by calculating correlation coefficients between the simulated AOD with MODIS AOD. We found that the most of the correlation grid (see Figures 13A,B) shows that the values above 0.5 with some exception. We observed that the correlation coefficient between MODIS observation and model simulation is higher at lower elevations in Nepal. As the model is developed by using AERONET data, systematic error between MODIS observation and AERONET observation also leads to lower correlation coefficient. In addition to systematic errors between the two observations, the impact of clouds and coarse spatial resolution of MODIS (10 × 10 km) may also introduce other uncertainties. Figure 12 shows the seasonal distribution of model AOD over Nepal. It follows the same seasonal distribution as we found from MODIS observation.

4.5.3. Seasonal AOD From MODIS and Model

Figures 6, 12 show the seasonal distribution of dry AOD from MODIS and model. From both, highest AOD is observed during winter and pre-monsoon season and lower AOD during the monsoon and post-monsoon season. In winter spatial distribution of dry AOD shows that for most of the southern part of Nepal, moderate to high values prevail. Cold surface conditions in winter, mostly in the southern plain region of Nepal, produces very dense mist, haze, and fog, through water vapor condensation on carbonaceous aerosol particles from biomass burning leading to elevated AOD, which is well-captured by MODIS observation. Higher average AOD values are observed over the Terai region which may be attributed to biomass burning activity during colder weather (Wang et al., 2013). Increased burning activity, in association with emission from industrial and fossil fuel burning (especially in Terai), also increases atmospheric AOD loading.

In pre-monsoon season, the shifting of the inter tropical convergence zone to Indo-Gangetic plain produces intense heating of the surface resulting in moderate to strong westerlies winds (Nayaya, 1980). These westerlies are also associated with strong dust storms (Acharya and Sreekesh, 2013) occurring mostly over the southern part of Nepal, transferring large amounts of dust to the air leading to higher AOD in the atmosphere (Figure 6). The spatial variability of AOD is also controlled by the surface moisture content during this period. Intense temperature, in association with strong surface winds during pre-monsoon, plays an important role in heating and lifting the loose soil. The onset of gusty winds during the pre-monsoon (Shea et al., 2015) caused by convective instability, produces a large amount of dust.
aerosol locally leads to an increase in AOD during this season (Flossmann et al., 1985).

The monsoon climate over Nepal controls the seasonal aerosol concentration in the atmosphere. Lower AOD during monsoon is observed from model and MODIS observation (Figures 6, 12). The lower AOD is due to the significant amount of rainfall which occurs during the monsoon, accounting for 60-90% of the total annual rainfall (Nayava, 1980). Higher amount of rainfall leads to higher relative humidity forcing model to predict lower AOD value. Therefore, the concentration of aerosol particles is reduced during this season as they are rapidly removed from the atmosphere through wet deposition (Gonçalves et al., 2010) which is well-captured by model and MODIS. Despite a strong dependence on the monsoon, interestingly the total average MODIS AOD value over Nepal during the monsoon season (0.02) is higher than during winter season (0.017). However, this potentially arises by choosing a fixed date for the monsoon season, rather than meteorological thresholds. Due to the fact that the timing for onset of monsoon in each year is not constant, the AOD values may be rather elevated if the monsoon rains have not yet initiated, leading to higher average values than in winter season. Figures 6, 12 clearly show that the AOD is higher in the western part than in eastern regions. This is a result due to the onset of monsoon from the east (Bhatt and Nakamura, 2005), and slowly moving toward the western region.

By the end of the monsoon, aerosol levels start to rise again during the post-monsoon seasons. It is the transition period between the wet and dry seasons. Retreat of the monsoon trough is accompanied by a high-pressure cell positioned over...
the Tibetan plateau, generally called the Tibetan high. The cold north-easterly wind from this high-pressure cell spreads over southern Nepal, and makes the surface air relatively dense, restricting the effective convection mechanism. As a result, aerosols are closer to the surface, yielding a large backscattering fraction that increases AOD levels in Terai region (Figures 6, 12). The spatial variability seen in the post-monsoon season shows typical AOD levels in the eastern and south-western parts of Nepal, respectively. As the eastern part of Nepal receives more precipitation compared to the western part (Department of Hydrology and Meteorology Government of Nepal, 2017), AOD distribution over this region is also different, as can be seen in Figures 1, 13. Variability in weather patterns, as well as the nature, and intensity of emissions, produces such dramatic variation in the spatial profile.

After the simulated AOD from model are compared with the MODIS AOD data, model is further used to extrapolate AOD over the cryospheric portion of Nepal, and presented in Figure 13A. It is evident that the model is also able to predict relatively higher AOD value in river valleys (line toward north at around N29, E83.5) with settlements, as compared with surrounding mountainous peaks with snow, which was not detected from MODIS instrument, ascertaining its ability to predict with ground reality.

4.5.4. Uncertainty in the Model Prediction
Effective AOD prediction is a complex issue that is easily affected by various factors, including weather and climatic conditions, and emission fluxes. Different input factors as an independent variables also have different degree of impact on the regression results used for dry AOD prediction. In the studied region, AOD is highly seasonal and altitude dependent, inclusion of altitudinal correction factor improves the predictability of the model. However, this study only predicted dry AOD for Nepal, but did not take into consideration any regional differences in the atmospheric environment. Although the results were good, few uncertainties are remained. First, uncertainties in the AOD data sources: on the one hand, this was a reflection of the uneven spatial distribution and few monitoring AERONET station, which are mainly in lower elevation region. Second, uncertainties in the ERA Interim data which were more prone to generating random noise, which affected prediction accuracies. And the third is the uncertainties in the proposed model itself. This study only assumed a possible linear relationship between AOD and three meteorological variables but did not consider the emission fluxes and sources, which would have an impact on the effectiveness of the model. Depending on the region and time period, significant differences exist in dry AOD. The good predictions achieved by the model proposed here were limited to Nepal and over short duration. Further examination would be needed to determine whether the developed proxy model could be applied to dry AOD predictions at other mountain region with longer time periods.

5. CONCLUSIONS
This paper presented dry AOD retrieval methods over the cryospheric portion of Nepalese Himalayas. Multivariate regression analysis is carried out to develop proxy an empirical model to predict AOD in dry forms. Three meteorological variables (relative humidity and 10 m wind velocity components) from ERA-Interim, and AOD observation at Pokhara AERONET station (in dry forms) were taken into account for model development. We have presented the results from the model corrected using average regression slope from MODIS (M1). Simulated dry AOD from developed empirical proxy model is validated with AERONET observations. Results showed that the presented model (M1) is able to simulate dry AOD over the different regions of Nepal, indicating model adequacy and establishing the model as an efficient AOD prediction model for Himalayan region. From both MODIS and model, highest dry AOD are observed during winter and pre-monsoon seasons. There are some discrepancies between observed and model values. The possible reason behind this discrepancies are because of the observed AOD represent sample of the population whose mean should correspond to the values predicted by the model. In this empirical proxy model, emission flux and sources are not considered while developing model. Therefore, prediction from model have biases with observations, and also in the case of few measurement data, it is very difficult to validate model results with the fewer observation. With limited observed data, presented proxy model shows sufficient number of grid cells with higher correlation coefficients indicating its adequacy as proxy AOD prediction model for Himalayan region. Identifying the source of aerosol is beyond the scope of this study but as mention in several studies, aerosols over Nepal are transported from Indo-Gangetic plain. The repeated occurrence of forest fires (especially in hilly plain region during pre-monsoon season) contributed to AOD over Nepal. Thus, it can be concluded that the spatial variability of AOD depends upon weather condition, and emission sources, which are subject to change with seasons. The results of this paper confirms that the use of even the simplest linear regression model will yield very good estimation results for daily average dry AOD data in Nepal. These simple models are indispensable prediction tools for scientist requiring AOD information in the data sparse Himalayan region. It is evident that the model from multivariate regression analysis has universality in statistics, and it can be able to predict most of the spatial AOD variability in the Himalayan regions. It is especially useful for the situation where we do not have explicit knowledge about the AOD. This includes modeling in very complex terrain where very limited observed data sources exist. We tested the model on a few available stations, but it can also be used for other regions for AOD estimation. Our proxy model only uses ASTER DEM and meteorological data from ECMWF, and therefore, these models could be easily applied to other regions with mountain environment, and in related climate research in mountain region.
AOD Over Nepalese Cryosphere

**AUTHOR CONTRIBUTIONS**

BB and JB designed the analysis. BB performed the experiments, derived the models, and analyzed the data. BB and JB wrote the manuscript with input from FS and C-YX who also provided expertise on the statistical modeling.

**FUNDING**

This work was conducted within the Strategic Research Initiative Land Atmosphere Interaction in Cold Environments (LATICE) of the University of Oslo and partially supported through the Norwegian Research Council’s INDNOR program under the Hydrological sensitivity to Cryosphere-Aerosol interaction in Mountain Processes (HyCAMP) project (NFR no. 222195).

**ACKNOWLEDGMENTS**

We were thankful to entire MODIS science team and NASA providing aerosol data.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2019.00178/full#supplementary-material
Bhattarai et al. AOD Over Nepalese Cryosphere


Pearson, K. (1896). The measurement of organs arise when indices are used in spurious correlation which may of theory of evolution–on a form mathematical contributions to the email alerting service. Proc. R. Soc. Lond. 1, 489–498.


Paper II: Evaluation of global forcing data sets for hydropower inflow simulation in Nepal
Evaluation of global forcing datasets for hydropower inflow simulation in Nepal
Bikas Chandra Bhattarai, John Faulkner Burkhart, Lena M. Tallaksen, Chong-Yu Xu and Felix Nikolaus Matt

ABSTRACT
Discharge over the Narayani river catchment of Nepal was simulated using Statkraft’s Hydrologic Forecasting Toolbox (Shyft) forced with observations and three global forcing datasets: (i) ERA-Interim (ERA-I), (ii) Water and Global Change (WATCH) Forcing Data ERA-I (WFDEI), and (iii) Coordinated Regional Climate Downscaling Experiment with the contributing institute Rossy Centre, Swedish Meteorological and Hydrological Institute (CORDEX-SMHI). Not only does this provide an opportunity to evaluate discharge variability and uncertainty resulting from different forcing data but also it demonstrates the capability and potential of using these global datasets in data-sparse regions. The fidelity of discharge simulation is the greatest when using observations combined with the WFDEI forcing dataset (hybrid datasets). These results demonstrate the successful application of global forcing datasets for regional catchment-scale modeling in remote regions. The results were also promising to provide insight of the interannual variability in discharge. This study showed that while large biases in precipitation can be reduced by applying a precipitation correction factor (p_corr_factor), the best result is obtained using bias-corrected forcing data as input, i.e. the WFDEI outperformed other forcing datasets. Accordingly, the WFDEI forcing dataset holds great potential for improving our understanding of the hydrology of data-sparse Himalayan regions and providing the potential for prediction. The use of CORDEX-SMHI- and ERA-I-derived data requires further validation and bias correction, particularly over the high mountain regions.

Key words | discharge, global forcing dataset, Himalaya, hydrological modeling, hydropower inflow simulation

INTRODUCTION
The Himalayan and adjacent Tibetan Plateau water supply is intricately linked to the livelihoods, economic and social, to millions of people (Bookhagen & Burbank 2010; Immerzeel et al. 2010; Remesan et al. 2019; Zhang et al. 2019). This region provides for drinking water, hydropower generation, agricultural demands, as well as water-powered grain mills and other agro-economic activities (Ménégoz et al. 2013). Large Asian rivers, such as the Indus, Sutlej, Ganges, and Brahmaputra, are widely acclaimed for their great cultural, spiritual, economic, and ecological significance (Kumar 2017). The Ganges River and its tributaries alone fulfill significant water demands for more than 250 million people. Originating in the Himalayas, the river travels over 2,500 km, aggregating water from its tributaries...
to become the third largest freshwater delta (Chowdhury & Ward 2004).

Due to the high precipitation rates (1,500–2,500 mm/year; Dahal & Hasegawa 2008) and steep elevation gradients, hydropower has significant potential in the region and, if well managed, could provide a resource for economic growth. Hydropower potential depends on climatic conditions, particularly on precipitation, evaporation, temperature, and snow/ice in the catchment (Edenhofer et al. 2011). However, the barriers for sustainable water resource management in this region are manifold. First, there are numerous socio-political issues related to trans-border management issues (Biswas 2011). Despite efforts, a comprehensive framework to guide the development and international cooperation has yet to be ratified (Biswas 2011). Second, portions of the region are heavily glaciated and climate change is exacting an immediate and tangible impact (Immerzeel et al. 2010; Dehecq et al. 2019).

In this paper, we focus on a third significant barrier: providing robust and reliable analysis of the water resource potential and impacts from climate change. Data scarcity and the complexity of the terrain creating large spatial gradients and variability in weather and climate make such analyses a challenging task. The extremely heterogeneous topography in the region presents a great challenge for the accurate measurement of meteorological variables, giving rise to data scarcity (Pellicciotti et al. 2012). Although there are some meteorological stations, they are not well distributed in space. To overcome these challenges, regional climate models and reanalysis data offer gridded datasets of many meteorological variables, although with a rather coarse spatial resolution (Guo & Su 2019). The Coordinated Regional Climatic Downscaling Experiment (CORDEX) (Giorgi et al. 2009) provides extensively used regional datasets for past and future climate. For a historical perspective, the reanalysis dataset ERA-Interim (ERA-I) from the European Center for Medium-Range Weather Forecasts (ECMWF) (Dee et al. 2011) is widely used (Li et al. 2013, 2018; Xu et al. 2016). Bharti & Singh (2015) reported that the ERA-I precipitation is largely overestimated over the Indian Himalayan region. To address the bias inherited in ERA-I, bias- and elevation correction-based monthly observations were carried out when creating the Water and Global Change Forcing Data ERA-I (WFDEI) datasets (Weedon et al. 2014; Kim et al. 2019).

Various factors make hydrological modeling in the region challenging, including a large spatial variability in hydrometeorological variables, steep gradients, marked seasonality driven by the Indian Monsoon, and contrasting moisture regimes between the high elevation Tibetan Plateau and regions in the vicinity of the Indian Ocean. Changes in the climatic condition in the region may lead to changes in regional water balance components impacting hydrological regimes (Koch et al. 2011). Moreover, snow and glacier storage and melt play an important role for the river discharge generation (Radić & Hock 2014; Li et al. 2015, 2016). The effective assessment and implementation of hydropower projects and other water resource management projects depend on a thorough analysis of the discharge and hydrological storages (Bakken et al. 2015).

Hydrologic simulation of discharge and other water balance components (evapotranspiration, snow, and groundwater storage) (Bhattarai & Regmi 2015; Matt & Burkhart 2018; Li et al. 2019) is used for the analysis of available water resources both in the past and future analyses, but the quality of discharge simulation remains a challenge (Rochester 2010; Engeland et al. 2016; Kauffeldt et al. 2016). The choice of a suitable hydrological model and appropriate forcing data are critical for any analysis and will greatly affect the outcome (Kauffeldt et al. 2016). Currently, no standardized or community modeling framework within hydrology and prior studies of the water resources of the Himalayan region have been conducted using different types of hydrologic models ranging from simple conceptual models (e.g. Pradhananga et al. 2014; Bhattarai & Regmi 2015; Skaugen & Weltzien 2016; Bhattarai et al. 2018) to more advanced, distributed models (e.g. Pellicciotti et al. 2012; Jain et al. 2017). Hydrological modeling and water balance studies in Himalayan regions have taken a range of approaches and used different models to focus on topics such as glacier melt and retreat, water balance, flooding, and the impact of climate change (Bookhagen & Burbank 2010; Pellicciotti et al. 2012; Wortmann et al. 2014; Khanal et al. 2015). Recent studies have addressed the performance of large-scale forcing datasets, including the WFDEI and ERA-I, for discharge simulation in various regions across the world (e.g. Li et al. 2015, 2016; Weedon et al. 2014; Essou et al. 2016; Nkiaka et al. 2017), but, to date, the use of these
global forcing datasets to simulate discharge in a Himalayan catchment is limited.

In this study, we use a distributed, conceptual hydrologic framework with an energy balance-based snow routine that has been demonstrated to perform well in the region (Xu et al. 2015; Hegdahl et al. 2016; Matt & Burkhart 2018). The main goal of this work is to evaluate the impact of four different forcing datasets on the simulated discharge and associated water balance assessment for the Narayani river catchment in Nepal. The forcing datasets include WFDEI, ERA-I, CORDEX, and a hybrid dataset referred to as Observed + WFDEI (all datasets are described in the Data and Methods section). Our specific objectives are to evaluate: (i) the discharge sensitivity to the forcing dataset for hydrologic modeling inflow, (ii) model performance using bias-corrected versus non-bias-corrected forcing as input to a (calibrated) hydrological model, and (iii) the sensitivity to forcing data for water balance assessment. Furthermore, we assess the overall ability of the Statkraft Hydrological Forecasting Toolbox (Shyft) (Burkhart et al. 2016) to simulate discharge in this complex and data-sparse regions.

**STUDY AREA**

The Narayani river catchment lies in the central part of Nepal. About 13% of the total area lies in China (Figure 1), and thus, it is a transboundary catchment. The main Narayani river gauging station is located in Narayanghat (27°42’30”N, 84°25’50”E) and is operated by the Department of Hydrology and Meteorology, Government of Nepal (DHM, GoN). The catchment area is 31,692 km² and is partly glacier covered (~8%; Omani et al. 2017) with elevation ranging from 175 m a.s.l. in the south to 8,148 m a.s.l. in the north. The hydrological regime is heavily influenced by the season. The seasons are defined as monsoon (June to September), post-monsoon (October to November), winter (December to February), and pre-monsoon (March to May) (Shrestha et al. 1999; Bhattarai et al. 2019). Almost 80% of total annual precipitation occurs during the monsoon period (Nayava 1974; Kripalani & Sontakke 1996). Tributaries to the Narayani river are either monsoon fed (those originating in middle and high mountain regions) or glacial and snow melt fed (those originating in a higher Himalayan...
region). The Narayani river catchment lies also in the primary hydropower development region of Nepal, providing 44% of the total electric generation (Adhikari 2006).

DATA AND METHODS

Meteorological forcing data

Forcing variables used in the study include air temperature (T), precipitation (P), relative humidity (RH), wind speed (WS), and shortwave incoming solar radiation (S) at a daily resolution. Details of each dataset are described in the following sections.

Observed data

The DHM, GoN is responsible for collecting and disseminating hydrological and meteorological information for water resources, agriculture, energy, and other development activities in the country. Most of the measured data are based on the conventional manual observation. Measurements from the automatic weather station (AWS) are available only after the year 2000 (Karki 2010). The geographic location and installation date for each station are available from http://dhm.gov.np/meteorological-station/.

Observed temperature and precipitation station data which are located inside or in the vicinity of the catchment (Figure 1) were collected from the DHM. Data from the DHM were manually plotted for each station and screened for data quality. Clearly, erroneous departures from the historical pattern were removed manually from the datasets. The highest temperature station elevation is 3,870 m a.s.l., located in Chhoser (DHM st. no 633), and the highest precipitation station elevation is 3,705 m a.s.l., located in Mustang (DHM st. no 612).

To accommodate missing data, stations with less than 10 years of record or missing more than 15% of the observations were removed from the datasets. Ten precipitation stations (out of 82) did not meet this criterion and were removed from the datasets. Similarly, seven (out of 30) temperature stations were removed from datasets. This resulted in 72 precipitation stations and 23 temperature stations that were used for further analysis. Maximum numbers of stations with temperature and precipitation data were observed in the 2000–2009 period. Therefore, in this study, data for the period 2000–2009 were used.

All the stations are located below 4,000 m a.s.l., covering only 60% of the total catchment area (Figure 2(a)). Generally, temperature and precipitation data from higher elevations are missing, which add uncertainty to the current study. Daily mean temperature and mean annual precipitation for each station are plotted versus station elevation in Figure 2(b). Temperature shows as expected a consistent decrease with elevation, whereas precipitation shows a more mixed picture, with no clear trend below approximately 2,000 m a.s.l., but a notable decrease in precipitation above, i.e. from about 2,000 to 3,500 m a.s.l. Normally precipitation increases with elevation (Daly et al. 1994) in the mountainous region due to the orographic effect, but in the Himalayan region, an opposite pattern has been reported after a certain elevation level (e.g. Nayava 1980; Kansakar et al. 2004), in agreement with what is observed for the Narayani river catchment. From the analysis, station average daily temperature and mean annual precipitation for the period 2000–2009 were 10.7 °C and 1,292 mm/year, respectively.

Discharge observed at the Narayanghat station is available from the year 2000–2009 with no missing values, and this is also the period used for the hydrological model calibration and validation. Daily average discharge for the period 2000–2009 was 1,482.5 m³/s.

Reanalysis and regional climate model data

A summary of the available time periods and the resolution of the forcing datasets are provided in Table 1. All datasets are well documented and are freely available.

ERA-I is a reanalysis global forcing dataset available from 1979, produced by the ECMWF. ERA-I temperature results from the assimilated surface temperature (Essou et al. 2016), while precipitation data are based on a reanalysis of precipitation fields generated with a meteorological model (Berrisford et al. 2011; Dee et al. 2011). The obtained precipitation data are not scaled using observation data. ERA-I is continuously updated once per month, with a delay of 2 months, and is freely available from http://apps.ecmwf.int/datasets/. Monthly mean meteorological variables
averaged over the Narayani river catchment using the Shyft interpolation routines (see the section ‘Spatial-temporal interpolation of forcing data’) are shown in Figure 3. The catchment average mean temperature for the period (2000–2009) was 7.6 °C, while catchment average annual precipitation was 4,660 mm/year.

The WFDEI is a global forcing dataset obtained by downscaling and bias-correcting ERA-I data (Weedon et al. 2014; Raimonet et al. 2017). The temporal and horizontal resolution of the dataset is shown in Table 1. WFDEI has two sets of rainfall generated by using either Climate Research Unit (CRU) or Global Precipitation Climatology Centre (GPCC) precipitation correction methods (Weedon et al. 2014). In this study, GPCC-corrected data were preferred to CRU because of their higher resolution and data quality (Weedon et al. 2014; Raimonet et al. 2017). The WFDEI dataset is freely available online from ftp.iiasa.ac.at. Catchment average mean daily temperature and annual precipitation (for the period 2000–2009) were found to be 7.9 °C and 1,764 mm/year, respectively.

CORDEX is a program sponsored by the World Climate Research Programme (WCRP), to produce an improved generation of regional climate change projections (Giorgi et al. 2013). CORDEX has two datasets, referred to Evaluation and Historical. Evaluation is run within reanalysis and is used to ‘mimic’ observations (i.e. represent real weather), whereas Historical is run within a climate model, and the results can only be used in a climatological sense. Daily datasets of the Evaluation product over the study area are available for the South Asia CORDEX.

Table 1 | Summary of selected forcing datasets (n refers to the number of grid cells in the Narayani river catchment)

<table>
<thead>
<tr>
<th>Forcing dataset</th>
<th>Data period</th>
<th>Spatial resolution (degrees)</th>
<th>Temporal resolution</th>
<th>n</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERA-I</td>
<td>1979–2015</td>
<td>0.75 × 0.75</td>
<td>3 h, 6 h</td>
<td>12</td>
<td>Dee et al. (2011)</td>
</tr>
<tr>
<td>WFDEI</td>
<td>1979–2016</td>
<td>0.5 × 0.5</td>
<td>Daily</td>
<td>20</td>
<td>Weedon et al. (2014)</td>
</tr>
<tr>
<td>CORDEX-SMHI</td>
<td>1980–2010</td>
<td>0.44 × 0.44</td>
<td>Daily</td>
<td>25</td>
<td>Sanjay et al. (2017)</td>
</tr>
<tr>
<td>Observed + WFDEI</td>
<td>1999–2010</td>
<td>–</td>
<td>Daily</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
domain (Region 6) from two institutional runs: Indian Institute of Tropical Meteorology (CORDEX-IITM) and Swedish Meteorological and Hydrological Institute (CORDEX-SMHI) (Sanjay et al. 2011). Based on a comparison of seasonal precipitation patterns over the study area, the CORDEX-SMHI was selected for this study. The horizontal resolution and available periods for the CORDEX-SMHI data are shown in Table 1. CORDEX-SMHI data can be freely downloaded on request from http://cccr.tropmet.res.in/home/index.jsp. Catchment average mean daily temperature and annual precipitation (for the period 2000–2009) are found to be 6.1 °C and 5,431 mm/year, respectively.

**Topographical and land cover datasets**

In this study, a digital elevation model (DEM) of 90 m spatial resolution from NASA’s Shuttle Radar Topography Mission (NASA-SRTM) was used. The NASA-SRTM DEM is freely available for download from https://eros.usgs.gov/. The hydrology tool available in the software package, System for Automated Geoscientific Analysis (SAGA) (Conrad et al. 2015) under Qgis-2.18 (QGIS Development Team 2016), was used for automatic catchment delineation. Catchment delineation was based on the gauge point at Narayanghat and NASA-SRTM DEM. Catchment slope
and aspect data for Shyft (see the Hydrological Model section) were generated during the catchment delineation sub-processes in Qgis. The calculated catchment domain was gridded into 2 km × 2 km cells for input to Shyft. The centroid of each vector grid cell was calculated using Qgis as further input to the model.

A land cover map (0.5° × 0.5°, resolution) providing forest, lake, glacier, and reservoir cover of the study area was extracted from the Moderate Resolution Imaging Spectroradiometer (MODIS) land cover data (Chhanan et al. 2014). MODIS land cover data are available free online at http://glcf.umd.edu/data/lc/. Land cover fractions of forest, lake, glacier, and reservoir cover area at the centroid point per each grid cell within the catchment were also calculated. The python code and algorithm for calculating land cover fraction per each vector grid cell at centroid are available online at https://github.com/felix-matt/shyft-gis.

Spatial interpolation of observed and gridded forcing data

Shyft is able to ingest daily spatially distributed meteorological data as input variables. Here, we interpolate both observations (station data) and gridded forcing data (WFDEI, ERA-I, and CORDEX-SMHI).

The gridded forcing data have a coarser resolution than the simulation domain in the hydrological modeling framework of Shyft (Table 1). Accordingly, Shyft provides interpolation routines used for downsampling the forcing data sets to the model grid cell scale. Data interpolation methods were selected based on prior studies by Sapkota (2016), Lombraña (2017), Matt et al. (2018), Matt & Burkhart (2018), and Teweldebrhan et al. (2018). Observed temperature and gridded data were extrapolated/interpolated by three-dimensional Kriging where elevation is the third dimension (Oliver & Webster 1990). Observed precipitation, gridded precipitation, global radiation, WS, and RH data were extrapolated/interpolated by inverse distance weighting (IDW) (Shepard 1968). A precipitation gradient of −0.07 mm/100 m (Table 2) and a temperature gradient of −0.6°C/100 m are used during interpolation. The precipitation gradient was calculated from the observed precipitation data. The different interpolation methods along with their parameter values are presented in Table 2.

### Hydrological model

The Shyft modeling framework was used to simulate daily discharge from the catchment. The modeling framework has three main hydrologic modeling routines (https://gitlab.com/shyft-os). These three models are different in the way to calculate evapotranspiration and snow estimation and melt.

In this study, the PT_GS_K model was used for discharge simulation. The PT_GS_K model uses the Priestley–Taylor (PT) method (Priestley & Taylor 1972) for estimating potential evapotranspiration, the Gamma snow routine (GS; described in the Gamma Snow section) for snowmelt, sub-grid snow distribution and mass balance calculation, and a simple storage–discharge function (K) for catchment response calculation (Lambert 1972; Kirchner 2009). Actual evapotranspiration (AE) was assumed to take place from the snow-free area and was estimated as a function of potential evapotranspiration and a scaling factor.

The catchment response function ‘K’ is based on the storage–discharge relationship concept described in Kirchner (2009) and represents the sensitivity of discharge to storage changes as given by the following equation. Generally, discharge is nonlinear and varies by many orders of magnitude, and it is recommended to use log-transformed

![Table 2](https://iwaponline.com/hr/article-pdf/51/2/202/682207/nh0510202.pdf)
discharge values to avoid numerical instability.

$$\frac{d(ln(Q))}{dt} = g(Q) \left( \frac{P - E}{Q} - 1 \right)$$  \hspace{1cm} (1)$$

where $P$, $Q$, and $E$ are the rate of precipitation (input from either snow melt or rain), discharge, and $AE$ (from the snow-free area), respectively, in units of depth per time.

The idea behind this method is that the discharge sensitivity to changes in storage, i.e. $g(Q)$, can be estimated from the time series of the discharge alone through fitting empirical functions to the data, such as the quadratic equation (Kirchner 2009), which is given by:

$$g(Q) = e^{c_1 \cdot c_2 (ln(Q)) + c_3 (ln(Q))^2}$$  \hspace{1cm} (2)$$

where $c_1$, $c_2$, and $c_3$ are the catchment-specific outlet parameters (hereafter called Kirchner parameters) obtained during the model calibration. There is no routing function in the model, but the Kirchner response function represents a delayed outflow from storage within the catchment.

**Gamma snow**

In Shyft, the Gamma snow routine, an energy balance approach (Equation (3)) for snow ablation, and the snow depletion curve following a gamma distribution are combined into a single routine. The energy balance approach in the Gamma snow routine is based on DeWalle & Rango (2008) and also briefly explained in Hegdahl et al. (2016).

The net energy flux ($\Delta E$) at the surface available for snow ablation is expressed as follows:

$$\Delta E = S.(1 - \alpha) + L_{in} + L_{out} + H_{SE} + H_{L} + E_G$$  \hspace{1cm} (3)$$

where $S$ is the net shortwave radiation, $L_{in}$ and $L_{out}$ are the incoming and outgoing longwave radiations, $H_{SE}$ and $H_{L}$ represent sensible and latent heat fluxes, and $E_G$ is the net ground heat flux calculated using a bulk-transfer approach. Two parameters defining the wind profile, intercept (wind constant), and slope (wind scale) are determined either by model calibration or as provided (Table 3). For a given time step $(t)$, the snow albedo $(\alpha)$ at each cell depends on the minimum $(\alpha_{min})$ and maximum albedo $(\alpha_{max})$ as well as the albedo decay rate, temperature, and snowfall as described in Hegdahl et al. (2016):

$$\alpha_t = \begin{cases} \alpha_{min} + (\alpha_{t-1} - \alpha_{min}) \cdot \left( \frac{1}{2} \right) \cdot \frac{1}{FDR} & \text{if } T_a > 0^\circ C \\ \alpha_{t-1} + (\alpha_{max} - \alpha_{min}) \cdot \left( \frac{1}{2 SDR} \right) & \text{if } T_a \leq 0^\circ C \end{cases}$$  \hspace{1cm} (4)$$

In Equation (4), FDR and SDR denote fast and slow snow cover decay rates, respectively. In this study, $\alpha_{max}$ and $\alpha_{min}$ are prescribed (refer Table 3).

Within the Gamma snow routine, precipitation falling in each cell is classified as solid or liquid depending on a threshold temperature $(tx)$ and the actual cell temperature. Snow distribution within each cell is estimated by using a three-parameter gamma probability distribution. The third parameter in the gamma probability distribution represents the bare ground fraction in the cell. Finally, snowmelt depth (mm/day) is calculated by multiplying $\Delta E$ (available energy) with the latent heat of fusion for water.

A temperature index model which does not require glacier ice albedo was used to calculate glacier melt (see Hock 2005). The glacier reservoir was assumed to be infinite, and the glacier area was assumed to be constant throughout the simulation periods. Within a glacier-covered cell, glacier melt only happens from the snow-free fraction.

**Parameters and calibration**

Hydrological simulation from distributed models such as Shyft generally requires the estimation of model parameters through calibration with measured data (Madsen 2003). In the PT_GS_K routine, there are 14 parameters, out of which eight parameters (Table 3: top eight parameters) have been found to be significantly more sensitive than the rest (Teweldebrhan et al. 2018) and were selected for calibration in this study. The remaining six parameters were prescribed (Table 3: lower six parameters). The precipitation correction factor $(p_{corr\_factor})$, which is used to correct bias, was also set as a calibration parameter.

Manual calibration can be time-consuming and subjective; therefore, an automatic calibration was carried out.
Upper and lower limits for each parameter are shown in Table 3. CREST v.2.1, the Shuffle Complex Evolution University of Arizona (SCE-UA) (Duan et al. 1992), is used as the kernel algorithm in the automatic calibration process. Typically, the procedure involves the selection of samples in the parameter space through the use of competitive evolution schemes, such as the simplex scheme, to reproduce better the observations. After several iterations, either due to convergence or when the maximum number of iteration is reached, the best set of parameter values based on the Nash–Sutcliffe coefficient of efficiency (NSE) is determined (Chu et al. 2016).

In this study, the model was calibrated and validated for each forcing dataset independently (hereafter named ‘independent calibration mode’). The model calibration was based on observed discharge data from 2000 to 2004, and validation was based on data from 2005 to 2009.

### Water balance estimation

A water balance analysis is a useful tool to describe the principal components of water in and out of a catchment (Rochester 2010), where the volume of water inflow should be balanced with water outflow, assuming no changes in storage ($\Delta S/\Delta t$). The following water balance components: mean annual precipitation, discharge, and AE were calculated. Water balance components were calculated using calendar years, so that the change in storage is mainly a

### Table 3  Model calibration parameters with upper and lower limits

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description and unit</th>
<th>Parameter used in the submodel</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>Outlet empirical coefficient 1 (-) Equation (4)</td>
<td>K</td>
<td>-8.0</td>
<td>0.0</td>
<td>Sapkota (2016); Lombraña (2017)</td>
</tr>
<tr>
<td>$c_2$</td>
<td>Outlet empirical coefficient 2 (-) Equation (4)</td>
<td>K</td>
<td>-1.0</td>
<td>1.2</td>
<td>Sapkota (2016); Lombraña (2017)</td>
</tr>
<tr>
<td>$c_3$</td>
<td>Outlet empirical coefficient 3 (-) Equation (4)</td>
<td>K</td>
<td>-0.15</td>
<td>-0.05</td>
<td>Sapkota (2016); Lombraña (2017)</td>
</tr>
<tr>
<td>wind scale</td>
<td>Slope in turbulent wind function (m/s)</td>
<td>GS</td>
<td>1.0</td>
<td>6.0</td>
<td>Sapkota (2016); Lombraña (2017)</td>
</tr>
<tr>
<td>$t_x$</td>
<td>Temperature threshold rain/snow (°C)</td>
<td>GS</td>
<td>-3.0</td>
<td>2.0</td>
<td>Sapkota (2016); Lombraña (2017)</td>
</tr>
<tr>
<td>$FDR$</td>
<td>Fast albedo decay rate during melt (days)</td>
<td>GS</td>
<td>5.0</td>
<td>15.0</td>
<td>Hegdahl et al. (2016); Sapkota (2016)</td>
</tr>
<tr>
<td>$SDR$</td>
<td>Slow albedo decay rate during cold condition (days)</td>
<td>GS</td>
<td>20.0</td>
<td>40.0</td>
<td>Hegdahl et al. (2016)</td>
</tr>
<tr>
<td>$p_{corr_factor}$</td>
<td>Scaling factor for precipitation (-)</td>
<td>GS</td>
<td>0.4</td>
<td>2.0</td>
<td>Sapkota (2016)</td>
</tr>
<tr>
<td>$ae_scale_factor$</td>
<td>Scaling factor for AE (-)</td>
<td>AE</td>
<td>1.0</td>
<td>1.0</td>
<td>Sapkota (2016); Lombraña (2017)</td>
</tr>
<tr>
<td>$Surface_magnitude$</td>
<td>Snow heat constant (mm SWE)</td>
<td>GS</td>
<td>30.0</td>
<td>30.0</td>
<td>Sapkota (2016)</td>
</tr>
<tr>
<td>wind constant</td>
<td>Intercept in turbulent wind function (-)</td>
<td>GS</td>
<td>1.0</td>
<td>1.0</td>
<td>Lombraña (2017)</td>
</tr>
<tr>
<td>min albedo</td>
<td>Minimum snow albedo used in snow routine</td>
<td>GS</td>
<td>0.6</td>
<td>0.6</td>
<td>Sapkota (2016)</td>
</tr>
<tr>
<td>max albedo</td>
<td>Maximum snow albedo used in snow routine</td>
<td>GS</td>
<td>0.9</td>
<td>0.9</td>
<td>Sapkota (2016)</td>
</tr>
<tr>
<td>max water</td>
<td>Frictional max water constant of snow (-)</td>
<td>GS</td>
<td>0.1</td>
<td>0.1</td>
<td>Hegdahl et al. (2016)</td>
</tr>
</tbody>
</table>

In the table, ‘$K$’ is the catchment response function; ‘$GS$’ is the Gamma snow; ‘$PT$’ is the Priestley and Taylor; and ‘$AE$’ is the actual evapotranspiration.
difference in storage between first and last days of simulation.

The annual change in storage was calculated according to the following equation:

\[
\frac{\Delta S}{\Delta t} = (P + G) - (Q + ET)
\]  

(5)

where \(P\) is precipitation, \(Q\) is the discharge, \(ET\) is actual evapotranspiration, \(G\) is the glacier melt, and the unit of measurement is mm/year. \(\Delta S/\Delta t\) is the change in storage per time unit.

Model performance evaluation

To determine the agreements between observed and simulated discharges, NSE (Nash & Sutcliffe 1970), Kling–Gupta efficiency (KGE) (Gupta et al. 2009), square-root-transformed NSE (NSEsqrt) (Seiller et al. 2012), benchmark series (Gbenchmark) (Seibert 2001), and percentage volume difference (\(D_v\)) (Martinec et al. 1998) were used (Table 4). The NSE, NSEsqrt, and KGE efficiency measures are used to access the predictive power of the hydrological model and can range from \(-\infty\) to 1. An efficiency of 1 corresponds to a perfect match of simulated and observed discharges. To improve the estimation of the performance error, KGE considers three components: bias (\(\alpha\)), variability (\(\beta\)), and linear correlation coefficient (\(r\)) to overcome the problems associated with NSE, i.e. NSE results in the underestimation of the streamflow variability and the runoff peaks (Gupta et al. 2009). NSEsqrt gives more emphasis to the overall agreement between observed and simulated discharges (Seiller et al. 2012; Peña-Arancibia et al. 2013) as compared to NSE, so it is also included as an evaluation criterion. Percentage volume difference (\(D_v\)) gives the percentage bias between the simulated and observed series and can range from \(-\infty\) to \(+\infty\). \(D_v\) equals to zero indicates a perfect agreement between simulated and observed discharges. Since the observed discharge shows strong seasonal patterns, we also used a benchmark series (Gbenchmark) for model evaluation. The monthly average discharge was used as a benchmark series. Gbenchmark is negative if the model performance is poorer than the benchmark, zero if the model performs as well as the benchmark, and positive if the model is superior, with a highest value of one for a perfect fit. In this study, we aimed to achieve \(NSE > 0.7\), \(KGE > 0.7\), \(NSE_{sqrt} > 0.7\), \(G_{benchmark} > 0.5\), and \(D_v\) within ±15%, during both the model calibration and validation periods.

### Table 4: Definitions of the evaluation criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Mathematical expression</th>
<th>Description</th>
<th>Best value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(D_v)</td>
<td>(\frac{V_{obs} - V_{sim}}{V_{obs}} \times 100)</td>
<td>Volume difference in percentage</td>
<td>0</td>
</tr>
<tr>
<td>NSE</td>
<td>(1 - \frac{\sum_{i=1}^{n} (Q_{sim,i} - Q_{obs,i})^2}{\sum_{i=1}^{n} (Q_{obs,i} - Q_{obs})^2})</td>
<td>Nash-Sutcliffe efficiency</td>
<td>1</td>
</tr>
<tr>
<td>NSEsqrt</td>
<td>(1 - \frac{\sum_{i=1}^{n} \left( \frac{\sqrt{Q_{sim,i}}}{\sqrt{Q_{obs,i}}} \right)^2}{\sum_{i=1}^{n} \left( \frac{\sqrt{Q_{obs,i}}}{\sqrt{Q_{obs}}\sqrt{Q_{obs}}} \right)^2})</td>
<td>Squared-root transformed Nash-Sutcliffe efficiency</td>
<td>1</td>
</tr>
<tr>
<td>KGE</td>
<td>(\sqrt{r - 1}^2 + (\alpha - 1)^2 + (\beta - 1)^2)</td>
<td>Modified Kling–Gupta efficiency</td>
<td>1</td>
</tr>
<tr>
<td>Here, (\beta = \frac{Q_{sim}}{Q_{obs}}); (\alpha = \frac{Q_{sim, std}}{Q_{obs, std}})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(G_{benchmark})</td>
<td>(\frac{\sum_{i=1}^{n} (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^{n} (Q_{obs,i} - Q_{benchmark,i})^2})</td>
<td>Goodness of fit with respect to the benchmark series</td>
<td>1</td>
</tr>
</tbody>
</table>

\(n\): total number of days in the evaluation period; \(Q_{sim}\) and \(Q_{obs}\): simulated and observed discharges; \(Q_{sim, std}\) and \(Q_{obs, std}\): standard deviation for simulated and observed discharges; \(Q_{sim}\) and \(Q_{obs}\): arithmetic mean for simulated and observed discharges; \(V_{sim}\) and \(V_{obs}\): total discharge volume for simulated and observed discharges; \(Q_{benchmark}\): monthly long-term average observed discharge.
RESULTS

Meteorological forcing data analysis

As discussed in the section ‘Reanalysis and regional climate model data’, the resolutions of gridded datasets are different. For the sake of comparison, Shyft-interpolated meteorological variables (for the period 2000–2009) for each forcing dataset were used. Temperature shows an overall agreement where all three datasets demonstrate similar seasonality and distinct cool or negative bias as compared to the observations during the winter and pre-monsoon seasons (Figure 3). The bias was the strongest in the CORDEX-SMHI data during the September–December period. Long-term spatially averaged temperature for July (Figure 4) shows that the CORDEX-SMHI was not able to capture higher daily average temperatures in the river valley. Daily average temperatures during July over the lower elevation region from ERA-I and CORDEX-SMHI were lower than the temperatures from the Observed and WFDEI interpolated datasets. What was most interesting is that negative catchment average temperatures, which were not at all captured by the interpolated observations (Figure 3), will have a significant impact on any simulation. We attribute the lack of negative catchment average temperatures to the lack of stations above 4,000 m a.s.l. combined with a too low-temperature lapse rate. This is a typical challenge when observations lack representativity to topography.

Precipitation varies notably among the forcing datasets. WFDEI precipitation agrees well with the observation (Figures 3 and 4). The strongest positive bias was observed for CORDEX-SMHI and ERA-I datasets (Figure 5). Long-term daily averaged precipitation for July–September shows the strongest deviation among forcing. The catchment precipitation for July was more than 150 mm/day in CORDEX-SMHI and ERA-I datasets, and the highest precipitation was found in the lower elevations. A seasonal comparison shows that the WFDEI was slightly higher than the observed precipitation during the monsoon season (Figure 6). Similar findings are reported for the Indus catchment by Dahri et al. (2016).

Monthly averaged wind speeds and incoming short-wave radiation from WFDEI and ERA-I were fairly similar but different for CORDEX-SMHI. Station
observations for RH, WS, and shortwave incoming global radiation were not available for the catchment, but previous global studies by Iizumi et al. (2014) and Weedon et al. (2014) showed that the WFDEI is more representative for the observations.

Model parameters

Shyft was calibrated for each forcing dataset, and the calibrated parameters are shown in Table 5. Normal ranges for all calibrated parameters are given in Table 5. The

Figure 5 | Distributed daily mean Shyft-interpolated precipitation (mm) for July over the Narayani river catchment for each forcing dataset.

Figure 6 | Seasonal precipitations from different forcing datasets.
highest range in values (greater than 0.15 variance in the normalized scale of $/C_0 1$ to 1) was seen for the parameters $p_{corr\_factor}$ ($0.4$–$1.34$), wind scale ($2.01$–$4.84$), threshold temperature ($tx$) ($/C_0 1.68$ to $0.94$), and slow albedo decay rate ($29.04$–$39.98$). A relatively small variation in parameter values (less than 0.05 variance in the normalized scale of $/C_0 1$ to 1) was seen for the Kirchner coefficients ($c_1$ ($/C_0 6.64$ to $5.51$), $c_2$ ($0.14$ to $0.44$), and $c_3$ ($/C_0 0.05$ to $0.07$)) and fast albedo decay rate ($9.94$–$12.03$). The highest variance (0.27) in a normalized scale of $1$ to $1$ was found for $p_{corr\_factor}$. The lowest and highest $p_{corr\_factor}$ values were observed for CORDEX-SMHI ($0.4$) and Observed + WFDEI ($1.34$) forcing dataset, respectively. The higher $p_{corr\_factor}$ for Observed + WFDEI is interesting and suggests an underprediction of precipitation in the catchment when gridded precipitation was based on Shyft-interpolated station observations. Unique to the CORDEX-SMHI dataset, the Gamma snow threshold temperature ($tx$) was found to be positive (0.94). Less sensitive parameters as reported by Teweldebrhan et al. (2018) were not calibrated but are listed with given values in Table 3.

**Evaluation of discharge simulation using different forcing datasets**

Figure 7 compares daily observed and simulated discharges using the different forcing datasets. Calibration and validation results in terms of error statistics are shown in Table 6. Simulations from Observed + WFDEI forcing give the best results in terms of error statistics during both calibration ($NSE = 0.90$, $NSE_{sqrt} = 0.91$, $KGE = 0.94$, $G_{bench} = 0.45$, and $D_v = -0.91$) and validation ($NSE = 0.90$, $NSE_{sqrt} = 0.93$, $KGE = 0.92$, $G_{bench} = 0.56$, and $D_v = -1.29$) periods. $D_v$ in both calibration and validation was also best for Observed + WFDEI and found to be less than $-2\%$. The second best performance was achieved for WFDEI (Table 6). For $NSE$, $NSE_{sqrt}$, and $KGE$, the ERA-I and WFDEI datasets were similar, but the volume differences during calibration were comparatively higher for ERA-I than WFDEI. Comparatively poorer model performance and higher volume differences (Table 6) were found

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unit</th>
<th>Observed + WFDEI</th>
<th>WFDEI</th>
<th>ERA-I</th>
<th>CORDEX-SMHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>–</td>
<td>$-5.51$</td>
<td>$-6.09$</td>
<td>$-6.64$</td>
<td>$-5.99$</td>
</tr>
<tr>
<td>$c_2$</td>
<td>–</td>
<td>$0.44$</td>
<td>$0.25$</td>
<td>$0.14$</td>
<td>$-0.20$</td>
</tr>
<tr>
<td>$c_3$</td>
<td>–</td>
<td>$-0.05$</td>
<td>$-0.05$</td>
<td>$-0.05$</td>
<td>$-0.07$</td>
</tr>
<tr>
<td>$tx$</td>
<td>◦C</td>
<td>$-1.68$</td>
<td>$-1.38$</td>
<td>$-0.36$</td>
<td>$0.94$</td>
</tr>
<tr>
<td><em>wind scale</em></td>
<td>m/s</td>
<td>$2.01$</td>
<td>$4.70$</td>
<td>$4.84$</td>
<td>$3.40$</td>
</tr>
<tr>
<td><em>FDR</em></td>
<td>days</td>
<td>$11.59$</td>
<td>$12.03$</td>
<td>$9.94$</td>
<td>$10.75$</td>
</tr>
<tr>
<td><em>SDR</em></td>
<td>days</td>
<td>$39.98$</td>
<td>$32.66$</td>
<td>$29.04$</td>
<td>$31.55$</td>
</tr>
<tr>
<td><em>$p_{corr_factor}$</em></td>
<td>–</td>
<td>$1.34$</td>
<td>$1.09$</td>
<td>$0.41$</td>
<td>$0.4$</td>
</tr>
</tbody>
</table>

Figure 7 | Calibrated and simulated daily discharge for the Narayani river catchment for the period 2000–2009.
for the model calibrated with the CORDEX-SMHI dataset. Goodness of fit with respect to the benchmark ($G_{\text{bench}}$) for ERA-I and CORDEX-SMHI during calibration periods was found negative, indicating poorer performance than the benchmark series. However, it should be noted that the peak discharge was best simulated with CORDEX-SMHI data (Figure 7).

In Figure 8, the year 2004 was plotted to show in more details the representation of the annual cycle by the different forcing datasets. Overall, all datasets were able to simulate the cycle well; however, the CORDEX-SMHI simulation deviates somewhat from the rest and shows a higher peak flow early in the wet season.

Quantile–Quantile (QQ) plots for the different forcing datasets are shown in Figure 9. The QQ-plot for the calibration period is shown in Figure 9(a), for the validation period in Figure 9(b). To highlight low flows, Figures 9(c) and 9(d) show the QQ-plot on log-scale. The QQ-plot is a graphical technique to determine if two datasets come from the same population with a common distribution (Renard et al. 2010). Departures from the 1:1 reference line indicate discrepancies between the simulated and the observed discharge. Figure 9(a) shows that all simulations slightly overestimate discharge (with respect to the observed discharge) up to a value of around 4,000 m³/s. Between 4,000 and 8,000 m³/s, the CORDEX-SMHI simulations were higher than observed, whereas other simulations were lower than observed. After 8,000 m³/s, CORDEX-SMHI simulations were able to capture peaks where others simulations were lower than observed. Although the

| Table 6 | Performance statistics for different forcing datasets |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|
| Error parameters | Observed - WFDEI | WFDEI | ERA-I | CORDEX-SMHI |
| | Calibration | Validation | Calibration | Validation | Calibration | Validation |
| NSE | 0.90 | 0.90 | 0.83 | 0.80 | 0.82 | 0.83 | 0.60 | 0.48 |
| KGE | 0.94 | 0.92 | 0.86 | 0.81 | 0.82 | 0.81 | 0.73 | 0.52 |
| $D_v$ | -0.91 | -1.29 | -1.72 | -5.68 | -4.01 | -5.53 | -19.9 | -42.4 |
| $NSE_{\text{mean}}$ | 0.91 | 0.93 | 0.88 | 0.84 | 0.88 | 0.87 | 0.70 | 0.57 |
| $G_{\text{bench}}$ | 0.43 | 0.56 | 0.01 | 0.07 | -0.03 | 0.19 | -1.34 | -1.47 |

Figure 8 | Daily observed and simulated discharges from different forcing models for the year 2004.

Downloaded from https://iwaponline.com/hr/article-pdf/51/2/202/682207/nh0510202.pdf by UNIVERSITY OF OSLO user on 12 May 2020
Observed + WFDEI forcing dataset resulted in discharge simulations with the best performance, it was not able to capture the highest observed discharge (simulated discharge points after 95% quantile). The simulations from CORDEX-SMHI were able to capture relatively high peak-discharge events (above the 99% quantile); however, they have large deviations from the observations between 90% and 95% quantiles. For the low flows (i.e. 1–10% quantile), the simulation from the WFDEI forcing captures observation better than the Observed + WFDEI forcing (Figure 9(c)). However, low-flow simulation from ERA-I until 5% quantile was better than the rest.

Similarly, in the validation period (Figure 9(b)), simulated discharge and observed daily discharge for values less than 90% quantile from all forcing datasets fit well to the observations. Above 95% quantile, most forcing datasets fail to capture these highest discharges. Interestingly, the CORDEX-SMHI, which generally provides lower performance (Table 6), notable in the lower flow range, manages to capture the highest flows best. However, these
simulations were mostly higher than the reference line, indicating a general overestimation during the full range of simulations. Similar to the calibration periods, low-flow simulation from WFDEI was also found closest to the reference (observations) during the validation periods (Figure 9(d)).

Water balance analysis

Figure 10 compares the catchment average (2000–2009) water balance components for the four different forcing datasets (the exact values are given in Table 7). Annual averages over the 10-year simulation period, 25.1%, 24.8%, 30.7%, and 17.8% of precipitation, were lost to evapotranspiration for Observed + WFDEI, WFDEI, ERA-I, and CORDEX-SMHI datasets, respectively. Among all forcing datasets, Observed + WFDEI shows lower runoff than the others. Higher runoff from CORDEX-SMHI forcing was observed, and it is associated with higher average precipitation during the monsoon period (Figure 3). The lowest (72.8 mm) total glacier melt during the simulation period was observed for the WFDEI dataset, while the highest (297.3 mm) was observed from ERA-I. The change in storage (Equation (5)) shows that 73.0, 29.2, 68.7, and 32.7 mm of water were surplus (a positive change in storage) in the catchment for Observed + WFDEI, CORDEX-SMHI, ERA-I, and WFDEI forcing data, respectively. The largest glacier melt contribution (19%) to total runoff was observed for ERA-I forcing, while the lowest was observed for WFDEI (4%). Different water balance components and their associated changes in storage are shown in Table 7.

DISCUSSION

Discussion on model parameters

As with any hydrologic model, the model used in this study is sensitive to the parameters; particularly the Kirchner coefficients, precipitation correction factor, and the threshold temperature (Teweldebrhan et al. 2018). While calibrating the model, different values of the parameters were obtained for different forcing datasets (Table 5). An examination of the model parameters revealed that the calibrated values of the precipitation correction factor, wind scale, threshold temperature, and slow albedo decay rates differ a lot

<table>
<thead>
<tr>
<th>Water balance components</th>
<th>Observed + WFDEI</th>
<th>WFDEI</th>
<th>ERA-I</th>
<th>CORDEX-SMHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>1,771.1</td>
<td>1,965.7</td>
<td>1,901.4</td>
<td>2,234.1</td>
</tr>
<tr>
<td>Evapotranspiration</td>
<td>449.6</td>
<td>487.4</td>
<td>584.6</td>
<td>399.5</td>
</tr>
<tr>
<td>Glacier melt</td>
<td>242.2</td>
<td>72.8</td>
<td>297.3</td>
<td>132.5</td>
</tr>
<tr>
<td>Runoff</td>
<td>1,490.7</td>
<td>1,521.9</td>
<td>1,545.4</td>
<td>1,934.4</td>
</tr>
<tr>
<td>Change in storage</td>
<td>73.0</td>
<td>29.2</td>
<td>68.7</td>
<td>32.7</td>
</tr>
</tbody>
</table>

Figure 10 | Annual average (2000–2009) water balance components (Evap, actual evapotranspiration; Prec, precipitation calculated by using p_corr_factor) for different forcing datasets in the Narayani catchment.
between the forcing datasets. Larson & Peck (1974) and Lakew et al. (2017) pointed out that the model simulations are highly sensitive to the precipitation correction factor, which helps adjusting systematic bias embedded in the precipitation forcing. The precipitation correction factor may account for both under catch and lack of representative stations (Lakew et al. 2017) and is directly related to the bias in precipitation data. Prior studies by Engeland et al. (2016) and Teweldebrhan et al. (2018) suggested that the Kirchner parameters are highly sensitive since they have a large influence on a given simulation. We found that the variance for Kirchner response parameters was less than 0.05. An explanation is that these parameters are not dependent on the forcing datasets but rather on the physical catchment characteristics and thereby the concentration time of the catchment. Kirchner response parameters are dependent on the physical catchment characteristics and thereby the concentration time of the catchment and are therefore less sensitive to the forcing.

**Potential factors controlling hydrological model efficiency during model calibration and validation**

Different forcing datasets lead to different levels of performance efficiency in terms of discharge simulations (Table 6). Explaining these differences is not easy, as each forcing dataset has different characteristics such as the spatial resolution and the assimilation methods. Figures 3–5 present the catchment average monthly mean forcing variables for the different forcing datasets. From the figures, we can see that in general, the spatio-temporal distribution of precipitation from WFDEI was quite similar to the interpolated observations. The extent of higher precipitation than observation during pre-monsoon and monsoon seasons was larger for ERA-I and CORDEX-SMHI. Overestimation of precipitation during the monsoon season identified for CORDEX-SMHI in the Himalayan region (Ghimire et al. 2018) and for the ERA-I over the Indus catchment (Dahri et al. 2016). Moulin et al. (2009) and Zhu et al. (2014) indicate that a poor representativeness of precipitation over the catchment is a major source of uncertainty in discharge simulation.

Despite some of the biases in the forcing data, the performance of the model during calibration and validation showed that all datasets performed reasonably well. The good agreement between simulated and observed discharges during the model calibration and validation periods (Figure 7) is to a large extent due to the application of the precipitation correction factor. The multiplicative correction factor to the precipitation removes a substantial proportion bias in both precipitation and simulated discharges. It should be clarified here that such an adjustment of bias in discharge (through the input forcing correction) is an ‘ad hoc’ procedure to overcome the quality of forcing data. The ‘p_corr_factor’ shown in Table 5 indicates that the CORDEX-SMHI was reduced by a factor of 0.4, and ERA-I precipitation was reduced by a factor of 0.41 during the calibration period to minimize bias in discharge estimation. These numbers indicate that there were significant biases in these two forcing datasets and suggest that some kind of bias correction should be considered in any application. We also notice that the precipitation correction factor for WFDEI was near one. This is reasonable and encouraging, as this dataset is already bias-corrected (using observed precipitation). The precipitation correction factor for observed datasets was 1.34, indicating that the areal precipitation is underestimated. This underestimation is likely caused by a combination of undercatch of the precipitation gauges and the lack of representative observation stations. The undercatch depends on the precipitation gauge, wind, and precipitation phase (Mekonnen et al. 2015; Zhao et al. 2019), and the non-representativeness is caused by the limited number of observation stations at high altitudes. Better performance in discharge simulation by the application of a correction factor to precipitation was also shown by previous studies Larson & Peck (1974), Lawrence et al. (2009), and Lakew et al. (2017) and is a standard practice in the hydrologic analysis. The improvement of simulation from all forcing datasets versus prior results presented by Bhattarai et al. (2018) using the HBV model (Bergström 1995) in the same catchment is likely attributed to the distributed nature of Shyft.

Though we achieved high NSE, NSE$_{exp}$, and KGE using WFDEI, ERA-I, and Observed + WFDEI forcing datasets with reasonable $D_v$ (Table 6), the simulations were still not able to capture the highest discharges as seen in the QQ-plots in Figure 9. As compared to high discharge, low discharge events were well captured by all forcing datasets.
Local simulated peak discharge during the monsoon season may be attributed to less representativeness of the input precipitation data. Particularly, the lack of ability to capture intense localized precipitation events could be the main reason for an inability to simulate high discharge concurrent with the observations. The CORDEX-SMHI simulation generated relatively higher discharge peaks as compared with the other datasets but has overall higher positive biases during the pre-monsoon season (Figure 3). The CORDEX-SMHI forcing datasets could provide an opportunity for the peak discharge analysis in the Himalayan catchment.

Discussion on the water balance analysis

The catchment water balance component analysis (Table 7) revealed notable differences among the forcing datasets. The discrepancies can, to some degree, be attributed to the uncertainty inherent in each forcing dataset. As mentioned by Nkiaka et al. (2017), different precipitation in each forcing dataset can strongly influence the optimized parameters that control the rates and threshold of hydrological processes in the catchment. Although measured evapotranspiration data are not available for comparison, the percentage of evapotranspiration estimates from WFDEI (24.79%) was found to be similar to the previous study by Sakai et al. (2004), Giertz et al. (2006), Sintondji et al. (2014), and Ragettli et al. (2015) in similar Himalayan catchments. The lowest evapotranspiration (399.5 mm/year) was found for the CORDEX-SMHI forcing dataset, which can be associated with high RH, and lower average shortwave radiation (Figure 3) than other forcing data. Lower evaporation (449.6 mm/year) and lower precipitation (1,771.1 mm/year) from the Observed + WFDEI than WFDEI may partly result from precipitation under-catch in the catchment.

Higher positive storage changes were found for the ERA-I and Observed + WFDEI forcing datasets as compared to CORDEX-SMHI and WFDEI. As compared to Observed + WFDEI forcing dataset, a lower storage change for ERA-I was observed and results from higher evapotranspiration in ERA-I. A change in storage was found smaller than the glacier melt for all forcing datasets. A smaller change in storage is due to the fact that it originates from snow that accumulates over the year and contributes to the glacier mass balance. Smaller changes in storage also indicate that the glacier mass balance was negative over the simulation period. Previous studies by Kulkarni (1992), Cogley et al. (2010), Bolch et al. (2012), Wagnon et al. (2013), and Gurung et al. (2016) also showed that the Himalayan glaciers are experiencing a higher retreat rate in recent decades. The highest glacier melt contribution was observed for ERA-I (19%), while the lowest glacier melt contribution is observed for WFDEI (4%). A study by Gupta et al. (2019) on the Marshyangdi River catchment (with 24% glacier cover area) shows that glacier melt contributes to 11.8% of the total discharge. A relatively lower percentage of snow and glacier melt contribution to the total discharge from the Modik River catchment (with 12% glacier cover area) is presented by Kayastha & Shrestha (2019). Similarly, a study by Nepal (2016) on the Dudhkoshi catchment (with 15% glacier cover area) shows that glacier melt contribution is 17% of the total discharge. A similar percentage of glacier melt contribution to the total discharge for Observed + WFDEI (16%) was observed for the Narayani river catchment, although the glacier cover area was only 8%. As suggested by Bonekamp et al. (2019), differences in glacier melt contribution to total discharge result from differences in meteorological forcing, as we also observed in our study.

Uncertainty in the model simulation and observation

Hydrological projections are subject to considerable uncertainty (Dobler et al. 2012) and are easily affected by various factors, including local and climatic conditions, optimized parameters (Shen et al. 2012), and the quality of forcing data (Teweldebrhan et al. 2018). Different factors also have a varying degree of impact on the discharge simulation. The challenge using a model with many model parameters is that we might get a good model fit but a less robust model for prediction and forecasting due to over-fitting. However, the uncertainty in the simulations will almost certainly be higher due to the increased uncertainty in the parameter values. In our study, the discharge was highly seasonal and particularly precipitation dependent. Inclusion of the precipitation correction factor as a calibrating parameter improves the predictability of the model at the expense of increases to the uncertainty in water balance components like evapotranspiration and snowmelt. However, only a single value of the precipitation correction
factor \( (p\_corr\_factor) \) was applied. More complex precipitation correction could include seasonality and impact of orographic effects, though we feel this would result in greater uncertainty. Nonetheless, uncertainties remain.

First, uncertainties in the observed data result from the uneven spatial distribution and few monitoring meteorological stations, which are mainly located in lower elevation regions (Figure 2). Furthermore, many observations are manual. This poses a unique source of uncertainty in measurements, not necessarily greater or less from automated. Furthermore, there are numerous sources of errors associated in the establishment of the river stage and parameterization of rating curves. During model calibration, we assumed that observed discharge is correct. Some of deviations between model simulations and observations might actually be explained by errors in the observations and not in the model or the forcings, to assess the uncertainty in streamflow observations with required detailed knowledge about local river profiles and data used to establish the rating curve and it is outside the scope of this paper. Second, uncertainties might originate from the estimation of AE and other water balance components, which again is influenced by the precipitation estimates and other forcing data used for driving the model. And thirdly, there are the uncertainties in the model formulation itself. The model in this study assumed that the discharge in the river depends on the amount of water stored in the catchment, and we did not consider the impact of industrial and households’ water consumption or any other regulation. Despite this, the simulation performance achieved by the model presented here is quite good, though the results are limited to the one catchment in Nepal. A further examination should evaluate whether the selected forcing datasets could be applied to simulated discharge at other mountain regions with longer time periods.

**CONCLUSIONS**

We aimed to identify the quality of discharge simulation for the Narayani catchment of Nepal based on different forcing datasets. The forcing datasets WFDEI, CORDEX-SMHI, ERA-I, and ground-based observation combined with WFDEI (i.e. Observed + WFDEI) were used to estimate the Narayani catchment discharge on a daily basis for the period of January 2000–December 2009. Because of the uniform coverage and data consistency, global forcing datasets from global and regional climate models were considered an important supplement to station data. In this study, the distributed Shyft hydrologic simulation platform was selected as a discharge simulation tool. The forcing data were interpolated using IDW for generating daily observed precipitation, while a gradient-based Kriging method was used for generating temperature fields matching the model resolution. In this study, Shyft was calibrated for the period of 2000–2004 and validation was done from 2005 to 2009.

Our analysis showed that large differences exist between different forcing datasets particularly in the amount of precipitation. Precipitation from ERA-I and CORDEX-SMHI was unrealistic leading to poor model performance. To improve the existing bias in precipitation particularly over the high mountain regions, further validation and algorithm improvements are required. With the application of a precipitation correction factor \( (p\_corr\_factor) \), a significant amount of bias in precipitation could be mitigated. Still, the performance of the bias-corrected forcing data, i.e. WFDEI, was better than the climate model datasets. Comparing model performance during model calibration and validation periods, relatively higher NSE, NSE\textsuperscript{sqrt}, and KGE with lower \( D_p \) were found for the Observed + WFDEI and WFDEI forcing datasets. The water balance analysis shows that higher evapotranspiration with large glacier melt is observed for the ERA-I forcing dataset, while the average evapotranspiration calculated from the WFDEI (24.79%) is similar to previous studies.

Therefore, based on the different results from the Shyft, we conclude that, in the data-poor Himalayan catchment, the WFDEI forcing dataset may be the best choice for water resource planning and hydropower inflow calculations. Discharge simulations resulting from the WFDEI forcing data were particularly promising for hydropower estimation and water resource assessment in data-scarce or ungauged regions. CORDEX-SMHI data captured higher peaks and may be suitable for the peak discharge analysis. However, to use CORDEX-SMHI and ERA-I in ungauged catchments for the water balance analysis, bias correction is required. Further analysis by implementing Shyft in the different Himalayan catchments for different forcing
datasets is a recommended further step to assess the regional extensibility of the current results.

**FUNDING**

This work is a contribution to the Strategic Research Initiative ‘Land Atmosphere Interaction in Cold Environments’ (LATICE) of the University of Oslo and partially supported through the Norwegian Research Council’s INDOM program under the Hydrologic sensitivity to Cryosphere-Aerosol interaction in Mountain Processes (HyCAMP) project (NRF No. 222195)

**ACKNOWLEDGEMENT**

We are thankful to the Department of Hydrology and Meteorology, Government of Nepal for providing observed hydrological and meteorological datasets. We also gratefully acknowledge Special Issue Editor Dr. Kolbjørn Engeland and anonymous reviewer for their comments on the research and manuscript.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this paper is available online at https://dx.doi.org/10.2166/nh.2020.079.

**REFERENCES**


System for automated geoscientific analyses (SAGA) v. 2.1.4. Geoscientific Model Development Discussions 8 (2), 2271–2312.


Lombrana, J. U. 2017 Evaluation of Snow Simulations in SHyFT. Norwegian University of Science and Technology (NTNU), Trondheim.


First received 6 June 2019; accepted in revised form 14 January 2020. Available online 21 February 2020
Paper III: Impact of catchment discretization and imputed radiation on model response: a case study from central Himalayan catchment
Impact of catchment discretization and imputed radiation on model response: a case study from central Himalayan catchment

Bikas C. Bhattarai¹, Olga Silantyeva¹, John F. Burkhart¹, Aynom T. Teweldebrhan¹, Sighjørn Helset², and Ola Skavhaug³

¹Department of Geosciences, University of Oslo, Oslo, Norway
²Statkraft AS, Lilleaker, NO-0216 Oslo Norway
³Expert Analytics, Norway

Correspondence: Bikas Chandra Bhattarai (b.c.bhattarai@geo.uio.no)

Abstract. Distributed and semi-distributed hydrological modelling approaches commonly involve the discretization of a catchment into several modelling elements. Although some modelling studies were conducted using triangulated irregular networks (TIN) previously, a little attention has been given to assess the impact of TIN as compared to the standard catchment discretization techniques. Here we examine how different catchment discretization approaches and radiation forcings influence hydrological simulation results. Three catchment discretization methods, i.e. elevation zones (Hypsograph) (HYP), regular square grid (sqGrid), and TIN were evaluated in a highly steep and glacierized Marsyangdi-2 river catchment, central Himalaya, Nepal. To evaluate the impact of radiation on model response, shortwave radiation was translated using two approaches, one with the measured solar radiation assuming a horizontal surface and another with a translation to slopes. The results indicate that the catchment discretization has a great impact on simulation results. Evaluation of the simulated streamflow value using Nash-Sutcliffe efficiency (NSE) and log-transformed Nash-Sutcliffe efficiency (LnNSE) shows that highest model performance was obtained when using TIN followed by HYP (during the high flow condition) and sqGrid (during the low flow condition). Similar order of precedence in relative model performance was obtained both during the calibration and validation periods. Snow simulated from the TIN based discretized models was validated with Moderate Resolution Imaging Spectroradiometer (MODIS) snow products. Critical Success Indexes (CSI) between TIN based discretized model snow simulation and MODIS snow were found satisfactory. Bias in catchment average snow cover area from the models with and without using translated radiation is less than two percent, but implementation of translated radiation into Shyft gives better CSI with MODIS snow.

Copyright statement. TEXT

1 Introduction

Accurate runoff prediction is one of the fundamental challenges for effective and sustainable water resource management in the Himalaya region. Reliable meteorological forcing data and accurate land-cover information at a range of spatial and temporal scales are critical to effective runoff prediction, and for water resource management (Xu, 1999). The main challenge
to estimate (or measure) meteorological forcing variables comes from the fact that they vary both in space and time as a function of meteorological inputs. Spatial variability of topography and land use cover types also adds to the complexity of the problem (Li et al., 2015).

In the past, several hydrological models have been developed for use in various applications. Some of these models include: SRM (Martinec et al., 1983), HEC-HMS (Halwatura and Najim, 2013), J2000 (Nepal et al., 2014), Statkraft’s Hydrological Forecasting Toolbox (Shyft) (Burkhart et al., 2020) etc. Most of the distributed hydrological models discretize the catchment and derive terrain attributes from digital elevation models (DEMs). Most commonly used catchment discretization methods in hydrological applications are grid-based e.g. (Kolberg et al., 2006; Bell et al., 2007; Hegdahl et al., 2016; Bhattarai et al., 2020), and elevation band (hypsography) based e.g. (Martinec et al., 2008; Pradhananga et al., 2014). Recently, triangulated irregular networks (TIN) based discretized hydrological models are also used for discharge simulations (Singh and Fiorentino, 1996; De Wulf et al., 2012), but till date TIN based distributed hydrological models are not tested in the Himalaya catchments. Although the use of grid and elevation based discretized models are more prevalent due to their simplicity and ease of implementation (Freer et al., 1997; Woods et al., 1997), TIN based discretized models have a better potential of accurately representing topographic features than grids and hypsography (Singh and Fiorentino, 1996). In this context, Shyft can be a valuable tool to evaluate different models as it have the ability to take the terrain heterogeneity into account, and provide distributed output variables. Shyft can be used to test different functioning hypotheses from which simplified and/or predictive models can be derived for more operational purposes. If reliable output variables are expected, Shyft parameters can be estimated a priori from available information. Verification of these models behaviour can be performed not only on the discharge at the outlet but also at intermediate variables (such as snow cover area, snow water equivalent etc) leading to a multi-objective verification.

The topographical information such as snow cover area is one of the most important variables for the application of hydrological model (Diffenbaugh et al., 2013; Lutz et al., 2014; Sam et al., 2018), and can dominate local and regional climate, and hydrology. Snow and glacier strongly influence the regional radiation budget with its high albedo and acts as an insulation, controlling snow accumulation and melt (Braakmann-Folgmann and Donlon, 2019). In this regard, accurate representation of snow cover at micro-scale of spatial variation is required to provide accurate and timely information on water availability and its management. Many distributed hydrological models are commonly employed to quantify snow cover area and snow water equivalent (SWE). All models have strengths and limitations (Parajka et al., 2012; Teweldebrhan et al., 2018). For example, current model-based snow simulation approaches in Shyft are limited to the radiation received on a horizontal plane surface, which introduce a certain degree of uncertainty in the radiation received by the surface. In view of this, we have implemented the imputed radiation algorithm following a similar approach as in Allen et al. (2006) in order to account for the effect of slope and aspect on incident radiation.

Remote sensing is an alternative powerful tool that offers the ability to examine quantitatively the physical properties of snow (Li et al., 2017). A well-known example is a MODerate Resolution Imaging Spectroradiometer (MODIS) instrument which can provide daily snow cover fraction with nearly global coverage at the resolution of 500 km (Hall et al., 2002). For regional snow cover mapping, the MODIS satellite sensors are particularly appealing due to their high Spatio-temporal
resolution (Urraca et al., 2017). In this regard, validation of snow cover area simulated from the Shyft with MODIS will also add the most important aspect to hydrological simulation.

Despite a number of previous modelling studies, the conversion of a highly-variable meteorological field into a spatially-distributed snow response remains an unanswered, open question. This is primarily due to the fact that meteorological processes operating in a watershed are highly heterogeneous and interconnected in the system (Fan et al., 2019). In this regard, we have implemented translated radiation following a similar approach to Allen et al. (2006) on different catchment discretized models (i.e. elevation band, regular spaced grid, and TIN) for calculating snow cover area using PT_GS_K stack of Shyft (https://shyft.readthedocs.io). The primary method for the simulation and distribution of snow cover area in PT_GS_K is in the matrix format of equally spaced elevation points. A data point in a matrix does not allow meaningful analysis of the terrain. Hence we hypothesize that the Shyft using the TIN for discretizing a watershed should enable a better prediction of the change in a runoff pattern caused by a change in a watershed characteristics. Finally, we tried to answer the question, what value does the TIN format offer over a matrix format for this particular application. This study also presents and evaluates the different catchment discretization models and the newly added translated radiation algorithm on the snow cover area and discharge simulations. As of our best knowledge, TIN based discretized model with imputed radiation has not been applied before to simulate discharge in the Himalayan river catchments.

2 Materials and Methods

2.1 Hydrologic model framework

The Statkraft Hydrological Forecasting Toolbox (Shyft, https://gitlab.com/shyft-os) was used in this study. Shyft provides an optimized platform for the implementation of well know hydrological models (Burkhart et al., 2020). High performance generic time series framework in Shyft allows for rapid calculations of hydrologic response at the regional to a cell scale.

The standard models in Shyft, i.e. PT_GS_K, PT_SS_K, and PT_HS_K are mainly differ in their snow mass balance methods calculations. The main model forcing variables are temperature, precipitation, wind speed, relative humidity, and shortwave radiation. Shyft works down from the regional level to the cell level by distributing the input variables and parameters to the georeferenced grid cell (described in catchment discretization technique). In the model, Bayesian kriging approach is used to distribute point temperature data over the catchment, while the inverse distance weighting approach is used for other forcing variables. The PT_GS_K model is the widely used as compared to the remaining models of Shyft (Teweldebrhan, 2019). Potential evapotranspiration in this model is calculated by Priestley-Taylor method (Priestley and Taylor, 1972). Actual evapotranspiration is assumed to take place from snow-free area and is estimated as a function of potential evapotranspiration, and a scaling factor. Gamma snow routine (Section 2.1.1) is used for snowmelt, sub-grid snow distribution and mass balance calculation. Simple storage-discharge function is used for catchment response calculation (Lambert, 1972; Kirchner, 2009). The catchment response function is based on the storage-discharge relationship concept described in (Kirchner, 2009) and
represents the sensitivity of discharge to storage changes given by Eq. 1. Since, discharge is generally nonlinear and varies by many orders of magnitude, the recommended approach is to use log-transformed discharge values to avoid numerical instability.

\[
\frac{d(\ln(Q))}{dt} = g(Q) \left( \frac{P - E}{Q} - 1 \right)
\]

where \( P, Q, \) and \( E \) are the rate of precipitation (from snow-melt and rain), discharge, and actual evapotranspiration (from snow-free area) in units of depth per time. The concept behind this method is that the catchment sensitivity of discharge to changes in storage i.e., \( g(Q) \), can be estimated from the time series of the discharge alone through fitting empirical functions to the data such as the quadratic equation, and is given by Eq. 2 (Kirchner, 2009).

\[
g(Q) = e^{c_1 + c_2(\ln(Q)) + c_3(\ln(Q))^2}
\]

where \( c_1, c_2, \) and \( c_3 \) are the catchment specific outlet parameters (hereafter called Kirchner parameters) obtained during the model calibration. There is no routing into the model, but Kirchner parameters response delays outflow from precipitation and snow/glacier melt.

2.1.1 Snow and glacier balance

The Gamma snow routine uses an energy balance approach (Eq. 3) for snow ablation, and the snow depletion curve represented by gamma distribution for snow distribution. The energy balance approach in the Gamma snow routine follows similar approach as used in DeWalle and Rango (2008). The net energy flux (\( \Delta E \)) at the surface available for snow ablation is expressed as:

\[
\Delta E = S(1 - \alpha) + L_{in} + L_{out} + H_{SE} + H_L + E_G
\]

where \( S \) is the net shortwave radiation, \( L_{in} \) and \( L_{out} \) are the incoming and outgoing longwave radiation, \( H_{SE} \) and \( H_L \) represent sensible and latent heat fluxes and \( E_G \) is the net ground heat energy flux and is calculated using bulk-transfer approach. Two parameters defining wind profile, intercept (wind constant) and slope (wind scale) are determined through either calibration or prescribed (Table 1). For a given time step (t), the snow albedo (\( \alpha \)) at each grid cell depends on the minimum (\( \alpha_{min} \)) and maximum albedo (\( \alpha_{max} \)) values as well as the albedo decay rate, temperature, and snowmelt e.g. Hegdahl et al. (2016):

\[
\alpha_t = \begin{cases} 
\alpha_{min} + (\alpha_{t-1} - \alpha_{min}).\left(\frac{1}{2^{|\frac{T_{\alpha}}{T_{\alpha}}|}}\right) & \text{if } T_{\alpha} > 0^\circ C \\
\alpha_{t-1} + (\alpha_{max} - \alpha_{min}).\left(\frac{1}{2^{|\frac{T_{\alpha}}{T_{\alpha}}|}}\right) & \text{if } T_{\alpha} \leq 0^\circ C 
\end{cases}
\]

In Eq. 4, FDR and SDR denote fast and slow snow cover decay rates, respectively. In this study, \( \alpha_{min} \) and \( \alpha_{max} \) are estimated during the model calibration.
Table 1. Model calibration parameters with upper and lower limits. Prescribed parameters are marked with *. Parameters only used in radiation model are marked with **.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description and unit</th>
<th>Lower limit</th>
<th>Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>kirchner parameter 1 (see Eq. 2) [-]</td>
<td>-8.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$c_2$</td>
<td>kirchner parameter 2 (see Eq. 2) [-]</td>
<td>-1.0</td>
<td>1.2</td>
</tr>
<tr>
<td>$c_3$</td>
<td>kirchner parameter 3 (see Eq. 2) [-]</td>
<td>-0.15</td>
<td>-0.05</td>
</tr>
<tr>
<td>p_corr_factor</td>
<td>scaling factor for precipitation [-]</td>
<td>0.2</td>
<td>2.0</td>
</tr>
<tr>
<td>tx</td>
<td>temperature threshold rain/snow [°C]</td>
<td>-3.0</td>
<td>2.0</td>
</tr>
<tr>
<td>wind_scale</td>
<td>determining wind profile [-]</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>FDR</td>
<td>fast albedo decay rate during cold condition [days]</td>
<td>5.0</td>
<td>15</td>
</tr>
<tr>
<td>SDR</td>
<td>slow albedo decay rate during melt [days]</td>
<td>20.0</td>
<td>40.0</td>
</tr>
<tr>
<td>surface magnitude*</td>
<td>snow heat constant [mm SWE]</td>
<td>30.0</td>
<td>30.0</td>
</tr>
<tr>
<td>max_water*</td>
<td>fractional max water content of snow [-]</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>max albedo*</td>
<td>maximum snow albedo used in snow routine [-]</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>min albedo*</td>
<td>minimum snow albedo used in snow routine [-]</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>ae_scale_factor*</td>
<td>scaling factor for actual evapotranspiration [-]</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>rad_albedo **</td>
<td>average albedo of the surrounding ground surface below the inclined surface [-]</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>rad_turbidity **</td>
<td>an empirical turbidity coefficient [-]</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Within the gamma snow routine, precipitation falling in each grid cell is classified as solid or liquid depending on a threshold temperature (tx), and the actual grid cell temperature. Snow distribution to each cell is estimated by using a three parameter gamma probability distribution (Kolberg et al., 2006). Snowmelt depth is calculated by multiplying $\Delta E$ (available energy) with the latent heat of fusion for water.

Temperature index model (Hock, 2003) is used in the model for glacier melt calculation. The glacier reservoir is assumed to be infinite and glacier area is assumed to be constant throughout the study periods. In a grid cell, snow is assumed to melt first from glacier cover/free areas, and glacier melt only happens from glacial areas that are snow-free. Obtained runoff depth per grid cell is summed with snowmelt depth from the gamma snow routine and converted to volume by multiplying the area of each grid cell before Kirchner routine.

2.1.2 Study catchment and data

The study has been conducted in the Marsyangdi-2 River catchment (Fig. 1), Nepal. It is located between 28°10'21″N to 28°54’11″N and 83°47’24″E to 84°48’04″E. The catchment has a total area of 3001 sq.km of which 80% lies above the elevation of 3,000 m asl. The catchment elevation ranges from 641 m asl to 7924 m asl with a mean elevation of 4377 m asl, and mean slope of 29.34°. Marsyangdi-2 river is a tributary of the Narayani river catchment (Fig. 1). It originates from the southern edge of the Tibetan Plateau, flows through Nepal to India, and drains into the Ganges River. About 24% of
area is covered by glaciers (Bajracharya et al., 2014). The hydrological regime is heavily influenced by seasonal weather variations. The seasons are defined as monsoon (June to September), post-monsoon (October to November), winter (December to February), and pre-monsoon (March to May) (Bhattarai et al., 2019). About 80% of the total annual precipitation occurs during the monsoon period (Nayava, 1974; Kripalani and Sontakke, 1996).

The Narayani River catchment lies also in the primary hydropower development region of Nepal, providing 44% of the total national electric supply (Adhikari, 2006). Some of the operating hydro-power development projects in this catchment include: Marsyangdi (69 MW) and middle marsyangdi (70 MW). The upper marsyangdi (600 MW), lower Manang marsyangdi (100 MW) and Nyadi (30 MW) are some of the hydropower project currently under different stages of development. Hence, the study area is an important catchment in Nepal from hydropower perspective.

2.1.3 Meteorological forcing and observed river gauge data

Water and Global Change (WATCH) Forcing Data ERA-Interim (WFDEI) is selected as forcing data for the model. Selection of the forcing data sets for the hydrological model was based on the prior study by Bhattarai et al. (2020). WFDEI is a global meteorological forcing data set at 0.5° x 0.5° horizontal resolution obtained by bias-correcting ERA-I data (Weedon et al., 2014; Raimonet et al., 2017). It covers the period 1979 – 2016 and contains eight meteorological variables at a 3-h time step (Essou et al., 2016). WFDEI then corrects ERA-Interim for precipitation biases using data from the Climatic Research Unit (CRU) or the Global Precipitation Climatology Center (GPCC). In this study, GPCC products were preferred to CRU because
of their higher resolution and data quality (Weedon et al., 2014; Raimonet et al., 2017). The WFDEI data set is freely available online from ftp.iiasa.ac.at.

Daily observed river discharge data (observation periods: 2000 – 2010) at Marsyangdi_2 river (28.2036 N, 84.403E) are obtained from the Department of Hydrology and Meteorology (DHM), Government of Nepal (GoN). The catchment is highly influenced by seasons as more than 70% of total annual discharge flows from June to September. Highest daily mean discharge is observed during August (386.7 m$^3$/s) and low mean discharge is observed in February (30.5 m$^3$/s), and daily mean observed discharge is 133.85 m$^3$/s.

2.1.4 MODIS snow product, topographical, and land cover data sets

In this study, the standard eight-day MODIS Terra and Aqua snow cover products, MOD10A1 and MYD10A1 are chosen for the analysis. They contain snow cover, snow albedo, fractional snow cover, and quality assessment (QA) data in a compressed HDF-EOS format along with corresponding metadata. Only snow cover information is used in the study. Both Aqua and Terra products are with 463.3 m spatial resolution in sinusoidal projection. Frequent cloud cover is one of the major challenges when using MODIS and other optical remote sensing data in the Himalayan region (Gurung et al., 2011). MOD10A1 was successfully used in earlier studies (Stroeve et al., 2006; Li et al., 2016; Williamson et al., 2018). However, Rittger et al. (2013) and Toure et al. (2018) showed that the MODIS data can be systematically biased compared to Landsat ETM+ snow cover data in the mountainous regions. The average root mean square error (RMSE) between Landsat ETM+ and MOD10A1 snow cover fraction over the Nepal Himalaya was 0.23 (Toure et al., 2018). Thus, a composite data set was formed, following the Muhammad and Thapa (2019) approach, using data retrieved from the Aqua and Terra products in order to minimize the effect of clouds and other sources.

A land cover map (0.3° x 0.3°, resolution) providing forest cover, lake cover, glacier cover, and reservoir cover of the study area is extracted from the GlobCover Project (Arino et al., 2010). Digital Elevation Model (DEM) of 90 meters spatial resolution from the Shuttle Radar Topography Mission (SRTM) (Van Zyl, 2001) is used in the study.

2.2 Catchment discretization technique

The study catchment was delineated using an automatic catchment delineation method described by Akram et al. (2012). The delineated catchment was further discretized into modelling elements using three methods.

2.2.1 Triangulated Irregular Network (TIN)

TIN is a vector-based data structure comprising a triangular network of vertices with associated coordinates (x, y, z) connected by edges. TIN utilizes the original sample points to constitute many non-overlapping triangles that cover the entire region. Spacing and shape of the triangles are determined by the terrain and by the desired degree of fit. The mathematical basis for a TIN construction is the Delaunay triangulation methods described in Field (1988) and Hjelle and Dæhlen (2006), and is based on a principle of maximizing the minimum angle of all triangles produced by connector lines to nearest neighbour points.
The TIN interpolation method is an altitude interpolation algorithm, this makes it possible to distribute hydro-meteorological forcing variables more correctly for hydrological simulations (Vivoni et al., 2004).

Rasputin (https://github.com/expertanalytics/rasputin) is used for generating TIN terrain models of the study area. It can convert a point set of coordinates to a TIN using Delaunay triangulation methods. Specifically, it has been developed to convert Digital Elevation Models (DEMs) into simplified triangulated surface meshes. Rasputin generates catchment TIN using the provided DEM and also calculates physical parameters of each TIN facet such as slope, area, and aspect, from the TIN geometry. The Rasputin software requires wkt-file describing the catchment polygon, which is easily obtained from shapefiles generated by Qgis. Rasputin is freely available under GNU GPLv3 license.

TINs with different number of cells are generated and tested before selecting a particular TIN data file. Implementing the spatially distributed parameters for the TIN with different number of cells yielded different simulation performance when compared to the observations over the calibration (2000–2004) and validation (2005-2009) periods. As expected, model efficiency in terms of NSE increased with increasing number of TINs (Fig. 2). However, after reaching a certain number of cells no further improvement in NSE was found, indicating the limit for Shyft to utilize terrain data. The number of cells corresponding to this limit was considered as an optimum i.e. 671 cells (with median area of 4km²) (see Fig. 2).

![Figure 2](image_url)  
*Figure 2. Scatter plot between NSE and number of TIN cells (Ncells).*

### 2.2.2 Regular space grid

Vector creation tools under Qgis geoalgorithms were used for creating regular space grid cells. Due to ease of comparison with TIN, grid mesh with an area of 4 km² was selected for this study. Unlike to TIN, each grid cell is connected to either four or eight adjacent neighbours, the single resolution inherent in raster grids. Physical parameters (slope, aspect, elevation,
latitude, and longitude) of each cell at the centroid (centre point) are calculated by using vector geometry tools under Qgis geoaLgorithms and further used for distributing the hydro-meteorological forcing.

2.2.3 Hypsography (Elevation zone)

In this study, the catchment area was divided into ten elevation zones. Mean elevation, aspect, and slope of each zone were calculated and further used for the distributing the hydro-meteorological forcing. Fig. 3 shows discretization of the study catchment using TIN, regular space grid and elevation bands.

![Figure 3. Catchment discretization by (A) hypsographic (B) Regular space grid, and (C) TIN methods.](image)

2.2.4 Implementation of new radiation algorithm in the Shyft

Daily incoming shortwave radiation (from forcing data) falling over the inclined surface with specified slope and aspect was calculated using the methods described by Allen et al. (2006). The procedure assumes each grid/TIN surface has uniform slope. Thus, effects of protruding surrounding terrain were not considered. This simplification facilitates implementation of the procedure within Shyft in energy balance calculation for snow mass balance and in Priestley-Taylor based potential evapotranspiration calculations.

2.3 Model Construction, Calibration, and Validation

Combination of different catchment discretization methods and radiation algorithms resulted in different models (Table 2). Each model was calibrated and validated separately. A cross-validation technique was used, where the optimal model identified in the first periods, i.e. years 2000-2004 was validated during the second period, i.e. 2005-2009 and vice-versa. For convention, different names were given to the model-discretization combinations (Table 2).
Table 2. Different models with descriptions.

<table>
<thead>
<tr>
<th>S.N.</th>
<th>Models</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PTGSK-HYP</td>
<td>Priestley–Taylor (PT) method for estimating potential evaporation; Gamma snow routine (GS) for snow melt, sub-grid snow distribution and mass balance calculations; a simple storage-discharge function (K) for catchment response calculation; and catchment discretization by hypsography (HYP) method.</td>
</tr>
<tr>
<td>2</td>
<td>RPTGSK-HYP</td>
<td>Imputed incoming shortwave radiation (R); Priestley–Taylor (PT) method for estimating potential evaporation; Gamma snow (GS) for snow melt, sub-grid snow distribution and mass balance calculations; a simple storage-discharge function (K) for catchment response calculation; and catchment discretization by hypsography (HYP) method.</td>
</tr>
<tr>
<td>3</td>
<td>PTGSK-SqGrid</td>
<td>Priestley–Taylor (PT) method for estimating potential evaporation; Gamma snow routine (GS) for snow melt, sub-grid snow distribution and mass balance calculations; and a simple storage-discharge function (K) for catchment response calculation, and catchment discretization by regular square grid (SqGrid) method.</td>
</tr>
<tr>
<td>4</td>
<td>RPTGSK-SqGrid</td>
<td>Imputed incoming shortwave radiation (R); Priestley–Taylor (PT) method for estimating potential evaporation; Gamma snow (GS) for snow melt, sub-grid snow distribution and mass balance calculations; and a simple storage-discharge function (K) for catchment response calculation, and catchment discretization by regular square grid (SqGrid) method.</td>
</tr>
<tr>
<td>5</td>
<td>PTGSK-TIN</td>
<td>Priestley–Taylor (PT) method for estimating potential evaporation; Gamma snow routine (GS) for snow melt, sub-grid snow distribution and mass balance calculations; a simple storage-discharge function (K) for catchment response calculation, and catchment discretization by Triangulated Irregular Network (TIN) method.</td>
</tr>
<tr>
<td>6</td>
<td>RPTGSK-TIN</td>
<td>Imputed incoming shortwave radiation (R); Priestley–Taylor (PT) method for estimating potential evaporation; Gamma snow (GS) for snow melt, sub-grid snow distribution and mass balance calculations; a simple storage-discharge function (K) for catchment response calculation, and catchment discretization by Triangulated Irregular Network (TIN) method.</td>
</tr>
</tbody>
</table>

2.4 Evaluation metrics

2.4.1 Nash-Sutcliffe Efficiency (NSE)

The Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970) was used (Eq. 5) to quantitatively evaluate the degree of agreement between observed and simulated discharge. NSE is commonly used to assess the predictive power of hydrological models and its value can range from $-\infty$ to 1. An efficiency of 1 corresponds to a perfect match of simulated to observed discharge values. However, the NSE calculated using raw values tends to overestimate model performance during peak streamflow and to underestimate during low-streamflow conditions (Krause et al., 2005; Teweldebrhan et al., 2018). To partly overcome this...
problem, NSE is often calculated with log-transformed observed and modelled values. In this study, both NSE and NSE with log-transformed streamflow values (LnNSE) were thus employed as streamflow efficiency metrics.

\[
NSE = 1 - \frac{\sum_{i=1}^{n} (Q_{\text{sim},i} - Q_{\text{obs},i})^2}{\sum_{i=1}^{n} (Q_{\text{obs},i} - \bar{Q}_{\text{obs},i})^2}
\]

(5)

where \(Q_{\text{sim}}\), \(Q_{\text{obs}}\), and \(\bar{Q}_{\text{obs}}\) are simulated discharge, observed discharges and arithmetic mean for observed discharges.

5 2.4.2 Critical Success Index (CSI)

We determine the agreement between observed (MODIS) and simulated snow via critical success index (CSI) (Schaefer, 1990). The CSI measures the patterns match between simulations and observations. CSI is the number of correct events simulated, divided by the number of cases observed. The value of the CSI score ranges between 0 and 1, with 1 being the ideal score. The formula for CSI is defined as Eq. 6, and two-dimensional contingency table fit for evaluating observed and simulated binary events is given in Table 3 (Teweldebrhan et al., 2018).

Shyft generates fractional snow cover area (fSCA) as one of its main output variables. Thus in order to convert the fractional snow cover area into binary values, a threshold fSCA value of 0.4 was used to distinguish between snow and snow free cells.

Table 3. Setup of the 2 x 2 contingency table for binary snow cover data comparison. O and S, respectively, represent observed and simulated binary snow cover and the subscripts refer to a snow -free (0) and snow- covered (1) grid cell.

<table>
<thead>
<tr>
<th>S.N.</th>
<th>S</th>
<th>S₀</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁</td>
<td>(n_{11})</td>
<td>(n_{x1})</td>
<td>(n_{x₁})</td>
</tr>
<tr>
<td>O₀</td>
<td>(n_{10})</td>
<td>(n_{00})</td>
<td>(n_{x0})</td>
</tr>
<tr>
<td>Sum</td>
<td>(n_{1x})</td>
<td>(n_{0x})</td>
<td>(n_{xx})</td>
</tr>
</tbody>
</table>

\[
CSI = \frac{n_{11}}{(n_{xx} - n_{00})}
\]

(6)

3 Results

3.1 Intercomparison of the catchment discretization methods

The different catchment discretization methods resulted in different number and distribution of computational cells with in the study catchment. The fixed median area of TIN cells (4km²) has yielded to lower number of computational cells (671) as compared to the regular square grid (839). The distribution of cells across the elevation range of the study catchment shows that most of the cells fall within the elevation range of 4000 to 6000 m asl. Most of the reductions in the number of computational...
cells when using TIN were also obtained in this elevation range (Fig. 4). TIN cell areas vary from 0.16 km$^2$ to 15.3 km$^2$, while in the hypsographic discretization method, zonal areas range from 6.8 km$^2$ to 726.8 km$^2$ (Table 4).

Figure 4. TIN and grid cell distribution across the elevation ranges in Marsyangadhi_2 catchment.

Table 4. Hypsographical discretization and zonal characteristics. In table CID represents catchment number.

<table>
<thead>
<tr>
<th>CID</th>
<th>Mean aspect (Degree from North)</th>
<th>Mean slope (Degree)</th>
<th>Mean elevation (m asl)</th>
<th>Glacier area (sq. km)</th>
<th>Zonal area (sq. km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>171.8</td>
<td>24.3</td>
<td>1084.0</td>
<td>0.0</td>
<td>96.7</td>
</tr>
<tr>
<td>2</td>
<td>174.3</td>
<td>31.3</td>
<td>1742.8</td>
<td>0.0</td>
<td>144.8</td>
</tr>
<tr>
<td>3</td>
<td>181.5</td>
<td>31.2</td>
<td>2483.6</td>
<td>0.0</td>
<td>193.3</td>
</tr>
<tr>
<td>4</td>
<td>186.7</td>
<td>31.8</td>
<td>3232.3</td>
<td>0.0</td>
<td>274.1</td>
</tr>
<tr>
<td>5</td>
<td>184.3</td>
<td>30.0</td>
<td>3948.2</td>
<td>0.5</td>
<td>487.6</td>
</tr>
<tr>
<td>6</td>
<td>179.2</td>
<td>28.2</td>
<td>4663.5</td>
<td>51.3</td>
<td>693.0</td>
</tr>
<tr>
<td>7</td>
<td>178.8</td>
<td>26.8</td>
<td>5358.1</td>
<td>290.7</td>
<td>726.8</td>
</tr>
<tr>
<td>8</td>
<td>181.9</td>
<td>32.4</td>
<td>6015.3</td>
<td>276.5</td>
<td>317.1</td>
</tr>
<tr>
<td>9</td>
<td>175.1</td>
<td>39.4</td>
<td>6728.4</td>
<td>40.1</td>
<td>59.7</td>
</tr>
<tr>
<td>10</td>
<td>208.4</td>
<td>37.6</td>
<td>7397.3</td>
<td>0.0</td>
<td>6.9</td>
</tr>
</tbody>
</table>

Remarkable differences are observed in grid/TIN area distributions with respect to aspect. In regular space square grid, higher percentage of area is facing towards south-west direction while in TIN, higher percentage of area is facing towards North-East.
direction (Fig. 5). Differences in area distribution with respect to aspect in different discretization methods ultimately result in differences in snow/glacier cover area in the catchment.

![Figure 5. Cell area distribution with respect to aspect.](image)

### 3.2 Evaluation of catchment discretization and imputed radiation on streamflow efficiency

Figure 6 shows the hydrograph of the measured and simulated discharges at Marsyangdi_2 station for calibration and validation periods and different model configurations. The different catchment discretization methods resulted in different model efficiency. Both the calibration and validation results show a good agreement between observed and modelled daily discharge values in terms of Nash-Sutcliffe efficiency (NSE > 0.75). However, annual peak values are underestimated by all models.

Low flow simulations are relatively well captured by PTGSK-TIN and RPTGSK-TIN as compared to the other models. Modelled discharge during pre-monsoon season from all models was higher than observation, which is caused by the positive biases in the WFDEI precipitation during this season (Bhattarai et al., 2020). From this figure it can be seen that there is no significant difference between the modelled discharges of the two radiation methods.
Effect of the catchment discretization methods and radiation algorithms on model performance was further evaluated using NSE and LnNSE (Fig. 7). The model efficiency metrics were computed for each hydrologic year constituting the validation period. As such, each factor had 10 samples with a total size of 60 samples (two radiation algorithms and three discretization types). For model efficiency under high-flow condition as estimated using the raw-data (NSE) (Fig. 7(a)), a more significant difference in model efficiency was observed between the catchment discretization types as compared to the radiation algorithms. The inter-comparison between discretized models shows that triangulated irregular network (TIN) was superior to both the square grid cells (SqGrid) and the elevation bands (HYP). While HYP has yielded better result as compared to SqGrid. A similar order of performance between the three catchment discretization types was observed under the two radiation algorithms. Although a slight improvement was observed in sqGrid as a result of using RPTGSK against PTGSK, a slight deterioration in model performance was observed when using HYP and TIN with RPTGSK. A superior result was also obtained.
in model performance during low-flow conditions (LnNSE) when using TIN as compared to other catchment discretization types (Fig. 7(b)). Unlike to the result obtained using NSE, a relatively better simulation result was obtained under SqGrid than HYP when assessed using LnNSE. No significant difference in LnNSE was also observed between the results obtained from PTGSK and RPTGSK.

Figure 7. Model efficiency during the high-flow (a) and low-flow (b) conditions under three catchment discretization types, i.e. elevation bands (HYP), square grids (SqGrid) and triangulated irregular networks (TIN) as well as under two radiation algorithms (with (RPTGSK) and without (PTGSK) corrections based on certain terrain parameters.

3.3 Effects of imputed radiation on snow simulations

Results presented in the previous sections show that RPTGSK-TIN and PTGSK-TIN performed better than the remaining models. Hence, these models were selected for snow cover analysis and validation of their snow cover simulations through
comparison against the MODIS snow product. Figure 8(a) demonstrates monthly average percentage of snow cover area from MODIS, PTGSK-TIN, and RPTGSK-TIN. MODIS observations and model simulations show higher percentage of snow cover area during winter periods (January-April). The observed time series with strong seasonality in snow cover are adequately represented with both models. However, simulated values tend to underestimate the observed values during the low snow cover period i.e. June-October. To find the agreement between the model simulation and MODIS, monthly correlation coefficients were calculated. Highest monthly correlation coefficients based on five years (2005-2009) simulation were observed between MODIS and RPTGSK-TIN snow during February (0.85), April (0.76), May (0.63). Similar to RPTGSK-TIN, highest correlation coefficients between PTGSK-TIN and MODIS were observed during February (0.75), April (0.64), and May (0.60). Negative correlation coefficients between MODIS and both models were observed during December and January.

Figure 8. (a) catchment scale 8-days maximum snow cover area (%) from MODIS, PTGSK-TIN and RPTGSK-TIN. (b) Snow cover area percentage differences between PTGSK-TIN and RPTGSK-TIN.

Snow cover areas from MODIS during July, August, and September are higher than from the models. Higher percentage of snow cover from MODIS might be connected to the fact that MODIS does not distinguish between snow and glacier, and considers both as a snow cover. Unlike to MODIS observations, glacier area is not simulated in Gamma snow, hence lower percentage of snow cover area during the snow ablation periods was found. The other possible source of uncertainty in the snow observations from MODIS during this period is related to cloud cover, which is obviously very high during monsoon period.

Figure 8(b) shows five year daily average snow cover area percentage differences between the two models simulations. With the implementation of translated radiation into the models, significant differences were observed in snow simulation. Highest differences (up to 12%) were observed during the winter season.

Figure 9 presents monthly CSI values of snow cover calculated from MODIS and model results for the simulation periods 2005 to 2009. The highest CSI from both models was obtained during pre-monsoon season (March- June) and the CSI score from both models are nearly the same. Median CSI score for both models during the pre-monsoon was around 0.75, indicating that 75% of the simulated snow from PTGSK-TIN and RPTGSK-TIN matches the remote observations. Similar to the correla-
tion coefficient values, the lowest CSI scores were observed during the winter and monsoon periods. There is a slight increase in CSI score for RPTGSK-TIN simulation relative to PTGSK-TIN during June, July, and August.

![Box plot of monthly Critical Success Index](image)

**Figure 9.** Monthly Critical Success Index calculated for models (PTGSK-TIN, and RPTGSK-TIN).

The scatterplots between 8-days snow cover area fraction from models and MODIS in different elevations are shown in Fig. 10. As can be seen from this figure, the topography seems to have a dominating effect on snow cover distribution. In the elevation ranges up to 3000 m asl, very few snow events were detected from both models and almost no-snow events were detected from the MODIS. The ranges from 3000 to 6000 m asl experience greater snow coverage whose spatial distribution is relatively better captured by both data sets. A slightly higher correlation coefficient between MODIS and RPTGSK-TIN simulation was observed as compared to PTGSK-TIN (Fig. 10). During the snow ablation period, the high mountain ranges (>6000 m asl.) have permanent glaciers which are not simulated by Gamma snow but detected in MODIS giving lower correlation coefficients for both models.
Discussion

The basic assumption in this study was that different model discretization methods were the reason for the differences in model performance. The results obtained through a variety of statistical models analysis undeniably follow the catchment discretization methods. We found all six models performed well (NSE > 0.75) for both calibration and validation. In general, all discretization methods provide an excellent representation of the general flow pattern and the overall water balance, while maintaining the significant interannual variability satisfying.

The predicted hydrographs for the calibration/validation, shown in Fig. 6, confirm that the overall hydrograph pattern is predicted quite well by all models. Similar to the previous studies by DeVantier and Feldman (1993); Singh and Fiorentino (1996) and Vivoni (2003), better performance was observed for the models with TINs. The poor performance of the grid-based models are partly attributed to complex topography of the region and the associated difference in distribution of land area (Fig. 11). According to the Nash–Sutcliffe efficiency (NSE) and the log-transformed Nash–Sutcliffe efficiency (InNSE) of the outlet discharge, the rank of discretized model efficiency, in the descending order is TIN, HYP, and SqGrid models both in the calibration and validation periods. Theoretically, the ability to account for spatial variability of meteorological forcing and physical features within a catchment should lead to better simulations (Brirhet and Benaabidate, 2016). However, similar to the prior studies by Reed et al. (2004) and Smith et al. (2004), we found that grid based distributed modeling approaches do not always provide improved discharge simulation compared to hypsography discretized conceptual models. As presented by Birhanu et al. (2019) and Xu et al. (2018), land cover types are highly dependent on elevation change, hence the catchment discretized by the elevation bands represent catchment better than the grid. As a result, in our study, better land representation in hypsography based models results in higher model efficiency compared to grid based model.

The Delaunay triangulation is a widely appreciated and investigated mathematical model to represent the topography and is highly efficient for building triangular irregular networks (TINs). Vivoni (2003) mention that in the regions of high terrain

Figure 10. Scatterplot between MODIS derived snow and models (PTGSK-TIN and RPTGSK-TIN) for each altitudinal zone. In each plot ‘r’ represent correlation coefficient.
variability the catchment discretized with TINs can be modeled more precisely due to higher flexibility of the mesh. The areas with little variation in topography are represented by larger triangles and the areas with more variation are represented by smaller ones leading to better representation of the ground shape (De Wulf et al., 2012). Figure 11 shows the distribution of land cover area in grid and TIN. Remarkable differences are seen for the glacier and forest cover area. Therefore higher efficiencies from the models with TIN are also attributed to the better representation of land cover area. One of the most useful characteristics of a TIN for hydrologic system is the ability to define stream in terms of triangle boundary segments. This allows a more continuous description of stream paths and networks in conjunction with the topography. By comparison, grid data tend to produce zig-zag meandering paths for stream on upslope portions of a watershed (DeVantier and Feldman, 1993). As indicated by (Jenness, 2004), Raster data sets such as DEMs and surface-area grids are inherently less accurate and precise than vector data sets such as TINs and polygons. Grid values do not reflect the actual measurements. The grid model is not adaptive: whereas TINs will naturally represent area with detailed relief information with a denser triangle pattern than area with a smoother relief, grids will be far less flexible to cope with variable levels of details (De Wulf et al., 2012).

The overall performance of the CSI estimated by the PTGSK-TIN is somewhat reduced compared to RPTGSK-TIN 5 years (2005-2009) simulation runs for the Marsyangdi_2 catchment as expected (Figs. 8 and 9). However, biases in catchment average snow cover area from PTGSK-TIN and RPTGSK-TIN are less then 1%. The central estimate (median) from RPTGSK-TIN is slightly better (Fig. 9; July-August-September) than PTGSK-TIN for studied catchment, giving better representation of observed simulated snow cover area.

Significant differences between snow simulation were observed by the implementation of translated radiation algorithm in to the Shyft. With the translated radiation, notable differences were also observed in evapotranspiration (see: Table 5). As seen in Fig. 12(A), unlike to Fig. 12(B), radiation is not distributed uniformly, and more radiation was received by south-east
and south-west aspect. Lower percentage of the radiation is distributed to the Northern aspect. It should be noted that small percentage of higher radiations are also seen on northern aspect, but similar to the study by Burtscher (2014), higher radiation in the northern aspect is because of the high elevation area there.

Although the overall efficiencies in terms of CSI and NSE are not significantly different (Figs. 6 and 8), effect of implemented imputed radiation can be clearly seen in water balance components (Table 5). As seen in Fig. 12(A), lower percentage of radiation is distributed towards East-South aspect resulting in lowering the total glacier melt. Furthermore, the lower percentage of evaporation is also observed for the model with imputed radiation and could be caused by reduced radiation.

Table 5. Annual (2005) water balance components for Marshyangdi_2 river catchment.

<table>
<thead>
<tr>
<th></th>
<th>PTGSK-TIN</th>
<th>RPTGSK-TIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evapotranspiration (%)</td>
<td>19.7</td>
<td>17.8</td>
</tr>
<tr>
<td>Glacier melt (mm)</td>
<td>535.9</td>
<td>513.1</td>
</tr>
<tr>
<td>Runoff (mm)</td>
<td>1355.9</td>
<td>1354.8</td>
</tr>
</tbody>
</table>

5 Conclusions

This study was focused on evaluation of the effect of three catchment discretization approaches; i.e. hypsography (HYP), square grid (sqGrid), and triangulated irregular network (TIN) on hydrological model simulation results, i.e. streamflow and snow cover. Further use of two different algorithms for imputed radiation and its impact on model simulation results was also assessed. The different configurations were successfully applied to discharge simulation in Marsyangdi_2 river catchment.
Generally a good agreement between the observed and modeled discharge was obtained from the different configurations. The model simulation results based on the different discretization approaches, i.e. TIN, HYP, and SqGrid were evaluated using graphical and certain efficiency metrics and they have shown a decreasing order of performance.

Selection of the catchment discretization methods depends upon the availability of computational resources and acceptable model efficiency. Regular sqGrid models are less flexible, but offer higher speed, lower memory requirements and easier implementation algorithms as most important assets, making them to be preferred when the studied area is more flat. For area with high relief, typical for hydrological modelling, TINs are a priori the preferred option.

Direct use of radiation from WFDEI is a good option for hydrological modeling in a catchment. Snow simulation results did not show any changes when using imputed radiation, the snow cover determined by the direct use of radiation approach in the high mountain region do not represent the real snow cover area. Additional work is still needed to test and validate the suitability of snow simulation determination at different spatial and temporal scales using imputed radiation. Snow simulation with the models with TIN and imputed radiation should also be validated by other models with a much wider range of terrain types.

Finally, our study area represent very complex regime in central Himalayas, Nepal and the results are encouraging. The TIN based models are particularly suited to the Himalayan catchments with steep descent/ascent because of the uniform slope along each triangle facet. If the study is constrained by time and computational memory, hypsography based discretization methods are better options than the grid based method.

Author contributions. Bikas C. Bhattarai, Olga Silantyeva, Aynom Teweldebrhan, John F. Burkhart designed and performed the analysis. BCB wrote the manuscript. John F. Burkhart supervised the study. Sigbjørn Helset, Ola Skavhaug and Olga Silantyeva implemented the relevant algorithms in Shyft. All authors revised the manuscript.

Acknowledgements. We are thankful to entire MODIS science team and NASA providing a massive data. This work was conducted within the Strategic Research Initiative “Land Atmosphere Interaction in Cold Environments” (LATICE) of the University of Olso and partially supported through the Norwegian Research Council’s INDNOR program under the Hydrologic sensitivity to Cryosphere-Aerosol interaction in Mountain Processes (HyCAMP) project (NFR no. 222195).
References


Li, H., Zhang, Y., and Zhou, X.: Predicting surface runoff from catchment to large region, Advances in Meteorology, 2015, 2015.


Table 6. Calibrated parameters for different model.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unit</th>
<th>PTGSK-HYP</th>
<th>RPTGSK-HYP</th>
<th>PTGSK-SqGrid</th>
<th>RPTGSK-SqGrid</th>
<th>PTGSK-TIN</th>
<th>RPTGSK-TIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>[-]</td>
<td>-6.25</td>
<td>-6.28</td>
<td>-6.17</td>
<td>-6.17</td>
<td>-6.26</td>
<td>-6.22</td>
</tr>
<tr>
<td>c2</td>
<td>[-]</td>
<td>-0.04</td>
<td>-0.10</td>
<td>-0.05</td>
<td>-0.02</td>
<td>0.15</td>
<td>0.20</td>
</tr>
<tr>
<td>c3</td>
<td>[-]</td>
<td>-0.09</td>
<td>-0.10</td>
<td>-0.11</td>
<td>-0.08</td>
<td>-0.11</td>
<td>-0.09</td>
</tr>
<tr>
<td>tx</td>
<td>[°C]</td>
<td>-1.86</td>
<td>-1.67</td>
<td>-1.92</td>
<td>-2.57</td>
<td>-2.19</td>
<td>-1.58</td>
</tr>
<tr>
<td>wind scale</td>
<td>[m/s]</td>
<td>5.26</td>
<td>4.28</td>
<td>6.00</td>
<td>6.00</td>
<td>4.00</td>
<td>3.60</td>
</tr>
<tr>
<td>FDR</td>
<td>[days]</td>
<td>8.96</td>
<td>9.05</td>
<td>7.55</td>
<td>7.58</td>
<td>8.99</td>
<td>9.90</td>
</tr>
<tr>
<td>SDR</td>
<td>[days]</td>
<td>25.83</td>
<td>33.04</td>
<td>36.12</td>
<td>30.25</td>
<td>34.33</td>
<td>31.75</td>
</tr>
<tr>
<td>p_corr_factor</td>
<td>[-]</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
</tr>
</tbody>
</table>
Part III

Appendices
A List of observation stations in the Narayani River catchment

A.1 Temperature observation stations

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Station Name</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Elevation (m a.s.l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>601</td>
<td>Jomsom</td>
<td>28.47</td>
<td>83.43</td>
<td>2744</td>
</tr>
<tr>
<td>607</td>
<td>Lete</td>
<td>28.38</td>
<td>83.36</td>
<td>2384</td>
</tr>
<tr>
<td>609</td>
<td>Benibajar</td>
<td>28.21</td>
<td>83.34</td>
<td>835</td>
</tr>
<tr>
<td>612</td>
<td>Mustang</td>
<td>29.11</td>
<td>83.58</td>
<td>3705</td>
</tr>
<tr>
<td>616</td>
<td>Gurja khani</td>
<td>28.36</td>
<td>83.13</td>
<td>2530</td>
</tr>
<tr>
<td>633</td>
<td>Chhoser</td>
<td>29.11</td>
<td>83.59</td>
<td>3870</td>
</tr>
<tr>
<td>706</td>
<td>Dumkauli</td>
<td>27.41</td>
<td>84.13</td>
<td>154</td>
</tr>
<tr>
<td>715</td>
<td>Khanchikot</td>
<td>27.56</td>
<td>83.09</td>
<td>1760</td>
</tr>
<tr>
<td>725</td>
<td>Tamghas</td>
<td>28.04</td>
<td>83.15</td>
<td>1530</td>
</tr>
<tr>
<td>802</td>
<td>Hudibajar</td>
<td>28.17</td>
<td>84.22</td>
<td>823</td>
</tr>
<tr>
<td>804</td>
<td>Pohara Airport</td>
<td>28.13</td>
<td>84</td>
<td>827</td>
</tr>
<tr>
<td>805</td>
<td>Syangja</td>
<td>28.06</td>
<td>83.53</td>
<td>868</td>
</tr>
<tr>
<td>809</td>
<td>Gorkha</td>
<td>28</td>
<td>84.37</td>
<td>1097</td>
</tr>
<tr>
<td>810</td>
<td>Chapkota</td>
<td>27.53</td>
<td>83.49</td>
<td>460</td>
</tr>
<tr>
<td>811</td>
<td>Malepatan(Phokhara)</td>
<td>28.07</td>
<td>84.07</td>
<td>856</td>
</tr>
<tr>
<td>814</td>
<td>Lumle</td>
<td>28.18</td>
<td>83.48</td>
<td>1740</td>
</tr>
<tr>
<td>817</td>
<td>Damauli</td>
<td>27.58</td>
<td>84.17</td>
<td>358</td>
</tr>
<tr>
<td>902</td>
<td>Rampur</td>
<td>27.37</td>
<td>84.25</td>
<td>256</td>
</tr>
<tr>
<td>905</td>
<td>Daman</td>
<td>27.36</td>
<td>85.05</td>
<td>2314</td>
</tr>
<tr>
<td>906</td>
<td>Hetauda</td>
<td>27.25</td>
<td>85.03</td>
<td>474</td>
</tr>
<tr>
<td>927</td>
<td>Bharatpur</td>
<td>27.4</td>
<td>84.26</td>
<td>205</td>
</tr>
<tr>
<td>1001</td>
<td>Timure</td>
<td>28.17</td>
<td>85.23</td>
<td>1900</td>
</tr>
</tbody>
</table>
### A.2 Precipitation observation stations

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Station Name</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Elevation (m a.s.l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1004</td>
<td>Nuwakot</td>
<td>27.55</td>
<td>85.1</td>
<td>1003</td>
</tr>
<tr>
<td>1005</td>
<td>Dhading</td>
<td>27.52</td>
<td>84.56</td>
<td>1420</td>
</tr>
<tr>
<td>1038</td>
<td>Dhunibesi</td>
<td>27.43</td>
<td>85.11</td>
<td>1085</td>
</tr>
<tr>
<td>1055</td>
<td>Dhunche</td>
<td>28.06</td>
<td>85.18</td>
<td>1982</td>
</tr>
<tr>
<td>1000</td>
<td>langtang</td>
<td>28.12</td>
<td>85.32</td>
<td>3670</td>
</tr>
<tr>
<td>601</td>
<td>Jomsom</td>
<td>28.783</td>
<td>83.717</td>
<td>2744</td>
</tr>
<tr>
<td>604</td>
<td>Thakmarpha</td>
<td>28.75</td>
<td>83.7</td>
<td>2566</td>
</tr>
<tr>
<td>605</td>
<td>Baglung</td>
<td>28.267</td>
<td>83.6</td>
<td>984</td>
</tr>
<tr>
<td>606</td>
<td>Tatopani</td>
<td>28.483</td>
<td>83.65</td>
<td>1243</td>
</tr>
<tr>
<td>609</td>
<td>Benibajar</td>
<td>28.35</td>
<td>83.567</td>
<td>835</td>
</tr>
<tr>
<td>610</td>
<td>Ghami</td>
<td>29.05</td>
<td>83.883</td>
<td>3465</td>
</tr>
<tr>
<td>612</td>
<td>Mustang</td>
<td>29.183</td>
<td>83.967</td>
<td>3705</td>
</tr>
<tr>
<td>613</td>
<td>karki neta</td>
<td>28.183</td>
<td>83.75</td>
<td>1720</td>
</tr>
<tr>
<td>614</td>
<td>Kusma</td>
<td>28.217</td>
<td>83.7</td>
<td>891</td>
</tr>
<tr>
<td>615</td>
<td>Bobang</td>
<td>28.4</td>
<td>83.1</td>
<td>2273</td>
</tr>
<tr>
<td>616</td>
<td>Gurja khani</td>
<td>28.6</td>
<td>83.217</td>
<td>2530</td>
</tr>
<tr>
<td>619</td>
<td>Ghorepani</td>
<td>28.4</td>
<td>83.733</td>
<td>2742</td>
</tr>
<tr>
<td>620</td>
<td>Tribeni</td>
<td>28.033</td>
<td>83.65</td>
<td>700</td>
</tr>
<tr>
<td>621</td>
<td>Darbang</td>
<td>28.383</td>
<td>83.4</td>
<td>1160</td>
</tr>
<tr>
<td>622</td>
<td>Rangkhani</td>
<td>28.15</td>
<td>83.567</td>
<td>1740</td>
</tr>
<tr>
<td>624</td>
<td>Samar Gaun</td>
<td>28.967</td>
<td>83.783</td>
<td>3570</td>
</tr>
<tr>
<td>625</td>
<td>Sanda</td>
<td>28.9</td>
<td>83.683</td>
<td>3570</td>
</tr>
<tr>
<td>626</td>
<td>Begha</td>
<td>28.467</td>
<td>83.6</td>
<td>1770</td>
</tr>
<tr>
<td>627</td>
<td>Kunun</td>
<td>28.383</td>
<td>83.483</td>
<td>1550</td>
</tr>
<tr>
<td>628</td>
<td>Muna</td>
<td>28.05</td>
<td>83.3</td>
<td>1970</td>
</tr>
<tr>
<td>629</td>
<td>Baghara</td>
<td>28.567</td>
<td>83.383</td>
<td>2330</td>
</tr>
<tr>
<td>630</td>
<td>Sirkon</td>
<td>28.133</td>
<td>83.617</td>
<td>790</td>
</tr>
<tr>
<td>701</td>
<td>Ridi Bajar</td>
<td>27.95</td>
<td>83.433</td>
<td>442</td>
</tr>
<tr>
<td>Station ID</td>
<td>Station Name</td>
<td>Latitude</td>
<td>Longitude</td>
<td>Elevation (m a.s.l)</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------</td>
<td>----------</td>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>702</td>
<td>Palpa</td>
<td>27.867</td>
<td>83.533</td>
<td>1067</td>
</tr>
<tr>
<td>704</td>
<td>Beluwo</td>
<td>27.683</td>
<td>84.05</td>
<td>150</td>
</tr>
<tr>
<td>706</td>
<td>Dumkauli</td>
<td>27.683</td>
<td>84.217</td>
<td>154</td>
</tr>
<tr>
<td>715</td>
<td>Khanchikot</td>
<td>27.933</td>
<td>83.15</td>
<td>1760</td>
</tr>
<tr>
<td>716</td>
<td>kapilvastu</td>
<td>27.55</td>
<td>83.067</td>
<td>94</td>
</tr>
<tr>
<td>722</td>
<td>Musikot</td>
<td>28.017</td>
<td>83.267</td>
<td>1280</td>
</tr>
<tr>
<td>725</td>
<td>Tamghas</td>
<td>28.067</td>
<td>83.25</td>
<td>1530</td>
</tr>
<tr>
<td>726</td>
<td>Garakot</td>
<td>27.867</td>
<td>83.8</td>
<td>500</td>
</tr>
<tr>
<td>802</td>
<td>Hudibajar</td>
<td>28.283</td>
<td>84.367</td>
<td>823</td>
</tr>
<tr>
<td>804</td>
<td>Pohara Airport</td>
<td>28.217</td>
<td>84</td>
<td>827</td>
</tr>
<tr>
<td>805</td>
<td>Syangja</td>
<td>28.1</td>
<td>83.883</td>
<td>868</td>
</tr>
<tr>
<td>807</td>
<td>Kunchha</td>
<td>28.133</td>
<td>84.35</td>
<td>855</td>
</tr>
<tr>
<td>808</td>
<td>Bandipur</td>
<td>27.933</td>
<td>84.417</td>
<td>965</td>
</tr>
<tr>
<td>809</td>
<td>Gorkha</td>
<td>28</td>
<td>84.617</td>
<td>1097</td>
</tr>
<tr>
<td>810</td>
<td>Chapkota</td>
<td>27.883</td>
<td>83.817</td>
<td>460</td>
</tr>
<tr>
<td>811</td>
<td>Malepatan(Phokhara)</td>
<td>28.117</td>
<td>84.117</td>
<td>856</td>
</tr>
<tr>
<td>813</td>
<td>Bhadaure Deurali</td>
<td>28.267</td>
<td>83.817</td>
<td>1600</td>
</tr>
<tr>
<td>814</td>
<td>Lumle</td>
<td>28.3</td>
<td>83.8</td>
<td>1740</td>
</tr>
<tr>
<td>815</td>
<td>Kharenitar</td>
<td>28.033</td>
<td>84.1</td>
<td>500</td>
</tr>
<tr>
<td>817</td>
<td>Damauli</td>
<td>27.967</td>
<td>84.283</td>
<td>358</td>
</tr>
<tr>
<td>818</td>
<td>Lamachaur</td>
<td>28.267</td>
<td>83.967</td>
<td>1070</td>
</tr>
<tr>
<td>820</td>
<td>Manang Bhot</td>
<td>28.067</td>
<td>84.017</td>
<td>3420</td>
</tr>
<tr>
<td>821</td>
<td>Ghandruk</td>
<td>28.383</td>
<td>83.8</td>
<td>1960</td>
</tr>
<tr>
<td>823</td>
<td>Gharedhunga</td>
<td>28.2</td>
<td>84.617</td>
<td>1120</td>
</tr>
<tr>
<td>824</td>
<td>Siklesh</td>
<td>28.367</td>
<td>84.1</td>
<td>1820</td>
</tr>
<tr>
<td>826</td>
<td>Walling</td>
<td>27.983</td>
<td>83.767</td>
<td>750</td>
</tr>
<tr>
<td>827</td>
<td>Rumjakot</td>
<td>27.867</td>
<td>84.133</td>
<td>660</td>
</tr>
<tr>
<td>829</td>
<td>Sallyan</td>
<td>28.267</td>
<td>83.75</td>
<td>1000</td>
</tr>
<tr>
<td>830</td>
<td>Pandur</td>
<td>28.267</td>
<td>83.783</td>
<td>1160</td>
</tr>
<tr>
<td>832</td>
<td>Dandaswanra</td>
<td>28.083</td>
<td>83.917</td>
<td>1432</td>
</tr>
<tr>
<td>833</td>
<td>Chheka kampar</td>
<td>28.483</td>
<td>85</td>
<td>3300</td>
</tr>
<tr>
<td>902</td>
<td>Rampur</td>
<td>27.617</td>
<td>84.417</td>
<td>256</td>
</tr>
<tr>
<td>903</td>
<td>Jhawani</td>
<td>27.583</td>
<td>84.533</td>
<td>270</td>
</tr>
</tbody>
</table>
## A.2 Precipitation observation stations

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Station Name</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Elevation (m a.s.l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>904</td>
<td>Chisapani Gadhi</td>
<td>27.55</td>
<td>85.133</td>
<td>1706</td>
</tr>
<tr>
<td>906</td>
<td>Hetauda</td>
<td>27.417</td>
<td>85.05</td>
<td>474</td>
</tr>
<tr>
<td>919</td>
<td>Makwanpur Gadi</td>
<td>27.417</td>
<td>85.017</td>
<td>1030</td>
</tr>
<tr>
<td>920</td>
<td>Beluwa (manahari)</td>
<td>27.55</td>
<td>84.817</td>
<td>274</td>
</tr>
<tr>
<td>925</td>
<td>Rajaya</td>
<td>27.433</td>
<td>84.983</td>
<td>332</td>
</tr>
<tr>
<td>927</td>
<td>Bharatpur</td>
<td>27.067</td>
<td>84.433</td>
<td>205</td>
</tr>
<tr>
<td>1001</td>
<td>Timure</td>
<td>28.283</td>
<td>85.383</td>
<td>1900</td>
</tr>
<tr>
<td>1002</td>
<td>Arughat D bajar</td>
<td>28.05</td>
<td>84.817</td>
<td>518</td>
</tr>
<tr>
<td>1004</td>
<td>Nuwakot</td>
<td>27.917</td>
<td>85.017</td>
<td>1003</td>
</tr>
<tr>
<td>1005</td>
<td>Dhading</td>
<td>27.867</td>
<td>84.933</td>
<td>1420</td>
</tr>
<tr>
<td>1007</td>
<td>Kakani</td>
<td>27.8</td>
<td>85.25</td>
<td>2064</td>
</tr>
<tr>
<td>1038</td>
<td>Dhunibesi</td>
<td>27.717</td>
<td>85.183</td>
<td>1085</td>
</tr>
<tr>
<td>1055</td>
<td>Dhunche</td>
<td>28.1</td>
<td>85.3</td>
<td>1982</td>
</tr>
<tr>
<td>1057</td>
<td>Pansayakhola</td>
<td>28.017</td>
<td>85.117</td>
<td>1240</td>
</tr>
<tr>
<td>1077</td>
<td>Sundarijal</td>
<td>27.75</td>
<td>85.417</td>
<td>1360</td>
</tr>
<tr>
<td>1079</td>
<td>Nagarjun</td>
<td>27.75</td>
<td>85.25</td>
<td>1690</td>
</tr>
<tr>
<td>1000</td>
<td>Langtang khola</td>
<td>28.209</td>
<td>85.547</td>
<td>3670</td>
</tr>
</tbody>
</table>