

1-1-2014

# Improving the Physical Processes and Model Integration Functionality of an Energy Balance Model for Snow and Glacier Melt

Avirup Sen Gupta  
*Utah State University*

---

## Recommended Citation

Sen Gupta, Avirup, "Improving the Physical Processes and Model Integration Functionality of an Energy Balance Model for Snow and Glacier Melt" (2014). *All Graduate Theses and Dissertations*. Paper 3875.  
<http://digitalcommons.usu.edu/etd/3875>

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact becky.thoms@usu.edu.



IMPROVING THE PHYSICAL PROCESSES AND MODEL INTEGRATION  
FUNCTIONALITY OF AN ENERGY BALANCE MODEL FOR SNOW AND  
GLACIER MELT

by

Avirup Sen Gupta

A dissertation submitted in partial fulfillment  
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Civil and Environmental Engineering

Approved:

---

David G. Tarboton  
Major Professor

---

Christopher M. U. Neale  
Committee Member

---

Jeffery S. Horsburgh  
Committee Member

---

David E. Rosenberg  
Committee Member

---

Jiming Jin  
Committee Member

---

Mark R. McLellan  
Vice President for the Research and  
Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY  
Logan, Utah

2014

Copyright © Avirup Sen Gupta 2014  
All Rights Reserved

## ABSTRACT

Improving the Physical Processes and Model Integration Functionality of an Energy  
Balance Model for Snow and Glacier Melt

by

Avirup Sen Gupta, Doctor of Philosophy

Utah State University, 2014

Major Professor: David G. Tarboton  
Department: Civil and Environmental Engineering

Quantification of the hydrologic water sources (snow, ice and rain) to river discharge in the Hindu-Kush Himalayan (HKH) region is important for decision-making in water sensitive sectors, such as flood protection. In this region, access to and monitoring of snow and glaciers and their melt outflow is challenging; thus modeling based on reanalysis and remote sensing data offers the potential for providing information to improve water resources decision making and management. Here I advanced the streamflow prediction capability in Himalayan watersheds by (1) developing a grid-based input/output framework for a point-based snow model to support its operational application in an integrated modeling system, (2) developing tools for spatial downscaling globally available reanalysis weather data to drive an energy balance model in the areas where meteorological observations are scarce and (3) extending an energy balance snowmelt model to include processes for quantifying melt from glaciers. In combination these provide the capability to model surface water supply and examine

changes in the contribution of glaciers to Himalayan water resources. This work uses the Utah Energy Balance (UEB) Snowmelt Model that simulates water inputs to runoff by modeling surface energy balance processes driven by weather inputs. The grid-based input-output framework extends UEB to run on a distributed mesh of grid cells with output aggregated over subwatersheds to facilitate integration into the EPA BASINS modeling system where it can be coupled with models such as the Geospatial Streamflow Forecast Model (GeoSFM) used to simulate streamflow in the Himalayan region. To overcome data scarcity in the HKH region, I developed an R-based procedure to downscale the NASA Modern-Era Retrospective Analysis for Research and Applications (MERRA) weather products and the NOAA Southern Asia Daily Rainfall estimate (RFE2) to obtain UEB inputs. The downscaling methods provide spatially and temporally continuous weather and incoming shortwave and longwave radiation data which are needed to drive energy and mass balance models such as UEB in hydrology. The ability to run models driven by these inputs supported examining questions related to the contribution of glaciers to water resources, thereby improving our understanding of hydrology in this area. This study also enhanced UEB by adding capability to quantify glacier melt. Direct physically based validation of this system is challenging due to the data scarcity in this region, but, to the extent possible, the model was validated through comparison to observed streamflow and to point measurements at the locations in the United States having available data.

(213 pages)

## PUBLIC ABSTRACT

### Improving the Physical Processes and Model Integration Functionality of an Energy Balance Model for Snow and Glacier Melt

by

Avirup Sen Gupta, Doctor of Philosophy

Utah State University, 2014

Major Professor: David G. Tarboton  
Department: Civil and Environmental Engineering

The Hindu-Kush Himalayan region possesses a large resource of snow and ice, which acts as a freshwater reservoir for irrigation, domestic water consumption or hydroelectric power for billions of people in South Asia. Monitoring hydrologic resources in this region is challenging because of the difficulty of installing and maintaining a climate and hydrologic monitoring network, limited transportation and communication infrastructure and difficult access to glaciers. As a result of the high, rugged topographic relief, ground observations in the region are extremely sparse. Reanalysis data offer the potential to compensate for the data scarcity, which is a barrier in hydrological modeling and analysis for improving water resources management. Reanalysis weather data products integrate observations with atmospheric model physics to produce a spatially and temporally complete weather record in the post-satellite era. This dissertation creates an integrated hydrologic modeling system that tests whether streamflow prediction can be improved by taking advantage of the National Aeronautics

and Space Administration (NASA) remote sensing and reanalysis weather data products in physically based energy balance snow melt and hydrologic models. This study also enhances the energy balance snowmelt model by adding capability to quantify glacier melt. The novelty of this integrated modeling tool resides in allowing the user to isolate various components of surface water inputs (rainfall, snow and glacier ice melt) in a cost-free, open source graphical-user interface-based system that can be used for government and institutional decision-making. Direct, physically based validation of this system is challenging due to the data scarcity in this region, but, to the extent possible, the model was validated through comparison to observed streamflow and to point measurements at locations in the United States having available data.

## DEDICATION

This work is dedicated to my parents, Shafali Sen Gupta and Asit Sen Gupta.

## ACKNOWLEDGMENTS

I dedicate this dissertation to my parents, Shafali Sen Gupta and Asit Sen Gupta, for their spiritual endless support. I am grateful to my sister, my wife and my family in Bangladesh for their unconditional love and support in every possible way throughout the process of this course, this dissertation and beyond. None of this would have been possible without their love and care.

My deepest gratitude is to my advisor, Dr. David G. Tarboton, for his generous availability despite his busy schedule to provide valuable suggestions and encouragements during the work. I thank him for setting up high academic standards for his students and guiding them to meet those standards. I have been amazingly fortunate to have an advisor like him; without his support this dissertation could not have been completed. I deeply appreciate his persistence in pushing me to improve my research and writing skills so as to produce the best work possible.

I want to thank my committee members, Dr. Christopher Neale, Dr. David Rosenberg, Dr. Jeffery Horsburgh and Dr. Jiming Jin, for their service in my committee and their critical reflections of my work. My special thanks to Ms. Carri Richards at Utah Water research Laboratory for providing language and writing help. Further thanks go to system administrator Kim Schreuders who assisted me to improve my programming knowledge skills in the earlier part of my PhD program.

I thank my fellow graduate students in the hydrology lab (USU Engineering Lab - 33) including Vinod, Ibrahim, Sulochan, John, Tseganeh, Nazmus, Ali and Tian for their friendship and their ability to alleviate the stresses of graduate school. A big part of the

graduate school experience has been interacting with fellow graduate students. I greatly value their friendship and I deeply appreciate their belief in me.

I am also thankful to Dr. Molly Brown, Dr. Shahid Habib from NASA and Dr. Mandira Shrestha from the International Centre for Integrated Mountain Development (ICIMOD) for providing me valuable data, suggestions and financial support. I would also like to thank Hindu Kush Himalayan (HKH) Cryosphere Monitoring Project implemented by ICIMOD and ICIMOD's glacier hydrologist Dr. Joseph Michael Shea for providing monthly temperature lapse rate data for Langtang Khola watershed.

This research was supported by the NASA Grant NNX11AK03G: Extending Utah Energy Balance Snowmelt Model to Glacial Melt Calculations in HIMALA

Avirup Sen Gupta

## CONTENTS

	Page
ABSTRACT .....	iii
PUBLIC ABSTRACT .....	v
DEDICATION .....	vii
ACKNOWLEDGMENTS .....	viii
LIST OF TABLES .....	xiii
LIST OF FIGURES .....	xiv
<b>CHAPTER</b>	
1. INTRODUCTION .....	1
1.1. Problem Statement.....	1
1.2. Objective.....	4
1.3. Literature Review .....	5
1.4. Summary.....	14
References.....	16
2. INTEGRATION OF AN ENERGY BALANCE SNOWMELT MODEL INTO AN OPEN SOURCE MODELING FRAMEWORK .....	21
Abstract.....	21
2.1. Introduction.....	21
2.2. Background.....	26
2.2.1. Utah Energy Balance Snowmelt Program .....	26
2.2.2. Geospatial Streamflow Model .....	28
2.2.3. BASINS .....	30
2.3. Integrated HIMALA BASINS System .....	31
2.3.1. UEB and GeoSFM Plugings to EPA BASINS .....	31
2.3.2. UEB Snowmelt Program Data Model .....	33
2.4. Langtang Khola Watershed Case Study .....	37
2.4.1. Data Sources .....	38
2.4.2. Model Setup.....	41

	xi
2.4.3. Case Study Results .....	42
2.5. Discussion.....	44
2.6. Conclusion .....	47
References.....	48
3. MSDH V1.0: A MERRA SURFACE CLIMATE AND RADIATION DATA SPATIAL DOWNSCALING TOOL FOR HYDROLOGICAL APPLICATION .....	62
Abstract.....	62
3.1. Introduction.....	63
3.2. Background.....	69
3.3. Data and Methods .....	73
3.3.1. Data Sources for Downscaling .....	73
3.3.2. Downscaling Methodology.....	74
3.3.3. Utah Energy Balance Snow and Glacier Melt Model .....	81
3.4. Software Implementation.....	82
3.4.1. Implementing Downscaling Algorithms in R.....	82
3.4.2. Output Data Storage in NetCDF.....	84
3.4.3. MSDH Graphical User Interface .....	86
3.5. Example Application .....	87
3.6. Discussion.....	92
3.7. Conclusion .....	93
References.....	96
4. ESTIMATING SNOW AND GLACIER MELT IN A HIMALAYAN WATERSHED USING AN ENERGY BALANCE SNOW AND GLACIER MELT MODEL.....	112
Abstract.....	112
4.1. Introduction.....	113
4.2. Study Area .....	117
4.3. Input Data .....	117
4.4. Model Description .....	119
4.4.1. Distributed Utah Energy Balance Snow and Glacier Melt Model (UEBGrid) .....	119
4.4.2. Geospatial Stream Flow Model (GeoSFM) .....	123
4.4.3. Model Setup.....	123
4.5. Results.....	126

4.5.1. Hydrologic Variability at Three Point Locations on Bare Ground, Clean Glacier and Debris-covered Glacier .....	126
4.5.2. Total Surface Water Input, Observed and Simulated Streamflow .....	129
4.6. Discussion.....	134
4.7. Conclusion .....	136
References.....	137
5. SUMMARY, CONCLUSION, AND RECOMMENDATIONS.....	154
5.1. Summary and Conclusion.....	154
5.2. Recommendations.....	162
References.....	165
APPENDICES .....	167
Appendix A.....	168
Appendix B.....	173
Appendix C.....	177
Appendix D.....	179
Appendix E.....	188
CURRICULUM VITAE.....	193

## LIST OF TABLES

Table	Page
3.1. Input MERRA variables used for downscaling .....	102
3.2. NRCS SNOTEL stations in Logan River Watershed .....	102
3.3. Nash-Sutcliffe Efficiency (NSE) of downscaled daily temperature, incoming shortwave, relative humidity, wind speed and precipitation data at USU Doc Daniel weather station.....	102
3.4. Nash-Sutcliffe Efficiency (NSE) of daily maximum, minimum and mean temperature and total monthly precipitation at six NRCS SNOTEL stations in the Logan River watershed .....	103
3.5. Nash-Sutcliffe Efficiency (NSE) of UEB simulated Snow Water Equivalent (m) compared to SNOTEL stations for water year 2010 .....	103
4.1. Physical characteristics of the Langtang Khola Watershed.....	142
4.2. Summary statistics of performance criteria between observed and simulated streamflow .....	142
4.3. Mass balance components.....	143

## LIST OF FIGURES

Figure	Page
2.1. Plug-in architecture of HIMALA BASINS .....	54
2.2. HIMALA BASINS Graphical User Interface.....	55
2.3. Organization of input and output files in the UEB program.....	56
2.4. Representing space and time in a netCDF file.....	57
2.5. Langtang Khola watershed. ....	58
2.6. Data workflow of input preprocessing and coupled UEB and GeoSFM modeling system in EPA BASINS for Langtang Khola case study. ....	59
2.7. Downscaling of MERRA temperature (°C) for Langtang Khola watershed at 3:00 am on January 1, 2003. (a) South Asian region temperature; (b) MERRA grid cells spanning Langtang Khola watershed (c) Downscaled temperature projected to Lambert Azimuthal Equal Area DEM grid.....	60
2.8. (a) Daily and (b) cumulative time series of UEB simulated surface water input components, streamflow measured at the Kyangin station and streamflow simulated by GeoSFM .....	61
3.1. Graphical User Interface for MERRA Spatial Downscaling for Hydrology (MSDH) .....	104
3.2. Logan River watershed. Blue lines indicate the stream network within the watershed and red dots symbolized the SNOTEL climate stations. Station numbers correspond with the numbers shown in table 3.2.....	105
3.3. Relationship between elevation (x-axis) and (a) mean monthly temperature (°C) (y-axis) and (b) mean monthly precipitation (m) (y-axis) at six SNOTEL stations in December for evaluating vertical temperature lapse rate ( $\Gamma$ ) and precipitation adjustment factor ( $\kappa_p$ ), respectively.....	106
3.4. Downscaling of MERRA temperature (o C) for the Logan River watershed 18:00 UTC on Dec 24, 2011 (a) temperature reported in MERRA for Contiguous United States of America (USA); (b) MERRA grid cells spanning Logan River watershed and surrounding areas and (c) downscaled temperature projected to Universal Transverse Mercator (UTM) projection at DEM grid resolution.....	107

3.5.	Comparison of downscaled daily mean observed temperature, shortwave radiation, wind speed, relative humidity, and precipitation, with respect to measured data at the USU Doc Daniel SNOTEL station. A time series plot (left) and scatter plot (right) of observed and downscaled data are shown for each variable.....	108
3.6.	Comparison of the downscaled data (y-axis) for daily maximum, minimum and mean temperature and monthly precipitation with observed data (x-axis) at six SNOTEL stations for water year 2010 (Oct 01 2009 - Sep 30 2010). SNOTEL station ID numbers are indicated at the top of each column .....	109
3.7.	Comparison between observed SWE and Utah Energy Balance (UEB) simulated SWE for water year 2010 (Oct 01 2009 - Sep 30 2010) at the six SNOTEL stations .....	110
3.8.	Comparison between the observed and UEB simulated snow water equivalent (SWE) (a) using observed temperature, precipitation, wind speed, relative humidity and shortwave radiation, (b) using downscaled temperature with observed data of other variables, (c) using downscaled precipitation with observed data of other variables, (d) using downscaled wind speed with observed data of other variables, (e) using downscaled relative humidity with observed data of other variables, (f) using downscaled shortwave radiation with observed data of other variables.....	111
4.1.	Location of Nepal and Langtang Khola watershed in South Asia. Bare ground, clean and debris covered glacier ice are shown in antique white, sky blue and dark orange colors, respectively. Three selected points in three substrates (A at bare ground, B at clean glacier and C at debris covered glacier) are shown by green points. Hatched area is the subwatershed depicted in Figure 4.3.....	144
4.2.	Glacier representation in UEBGrid.....	145
4.3.	(a) Substrate type and (b) substrate albedo. In Figure 4.1, this subwatershed of Langtang Khola is shown as the hatched area .....	146
4.4.	Inputs (i.e. (a) temperature ( $^{\circ}\text{C}$ ), (b) precipitation (m/hr) and (c) solar radiation ( $\text{KJ m}^{-2} \text{ hr}^{-1}$ )) and outputs (i.e. (d) SWE (m), (e) Total Surface water input (SWIT, m/hr) and (f) glacier melt (SWIGM, m/hr)) at three different points (A: bare ground, B: Clean glacier, C: Debris-covered glacier) aggregated (i.e., mean) at a daily scale for the year 10/1/2002-9/30/2003, the first year after the spin up period. Color legend shown in	

(d) for bare ground (red), clean (blue) and debris covered glacier (green) remains unchanged for other figures (i.e., a through f) .....	147
4.5. Comparison between melt estimated by UEBGrid and streamflow measured at Kyangjin hydrologic station .....	148
4.6. Comparison between observed and GeoSFM simulated streamflow driven by UEBGrid modeled surface water inputs. (a) Daily time series of observed and simulated streamflow, (b) scatter plot of daily observed and simulated streamflow, (c) monthly mean observed and simulated streamflow.....	149
4.7. Yearly contribution of glacier, snow, and rain in total surface water input (SWIT) .....	150
4.8. Monthly variation of SWIR (black bar), SWISM (dark grey bar), SWIGM (light grey bar), observed streamflow (black line), temperature (red line), and sublimation (dark golden line) averaged over the 10 years 2003-2012 (streamflow only available to 2010) .....	151
4.9. Rain, snowmelt and glacier melt from three different substrate types .....	152
4.10. Paired relationship between annual surface water input from glacier melt (SWIGM, meter), total surface water input (SWIT, meter), precipitation (meter), observed streamflow, (Observed Q, meter), simulated streamflow, (Simulated Q, meter), sublimation (m) and temperature (°C). The numerical values given at the top of each plot represents the correlation coefficient between the two variables. Linear regression lines are shown in red .....	153

## CHAPTER 1

### INTRODUCTION

#### 1.1. Problem Statement

High altitude watersheds in the Hindu-Kush Himalayan (HKH) region possess a large volume of snow and ice, which acts as a freshwater reservoir for irrigation, domestic water consumption or hydroelectric power for 1.4 billion people in South Asia (Immerzeel et al., 2010). During the dry years with less precipitation, glaciers act as natural buffers, releasing meltwater and sustaining relatively stable flow in the streams. With population growth and projected climate-induced changes in snow and ice, the region is at risk of experiencing water stress in the long-term (Immerzeel et al., 2010; 2012; Kaser et al., 2010), but is also vulnerable to high flow during the melting season in the near future. In particular, there are concerns about the effect of climate change on glacier retreat, and variation in runoff contributions from snow and glacier melt.

As a data driven scientific tools, hydrological models require a vast amount of data for advancing the prediction capability of hydrological quantities such as streamflow. However, monitoring hydrological resources including glaciers in the HKH region is challenging due to the difficulty in installing and maintaining weather and hydrologic stations, limited transportation infrastructure, difficulty in accessing glaciers and lack of financial support. As a result of the extreme rugged topography, ground observations in the region are extremely sparse and inadequate for hydrologic analysis (Lo et al., 2011). A limited number of weather stations are maintained by the Nepal Department of Hydrology and Meteorology (DHM) for collecting hydrological and meteorological variables such as discharge, temperature, and precipitation in Himalayan

headwaters in Nepal (Konz et al., 2007). These data are a unique resource to examine the physical processes of runoff generation (Konz et al., 2007) and make predictions of runoff from high altitude Himalayan watersheds; however, these measurements are often unable to represent the spatio-temporal variability of the complex terrain and suffer from prolonged periods of missing data. In recent years, the advancement in remote sensing data has greatly helped modelers to model snow and glacier melt in the Himalayas (Thayyen and Gergan, 2010) and has helped to improve water resources decision making and management in these data-scarce areas. Remote sensing data has been used for various hydrological and meteorological studies such as: estimating snowmelt runoff (Thapa, 1993), mapping and characterizing glaciers (Racoviteanu et al., 2008), and using MODIS snow cover for calibrating a distributed hydrological model (Konz et al., 2010). In this study, we used remote sensing data, along with weather reanalysis products and ground-based streamflow observations to model snow and glacier melt for a high altitude watershed in the HKH region.

While some progress has been made in understanding the contribution of snow and ice melt to streamflow, most studies in the HKH region use degree-day models or simple ablation models (Immerzeel et al., 2010; 2012; Racoviteanu et al., 2013). The ablation models analyze the distribution of area over the range elevation in a watershed. Ice area is typically calculated for 100 m altitudinal bands below the average elevation at which annual accumulation equals ablation, the equilibrium line altitude (ELA). These ablation models assume a separate yearly melt rate (mm/day) for clean and debris covered glacier for each elevation band over the watershed (Racoviteanu et al., 2013). Degree-day models, on the other hand, estimate runoff using mean daily or monthly air

temperature and precipitation (Braun et al., 1992; Immerzeel et al., 2010; 2012). Both of these work reasonably well given that they have limited data requirements; however, they do not account for topographic effects, surface albedo, solar radiation and turbulent heat exchanges in melt calculations (Hock, 2003). Recognizing these limitations, this study explored using the physically-based Utah Energy Balance (UEB) snowmelt model (Tarboton et al., 1995) to compute glacier melt. Compared to multi-layer energy balance snowmelt models, UEB is a relatively simple energy balance model that parameterizes the snowpack using lumped (depth averaged) state variables so as to avoid having to model the complex processes that occur within a snowpack. Physical differences between bulk (depth averaged) properties and the surface properties, which are important for calculating surface energy exchanges, are captured by modeling diurnally forced heat flow at the surface using the so-called force-restore parameterization where there is a forcing term related to the difference between surface and depth temperature and a restore term related to the temporal gradient of surface temperature (Deardorff, 1978; Luce and Tarboton, 2010). UEB was chosen for this purpose because the model is open source and provides a relatively simple, transferable, physically-based approach to the quantification of snowmelt.

UEB was initially designed to simulate snowmelt and track the energy and mass balance of snow to model snow accumulation and melt at a single point location such as a weather station (Tarboton and Luce, 1996; Mahat and Tarboton, 2012; 2013; Mahat et al., 2013). In this study, a spatially distributed version of the UEB snowmelt model was developed that applies the model separately at each point on a watershed grid. The physical processes represented by the model were extended to include glacier melt in

addition to snowmelt. The model represents glacier as a substrate layer and computes melt from glacier substrate when seasonal snow cover has melted. Glacier outlines and the albedo of clean glaciers and debris-covered glaciers are used as inputs to the model.

This work was part of a larger NASA sponsored project to apply NASA data products to advancing understanding the snow and glacier melt processes and their relative contribution in streamflow. Inclusion of glacier melt estimates in hydrological analysis in South Asian countries is hindered by a lack of inexpensive, easy-to-use operational hydrologic models and a lack of training. To address this shortcoming, UEB was configured for incorporation as a plug-in to the Environmental Protection Agency (EPA) Better Assessment Science Integrating point and Nonpoint Sources (BASINS) modeling system (EPA, 1998). BASINS is a free, open-source, graphical user interface-based streamflow modeling system that allows users to configure and run a broad set of hydrologic models. The specific version developed for this broader NASA project is referred to as HIMALA BASINS. The addition of UEB to HIMALA BASINS adds the capability incorporate snow and glacier melt information into streamflow simulations using the GeoSFM (Asante et al., 2008) rainfall-runoff model. UEB was integrated into the EPA BASINS simulation environment by AQUATERRA Consultants, the prime developer of EPA BASINS using a graphical user interface coupling approach (Brandmeyer and Karimi, 2000).

## **1.2. Objective**

The primary objective of this work was to develop a hydrological tool that includes modeling of both snow and glacier-melt contributions to river flow in the HKH region. Our goal was to improve streamflow prediction by taking advantage of modern

remote sensing and high resolution weather data products to compensate for the scarcity of ground-based observation data. The UEB model has the capability to predict snowmelt at a point driven by weather data at that point. Our first objective was to convert the point-based UEB to a distributed model that allows it be run over a grid to better represent spatial variability. The second objective was to enable coupling of UEB with other models and enhance usability by integration into the EPA Basins modeling framework. The third objective was to add a glacier substrate layer to the model to compute glacier melt when the seasonal snow cover is melted. Representation of glacier substrate and spatially explicit representation of variability on a grid allowed the model to estimate runoff generation from snow and glaciers. To overcome limitations in data availability, my fourth objective was to develop tools to retrieve distributed weather inputs from Internet data repositories, and downscale them to the spatial footprint of the model. We addressed the following questions by applying the model at a high altitude glacierized Himalayan watershed:

- (1) How well can glacier melt be quantified using adaptations of a simple energy balance model initially developed for snow?
- (2) What is the relative contribution of glacier melt, snowmelt and rain to the total surface water input?
- (3) Are there any changes in glacier mass balance during the model simulation period?

### **1.3. Literature Review**

Review of snow and glacier melt studies in Nepal Himalayas has indicated that temperature index or degree-day methods are the most widely used because of their simplicity and less demanding data requirements. Braun et al. (1992) applied a HBV-

based hydrologic model (Sten and Arne, 1973) over Langtang Khola watershed. HBV is a conceptual rainfall-runoff model originally developed in Scandinavia (Bergstrom, 1976). Braun (1988) enhanced the original HBV model by introducing a temperature index model. Daily melt was estimated based on a seasonally variable melt rate (mm/day) and the difference between the daily air temperature and snow-rain transition temperature. The model keeps track of snow accumulation. Where glacier is present, an ice melt factor is used to compute the glacier melt after the snow cover melts out. The ice melt factor is assumed to be higher than that of snow due to ice's lower albedo. However, 7% of the total basin area is covered by debris-cover glacier ice. On debris-covered glacier, the debris layer works as an insulator in heat exchange between the atmosphere and glacier surface; therefore, a lower melt factor was assumed for debris-covered glacier. The model was driven by temperature and precipitation data and was validated using streamflow data at the gauged outlet. The simulated streamflow showed reasonable agreement with observed streamflow on a yearly scale; while streamflow was overestimated during the post monsoon and underestimated during the pre-monsoon seasons. Following this study, degree-day models were used for quantifying ice melt beneath the debris-covered Kumbu glacier, Nepal (Kayastha et al., 2000), the Langtang Valley (Kayastha et al., 2005), four glaciers in Nepalese Himalayas and Qinghai-Tibetan Plateau (Matsuda, 2003), the Tamakoshi basin in Nepal (Shilpkar et al., 2009) and others. These studies have been used to examine ice melting and discharge and climate change impacts on snow and glacier melt.

Konz et al. (2007) used the distributed tracer aided catchment model (TAC<sup>D</sup>) (Uhlenbrook and Leibundgut, 2002; Uhlenbrook et al., 2004), a modified version of the

HBV model with an enhanced process-based runoff computation module, to simulate streamflow from the Langtang Khola watershed. The model is coded using the PCRaster programing language. Recognizing the data scarcity, the authors used a simplified approach, requiring only temperature and precipitation data, illustrated by Braun et al. (1992) instead of using sophisticated TAC<sup>D</sup>. The authors presented a statistical approach to fill the gaps in temperature and precipitation data. Konz et al. (2010) calibrated the snow and glacier routine of the TAC<sup>D</sup> model using snow cover data from the Moderate-resolution Imaging Spectroradiometer (MODIS).

Immerzeel et al. (2010) used the Snowmelt Runoff Model (SRM), originally developed by Martinec (Martinec et al., 1994; Rango and Martinec, 1994; 1995) to model discharge over one hundred basins in South Asia. The original version of the SRM model simulates discharge based on snowmelt and rainfall using a deterministic degree day approach. The SRM model has been extended to include glacier melt as an additional source of water by assigning a separate melt rate for ice, if the glacier is present. The model is driven by daily temperature, precipitation and snow-covered area.

Immerzeel et al. (2012) also developed a combined cryospheric hydrologic model to investigate hydrologic response to climate change in the Langtang River watershed. This model includes the simulation of glacier movement by basal sliding. This is the first model, to my best knowledge, to model glacier flow in Nepal Himalayas. Temperature and precipitation are the main forcing variables and glacier melting is estimated using a degree-day approach. Temperature was spatially differentiated using a vertical lapse rate. While temperature is the primary reason to cause melt in the model, there is considerable variability in General Climate Model (GCM) temperature projections from 2000 to 2100.

Therefore, the uncertainty in temperature projections directly gets translated into glacier melt predictions and glacier mass balance.

Racoviteanu et al. (2013) used an ice ablation model to compute the glacier contribution to annual streamflow. The ice ablation model, initially developed by Alford (1992), uses (1) an estimate of the equilibrium line altitude (ELA) giving the average elevation over the basin where the accumulation is at an equilibrium with the ablation, (2) 100 m elevation bands to represent the elevation distribution of the basin, and (3) separate melt factors for clean and debris covered glacier used to compute melt for every 100 meter altitude band under the ELA based on the digital elevation model and glacier map. The study quantifies the total glacier outflow by multiplying the glacierized area with the melt rate for each specific elevation band. For clean and debris covered glacier, two separate relationships between the elevation and ablation rate were used. The study estimated that 58.3% of the streamflow was generated from glacier melt for Langtang Khola watershed. A major limitation of this model is that it cannot predict the seasonality of the glacier melt and, therefore, cannot be used for decision making in water resources management (such as irrigation in rice or wheat crops) where monthly or seasonal information are required.

Kayastha et al. (1999) applied a mass balance model on a small glacier AX010, Shorong, Himal in Nepal. The model was driven by hourly temperature, precipitation, wind speed, relative humidity and cloud cover. The model estimated surface albedo, shortwave and longwave radiation internally and the screening effect of the surrounding mountains on radiation and multiple reflections were also taken into account. A temperature-relative humidity relationship was used to separate snow from rain during a

precipitation event. This study assumed a constant temperature lapse rate for calculating temperature at high altitudes and constant precipitation over the basin. The surface-elevation changes were validated at three points in the glacier from 25 May to 25 September, 1978. Modeled and observed surface elevation of the glacier were in good agreement except for some discrepancies in June when a thin snow layer covered the glacier due to light precipitation. The authors found that the model was highly sensitive to radiation and surface albedo, suggesting the importance of using energy balance models in glacier and snow melt calculation.

In assessing the models reviewed above it is apparent that there are three types of models used for glacier melt studies in Nepal Himalayas (1) degree day models, (2) ice ablation models and (3) energy and mass balance models. Most recent degree-day models are distributed and capable of explicitly representing the temperature and precipitation variability. Among all these models, Kayastha et al. (1999) used an energy balance model to simulate the change in glacier elevation and showed that the model was well capable of simulating the glacier melt by comparing with the observed elevation change in glacier thickness. However, this model was only run over a small glacier while water resource managers are more interested in streamflow simulation over an entire watershed. Even so, the success of this energy balance model is one factor motivating the use of UEB as a starting point for my work.

In distributed hydrologic modeling a number of data structures have been developed for the discretization of the spatial domain into model elements that explicitly represent spatial variability. Here, a number of the relevant approaches are reviewed to

provide ideas and information for extending the single point UEB model into a distributed model.

Julien et al. (1995) developed a raster based hydrologic model (CASC2D) for surface runoff simulation. CASC2D has the capability to use raster-based GIS and radar data and graphically display maps of different variables such as rainfall and cumulative infiltration. The spatial area of the watershed is divided into equal size raster grid cells, and a mask map holding values 0 or 1 is used to identify whether a particular grid cell falls within the watershed. The model uses raster inputs of classified soil texture initial soil moisture, elevation, retention depth and surface roughness coefficient to compute overland flow. The raster capability in CASC2D was established by linking the program with the publicly available Geographical Resources Analysis Support System (GRASS) GIS software (Mitasova et al., 1995; Neteler and Mitasova, 2004)

Johnson and Miller (1997) developed another distributed hydrologic model with a goal to use remote sensing and a Digital Elevation Model (DEM) to reduce the dependence on ground-based observations. The model uses a DEM to delineate the watershed, compute flow directions and build the stream network. Then precipitation is applied to each grid cell to compute accumulated precipitation, excess precipitation, infiltration and runoff at each time step. The model then routes the excess runoff through the stream network to compute a hydrograph at the outlet.

The PCRaster Environmental Modeling Language (Van Deursen and Wesseling, 1992; 1996; Wesselung et al., 1996) is a dynamic programming language for developing and running spatially distributed environmental models. Like other higher level programming languages, codes in PCRaster are highly optimized, shorter, simpler, and

easy to understand and modify as its interface hides lower level details such as memory management (Van Der Knijff et al., 2010). There are two types of file formats that can be imported into and exported from PCRaster: (1) PCRaster Maps and (2) ASCII files. The language also contains a large number of built-in functions for mathematical analysis and raster processing. Initially developed at Utrecht University in the Netherlands, PCRaster is used for modeling in many disciplines such as hydrology, ecology, and glaciology. Example uses of PCRaster include the hydrologic and soil erosion model (LISEM) (De Roo et al., 1996), ground water modeling by linking with MODFLOW (Schmitz et al., 2009), and comparison between hydrologic processes between a mountainous and valley catchment (Zhang et al., 2009). Konz et al. (2007) and Immerzeel et al. (2012), respectively, applied the TAC<sup>D</sup> model (tracer aided distributed catchment model) and combined cryospheric hydrologic model in the Langtang Khola watershed. Both of these models are coded entirely in PCRaster. Despite its many benefits, PCRaster comes with some limitations that restrict its applicability to host an operational model (Van Der Knijff et al., 2010). Input data stored in raster formats such as netCDF, GRIB and GeoTiff require additional software and substantial effort for converting to PCRaster format. Additional software is also required to visualize tabulated time series of model outputs. PCRaster does not provide functionality to customize which variables should be output from a model run (Van Der Knijff et al., 2010). It also lacks the flexibility to switch on/off a particular module of the model without modifying the source code (Van Der Knijff et al., 2010).

The Arc Hydro Geographic Information System (GIS) data model (Maidment, 2002; Strassberg et al., 2007; 2011) was developed to provide a standard way to organize

digital elevation and hydrography information in support of hydrology oriented GIS analyses. Arc Hydro combines grid approaches with vector representations of catchments, streams and water bodies. Data is stored in Esri geodatabase format to facilitate analysis from within ArcGIS. Connectivity between the landscape and hydrography network in Arc Hydro starts at the scale of a grid cell. Terrain analysis methods are used to derive flow directions and flow accumulation, which, by thresholding, produces streams that are represented as vector features and catchments that are represented as vector polygons connected to streams. There is a one-to-one relationship between the area represented by a catchment and the stream segment to which it drains. Stream segments are also connected upstream and downstream, and this provides a connected discretization of the land surface that can serve as a template for hydrologic modeling with flow generated over catchments being routed down through the stream network.

Kumar et al. (2010) developed an object-oriented, shared data model to support distributed application of the Penn State Integrated hydrologic Model (PIHM). This shared data model developed for integrating PIHM and GIS defines various objects, their attributes and various operations that can be performed. The data model has the capability to represent the same object at different spatial and temporal scales. Clusters of hydrological entities with similar attributes are defined by a class, which may be further enhanced through additional properties that may be added to some of the members of the class. Real world hydrological entities are the instances of a class. For example, a polygon is a class and a lake may be an instance of that class. Water level time series can

be added as an additional data type to the “lake” instance. The data model incorporates a wide range of data types from static objects to dynamic 3-dimensional lines and fluxes.

Along with raster-based hydrologic models and modeling platforms, software tools have also been developed to facilitate linkage between hydrologic models with data in standard file formats. For example, the modular library for raster based hydrological application (MOSAICO) (Ravazzani, 2013b), is a set of libraries developed to support raster-based hydrologic simulations using netCDF as a standard file format. Ravazzani (2013a) showed an example where monthly precipitation is read from an ASCII file and area averaged precipitation is written in a netCDF file. Such applications can also be extended for more complicated modeling applications.

Given recent improvement in computational capability and Geographical Information System (GIS) and related tools, many distributed hydrologic models have emerged with various degrees of complexity. However, only a limited number of studies address the data management aspects of raster-based distributed hydrologic modeling. Many hydrologic models store raster-based input and output data in ASCII format, while others use binary data formats (Ravazzani, 2013a). Storage of numerical data values in ASCII is inefficient as reading and writing it requires translation between the ASCII representation and internal binary representation of numbers. However; binary formats make the data less accessible as they are not immediately human readable. Hydrologists have not agreed on a standard data format for storing and sharing data (Ravazzani, 2013a). The trend towards coupled or integrated environmental modeling for addressing broader environmental problems motivates the need for standard formats for efficiently storing, accessing and sharing data.

This literature review has indicated that due to data scarcity most Himalayan glacier melt efforts have been temperature index or ablation models that rely on empirical relationships developed between the snow and glacier melt and temperature or elevation. However the statistical basis for these models may weaken under climate change conditions and it is difficult to explicitly represent topographic effects other than elevation (i.e., slope and aspect) in these models. Distributed energy balance models are physically-based and are likely to have the potential to improve melt simulation capability. However they have been used only rarely in the HKH region. In this study, we reconfigure a point-based snowmelt model (i.e., UEB) to a distributed model by computing the melt from the snow and glacier. This reconfiguration is done by developing a structured data model using netCDF and ASCII file formats. The data model also enables UEB model to be coupled with other hydrologic models such as GeoSFM. The data model developed in this study can be adopted, extended or modified for the development of raster-based distributed hydrological models. The UEB model was also extended to include computation of glacier melt by adding a static glacier substrate layer into the model. These two additional features in UEB make the integrated system applicable for streamflow simulation in glacierized watersheds where glaciers are significant sources of water.

#### **1.4. Summary**

This work was driven by the need for better prediction of streamflow by modeling snow and glacier melt in Himalayan watersheds. This dissertation contributes towards developing an integrated modeling system to compute the snow and glacier melt and accumulation and to assess their relative importance to streamflow in a high altitude

Himalayan watershed. Our work included: (1) a distributed UEB model capable of running on a grid over a watershed to produce snow and glacier ablation, (2) an R-based weather and radiation data downscaling tool for generating time-varying dynamic input data for UEB and (3) an extended version of the model that includes the representation of a glacier substrate layer to compute glacier melt using UEB's existing formulation of surface energy balance fluxes. Each of these contributions is detailed in a chapter (paper) that follows.

The chapters that present this work are ordered to first present the distributed data model (Chapter 2), then the downscaling tools (Chapter 3) followed by the glacier layer extensions (Chapter 4). Development of the computer codes for the data model described in Chapter 2 and the glacier representation in Chapter 4 proceeded simultaneously, and each does depend on the other. However it seemed best to present the data model and integrated modeling system first with results for just one year at Langtang Khola watershed used to illustrate the integration functionality. Then in Chapter 3 the downscaling model is presented and Chapter 4 that presents the glacier representation serves as the capstone for this work. In Chapter 4 the UEB model was applied for ten years test to compute the glacier contribution and glacier mass balance over an extended period relying on downscaled input data using the model from Chapter 3. Streamflow simulated by GeoSFM with inputs from UEB over a longer time series was compared with 8 years of available observed streamflow. The reasonable comparisons obtained indicate that in aggregate discrepancies offset and the effects of biases are small and do not propagate as the time progresses.

## References

- Alford, D., 1992, Hydrological aspects of the Himalayan region, International Centre for Integrated Mountain Development, ICIMOD Occas. Pap., Kathmandu.
- Asante, K. O., G. A. Artan, S. Pervez, C. Bandaragoda and J. P. Verdin, 2008, Technical manual for the Geospatial Stream Flow Model (GeoSFM), U. S. Geological Survey.
- Bergstrom, S., 1976, Development and application of a conceptual runoff model for Scandinavian catchments., Dept. of Water Resources Engineering, Lund Inst. of Technology / University of Lund. Bull. Ser. A no. 52.
- Brandmeyer, J. E. and H. A. Karimi, 2000. Coupling methodologies for environmental models, Environ. Modelling & Software 15(5), 479-488.
- Braun, L., W. Grabs and B. Rana, 1992. Application of a Conceptual Precipitation-runoff Model in the Langtang Kfaola Basin, Nepal Himalaya, Proceedings of the Kathmandu Symposium, p. 221-237.
- Braun, L. N., 1988. Parameterization of snow-and glacier melt, Geographisches Institut ETH, Zurich.
- De Roo, A., C. G. Wesseling, V. G. Jetten and C. J. Ritsema, 1996. LISEM: a physically-based hydrological and soil erosion model incorporated in a GIS, IAHS PUBLICATION 235, 395-403.
- Deardorff, J. W., 1978. Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation, J. Geophys. Res. 83(C4), 1889-1903.
- EPA, B. V., 1998. 2.0 User's manual: better assessment science integrating point and nonpoint sources, US Environmental Protection Agency, Washington, DC.
- Hock, R., 2003. Temperature index melt modelling in mountain areas, J. Hydrol. 282(1), 104-115.
- Immerzeel, W. W., L. Van Beek, M. Konz, A. Shrestha and M. Bierkens, 2012. Hydrological response to climate change in a glacierized catchment in the Himalayas, Clim. Chang. 110(3-4), 721-736.
- Immerzeel, W. W., L. P. van Beek and M. F. Bierkens, 2010. Climate change will affect the Asian water towers, Science 328(5984), 1382-1385.
- Johnson, D. L. and A. C. Miller, 1997. A spatially distributed hydrologic model utilizing raster data structures, Comput. & Geosci. 23(3), 267-272.
- Julien, P. Y., B. Saghafian and F. L. Ogden, 1995. Rater-based hydrologic modeling of spatially-varied surface runoff, J. Am. Water Resour. Assoc. 31(3), 523-536.

- Kaser, G., M. Großhauser and B. Marzeion, 2010. Contribution potential of glaciers to water availability in different climate regimes, *Proceedings of the National Academy of Sciences* 107(47), 20223-20227.
- Kayastha, R. B., Y. Ageta and K. Fujita, 2005. Use of positive degree-day methods for calculating snow and ice melting and discharge in glacierized basins in the Langtang Valley, *Clim. and Hydrol. of Mt. Ar.*, 7-14 pp.
- Kayastha, R. B., T. Ohata and Y. Ageta, 1999. Application of mass balance model to a Himalayan glacier, *J. Glaciol.* 45(151), 559-567.
- Kayastha, R. B., Y. Takeuchi, M. Nakawo and Y. Ageta, 2000. Practical prediction of ice melting beneath various thickness of debris cover on Khumbu Glacier, Nepal, using a positive degree-day factor, *Debris-covered Glaciers: Proceedings of an International Workshop Held at the University of Washington in Seattle, Washington, USA*, p. 71-82.
- Konz, M., D. Finger, C. Buergi, S. Normand, W. Immerzeel, J. Merz, A. Giriraj and P. Burlando, 2010. Calibration of a distributed hydrological model for simulations of remote glacierised Himalayan catchments using MODIS snow cover data, *Global Change: Facing Risks and Threats to Water Resources*, p. 465-473.
- Konz, M., S. Uhlenbrook, L. Braun, A. Shrestha and S. Demuth, 2007. Implementation of a process-based catchment model in a poorly gauged, highly glacierized Himalayan headwater, *Hydrol. and Earth Syst. Sci.* 11(4), 1323-1339.
- Kumar, M., G. Bhatt and C. J. Duffy, 2010. An object-oriented shared data model for GIS and distributed hydrologic models, *Int. J. Geogr. Inform. Sci.* 24(7), 1061-1079.
- Lo, Y.-H., J. A. Blanco, B. Seely, C. Welham and J. P. Kimmins, 2011. Generating reliable meteorological data in mountainous areas with scarce presence of weather records: the performance of MTCLIM in interior British Columbia, Canada, *Environ. Modelling & Software* 26(5), 644-657.
- Luce, C. H. and D. G. Tarboton, 2010. Evaluation of alternative formulae for calculation of surface temperature in snowmelt models using frequency analysis of temperature observations, *Hydrol. and Earth Syst. Sci.* 14(3), 535-543.
- Mahat, V. and D. G. Tarboton, 2012. Canopy radiation transmission for an energy balance snowmelt model, *Water Resour. Res.* 48, W01534.
- Mahat, V. and D. G. Tarboton, 2013. Representation of canopy snow interception, unloading and melt in a parsimonious snowmelt model, *Hydrol. Process.* in press
- Mahat, V., D. G. Tarboton and N. P. Molotch, 2013. Testing above and below canopy representations of turbulent fluxes in an energy balance snowmelt model, *Water Resour. Res.* 49(2), 1107-1122.

- Maidment, D. R., 2002. ArcHydro: GIS for water resources, ESRI, Inc., Redlands, CA.
- Martinec, J., A. Rango and R. Roberts, 1994, The Snowmelt-Runoff Model (SRM) Users Manual, Updated Edition 1994, Version 3.2, Department of Geography - University of Bern, Bern, Switzerland.
- Matsuda, Y., 2003. Positive degree-day factors for ice ablation on four glaciers in the Nepalese Himalayas and Qinghai-Tibetan Plateau, *Bull. Glaciol. Res.* 20, 7-14.
- Mitasova, H., L. Mitas, W. M. Brown, D. P. Gerdes, I. Kosinovsky and T. Baker, 1995. Modelling spatially and temporally distributed phenomena: new methods and tools for Grass GIS, *Int J. Geogr. Inform. Syst.* 9(4), 433-446.
- Neteler, M. and H. Mitasova, 2004. Open source GIS: a GRASS GIS approach, Vol 2, Springer, New York.
- Racoviteanu, A. E., R. Armstrong and M. W. Williams, 2013. Evaluation of an ice ablation model to estimate the contribution of melting glacier ice to annual discharge in the Nepal Himalaya, *Water Resour. Res.* 49(9), 5117-5133.
- Racoviteanu, A. E., M. W. Williams and R. G. Barry, 2008. Optical remote sensing of glacier characteristics: a review with focus on the Himalaya, *Sensors* 8(5), 3355-3383.
- Rango, A. and J. Martinec, 1994. Areal extent of seasonal snow cover in a changed climate, *Nordic Hydrol.* 25(1994), 233-246.
- Rango, A. and J. Martinec, 1995. Revisiting the degree-day method for snowmelt computations, *Water Resour. Bull.* 31(4), 657-669.
- Ravazzani, G., 2013a. MOSAICO, a Fortran 90 Free Open Source library for raster based hydrological applications, EGU General Assembly Conference, 15, p. 8761.
- Ravazzani, G., 2013b. MOSAICO, a library for raster based hydrological applications, *Comput. & Geosci.* 51, 1-6.
- Schmitz, O., D. Karssenberg, W. P. van Deursen and C. Wesseling, 2009. Linking external components to a spatio-temporal modelling framework: Coupling MODFLOW and PCRaster, *Environ. Modelling & Software* 24(9), 1088-1099.
- Shilpkar, R. B., N. M. Shakya and A. Hiratsuka, 2009. Impact of climate change on snowmelt runoff: A case study of Tamakoshi basin in Nepal, International Symposium on Society for Social Management Systems, SSMS2009, Kochi, Japan, Internet Journal, SMS09-124, p.
- Sten, B. and F. Arne, 1973. Development of a conceptual deterministic rainfall-runoff model, *Nordic Hydrol.* 4(3), 147-170.

- Strassberg, G., N. L. Jones and D. R. Maidment, 2011. Arc hydro groundwater, Esri Press, Austin, Texas.
- Strassberg, G., D. R. Maidment and N. L. Jones, 2007. A geographic data model for representing ground water systems, *Ground Water* 45(4), 515-518.
- Tarboton, D. G., T. G. Chowdhury and T. H. Jackson, 1995. A Spatially Distributed Energy Balance Snowmelt Model, in *Biogeochemistry of Seasonally Snow-Covered Catchments (Proceedings of a Boulder Symposium, July 1995)*, Edited by K. A. Tonnessen, M. W. Williams and M. Tranter, IAHS Publ. no. 228, Wallingford, p. 141-155.
- Tarboton, D. G. and C. H. Luce, 1996, Utah Energy Balance Snow Accumulation and Melt Model (UEB), Computer model technical description and users guide, Utah Water Research Laboratory and USDA Forest Service Intermountain Research Station, Logan, UT, 64 pp.
- Thapa, K., 1993. Estimation of snowmelt runoff in Himalayan catchments incorporating remote sensing data, *IAHS Publications-Publications of the International Association of Hydrological Sciences* 218, 69-74.
- Thayyen, R. and J. Gergan, 2010. Role of glaciers in watershed hydrology: a preliminary study of a "Himalayan catchment", *The Cryosphere* 4(1), 115-128.
- Uhlenbrook, S. and C. Leibundgut, 2002. Process-oriented catchment modelling and multiple-response validation, *Hydrol. Process.* 16(2), 423-440.
- Uhlenbrook, S., S. Roser and N. Tilch, 2004. Development of a distributed, but conceptual catchment model to represent hydrological processes adequately in the meso scale, *J. Hydrol.* 291, 278-296.
- Van Der Knijff, J., J. Younis and A. De Roo, 2010. LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation, *Int. J. Geogr. Inform. Syst.* 24(2), 189-212.
- Van Deursen, W. and C. Wesseling, 1992. The PCRASTER package, Department of Physical Geography, University of Utrecht, Utrecht, The Netherlands
- Van Deursen, W. and C. Wesseling, 1996. PCRaster version 2 manual, Department of Physical Geography, Utrecht University, Utrecht, The Netherlands
- Wesselung, C. G., D. J. KARSSENBERG, P. A. Burrough and W. DEURSEN, 1996. Integrating dynamic environmental models in GIS: the development of a Dynamic Modelling language, *Trans. in GIS* 1(1), 40-48.

Zhang, X., G. Hörmann and N. Fohrer, 2009. Hydrologic comparison between a lowland catchment (Kielstau, Germany) and a mountainous catchment (XitaoXi, China) using KIDS model in PCRaster, *Adv. Geosci.* 21, 125-130.

## CHAPTER 2

# INTEGRATION OF AN ENERGY BALANCE SNOWMELT MODEL INTO AN OPEN SOURCE MODELING FRAMEWORK<sup>1</sup>

### **Abstract**

This paper presents a data model for organizing the inputs and outputs of an energy balance snowmelt model (the Utah Energy Balance Model, UEB) that provides a foundation for its integration into the EPA BASINS modeling framework and enables its coupling with other hydrologic models in this system. Having UEB as a BASINS component has facilitated its coupling with the Geospatial Streamflow Forecast Model (GeoSFM) to compute the melting of glaciers and subsequent streamflow in the Himalayas. The data model uses a combination of structured text and network Common Data Form (netCDF) files to represent parameters, geographical, time series, and gridded space-time data. We describe the design and structure of this data model, integration methodology of UEB and GeoSFM and illustrate the effectiveness of the resulting coupled models for the computation of surface water input and streamflow for a glaciated watershed in Nepal Himalayas.

### **2.1. Introduction**

Snow and Glaciers provide significant contributions to streamflow in the Hindu Kush Himalayan (HKH) region. Managing water resources and protecting against flooding requires an ability to model streamflow that is driven by precipitation and

---

<sup>1</sup>Coauthored by Avirup Sen Gupta, David G. Tarboton, Paul Hummel, Molly E. Brown and Shahid Habib

snowfall, snow and glacier melt, and hydrologic processes involved in runoff generation. This motivates the need to couple snow/glacier and streamflow models. In this study, we combined two independent models (1) the Utah Energy Balance (UEB) Snowmelt Model (Tarboton et al., 1995; Tarboton and Luce, 1996; Mahat and Tarboton, 2012) and (2) the United State Geological Survey (USGS) Geospatial Streamflow Forecast Model (GeoSFM) (Asante et al., 2008) to simulate streamflow using snow- and glacier-melt information. This was done within the US Environmental Protection Agency (EPA) Better Assessment Science Integrating point and Nonpoint Sources (BASINS) modeling system (Kinerson et al., 2009) modeling framework as part of a NASA project to provide integrated modeling capability to take advantage of NASA data products in this region. To facilitate the integration of UEB into the EPA BASINS so that it could be coupled and made interoperable with other models in BASINS and so it could be applied over a grid, the input/output data model of UEB was redesigned and generalized to have greater flexibility in its text inputs/outputs and to exploit the capability of network Common Data Form (netCDF) to hold gridded space (two dimensions) and space-time (three dimensions) data.

The Himalayan region is one of the world's largest reservoirs of snow and glaciers and is a major freshwater source for 1.4 billion people in Asia (Immerzeel et al., 2010). A recent study by Racoviteanu et al. (2013) shows glaciers contribute over 58% of total annual streamflow in a high altitude Himalayan watershed (Langtang Khola) in the Trishuli basin in Nepal. Glacier-melt models and water balance studies are often complicated by limited access to glacierized area, insufficient data measurement infrastructure, and a lack of financial support (Konz et al., 2007). Yet, some research

advances have been made towards the understanding of glacier melt contributions to streamflow mostly by using temperature index models (Kayastha et al., 2000; 2005; Immerzeel et al., 2010; 2012; Konz et al., 2010). There are, however, challenges associated with these models. Estimates of the glacier-melt contribution to streamflow by different models vary greatly (Racoviteanu et al., 2013), and inconsistent assessment methods and a high degree of uncertainty in modeling glacier change and weather input data make melt estimates inconclusive (Immerzeel et al., 2012). Furthermore, temperature index models are limited in their ability to quantify energy balance processes in complex topography where the interactions between radiation and topography (slope and aspect) play significant roles in snow and glacier melting. Also, in developing countries in South Asia there is limited expertise and access to tools needed to integrate models and translate research knowledge into policy and water resources management decisions. Thus, there is a need for better models as well as a system that can be used in the local institutions involved in water resources management to conduct their own local analyses. The work here that incorporates a physically based model into EPA BASINS strives to address these needs.

The Utah Energy Balance (UEB) Snowmelt Model (Tarboton et al., 1995; Mahat and Tarboton, 2012; 2013; Mahat et al., 2013) is a physically based model that simulates surface melt by estimating surface energy fluxes from weather inputs. UEB parameterizes the snowpack using lumped (depth averaged) state variables so as to avoid having to model the complex processes that occur within a snowpack. Physical differences between bulk (depth averaged) properties and the surface properties, which are important for calculating surface energy exchanges, are captured by modeling diurnally forced heat

flow at the surface using the so-called force-restore parameterization where there is a forcing term related to the difference between surface and depth temperature and a restore term related to the temporal gradient of surface temperature (Deardorff, 1978; Luce and Tarboton, 2010). Prior to this study, UEB did not have a glacier melt component and was configured as a point model that could be applied at a single site with homogeneous (or average) terrain characteristics and weather input data.

GeoSFM is a spatially semi-distributed, physically based streamflow simulation program. GeoSFM incorporates terrain analysis tools, hydrologic simulation routines, and tools for time series post-processing. GeoSFM's ability to simulate streamflow using remotely sensed data to compensate for lack of ground-based observations makes it especially suitable for the data scarce HKH region. However, GeoSFM does not have an explicit way to represent snow and glaciers or to estimate melt from these sources, which limits its use to only low-elevation, non-glaciated, rain-fed downstream watersheds in HKH. This limitation can be overcome by the addition of a model to estimate snow and glacier melt. UEB was chosen for this purpose because the model was open source and provided a relatively simple, transferable, physically-based approach to the quantification of snowmelt. We envision that the coupling between GeoSFM and UEB in an integrated framework will enhance streamflow prediction information in glaciated watersheds in the HKH region and elsewhere.

UEB and GeoSFM were developed independently with no prior means of interoperability. Prior to this study, these programs were incompatible in terms of data format, scale, and could not be easily coupled. The US EPA's BASINS software is an open source framework that facilitates the integration of programs. BASINS has as plug-

in components a number of hydrologic programs as well as data preparation and results analysis tools (EPA, 1998; Parisi et al., 2003). Given the target application in the HKH region, the plug-in framework and free and open source aspects of BASINS made it an ideal choice for coupling UEB and GeoSFM. UEB and GeoSFM were each configured as BASINS plug-ins to accomplish their integration. The resulting operational software application is referred to as HIMALA BASINS. This software application retains all of the original BASINS 4.0 analysis, data downloading functionalities, preexisting plug-in and adds UEB and GeoSFM as two new plug-ins. To overcome the scarcity of available input data in the HKH region, we developed a capability to derive inputs for UEB from the Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Bosilovich, 2008; Rienecker et al., 2011) and Southern Asia Daily Rainfall Estimate (RFE2) (Xie et al., 2002; Xie and Arkin, 1996) products provided by NASA and NOAA, respectively. The integrated system can run UEB to simulate total surface water input, and then run GeoSFM to simulate streamflow, perform sensitivity analysis, and enable parameter calibration using streamflow data at a gauged outlet.

The contributions of this paper are (1) the development of a data model to structure the input and output of UEB to enable its extension from a point-based research model to a spatially distributed operational model capable of running over a watershed to simulate snow and glacier ablation; and (2) the integration of UEB into the EPA BASINS simulation environment for coupling with the GeoSFM model. "Data model" here refers to the specific data structures used to represent UEB parameters, site variables, state variables and dynamic inputs and outputs. The data model is distinct from the computer code or program often also referred to as a "model." In the remainder of this paper, we

use the word “program” to refer to the computer program and “model” to refer to the data model to keep these concepts distinct.

Section 2.2 provides a brief background on UEB, GeoSFM, and BASINS. In Section 2.3, we describe UEB and geoSFM plugins developed within EPA BASINS and the data model developed to extend UEB from a point program to a grid program to facilitate its integration into BASINS and coupling with GeoSFM. In Section 2.4, we use the coupled models to simulate snow and glacier melt and the generation of streamflow from the glaciated Langtang Khola watershed in Nepal. This paper concludes with a summary of research contributions and ideas for consideration in future work.

## **2.2. Background**

### **2.2.1. Utah Energy Balance Snowmelt Program**

The Utah Energy Balance (UEB) program was originally written in Fortran 77 to produce snowmelt outputs at a point (such as a weather station) driven by the inputs at that location. UEB (Tarboton et al., 1995; You, 2004; Mahat and Tarboton, 2012) is physically-based and tracks point energy and mass balances to quantify snow accumulation and melt. To enhance the capability of UEB to quantify snow processes in a forest covered area, Mahat and Tarboton (2012) developed a two-stream radiation transfer process that explicitly accounts for canopy scattering, absorption, and refection. They also added the capability to represent turbulent exchanges within and above a forest canopy (Mahat et al., 2013) and to represent snow interception (Mahat and Tarboton, 2013). The most recent version of UEB with forest canopy additions has four state variables: surface snow water equivalent,  $W_s$  (m); surface snow and substrate energy

content,  $U_s$  ( $\text{kJ m}^{-2}$ ); the dimensionless age of the snow surface; and the snow water equivalent of canopy intercepted snow,  $W_c$  (m). The dimensionless age of snow is a surface condition variable defined by Dickinson et al. (1993) to parameterize the sensitivity of the decrease of albedo over time due to the increase in snow grain size and accumulation of dirt to environmental conditions such as temperature. The UEB program is driven by inputs of air temperature, precipitation, wind speed, relative humidity, and incoming shortwave and longwave radiation at time steps (i.e., typically less than 6 hours) sufficient to resolve the diurnal cycle.

UEB was initially tested by simulating snow water equivalent (SWE) and snow melt at point locations such as experimental weather stations in the Western United States. These included stations in the Reynolds Creek Experimental Watershed, Boise, Idaho; Utah State University drainage and evapotranspiration research farm; and the TW Daniels Experimental Forest (TWDEF) in Logan, Utah (Tarboton and Luce, 1996; Luce and Tarboton, 2010; Mahat and Tarboton, 2012). UEB has also successfully been used in several snow studies such as estimating snowmelt and sublimation in the high Atlas mountains in Morocco (Schulz and De Jong, 2004), examining climate change impacts in Sacramento/San Joaquin watershed in the Western US (Knowles and Cayan, 2002), and assessing the importance of meteorological variables in snowmelt processes (Raleigh et al., 2008).

To account for glacier melt, which is primarily driven by surface energy exchanges, we extended the representations of surface energy balance fluxes in the UEB snowmelt program to include the melting and generation of surface water from a glacier surface. Maps of glaciers and their surface albedo, determined from remote sensing, are

used as inputs. Seasonal snow may accumulate and melt on a surface of glacier ice or bare ground. When seasonal snow cover on top of glacier ice completely melts the surface energy balance switches to the glacier substrate surface and may generate melt from the glacier ice. The surface, or substrate, beneath seasonal snow may be input as one of four types: (1) Ground/non-glacier, (2) Clean glacier ice, (3) Debris covered glacier ice and (4) Glacier accumulation zone. For grid cells with ground/non-glacier substrate, the program computes energy content and simulates snow melt as it did prior to the implementation of glacier melting. For grid cells in the glacier accumulation zone, the program is bypassed, as all precipitation is presumed to add to glacier accumulation. For clean or debris-covered glacier, the program tracks seasonal snow accumulation and ablation, but when seasonal snow water equivalent is zero, the energy balance at the surface is used to calculate the melting of glacier ice, which becomes a component of the surface water input. The difference in functionality between a debris-covered and clean glacier ice surface is the substrate albedo, which quantifies the fraction of incoming solar radiation reflected from the surface when the substrate is exposed. This albedo is provided as a separate input layer. Debris covered glacier albedo is generally lower than that of clean glacier ice, resulting in larger energy inputs and higher melt rates.

Additional details on the glacier implementation are reported in Chapter 4.

### **2.2.2. Geospatial Streamflow Model**

The USGS Geospatial Streamflow Model (GeoSFM) is a semi-distributed, physically-based hydrologic program developed to monitor flood hazards and provide early warning across Africa and other data scarce regions around the globe (Asante et al., 2008). It was originally configured to operate as an extension within ArcView 3.2 to take

advantage of existing spatial analysis algorithms. GeoSFM uses a wide range of input data, including digital elevation model (DEM), topographical, land cover, and soil data, daily estimates of precipitation, and potential evapotranspiration to predict daily streamflow at in-situ gauge stations. GeoSFM is designed to use remotely sensed and satellite data to compensate for the data scarcity in data sparse parts of the world. The program has six components: (1) the terrain analysis module, (2) the parameter estimation module, (3) the data preprocessing module, (4) the water balance module, (5) the flow routing module, and (6) the post-processing module.

GeoSFM has been used in the HKH region for the past ten years to take advantage of satellite-derived precipitation data products to help with water management and flood prediction (Shrestha et al., 2008; Shrestha, 2011). The International Centre for Integrated Mountain Development (ICIMOD), a HKH regional organization for water resources management, used GeoSFM for streamflow simulation in large watersheds such as Brahmaputra (Pervez et al., 2008) and Bagmati rivers (Shrestha et al., 2008) using satellite-based rainfall data.

GeoSFM's terrain analysis module uses DEM data to delineate subwatersheds and stream networks to establish the connectivity among various subwatersheds and to compute topographical parameters such as slope and aspect. Using United Nations Food and Agriculture Organization (FAO) and United Nations Educational, Scientific and Cultural Organization (UNESCO) soil data (Batjes, 1997), the parameter estimation module estimates soil parameters such as water holding capacity, hydrological active depth, texture, and saturated hydraulic conductivity. The land cover data is used to compute an impervious area grid and vegetation roughness of each subwatershed. The

soil and land cover data are used together to determine Soil Conservation Service (SCS) runoff curve numbers.

The GeoSFM data preprocessing module converts ground-based and satellite data into a common ASCII format where each subwatershed contains a single data point at a particular time step. Input satellite rainfall estimates are spatially distributed raster grids, and these need to be aggregated on a subwatershed scale. The water balance module separates rainfall into various components of the hydrologic cycle, such as evapotranspiration, interflow, baseflow, groundwater, and surface runoff at each time step. The flow routing module aggregates runoff at each subwatershed at the subwatershed outlet and then routes the flow through the stream network. The post-processing module calculates and displays a statistical summary of streamflow and enables output visualization.

### **2.2.3. BASINS**

Better Assessment Science Integrating Point and Nonpoint Sources (BASINS), is an open source, freely-distributable, GIS-enabled tool for environmental analysis and monitoring. Developed and supported by the US Environmental Protection Agency, it was designed to facilitate modeling of environmental systems and analysis of management alternatives by integrating environmental and geospatial data and programs. The primary BASINS interface is enabled by an open source Geographic Information System (GIS) called MapWindow (Ames et al., 2007; 2008; Kinerson et al., 2009). BASINS provides a database management system that enables seamless interaction between data and programs and provides capabilities to analyze, organize, and display spatial data as maps, tables, or graphics. It encompasses a suite of hydrological and water

quality programs and also provides a framework for adding preexisting programs as additional software components or plug-ins. Notable examples of BASINS plug-ins include EPA's storm water management model (SWMM) (Rossman, 2010), Soil Water Assessment Tool (SWAT) (Luzio et al., 2002; Arnold and Fohrer, 2005; Zhao et al., 2009), and Hydrologic Simulation Programme-Fortran (HSPF) (Singh et al., 2005). Simpler analysis tools, such as importing time series data, finding maximum and minimum from a time series data, etc., can also be added as BASINS scripts. This functionality allows a user to customize and extend the system without the complexity of writing a plug-in. BASINS also supports preexisting tools such as Climate Assessment Tools (CAT), data visualization, and explore and query tools.

### **2.3. Integrated HIMALA BASINS System**

#### **2.3.1. UEB and GeoSFM Plug-ins to EPA BASINS**

The BASINS plug-in interface was used to facilitate the integration between UEB and GeoSFM (Figure 2.1). Both UEB and GeoSFM retained their independent program executables. UEB FORTRAN code was modified to support the generalized input/output data model described below. The source code of the GeoSFM rainfall-runoff model was not modified at all. These executables each operate on their own separate input and output files. A plug-in module was written for each that provided a graphical user interface and supported some necessary data manipulations, such as aggregation of time series from 3-hourly to daily using functionality of BASINS.

UEB and GeoSFM plug-ins were developed in Visual Basic using Visual Studio .NET 2010. The plug-ins were implemented by developing classes that define a small

number of key properties and methods that allow them to be identified by the BASINS framework. The BASINS plug-in interface provides a link between the plug-ins and BASINS plug-in manager, enabling plugin-ins to interact with all preexisting functionality of BASINS. Here, the GeoSFM plug-in took advantage of existing potential evaporation functionality. The UEB and GeoSFM plug-ins were programmed to include several user interface forms that provided them with a Graphical User Interface (GUI) for the users to interact with the underlying programs. The HIMALA BASINS GUI (Figure 2.2) is separate from the BASINS parent window and provides a seamless integration between two programs. This new GUI window contains a series of tabs to perform UEB and GeoSFM simulation tasks, including terrain analysis, GeoSFM parameter estimation, UEB program set up and run, evaporation estimation, soil water balance and streamflow computation, sensitivity analysis, calibration, and output visualization. For UEB program setup, we created an additional window for creating and editing UEB control files. This gives the users a file browsing option to select files and modify UEB's start and end dates, time resolution, and the parameters. Also, the BASINS visualization tool is able to display UEB and GeoSFM outputs, which improves both the post-processing capabilities and the ability to test the linking of the models. Thus, the HIMALA BASINS GUI improves usability of both UEB and GeoSFM. The GUI coupling methodology allowed for plug-in testing to focus on pre- and post-processing functionality without altering the code of the underlying executable, recognizing that prior to this development of the plug-in UEB code was modified to generalize its input/output data model.

Subwatersheds or topographical variables such as slope and aspect are inputs for both UEB and GeoSFM. GeoSFM contains a terrain analysis module that produces these raster files; however, UEB does not. Traditionally, UEB requires a separate GIS-enabled watershed delineation tool to accomplish these tasks. In HIMALA BASINS, GeoSFM's terrain analysis task is performed at the beginning to create these variables, thus eliminating the need for a separate GIS tool.

### **2.3.2. UEB Snowmelt Program Data Model**

A netCDF and text file-based input/output file schema (Figure 2.3) was developed to enable UEB to run as a point-based as well as a distributed program to predict snow and glacier melt at a point or over a watershed. UEB input variables are classified into three groups: (1) parameters that are spatially constant and constant in time, (2) site variables that are constant in time but may be spatially variable and (3) dynamic inputs that vary in time. Here site refers to the area or footprint for which the model is run and is either a grid cell or a point location. Site variables include quantities such as slope, aspect and vegetation that characterize each point location where UEB is being applied. On the other hand, dynamic inputs are typically the weather input variables such as precipitation, air temperature, wind, and humidity, but they may also include quantities such as albedo. Generally the dynamic inputs are also spatially variable, although the program is configured to allow these to be either spatially variable or spatially constant to accommodate configurations where no information about spatial variability exists or the variability is at a scale larger than the UEB domain so that, for efficiency, they may be represented as spatially constant. All UEB outputs are assumed to be dynamic (i.e., both

space and time varying). Descriptions of UEB parameters, site, time-varying input and output variables are provided in Appendixes A through D.

UEB's input/output data model (Figure 2.3) starts with an overall control file that specifies whether the model is run at a point location or over a grid and gives names of other input files. Input files comprise the watershed file, parameter file, a file specifying site and initial conditions, and a file specifying time varying inputs (top of input files box, Figure 2.3). The output control file specifies the variables to be output, at a point, on a grid or aggregated (output files box, Figure 2.3).

NetCDF was chosen as a standard input-output multidimensional data format for UEB. A detailed description of netCDF file formats is available in the netCDF user's guide (Rew et al., 1993). Figure 2.4 is a simple illustration of the organization of a netCDF file used in UEB. Two dimensional (2-D) netCDF files are used to store variables that are constant in time, while three dimensional netCDF files are used to store variables that change in time. A watershed file, for example, is a 2-D netCDF file with an X and Y coordinate system. In this case, we can ignore the "time" dimension shown in Figure 2.4. Three different subwatersheds are shown in three different colors and each of these watersheds is represented by a unique integer number in the netCDF file. The white grid cells in Figure 2.4 indicate the area outside of the spatial domain of interest and may be represented with a missing value following the Climate and Forecast (CF) convention (Eaton et al., 2003). Streamlines are also shown for illustration purposes (Figure 2.4); however, these are not represented within the file. Other 2-D netCDF files will contain slope, aspect, canopy coverage and other site variables. In a three-dimensional (3-D) netCDF file used in UEB, time is the third dimension. For a single grid, specified by X

and Y coordinates, time series data is stored on the time-axis and each value is associated with a specific time instant. All grid files input to and output from UEB need to have the same cell size and extent.

UEB parameters are stored in a text file tagged by their identifiers or “flags.” Some parameters serve to “switch on” functionality within UEB. For example, the “ireadalb” flag indicates to the program whether surface albedo will be provided as an input time series or needs to be calculated by the UEB. If the user sets “ireadalb” to 1, the UEB turns on its albedo calculation module and calculates albedo, and if is set to 0, this function remains turned off and reads albedo as a user-provided input from the list of time varying input variables.

The site initial file contains a list of site variables and initial conditions of UEB’s state variables, information pertinent to their spatial variability (whether spatially varying or constant) and their value or the locations of files that store their values. Site variables and initial conditions may be spatially constant or variable. The strategy for these is to have a text file that gives either the value for the variable, if it is spatially constant, or the corresponding spatial grid (netCDF) file if it is a spatial variable (Figure 2.3). The strategy for dynamic input variables (such as temperature) is to have a text file that gives either the value for the variable for each time step or the name of a 3-D netCDF file holding the value for each time step (Figure 2.3).

The time varying input file specifies the time step and start and end times and has a flag for each variable that specifies whether or not the variable is spatially and/or temporally variable. The file then gives the name of the corresponding text file containing the time series for variables that are spatially constant, or points to a list of

netCDF files containing space-time data for spatially and temporally varying inputs. To accommodate input data that may have occasional missing data or may have time steps that are a bit irregular or do not coincide with model time steps exactly, the program adopts the approach that any time-varying value persists until another later time value is available. This means that if any input time series data value is missing, either for a single grid cell or for the entire spatial domain, the previous time step value will be repeated. Solar radiation is the only exception to this rule. If a solar radiation value at a particular times step is unavailable, the index value becomes -9999 and the program calculates solar radiation using slope, aspect, date and time, and temperature. Note that this persistence approach is not intended to be a comprehensive solution to the missing data problem. Rather it is a fail over to ensure that the program runs. We feel that it is better for gap filling to be implemented separately as a part of input data preparation, rather than in the program, as this provides greater flexibility to accommodate improvements and best practices in gap filling methods.

In the UEB program, a significant part of the run time is reading and writing data from and to the netCDF files. We implemented the following techniques to enhance efficiency by reducing this data access time.

- a) All the dimensions, variables, and attributes are defined before writing the data in output netCDF files.
- b) UEB runs through a nested loop of three dimensions where the sequence of outer to inner loop is longitude, latitude, and time. Hence, time is the “most rapidly varying dimension.” Array declarations in space- and time-varying dynamic input netCDF files

for any variable begin with the least rapidly varying dimension and end with most rapidly varying dimension.

- c) For a single grid cell, the entire time series is read at the beginning instead of reading a single value at each time step.

The program runs separately for each grid cell and stores outputs in netCDF files. No computations are performed for the grids outside of the study domain or predefined watershed. Most modelers may only require total surface water input and its components (i.e., snow melt, glacier melt and input from rain) aggregated over a watershed, while others may also be interested in obtaining outputs in a gridded format. Both options are available; however, the latter option comes at the expense of higher writing time. UEB produces 67 output variables, but reporting all of these will significantly increase the program execution time and the space occupied by the output files. The output control file allows the user to specify the space-time outputs to be written as netCDF files. Similarly, it specifies which aggregate variables to output into text files. A third option allows the user to specify points where detailed point output is required. This option is particularly useful for analyzing results at an individual point, better understanding the system, and identifying potential sources of errors.

#### **2.4. Langtang Khola Watershed Case Study**

UEB and GeoSFM were used in HIMALA BASINS to evaluate the contribution of glacier and snow melt to total surface water input in Langtang Khola, a Himalayan watershed in Nepal. This medium sized watershed, with an area of 360 km<sup>2</sup>, is situated approximately 100 km north of Kathmandu. It is a high-altitude basin with elevation ranging from 3700 m to 7184 m and an average elevation of 5176 m. Fifty seven percent

of the watershed is non-glaciated, 35% of the area is occupied by clean glacier ice, and 8% of area is covered by debris-cover glacier ice (Figure 2.5). One discharge measuring station is located at the outlet of the watershed (elevation 3800, Figure 2.5).

#### 2.4.1. Data Sources

Figure 2.6 presents the workflow used to obtain and prepare the data needed to run UEB and GeoSFM in the Langtang Khola watershed. Data sources are shown on the left, data preprocessing activities are illustrated in the mid-section, and the integrated BASINS framework that runs UEB and GeoSFM is shown in the right section.

UEB requires climate, hydrologic, land cover, and topographic data. These data can be collected from any sources that meet the following criteria: (1) the data must cover the spatial domain of study area or watershed, (2) time varying dynamic variables must be collected-produced at time resolutions sufficient to resolve the diurnal cycle, and (3) elevation and land-cover data must reasonably capture the spatial variability of the region. The digital elevation model (DEM) was obtained from the Space Shuttle Radar and Topography Mission (SRTM) (Gesch et al., 2006) at 3 arc second (or  $3/3600^{\circ}$ ) cell resolution. UEB works with spatially distributed raster data in any projection. However, evaluation of slope and aspect and the delineation of watersheds requires data in projected coordinates, so this data was projected. GeoSFM's terrain analysis tool was then used to create subwatersheds from the DEM. These were used to aggregate distributed UEB outputs for input to GeoSFM that had subwatersheds as its modeling element. Vegetation related variables such as canopy coverage, leaf area index, canopy height and canopy structure were estimated from various land cover datasets, such as

Global Land Cover 2000 (GLC-2000) and 500-m MODIS global land cover (Cohen et al., 2003).

Glacier outline maps for Langtang Khola watershed were derived from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images from October 2003 orthorectified products (Racoviteanu et al., 2013). The scenes acquired at the end of the seasonal snow melting season (for minimal snow cover) with high contrast over the glaciers and minimal cloud cover are preferred for glacier mapping. Glacier maps for other areas of the world can be obtained from the Global Land Ice Measurements from Space (GLIMS) project (Bishop et al., 2004) that has a glacier inventory storing critical information about the extent and rates of change of the world's estimated 160,000 glaciers. Substrate albedo was derived from the atmospherically corrected surface reflectance product from ASTER.

Time-varying input data such as temperature, relative humidity, wind speed, and longwave and shortwave radiation were derived from MERRA. MERRA is a near-real-time global climate reanalysis product developed at NASA and is available from 1979 to the present (Rienecker et al., 2011; Suarez et al., 2008; Wu et al., 2002; Lucchesi, 2012). Hourly temperature, wind speed, and relative humidity are reported at a height of 2 m from the ground at a spatial resolution of  $0.67^\circ$  longitude  $\times 0.5^\circ$  latitude, and incoming shortwave and longwave radiation are available at three-hourly time steps at a coarser resolution of  $1.25^\circ \times 1.0^\circ$  (Lucchesi, 2012).

Precipitation data were derived from the near-real time Southern Asia Daily Rainfall estimate (RFE2) data product available for the Southern Asian domain ( $70^\circ$ - $110^\circ$  East,  $5^\circ$ - $35^\circ$  North) at a spatial resolution of  $0.1^\circ$  by  $0.1^\circ$  beginning on May 01, 2001.

RFE2 data are constructed using four observational input data sources: approximately 280 Global Telecommunications System (GTS) stations, geostationary infrared cloud top temperature fields, polar orbiting satellite precipitation estimate data from SSM/I, and AMSU-B microwave sensors (Xie et al., 2002).

To capture the local variability of snow and glacier melt, UEB was run at much finer spatial resolution than that of the available data. We developed a data preprocessing tool referred to as MERRA Spatial Downscaling for Hydrology (MSDH), to downscale temperature, precipitation, wind speed, relative humidity, and shortwave and longwave radiation from the coarse resolution at which they are available to the scale of the projected SRTM DEM. MSDH was written in the R statistical and programming environment (R Development Core Team, 2009). It takes advantage of several R libraries, including “raster,” “ncdf,” and “rgdal” and other independently developed netCDF data manipulation and analysis toolkits such as netCDF Operators (NCO) (Zender, 2008) and Climate Data Operators (CDO) (Schulzweida et al., 2006) for raster and netCDF manipulations and analyses. Based on local topography, MERRA and RFE2 data are adjusted using micrometeorological parameterizations for how the variables vary with elevation, slope, aspect, curvature and cloudiness (Liston and Elder, 2006)+ to obtain local meteorological variables required to run UEB at a watershed scale. For each month, we stored a complete set of time-varying dynamic input variables in a single netCDF file, resulting in a small number of files to manage. This approach also has the advantage that additional simulation months can be added without having to edit existing files.

### 2.4.2. Model Setup

The Langtang Khola watershed was divided into eighteen subwatersheds and each watershed was marked by a unique ID number. MERRA temperature data was downloaded for the South Asian region (Figure 2.7 (a)) and the four grid cells spanning the Langtang Khola watershed (Figure 2.7 (b)) were downscaled to obtain gridded temperature at the scale of the DEM (Figure 2.7(c)). This involved using R's raster library projection transformation capability to transform the data to the DEM's Lambert Azimuthal Equal Area projection and clip it to the extent of the DEM. This raster layer contains resampled temperature data, and its spatial domain, number of rows, and number of columns are exactly the same as the DEM. This consistency is important since UEB requires the same spatial domain and the same number of rows and columns in all netCDF files. Next, the monthly lapse rate and difference between MERRA elevation and DEM elevation were used to adjust temperature at each grid cell to the elevation of the DEM. This procedure was repeated for all the time steps. Other UEB inputs such as incoming shortwave radiation, wind speed were also downscaled to the DEM spatial scale using physically based elevation, slope and aspect micrometeorology adjustments as mentioned above.

Initial conditions of UEB's state variables, comprised of snow water equivalent, the internal energy of the snowpack and top layer of soil, and the dimensionless age of the snow surface were unknown. These were initialized using a one year spin up period. At the beginning of this period, state variables were set to zero. Errors due to this assumption diminish with time as the model adjusts to the driving inputs. Therefore, results from the spin up period, October 2002 to September 2003, were discarded and

only the output from the one year period, 10/1/2003 to 9/30/2004, are reported in the results section and shown in Figure 2.8.

UEB total surface water input, which is the combination of rainfall, snowmelt and glacier melt, was used as input to GeoSFM, and hydrologic losses (i.e., evaporation, change in storage) were modeled by GeoSFM. Mean daily potential evapotranspiration (PET) for each subwatershed was estimated using the Hamon method (Hamon, 1961), the implementation of which was available as functionality within BASINS. This was provided as an input to GeoSFM. GeoSFM was configured to use its nonlinear soil moisture accounting routine that combines the SCS runoff curve number method with the Green-Ampt equation (Green and Ampt, 1911) to compute runoff, interflow and baseflow. Parameters such as curve number and soil hydraulic properties were based on the soil data described above. GeoSFM was configured to use the Muskingum-Cunge channel routing method (Ponce and Yevjevich, 1978). Sensitivity analysis showed that the model was most sensitive to soil water holding capacity (mm), total soil depth (cm) and baseflow reservoir residence time (days). Therefore, the model was manually calibrated by adjusting these three parameters within their plausible ranges as suggested by Asante et al. (2008) to match the observed streamflow data at the hydrologic station.

#### **2.4.3. Case Study Results**

Outputs such as surface water input from snow melt, glacier melt and rain were aggregated over the watershed. Since snow- and glacier-melt and rain are the inputs for streamflow generation, these are called “surface water input components” and the sum of these is referred as “total surface water input” as shown in equation (1). Annual cumulative total surface water input (SWIT) was 1.2 m (Figure 2.8(b)) and is comprised

of glacier melt (SWIGM), snowmelt (SWISM), and rain (SWIR), each contributing about 60 %, 31 % and 9%, respectively. In rainfall-runoff applications GeoSFM is driven by input rainfall. Here GeoSFM was driven by UEB-derived total surface water input that is the sum of glacier melt, snowmelt and rain aggregated over each of eighteen subwatersheds.

$$\text{SWIT} = \text{SWISM} + \text{SWIGM} + \text{SWIR} \quad (2.1)$$

GeoSFM computes hydrologic losses (i.e., evaporation, change in storage), and, after calibration, the results indicate a daily mean bias of -6 % between the observed (at Kyangin hydrologic station) and simulated streamflow. Correlation and Nash-Sutcliffe Efficiency were 0.92 and 0.82, respectively, indicating satisfactory agreement between the simulated and observed streamflow. The root mean square error (RMSE) was 0.47 mm/day (about 25 % of the observed daily mean streamflow) also demonstrating that the integrated system is capable of capturing the variability of observed streamflow reasonably well. Both Figures 2.8 (a) and 2.8 (b) show that during the winter, simulated streamflow captures both the seasonal pattern and magnitude of the observed streamflow. During early- to mid-monsoon (i.e., May through June) streamflow was slightly overestimated, and during the late-monsoon (i.e., September) streamflow was underestimated by GeoSFM driven by UEB. Overall, yearly aggregated simulated streamflow depth (total yearly streamflow divided by the watershed area) was 0.65 m compared to 0.69 m of observed streamflow. This is a relatively small discrepancy given that the integrated system involves multiple models and is driven entirely by reanalysis and remote sensing input data.

## 2.5. Discussion

In the process of selecting netCDF as a standard input-out data format, we investigated a series of data formats for UEB's input-output data storage. First, we investigated the feasibility of using time series of tabulated data for each grid cell within a watershed. This approach was driven by the simplicity of tabulated text files and readily available tools to read and write text files. Establishing a standardized way to map between the grid cells with associated text files was the only major task. Although it was easy to implement, we noticed the following shortcomings: (1) storage of tabulated text data was inefficient, (2) adding additional time steps in files required modifying all of the existing files, and (3) output post-processing, such as aggregation, was difficult as it required opening all of the text files associated with the grid cells that fall within a subwatershed.

Since NASA two-dimensional or multidimensional remote sensing and reanalysis climate data products are the primary data sources for this study, we also considered two-dimensional raster formats such as ASCII and GeoTIFF. Plain ASCII files are the simplest and the most portable gridded data format (Ravazzani, 2013), and sequential ASCII files are often used in traditional scientific data management (Treinish, 1999). However, like tabulated time series text files, ASCII files are inefficient to read and write. On the other hand, 2-dimensional gridded binary formats are efficient, but each file can accommodate only a single time step for a single variable. UEB runs cell-by-cell and requires all of the input variables for all time steps at each cell, so this format requires opening all of the gridded binary files in order to get data for even a single point. Our

preliminary implementation to evaluate this option found the large number of files involved to be difficult to manage and inefficient.

GeoTIFF raster files can accommodate multiple time steps as bands. However there were two drawbacks to working with GeoTIFF: (1) GeoTIFF does not provide an easy way to hold the time associated with each band if bands represent time steps, and (2) we could not locate an open source Fortran 90 library for GeoTIFF. Options were thus either to develop GeoTIFF functionality for FORTRAN, or to use mixed language programming, neither of which seemed attractive.

In the data framework we developed, we chose netCDF for data management for a number of reasons. NetCDF is a very common data format used by the oceanographic and atmospheric scientific community for creating, managing, storing, and distributing scientific data. A single netCDF file can accommodate multiple variables and can store 2 GB of data without implementing netCDF's large file support (<https://www.unidata.ucar.edu/software/netcdf/docs/netcdf/Large-File-Support.html>). Also, during various regional conferences and training meetings, we determined that researchers from South Asia are familiar with the netCDF file format. Thus, due to netCDF's wide application, availability of a FORTRAN netCDF library (<http://www.unidata.ucar.edu/>), and the wide range of software tools available for pre-and post-processing, we chose netCDF as the standard input-output multidimensional data format for UEB.

The HIMALA BASINS system can run UEB and GeoSFM independently or as an integrated system, which provides additional flexibility to users. Because of its distributed nature, UEB is often expensive in input data preprocessing and run time.

Similarly, GeoSFM calibration is a time consuming task. We were able to take advantage of this flexibility by first running UEB coupled to GeoSFM once and saving the UEB outputs/GeoSFM inputs. Then calibration of GeoSFM involved repeated iteration of GeoSFM without the need to rerun UEB.

The UEB program's run time varies significantly depending on the number of variables output in gridded netCDF format. This is also a consideration in the management of computer disk space. The case study involved a space domain consisting of  $319 \times 330$  grid cells, and the program was run at 3-hour time steps for one year. This resulted in a simulation of over 300 million data values ( $319 \times 330 \times 2920$  time steps) for each variable. In netCDF format, these data values occupied 1.3 GB of disk space, and writing all 67 variables for each grid cell in the aforementioned watershed required over 87 GB of disk space. The capability for a user to choose only a small subset of the output variables is useful to reduce the model runtime and manage the disk space efficiently.

Despite the many benefits of integrating UEB with GeoSFM, UEB suffers from efficiency issues. Running UEB in the Langtang Khola watershed with the grid above for a 10 year period takes about 30 hours to complete on a common commodity workstation (Dell Optiplex 780, with Intel Q9650 processor @ 3.0 GhZ and 8 GB RAM). Our experience indicates that UEB runs faster if the data is stored in a smaller number of large netCDF files rather than a large number of small netCDF files. Therefore, enabling "large netCDF" files (i.e., larger than 2 GB) may increase the program's efficiency. We also envision that netCDF-4, which implements HDF-5 for parallel access, may improve efficiency. However, the netCDF library that we are currently using (netCDF 3.6.1 FORTRAN version) does not have any of these implementations.

## 2.6. Conclusion

In this study, the UEB snowmelt program has been extended and refactored so that it can be applied over a grid using netCDF files to manage input-output workflow and data storage. This refactoring provides the capability to drive UEB using NASA remote sensing and earth science data products. Integration of GeoSFM and UEB in the BASINS framework contributes to an improved hydrologic information system that enhances the usability and applicability of UEB and GeoSFM programs.

Although the example application of the system at Langtang Khola watershed only simulates one year of streamflow at the hydrologic station, it shows UEB's capability to estimate snowmelt and glacier melt, snow accumulation and GeoSFM's capability to simulate streamflow at a satisfactory level using UEB-derived melt information. Watershed-scale total surface water input aggregation indicates the importance of the glacier melt (i.e., 60 % of total surface water input is generated from glacier melt) for Langtang Khola. By coupling UEB and GeoSFM, a better understanding of the overall contribution of ice and snow melt to streamflow has been obtained in this region with sparse data and limited observation of glacier dynamics.

The data model presented here, in the context of UEB has broad generality. It is common for models to have parameters that represent time and space invariant properties in the processes involved. It is also common for models to have inputs that quantify the spatial properties of the modeling domain. These are site variables. The data model provides the flexibility for these to be set as spatially constant or spatially variable. They would be set as spatially variable where explicitly representing their variability is deemed important, and spatially constant where their variability is deemed unimportant at the

scale of the model, or is perhaps unknown. Spatially constant site variables can be read once quite efficiently from text files, while the two dimensional netCDF format proved effective for the spatially variable site variables. The third category of input variable may be dynamically varying in space and time. The model accommodates these being spatially constant or spatially variable, with the constant case an option for variables whose scale of spatial variability is larger than the modeling domain, or unknown, or deemed unimportant and neglected for efficiency. Spatially constant dynamic inputs can be represented using text files while spatially variable dynamic inputs are efficiently represented as netCDF files. The flexibility afforded by this approach provides the modeler with latitude to accommodate varying degrees of availability of input information. The approach taken here is an example of a general path for extending a point based program into a spatially distributed program over a grid in a way that enables its coupling with other programs.

## References

- Ames, D. P., C. Michaelis, A. Anselmo, L. Chen and H. Dunsford, 2008. MapWindow GIS, in Encyclopedia of GIS, Springer, p. 633-634.
- Ames, D. P., C. Michaelis and T. Dunsford, 2007. Introducing the MapWindow GIS project, OSGeo Journal 2(1), 1-5 pp.
- Arnold, J. G. and N. Fohrer, 2005. SWAT2000: current capabilities and research opportunities in applied watershed modelling, Hydrological Processes 19(3), 563-572.
- Asante, K. O., G. A. Artan, S. Pervez, C. Bandaragoda and J. P. Verdin, 2008, Technical manual for the Geospatial Stream Flow Model (GeoSFM), U. S. Geological Survey.
- Batjes, N., 1997. A world dataset of derived soil properties by FAO-UNESCO soil unit for global modelling, Soil Use and Management 13(1), 9-16.

- Bishop, M. P., J. A. Olsenholter, J. F. Shroder, R. G. Barry, B. H. Raup, A. B. Bush, L. Copland, J. L. Dwyer, A. G. Fountain and W. Haeberli, 2004. Global Land Ice Measurements from Space (GLIMS): Remote sensing and GIS investigations of the Earth's cryosphere, *Geocarto International* 19(2), 57-84.
- Bosilovich, M., 2008. NASA's modern era retrospective-analysis for research and applications: integrating Earth observations, *EarthZine*, E-Zine article
- Cohen, W. B., T. K. Maiersperger, Z. Yang, S. T. Gower, D. P. Turner, W. D. Ritts, M. Berterretche and S. W. Running, 2003. Comparisons of land cover and LAI estimates derived from ETM+ and MODIS for four sites in North America: a quality assessment of 2000/2001 provisional MODIS products, *Remote Sensing of Environment* 88(3), 233-255.
- Deardorff, J. W., 1978. Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation, *Journal of Geophysical Research* 83(C4), 1889-1903.
- Dickinson, R. E., A. Henderson-Sellers and P. J. Kennedy, 1993. Biosphere-Atmosphere Transfer Scheme (BATS) Version 1e as Coupled to the NCAR Community Climate Model, National Center for Atmospheric Research, Boulder, CO, p. 71.
- Eaton, B., J. Gregory, B. Drach, K. Taylor, S. Hankin, J. Caron, R. Signell, P. Bentley, G. Rappa and H. Höck, 2003. NetCDF Climate and Forecast (CF) Metadata Conventions, Version 1.0.
- EPA, B. V., 1998. 2.0 User's manual: better assessment science integrating point and nonpoint sources, US Environmental Protection Agency, Washington, DC.
- Gesch, B., J. Muller and T. G. Farr, 2006. The shuttle radar topography mission-Data validation and applications, *Photogrammetric Engineering and Remote Sensing* 72(3), 233.
- Green, W. H. and G. Ampt, 1911. Studies of Soil Physics. Part I - The flow of air and water through soils, *The Journal of Agricultural Science* 4, 1-24.
- Hamon, R. W., 1961. Estimating potential evapotranspiration, *Proceedings of the American Society of Civil Engineers, Journal of the Hydraulic Division*, 87, p. 107-120.
- Immerzeel, W. W., L. Van Beek, M. Konz, A. Shrestha and M. Bierkens, 2012. Hydrological response to climate change in a glacierized catchment in the Himalayas, *Climatic Change* 110(3-4), 721-736.
- Immerzeel, W. W., L. P. van Beek and M. F. Bierkens, 2010. Climate change will affect the Asian water towers, *Science* 328(5984), 1382-1385.

- Kayastha, R. B., Y. Ageta and K. Fujita, 2005. Use of positive degree-day methods for calculating snow and ice melting and discharge in glacierized basins in the Langtang Valley, Climate and Hydrology of Mountain Areas, 7-14 pp.
- Kayastha, R. B., Y. Takeuchi, M. Nakawo and Y. Ageta, 2000. Practical prediction of ice melting beneath various thickness of debris cover on Khumbu Glacier, Nepal, using a positive degree-day factor, IAHS PUBLICATION 71-82
- Kinerson, R. S., J. L. Kittle and P. B. Duda, 2009. BASINS: better assessment science integrating point and nonpoint sources, in Decision Support Systems for Risk-Based Management of Contaminated Sites, Springer, Joplin, MO, p. 1-24.
- Knowles, N. and D. R. Cayan, 2002. Potential effects of global warming on the Sacramento/San Joaquin watershed and the San Francisco estuary, Geophysical Research Letters 29(18), 1-4.
- Konz, M., D. Finger, C. Buergi, S. Normand, W. Immerzeel, J. Merz, A. Giriraj and P. Burlando, 2010. Calibration of a distributed hydrological model for simulations of remote glacierised Himalayan catchments using MODIS snow cover data, Global Change: Facing Risks and Threats to Water Resources, p. 465-473.
- Konz, M., S. Uhlenbrook, L. Braun, A. Shrestha and S. Demuth, 2007. Implementation of a process-based catchment model in a poorly gauged, highly glacierized Himalayan headwater, Hydrology and Earth System Sciences 11(4), 1323-1339.
- Liston, G. E. and K. Elder, 2006. A meteorological distribution system for high-resolution terrestrial modeling (MicroMet), Journal of Hydrometeorology 7(2), 217-234.
- Lucchesi, R., 2012. File Specification for MERRA Products. GMAO Office Note No. 1 (Version 2.3), 82 pp.
- Luce, C. H. and D. G. Tarboton, 2010. Evaluation of alternative formulae for calculation of surface temperature in snowmelt models using frequency analysis of temperature observations, Hydrology and Earth System Sciences 14(3), 535-543.
- Luzio, M., R. Srinivasan and J. G. Arnold, 2002. Integration of watershed tools and SWAT model into BASINS1, JAWRA Journal of the American Water Resources Association 38(4), 1127-1141.
- Mahat, V. and D. G. Tarboton, 2012. Canopy radiation transmission for an energy balance snowmelt model, Water Resources Research 48, W01534.
- Mahat, V. and D. G. Tarboton, 2013. Representation of canopy snow interception, unloading and melt in a parsimonious snowmelt model, Hydrological Processes in press

- Mahat, V., D. G. Tarboton and N. P. Molotch, 2013. Testing above and below canopy representations of turbulent fluxes in an energy balance snowmelt model, *Water Resources Research* 49(2), 1107-1122.
- Parisi, D., M. Taquino, S. MICHAEL GRICE and D. A. Gill, 2003. Promoting environmental democracy using GIS as a means to integrate community into the EPA-BASINS approach, *Society & Natural Resources* 16(3), 205-219.
- Pervez, S., G. Artan and M. Shrestha, 2008. Stream flow simulation from remotely sensed data: Brahmaputra River, AGU Fall Meeting 1, p. 1043.
- Ponce, V. M. and V. Yevjevich, 1978. Muskingum-Cunge method with variable parameters, *Journal of the Hydraulics Division* 104(12), 1663-1667.
- R Development Core Team, 2009, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, <http://www.R-project.org>.
- Racoviteanu, A. E., R. Armstrong and M. W. Williams, 2013. Evaluation of an ice ablation model to estimate the contribution of melting glacier ice to annual discharge in the Nepal Himalaya, *Water Resources Research* 49(9), 5117-5133.
- Raleigh, M., F. Lott and J. Lundquist, 2008. Most Critical Surface Meteorological Measurements for Modeling Distributed Snowmelt in the Sierra Nevada, California, AGU Fall Meeting Abstracts, 1, p. 0971.
- Ravazzani, G., 2013. MOSAICO, a library for raster based hydrological applications, *Computers & Geosciences* 51, 1-6.
- Rew, R., G. Davis, S. Emmerson, H. Davies and E. Hartnett, 1993. NetCDF User's Guide, Unidata Program Center, Boulder, CO.
- Rienecker, M. M., M. J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D. Schubert, L. Takacs and G.-K. Kim, 2011. MERRA: NASA's modern-era retrospective analysis for research and applications, *Journal of Climate* 24(14), 3624-3648.
- Rossman, L. A., 2010. Storm water management model user's manual, version 5.0, National Risk Management Research Laboratory, Office of Research and Development, US Environmental Protection Agency.
- Schulz, O. and C. De Jong, 2004. Snowmelt and sublimation: field experiments and modelling in the High Atlas Mountains of Morocco, *Hydrology and Earth System Sciences Discussions* 8(6), 1076-1089.
- Schulzweida, U., L. Kornblueh and R. Quast, 2006, CDO User's Guide, Climate Data Operators, Version 1.0.1, Max-Planck-Institute for Meteorology, Hamburg, Germany.

- Shrestha, M., G. Artan, S. Bajracharya and R. Sharma, 2008. Using satellite-based rainfall estimates for streamflow modelling: Bagmati Basin, J. Flood Risk Manage. 1(2), 89-99.
- Shrestha, M. S., 2011, Bias-adjustment of satellite-based rainfall estimates over the central Himalayas of Nepal for flood prediction, PhD Thesis, Kyoto University, Japan.
- Singh, J., H. V. Knapp, J. Arnold and M. Demissie, 2005. Hydrological modeling of the iroquois river watershed using HSPF and SWAT1, Journal of the American Water Resources Association 41, 343–360.
- Suarez, M. J., M. Rienecker, R. Todling, J. Bacmeister, L. Takacs, H. Liu, W. Gu, M. Sienkiewicz, R. Koster and R. Gelaro, 2008. The GEOS-5 Data Assimilation System- Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0,
- Tarboton, D. G., T. G. Chowdhury and T. H. Jackson, 1995. A Spatially Distributed Energy Balance Snowmelt Model, in Biogeochemistry of Seasonally Snow-Covered Catchments (Proceedings of a Boulder Symposium, July 1995), Edited by K. A. Tonnessen, M. W. Williams and M. Tranter, IAHS Publ. no. 228, Wallingford, p. 141-155.
- Tarboton, D. G. and C. H. Luce, 1996, Utah Energy Balance Snow Accumulation and Melt Model (UEB), Computer model technical description and users guide, Utah Water Research Laboratory and USDA Forest Service Intermountain Research Station, Logan, UT, 64 pp.
- Treinish, L., 1999. A function-based data model for visualization, Proceedings of the IEEE Visualization 1999 Conference Late Breaking Hot Topics, p. 73-76.
- Wu, W.-S., R. J. Purser and D. F. Parrish, 2002. Three-dimensional variational analysis with spatially inhomogeneous covariances, Monthly Weather Review 130(12), 2905-2916.
- Xie, P. and P. A. Arkin, 1996. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions, Journal of Climate 9(4), 840-858.
- Xie, P., Y. Yarosh, T. Love, J. Janowiak and P. A. Arkin, 2002. A real-time daily precipitation analysis over South Asia, 16th conference on hydrology, American Meteorological Society, Orlando, FL, p. 3.15.
- You, J., 2004, Snow Hydrology: The parameterization of subgrid processes within a physically based snow energy and mass balance model, PhD Thesis, Civil and Environmental Engineering, Logan, UT.

- Zender, C. S., 2008. Analysis of self-describing gridded geoscience data with netCDF Operators (NCO), *Environmental Modelling & Software* 23(10), 1338-1342.
- Zhao, L., K. S. K. Subramanian, C. X. Song, S. Kumar, V. Merwade, I. Chaubey, R. S. Govindaraju and M. Sayeed, 2009. A web based interface for SWAT modeling on the TeraGrid, 5th International SWAT Conference, Boulder, CO, p.

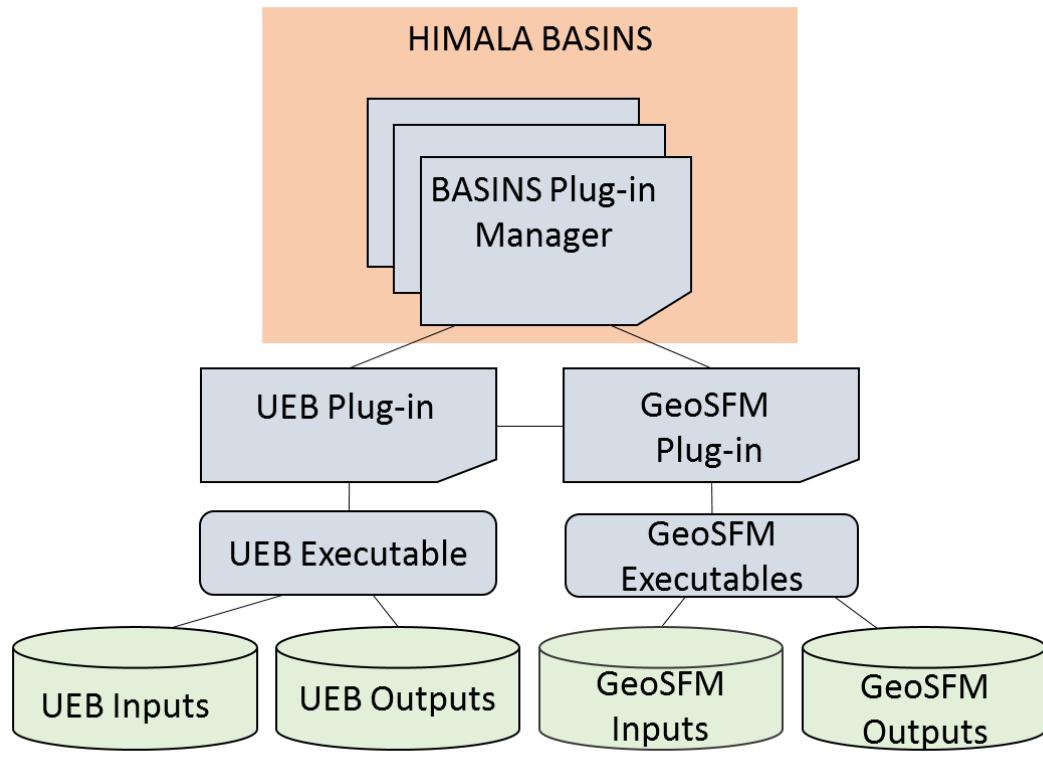


Figure 2.1. Plug-in architecture of HIMALA BASINS.

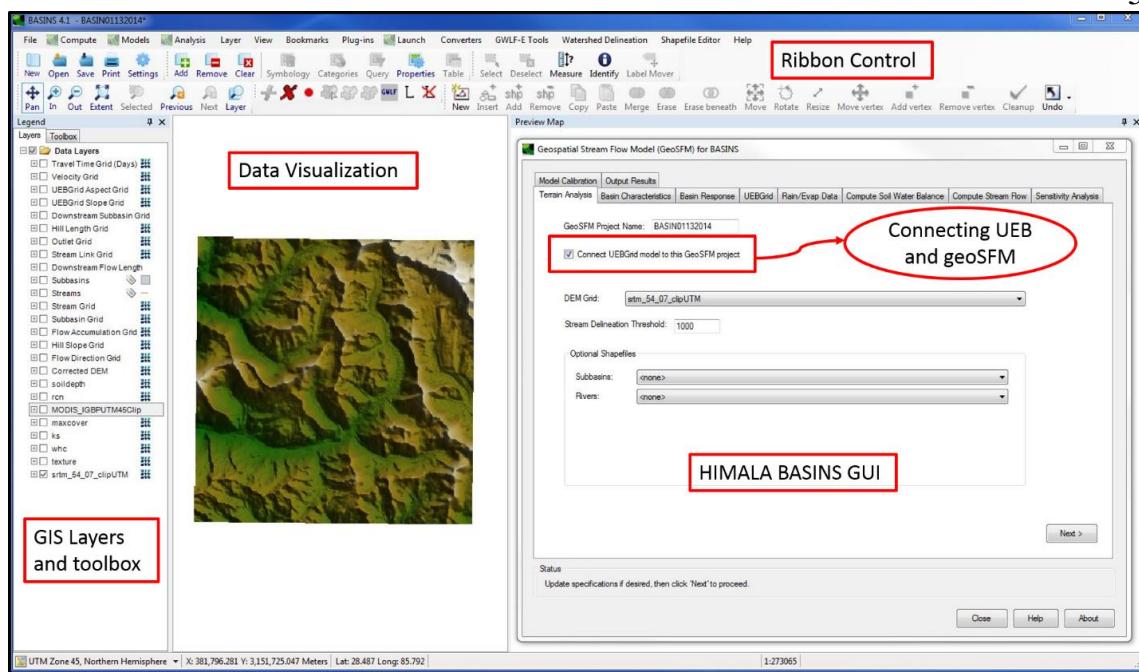


Figure 2.2. HIMALA BASINS Graphical User Interface

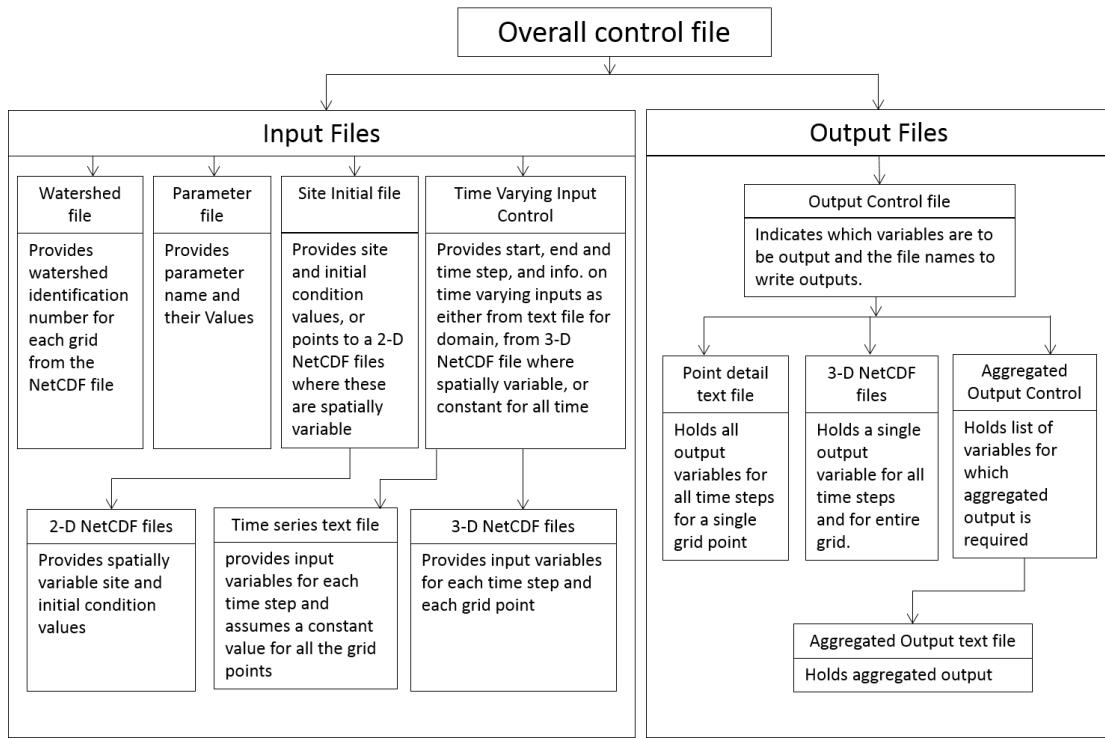


Figure 2.3. Organization of input and output files in the UEB program

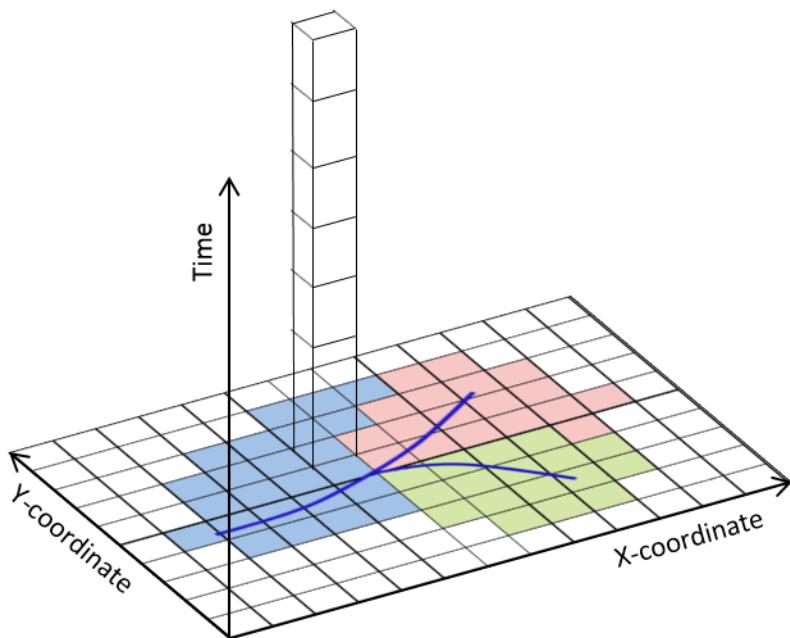


Figure 2.4. Representing space and time in a netCDF file

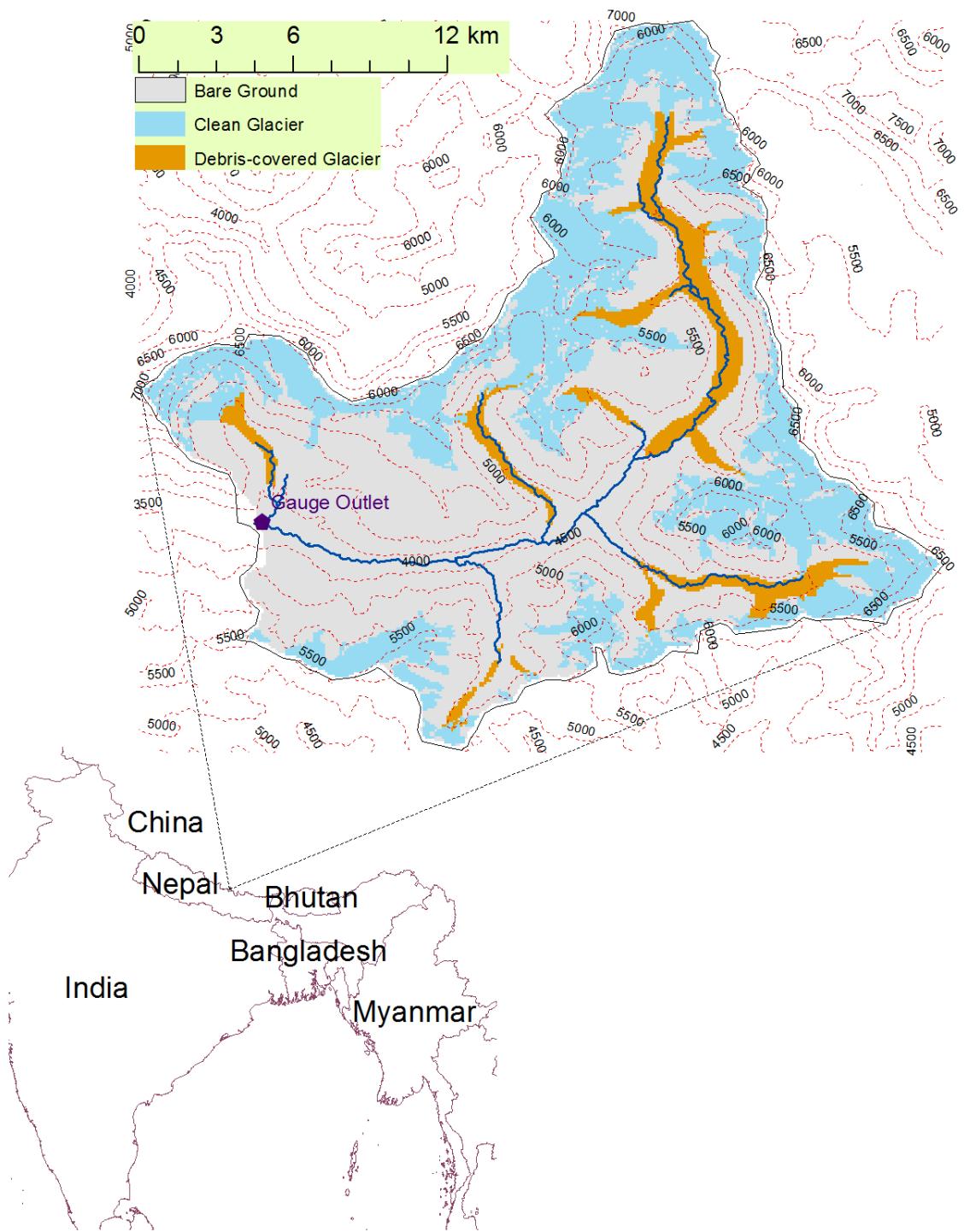


Figure 2.5. Langtang Khola Watershed.

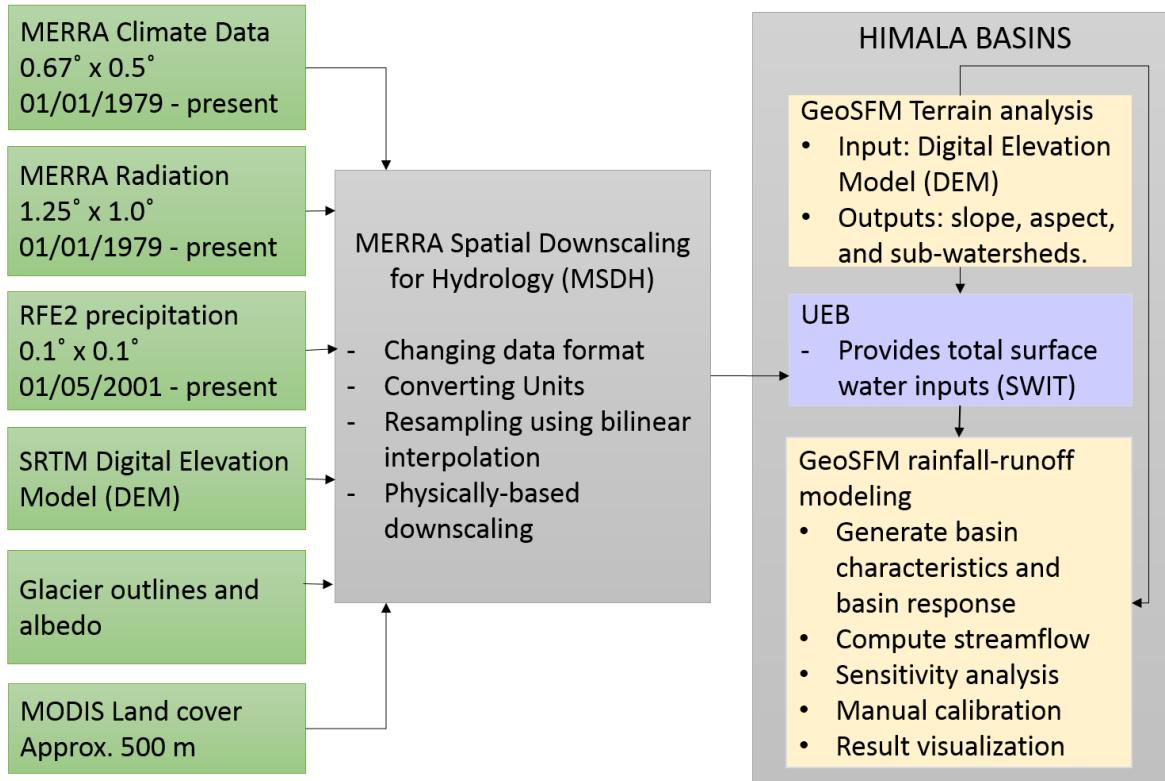


Figure 2.6. Data workflow of input preprocessing and coupled UEB and GeoSFM modeling system in EPA BASINS for Langtang Khola case study.

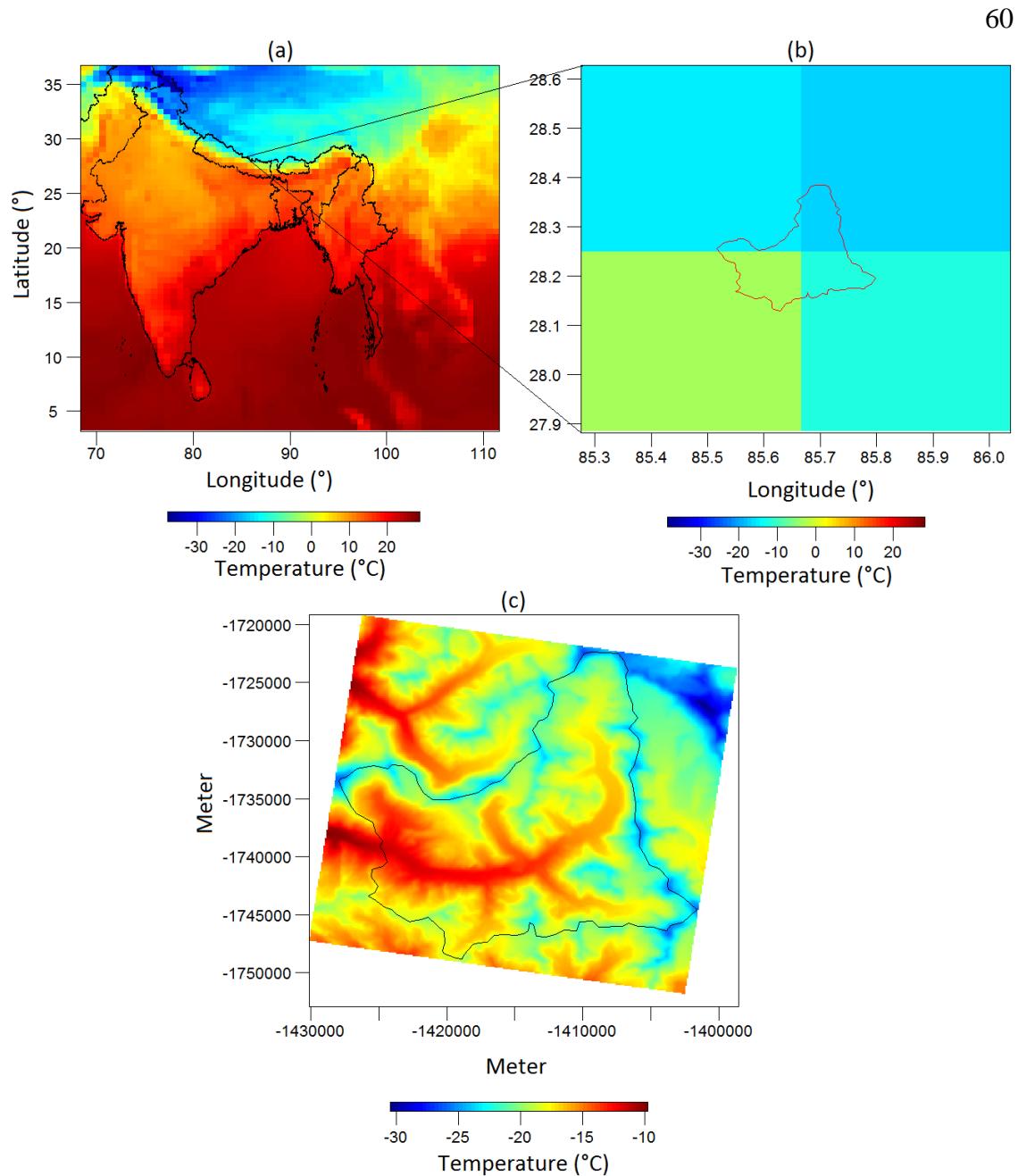


Figure 2.7. Downscaling of MERRA temperature (°C) for Langtang Khola watershed at 3:00 am on January 1, 2003. (a) South Asian region temperature; (b) MERRA grid cells spanning Langtang Khola watershed (c) Downscaled temperature projected to Lambert Azimuthal Equal Area DEM grid.

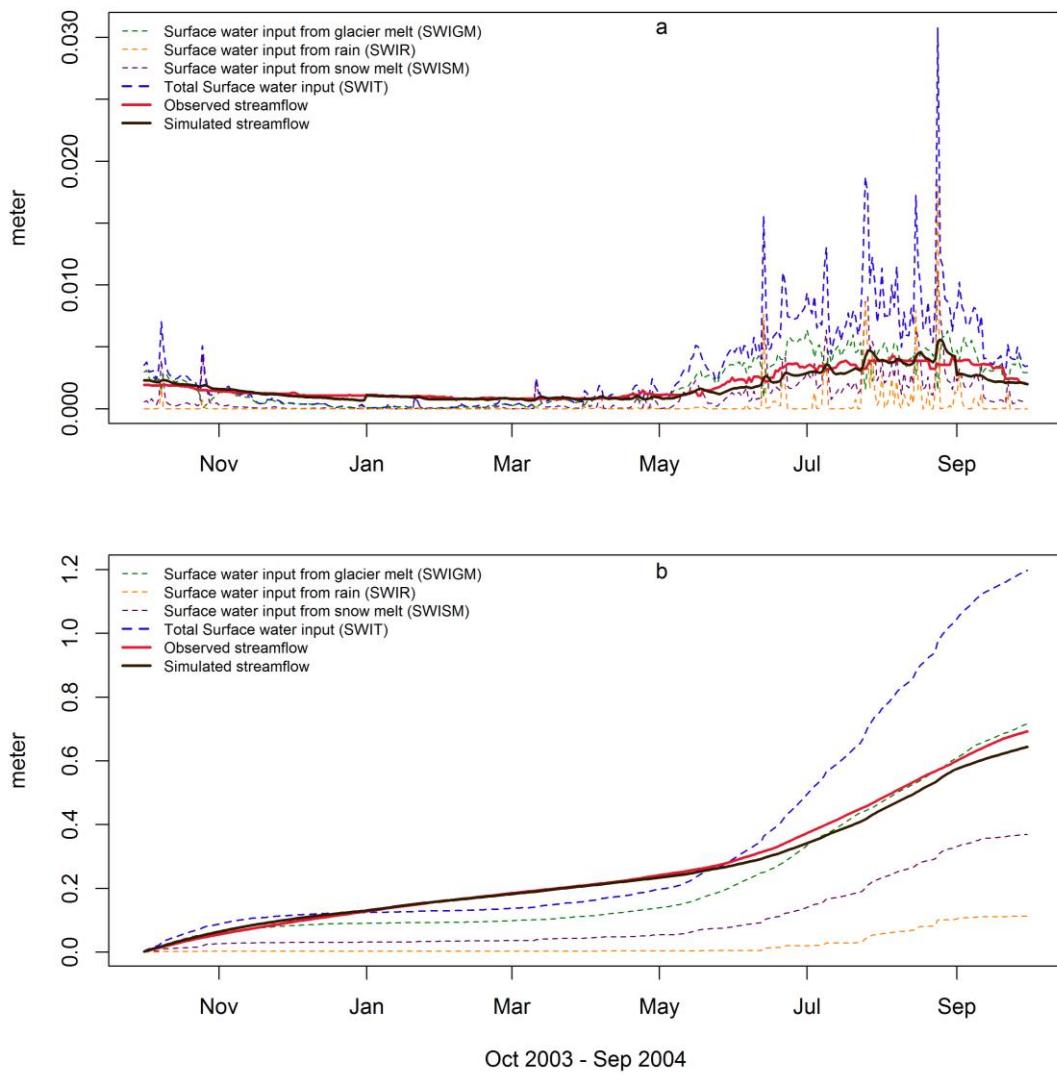


Figure 2.8. (a) Daily and (b) cumulative time series of UEB simulated surface water input components, streamflow measured at the Kyangin station and streamflow simulated by GeoSFM.

## CHAPTER 3

MSDH V1.0: A MERRA SURFACE WEATHER AND RADIATION SPATIAL  
DOWNSCALING TOOL FOR HYDROLOGICAL APPLICATION<sup>1</sup>

**Abstract**

There is growing interest in generating high resolution climate data to simulate catchment responses to different climate conditions. Constructing a reliable meteorological dataset in complex terrain is particularly challenging due to limited observational data, limitations on accessibility and high climate variability in areas with extreme topography. In this study, we developed and implemented a quasi-physically-based spatial downscaling tool to generate 3-hourly surfaces of weather variables at a grid scale of ~ 100 m over a watershed with complex terrain from 2/3° longitude by 1/2° latitude, and hourly Modern Era Retrospective-Analysis for Research and Applications (MERRA) climate and radiation data products. The weather variables downscaled were temperature, precipitation, relative humidity, wind speed, and shortwave and longwave radiation and the tool was developed using the R scripting language. First, we bilinearly interpolated MERRA data to the scale of a high resolution Digital Elevation Model (DEM), then, we made topographic adjustments using well-established relationships: precipitation and temperature with elevation; wind with slope, curvature and aspect; atmospheric transmission with air pressure and vapor pressure; and humidity with cloudiness. The application of the software is demonstrated in the 570 km<sup>2</sup> Logan River Watershed in Northern Utah. The downscaled climate variables were compared with daily observations at the Natural Resources Conservation Service (NRCS) USU Doc

---

<sup>1</sup>Coauthored by Avirup Sen Gupta and David G. Tarboton

Daniel SNOWpack TELemetry (SNOTEL) station in the Logan River Watershed during October 2009 to June 2010 where Utah State University (USU) has measurements of radiation, humidity and wind speed that are beyond the standard set of SNOTEL measurements. The daily mean, maximum and minimum temperature and monthly precipitation were also compared at a total of six SNOTEL stations in the Logan River Watershed including the USU Doc Daniel station. A distributed snowmelt model was then applied using the downscaled data to simulate spatial and temporal variability of Snow Water Equivalent (SWE) in the Logan River watershed, Utah. The results showed reasonably good agreement (i.e., average Nash-Sutcliffe efficiency = 0.6) between the SNOTEL observations and the downscaled data and Utah Energy Balance Snowmelt Model-simulated SWE. This work showed that it is possible to obtain the input variables required to drive the UEB model entirely from climate reanalysis data extending its applicability to data scarce regions of the world. The impact of discrepancies in this data on the overall model simulations was quantified and needs to be factored into the use of simulations driven by downscaled results for hydrological modeling and analysis.

### **3.1. Introduction**

High resolution climate data are increasingly used in distributed hydrologic modeling studies to simulate hydrological responses in heterogeneous areas. The outcomes of these studies are critical for water resources management decisions related to agricultural water supply, ecosystem services and hydropower production (Daly, 2006). While computer models in hydrology vary widely in purpose, complexity and spatial-temporal scale, most models require a continuous time-varying climate dataset with minimal missing data at a site or grid point (Jeffrey et al., 2001). Moreover, physically

based energy balance models often require incoming radiation fluxes and wind speed, which are not measured at most climate stations, especially in developing countries. Furthermore, observational climate datasets may not cover the complete time span of the model simulation and may contain systematic and random errors (Jeffrey et al., 2001). Due to limited availability of complete observed datasets, which are required for sophisticated models, research efforts are often restricted to either shorter simulated time spans or simpler, less data-demanding models.

Climate reanalysis datasets are commonly used to complement a limited observational record. Climate reanalysis data is produced by re-analyzing historic observations using a climate model that has unchanging parameters and equations based on known physics. They assimilate measurements of different atmospheric variables (temperature, pressure, precipitation etc.) from many sources to produce spatially complete, gridded meteorological variables at a continental or global scale (Rienecker et al., 2011). Most reanalysis data are also temporally complete during the satellite era (1979 to present) and are typically generated at a resolution (hourly, 3-hourly and 6-hourly) sufficient to capture the diurnal variability (Rienecker et al., 2011). Temporally complete reanalysis data has great appeal to the scientific community (Rienecker et al., 2011) and has proven to be a valuable research tool in meteorology, climatology, hydrology, and ecology. The Modern-Era Retrospective Analysis for Research and Applications (MERRA) is a new generation reanalysis dataset developed by NASA's Global Modeling and Assimilation Office to improve water cycle representation, which is a known problem in other reanalysis datasets such as the European Centre for Medium-Range Weather (Trenberth and Olson, 1988), the NOAA/NCEP (Kanamitsu et al., 2002),

and the Japanese 55-year Reanalysis (Ebita et al., 2011). Unlike observational data, MERRA does not suffer from spatio-temporal discontinuity but does contain uncertainty in precipitation and surface fluxes because of model biases in long term climatology and limitations in reproducing the diurnal cycle. Even with the known limitations, reanalysis data is a valuable resource for obtaining forcing variables to drive hydrological models in data scarce regions such as the Himalayas in South Asia (Xie et al., 2007) and the Blue Nile Basin in Africa (Dile and Srinivasan, 2014).

The spatial resolution of reanalysis data is typically fifty to a few hundred kilometers, similar to the General Circulation model (GCM) scale. While GCMs successfully simulate climatic conditions at the continental and hemispheric spatial scales, their performance decreases when representing regional and local scale dynamics (Carter et al., 1994; Wigley et al., 1990; Xu, 1999). On the other hand, using hydrologic models to simulate river flow requires information about local climate variability. Moreover, improvements in computer technologies over the past few decades has enabled researchers to design distributed hydrologic models capable of running at a very high resolution (30 m to 1 km horizontal grid) (Liston and Elder, 2006). These models also need scale-appropriate weather input data (Liston and Elder, 2006). Thus, methods and tools are needed to produce high resolution downscaled reanalysis and GCM outputs.

Past studies (McMurtrie et al., 1992; Running et al., 1987; Thornton et al., 1997) have shown that daily air temperature, precipitation, humidity, and shortwave radiation are the minimum required variables to accurately simulate the hydrologic conditions. Accordingly, the MTCLIM (Hungerford et al., 1989) and DAYMET (Thornton et al., 2012; Thornton et al., 1997) models were developed to downscale these variables at a

single site or distributed grid over a complex terrain using the observations collected at a station network. Liston and Elder (2006) concluded that a terrestrial model may require wind speed and direction, surface pressure, and longwave radiation in addition to the above-mentioned variables. They subsequently developed the MicroMet model to construct spatially distributed temperature, precipitation, humidity, surface pressure, wind speed and direction, and incoming shortwave and longwave radiation data using relationships between these variables and the surrounding topography (Liston and Hiemstra, 2011). These models all take point observations as inputs. There is a need to adapt the ideas from these models to downscale climate reanalysis data.

In this study, we developed a spatial downscaling tool called MERRA Spatial Downscaling for Hydrology (MSDH) for generating 3-hourly grid surfaces of temperature, precipitation, relative humidity, wind speed, and shortwave and longwave radiation over a complex terrain watershed using MERRA reanalysis and Rain Fall Estimates (RFE2) (Xie et al., 2002) data. MERRA assimilates a vast number of meteorological variables at 72 vertical levels, of which only a limited selection at the earth's surface are relevant to hydrological studies. In this study, we used MERRA temperature, wind speed, specific humidity, pressure, and shortwave radiation. Daily precipitation estimates from RFE2 are also included as an alternative source of precipitation data only for the South Asian region.

This automated spatial downscaling approach, drawing upon ideas from previous models for point observations (MTCLIM, Hungerford et al., 1989; MicroMet, Liston and Elder, 2006; DAYMET, Thornton et al., 2012), only requires a high-resolution digital elevation model (DEM) of the target area or watershed. The choice of DEM resolution is

left to the user based on the watershed area, source of the DEM, availability of computer disk space, resource constraints, and use of the data. Our choice of temporal resolution is largely influenced by the need for the input variables in a physically based energy balance snowmelt model to quantify the diurnal cycle. This is a common requirement in the computation of surface energy balance so we anticipate that this approach has broad applicability. To capture the diurnal pattern of a variable, it is preferable to obtain data at a time resolution of 6 hours or less. In this application, we have chosen a 3-hourly time step that, in our judgment, holds sufficient information to resolve the diurnal cycle but does not increase the data volume to an unmanageable level for a desktop application. The model is capable of producing spatially distributed climate data without requiring any ground-based observations, which, once it has been validated at locations with observations, makes it suitable for data scarce watersheds. However, when observed data is available it can be used to derive location specific topographic adjustment coefficients that improve the quality of the downscaled data.

While developing the tool, we considered the following criteria.

- (1) Given the target application in data scarce remote locations, often in developing countries, the tool should be based on a free and open source software solution.
- (2) The tool should have an easy-to-use graphical user interface to hide internal codes and file-folder complexity and to provide an intuitive visual environment.
- (3) The data should be stored in a standard file format that can be accessed by readily available software tools.
- (4) The computational complexity should be limited so that the software tool can be used on a personal computer (PC).

The downscaling method was implemented in the R scripting language (R Development Core Team, 2009), providing a free and open source platform. All other supporting tools required for running the software, such as Climate Data Operators (CDO) (Schulzweida et al., 2006), GTK+ (Krause, 2007) and netCDF Operators (NCO) (Zender, 2008), are also freely available. A Graphical User Interface (GUI) provides a visual environment for users to interact with the underlying code, requiring only a small set of inputs from the users, such as a directory of input and output files, start and end date, and the extent of the spatial domain. Downscaled data is saved in Climate and Forecast (CF) convention (Eaton et al., 2003) compatible multidimensional Network Common Data Form (netCDF) format (Rew et al., 1993), which can be accessed and visualized in a number of freely available software tools such as ncBrowse, ncview, and Integrated Data Viewer (IDV), in addition to R itself. The spatial downscaling method is relatively light weight and of moderate complexity and can be run on a PC with low performance computing capability. This is a great advantage for applications in developing countries where students, researchers and engineers may not have access to the latest advanced computing facilities.

This work was driven by the need to apply the Utah Energy Balance Snowmelt Model (UEB) to the melting of glaciers in the Himalaya region as reported elsewhere (see Chapter 4). However there is insufficient data there to evaluate and validate the downscaling approaches described here. Instead, the software was evaluated in the mountainous 570 km<sup>2</sup> Logan River watershed in Northern Utah at a 120-m grid resolution. This is a spatial scale judged to be fine enough to quantify the explicit effects of elevation, slope and aspect on snowmelt model inputs. The downscaled climate

variables were validated using daily observations at the USU Doc Daniel station (NRCS, 2014) in the Logan River watershed from October 2009 to June 2010. The downscaled data were then used to drive an energy balance snowmelt model (Utah Energy Balance, UEB) to simulate the spatial and temporal variability of Snow Water Equivalent (SWE) for one water year (October 01, 2009 to Sep 30, 2010). MSDH downscals all of the input variables required by UEB. In the past, the UEB model has been successfully used for snow accumulation and melt computation in the state of Utah (Luce and Tarboton, 2010; Mahat and Tarboton, 2013). Comparison between the measured and simulated SWE at six sites in this study shows that the UEB model was able to simulate snow accumulation and melt with reasonable success.

Section 3.2 provides a brief background on existing climate data interpolation and downscaling techniques. In Section 3.3, we describe the downscaling techniques we developed and adopted in MSDH and provide a brief background on the Utah Energy Balance (UEB) Snowmelt Model. We then give software implementation details in Section 3.4. These include the implementation of the downscaling algorithm in R, strategies for efficient storage of output data in NetCDF, and the graphical user interface. In Section 3.5, we use the MSDH software tool to produce climate data for Logan River Watershed in Utah, simulate snow accumulation and melt using UEB, and compare the results with the observations. This paper concludes with a summary of contributions and limitations of the research.

### **3.2. Background**

Downscaling is the process of adjusting information at a coarse scale in space or time to a finer scale for use in a model at the finer scale. Since MERRA climate variables

and radiation fluxes sufficiently capture the diurnal cycle required for energy balance hydrologic models, temporal downscaling was not necessary. We therefore reviewed the previous studies dedicated to spatial downscaling, also often treated as “spatial interpolation methods,” to guide our downscaling methodology and implementation of software for production of high resolution climate data from lower resolution gridded climate reanalysis products or Global Circulation Model (GCM) outputs. In the context of GCMs, two common downscaling approaches are used to translate information from the GCM grid scale to smaller local scales: (1) dynamic downscaling and (2) statistical downscaling (Wilby et al., 2002). Dynamic downscaling can be achieved by nesting a high-resolution Regional Climate Model (RCM) within a course resolution GCM model (Wilby et al., 2002; Xu, 1999) or by using a variable resolution GCM, in which the area of interest is represented with high-resolution spatial grid cells (Xu, 1999). In the RCM approach, the GCM provides time-varying physical boundary conditions to the RCM models to simulate the climatic quantities at a resolution of 20-50 km (Wilby et al., 2002). The RCM is capable of better capturing the small scale atmospheric or orographic effects than the GCM (Wilby et al., 2002); however, similar to GCMs, RCMs are computationally expensive (Fowler et al., 2007; Wilby et al., 2002; Xu, 1999) and require substantial knowledge to apply (Benestad, 2004). Statistical spatial downscaling techniques involve interpolation and extrapolation approaches such as inverse distance, kriging, and smoothing splines (Lo et al., 2011). These methods are quite reliable for flat

terrain watersheds, but they fail to produce reliable meteorological data in more complex terrains (Lo et al., 2011).

There are also physically based meteorological downscaling techniques that consider the physical properties of the quantity being used. These techniques distribute point-measured information over a modeling domain or downscale from either regional or global information to a distributed local modeling domain. The mountain climate simulation model, MTCLIM (Hungerford et al., 1989) provides algorithms for extrapolating meteorological forcing variables such as daily air temperature, precipitation, solar radiation, and relative humidity at a location of interest by using point measurements at weather stations (Zimmermann and Roberts, 2001). This approach constructs climate data at any elevation by adjusting the observed data collected at lower elevation climate stations. Meteorological variables are adjusted for elevation difference between the weather station and target site, slope, aspect, east-west orientation and leaf area index (LAI). The main objective of developing MTCLIM was to provide inputs to an ecological model for simulating plant growth in mountainous regions where observed data is sparse. DAYMET extends MTCLIM algorithms to produce gridded daily meteorological variables by interpolating observations at multiple stations across larger regions (Thornton et al., 2012; Thornton et al., 1997; Zimmermann and Roberts, 2001).

The Parameter-elevation Regressions on Independent Slopes Model (PRISM) is another widely used approach to produce high-resolution climate data in North America. PRISM generates gridded estimates of annual, monthly, and event-based climatic variables such as maximum and minimum temperature, precipitation, and humidity using observational data at point locations, DEM, other spatial data, and local information

(Daly et al., 1994; 1997; 2000). Variables at a target site are calculated by using linear regression, with regression weighting factors estimated based on elevation, terrain aspect, coastal proximity, and vertical air mass layering (Hunter and Meentemeyer, 2005).

MicroMet, a quasi-physically based spatial and temporal downscaling process, is capable of producing high-resolution (30- to 1000-m) climate data over a wide range of landscapes (Liston and Elder, 2006). Using ground-based observations of air temperature, precipitation, relative humidity, wind speed, and direction within or near the area of interest, MicroMet is capable of constructing high-resolution gridded air temperature, precipitation, pressure, relative humidity, wind speed and direction, and shortwave and longwave radiation estimates. Spatial interpolations are performed using the Barnes objective analysis scheme and adjustments are made for elevation, topography, and cloudiness (Liston and Elder, 2006).

In this study, we developed physically-based spatial downscaling techniques designed to take gridded reanalysis data (specifically MERRA and RFE2) as input. In contrast to approaches designed to take point observations as input, such as MicroMet (Liston and Elder, 2006), this requires different approaches to interpolation and elevation adjustments that are based on the coarse scale elevation of the gridded reanalysis data. Unlike statistical techniques, physically-based approaches do not require long-term point observations, although observations can be used for estimating local precipitation and temperature adjustment coefficients (lapse rates). Physically based downscaling is also less complex and computationally expensive than dynamic downscaling using a regional atmospheric model, which would not be feasible for the scales for most hydrologic studies (i.e. 30 m to 1 km). Considering these factors, we used physically-based

meteorological downscaling methods, drawing heavily on the ideas in MicroMet (Liston and Elder, 2006).

### **3.3. Data and Methods**

#### **3.3.1. Data Sources for Downscaling**

MERRA is a recent near-real-time global climate reanalysis product developed at NASA during the satellite era (1979 to present) and derived from the Goddard Earth Observing System version 5 (GEOS-5), NASA general circulation model (Rienecker et al., 2011; Suarez et al., 2008) and National Centers for Environmental Prediction (NCEP) Gridpoint Statistical Interpolation (GSI) analysis (Wu et al., 2002). Hourly temperature, wind speed, and relative humidity are available at a spatial resolution of  $2/3^{\circ}$  longitude by  $1/2^{\circ}$  latitude, and 3-hourly incoming shortwave and longwave radiation are available at a coarser resolution of  $1.0^{\circ}$  by  $1.25^{\circ}$  (Lucchesi, 2012). MERRA data can be accessed and downloaded via NASA's Goddard Earth Science Data and Information Services Center website, where users are provided an option to choose a range of dates, vertical level, spatial bounding box, list of data products, and variables. Daily data are available in standard HDF5 (Folk et al., 1999) and netCDF formats. The MERRA variables used in this study are listed in table 3.1.

Rainfall Estimation (RFE2) daily total precipitation estimates are constructed using four observational input data sources: approximately 280 GTS stations, geostationary infrared cloud top temperature fields, polar orbiting satellite precipitation estimate data from SSM/I, and AMSU-B microwave sensors (Xie et al., 2002). Near real-time daily rainfall estimations are available for the Southern Asian domain ( $70^{\circ}$ - $110^{\circ}$

East; 5°-35° North) at a spatial resolution of 0.1° by 0.1° beginning on May 01, 2001. The data are available in gridded binary format via NOAA's National Centers for Environmental Protection (NCEP) ftp website (<ftp://ftp.cpc.ncep.noaa.gov/fews/S.Asia/>). The merits of precipitation data from these two datasets (i.e. MERRA and RFE2) are demonstrated by Shrestha et al. (2008) and Reichle et al. (2011), respectively.

### **3.3.2. Downscaling Methodology**

Variables listed in table 3.1 correspond to the elevations that are specified by geopotential height in MERRA's NASA general circulation model (Rienecker et al., 2011). Geopotential height is reported at the same spatial resolution with the corresponding variable and is constant over time. MSDH downscaling techniques follow a four-step procedure: (1) perform temporal averaging of MERRA hourly temperature, precipitation, eastward and northward wind speed, specific humidity, and pressure in three hour blocks, (2) project MERRA data to the spatial projection of the DEM, (3) distribute the MERRA elevations and meteorological variables from MERRA resolution to DEM resolution using bilinear interpolation and (4) use known relationships between climate variables with elevation, slope, aspect, curvature and cloudiness to parameterize the effect of topography. RFE2 precipitation is reported as total daily values; thus, to obtain 3-hourly precipitation, we distribute the total daily precipitation equally, assuming uniform precipitation throughout the day. In the third step, bilinear interpolation at any point on the DEM grid uses four surrounding MERRA grid cells to apply linear interpolation. The values at any grid cell of a bilinearly interpolated surface at DEM resolution always remains within the minimum and maximum range of surrounding MERRA grid points, resulting in smoother high resolution MERRA data. In the

following section, where we describe the procedures implemented to adjust the selected variables, bilinearly interpolated high resolution MERRA data are subscripted as “MERRA” and physically (e.g., topographical) adjusted climate variables at DEM resolution are subscripted as “DEM.”

## Temperature

Past studies (Dodson and Marks, 1997; Liston and Elder, 2006) have demonstrated that a constant vertical lapse rate representing the decrease in temperature with elevation as a linear function is a simple yet effective way to successfully reproduce the temperature distribution in complex terrains. We therefore apply a monthly varying temperature lapse rate to adjust the MERRA temperature using the equation:

$$T_{DEM} = T_{MERRA} - \Gamma (z_{DEM} - z_{MERRA}) \quad (3.1)$$

where  $T_{DEM}$  is topographically adjusted temperature at DEM resolution,  $T_{MERRA}$  is the bilinearly interpolated MERRA temperature at DEM resolution,  $z_{DEM}$  is DEM elevation,  $z_{MERRA}$  is the elevation from MERRA geopotential height bilinearly interpolated to DEM resolution and  $\Gamma$  is the monthly varying lapse rate. A global averaged monthly lapse rate obtained from Table 1 from Liston and Elder (2006),  $\Gamma$ , is provided as the default lapse rate for each month. Recognizing the high variability of temperature lapse rate with both space and time, MSDH also allows users to calculate lapse rate from local ground-based data.

## Shortwave Radiation

First, we evaluate top of the atmosphere solar radiation ( $SW_{top}$ ) for the three hour interval based on solar constant ( $S^*$ ), and the zenith angle ( $Z$ ) of the sun, which is a

function of latitude, date, and time (Dingman, 2002). A single value was assumed for the whole domain based on a central latitude and longitude.

$$SW_{top} = S^* \cos(Z) \quad (3.2)$$

We evaluated attenuation of solar radiation as the ratio of MERRA shortwave radiation ( $SW_{MERRA}$ ) to the top of the atmosphere solar radiation ( $SW_{top}$ ), expressed as a transmission factor,  $T_{f-MERRA}$ .

$$T_{f-MERRA} = \frac{SW_{MERRA}}{SW_{top}} \quad (3.3)$$

We parameterize the attenuation of solar radiation using Beer's atmospheric transmission law assuming that the optical thickness above a point is based on the atmospheric pressure.

$$SW(P) = SW_{top} e^{-k \cdot P} \quad (3.4)$$

where  $k$  is the atmospheric attenuation coefficient,  $P$  atmospheric pressure and  $SW$  shortwave radiation at a height with atmospheric pressure  $P$ . The following standard atmospheric pressure versus elevation function is used:

$$P_{DEM} = P_0 \left( \frac{T_0 + z \lambda}{T_0} \right)^{-\frac{g}{R\lambda}} \quad (3.5)$$

where  $P_0$  is standard sea level pressure (101,325 Pa),  $T_0$  is standard sea level temperature (288.15 K),  $g$  is earth gravitational acceleration ( $9.81 \text{ m s}^{-2}$ ),  $R$  is the gas constant for dry air ( $287.04 \text{ J kg}^{-1} \text{ K}^{-1}$ ) and  $\lambda$  is the temperature lapse rate calculated by MSDH or provided by the user. The atmospheric attenuation coefficient is determined by solving (3.4) for  $k$  and using the transmission factor evaluated in (3.3),

$$k = \frac{-\log(T_{f-MERRA})}{P_{MERRA}} \quad (3.6)$$

Then we evaluate shortwave radiation at DEM grid resolution ( $SW_{DEM}$ ) using the following equation.

$$SW_{DEM} = SW_{top} e^{-k P_{DEM}} \quad (3.7)$$

## Relative Humidity

MERRA specific humidity is used to calculate actual vapor pressure at MERRA elevations that are specified by geo-potential height.

$$e_{MERRA} = \frac{q_{MERRA} * P_{MERRA}}{(0.622 + q_{MERRA})} \quad (3.8)$$

where  $q_{MERRA}$  is bilinearly interpolated MERRA specific humidity at DEM resolution,  $P_{MERRA}$  is bilinearly interpolated MERRA pressure at DEM resolution and  $e_{MERRA}$  is actual air vapor pressure at DEM resolution.

This is then used to evaluate dew point temperature at MERRA elevation ( $T_{d-MERRA}$ ).

$$T_{d-MERRA} = \frac{c \ln \left[ \frac{e_{MERRA}}{a} \right]}{b - \ln \left[ \frac{e_{MERRA}}{a} \right]} \quad (3.9)$$

where for ice/snow,  $a = 611.21 \text{ Pa}$ ,  $b = 22.452$  and  $c = 272.55 \text{ }^{\circ}\text{C}$ . Dew point is then adjusted for DEM elevation using a monthly vapor pressure coefficient  $\lambda (\text{m}^{-1})$  provided by Liston and Elder, table 1 (2006).

$$T_{d-DEM} = T_{d-MERRA} + (z_{DEM} - z_{MERRA}) \lambda \frac{c}{b} \quad (3.10)$$

where  $T_{d-MERRA}$  and  $T_{d-DEM}$  are dew point temperature at MERRA elevation and DEM grid elevation, respectively. The following function is used to relate saturation vapor pressure and temperature:

$$e_s(T) = a \exp \left( \frac{b T}{c + T} \right) \quad (3.11)$$

This is used to evaluate relative humidity as the ratio of actual and saturated air vapor pressure from dew point and air temperatures at DEM elevation.

$$RH_{DEM} = \frac{e_s(T_d-DEM)}{e_s(T_{DEM})} \quad (3.12)$$

## Wind Speed

MERRA eastward, E-W ( $U_{MERRA}$ ), and northward, N-S ( $V_{MERRA}$ ) wind components are combined by Pythagoras' equation (equation 3.13) to obtain the horizontal wind speed magnitude.

$$W_{MERRA} = \sqrt{(U_{MERRA})^2 + (V_{MERRA})^2} \quad (3.13)$$

Wind direction, terrain slope and terrain aspect are calculated using equations (3.14), (3.15) and (3.16) as suggested by Liston and Sturm (1998) and Liston and Elder (2006).

$$\theta = \frac{3\pi}{2} - \tan^{-1} \left( \frac{V_{MERRA}}{U_{MERRA}} \right) \quad (3.14)$$

$$\beta = \tan^{-1} \sqrt{\left[ \left( \frac{\Delta z_x}{\Delta x} \right)^2 + \left( \frac{\Delta z_y}{\Delta y} \right)^2 \right]} \quad (3.15)$$

$$\gamma = \frac{3\pi}{2} - \tan^{-1} \frac{\left( \frac{\Delta z}{\Delta y} \right)}{\left( \frac{\Delta z}{\Delta x} \right)} \quad (3.16)$$

Both slope and aspect are computed using “Four nearest” method where  $\Delta z_x$  and  $\Delta z_y$  are the elevation difference between the two nearest cells of the target cell in horizontal and vertical directions, respectively.

Equation (3.17) parameterizes the effect of the terrain slope and curvature on the MERRA wind speed ( $W_{MERRA}$ ) (Liston and Elder, 2006).

$$W_{DEM} = W_{MERRA} (1 + \gamma_s \Omega_s + \gamma_c \Omega_c) \quad (3.17)$$

where  $\Omega_c$  (equation 3.18) and  $\Omega_s$  (equation 3.19) are the curvature and slope in the direction of the wind, respectively.

$$\Omega_c = \frac{1}{4} \left[ \frac{z - 0.5(z_w - z_e)}{2\eta} + \frac{z - 0.5(z_s - z_n)}{2\eta} + \frac{z - 0.5(z_{sw} - z_{ne})}{2\sqrt{2}\eta} + \frac{z - 0.5(z_{nw} - z_{se})}{2\sqrt{2}\eta} \right] \quad (3.18)$$

$$\Omega_s = \beta \cos(\theta - \gamma) \quad (3.19)$$

where  $z_e, z_w, z_n, z_s, z_{sw}, z_{ne}, z_{nw}, z_{se}$  are the elevations at eight possible neighboring cells at the east, west, north, south, south-west, north-east, north-west and south-east direction from the target cell.  $\eta$  is the distance between the center of two neighboring cells. Both curvature and slope are normalized such that their values range between -0.5 to 0.5 over the watershed or target domain. In equation 3.17,  $\gamma_c$  and  $\gamma_s$  are weight factors that adjust wind magnitude based on curvature and slope, respectively. Liston and Elder (2006) suggested that the valid range of  $\gamma_c$  and  $\gamma_s$  is between 0 to 1 such that  $\gamma_c + \gamma_s = 1.0$ . In MSDH, we approximated both of these quantities as 0.5 assuming equal weight for slope and curvature adjustments.

## Precipitation

After distributing the reanalysis precipitation over the domain distributed at DEM spatial resolution using bilinear interpolation, topographical adjustments were made using the non-linear relationship between elevation and precipitation expressed as follows.

$$P_{DEM} = P_{MERRA} \left[ \frac{1 + \kappa_p (Z_{DEM} - Z_{MERRA})}{1 - \kappa_p (Z_{DEM} - Z_{MERRA})} \right] \quad (3.20)$$

$P_{MERRA}$  is the MERRA or RFE2 reanalysis precipitation interpolated at DEM solution, and  $\kappa_p$  is a coefficient that quantifies how precipitation varies with elevation in this function. Like temperature lapse rate ( $\Gamma$ ), a global averaged monthly  $\kappa_p$  is provided

as the default value.  $\kappa_p$  can also be calculated by fitting equation (3.20) to precipitation and elevation from point observations in an iterative process. After the precipitation data is produced for a given time period, a bias coefficient ( $B_c$ ) is calculated for bias adjustment.  $B_c$  is the ratio of the observed data at a precipitation measuring station and the downscaled data at the grid cell in which the station is located.

$$B_c = \frac{P_o}{P_d} \quad (3.21)$$

where  $P_o$  and  $P_d$  are mean annual observed precipitation (m) and downscaled precipitation (m), respectively.

If multiple stations are located in or near the target spatial domain,  $B_c$  is calculated for each station, and an average value is taken. Finally, downscaled data is corrected by multiplying by the bias coefficient ( $B_c$ ).

### **Longwave Radiation**

We estimated incoming longwave radiation based on downscaled air temperature following the methods of Liston and Elder (2006). First we evaluate the elevation at 700 millibar pressure level using (3.5). And then air and dew point temperatures at this elevation using (3.1) and (3.10), and finally, relative humidity at this elevation using (3.12).

Implied cloud fraction  $\sigma_c$  and then emissivity  $\epsilon$  is parameterized by Walcek (1994) using equation (3.22) and by Iziomon et al. (2003) using equation (3.23), respectively.

$$\sigma_c = 0.832 \exp\left(\frac{RH_{700} - 100}{41.6}\right) \quad (3.22)$$

$$\varepsilon = \kappa_e (1 + Z_s \sigma_c^2) \left( 1 - X_s \exp \left( \frac{-Y_s e_{DEM}}{T_{DEM}} \right) \right) \quad (3.23)$$

where  $e_{DEM}$  is the atmospheric vapor pressure at DEM resolution and  $\kappa_e$  is 1.08 (Liston and Elder, 2006).  $X_s$ ,  $Y_s$  and  $Z_s$  are coefficients that vary depending on elevation. At elevations below 200 m,  $X_s$ ,  $Y_s$  and  $Z_s$  are 0.35, 0.1 K Pa<sup>-1</sup> and 0.224, respectively.  $X_s$ ,  $Y_s$  and  $Z_s$  are 0.51, 0.13 K Pa<sup>-1</sup> and 1.1, respectively, at elevations above 3000 m. These coefficients vary linearly between these values for elevations from 200 to 3000 m. We then calculate incoming longwave radiation using the Stefan-Boltzmann equation.

$$Q_{li-DEM} = \varepsilon \sigma_c (T_{DEM})^4 \quad (3.24)$$

where  $\sigma$  is the Stefan-Boltzmann constant ( $5.670373 \times 10^{-8} \text{ kg s}^{-3} \text{ K}^{-4}$ ).

The downscaling parameterizations detailed above have been either drawn from the literature or developed in this study based on physical principles and reflect our judgment as to the most appropriate parameterizations to use, given the information available, for downscaling from the relatively coarse grid scale of MERRA variables at the MERRA geopotential height to the elevation associated with the fine scale grid used by a distributed hydrologic model.

### **3.3.3. Utah Energy Balance Snow and Glacier Melt Model**

The Utah Energy Balance model is a spatially distributed model that uses energy balance formulations to simulate the snowmelt and SWE over a watershed, driven by the gridded climate variables (Luce and Tarboton, 2010; Mahat and Tarboton, 2012; Tarboton et al., 1995; Tarboton and Luce, 1996; You, 2004). UEB is physically based and tracks point energy and mass balances to model snow accumulation and melt. UEB has four state variables: surface snow water equivalent,  $W_s$  (m); surface snow and

substrate energy content,  $U_s$  ( $\text{kJ m}^{-2} \text{hr}^{-1}$ ); the dimensionless age of the snow surface  $\eta$ ; and the snow water equivalent of canopy intercepted snow,  $W_c$ , (m). The model is driven by time-varying air temperature, precipitation, wind speed, relative humidity, and incoming shortwave and longwave radiation at time steps sufficient to resolve the diurnal cycle. A detailed description of the distributed version of UEB is provided by Sen Gupta and Tarboton (2013) and also in Chapter 4.

### **3.4. Software Implementation**

#### **3.4.1. Implementing Downscaling Algorithms in R**

R is a statistical software and scripting language initially developed for statistical analysis such as hypothesis testing, time series analysis and plotting, and linear and nonlinear modeling (Carslaw and Ropkins, 2012). R is also extensively used in environmental data analysis, visualization, and modeling. Open source, highly optimized coding functionality, extensibility, and simplicity contributed significantly to the large popularity of R. Users can extend its functionality by writing R packages, collections of well-structured reusable functions and data. These packages can be distributed to the entire R user group through a single web repository (Horsburgh and Reeder, 2014; Pinheiro et al., 2011). In this study, we used several existing R packages such as utils, ncdf (Pierce, 2011), rgdal (Keitt et al., 2011), and raster (Hijmans et al., 2013). We also used NetCDF Operators (NCO) (Zender, 2008) and Climate Data Operators (CDO) (Schulzweida et al., 2006) tools for efficient manipulation of netCDF files. NCO and CDO are both collections of operators for statistical and arithmetic processes, subsetting, interpolation, extrapolation, and transformation of geospatial time series data stored in

netCDF files. The windows version of NCO and CDO program executables are called from R using the system() function.

First, a R function was developed to download MERRA and RFE2 data for the variables listed in Table 3.1 for a specified spatial and temporal extent using the binary file transfer method provided in the function download.file() from utils package. Three files are downloaded for a single day, where one file contains temperature, wind speed, and specific humidity data and the other two files contain precipitation and shortwave radiation, respectively. MERRA and RFE2 files are downloaded in netCDF and zipped binary grid format by our code. RFE2 binary grid files are converted into netCDF to achieve a uniform file format for the datasets. This is done by unzipping the files using R's gunzip() function and then converting binary grid files to netCDF using the CDO import\_binary command.

Next, for each netCDF file, all the MERRA and RFE2 variables are aggregated into three hourly time steps. Hourly MERRA data, such as temperature, is averaged over a three-hourly time step using NCO's ncra command. Then, daily RFE2 precipitation is uniformly distributed into three-hourly time steps by creating a separate netCDF file for each day containing eight time steps using CDO's arithmetic process capability on netCDF datasets (Schulzweida et al., 2006).

A TIFF or image file of the DEM is read into R using rgdal's readGDAL() function and converted into a RasterLayer object. A RasterLayer object is single layer of raster data described by a set of parameters, such as number of columns and rows, spatial resolution, the coordinates of its spatial extent, and map projection. The DEM RasterLayer represents the domain and modeling grid that is the target for the

downscaling. Then MERRA and RFE2 variables such as temperature and precipitation are read from netCDF files for each time step as a two-dimensional array. Using latitude and longitude bounding box information, the array is projected into another RasterLayer, then the netCDF RasterLayer is projected to the DEM RasterLayer using the projectRaster() function from raster package. This function of the raster package bilinearly interpolates the values of the netCDF RasterLayer to the extent and resolution of the DEM and transforms its projection to the DEM's projection (coordinate reference system, CRS). MERRA Geo-potential height in netCDF files are converted to a MERRA height RasterLayer with the resolution and spatial extent of the DEM. The conversion of multiple two-dimensional data objects to a uniform RasterLayer eases the implementation of the topographical adjustment algorithms described in Section 3.2. Once the adjustment algorithms are implemented, the final RasterLayer of each output variable is converted into a two-dimensional matrix in R and appended onto a designated netCDF file that holds the downscaled result.

### **3.4.2. Output Data Storage in NetCDF**

The input and output gridded data used in MSDH are stored in netCDF files. NetCDF is a binary, multidimensional format commonly used by the oceanographic and atmospheric scientific communities for storing and managing scientific data. NetCDF3 (Rew and Davis, 1990) is a machine-independent format that allows direct access, shared access, visualization, and appending of new data to portable binary files. The output netCDF files of MSDH are always three-dimensional: (a) X (m), (b) Y (m) and (c) time (hours). Since the climate variables are produced at the surface, altitude is not a required dimension.

The ordering of the dimensions is important in order to minimize the time necessary to retrieve the data from a netCDF file. For a netCDF file, the “most rapidly varying dimension” is the dimension that corresponds to the most rapidly changing index of the data array in computer memory. In applications written in column-major languages such as FORTRAN, the first index is the most rapidly varying dimension when reading through a multidimensional array. On the other hand, in C and C++, which are row-major languages, the last index reading is the fastest. In hydrology, time series data are of great importance for running hydrologic models and performing hydrologic analyses, and fastest access to data on the time dimension appears to be the most desired. Since the data produced by MSDH may be used in a wide range of applications written in both column-major and row-major languages, we provide an option to the users to choose the order of the time dimension in the file.

Each variable is associated with several attributes, such as short name, long name, units, a numeric value to represent the missing data, and a plausible range of values. All six variables are stored in the same netCDF file with a data array for each variable corresponding to the same set of dimension vectors. A large volume of data might be generated if the program is run for multiple years or at a very high spatial resolution or combination of these two. To avoid storing a large volume of data in a single netCDF file, a separate file is created for each month. The temporal sequence of the data between multiple files is maintained by incrementing the time dimension from “time of origin” or start time. The units of time dimension stores the start time in each file.

### 3.4.3. MSDH Graphical User interface

Using R packages is a relatively straightforward task for experienced users, but it can be challenging, with a steep learning curve, for beginners with no prior programming experience. We, therefore, developed a GUI in order to create a visual environment for the users to enter inputs and execute the R functions. The GUI is also coded in R using the RGtk package and the R script runs from a C# wrapper program. Thus, the MSDH GUI hides the R code from the user and enables data downloading and downscaling tasks while eliminating the complexity of creating or editing codes, files, and folders.

The MSDH has three main tabs: (1) data download, (2) coefficient calculations and (3) data downscale. The “data download” tab (Figure 3.1) provides an option for the users to download data for the variables listed in Table 3.1 using R’s utils package. Precipitation can be downloaded from either RFE2 or MERRA. MERRA data is available globally, while RFE2 covers only the South Asian region, but with better resolution.

The “coefficients calculations” tab performs the task of calculating monthly temperature lapse rate and precipitation adjustment coefficient using the observational data from the station network within the target domain or a watershed.

The “data downscale” tab performs the four-step downscaling methodology described in Section 3.2. The user only needs to specify a DEM of the target spatial domain (in image/TIFF format). The user is provided with a capability to choose the source of the temperature lapse rate or precipitation adjustment factor from a set of options, such as (1) default specified by Liston and Elder (2006), (2) calculated from the “Coefficients calculations” and (3) user input.

### 3.5. Example Application

To test the downscaling methodology, software implementation, and test that the GUI functioned as intended, MSDH was run for one water year starting from October 2009 over the 570 km<sup>2</sup> Logan River watershed (41.71° to 42.09° N latitude and –111.82° to –111.47° W longitude, Figure 3.2) at 120 m resolution. The Logan River flows southwesterly through the mountains of Cache County, Utah. The elevation of the watershed ranges from 1382 m to 3040 m, with an average elevation of 2294 m. Five U.S. Department of Agriculture snowpack telemetry (SNOTEL) stations are located inside the watershed and one SNOTEL station is about 550 m outside the eastern boundary of the watershed (Table 3.2). Daily historical minimum, mean, and maximum temperature; daily total precipitation; snow depth; and SWE data are available at these stations.

Daily mean temperature and total precipitation data were downloaded for each selected SNOTEL station and aggregated to monthly time steps. For each month, a linear regression model was developed between the elevation and mean monthly temperature (e.g., Figure 3.3 (a) for December). In Figure 3.3 (a), the slope of the regression model (-0.00421 °C/m) is the lapse rate for December in the Logan River watershed. Similarly, Equation (3.20) represents the relation between elevation and mean monthly precipitation illustrated in Figure 3.3 (b). The curve fitting coefficient of the nonlinear least-square model (blue line) is the precipitation adjustment coefficient ( $\kappa_p$ ).

MERRA temperature data was downloaded for the contiguous United States (Figure 3.4 (a)) and the six grid cells spanning the Logan River watershed (Figure 3.4 (b)) were used in bilinear interpolation to obtain gridded temperature at the scale of the

DEM (Figure 3.4 (c)). This involved using R's raster library projection transformation capability to transform the data into the DEM's Universal Transverse Mercator (UTM) projection system and clip it to the extent of the DEM. This raster layer contains bilinearly resampled temperature data, while its spatial domain, resolution, and number of rows and columns are exactly the same as the DEM. Next, temperature was adjusted using the monthly lapse rate and the difference between MERRA elevation and DEM elevation using the methodology described in Section 3.3.2. This procedure was repeated for all time steps. Other variables, such as incoming shortwave radiation and wind speed, were also downscaled to the DEM spatial scale using the physically based methodology described in Section 3.3.2.

Daily mean wind speed, relative humidity, and incoming shortwave radiation are only available at USU Doc Daniel (table 3.2) from October 2009 to June 2010 from a separate study by Mahat and Tarboton (2012; 2013) and Mahat et al. (2013). The observations at this station were compared with the downscaled data at the grid cell where the station is located, to test how closely the model reproduces the observation (Figure 3.5). Table 3.3 reports the Nash-Sutcliffe Efficiency (NSE, equation 3.25) of the downscaled data with respect to the observations.

$$\text{NSE} = 1 - \frac{\sum_{t=1}^n (\text{Obs}_t - \text{Sim}_t)^2}{\sum_{t=1}^n (\text{Obs}_t - \text{Obs}_{\text{mean}})^2} \quad (3.25)$$

where  $\text{Obs}_t$  and  $\text{Sim}_t$  are observed and simulated values at any time step  $t$  and  $\text{Obs}_{\text{mean}}$  is the mean of observed values.

Figure 3.5 shows that downscaled data captures the seasonal pattern quite successfully, such as the low temperature phases in December and the high temperature phases in June. High NSE (Table 3.3) demonstrates the model's strength to successfully

reproduce the observed temperature. Both downscaled incoming shortwave radiation and relative humidity capture the seasonal cycle of the observed data reasonably well; however, they fail to reproduce some short-term changes and appear to fluctuate at smaller amplitude than the observations at short time scales for some months. This is reflected in their somewhat lower NSE. Nevertheless, the NSE values obtained (Table 3.3; 0.68 for shortwave radiation and 0.65 for relative humidity) indicate the program's capability to reproduce these two variables reasonably well. Compared to these variables, wind speed and precipitation perform rather poorly (i.e., wind speed NSE = 0.16, precipitation NSE = 0.05). The wind discrepancies likely reflect the challenge in representing local (DEM grid scale) wind variability from regional information, while precipitation discrepancies originate both in the driving MERRA data and downscaling. Although 96% of precipitation events were simulated successfully by MERRA, it produces a considerable number of non-observed rainfall events with low magnitudes and fails to simulate the magnitude of observed rainfall events at a satisfactory level (Figure 3.5). Less intense precipitation events are often overestimated, and moderately heavy events are underestimated.

Secondly, we tested whether the program is capable of reproducing the measurements of daily maximum temperature ( $T_{max}$ ), daily minimum temperature ( $T_{min}$ ), daily mean temperature ( $T_{mean}$ ), and monthly precipitation at SNOTEL stations for water year 2010. Figure 3.6 displays the scatter plots of observed data at SNOTEL stations and downscaled data at grid cells where those stations are located. Table 3.4 shows NSE values at each of these six stations. Both daily maximum and mean temperature show very good simulation with NSE of 0.9. Daily minimum temperature

shows slightly lower NSE, ranging from 0.77 to 0.85 indicating slightly lower performance in reproducing daily minimum temperature compared to daily mean and maximum temperature. Overall, the downscaled temperature captures temporal variation quite satisfactorily in both the short- and long-term. Precipitation is reasonably well predicted at SNOTEL stations on a yearly scale, with differences ranging between -20% and 12% of the observed data. However, at daily or monthly time steps, precipitation simulation incorporates considerable uncertainty, especially during the late winter and early spring season. Relatively low NSE values for monthly total precipitation (Table 3.4) also indicate high uncertainty in precipitation downscaling.

Finally, the Utah Energy Balance (UEB) snowmelt model was run using the downscaled data to demonstrate the applicability of the data to produce realistic simulations of snow accumulation and melt variability. The observed SWE at SNOTEL stations was compared with the UEB simulated SWE at the cells where the stations are located (Figure 3.7). Two stations, Garden City Summit and Temple Fork, match the snow accumulation and melting pattern with high accuracy (NSE: 0.97 and 0.96, respectively, see Table 3.5). The accumulation and melt pattern is also captured reasonably well at Tony Grove Lake (NSE: 0.76), although with about 20% underestimation of the peak. The peak SWE is underestimated in the other three stations which leads to modeled snow disappearing before observed and lower NSE coefficients at these sites (Table 3.5). The performance was most unsatisfactory at Klondike Narrows, where UEB modeled SWE completely melts about one and a half months before the observation. In Figure 3.3 Klondike Narrows is the point second from the left with elevation of 2210 m (Table 3.2). It has mean monthly precipitation for December

slightly higher than the other two stations at slightly higher elevations (Temple Fork and Garden City Summit). A similar pattern was also observed for other winter months. This phenomena was not captured by the precipitation downscaling method, which resulted in poor simulation of SWE. The mean of the NSE values between the observed and simulated SWE for the six stations (Table 3.5) was 0.6, which is satisfactory, given that the model was entirely run using downscaled reanalysis data.

To test the impact of downscaled variables on SWE simulation, we ran UEB at the USU Doc Daniel SNOTEL station with observed temperature, precipitation, relative humidity, wind speed, shortwave and longwave radiation. We then replace the observed data for each of these variables by the downscaled data to study the error introduced by the downscaled data. In each case only one variable was replaced and simulations were compared with observations (Figure 3.8). Root Mean Square Error (RMSE) was used to quantify the errors.

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (obs_t - sim_t)^2} \quad (3.26)$$

where N is the number of observations, and Obs<sub>t</sub> and Sim<sub>t</sub> are observed and simulated values at any time step t. The SWE simulation using the observed data matches the seasonal accumulation and ablation pattern nicely (Figure 3.8 a). However, the model underestimates the SWE during the accumulation period and overestimates the SWE during the melting season. The root mean square error (RMSE) between the observed and simulated SWE is 0.08 m. Surprisingly the RMSE (0.06 m) in simulated SWE reduces when observed temperature were replaced by the downscaled data (Figures 3.8 b). The seasonal pattern and RMSE do not change significantly while observed precipitation,

wind speed and relative humidity data, respectively, were replaced by downscaled data (Figure 3.8 c, d and e). RMSE values between the observed and simulated SWE were 0.085 m, 0.08 m and 0.09, respectively. However, the performance drastically decreases (RMSE 0.27 m) while solar radiation downscaled data were used for the simulation (Figure 8 f). The possible explanation for this is observed low shortwave radiation (ranging from 30 to 562 W/m<sup>2</sup>) are overestimated by twice or even more at a daily scale for several days in winter and early spring. The overestimated shortwave radiation causes the snow to melt quickly and results lower SWE peak.

The program's run time varied significantly depending on the number of rows and columns in the DEM raster file, as the process takes the majority of runtime to interpolate the variables from MERRA to DEM resolution. The Logan River watershed used here consisted of 420 × 254 grid cells. Constructing data for six variables at 3-hourly time steps for a single month on this grid takes about an hour on a common commodity workstation (Dell Optiplex 780, with Intel Q9650 processor @ 3.0 GhZ and 8 GB RAM).

### **3.6. Discussion**

While developing MSDH, we recognized a number of limitations in downscaling methodologies and input data. As described by Liston and Elder (2006), this is a one-way approach where the vertical feedback between the near-land surface and atmosphere is completely ignored. While surface conditions such as presence of the canopy, soil moisture, and proximity to the water can have substantial impact on the local climate, MSDH adjusts the variables using mainly topographical information. Rienecker et al. (2011) explained many limitations of MERRA data including: (1) poor performance in capturing the diurnal temperature pattern by underestimating daily maximum and

overestimating daily minimum temperature, (2) deviation of 3°C or more from the observation in daily temperature estimates, (3) short heavy precipitation events often simulated as precipitation drizzles, and (4) low solar radiation during daytime precipitation events often over estimated. These inaccuracies in MERRA are directly translated into the downscaled data and are responsible for some of the discrepancies described in the previous section (Section 3.5). From our single watershed investigation it was apparent that reproduction of precipitation with a reasonable accuracy at a daily scale, or even at a monthly scale, was a challenge as manifested by the low NSE values received for precipitation. Interestingly, sensitivity analysis of downscaled variables revealed that despite discrepancies in precipitation, reasonably good simulation of seasonal accumulation of snow water equivalent results in satisfactory simulations with downscaled precipitation inputs and other inputs observed (Figure 3.8 c). On the other hand, when the only downscaled variable used as input to the model was solar radiation (Figure 3.8 f) the start of snow accumulation is delayed and overall there is an under simulation of accumulation. From this we infer that even though the NSE for incoming solar radiation is good overall, and that discrepancies in Figure 3.5 are hard to discern, there is a cumulative discrepancy in downscaled incoming solar radiation that results in erroneous melting too early and hence under simulation of the peak snow water equivalent. This indicates a need to examine ways to improve incoming solar radiation downscaling in addition to precipitation downscaling.

### **3.7. Conclusion**

We have developed spatial downscaling methods that adapt approaches from the MicroMet model by Liston and Elder (2006), DAYMET and MTCLIM to address the

problem of downscaling climate reanalysis data. Variables downscaled include: temperature, precipitation, wind speed, relative humidity, shortwave and longwave radiation. The model produces 3-hourly, high resolution, gridded weather data for input to a spatially distributed hydrologic model. NASA Modern-Era Retrospective Analysis for Research and Applications (MERRA) climate products and Southern Asia Daily Rainfall estimate (RFE2) data are the major inputs to the program. In the first step of a two-step downscaling approach, we bilinearly interpolate RFE2 or MERRA reanalysis data to a high resolution digital elevation model (DEM) grid. In the second step, we make topographic adjustments using well-established relationships of elevation, slope, aspect, curvature, and cloudiness with the selected variables.

Development of MSDH was necessary for constructing topographically adjusted high resolution meteorological data to drive hydrological models in data scarce regions. Reanalysis data such as MERRA were developed to analyze the earth system at global or continental scales, whereas hydrological decision making for water availability and flood forecasting, for example, are studied at the watershed level. MSDH can be used as a tool to bridge the gap between the spatial scales of data and used in these two scientific domains. MSDH is capable of producing data at any grid resolution specified in an input DEM. The example application of the system produced the gridded surface of six variables at 120 m resolution and 3-hourly time steps for the Logan River watershed for 1 year starting on October 1, 2009. The data was then used to drive the Utah Energy Balance (UEB) snowmelt model to simulate one year of snow accumulation and melt. Downscaled variables and simulated SWE showed reasonably good agreement with the

observations, indicating MSDH's capability to produce reasonably good quality high resolution climate data using very limited observational data.

This work showed that it is possible to obtain the input variables required to drive the UEB model entirely from climate reanalysis data extending its applicability to data scarce regions of the world. The discrepancies that result due to errors in the reanalysis data and downscaling model were quantified for a location in the US where there is detailed data available. Comparison (i.e., Nash-Sutcliffe efficiency = 0.6) between SNOTEL observations and the Utah Energy Balance Snowmelt Model-simulated snow water equivalent indicates the degree to which this method is effective. Sources of discrepancies, in terms of precipitation and solar radiation uncertainty were identified and motivate opportunities for future research to reduce uncertainty and improve simulations. These discrepancies need to be factored into the use of simulations driven by downscaled results for hydrological modeling and analysis.

The tool was developed using open source, freely available scripting language and programs. The R code is publically available in bitbucket (<https://bitbucket.org/AvirupSenGupta/msdh.usu>) so that the user community outside the initial development team can participate in future improvements of the software by integrating new approaches and analysis techniques. The program has a GUI to make it accessible to users unfamiliar with R. Downscaled data is saved in CF-convention compatible three dimensional self-describing netCDF format, which makes the data portable across operating systems and accessible and displayable in a number of freely available software tools such as ncdump, ncBrowse, and Integrated Data Viewer (IDV).

The application demonstrated in this paper was successfully run on a PC with the Windows operating system. This is particularly advantageous for developing countries where students, engineers, or even researchers may not have access to the latest model high performance computing systems. Presently, MDSH is only available in windows-based systems. The availability of R and all other required programs, such as NCO and CDO in UNIX/Linux operating systems suggests that the program could be ported to UNIX/Linux based computers with little code modification.

## References

- Benestad, R.: Empirical-statistical downscaling in climate modeling, *Eos, Trans. Am. Geophys. Union*, 85, 417-422, 2004.
- Carslaw, D. C. and Ropkins, K.: openair—An R package for air quality data analysis, *Environ. Modelling & Software*, 27, 52-61, 2012.
- Carter, T., Parry, M., Harasawa, H., and Nishioka, S.: IPCC technical guidelines for assessing climate change impacts and adaptations, Part of the IPCC Special Report to the First Session of the Conference of the Parties to the UN Framework Convention on Climate Change, Intergovernmental Panel on Climate Change. Department of Geography, University College London, UK and Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan, 1994.
- Daly, C.: Guidelines for assessing the suitability of spatial climate data sets, *Int. J. Climatol.*, 26, 707-721, 2006.
- Daly, C., Neilson, R. P., and Phillips, D. L.: A Statistical-Topographic Model for mapping Climatological Precipitation over Mountainous Terrain, *J. Appl. Meteorol.*, 33, 140-150, 1994.
- Daly, C., Taylor, G., and Gibson, W.: The PRISM approach to mapping precipitation and temperature, *Proceedings, 10th Conference on Applied Climatology*, Am. Meteorol. Soc., 10-12, 1997.
- Daly, C., Taylor, G., Gibson, W., Parzybok, T., Johnson, G., and Pasteris, P.: High-quality spatial climate data sets for the United States and beyond, *Trans. ASAE-Am. Soc. of Ag. Engineers*, 43, 1957-1962, 2000.

- Dile, Y. T. and Srinivasan, R.: Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin, JAWRA J. Am. Water Resour. Assoc., 50(5), 1226-1241, 2014.
- Dingman, S. L.: Physical Hydrology, Prentice Hall, 2002.
- Dodson, R. and Marks, D.: Daily air temperature interpolated at high spatial resolution over a large mountainous region, Clim. Res., 8, 1-20, 1997.
- Eaton, B., Gregory, J., Drach, B., Taylor, K., Hankin, S., Caron, J., Signell, R., Bentley, P., Rappa, G., and Höck, H.: NetCDF Climate and Forecast (CF) Metadata Conventions, Version 1.0. 2003.
- Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K., and Takahashi, K.: The Japanese 55-year Reanalysis "JRA-55": an interim report, Sola, 7, 149-152, 2011.
- Folk, M., McGrath, R. E., and Yeager, N.: HDF: an update and future directions, Geoscience and Remote Sensing Symposium, 1999. IGARSS'99 Proceedings. IEEE International, 273-275, 1999.
- Fowler, H. J., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling, Int. J. Climatol., 27, 1547-1578, 2007.
- Hijmans, Robert J., Etten, J. v., Mattiuzzi, M., Summer, M., Greenberg, J. A., Lamigueiro, O. P., Bevan, A., Racine, E. B., and Shortridge, A.: Raster: Geographic data analysis and modeling. In: 2.1-66, 2013.
- Horsburgh, J. S. and Reeder, S. L.: Data visualization and analysis within a Hydrologic Information System: Integrating with the R statistical computing environment, Environ. Modelling & Software, 52, 51-61, 2014.
- Hungerford, R. D., Nemani, R. R., and Running, S. W.: MTCLIM: A mountain microclimate simulation model. In: Forest Service Research Paper INT-414, U.S. Dept. of Agriculture, Washington, D. C., 1989.
- Hunter, R. D. and Meentemeyer, R. K.: Climatologically aided mapping of daily precipitation and temperature, J. Appl. Meteorol., 44, 2005.
- Iziomon, M. G., Mayer, H., and Matzarakis, A.: Downward atmospheric longwave irradiance under clear and cloudy skies: Measurement and parameterization, J. Atmos. and Solar-Terrestrial Phys., 65, 1107-1116, 2003.
- Jeffrey, S. J., Carter, J. O., Moodie, K. B., and Beswick, A. R.: Using spatial interpolation to construct a comprehensive archive of Australian climate data, Environ. Modelling & Software, 16, 309-330, 2001.

- Kanamitsu, M., Kumar, A., Juang, H.-M. H., Schemm, J.-K., Wang, W., Yang, F., Hong, S.-Y., Peng, P., Chen, W., and Moorthi, S.: NCEP dynamical seasonal forecast system 2000, *Bull. Am. Meteorol. Soc.*, 83, 1019-1037, 2002.
- Keitt, T. H., Bivand, R., Pebesma, E., and Rowlingson, B.: rgdal: bindings for the Geospatial Data Abstraction Library, R package version 0.7-1, URL <http://CRAN.R-project.org/package=rgdal>, 2011.
- Krause, A.: Foundations of GTK+ Development. Springer, 2007.
- Liston, G. E. and Elder, K.: A meteorological distribution system for high-resolution terrestrial modeling (MicroMet), *J. Hydrometeorol.*, 7, 217-234, 2006.
- Liston, G. E. and Hiemstra, C. A.: Representing grass-and shrub-snow-atmosphere interactions in climate system models, *J. Clim.*, 24, 2061-2079, 2011.
- Liston, G. E. and Sturm, M.: A snow-transport model for complex terrain, *J. Glaciol.*, 44, 498-516, 1998.
- Lo, Y.-H., Blanco, J. A., Seely, B., Welham, C., and Kimmens, J. P.: Generating reliable meteorological data in mountainous areas with scarce presence of weather records: the performance of MTCLIM in interior British Columbia, Canada, *Environ. Modelling & Software*, 26, 644-657, 2011.
- Lucchesi, R.: File Specification for MERRA Products. GMAO Office Note No. 1 (Version 2.3), 82 pp, 2012.
- Luce, C. H. and Tarboton, D. G.: Evaluation of alternative formulae for calculation of surface temperature in snowmelt models using frequency analysis of temperature observations, *Hydro. and Earth Syst. Sci.*, 14, 535-543, 2010.
- Mahat, V. and Tarboton, D. G.: Canopy radiation transmission for an energy balance snowmelt model, *Water Resour. Res.*, 48, W01534, 2012.
- Mahat, V. and Tarboton, D. G.: Representation of canopy snow interception, unloading and melt in a parsimonious snowmelt model, *Hydrol. Process.*, in press, 2013.
- Mahat, V., Tarboton, D. G., and Molotch, N. P.: Testing above and below canopy representations of turbulent fluxes in an energy balance snowmelt model, *Water Resour. Res.*, 49, 1107-1122, 2013.
- McMurtrie, R., Leuning, R., Thompson, W., and Wheeler, A.: A model of canopy photosynthesis and water use incorporating a mechanistic formulation of leaf CO<sub>2</sub> exchange, *For. Ecol. Manage.*, 52, 261-278, 1992.

NRCS: Natural Resource Conservation Service: USU Doc Daniel SNOTEL Site, <http://www.wcc.nrcs.usda.gov/nwcc/site?sitenum=1098&state=ut>, last access: Oct 2, 2014.

Pierce, D.: ncdf: Interface to Unidata netCDF data files, URL <http://CRAN.R-project.org/package=ncdf>, R package version, 1, 2011.

Pinheiro, J., Bates, D., DebRoy, S., and Sarkar, D.: R Development Core Team. 2010. nlme: linear and nonlinear mixed effects models. R package version 3.1-97, R Foundation for Statistical Computing, Vienna, 2011.

R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2009.

Reichle, R. H., Koster, R. D., De Lannoy, G. J., Forman, B. A., Liu, Q., Mahanama, S. P., and Touré, A.: Assessment and enhancement of MERRA land surface hydrology estimates, *J. Clim.*, 24, 6322-6338, 2011.

Rew, R. and Davis, G.: NetCDF: an interface for scientific data access, *Computer Graphics and Applications*, IEEE, 10, 76-82, 1990.

Rew, R., Davis, G., Emmerson, S., Davies, H., and Hartnett, E.: NetCDF User's Guide, Unidata Program Center, Boulder, Colorado, 1993.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., and Kim, G.-K.: MERRA: NASA's modern-era retrospective analysis for research and applications, *J. Clim.*, 24, 3624-3648, 2011.

Running, S. W., Nemani, R. R., and Hungerford, R. D.: Extrapolation of Synoptic Meteorological Data in Mountainous Terrain and its use for Simulating Forest Evapotranspiration and Photosynthesis, *Can. J. For. Res.*, 17, 472-483, 1987.

Schulzweida, U., Kornblueh, L., and Quast, R.: CDO User's Guide, Climate Data Operators, Version 1.0.1, Max-Planck-Institute for Meteorology, Hamburg, Germany., 2006.

Sen Gupta, A. and Tarboton, D. G.: Using the Utah Energy Balance Snow Melt Model to Quantify Snow and Glacier Melt in the Himalayan Region, 81<sup>st</sup> Annual Western Snow Conference, Jackson Hole, Wyoming, 2013.

Shrestha, M., Artan, G., Bajracharya, S., and Sharma, R.: Using satellite-based rainfall estimates for streamflow modelling: Bagmati Basin, *J. Flood Risk Manage.*, 1, 89-99, 2008.

Suarez, M. J., Rienecker, M., Todling, R., Bacmeister, J., Takacs, L., Liu, H., Gu, W., Sienkiewicz, M., Koster, R., and Gelaro, R.: The GEOS-5 Data Assimilation System—Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0, 2008.

Tarboton, D. G., Chowdhury, T. G., and Jackson, T. H.: A Spatially Distributed Energy Balance Snowmelt Model. In: Biogeochemistry of Seasonally Snow-Covered Catchments (Proceedings of a Boulder Symposium, July 1995), Tonnessen, K. A., Williams, M. W., and Tranter, M. (Eds.), IAHS Publ. no. 228, Wallingford, 1995.

Tarboton, D. G. and Luce, C. H.: Utah Energy Balance Snow Accumulation and Melt Model (UEB), Utah Water Research Laboratory and USDA Forest Service Intermountain Research Station, Logan, UT, 64 pp., 1996.

Thornton, P., Thornton, M., Mayer, B., Wilhelmi, N., Wei, Y., and Cook, R.: DAYMET: Daily Surface Weather on a 1 km Grid for North America. 1980–2008, Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, T, N. doi, 10, 2012.

Thornton, P. E., Running, S. W., and White, M. A.: Generating surfaces of daily meteorological variables over large regions of complex terrain, *J. Hydrol.*, 190, 214-251, 1997.

Trenberth, K. E. and Olson, J. G.: An evaluation and intercomparison of global analyses from the National Meteorological Center and the European Centre for Medium Range Weather Forecasts, *Bull. Am. Meteorol. Soc.*, 69, 1047-1057, 1988.

Walcek, C. J.: Cloud cover and its relationship to relative humidity during a springtime midlatitude cyclone, *Mon. Weather Rev.*, 122, 1021-1035, 1994.

Wigley, T., Jones, P., Briffa, K., and Smith, G.: Obtaining sub-gridscale information from coarse-resolution general circulation model output, *J. Geophy. Res.: Atmospheres* (1984–2012), 95, 1943-1953, 1990.

Wilby, R. L., Dawson, C. W., and Barrow, E. M.: SDSM—a decision support tool for the assessment of regional climate change impacts, *Environ. Modelling & Software*, 17, 145-157, 2002.

Wu, W.-S., Purser, R. J., and Parrish, D. F.: Three-dimensional variational analysis with spatially inhomogeneous covariances, *Mon. Weather Rev.*, 130, 2905-2916, 2002.

Xie, A., Ren, J., Qin, X., and Kang, S.: Reliability of NCEP/NCAR reanalysis data in the Himalayas/Tibetan Plateau, *J. Geogr. Sci.*, 17, 421-430, 2007.

Xie, P., Yarosh, Y., Love, T., Janowiak, J., and Arkin, P. A.: A real-time daily precipitation analysis over South Asia, 16th conference on hydrology, Am. Meteorological Soc., Orlando, FL, 3.15,2002.

Xu, C.-y.: From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches, *Prog. Phys. Geogr.*, 23, 229-249, 1999.

You, J.: Snow Hydrology: The parameterization of subgrid processes within a physically based snow energy and mass balance model, PhD Thesis, Civil and Environmental Engineering, Utah State University, Logan, UT, 2004.

Zender, C. S.: Analysis of self-describing gridded geoscience data with netCDF Operators (NCO), *Environ. Modelling & Software*, 23, 1338-1342, 2008.

Zimmermann, N. E. and Roberts, D. W.: Final report of the MLP climate and biophysical mapping project, Swiss Federal Research Institute WSL/Utah State University, Birmensdorf, Switzerland/Logan, USA, 2001.

Table 3.1. Input MERRA variables used for downscaling

MERRA Variable	description	Spatial resolution (longitude × latitude)	Temporal Resolution
t2m	Temperature at 2 m above the ground (K)	0.67°× 0.5°	hourly
v2m	Northward wind at 2 m above the ground (m s <sup>-1</sup> )	0.67°× 0.5°	hourly
u2m	Eastward wind at 2 m above the ground (m s <sup>-1</sup> )	0.67°× 0.5°	hourly
ps	Time averaged surface pressure (Pa)	0.67°× 0.5°	hourly
qv2m	Specific humidity at 2 m above the ground (kg kg <sup>-1</sup> )	0.67°× 0.5°	hourly
swgdn	Surface downward shortwave flux (W m <sup>-2</sup> )	1.25°× 1.0°	3-hourly

Table 3.2. NRCS SNOTEL stations in the Logan River watershed

Site Name	Latitude (°)	Longitude (°)	Elevation (m)
Garden City Summit	41.9215	-111.41915	2348
Klondike Narrows	41.967689	-111.59713	2210
Temple Fork	41.793	-111.54605	2257
Tony Grove lake	41.898333	-111.62957	2583
Tony Grove RS	41.885733	-111.56918	1930
USU Doc Daniel	41.86425	-111.50603	2521

Table 3.3. Nash-Sutcliffe Efficiency (NSE) of downscaled daily temperature, incoming shortwave, relative humidity, wind speed and precipitation data at USU Doc Daniel weather station.

Variable	Nash-Sutcliffe Efficiency (NSE)
Temperature	0.90
Incoming shortwave radiation	0.68
Relative humidity	0.65
Wind Speed	0.16
Precipitation	0.05

Table 3.4. Nash-Sutcliffe Efficiency (NSE) of daily maximum, minimum and mean temperature and total monthly precipitation at six NRCS SNOTEL stations in the Logan River watershed.

Variable Name	Garden City Summit	Klondike Narrows	Temple Fork	Tony Grove Lake	Tony Grove RS	USU Doc Daniel
T <sub>max</sub>	0.92	0.93	0.93	0.89	0.9	0.93
T <sub>min</sub>	0.79	0.78	0.85	0.84	0.77	0.87
T <sub>mean</sub>	0.9	0.91	0.91	0.9	0.92	0.91
P <sub>month</sub>	0.44	0.22	0.38	-0.01	0.22	0.49

Table 3.5. Nash-Sutcliffe Efficiency (NSE) of UEB simulated Snow Water Equivalent (m) compared to SNOTEL stations for water year 2010.

Stations	Elevation (m)	NSE
Garden City Summit	2348	0.97
Klondike Narrows	2210	0.06
Temple Fork	2257	0.96
Tony Grove Lake	2583	0.76
Tony Grove RS	1930	0.46
USU Doc Daniel	2521	0.39

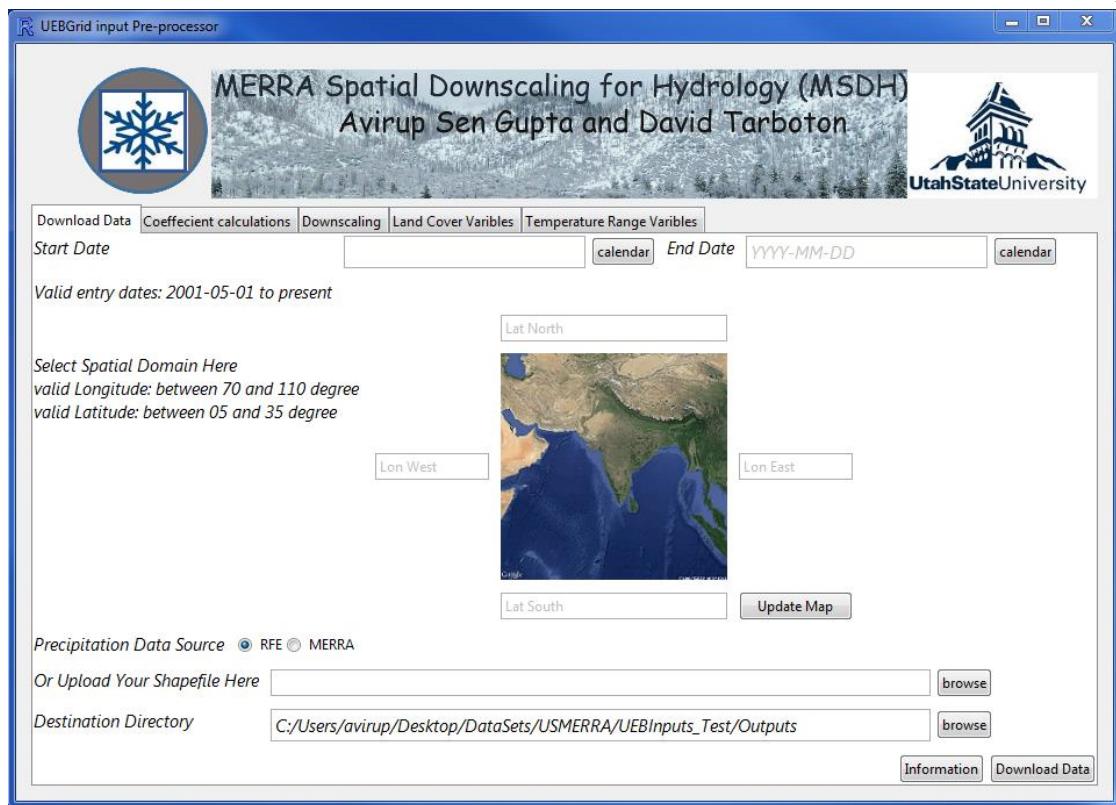


Figure 3.1. Graphical User Interface for MERRA Spatial Downscaling for Hydrology (MSDH).

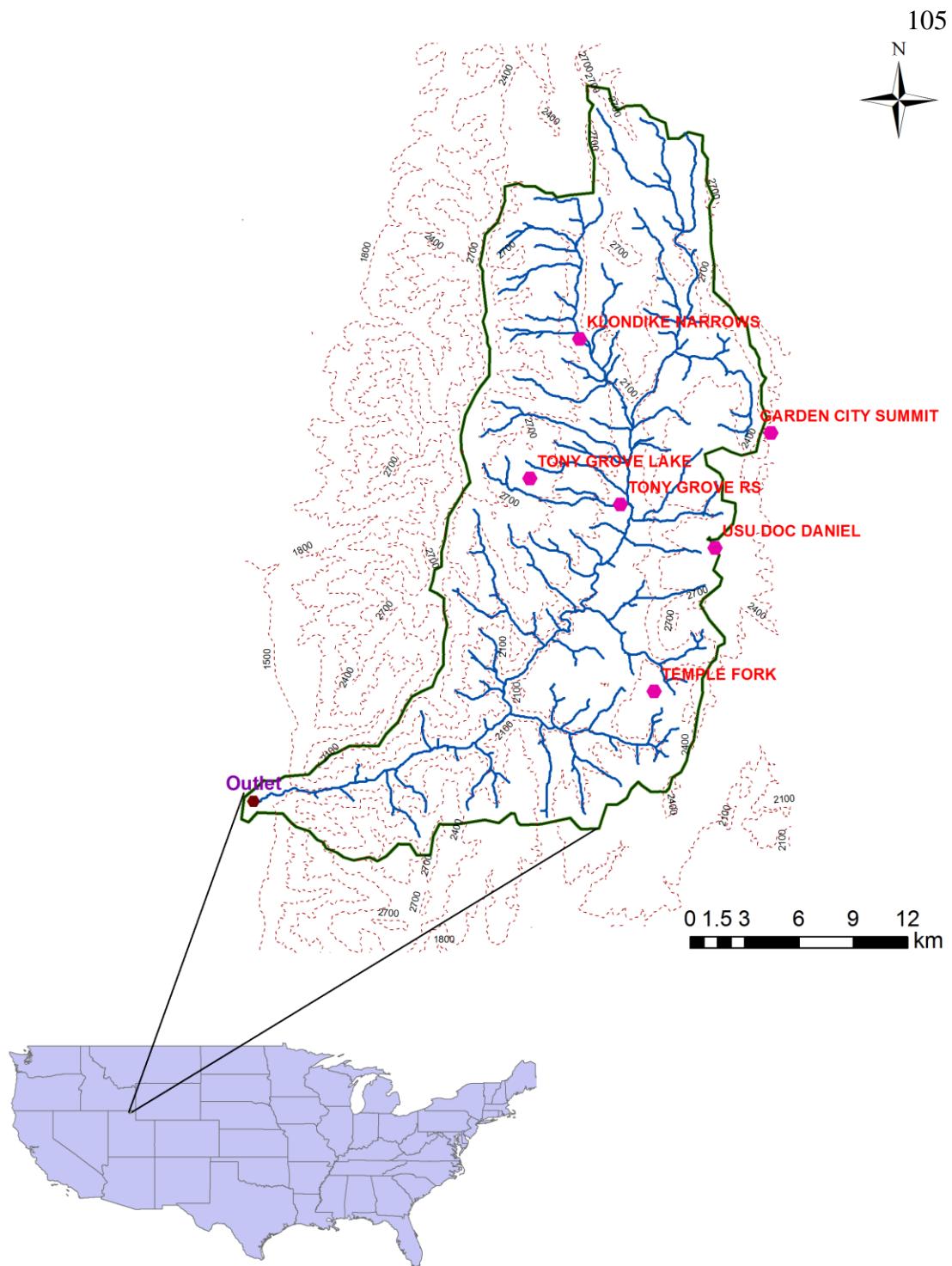


Figure 3.2. Logan River watershed. Blue lines indicate the stream network within the watershed and red dots symbolized the SNOTEL climate stations. Station numbers correspond with the numbers shown in table 3.2.

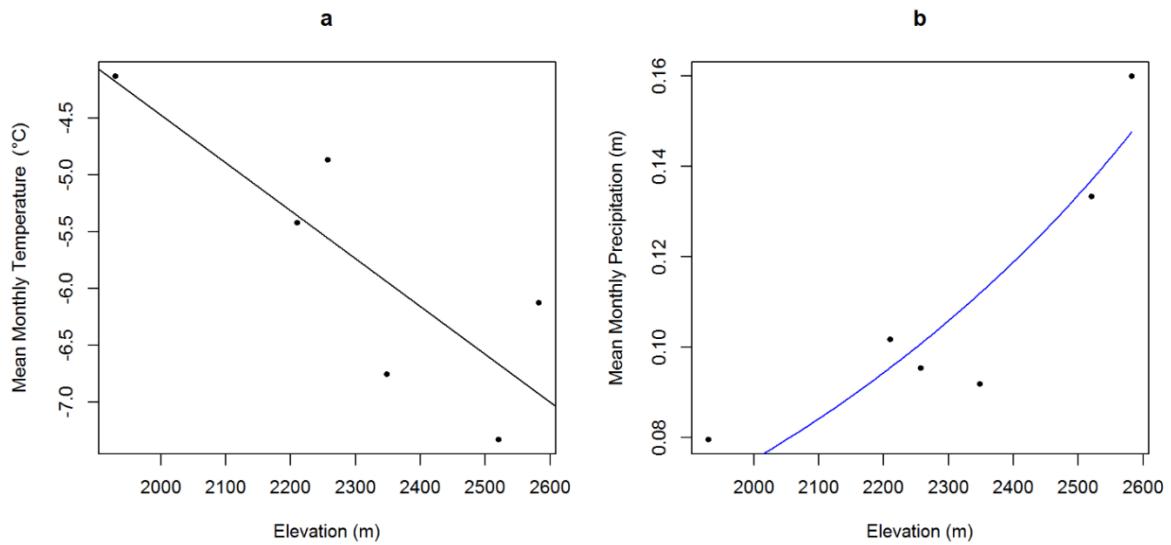


Figure 3.3. Relationship between elevation (x-axis) and (a) mean monthly temperature ( $^{\circ}\text{C}$ ) (y-axis) and (b) mean monthly precipitation (m) (y-axis) at six SNOTEL stations in December for evaluating vertical temperature lapse rate ( $\Gamma$ ) and precipitation adjustment factor ( $\kappa_p$ ), respectively.

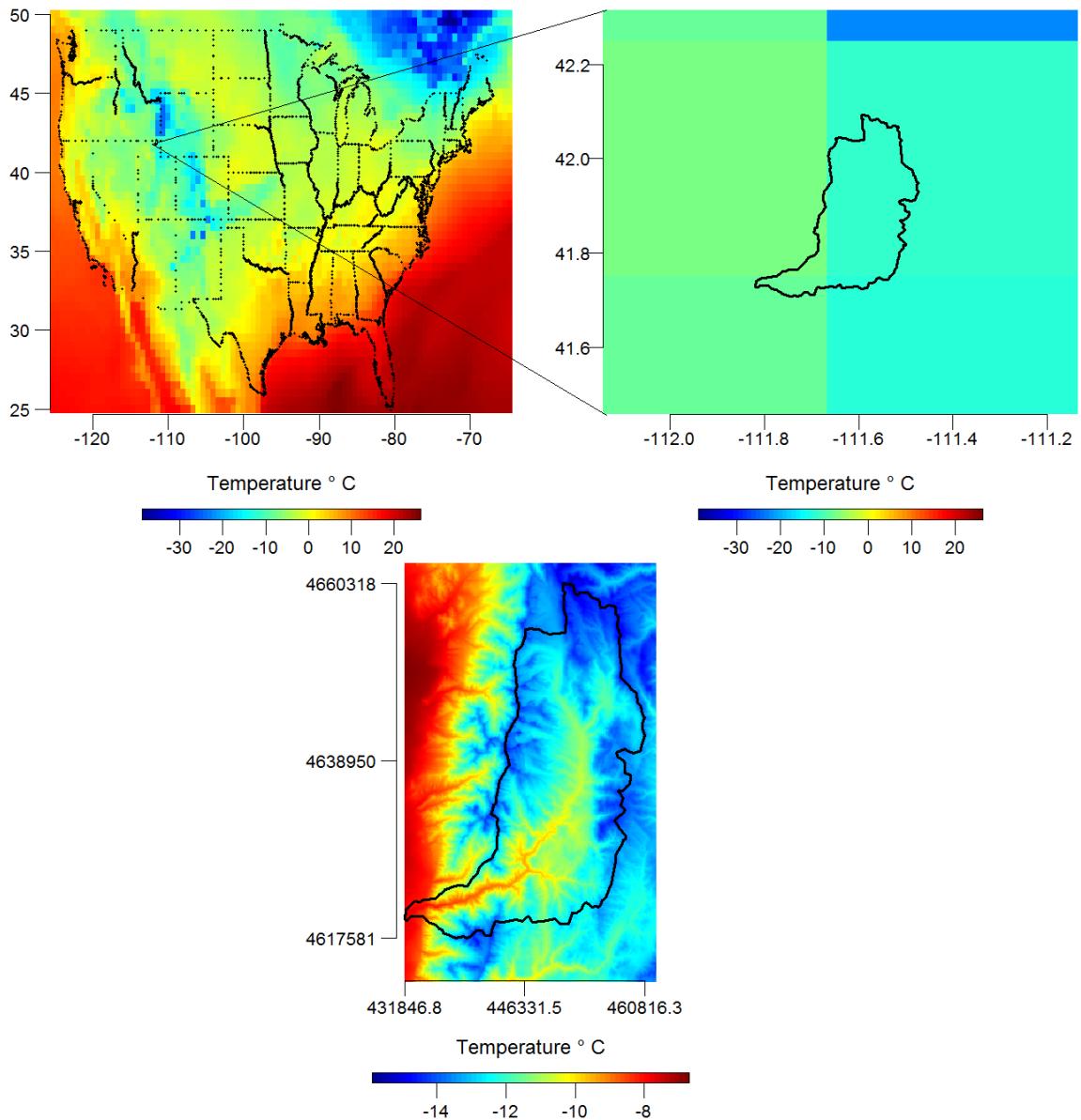


Figure 3.4. Downscaling of MERRA temperature ( $^{\circ}\text{C}$ ) for the Logan River watershed 18:00 UTC on Dec 24, 2009 (a) temperature reported in MERRA for Contiguous United States of America (USA); (b) MERRA grid cells spanning Logan River watershed and surrounding areas and (c) downscaled temperature projected to Universal Transverse Mercator (UTM) projection at DEM grid resolution.

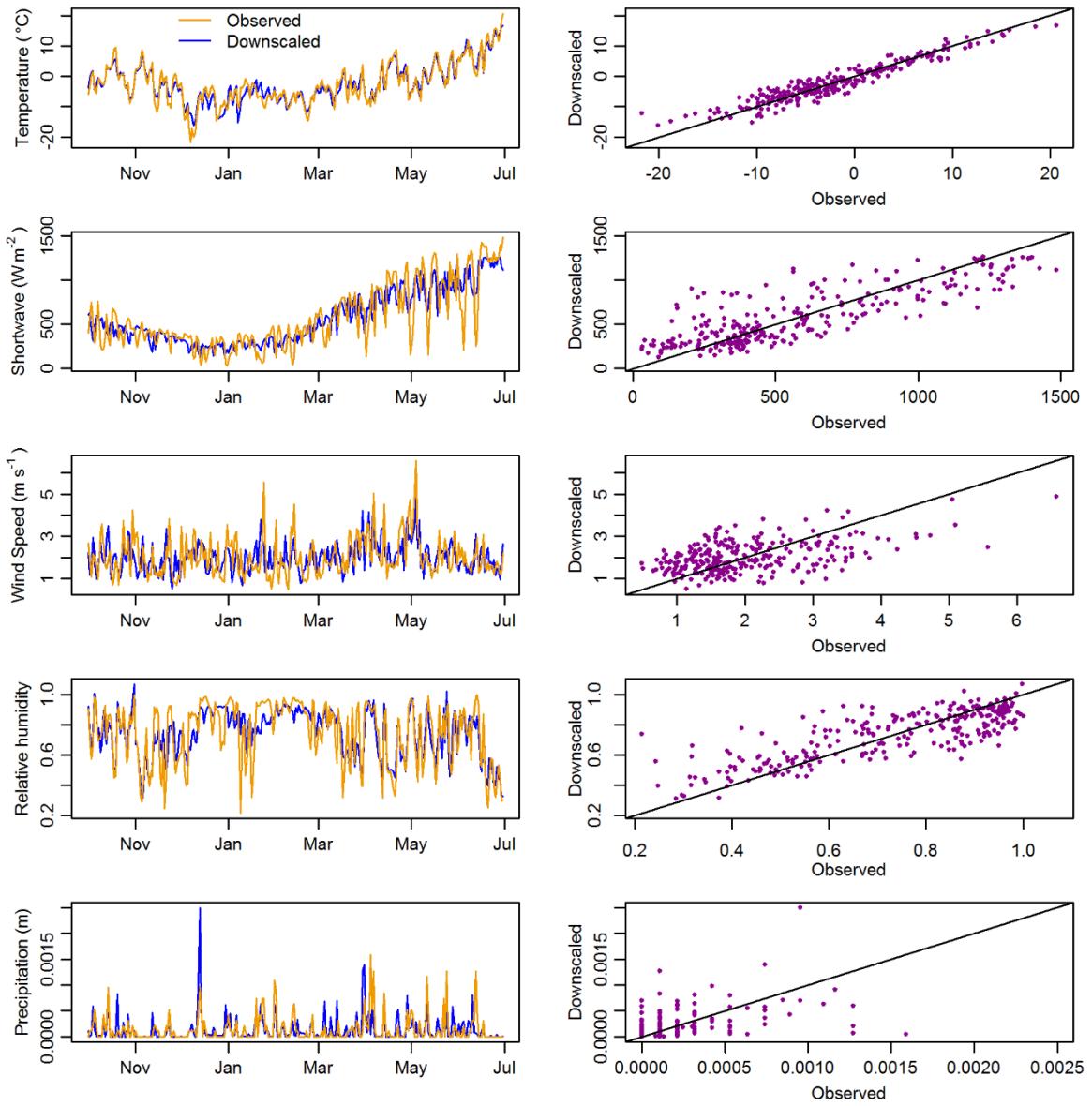


Figure 3.5. Comparison of downscaled daily mean observed temperature, incoming shortwave radiation, wind speed, relative humidity, and precipitation, with respect to measured data at the USU Doc Daniel SNOTEL station. A time series plot (left) and scatter plot (right) of observed and downscaled data are shown for each variable.

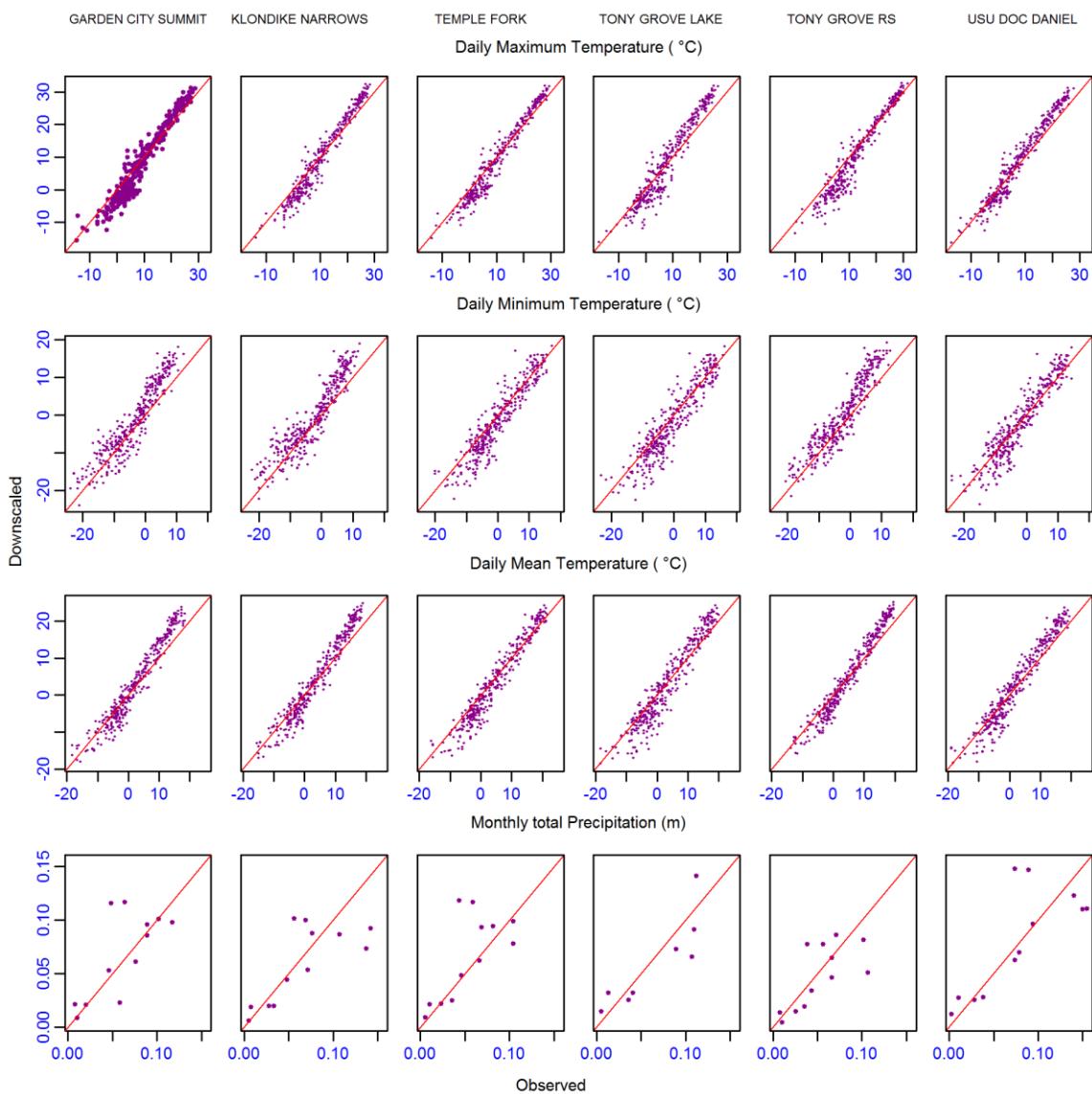


Figure 3.6. Comparison of the downscaled data (y-axis) for daily maximum, minimum and mean temperature and monthly precipitation with observed data (x-axis) at six SNOTEL stations for water year 2010 (Oct 01 2009 - Sep 30 2010). SNOTEL station names are indicated at the top of each column.

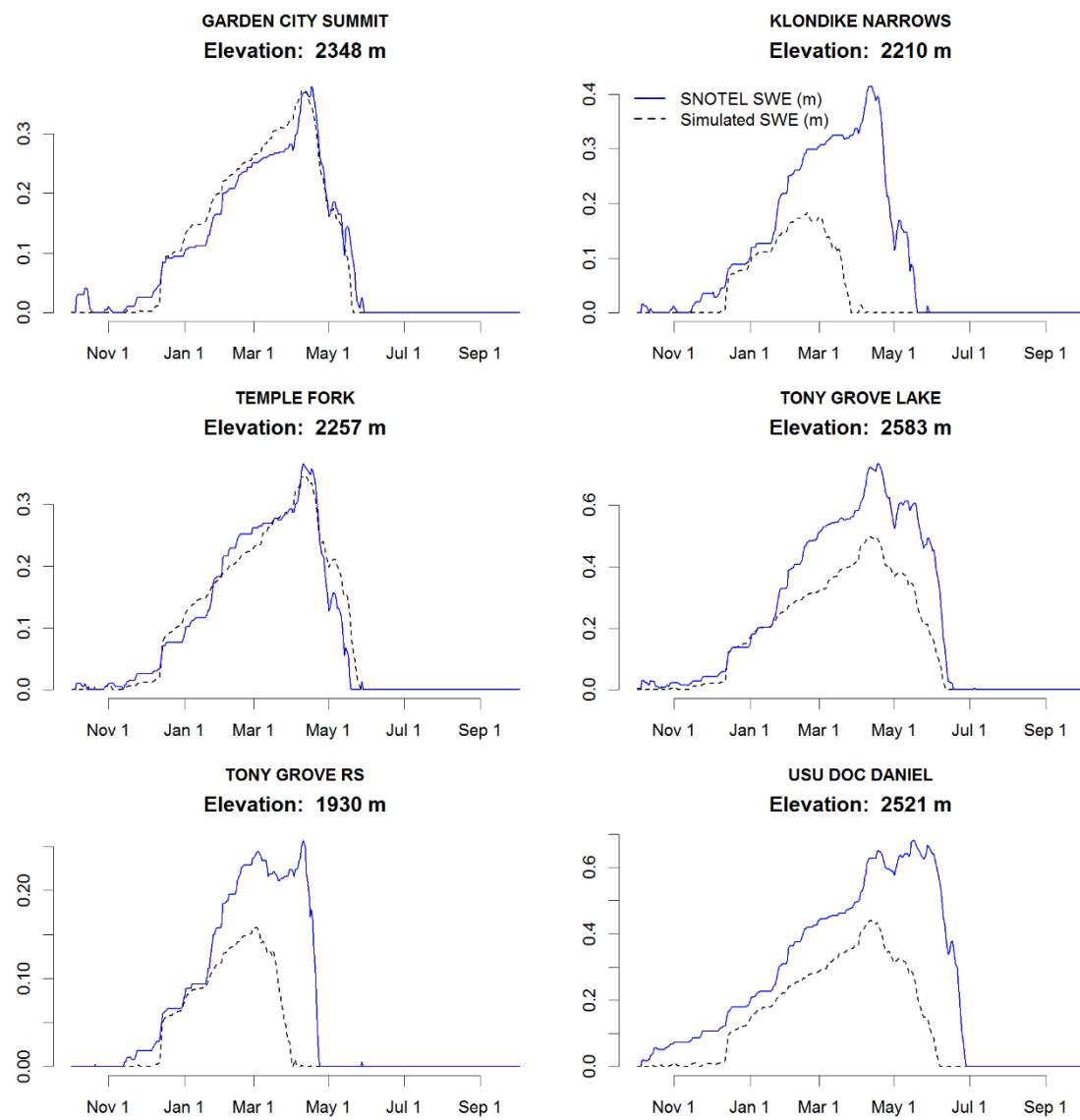


Figure 3.7. Comparison between observed SWE and Utah Energy Balance (UEB) simulated SWE for water year 2010 (Oct 01 2009 - Sep 30 2010) at the six SNOTEL stations.

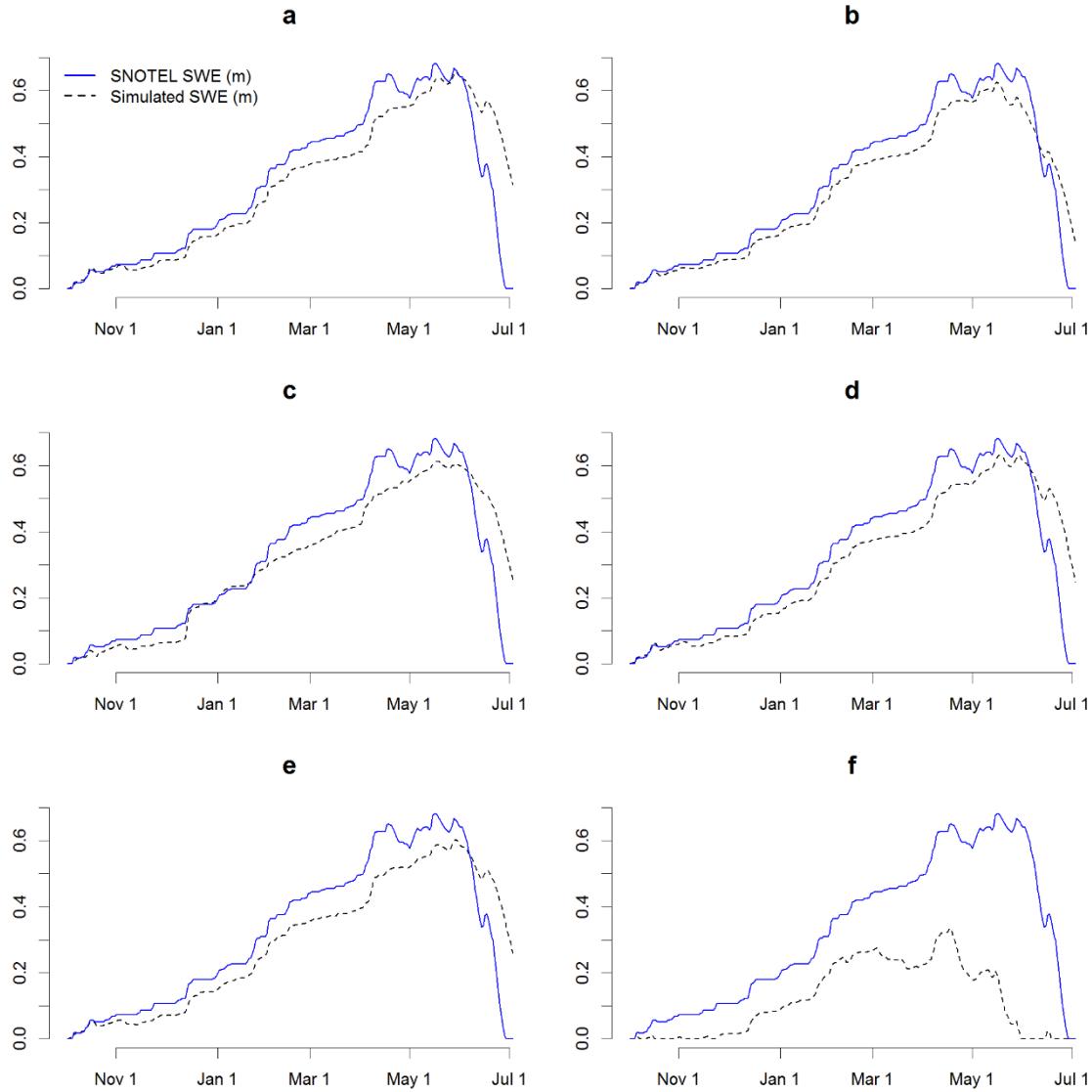


Figure 3.8. Comparison between the observed and UEB simulated snow water equivalent (SWE) at the USU Doc Daniel SNOTEL station using (a) observed temperature, precipitation, wind speed, relative humidity and shortwave radiation, (b) downscaled temperature with observed data of other variables, (c) downscaled precipitation with observed data of other variables, (d) downscaled wind speed with observed data of other variables, (e) downscaled relative humidity with observed data of other variables, (f) downscaled shortwave radiation with observed data of other variables.

## CHAPTER 4

## ESTIMATING SNOW AND GLACIER MELT IN A HIMALAYAN WATERSHED

USING AN ENERGY BALANCE SNOW AND GLACIER MELT MODEL<sup>1</sup>**Abstract**

This study enhances an energy balance snowmelt model (Utah Energy Balance, UEB) to include the capability to quantify glacier melt. To account for clean and debris covered glaciers, substrate albedo and glacier outlines determined from remote sensing, are taken as inputs. The model uses the surface energy balance to compute the melting of seasonal snow and glacier substrate once the seasonal snow has melted. In this application the model was run over a 360 km<sup>2</sup> glacierized watershed, Langtang Khola, in the Nepal Himalaya for a 10-year simulation period starting in water year 2003. The model was run on a distributed mesh of grid cells providing the capability to quantify both timing and spatial variability in snow and glacier melt. The distributed UEB melt model has a relatively high data demand, while the Hindu-Kush Himalayan region is a data-scarce region, a limitation that affects most water resources impact studies in this region. In this study, we determined model inputs from the Modern Era Retrospective-Analysis for Research and Applications (MERRA) and Southern Asia Daily Rainfall Estimate (RFE2) data products. The model estimates that roughly 57% of total surface water input is generated from glacier melt, while snowmelt and rain contribute 34% and 9%, respectively over the simulation period. The melt model provided input to the USGS

---

<sup>1</sup>Coauthored by Avirup Sen Gupta, David G. Tarboton, Adina Racoviteanu, Molly E.

Brown and Shahid Habib

Geospatial Stream Flow Model (GeoSFM) for the computation of streamflow and produced reasonable streamflow simulations at daily scale with some discrepancies, while monthly and annual scale comparisons resulted in better agreement. The result suggests that this approach is of interest for water resources applications where monthly or longer scale streamflow estimates are needed. Mean annual streamflow was positively correlated with the total annual surface water input. However, mean annual streamflow was not correlated with total annual precipitation, highlighting the importance of energy balance melt calculation, in comparison to just using precipitation when considering streamflow availability. Overall, for a 10-year model run, the water equivalent of snow accumulation is 2.46 m compared to 7.13 m of glacier melt over the basin, suggesting a net loss in glacier mass.

#### **4.1. Introduction**

South Asian countries may face water insecurity due to high population and economic growth, and potential climate-induced changes in water availability. Millions of people in the South Asian region depend on the fresh water generated from snow and glacier melt (Kehrwald *et al.*, 2008), but are also at a risk of being subject to flood hazards due to high flow during the melting season. Climate change may disrupt the hydrological balance in snow- and glacier-fed rivers (Kaser *et al.*, 2010). High uncertainty in glacier melt has generated intense debate over the Himalayan glacier mass balance. While studies on the central and eastern Himalayas report glacier shrinkage, studies of the mass balance in Karakoram and the western Himalayas suggest no change or a small increase in mass (Gardelle *et al.*, 2012; 2013). The contribution of glacier melt

to streamflow is poorly understood due to limited climate and hydrological data (Konz *et al.*, 2007). Accurate predictions of melt quantity, timing, and spatial pattern are important for sustainable development in South Asian countries (Viviroli and Weingartner, 2004; Konz *et al.*, 2007). Some research efforts have been directed towards understanding glacier contributions to streamflow using temperature index (Kayastha *et al.*, 2000; 2005; Immerzeel *et al.*, 2010; Konz *et al.*, 2010) and ice ablation models (Racoviteanu *et al.*, 2013). The reliance by these models on empirical relationships between melt rate and temperature or elevation in the case of the ice ablation model limit their ability to address questions related to climate change effects on radiation inputs. Such empirical models are also limited in their ability to quantify energy balance processes in complex topography where the interactions between radiation and topography (slope and aspect) play significant roles in snow and glacier melting. Therefore, more physically based modeling is needed to get a better understanding of glacier mass balance and the contribution of glacier melt to the total surface water input in these settings.

The study examines the following specific questions: (1) How well can glacier melt be quantified using adaptations of a simple energy balance model initially developed for snow?, (2) what is the relative contribution of glacier melt, snowmelt and rain to the total surface water input? and (3) can we infer any changes in glacier mass balance during the model simulation period using an energy balance model for snow and glacier melt?

Previously, the Utah Energy Balance (UEB) model, a physically based energy and mass balance model was designed to track energy and snow mass balance to model snow

accumulation and ablation at a point (Tarboton *et al.*, 1995; Mahat and Tarboton, 2012; 2013; Mahat *et al.*, 2013). For this study, the UEB model was reconfigured to be spatially distributed over a grid, and to include calculations of glacier melt in addition to snowmelt. The distributed version of UEB enables the explicit representation of spatially varying input fields, while retaining the process physics of the previous version. The model enhancement for this study represents glacier as a substrate layer and computes melt from glacier substrate when seasonal snow has melted. Glacier outlines and the albedo of clean glaciers, and debris-covered glaciers are used as inputs to the model.

This work was part of an integrated modeling project to extend remote sensing driven hydrologic modeling capability for the Hindu-Kush Himalayan (HKH) region (Brown *et al.*, 2010; 2014). This project involved the extension to the physics of UEB to include the melting of glaciers developed and evaluated here. This project also included the development of a data model to support the inclusion of UEB into the EPA Better Assessment Science Integrating point and Nonpoint Sources (BASINS) integrated modeling system (EPA, 1998) reported in chapter 2 and the development of downscaling methods to prepare input to UEB from NASA Modern Era Retrospective-Analysis for Research and Applications (MERRA) weather data products (Rienecker *et al.*, 2011) and the NOAA Rainfall Estimate (RFE2) product (Xie *et al.*, 2002) reported in chapter 3. The new version of the UEB model with new capabilities (glacier and grid) developed here is referred to here as UEBGrid.

UEBGrid produces melt outputs from snow or glacier that serve as inputs to the surface hydrologic system. To use these outputs for hydrologic modeling and water

resources analyses, as well as to evaluate the outputs against measured streamflow, a hydrologic model is required to simulate the filtering of surface water inputs through the drainage basin to produce streamflow. Here the United States Geological Survey (USGS) Geospatial Stream Flow Model (GeoSFM) (Asante *et al.*, 2008) rainfall-runoff model was used because of its prior use in the Himalayan region (Shrestha *et al.*, 2008; Shrestha, 2011), and because of its inclusion as a plug-in to the EPA BASINS system.

To evaluate the snow and glacier melt quantity, timing, and spatial pattern simulated using the UEB extensions developed here, we applied UEBGrid and GeoSFM in the Langtang Khola watershed, with inputs downscaled from MERRA and RFE2 data products to compensate for the scarcity of ground-based observation data.

This paper first briefly reviews the processes used in the UEB snowmelt model, describes changes introduced to accommodate glacier melting and convert UEB from a point-based model to a fully distributed model. Then it describes the Langtang Khola watershed, data sources, and downscaling approaches for data preparation prior to applying UEB to this watershed. In the results section, we show the relative contribution of the three components of sources of water: rain, snow, and glacier in total surface water input. We also examine glacier mass balance, the seasonal variability of the simulation of streamflow from the GeoSFM hydrologic model driven by the UEB snow- and glacier-melt and rain outputs. We finally discuss the findings, uncertainties of this study and ideas for consideration in future work.

#### **4.2. Study Area**

The Langtang Khola watershed, approximately 100 km north of Kathmandu Nepal (Figure 4.1), has an area of 360 km<sup>2</sup>. The watershed was divided into grid cells with spatial resolution of 98 m coinciding with grid cells of the projected Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) used to delineate the watershed. Elevation ranges from 3700 m to 7184 m, with an average elevation of 5176 m. One weather station, maintained by Department of Hydrology and Meteorology (DHM), is situated within the watershed at Kyangjing (3920 m, Figure 4.1) and a discharge measuring station is located at the outlet of the watershed (3800 m, Figure 4.1).

#### **4.3. Input Data**

Glacier outline maps for the Langtang Khola watershed were derived from October 2003 ASTER orthorectified products (ASTDMO14) (Racoviteanu *et al.*, 2013). The scenes had a good contrast over glaciers, and as they were acquired at the end of the ablation season had minimal seasonal snow. Based on these glacier outlines, the watershed is 57% non-glacierized (i.e. sparsely vegetated or bare rock), 35% occupied by clean glacier and 8% covered by debris-cover glacier ice (Table 4.1). No area was mapped as accumulation zone to avoid the uncertainty in its mapping. This has the effect of making the model run at each grid cell with the result that the model naturally represents the accumulation zone as those grid cells where snow accumulates indefinitely and does not melt. Substrate albedo (Table 4.1) was derived from the ASTER atmospherically corrected surface reflectance product AST07\_XT, as described in Racoviteanu *et al.* (2013). Slope and aspect were derived from SRTM elevation data.

Land cover variables such as leaf area index (LAI), canopy coverage and height of the canopy were derived from the MODIS MOD12Q1 500-meter land cover product (Sulla-Menashe and Friedl, 2007), although only 1.6 % of total area of Langtang Khola watershed is vegetated.

Time-varying weather and radiation forcing data were derived from MERRA and RFE2 data products. MERRA is a recent near-real-time global climate reanalysis product developed at NASA and derived from the Goddard Earth Observing System version 5 (GEOS-5) NASA general circulation model (Suarez *et al.*, 2008; Rienecker *et al.*, 2011), and National Centers for Environmental Prediction (NCEP) Gridpoint Statistical Interpolation (GSI) analysis (Wu *et al.*, 2002). It is available from 1979 to the present (Lucchesi, 2012). Temperature, wind speed, and relative humidity are available at a spatial resolution of 2/3° longitude by 1/2° latitude, and incoming shortwave radiation is available at a coarser resolution of 1.25° by 1.0° (Lucchesi, 2012).

RFE2 daily total precipitation estimates are constructed using four observational input data sources: approximately 280 ground based stations, geostationary infrared cloud top temperature fields, and precipitation estimated from SSM/I, and AMSU-B microwave sensors (Xie *et al.*, 2002). Near real-time daily rainfall estimates are available for the Southern Asian domain (70°-110° East, 5°-35° North) at a spatial resolution of 0.1° by 0.1° beginning on May 1, 2001.

MERRA and RFE data were downscaled to the 98 m grid scale of the SRTM DEM and 3 hour time steps using the MERRA Spatial Downscaling for Hydrology tool described in chapter 3. In this tool temperature was adjusted for elevation differences

between the effective elevations determined from the geo-potential height that MERRA used and SRTM DEM elevation using a monthly lapse rate obtained from field observations of HKH Cryosphere Monitoring Project (<http://www.icimod.org/?q=8408>) (Immerzeel *et al.*, 2014).

#### **4.4. Model Description**

##### **4.4.1. Distributed Utah Energy Balance Snow and Glacier Melt Model (UEBGrid)**

UEB was originally configured as a point model to produce snowmelt and related outputs driven by the climate inputs at that point (Tarboton *et al.*, 1995; Tarboton and Luce, 1996; You, 2004; Luce and Tarboton, 2010). To enhance the capability of UEB in order to quantify snow processes in a forested area, Mahat and Tarboton (2012) developed a two stream canopy radiation model that explicitly accounts for canopy scattering, absorption, and reflection. They also added capabilities to represent turbulent exchanges within and above a forest canopy (Mahat *et al.*, 2013) and to model snow interception (Mahat and Tarboton, 2013). The UEBGrid model developed here included these forest canopy additions.

The UEB model has four state variables: the surface snow water equivalent,  $W_s$  (m), the surface snow and substrate energy content,  $U_s$  (kJ m<sup>-2</sup>), the dimensionless age of the snow surface, and the snow water equivalent of canopy intercepted snow,  $W_c$  (m). The dimensionless age of snow is a surface condition variable defined by Dickinson *et al.* (1993), to parameterize the sensitivity of the decrease of albedo over time due to the increase in snow grain size and accumulation of dirt to environmental conditions such as temperature. The model is driven by air temperature, precipitation, wind speed, relative

humidity, and incoming shortwave and longwave radiation at time steps (typically less than 6 hours) sufficient to resolve the diurnal cycle. UEB models the surface snowpack as a single layer to avoid the complexity of over-parameterization. The amount of melt from the snowpack at each time step is driven by the exchange of energy between the snowpack, the atmosphere, and the ground below the snowpack. A modified Force-Restore parameterization of snow surface temperature accounts for differences between snow surface temperature and average snowpack temperature (You, 2004; Luce and Tarboton, 2010). The model evaluates the following energy balance and mass balance equations.

$$\frac{dU_s}{dt} = Q_{sns} + Q_{snl} + Q_{ps} + Q_{hs} + Q_{es} + Q_g - Q_{ms}, \text{ (KJ m}^{-2} \text{ h}^{-1}\text{)} \quad (4.1)$$

$$\frac{dW_s}{dt} = P_r + P_s - i + R_m + M_c + E_s + M_s \text{ (m h}^{-1}\text{)} \quad (4.2)$$

$$\frac{dW_c}{dt} = i - R_m - M_c - E_c \text{ (m h}^{-1}\text{)} \quad (4.3)$$

In Equation (4.1),  $\frac{dU_s}{dt}$  is the change in energy content,  $Q_{sns}$  is the below-canopy net shortwave radiation,  $Q_{snl}$  is the below-canopy net longwave radiation,  $Q_{ps}$  is the advected heat from precipitation,  $Q_{hs}$  is the sensible heat flux, and  $Q_{es}$  is the latent heat flux due to sublimation/condensation,  $Q_g$  is the ground heat flux to the snow, and  $Q_{ms}$  is the advected heat removed by meltwater. In Equations (4.2) and (4.3),  $P_r$  is the rate of precipitation as rain;  $P_s$  is the rate of precipitation as snow,  $i$  is canopy interception,  $R_m$  is mass release from the canopy,  $M_c$  is melt water drip from the canopy snow,  $E_s$  is sublimation from the surface snow,  $M_s$  is melt from the surface snow, and  $E_c$  is sublimation from the canopy snow.  $\frac{dW_s}{dt}$  and  $\frac{dW_c}{dt}$  are the changes in surface and canopy

snow water equivalent, respectively. Intercepted snow energy content is assumed to be negligible, so canopy energy balance is quantified using the following closure equation.

$$Q_{cns} + Q_{cnl} + Q_{pc} + Q_{hc} + Q_{ec} - Q_{mc} = 0 \quad (4.4)$$

In Equation (4.4),  $Q_{cns}$  is the canopy net shortwave radiation exchange,  $Q_{cnl}$  is the canopy net longwave radiation,  $Q_{pc}$  is the net advected heat from precipitation to the canopy,  $Q_{hc}$  is the sensible heat to the canopy,  $Q_{ec}$  is the latent heat to the canopy, and  $Q_{mc}$  is the advected heat removed by melt water from the canopy.

In this study, we extended the UEB snow surface energy balance calculation to compute the generation of melt at the surface of glaciers. Where the substrate is glacier we add an artificial one meter water equivalent of glacier ice beneath any seasonal snowpack (Figure 4.2). Seasonal snow may accumulate and melt on this glacier substrate. When seasonal snowpack disappears, as indicated by the combined snow and ice water equivalent becoming less than one meter, the additional melt is considered to be glacier melt. This calculation is performed at each time step, and the glacier substrate is reset to one meter water equivalent at the beginning of each time step.

Glaciers are generally much thicker than one meter, but the entire depth of a glacier does not interact thermally with the surface energy balance at the time scale of diurnal energy cycles involved in the generation of surface melt. The one meter thickness is assumed to provide a reasonable thermal buffering due to the presence of glacier ice without going into the detail of modeling heat and mass transport within the glacier, which would, at the level of detail of this model, add complexity that is unwarranted, given the other uncertainties involved. Sensitivity analysis showed that as

long as this thickness is greater than 0.7 m, sensitivity to the particular thickness chosen is small.

Glacier substrate determined from remote sensing on the basis of albedo, is encoded using one of four unique values at each grid cell (Figure 4.3). These represent: bare ground or non-glaciated substrate, clean glacier, debris covered glacier, and accumulation zone (i.e., area where snow accumulates and does not contribute to melting), respectively. For non-glaciated surface, UEB evaluates the energy balance equation and computes melting from seasonal snow as it did previously. For clean and debris covered glacier as a substrate the one meter surface ice layer is used. The difference between clean and debris covered glacier consists only in the value of substrate albedo. Ranges of substrate albedo for clean and debris-covered glaciers at Langtang Khola watershed are listed in table 4.1. Due to the lower substrate albedo, debris covered glacier ice will absorb more energy when the seasonal snow cover is thin or absent and will thus generate more melt. For the accumulation zone, UEB is bypassed, since snow and glacier melt is assumed to not occur in these areas; rather, all precipitation is presumed to add mass to the glacier. Accumulation zone should thus only be used for grid cells that are definitively always in the accumulation zone. Transitional grid cells should be represented as glacier to allow the model to then determine, based on inputs, whether snow accumulates indefinitely or contributes to melt.

The model computes surface water input components from rain (SWIR), snowmelt (SWISM), and glacier melt (SWIGM), the sum of which comprise the total surface water input (SWIT), which is the input for streamflow generation:

$$\text{SWIT} = \text{SWIR} + \text{SWISM} + \text{SWIGM} \quad (4.5)$$

Although the model runs separately for each grid cell, outputs can also be aggregated over subwatersheds defined, for example, from a digital elevation model.

#### **4.4.2. Geospatial Stream Flow Model (GeoSFM)**

The USGS Geospatial Streamflow Model (GeoSFM) is a semi-distributed, physically based hydrologic program developed to monitor flood hazard and provide early warning across Africa and other data scarce regions around the globe (Asante *et al.*, 2008). The model has six components: (1) the terrain analysis module, (2) the parameter estimation module, (3) the data preprocessing module, (4) the water balance module, (5) the flow routing module, and (6) the post-processing module. GeoSFM uses a wide range of input data, including digital elevation model (DEM), land cover, and soil data, daily estimates of precipitation and potential evapotranspiration to predict daily streamflow. It includes methods for examination of parameter sensitivity and calibration of parameters using streamflow measured at gage stations.

#### **4.4.3. Model Setup**

The GeoSFM rainfall-runoff model was parameterized using SRTM DEM, MODIS land cover, and United Nations Food and Agriculture Organization (FAO) and United Nations Educational, Scientific and Cultural Organization (UNESCO) soil data (Batjes, 1997). GeoSFM's terrain analysis module was used to delineate subwatersheds and create stream networks to establish the connectivity among various subwatersheds and to compute topographical parameters such as watershed area, slope and river length. The parameter estimation module was used to estimate soil parameters such as water

holding capacity, hydrological active depth, texture, and saturated hydraulic conductivity from the soil data. The land cover data was used to compute impervious area and vegetation roughness for each subwatershed. The soil and land cover data were used together to determine Soil Conservation Service (SCS) runoff curve numbers.

Initial conditions of the UEB's state variables, such as snow water equivalent,  $W_s$  (m), the internal energy of the snowpack and top layer of soil,  $U_s$  (kJ m<sup>-2</sup>), and the age of the snow surface are unknown. These were initialized using a one year spin up period. Input data was available to run the model for 11 years from October 2001 to September 2012. At the beginning of this period state variables were set to zero. Errors due to this assumption diminish with time as the model adjusts to the driving inputs. Outputs from the spin up period 10/1/2001 to 9/30/2002 were discarded and only the output from the period 10/1/2002 to 9/30/2012 was examined in the results. Since measured streamflow was not complete beyond 2010, GeoSFM was only run for the 8 years from 10/1/2003 to 9/30/2010.

Driven by inputs of precipitation and other weather variables UEB models the accumulation and melting of snow, and with the extension introduced above the melting of glaciers. UEB also partitions precipitation into rain and snow and tracks whether this occurs on bare ground, a snow surface or a glacier surface. UEB maintains a mass balance between storage in snow and ice, and inputs and outputs. Its output is the total surface water input (SWIT) and a partitioning of this into components from rain, snow and glacier. The output from UEB is taken as an input to GeoSFM, with GeoSFM

treating SWIT as it would normally treat rainfall, i.e. water applied at the surface of the soil in a drainage basin.

GeoSFM is driven by inputs of the total surface water input (SWIT) and potential evapotranspiration (PET) for each subwatershed. In the application of GeoSFM here PET was calculated using the Hamon method (Hamon, 1961; Hummel *et al.*, 2001) as implemented within EPA BASINS. GeoSFM models hydrologic processes such as infiltration, surface runoff, interflow, baseflow and actual evapotranspiration which may be less than PET due to soil moisture limitations. GeoSFM maintains its own water balance between its representations of surface and subsurface storage and inflows and outflows. GeoSFM was configured to use its nonlinear soil moisture accounting routine that combines the SCS runoff curve number method with the Green-Ampt equation (Green and Ampt, 1911) to compute runoff, interflow and baseflow. Parameters such as curve number and soil hydraulic properties were based on the soil data described above. In GeoSFM the soil water balance and streamflow were estimated using the non-linear soil moisture accounting (Asante *et al.*, 2008) and Muskingum-Cunge method (Ponce and Yevjevich, 1978) options, respectively.

GeoSFM was initially run using the default parameters specified in the model input. With these parameters the model over predicted the peak flow during the summer months and under predicted the flow during low-flow season. Such discrepancies are often associated with parameterization of infiltration, interflow and baseflow. It was also observed that during first few months of the simulation period, the model under-predicted the flow, but then gradually converged towards the observed flow patterns. Specifying a

low value of initial soil moisture may be to blame for this. We performed a sensitivity analysis using the One-At-a-Time Sensitivity Analysis method (Asante *et al.*, 2008). This sensitivity analysis showed that the model was most sensitive to soil water holding capacity (mm), total soil depth (cm) and baseflow reservoir residence time (days). The model was calibrated to match the observed streamflow data by adjusting these three parameters within their plausible ranges using a trial and error approach as suggested by Asante *et al.* (2008). In this trial and error approach the performance of the simulation was evaluated graphically and using the Nash-Sutcliffe Efficiency (NSE) measure based on differences between model results and observed data.

## **4.5. Results**

### **4.5.1. Hydrologic Variability at Three Point Locations on Bare Ground, Clean Glacier and Debris-covered Glacier**

Model output was examined at three points representing “bare ground,” “clean glacier ice,” and “debris covered glacier ice” (Figure 4.4). The specific points (A), (B), and (C) are located in Figure 4.1. Point A (4559 m) is located about 1.6 km north-east of the weather station at Kyangjin, point (B) (6189 m) is on the valley’s southern upper ridges, and Point C (4782 m) is located on the Langtang glacier (Figure 4.1). These points were selected arbitrarily to illustrate the model performance and are not intended to represent aggregate or average conditions on these surfaces throughout the study area.

Major input variables (air temperature, precipitation and incoming shortwave radiation) and major outputs (snow water equivalent, total surface water input, and surface water input from glacier melt) for a complete water year (Oct. 1, 2002 – Sept. 30, 2003) are

shown in Figure 4.4. All three points showed a similar seasonal temperature pattern: high temperature in summer months and low in winter. Mean monthly temperature over the watershed was lowest in February (-14.2 °C) and highest in July (-1 °C). Since temperature is elevation dependent, the highest point (B) consistently showed the lowest temperature of the three points throughout the year. The elevation difference between points (A) and (C) is relatively small, and temperature variation between these two locations was mainly due to the variation in linearly interpolated MERRA data. During winter months, daily maximum temperature at point (C) was higher than that at point (A), and during summer, the opposite trend was observed in Figure 4.4 (a). At point (B), the daily maximum temperature only rose above the freezing point for a few days during the summer months (i.e. mid-June through August), and daily average temperature always remained below the freezing point. On the other hand, the daily maximum temperatures at points (A) and (C) only fell below the freezing point during the winter months (i.e. Dec through May).

Over the watershed, annual precipitation varied from 0.51 m in the driest year to 2.42 m in the wettest year. Our results showed that over 70% (0.68 m) of mean annual precipitation (1.07 m) in Langtang Khola occurs during the monsoon season (May through September). Among the three points, (A) received the highest amount of precipitation (1.86 m for Oct, 2002 – Sep, 2003), while (C) at similar elevation received the least (1.41 m for this water year). The highest selected point (B) received about 1.59 m of precipitation for this year, which was comparable to the average precipitation over the watershed (1.65 m). One possible explanation for point (C) receiving the least

precipitation is its location at the bottom of a valley. The northward movement of the monsoon wind is blocked by the east-west mountain ridges to the south, so the bottom of the valley to the north of this ridge receives less precipitation (Shiraiwa *et al.*, 1992; Konz *et al.*, 2007). These spatial variations in precipitation are relatively small, due in part to the close spatial proximity of these locations relative to the spatial scale of precipitation variability at an annual scale and also due to the coarseness of the RFE2 data (0.1 deg grid scale) that determines this variability.

Figure 4.4 (c) showed that incoming daily maximum shortwave radiation increased during January, peaks during March and April, and dropped during the summer months (i.e. July-September). This may seem surprising, but it is probably due to the lower daily temperature range during these monsoon months with considerable cloud cover. For example, average diurnal temperature ranges were 11.2°C and 7.1°C for January and August 2003, respectively. Incoming solar radiation was least affected by the atmosphere at point (B) and most affected at point (A). Therefore, point (A) received the least amount of shortwave radiation and (B) receives the highest. However, this variation was also season dependent. Shortwave radiation at point B peaks over 1400 KJ hr<sup>-1</sup> m<sup>-2</sup> during March and April and drops to a maximum daily value of about 800 KJ hr<sup>-1</sup> m<sup>-2</sup> during summer. Unlike point (B), points (A) and (C) showed smaller seasonal variations in incoming solar radiation.

The plot for snow water equivalent in Figure 4.4 (d) shows that snow accumulation and melt at points (A) and (C) show similar patterns, although peak snow accumulation at point (A) (1.03 m) is 2.78 times higher than that at point C (0.37 m), due

to the higher precipitation at (A). The snow accumulation rate was low from November to January due to low winter precipitation. From February to May, snow accumulated faster. Beginning in June, snow started disappearing at a fast rate at both points. Due to higher accumulation, snow at point (A) takes a longer time to melt. At point B, 0.59 m of snow remained on the glacier surface from the previous year's simulation. This then increased rapidly during October and then remained nearly unchanged until the start of the monsoon, which brought an additional 0.77 m snow accumulation during the five months from March through July. From August to September there was only a small amount of ablation due to snowmelt, indicating that (B) is a point where the primary effect of precipitation is contribution to the accumulation of glacier ice. Other high elevation points behaved similarly. At point (A), surface water input was generated from May to October, with the highest rate during the snowmelt season (Figure 4.4 (e)). At Point (C), on the other hand, total surface water input was generated from all three possible sources: rain fall, snowmelt, and glacier melt. Precipitation in the form of rain occurred mostly during May to October; a large amount of glacier melt also occurred during that time. Thus, we can conclude that, during the monsoon, the combination of rain, snowmelt, and glacier melt contribute significant surface water inputs at lower elevations, while the monsoon snow accumulated and added to glacier ice at high elevations.

#### **4.5.2. Total Surface Water Input, Observed and Simulated Streamflow**

Figure 4.5 compares total surface water input (SWIT) and observed streamflow. Mean daily SWIT and daily observed streamflow were highly correlated (0.80) and

SWIT captured the seasonal pattern of the streamflow, both showing high values during the Monsoon and low values in winter (Figure 4.5 a). For any particular year, SWIT was higher than the streamflow during the monsoon and lower during the winter. Since there is little melt in winter, baseflow is the most likely source of streamflow in the winter months (Immerzeel *et al.*, 2012). Streamflow gradually receded from October to March due to groundwater recession. As the summer approaches, SWIT slowly starts to increase, and soon after, the increase also appeared in streamflow. Overall cumulative SWIT was more than cumulative streamflow (Figure 4.5 b), reflecting hydrologic losses, primarily evapotranspiration of about 0.5 m per year. Hydrologic losses that occur during the transformation between SWIT and streamflow also explain why the peaks in Figure 4.5 a are higher than streamflow.

Figure 4.6 compares observed and simulated streamflow obtained as output from the GeoSFM model. Three performance indicators: (1) Root mean square error to the standard deviation of measured data ratio (RSR) (2) Nash-Sutcliffe efficiency (NSE) and (3) Percent bias (PBIAS) described by Moriasi *et al.* (2007) were calculated using time series of observed and simulated streamflow (Table 4.2). The general performance guideline recommended by (Moriasi *et al.*, 2007) indicates that  $NSE \leq 0.5$ , 0.5 to 0.65, 0.65 to 0.75 and  $\geq 0.75$  are considered as “poor,” “satisfactory,” “good” and “very good” ratings, respectively.  $RSR \leq 0.5$  and 0.5 to 0.6 are considered as “good” and “very good” ratings, respectively. Similarly, the guideline also specifies PBIAS less than 10% is considered as “very good” (Moriasi *et al.*, 2007). The combined (UEB and GeoSFM)

model's ability to simulate daily streamflow may thus be interpreted as "good" and the ability to simulate monthly streamflow as "very good."

We examined the mass balance of the watershed and the partitioning of input precipitation into surface water input, storage and flux components. UEBGrid quantified the snow mass balance error (Equation 4.6) while the watershed mass balance is presented in Equation (4.7). Although the model was run for ten years, the lack of streamflow data in 2011 and 2012 limited the watershed mass balance to the 8 years from 2003 to 2010 (Table 4.3).

$$\text{Mass Balance Error} = \text{Precipitation (P)} - \text{Surface water input from snow melt (SWISM)} - \text{Surface water input from rain (SWIR)} - \text{Snow Sublimation (E)} - \text{Snow Accumulation (SWE)} \quad (4.6)$$

$$\text{Change in storage } (\Delta S) = \text{Total surface water input (SWIT)} - \text{Streamflow (q)} - \text{Evapotranspiration Loss (ET)} \pm \text{Groundwater exchange } (\Delta G) \quad (4.7)$$

During the 10-year simulation period, precipitation was 9.53 m, SWISM was 4.23 m, SWIR was 1.1 m, sublimation (E) was 2.02 m, and snow accumulation (SWE) was 2.46 m. In Equation (4.6), the snow mass balance error was -0.28 m, which was about 3% of the total precipitation and is therefore negligible. During this period, 7.13 m of glacier melted while there was 2.46 m of snow accumulation indicating a net glacier loss of 4.67 m over the watershed. In the watershed mass balance (Equation 7), UEBGrid estimated the SWIT over 8 years (9.43 m), while the hydrologic station measures streamflow (q) (5.02 m). This leaves a residual of 4.41 m (i.e. 47% of SWIT) that represents the combined effect of evaporation and transpiration from non-glacier and snow-covered

areas and net changes in storage either in the remaining snow/glacier ice or groundwater. The GeoSFM rainfall-runoff model simulated an 8 year streamflow of 5.31 m which is in reasonable agreement with the measured flow (5.02 m).

Figure 4.7 showed yearly variations in streamflow and the components of SWIT. Apart from 2011, SWIT remained relatively stable between years, varying between 1.0 and 1.3 m. Rain contributed the least of all three components (about 8%) of SWIT and did not vary substantially from year to year. Glacier melt and snowmelt generated the rest of the SWIT (92%). In 2011, SWIR (0.27 m) and SWISM (0.96 m) were unusually high compared with mean yearly values (0.11 m and 0.42 m, respectively), and these two components, along with SWIGM, cause a very high SWIT in that year. From 2004 to 2009, observed annual streamflow was relatively constant and then in 2010, observed annual streamflow was 28% higher than the average although SWIT is a little over the average yearly value. This may be due to a higher fraction of glacier melt in total surface water input (SWIT), but it is unclear why observed streamflow increases in only 2010 while high fractions of glacier melt are simulated for three years (2008-2010). The difference between the observed and simulated total annual streamflow showed a strong positive correlation (0.78) with % contribution of yearly mean glacier melt in SWIT indicating either overestimation of glacier melt by UEB in later years or GeoSFM's inability to simulate streamflow in a year when glacier contribution is relatively higher or combination of these two.

Figure 4.8 depicts monthly variation of SWIT, streamflow, sublimation, and temperature averaged over the 10 years of simulation. This showed the seasonal cycle

with high values during summer and low in winter. SWIT and temperature peak in July while streamflow and sublimation peak in August. Sublimation is also slightly higher during July and August than the rest of the year. From December through March, streamflow is higher than SWIT, indicating that the majority of the streamflow is generated from baseflow contribution. Generally, seasonality of measured streamflow is well captured by the GeoSFM simulation.

In Figure 4.9, the spatial sources of surface water input were examined. This figure showed that debris covered glacier, representing only 8% of the basin area, was responsible for 49% of the surface water input over the watershed; of which 43%, 4% and 2% were generated from glacier, snow and rain, respectively. Seventy seven percent of Rain occurred on bare ground, which although 57% of the area contributed only 7% of SWIT. Melting of snow on bare ground contributed 26% of total SWIT.

Figure 4.10 examines correlations among annual variables such as SWIT, precipitation, observed and simulated streamflow, sublimation, and temperature. Precipitation is positively correlated with SWIT and negatively correlated with sublimation and temperature. A negative correlation between precipitation and temperature indicates that a typical wet year is a relatively colder. Sublimation showed a modest positive correlation with temperature, indicating that increases in temperature increase the sublimation. Surprisingly, both observed and simulated streamflow was not correlated with precipitation; however, both of these quantities showed a modest positive correlation with SWIT. Glacier melt is strongly negatively correlated with precipitation (-0.89). A possible reason for this is that during a typical wet year snow accumulation is

high on the glacier surface and melting of the seasonal snow takes longer so the glacier surface becomes exposed for melting later part of the melting season.

#### **4.6. Discussion**

In this study, an energy balance model driven by downscaled weather inputs was able to represent glacier melt and drive simulations of streamflow that compare reasonably well to the observations (i.e., NSE = 0.72 at daily scale). The differences noted may be due to limitations and uncertainties such as modeling the melt from debris covered glacier ice, uncertainty in reanalysis inputs and errors due to processes not represented or simplified in the model.

Melt from a debris-covered glacier is a complicated process due to the interplay of albedo and debris thickness in glacier melting process. Our model may overestimate melt from glacier due to debris inhibiting the surface energy exchanges where there is a thick debris-layer present on a glacier (Brock *et al.*, 2010). On the other hand, presence of a thin debris layer may increase the melt, as debris-cover has low albedo and thus absorbs more energy (Kayastha *et al.*, 2000). Due to lack of observational data on debris thickness (Mihalcea *et al.*, 2008a; 2008b; Foster *et al.*, 2012), we were unable to include debris thickness into the model. Lack of ability to quantify these effects introduces uncertainty into the melt from debris-covered glaciers.

MERRA and RFE2 reanalysis input data are available at coarse scale and may have biases and downscaling errors. Temperature was adjusted using a constant linear monthly lapse rate. While the constant linear lapse rate method most successfully reproduces temperature variability (Dodson and Marks, 1997; Liston and Elder, 2006), in

reality lapse rate varies in both space and time. Lapse rate also depends on other factors, such as hill shading, complex heat balance on debris-cover glaciers, and the diurnal pattern of lapse in the Himalayas (Fujita and Sakai, 2000). For precipitation, we assumed a uniform precipitation rate throughout the day to obtain three hourly model inputs from daily RFE2 precipitation estimates. While calculating shortwave and longwave radiation, we used a standard elevation pressure relationship to compute pressure at a point and we assumed that the attenuation coefficient was constant. This ignores cloud variability that may affect solar radiation at a particular time step. There may also be uncertainty in the downwelling longwave radiation where we used the Stefan-Boltzmann equation and emissivity assumptions provided in Liston and Elder (2006). All these assumptions incorporate uncertainty at various degrees to the model forcing data. Calibration of streamflow may have compensated for some reanalysis data issues and downscaling errors which together with streamflow being an aggregate quantity results in reasonably good streamflow simulations in spite of aforesaid uncertainties.

There are a number of processes not modeled or simplified in the model. The model does not represent glacier movement and the presence or absence of glacier is treated as a fixed input precluding the ability to model retreat of glaciers that melt completely. We assumed a 1 m thick glacier layer that interacted thermally with the seasonal snowpack and was used to calculate glacier melt. While the sensitivity of this assumption was evaluated, it has not been tested versus observations.

There may also be errors in the measurements used to evaluate the model. The streamflow data used was based on daily measurements (one measurement/day) at around

9:00 AM and published as a daily mean discharge. This procedure may produce systematic error in the discharge measurement as it is unable to capture the diurnal streamflow cycles resulting from diurnal melt inputs. Given all the potential sources of error, it is actually encouraging how good the degree of agreement between observed and simulated streamflow in Figure 4.6 actually is with NSE of 0.72 for daily flows and NSE of 0.78 for monthly flows and peaks that are typically within 15 % of observed. This means that a model such as this can be used with this degree of confidence to quantify the streamflow from similar, but ungauged glaciated basins for planning and water resources management purposes. The sensitivity to changes in glacier area and climate can be quantified to help inform water resources planning in the future.

#### **4.7. Conclusions**

In this paper, we also introduced a simple way to represent glaciers using glacier outlines and albedo maps in a physically based energy balance model. We applied the model at the Langtang Khola watershed for a ten-year simulation period (water years 2003 to 2012). The model showed that in the highly glaciated sub-basin, precipitation is a not a strong predictor of streamflow. During a dry year, surface water input is compensated by the glacier melt component which contributes to steady flow of water in the streams; therefore, glaciers serve as an important source of water to the outflow from the Langtang Khola basin. The model estimates an average of 0.71 m glacier ice melt per year. Glacier mass balance showed a negative trend, with an average snow accumulation of 2.46 m compared to 7.13 m of glacier melt during water year 2003 to 2012.

In this study the only ground-based observational data used was streamflow. Without additional ground based measurements, results from this study should be considered to be an initial baseline to provide hydrological insights on the hydrology and water balance of the Langtang Khola watershed. These insights would be further strengthened if there was an opportunity to compare additional outputs measurements. Rohrer *et al.* (2013) advocated the idea of development and maintenance of stations to collect snow related variables in the Himalayas. We agree that there is a need to establish high altitude stations with weather, snow and glacier melt measuring capability for model testing and improvement.

## References

- Asante KO, Artan GA, Pervez S, Bandaragoda C, Verdin JP. 2008. Technical manual for the Geospatial Stream Flow Model (GeoSFM). U. S. Geological Survey.
- Batjes N. 1997. A world dataset of derived soil properties by FAO–UNESCO soil unit for global modelling. *Soil Use and Management* **13**: 9-16.
- Brock BW, Mihalcea C, Kirkbride MP, Diolaiuti G, Cutler ME, Smiraglia C. 2010. Meteorology and surface energy fluxes in the 2005–2007 ablation seasons at the Miage debris-covered glacier, Mont Blanc Massif, Italian Alps. *Journal of Geophysical Research: Atmospheres* (1984–2012) **115**: 106.
- Brown M, Racoviteanu A, Tarboton D, Gupta AS, Nigro J, Policelli F, Habib S, Tokay M, Shrestha M, Bajracharya S. 2014. An integrated modeling system for estimating glacier and snow melt driven streamflow from remote sensing and earth system data products in the Himalayas. *Journal of Hydrology* **519**: 1859–1869.
- Brown ME, Ouyang H, Habib S, Shrestha B, Shrestha M, Panday P, Tzortziou M, Policelli F, Artan G, Giriraj A. 2010. HIMALA: Climate Impacts on Glaciers, Snow, and Hydrology in the Himalayan Region. *Mountain Research and Development* **30**: 401-404.
- Dickinson RE, Henderson-Sellers A, Kennedy PJ. 1993. Biosphere-Atmosphere Transfer Scheme (BATS) Version 1e as Coupled to the NCAR Community Climate Model. National Center for Atmospheric Research, Boulder, CO, p. 71.

- Dodson R, Marks D. 1997. Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Climate Research* **8**: 1-20.
- EPA BV. 1998. 2.0 User's manual: better assessment science integrating point and nonpoint sources. US Environmental Protection Agency, Washington, DC.
- Foster L, Brock B, Cutler ME, Diotri F. 2012. A physically based method for estimating supraglacial debris thickness from thermal band remote-sensing data. *Journal of Glaciology* **58**: 677-691.
- Fujita K, Sakai A. 2000. Air temperature environment on the debris-covered area of Lirung Glacier, Langtang Valley, Nepal Himalayas. *IAHS PUBLICATION*: 83-88.
- Gardelle J, Berthier E, Arnaud Y. 2012. Slight mass gain of Karakoram glaciers in the early twenty-first century. *Nature Geoscience* **5**: 322-325.
- Gardelle J, Berthier E, Arnaud Y, Kääb A. 2013. Region-wide glacier mass balances over the Pamir-Karakoram-Himalaya during 1999-2011. *Cryosphere* **7**.
- Green WH, Ampt G. 1911. Studies of Soil Physics. Part I - The flow of air and water through soils. *The Journal of Agricultural Science* **4**: 1-24.
- Hamon RW. 1961. Estimating potential evapotranspiration. In Proceedings of the American Society of Civil Engineers, *Journal of the Hydraulic Division*, pp: 107-120.
- Hummel P, Kittle Jr J, Gray M, Consultants AT, Decatur G, No WA, Wellman M. 2001. A Tool for Managing Watershed Modeling Time-Series Data. AQUA TERRA Consultants, Decatur, GA, p. 157.
- Immerzeel W, Petersen L, Ragettli S, Pellicciotti F. 2014. The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas. *Water Resources Research* **50**: 2212-2226.
- Immerzeel WW, Van Beek L, Konz M, Shrestha A, Bierkens M. 2012. Hydrological response to climate change in a glacierized catchment in the Himalayas. *Climatic Change* **110**: 721-736.
- Immerzeel WW, van Beek LP, Bierkens MF. 2010. Climate change will affect the Asian water towers. *Science* **328**: 1382-1385.
- Kaser G, Großhauser M, Marzeion B. 2010. Contribution potential of glaciers to water availability in different climate regimes. *Proceedings of the National Academy of Sciences* **107**: 20223-20227.

- Kayastha RB, Ageta Y, Fujita K. 2005. Use of positive degree-day methods for calculating snow and ice melting and discharge in glacierized basins in the Langtang Valley. *Climate and Hydrology of Mountain Areas* **7**: 7-14.
- Kayastha RB, Takeuchi Y, Nakawo M, Ageta Y. 2000. Practical prediction of ice melting beneath various thickness of debris cover on Khumbu Glacier, Nepal, using a positive degree-day factor. *IAHS PUBLICATION*: 71-82.
- Kehrwald NM, Thompson LG, Tandong Y, Mosley-Thompson E, Schotterer U, Alfimov V, Beer J, Eikenberg J, Davis ME. 2008. Mass loss on Himalayan glacier endangers water resources. *Geophysical Research Letters* **35**.
- Konz M, Finger D, Buergi C, Normand S, Immerzeel W, Merz J, Giriraj A, Burlando P. 2010. Calibration of a distributed hydrological model for simulations of remote glacierised Himalayan catchments using MODIS snow cover data. In: *Global Change: Facing Risks and Threats to Water Resources*, p. 465-473.
- Konz M, Uhlenbrook S, Braun L, Shrestha A, Demuth S. 2007. Implementation of a process-based catchment model in a poorly gauged, highly glacierized Himalayan headwater. *Hydrology and Earth System Sciences* **11**: 1323-1339.
- Liston GE, Elder K. 2006. A meteorological distribution system for high-resolution terrestrial modeling (MicroMet). *Journal of Hydrometeorology* **7**: 217-234.
- Lucchesi R. 2012. File Specification for MERRA Products. GMAO Office Note No. 1 (Version 2.3). 82 pp.
- Luce CH, Tarboton DG. 2010. Evaluation of alternative formulae for calculation of surface temperature in snowmelt models using frequency analysis of temperature observations. *Hydrology and Earth System Sciences* **14**: 535-543.
- Mahat V, Tarboton DG. 2012. Canopy radiation transmission for an energy balance snowmelt model. *Water Resources Research* **48**: W01534.
- Mahat V, Tarboton DG. 2013. Representation of canopy snow interception, unloading and melt in a parsimonious snowmelt model. *Hydrological Processes* **in press**.
- Mahat V, Tarboton DG, Molotch NP. 2013. Testing above and below canopy representations of turbulent fluxes in an energy balance snowmelt model. *Water Resources Research* **49**: 1107-1122.
- Mihalcea C, Brock B, Diolaiuti G, D'Agata C, Citterio M, Kirkbride M, Cutler M, Smiraglia C. 2008a. Using ASTER satellite and ground-based surface temperature measurements to derive supraglacial debris cover and thickness patterns on Miage Glacier (Mont Blanc Massif, Italy). *Cold Regions Science and Technology* **52**: 341-354.

- Mihalcea C, Mayer C, Diolaiuti G, D'agata C, Smiraglia C, Lambrecht A, Vuillermoz E, Tartari G. 2008b. Spatial distribution of debris thickness and melting from remote-sensing and meteorological data, at debris-covered Baltoro glacier, Karakoram, Pakistan. *Annals of Glaciology* **48**: 49-57.
- Moriasi D, Arnold J, Van Liew M, Bingner R, Harmel R, Veith T. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions ASABE* **50**: 885-900.
- Ponce VM, Yevjevich V. 1978. Muskingum-Cunge method with variable parameters. *Journal of the Hydraulics Division* **104**: 1663-1667.
- Racoviteanu AE, Armstrong R, Williams MW. 2013. Evaluation of an ice ablation model to estimate the contribution of melting glacier ice to annual discharge in the Nepal Himalaya. *Water Resources Research* **49**: 5117-5133.
- Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim G-K. 2011. MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate* **24**: 3624-3648.
- Rohrer M, Salzmann N, Stoffel M, Kulkarni AV. 2013. Missing (in-situ) snow cover data hampers climate change and runoff studies in the Greater Himalayas. *Science of The Total Environment* **468**: S60-S70.
- Shiraiwa T, UENO K, Yamada T. 1992. Distribution of mass input on glaciers in the Langtang Valley, Nepal Himalayas. *Bulletin of Glacier Research* **10**: 21-30.
- Shrestha M, Artan G, Bajracharya S, Sharma R. 2008. Using satellite-based rainfall estimates for streamflow modelling: Bagmati Basin. *Journal of Flood Risk Management* **1**: 89-99.
- Shrestha MS, 2011, Bias-adjustment of satellite-based rainfall estimates over the central Himalayas of Nepal for flood prediction, PhD Thesis, Kyoto University, Japan.
- Suarez MJ, Rienecker M, Todling R, Bacmeister J, Takacs L, Liu H, Gu W, Sienkiewicz M, Koster R, Gelaro R. 2008. The GEOS-5 Data Assimilation System- Documentation of Versions 5.0, 5.1, 0, and 5.2. 0.
- Sulla-Menashe D, Friedl M. 2007. MODIS Collection 5 Land Cover Type and Land Cover Dynamics: Algorithm Refinements and Early Assessment. AGU Fall Meeting; p. 34.
- Tarboton DG, Chowdhury TG, Jackson TH. 1995. A Spatially Distributed Energy Balance Snowmelt Model. In: Biogeochemistry of Seasonally Snow-Covered Catchments (Proceedings of a Boulder Symposium, July 1995), Tonnessen KA, Williams MW, Tranter M (eds) IAHS Publ. no. 228, 141-155.

- Tarboton DG, Luce CH. 1996. Utah Energy Balance Snow Accumulation and Melt Model (UEB). Utah Water Research Laboratory and USDA Forest Service Intermountain Research Station, Logan, UT, p. 64.
- Viviroli D, Weingartner R. 2004. The hydrological significance of mountains: from regional to global scale. *Hydrology and Earth System Sciences Discussions* **8**: 1017-1030.
- Wu W-S, Purser RJ, Parrish DF. 2002. Three-dimensional variational analysis with spatially inhomogeneous covariances. *Monthly Weather Review* **130**: 2905-2916.
- Xie P, Yarosh Y, Love T, Janowiak J, Arkin PA. 2002. A real-time daily precipitation analysis over South Asia. 16th conference on hydrology, American Meteorological Society, p. 3.15.
- You J, 2004, Snow Hydrology: The parameterization of subgrid processes within a physically based snow energy and mass balance model, PhD Thesis, Civil and Environmental Engineering, Utah State University, Logan, UT.

Table 4.1. Physical characteristics of the Langtang Khola Watershed

Substrate Type	Elevation (m)			Albedo			watershed area %
	Minimum	Maximum	Mean	Minimum	Maximum	Mean	
Bare ground	3700	6800	4882	0.09	0.74	0.26	57
Debris-cover glacier	3997	5554	4823	0.15	0.71	0.25	8
Clean glacier	4390	7104	5670	0.17	0.87	0.55	35

Table 4.2. Summary statistics of performance criteria between observed and simulated streamflow

Performance Criteria	Equation	Daily simulation	Monthly simulation
Root mean square error to the standard deviation of measured data ratio (RSR)	$RSR = \frac{\sqrt{\sum_{i=1}^n (Obs_i - Sim_i)^2}}{\sqrt{\sum_{i=1}^n (Obs_i - Obs_{mean})^2}}$	0.53 (good)	0.47 (very good)
Nash-Sutcliffe efficiency (NSE)	$NSE = 1 - \frac{\sum_{i=1}^n (Obs_i - Sim_i)^2}{\sum_{i=1}^n (Obs_i - Obs_{mean})^2}$	0.72 (good)	0.78 (very good)
Percent bias (PBIAS)	$PBIAS = \frac{\sum_{i=1}^n (Obs_i - Sim_i)}{\sum_{i=1}^n Obs_i} \times 100$	6.1 % (very good)	6% (very good)

Table 4.3. Mass balance components.

Years	Precipitation (m)	Surface water input from glacier melt (m)	Surface water input from snow melt (m)	Surface water input from rain (m)	Total surface water input (m)	Snow Water Equivalent (m)	Sublimation (m)	Observed streamflow (m)	Simulated streamflow (m)
2003	1.65	0.41	0.8	0.11	1.32	0.57	0.18	0.64	0.66
2004	0.78	0.72	0.37	0.11	1.2	0.11	0.21	0.67	0.62
2005	0.59	0.66	0.26	0.09	1.0	0.08	0.18	0.59	0.51
2006	0.8	0.74	0.33	0.1	1.17	0.17	0.22	0.59	0.57
2007	0.91	0.65	0.4	0.1	1.16	0.25	0.18	0.58	0.56
2008	0.59	0.85	0.25	0.08	1.17	0.09	0.21	0.58	0.69
2009	0.51	0.91	0.21	0.07	1.18	0.06	0.21	0.55	0.73
2010	0.53	0.95	0.2	0.09	1.23	0.08	0.2	0.82	0.97
2011	2.42	0.55	0.96	0.27	1.78	1.02	0.21	NA	NA
2012	0.75	0.69	0.45	0.08	1.21	0.03	0.22	NA	NA
Total	9.53	7.13	4.23	1.1	12.42	2.46	2.02	5.02	5.31
Mean	0.95	0.71	0.42	0.11	1.24	0.25	0.2	0.63	0.66

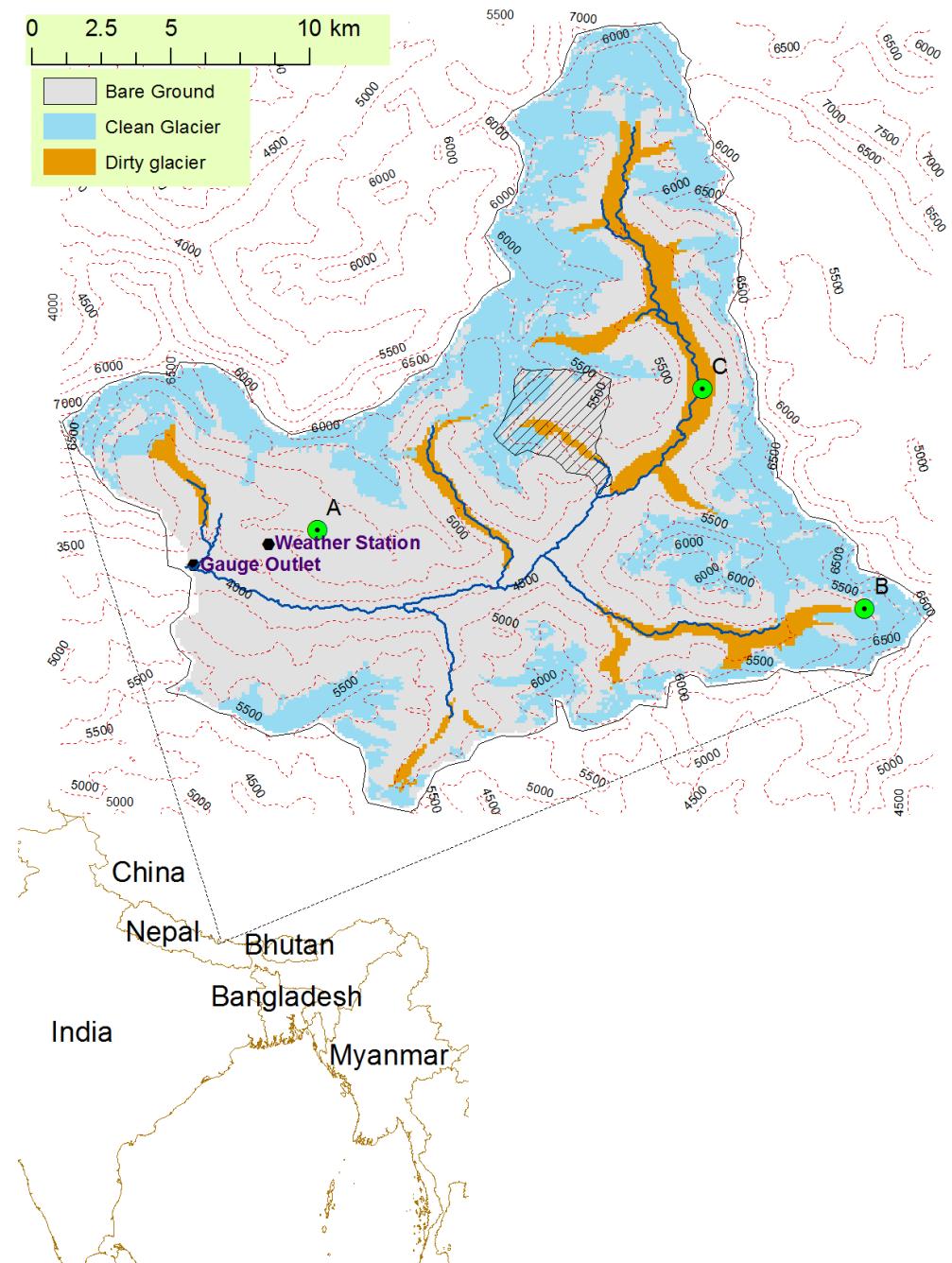


Figure 4.1. Location of Nepal and Langtang Khola watershed in South Asia. Bare ground, clean and debris covered glacier ice are shown in antique white, sky blue and dark orange colors, respectively. Three selected points in three substrates (A at bare ground, B at clean glacier and C at debris covered glacier) are shown by green points. Hatched area is the subwatershed depicted in Figure 4.3.

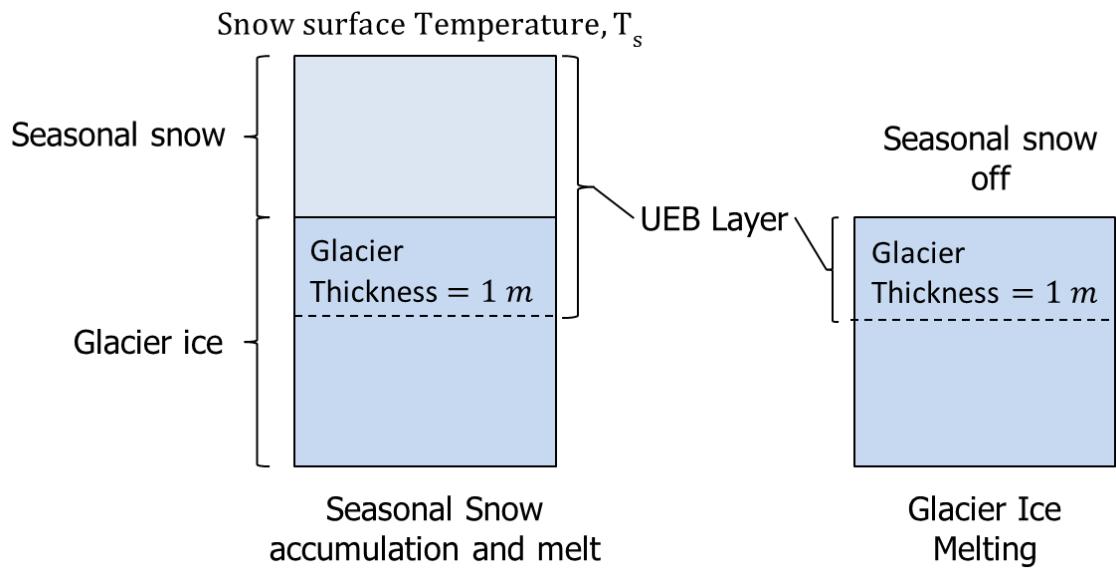


Figure 4.2. Glacier representation in UEBGrid.

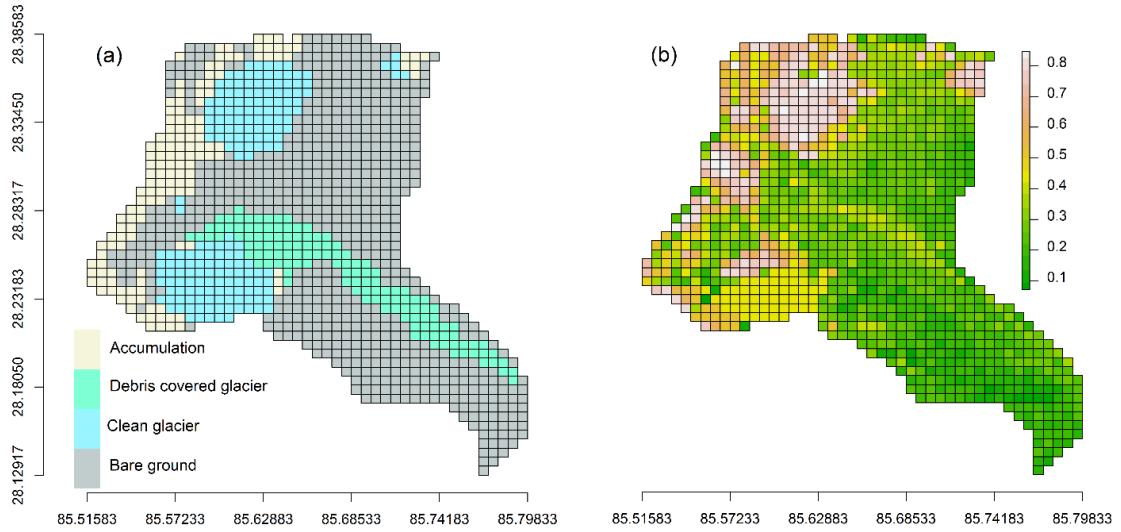


Figure 4.3. (a) Substrate type and (b) substrate albedo. In Figure 4.1, this subwatershed of Langtang Khola is shown as the hatched area.

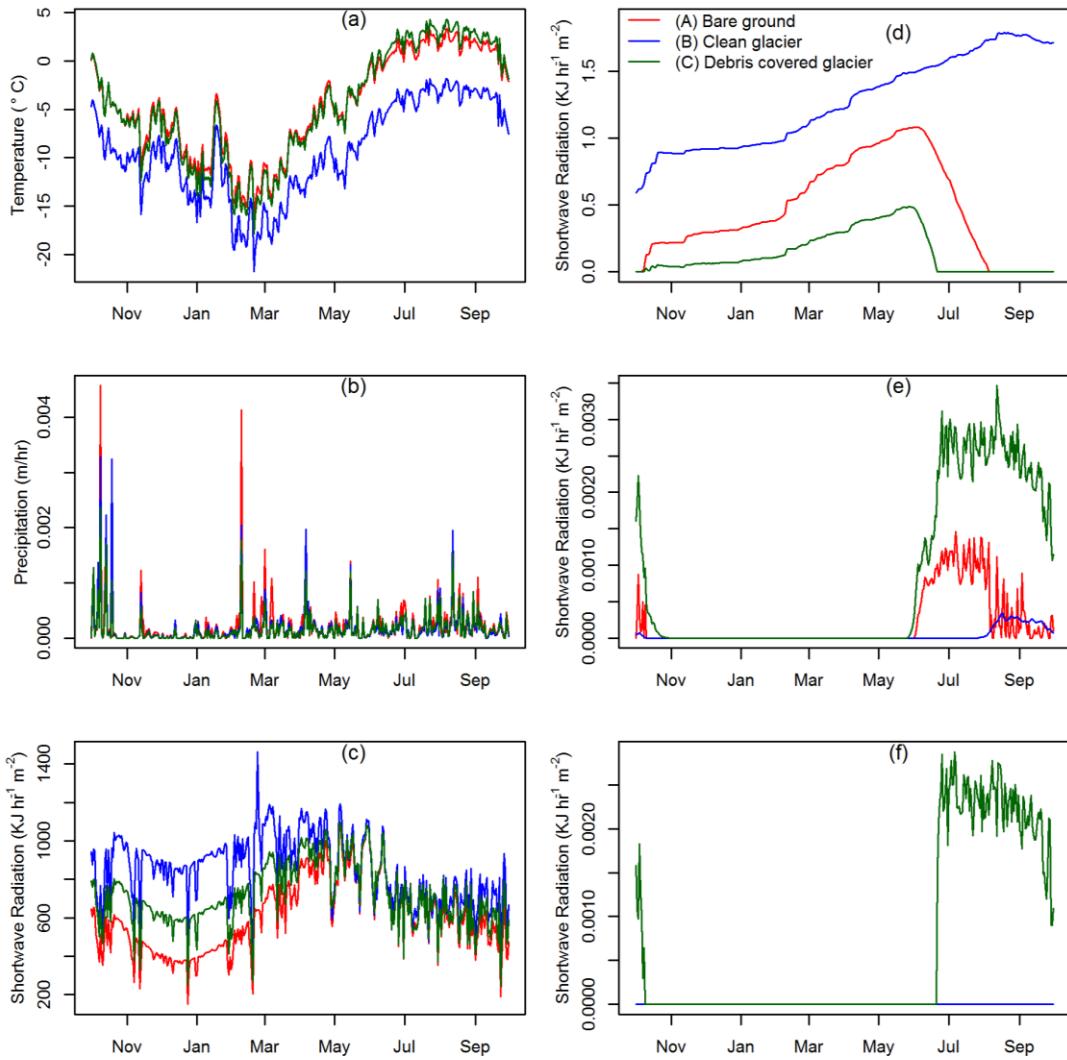


Figure 4.4. Inputs (i.e. (a) temperature ( $^{\circ}\text{C}$ ), (b) precipitation (m/hr) and (c) solar radiation ( $\text{KJ m}^{-2} \text{hr}^{-1}$ )) and outputs (i.e. (d) SWE (m), (e) Total Surface water input (SWIT, m/hr) and (f) glacier melt (SWIGM, m/hr)) at three different points (A: bare ground, B: Clean glacier, C: Debris-covered glacier) aggregated (i.e., mean) at a daily scale for the year 10/1/2002-9/30/2003, the first year after the spin up period. Color legend shown in (d) for bare ground (red), clean (blue) and debris covered glacier (green) remains unchanged for other figures (i.e., a through f).

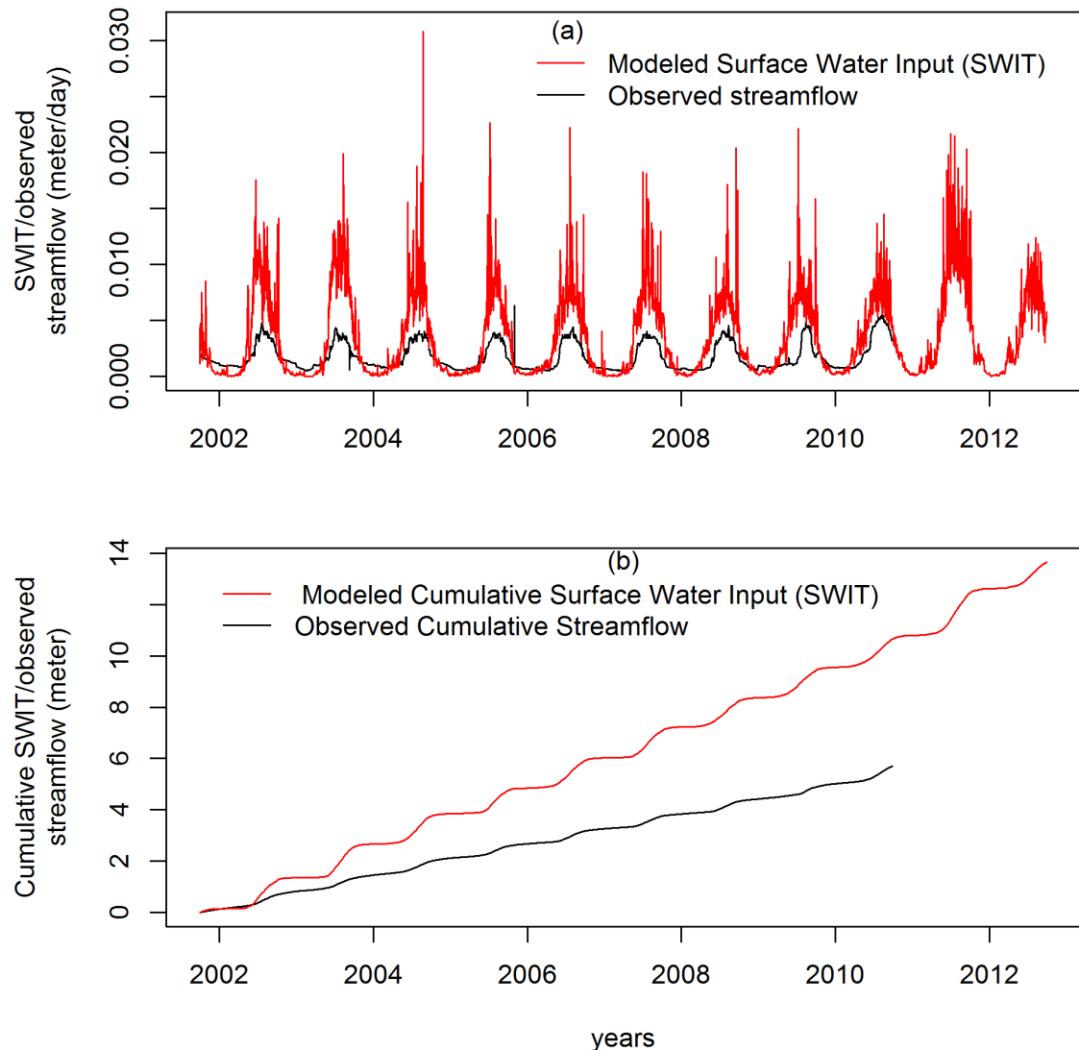


Figure 4.5. Comparison between melt estimated by UEBGrid and streamflow measured at Kyangjin hydrologic station.

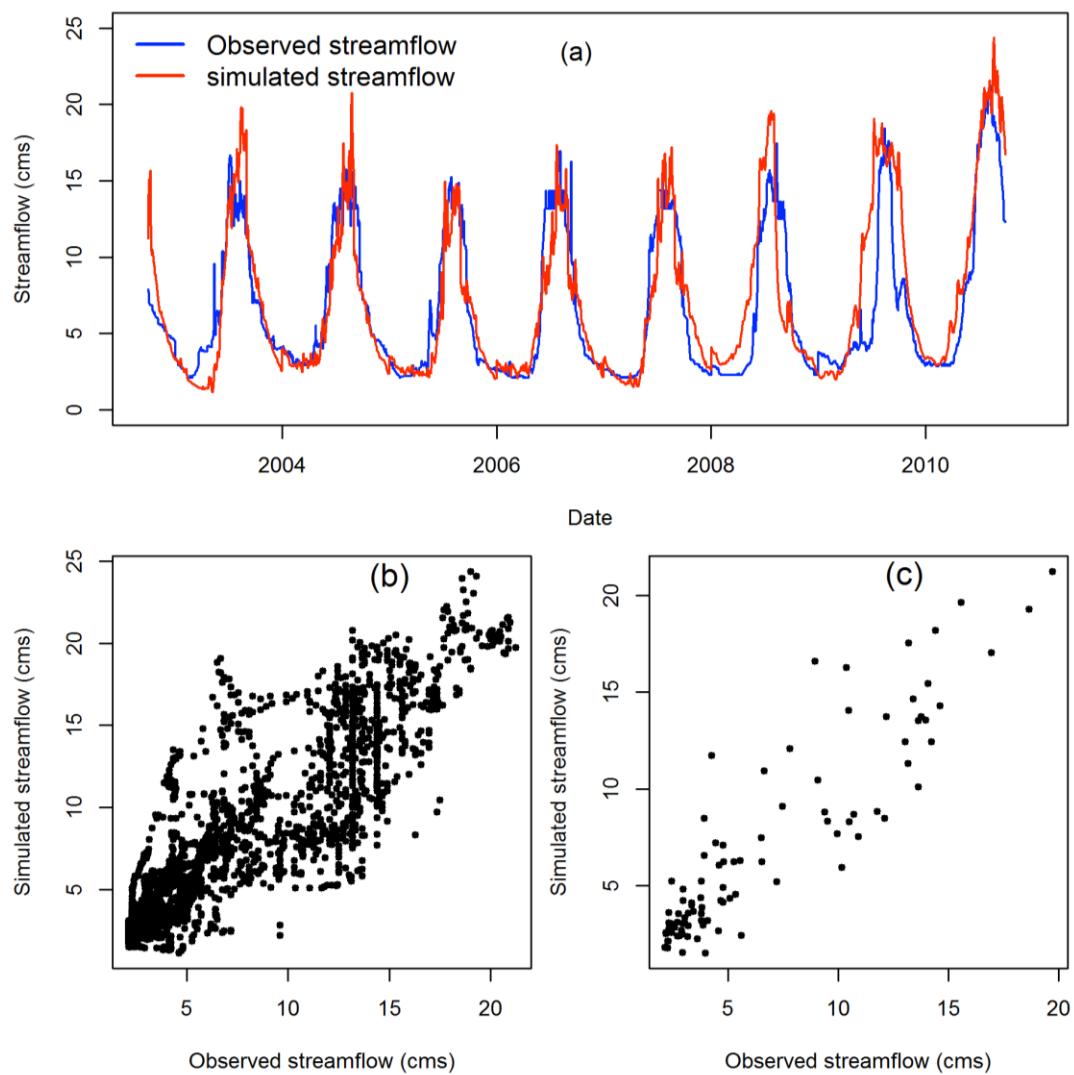


Figure 4.6. Comparison between observed and GeoSFM simulated streamflow driven by UEBGrid modeled surface water inputs. (a) Daily time series of observed and simulated streamflow, (b) scatter plot of daily observed and simulated streamflow, (c) monthly mean observed and simulated streamflow.

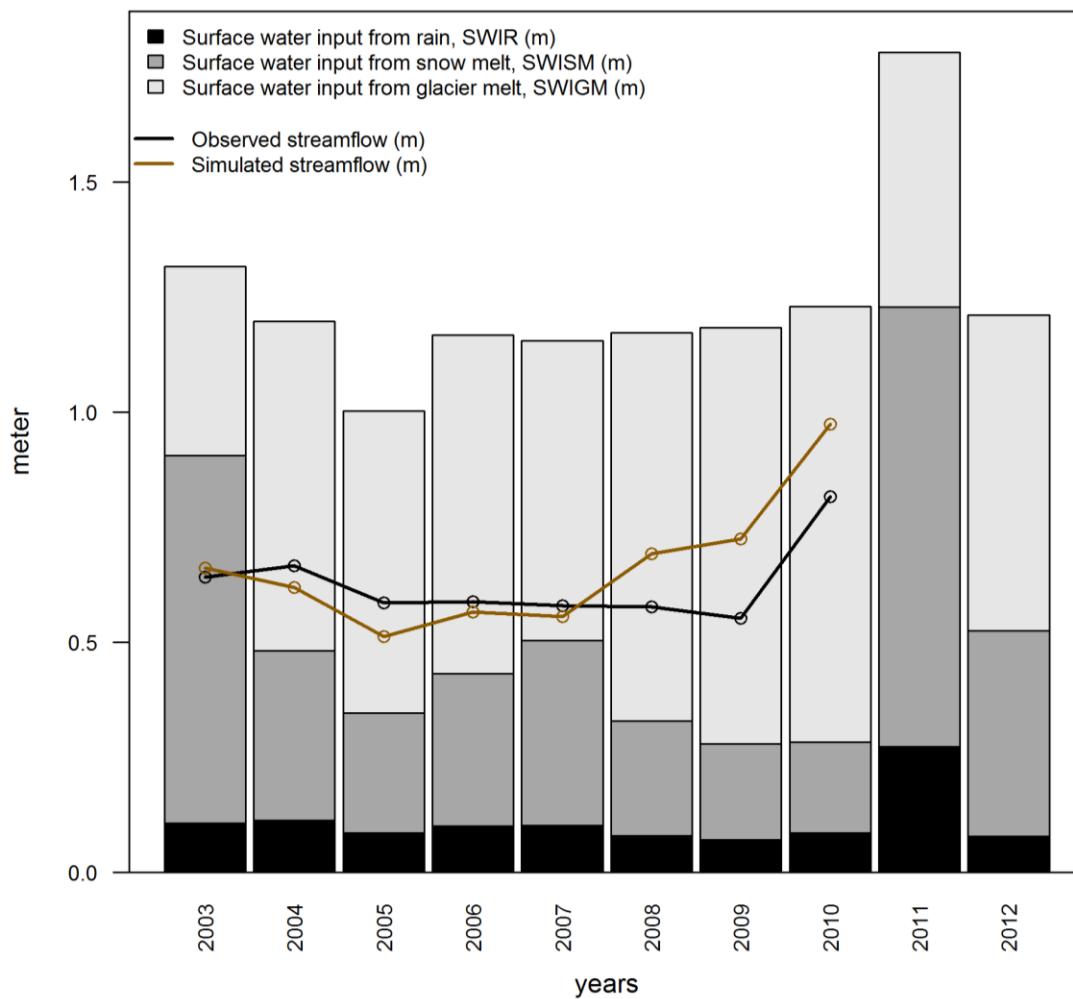


Figure 4.7. Yearly contribution of glacier, snow, and rain in total surface water input (SWIT).

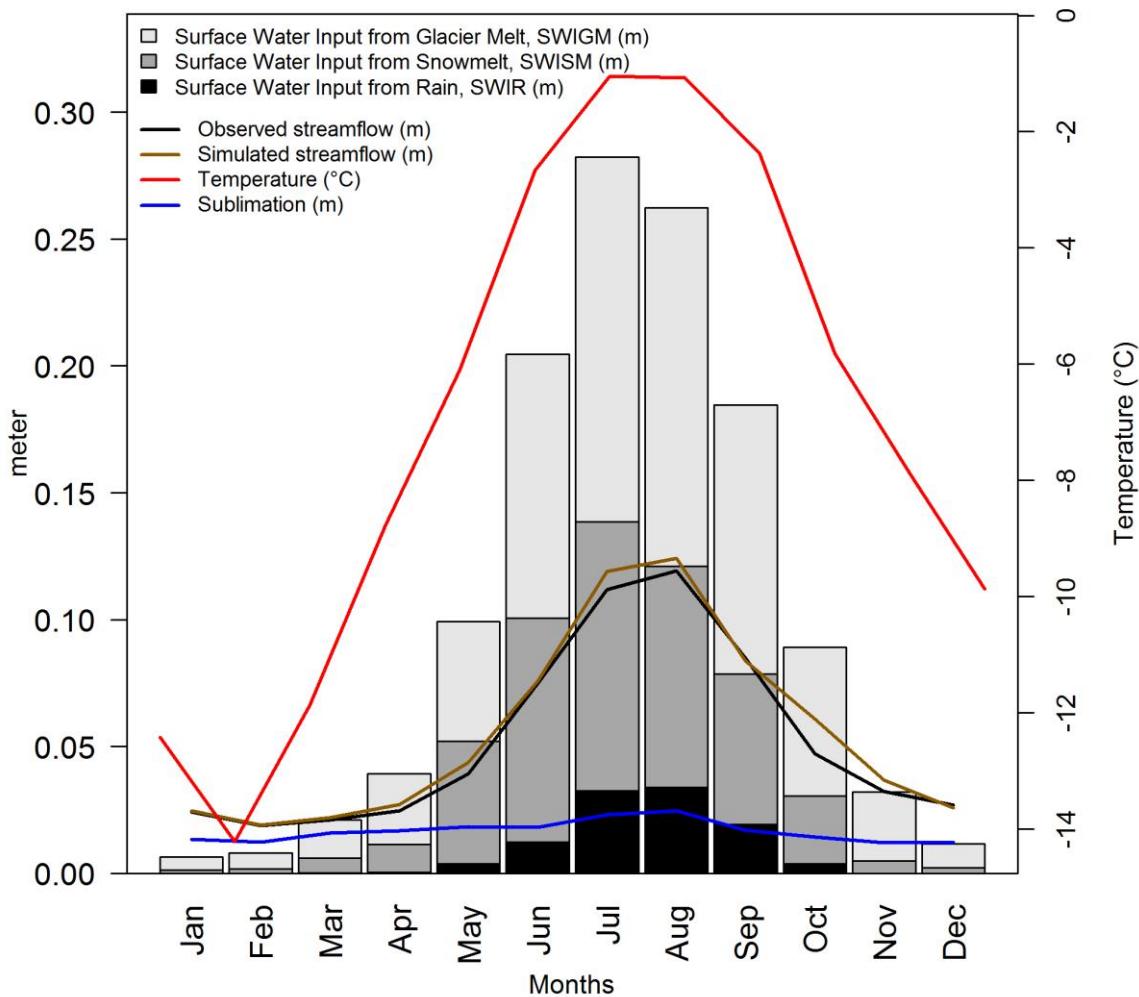


Figure 4.8. Monthly variation of SWIR (surface water input from rain, black bar), SWISM (surface water input from snow melt, dark grey bar), SWIGM (surface water input from glacier melt, light grey bar), observed streamflow (black line), temperature (red line), and sublimation (dark golden line) averaged over the 10 years 2003-2012 (streamflow only available to 2010).

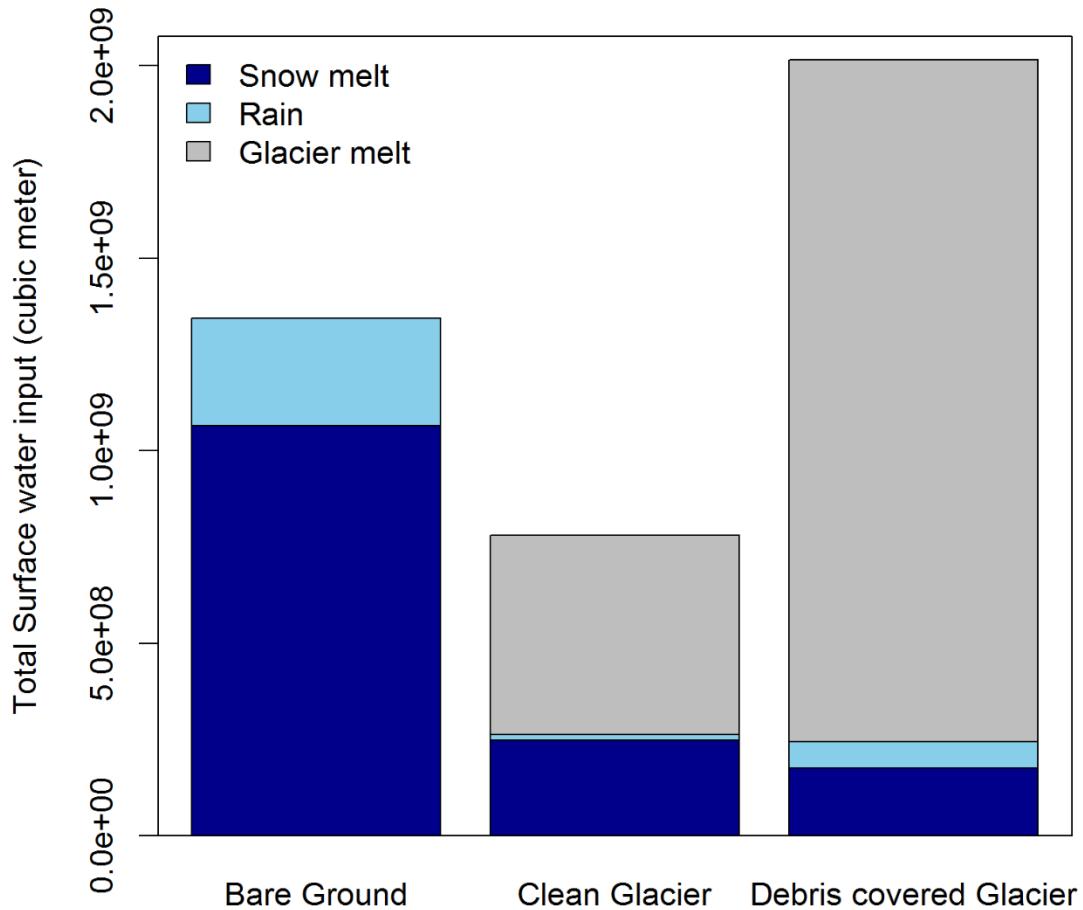


Figure 4.9. Rain, snowmelt and glacier melt from three different substrate types.

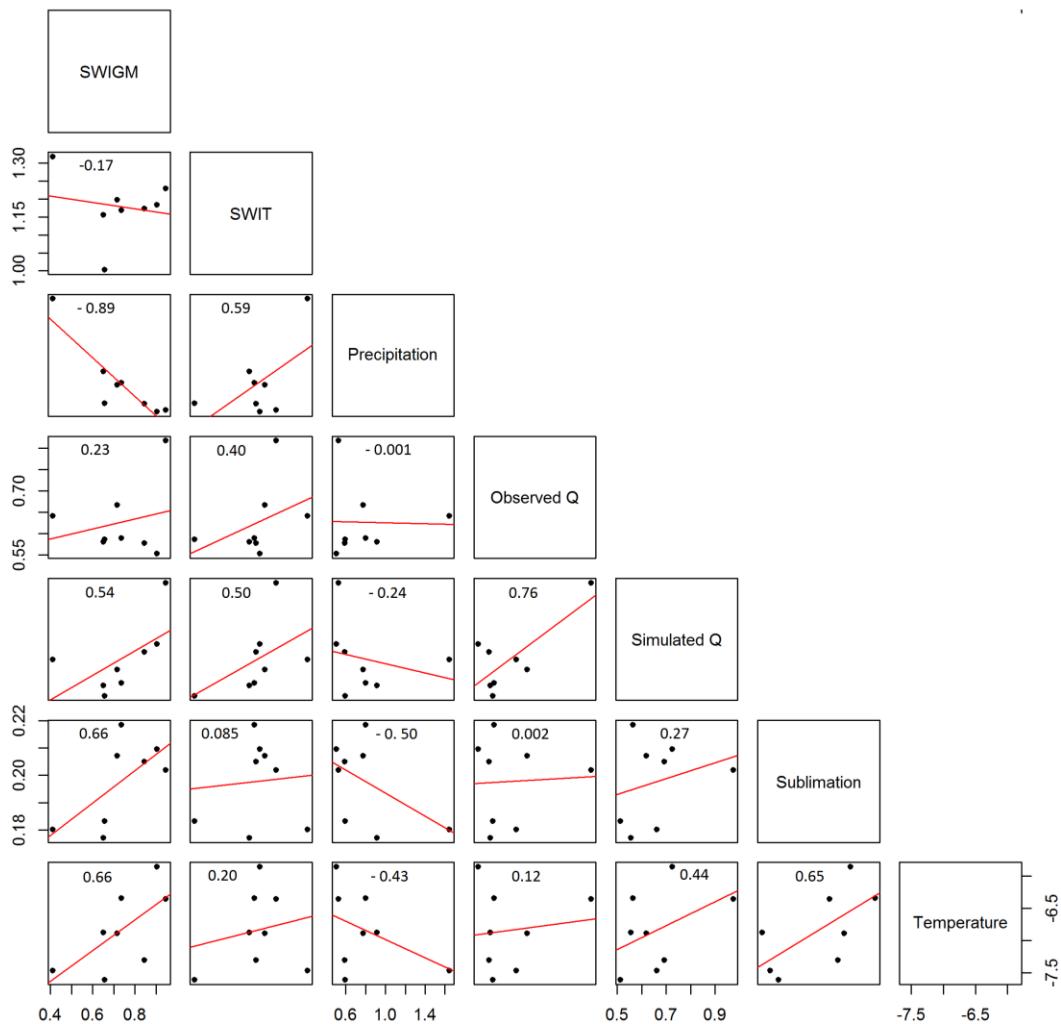


Figure 4.10. Paired relationship between annual surface water input from glacier melt (SWIGM, meter), total surface water input (SWIT, meter), precipitation (meter), observed streamflow, (Observed Q, meter), simulated streamflow, (Simulated Q, meter), sublimation (m) and temperature ( $^{\circ}\text{C}$ ). The numerical values given at the top of each plot represents the correlation coefficient between the two variables. Linear regression lines are shown in red.

## CHAPTER 5

### SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

#### **5.1. Summary and Conclusions**

This dissertation addressed the need for improving streamflow simulation due to snow and glacier melt using physically based energy balance approaches. Understanding the far-reaching societal consequences of future climate change in South Asia requires monitoring and prediction of snow and glacier ice melt runoff at watershed scale. In developing countries in this region, there is limited expertise and access to tools needed to integrate models and translate research knowledge into policy and water resources management decisions. Thus, the local institutions involved in hydrologic research and water resources management needs easy-to-use better modeling system to conduct analysis and simulation of streamflow in glacierized watersheds.

This dissertation extended the capability of the Utah Energy Balance (UEB) Snowmelt model for snow and glacier melt simulation in a watershed distributed over a grid by developing a grid-based input/output data model to support the operational application of UEB and by adding the capability to estimate the glacier melt quantity. A spatial downscaling tool was developed to produce high resolution climate data using globally available Modern Era Retrospective-Analysis for Research and Applications (MERRA) reanalysis climate data and South Asian Rain Fall Estimates (RFE2) satellite-based precipitation to drive UEB for its application in data sparse regions. Chapters 2 through 4 present the model development and scientific results of this dissertation. In this chapter I summarize the contributions in each of these chapters.

The first paper (Chapter 2) focuses on the development of a data model to structure the input and output of UEB to enable its extension from a point-based research program to a spatially distributed operational program capable of running over a watershed to produce snow and glacier ablation and feeding this as input to a rainfall-runoff model. For this, the UEB model's input and output variables were classified into four groups based on their spatial and temporal variability. Then, text files were chosen for non-spatial and netCDF files were chosen for geospatial data input and output storage. This was after investigating a series of data formats and data storage strategies that included (1) tabulated data for each grid cell within a watershed, (2) two-dimensional raster formats such as ASCII and geoTIFF, (3) sequential two-dimensional ASCII grid files. Ultimately, netCDF was chosen because of its wide application, relatively efficient access, capability of accommodating multiple variables with multiple time steps with sufficient metadata, availability of a FORTRAN netCDF library (Wessel and Smith, 1991), and the wide range of software tools for pre-and post-processing.

The UEB model was coupled with the USGS Geospatial Stream flow Model (GeoSFM) into an integrated framework to enhance streamflow prediction information in glaciated watersheds in the Hindu Kush Himalayan region and elsewhere. UEB and GeoSFM programs were each configured as freely available EPA Better Assessment Science Integrating point and Nonpoint Sources (BASINS) plug-ins to accomplish the integration. This software application provides an integrated modeling environment enhanced by the addition of UEB and GeoSFM with all of the original BASINS 4.0 analysis capability, and other preexisting plug-ins, for the coupling of models,

preparation of model input and analysis of model results. The UEB and GeoSFM plugin includes a separate Graphical User Interface containing a series of tabs to perform UEB and GeoSFM simulation tasks and visualize output. An additional window for UEB provides an easy to use environment for creating and editing UEB control files. The integrated UEB and GeoSFM models were applied to the Langtang Khola watershed for one year (October 1, 2003 to September 30, 2004) excluding the first year as model spin-up period. Snow and glacier melt contribute about 60% and 31%, respectively of total surface water input indicating the importance of these hydrological resources as streamflow contributor.

The second paper (Chapter 3) focused on how the inputs needed for UEB comprised of surface climate, radiation and precipitation data can be downscaled from MERRA climate products and Southern Asian RFE2 data. We developed spatial downscaling methods for temperature, precipitation, wind speed, relative humidity, shortwave and longwave radiation to match the scale of the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) that was chosen as the grid scale for this modeling. The temperature was adjusted for elevation differences between the effective elevation determined from the geo-potential height that MERRA used and SRTM DEM elevation using a monthly lapse rate from Liston and Elder (2006). MERRA specific humidity was used to calculate the dew point temperature, which was then adjusted for DEM elevations using a monthly vapor pressure coefficient and parameters in the saturation vapor pressure function for ice, relying on the relatively linear relationship between dew point temperature and elevation (Liston and Elder, 2006). We

then evaluated actual vapor pressure from air temperature and saturated vapor pressure from dew point temperature. Relative humidity was quantified as the ratio of these two quantities. Horizontal wind speed magnitude was obtained from eastward and northward wind components from MERRA and was interpolated bilinearly and projected to the model grid resolution. Then the effect of slope, aspect and curvature on wind speed was accounted for following Liston and Sturm (1998). MERRA reports three hourly incoming solar radiation at an elevation corresponding to the MERRA geo-potential height. A pressure based atmospheric attenuation coefficient was calculated for each time step and used to adjust MERRA incoming solar-radiation to the grid SRTM DEM elevation using a standard atmosphere pressure elevation relationship. Incoming longwave radiation was estimated based on downscaled air temperature following the methods of Liston and Elder (2006). These procedures were all coded in a computer application developed in R statistical computing language (R Development Core Team, 2009) that includes a Graphical User Interface (GUI). This application is referred to as “MERRA Spatial Downscaling for Hydrology (MSDH).”

MSDH was developed driven by the need to apply the Utah Energy Balance Snowmelt Model (UEB) to the melting of glaciers in the Himalaya region. However, there is insufficient data there to evaluate and validate the downscaling approaches so the software and methods were evaluated in the mountainous 570 km<sup>2</sup> Logan River watershed in Northern Utah at a 120-m grid resolution. The downscaled climate variables were compared with daily observations at the Natural Resources Conservation Service (NRCS) USU Doc Daniel SNOTEL station in the Logan River Watershed during October

2009 to June 2010 where USU has measurements of radiation, humidity and wind speed that are beyond the standard set of SNOTEL measurements. The daily mean, maximum and minimum temperature and monthly precipitation were also compared at a total of six SNOTEL stations in the Logan River Watershed including the USU Doc Daniel station. Then the distributed UEB model was applied to the downscaled data in the Logan River Watershed for one water year starting from October 1, 2009. Downscaled variables and simulated SWE show reasonably good agreement with the observations, indicating MSDH's capability to produce to good quality high resolution climate data using very limited observational data.

My third paper (Chapter 4) addresses the representations of surface energy balance fluxes in the UEB snowmelt model and the extension of the model to the computation of glacier melt. The model was applied in the highly glacierized Langtang Khola watershed in the Nepal Himalaya. The presence of glaciers was parameterized by adding a glacier surface ice layer beneath seasonal snowpack. When seasonal snowpack disappears, available surface energy may generate additional melt from the glacier ice substrate. The model focuses on the generation of melt at the surface using a surface ice layer of fixed thickness, avoiding the complexity of modeling the full thickness of the glacier. The model computes surface water input components from rain, snowmelt, and glacier melt, the sum of which comprise the total surface water input, which is the input for streamflow generation. Although the model runs separately for each grid cell, outputs can also be aggregated over subwatersheds defined, for example, from a DEM.

The distributed UEB snow and glacier melt model was applied in the Langtang Khola watershed for a ten year simulation period starting in water year 2003 to examine glacier and snow melt contributions in the total surface outflow and glacier mass balance. Model input preparation took advantage of the bundling of UEB as an EPA BASINS plugin that provided access to other tools in the EPA BASINS system. In particular, the GeoSFM terrain analysis tools were used to calculate slope and aspect. TauDEM tools (Tarboton and Ames, 2001) also packaged as a plug-in to EPA BASINS were used to delineate the stream network, and subwatersheds from the SRTM DEM. UEB was then run for 10 years from October 2002 to September 2012 to compute total surface water input (SWIT). The UEB-simulated total surface water input was then fed as input to GeoSFM in the place of rainfall and hydrologic losses (i.e., evaporation, change in storage) were modeled. In GeoSFM, the soil water balance and streamflow are estimated by a non-linear soil moisture accounting (Asante et al., 2008) and Muskingum-Cunge method (Ponce and Yevjevich, 1978), respectively. The model was calibrated both manually and automatically by adjusting the parameters within their plausible ranges as suggested by Asante et al. (2008) to match the observed streamflow data at Kyanjing hydrologic station provided by the Nepal Department of Hydrology and Meteorology (DHM), Nepal. The relative contribution of the three components of surface water input, namely: rain, snow, and glacier melt were estimated. Glacier mass balance and the seasonal variability of observed and simulated streamflow was also examined.

Daily mean UEB-simulated total surface water input and daily observed streamflow were found to be highly correlated and modeled total surface water input

captured the seasonal pattern of the streamflow. UEB estimates that roughly 57% and 34% of total surface water input is generated from glacier and snow. Examination of the spatial sources of surface water input shows that debris covered glaciers are the largest contributor to the surface water input. Debris covered glaciers occupying only 8% of the basin area were responsible for about half of the surface water input over the watershed. Glacier mass balance showed a negative trend, with an average snow accumulation of 2.46 m compared to 7.13 m of glacier melt during water year 2003 to 2012. To evaluate the “goodness of fit” between the observed and simulated streamflow we calculated root mean square error to the standard deviation of measured data ratio, Nash-Sutcliffe efficiency, and Percent bias. The combined (UEB and GeoSFM) model's ability to simulate daily streamflow was interpreted as “good” and the ability to simulate monthly streamflow as “very good” according to the guidance provided by Moriasi et al. (2007) for the interpretation of these fit metrics.

This work contributes to the area of physically-based glacier melt modeling and development of distributed hydrologic models. The physical basis of the energy balance approach gives it a better potential for prediction under changed conditions (e.g. climate or land cover change) where the statistical basis for empirical temperature-index models diminishes.

Another contribution of this work is tools developed to access weather data and address the disparity in scale between atmospheric and hydrologic models. NASA MERRA and NOAA RFE2 are at coarse scale and in formats not readily useable by watershed scale models. This is effectively a barrier for climate and weather information

to cross the disciplinary domain into hydrologic modeling and watershed management.

MSDH is one form of a tool that helps bridge this scale and disciplinary divide.

The development of open source, freely available and transparent modeling work (such as UEB, MSDH) is another broad contribution in that it empowers analysts and decision makers to use these tools, and the data it provides access to, in their own work. This is important in South Asian and other developing counties as it helps local planners and policy makers to blend their local hydro-climatic knowledge with the data, and scientific information provided by these tools in their water resources decision making.

The comparisons of model outputs with observations in this study provided insights on the strengths and weaknesses of this modeling. I learned in the downscaling work which variables can be downscaled relatively accurately (temperature) and which give trouble (wind and precipitation). I learned that even though direct variable comparisons may appear good (e.g. solar radiation) the sensitivity to apparent small errors may manifest significantly in model predictions. This underscores the importance of testing as many as possible model outputs against observations, highlighting the need for and importance of ground based observations. This study should thus be considered to be an initial baseline to provide hydrological insights on the hydrology and water balance of the Langtang Khola watershed and support the notion of establishing high altitude stations with weather, snow and glacier melt measuring capability for model testing and improvement.

## 5.2. Recommendations

This dissertation has successfully developed a new approach that advances UEB for representing and simulating glacier melt in addition to snowmelt capability. The model was indirectly evaluated by comparing the simulated streamflow in GeoSFM driven by UEB outputs with the measured streamflow at the watershed outlet. The results and insights from this study would be strengthened if there was an opportunity to further compare outputs to measurements such as Snow Water Equivalent (SWE) in glaciated watersheds and to compare modeled glacier melt to elevation changes in glaciers. Thus, further work is highly advocated for establishing high altitude stations with weather, snow and glacier melt measuring capability in the Himalayas for model testing and improvement.

The MSDH spatial downscaling tool needs further testing and validation. It's more detailed radiation, humidity and wind downscaling was tested and validated at only one location, with temperature UEB-simulated snow water equivalent driven by the downscaled data tested at six SNOTEL stations. Further testing of the transferability of these downscaling methods in other watersheds in different regions is recommended.

In Chapter 2, we demonstrated a successful integration of a distributed snow and glacier melt model with a semi-distributed streamflow model for the inclusion of ablation information in streamflow for glacierized watersheds using a data model based on netCDF for data distributed in space and time. However, despite the many benefits of the netCDF-based approach, there are outstanding computational input and output efficiency issues that need further study. Our initial evaluation on model run time showed that about

80% of the total run time was spent on input/output reading and writing operations. Because the computations in each grid cell proceed independently, the model is suitable for being configured to run in parallel on multiple processors by partitioning of the watershed into number of smaller spatial domains. This will require implementation of the parallel NetCDF library and parallel computation in the UEB model. Further research is also recommended to understand the effect of size and number of netCDF files on model runtime. Our experience indicates that UEB runs faster if the data is stored in a smaller number of large netCDF files rather than a large number of small netCDF files. Therefore, enabling “large netCDF” files (i.e., larger than 2 GB) may increase the program’s efficiency. Further studies can be done to explore the effect of number of variables in each netCDF on the access time.

The quasi-physically based downscaling products described in Chapter 3 are, in theory, applicable to GCM outputs and other reanalysis data products such as the European Centre for Medium-Range Weather (Trenberth and Olson, 1988), the NOAA/NCEP (Kanamitsu et al., 2002), and the Japanese 55-year Reanalysis (Ebita et al., 2011). Having multiple data sources in the system will enable us to evaluate relative strengths and weaknesses and take into account the uncertainty in the datasets. However, presently the MSDH program is capable to downscale only with MERRA and RFE2 data. Thus, extending the automated downscaling process to other reanalysis datasets is recommended.

Thin debris cover on glaciers enhances the melt due to high absorption of radiation while thick debris reduces the melting due to insulation effect, as noted in other

studies (Brock et al., 2010). In a temperature index model, Braun (1991) introduced a separate melt rate for debris covered glaciers as a calibration parameter. In an energy balance model, melt simulation in debris covered glaciers will require better understanding of melt process and information on debris thickness, which was not available for this study. We therefore recommend further research on the effect of debris on glacier melt processes for a physically based energy balance model.

The effect of glacier movement on streamflow response in Langtang Khola Watershed was not modeled, as glacier dynamics are not represented in UEB's glacier modeling module. Inclusion of glacier movement will enable the model to represent anticipated shrinkage in glacier area and volume in future climate change conditions. Development of, or coupling with, a model that represents glacier dynamics is recommended to have the capability to address these longer term questions.

In summary, we developed a data model to make UEB snow melt model interoperable with other models in BASINS and to apply over a watershed distributed on a grid. The input/output data model of UEB was redesigned and generalized to have greater flexibility in its text inputs/outputs and to exploit the capability of netCDF to hold gridded space and space-time geospatial data. Model physics was extended was to include a capability to simulate glacier melt. The UEB model and GeoSFM have been added to the BASINS toolset and coupled to estimate the contributions of glacier, snow melt and rain to streamflow in a seamless fashion using remote sensing and reanalysis weather data products. This integrated modeling system was demonstrated by successful

simulation of streamflow using glacier and snow melt contribution in the Langtang Khola watershed.

## References

- Asante, K.O., Artan, G.A., Pervez, S., Bandaragoda, C., Verdin, J.P., 2008. Technical manual for the Geospatial Stream Flow Model (GeoSFM). U. S. Geological Survey.
- Braun, L.N., 1991. Modelling of the Snow-Melt Equivalent in the Mountain Environment, In: Bergmann, H., Lang, H., Frey, W., Issler, D., Salm, B. (Eds.), Snow, Hydrology and Forests in High Alpine Areas. IAHS Publ. no. 205: Vienna, Austria, pp. 3-177.
- Brock, B.W., Mihalcea, C., Kirkbride, M.P., Diolaiuti, G., Cutler, M.E., Smiraglia, C., 2010. Meteorology and surface energy fluxes in the 2005–2007 ablation seasons at the Miage debris-covered glacier, Mont Blanc Massif, Italian Alps. *J. Geophys. Res.: Atmospheres* 115(D9), 106.
- Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K., Takahashi, K., 2011. The Japanese 55-year Reanalysis "JRA-55": an interim report. *Sola* 7, 149-152.
- Kanamitsu, M., Kumar, A., Juang, H.-M.H., Schemm, J.-K., Wang, W., Yang, F., Hong, S.-Y., Peng, P., Chen, W., Moorthi, S., 2002. NCEP dynamical seasonal forecast system 2000. *Bull. Am. Meteorol. Soc.* 83(7), 1019-1037.
- Liston, G.E., Elder, K., 2006. A meteorological distribution system for high-resolution terrestrial modeling (MicroMet). *J. Hydrometeorol.* 7(2), 217-234.
- Liston, G.E., Sturm, M., 1998. A snow-transport model for complex terrain. *J. Glaciol.* 44(148), 498-516.
- Moriasi, D., Arnold, J., Van Liew, M., Bingner, R., Harmel, R., Veith, T., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* 50(3), 885-900.
- Ponce, V.M., Yevjevich, V., 1978. Muskingum-Cunge method with variable parameters. *J. Hydraul. Div.* 104(12), 1663-1667.
- R Development Core Team, 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria, <http://www.R-project.org>.

Tarboton, D.G., Ames, D.P., 2001. Advances in the mapping of flow networks from digital elevation data, World Water and Environmental Resources Congress. ASCE: Orlando, Florida, <http://www.engineering.usu.edu/dtarb/asce2001.pdf>.

Trenberth, K.E., Olson, J.G., 1988. An evaluation and intercomparison of global analyses from the National Meteorological Center and the European Centre for Medium Range Weather Forecasts. Bull. Am. Meteorol. Soc. 69(9), 1047-1057.

Wessel, P., Smith, W.H., 1991. Free software helps map and display data. Eos, Trans. Am. Geophy. Union 72(41), 441-446.

## APPENDICES

## Appendix A

### Parameters in Utah Energy Balance Snow and Glacier Melt Model

In the Utah Energy Balance Snow and Glacier Melt Model parameters are the inputs that are spatially constant and constant in time. They include quantities intended to represent the unchanging physics of snow and glacier melt processes that are the same everywhere, as well as variables used to control the configuration of the model. They are stored in text format in the UEB model parameter file indexed by the code given in column 1.

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>	<b>Suggested value</b>
irad	Radiation control flag	This is an integer value that controls how the program calculates radiation. Values of 0, 1, 2 or 3 are valid and should be used as follows. 0: No radiation inputs are used. The model calculates radiation based on air temperature diurnal range. 1: Shortwave radiation is input. 2: Both longwave and shortwave radiation data is input. 3: Net radiation is input.		0, 1, 2 depending on data available. 3 is not recommended as it circumvents the sensitivity of outgoing radiation to modeled surface temperature (and has undergone limited testing)
ireadalb	Albedo reading control flag	This is an integer that controls whether snow surface albedo is to be input or computed internally. Values of 0 or 1 are valid and should be used as follows: 0: Model computes albedo. 1: Albedo is input.		0
tr	Rain threshold temperature	Temperature above which all precipitation occurs in form of rain	°C	3
ts	Snow threshold temperature	Temperature below which all precipitation occurs in form of snow	°C	-1

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>	<b>Suggested value</b>
ems	Emissivity of snow	Snow emissivity quantifies the emission of longwave radiation energy from the snow surface relative to black body radiation.		0.99
cg	Ground heat capacity	Ground heat capacity is the amount of heat required to change one kilogram of ground beneath the snow by 1 °C. This applies to ground in the thermally interacting layer beneath the snow.	KJ/kg/°C	2.09
z	Air measurement height	This is the height above the top of the canopy where air temperature, humidity and wind speed were measured or assumed to be effective. If no canopy, z is height above the ground or snow surface.	m	2
zo	Roughness length	Surface aerodynamic roughness length in logarithmic boundary layer wind profile	m	0.010
rho	Snow density	Density of snow is its mass per unit volume.	kg/m <sup>3</sup>	450
rhog	Soil density	Density of soil is its mass per unit volume.	kg/m <sup>3</sup>	1700
lc	Liquid holding capacity	The liquid retention capacity of the snowpack as a fraction of snow in ice (solid) phase. This quantifies the amount of liquid water that the snow can hold by capillary forces that has to be filled prior to melt outflow from the base of the snowpack.		0.05
Ks	Snow saturated hydraulic conductivity	Hydraulic conductivity parameter used in computing snow melt outflow as function of liquid relative saturation in excess of liquid holding capacity	m/hr	20

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>	<b>Suggested value</b>
de	Thermally active soil depth	The depth of substrate (ground/ice) beneath the modeled snow layer included in energy conservation calculations. This substrate depth is assumed to interact thermally and have the same average temperature as the snowpack.	m	0.1
avo	Visual new snow albedo	The fraction of the visual part of shortwave radiation (380nm-750nm) reflected by a new snow surface.		0.95
anir0	NIR new snow albedo	The fraction of the near infrared radiation (NIR) part of solar radiation (800 nm to 2500 nm) reflected by a new snow surface.		0.65
lans	Thermal conductivity of surface snow	Parameter that quantifies the rate of conduction of energy into the snow as a function of the temperature gradient.	$\text{kJ m}^{-1}\text{C}^{-1}$ $\text{hr}^{-1}$	1.0
lang	Thermal conductivity of soil	Parameter that quantifies the rate of conduction of energy into the substrate as a function of the temperature gradient.	$\text{kJ m}^{-1}\text{C}^{-1}$ $\text{hr}^{-1}$	4.0
wlf	Low frequency surface temperature parameter	Frequency of slow time scale air temperature fluctuation used in modeling surface temperature. ( $0.0654 = 2 \pi \text{ rad}/96 \text{ hour}$ for 4 day cycles). This is intended to quantify time scales longer than a day.	$\text{rad hr}^{-1}$	0.0654
rd1	Damping depth adjustment parameter	Parameter used to adjust the dampening depth for the amplitude of diurnal fluctuations in surface temperature parameterization.		1
dnews	New snow threshold depth	New snow depth (expressed as water equivalent) required for albedo to be reset to the albedo of fresh snow. For new snow depths less than this the age of the snow surface is proportionally reduced.	m	0.001

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>	<b>Suggested value</b>
emc	Canopy emissivity	Canopy emissivity quantifies the emission of longwave radiation energy from the canopy surface relative to black body radiation.		0.98
alpha	Shortwave leaf scattering coefficient	Scattering coefficient for shortwave radiation passing through the canopy.		0.5
alphal	Scattering coefficient for long wave radiation	Scattering coefficient for longwave radiation passing through the canopy.		0.0
g	Leaf orientation geometry factor	Geometry factor quantifying the fraction of leaf area that intersects a light beam penetrating the canopy. The model takes this to be constant, neglecting changes with solar incidence angle. 0.5 assumes random leaf orientation.		0.5
uc	Unloading rate coefficient	Parameter used in determining the rate of unloading of intercepted snow. Unloading rate is this coefficient time's water equivalent of intercepted snow.	hr <sup>-1</sup>	0.00463
as	Cloudy atmospheric transmissivity	Fraction of extraterrestrial radiation incident at surface on cloudy day, Shuttleworth (1993)		0.25
bs	Clear sky atmospheric transmissivity increment	Additional fraction of extraterrestrial radiation received at surface on clear day. The total radiation received at the surface on a clear day is as+bs, Shuttleworth (1993)		0.5
lambda	Clear sky direct radiation fraction	The fraction of incident radiation at the surface that is taken as direct radiation in clear sky conditions used to partition radiation into direct and diffuse fractions		0.857
rimax	Richardson number upper bound	Maximum value of Richardson number used in atmospheric stability correction		0.16

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>	<b>Suggested value</b>
wcoeff	Forest wind decay coefficient	Parameter quantifying decay of wind speed through forest canopy. Within canopy wind speed is represented as $u = u_h \exp(-n(1-z/h))$ where z is height above surface, h canopy height, $u_h$ wind speed at the top of the canopy and $n=wcoeff * LAI$ where LAI is leaf area index.		0.5
a	Transmissivity parameter	Parameter A in Bristow-Campbell formula for atmospheric transmissivity		0.8
c	Transmissivity exponent	Parameter C in Bristow-Campbell formula for atmospheric transmissivity		2.4

## Appendix B

### **Site Variables and Initial Conditions in UEB Snow and Glacier Melt Model**

In the Utah Energy Balance Snow and Glacier Melt Model site variables are the inputs that are constant in time but depend upon conditions at the location being modeled so may vary in space. They are comprised of quantities intended to represent the environment or setting being modeled such as elevation, slope, aspect and vegetation or land cover. Initial conditions are quantities that represent the initial state of the model at the beginning of a simulation. They may also vary in space, but do not vary in time in the sense that they apply to the specific time that the model is initialized. Site variables and initial conditions are stored in the same set of input files. The model allows for site variables and initial conditions to be spatially constant or spatially variable. A text file indexed by the code given in column 1 gives either the specific value of the site variable/initial condition (for spatially constant), or the name and details of a netCDF file from which spatially variable site variable/initial conditions are to be read.

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
USic	Snow energy content	Initial value of energy content state variable giving the energy content of the snow pack plus thermally active soil per unit of horizontal area defined with respect to solid (ice) phase snow at 0 °C.	kJ/m <sup>2</sup>
WSis	Snow water equivalent	Initial value of the snow water equivalent state variable giving the water equivalent of snow on the surface (ground or glacier)	m
Tic	Age of snow surface	Initial value of the dimensionless age of the snow surface state variable used in albedo calculation	

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
WCic	Canopy snow water equivalent	Initial value of intercepted snow state variable giving the water equivalent of snow held as interception in the canopy	m
df	Drift multiplier	A factor that precipitation in the form of snow is multiplied by to account for drift accumulation	
apr	Atmospheric pressure	Atmospheric pressure of a grid or a particular site. An average is sufficient as the model uses a constant value (does not accommodate weather fluctuations) in its sensible and latent heat flux calculations	Pa
Aep	Albedo extinction coefficient	Depth threshold used to interpolate albedo for shallow snow. When snow depth is shallower than apr, albedo is interpolated between snow value and substrate value. This should reflect the surface roughness or shrub height in combination with penetration depth of solar radiation into snow.	m
cc	Canopy cover fraction	The fraction of ground area covered by the vertical projection of tree crown	
hcan	Canopy height	Height of canopy	m
lai	Leaf area index	Leaf area index (LAI) is defined as one half the total leaf area per unit of horizontally projected surface area	
sbar	Interception capacity	Maximum snow load held per unit leaf area	kg/m <sup>2</sup>
ycage	Forest canopy structure flag	A parameter required for wind speed profile parameterization. Valid values are 1, 2 or 3 reflecting canopy structure 1: young coniferous 2: deciduous 3: mature coniferous (based on Paw U and Meyers, 1987)	
Slope	Slope	The slope angle measured from horizontal	degrees
Aspect	Aspect	Aspect is the direction the slope faces measured clockwise from North.	degrees
Latitude	Latitude	Geographic Latitude in decimal degrees	degrees
Longitude	Longitude	Geographic Longitude in decimal degrees (West is negative)	degrees
subalb	Substrate albedo	The fraction of shortwave radiation (fraction 0-1) reflected by the substrate beneath the snow (ground or glacier)	
subtype	Snow substrate type	Type of beneath snow substrate encoded as: 0 = Ground/Non Glacier, 1=Clean Ice/glacier, 2=	

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
		Debris covered ice/glacier, 3= Excluded area. Excluded area may include the glacier accumulation zone or any other area such as lakes where the snowmelt model is not run. No output is produced over excluded area.	
gsurf	Fraction of surface melt	The fraction of surface melt that runs directly off without infiltrating the snowpack (e.g. from a glacier)	
b01	January mean diurnal temperature range	Monthly mean of daily temperature range for January used in Bristow Campbell formulas for atmospheric transmissivity	°C
b02	February mean diurnal temperature range	Monthly mean of daily temperature range for February used in Bristow Campbell formulas for atmospheric transmissivity	°C
b03	March mean diurnal temperature range	Monthly mean of daily temperature range for march used in Bristow Campbell formulas for atmospheric transmissivity	°C
b04	April mean diurnal temperature range	Monthly mean of daily temperature range for April used in Bristow Campbell formulas for atmospheric transmissivity	°C
b05	May mean diurnal temperature range	Monthly mean of daily temperature range for May used in Bristow Campbell formulas for atmospheric transmissivity	°C
b06	June mean diurnal temperature range	Monthly mean of daily temperature range for June used in Bristow Campbell formulas for atmospheric transmissivity	°C
b07	July mean diurnal temperature range	Monthly mean of daily temperature range for July used in Bristow Campbell formulas for atmospheric transmissivity	°C
b08	August mean diurnal temperature range	Monthly mean of daily temperature range for August used in Bristow Campbell formulas for atmospheric transmissivity	°C
b09	September mean diurnal	Monthly mean of daily temperature range for September used in Bristow Campbell formulas for atmospheric transmissivity	°C

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
	temperature range		
b10	October mean diurnal temperature range	Monthly mean of daily temperature range for October used in Bristow Campbell formulas for atmospheric transmissivity	°C
b11	November mean diurnal temperature range	Monthly mean of daily temperature range for November used in Bristow Campbell formulas for atmospheric transmissivity	°C
b12	December mean diurnal temperature range	Monthly mean of daily temperature range for December used in Bristow Campbell formulas for atmospheric transmissivity	°C

## Appendix C

### **Time-varying Variables in Utah Energy Balance Snow and Glacier Melt Model**

In the Utah Energy Balance Snow and Glacier Melt Model time varying variables are the dynamics inputs that vary in time. Dynamic inputs are typically the weather input variables such as precipitation, air temperature, wind, and humidity, but they may also include quantities such as albedo. Generally the dynamic inputs are also spatially variable, although the program is configured to allow these to be either spatially variable, spatially constant or fixed. Fixed is used for quantities that may be time varying in principle but are to be held constant for a particular model run (such as ground heat flux). A text file indexed by the code given in column 1 gives either the specific value of the input (for fixed), the name of a text file holding the time series (for spatially constant), or the name of a text file listing netCDF file or files from which spatially variable dynamic inputs are to be read.

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
Ta	Air temperature	Air temperature	°C
Prec	Precipitation	Precipitation that is the sum of both rain and snowfall expressed as water equivalent	m/hr
V	Wind Speed	Wind Speed at a point z m above the snow surface or top of canopy if present	m/s
RH	Relative humidity	Relative humidity at a point z m above the snow surface or top of canopy if present	
Qsi	Shortwave radiation	Incoming shortwave radiation measured or that would be measured on a horizontal surface above the snow and canopy if present	kJ/m <sup>2</sup> /hr
Qli	Longwave radiation	Incoming longwave radiation that would be measured above the snow and canopy if present	kJ/m <sup>2</sup> /hr

Qnet	Net radiation	Net radiation that would be measured on a horizontal surface above the snow and canopy if present. This is only required if irad=3.	kJ/m <sup>2</sup> /hr
Snowalb	Snow albedo	The fraction of incident solar radiation reflected by the snow surface (in the range 0 to 1). This is only required as an input if ireadalb=1. For other values of ireadalb, the snow albedo is calculated internally based on snow surface age.	
Qg	Ground heat flux	Ground heat flux	kJ/m <sup>2</sup> /hr

## Appendix D

### **Output Variables in Utah Energy Balance (UEB) Snow and Glacier Melt Model**

UEB computes a large number of quantities at each grid cell and each time step. These may all be output, but do not have to be output, to save space and input/output time. The table below lists outputs that UEB can produce. The output control and aggregate output control files use the codes given in column 1 to designate the variables to be output, either at specific locations, for the entire grid or aggregated over subwatersheds.

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
Year	Model year	Year of beginning of time step (integer)	
Month	Model month	Month of beginning of time step (integer)	
Day	Model day	Day of beginning of time step (integer)	
Hour	Model hour	Hour of beginning of time step (may be fraction)	hr
ATF-BC	Atmospheric transmission factor	The fraction of radiation at the top of the atmosphere that reaches the top of the canopy or in its absence, the snow surface.	
HRI	Radiation index	Integration of solar radiation incident angle cosine over time step. When radiation data is not input, IRAD flag (in param.dat file) set to 0, incoming solar radiation is calculated as Tf * HRI * Solar constant.	
Eacl	Clear sky emissivity	Clear sky emissivity quantifies the emission of longwave radiation energy from a cloud free atmosphere towards the surface relative to black body radiation at the air temperature	
Ema	Atmospheric emissivity	Atmospheric emissivity quantifies the emission of longwave radiation energy from the atmosphere towards the surface relative to black body radiation at the air temperature. The emission from clouds is included	

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
Ta	Air temperature	Air temperature at a point z m above the snow surface or top of canopy if present	C
P	Precipitation	Precipitation that is the sum of both rain and snowfall expressed as water equivalent	m/hr
V	Wind speed	Wind speed at a point z m above the snow surface or top of canopy if present	m/s
RH	Relative humidity	Relative humidity at a point z m above the snow surface or top of canopy if present	RH
Qsi	Shortwave radiation	Modeled incoming shortwave radiation accounting for slope and aspect of the surface. This may be different from input Qsi for sloping surfaces	kJ/m <sup>2</sup> /hr
Qli	Longwave radiation	Modeled incoming longwave radiation	kJ/m <sup>2</sup> /hr
QnetOb	Observed net radiation	Observed net radiation that was input to the model	kJ/m <sup>2</sup> /hr
Cos	Cosine of illumination angle	Cosine of solar illumination angle (accounts for slope)	Degree
Ub	Energy content	State variable that gives the energy content of the snow pack plus thermally active soil per unit of horizontal area defined with respect to solid (ice) phase snow at 0 °C	kJ/m <sup>2</sup>
SWE	Surface snow water equivalent	State variable that gives the Snow Water Equivalent (SWE) of snow on the surface. It can be considered as the depth of water that would theoretically result if the whole snow pack instantaneously melts. This tracks snow accumulation and ablation on top of a substrate layer which may be ground or glacier. In the case that the substrate is glacier this does not track the quantity of glacier ice.	m
tausn	Dimensionless snow surface age	Dimensionless age of the snow surface state variable to account for aging of the snow surface dependent on snow surface temperature and snowfall	

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
Prain	Precipitation in the form of rain	Amount of precipitation that occurred in the form of rain at any time step	m/hr
Psnow	Precipitation in the form of snow	Amount of precipitation that occurred in the form of snow at any time step expressed as water equivalent	m/hr
Albedo	Snow surface albedo	The fraction of shortwave radiation reflected by the snow surface.	
Qh	Surface Sensible heat flux	Surface sensible heat flux is the flux of energy transferred from the snow surface to the atmosphere by air movement (wind and turbulence).	kJ/m <sup>2</sup> /hr
Qe	Surface Latent heat flux	Surface latent heat flux is the flux of energy transferred from the snow surface to the atmosphere by water vapor carried in air movement (wind and turbulence).	kJ/m <sup>2</sup> /hr
E	Surface sublimation	Amount of water removed from the snow surface by sublimation	m
SWIT	Total outflow	Total outflow from the base of the snowpack (and glacier). This includes rainfall, melt from seasonal snow and melt from glaciated surface.	m/hr
Qm	Outflow energy flux	Energy removed from the snowpack by total outflow	kJ/m <sup>2</sup> /hr
Q	Net surface energy exchange	The net sum of all surface layer (snow plus thermally interacting substrate) energy fluxes	kJ/m <sup>2</sup> /hr
dM/dt	Net surface mass exchange	The net sum of all surface layer mass fluxes	m/hr
Tave	Average snow temperature	Average temperature of the snow and thermally interacting substrate.	Degree C
Ts	Surface snow temperature	Temperature at the surface of the snow	Degree C
CumP	Cumulative precipitation	Cumulative precipitation from beginning of model run	m
CumE	Cumulative surface sublimation	Cumulative sublimation from beginning of model run	m
CumMelt	Cumulative surface melt	Cumulative melt outflow from beginning of model run	m

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
NetRads	Surface net radiation	Modeled net radiation exchange between the snow surface and atmosphere above and canopy above if present.	kJ/m <sup>2</sup> /hr
Smelt	Melt generated at surface	Amount of melt generated at the snow surface due to rain, snowmelt or glacier melt. Smelt does not include snow melt from the canopy. Smelt also does not equate to melt outflow since it infiltrates into the snow and is subject to refreezing or liquid retention depending on the thermal state of the snow.	m/hr
RefDepAct	Active refreezing front depth	The depth of a refreezing front that is active in impacting surface temperature. This quantifies the depth that refreezing has propagated into the snowpack where liquid water is present. This is reset to 0 when it exceeds the depth to which diurnal temperature fluctuations propagate and refreezing becomes inactive in snow surface temperature and energy exchange.	m
RefDep	Refreezing front depth	The depth the refreezing front has propagated into the snowpack where liquid water is present. This is physically the same as RefDepAct but is not set to 0 when it exceeds the depth to which diurnal temperature fluctuations propagate, so records refreezing depth whenever there has been refreezing and there is still liquid water present.	m
Cf	Cloudiness fraction	The fraction (between 0 and 1) of the sky occupied by clouds.	
Taufb	Direct solar radiation atmospheric transmissivity	The part of the atmospheric transmissivity that quantifies direct solar radiation, defined as the ratio of top of atmosphere radiation to direct solar radiation at the surface or top of canopy if present.	

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
Taufd	Diffuse solar radiation atmospheric transmissivity	The part of the atmospheric transmissivity that quantifies diffuse solar radiation, defined as the ratio of top of atmosphere radiation to diffuse solar radiation at the surface or top of canopy if present.	
Qsib	Direct solar radiation	The incident solar radiation received at the surface or top of canopy if present as direct solar radiation.	kJ/m <sup>2</sup> /hr
Qsid	Diffuse solar radiation	The incident solar radiation received at the surface or top of canopy if present as diffuse solar radiation.	kJ/m <sup>2</sup> /hr
Taub	Direct solar radiation canopy transmission fraction	The fraction of direct solar radiation incident at the top of the canopy that is transmitted through the canopy as direct solar radiation without being scattered or absorbed.	
Taud	Diffuse solar radiation canopy transmission fraction	The fraction of diffuse solar radiation incident at the top of the canopy that is transmitted through the canopy without being scattered or absorbed.	
Qsns	Surface shortwave absorption	Amount of solar radiation absorbed at snow surface	kJ/m <sup>2</sup> /hr
Qsnc	Canopy shortwave absorption	Amount of solar radiation absorbed in canopy	kJ/m <sup>2</sup> /hr
Qlns	Surface longwave absorption	Amount of longwave radiation absorbed at snow surface	kJ/m <sup>2</sup> /hr
Qlnc	Surface longwave absorption	Amount of longwave radiation absorbed in canopy	kJ/m <sup>2</sup> /hr
Vz	wind speed beneath canopy	Modeled wind speed beneath canopy at height z above the surface	m/s
Inmax	Interception capacity	Maximum amount of snow that a canopy can hold during a snowfall. This is a function of maximum snow load per unit leaf area, leaf area index and the density of fresh snow.	m

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
int	Interception flux	The flux if precipitation that is intercepted by the canopy. This is a function of the interception capacity and intercepted snow state variable.	m/hr
ieff	interception efficiency	Fraction of precipitation intercepted by the canopy	
Ur	Canopy unloading rate	The flux of snow unloaded from the canopy. Unloading rate is the intercepted snow state variable times the unloading rate coefficient and represents the transfer of snow from the canopy to the surface. It quantifies snow water equivalent removed from the canopy and added to the surface snow water equivalent.	m/hr
SWEc	Canopy snow water equivalent	Intercepted snow state variable giving the water equivalent of snow held as interception in the canopy.	m
Tc	Canopy temperature	Temperature of the leaves and branches within the canopy. This is used in the calculation of energy fluxes between the canopy and within canopy air.	Degree C
Tac	Air temperature within canopy	Temperature of air within the canopy. This is used in the calculation of energy fluxes between the canopy and within canopy air, and in the calculation of energy fluxes between within canopy air and the atmosphere above, and snow surface below.	Degree C
QHc	Canopy sensible heat flux	Energy flux from the air within the canopy to the canopy. This is positive towards the canopy and is calculated based on temperature gradient and bulk leaf boundary layer resistance.	kJ/m <sup>2</sup> /hr

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
Qec	Canopy latent heat flux	Latent energy flux from the air within the canopy to the canopy. This is positive towards the canopy and is calculated based on the vapor pressure gradient and bulk leaf boundary layer resistance. It represents the energy flux associated with the phase change due to sublimation (removal) or condensation/deposition (addition) of canopy intercepted snow from water vapor in the air.	kJ/m <sup>2</sup> /hr
Ec	Canopy sublimation	The flux, expressed as snow water equivalent, of removal of snow from canopy interception by sublimation. This is positive away from the canopy.	m/hr
Qpc	Precipitation energy flux to canopy	The flux of energy added to the canopy by interception. This represents the flux due to the energy difference between the phase and temperature of precipitation and the reference condition of 0 °C solid phase.	kJ/m <sup>2</sup> /hr
Qmc	Canopy melt energy	The flux of energy removed from the canopy due to melt. This represents the energy flux due to the latent heat of fusion energy difference between melt water and the reference condition of 0 °C solid phase. This is subtracted from the canopy and added to the surface snow energy content.	kJ/m <sup>2</sup> /hr
Mc	Melt from canopy	The flux, expressed as snow water equivalent, of removal of snow from canopy interception by melting. This is subtracted from the intercepted snow and added to surface snow.	m/hr
FMc	Net canopy mass exchange	The net sum of all canopy mass fluxes	m/hr

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
MassError	Mass balance closure error	A running total of the sum of all inputs to and outputs from the model. Theoretically this should be 0, but practically differs from 0 due to numerical precision and rounding errors in the computation. It is included as a check on the functioning of the model and if significantly different from 0 is indicative of a problem.	m
SWIGM	Glacier melt outflow	The part of outflow from the base of the snowpack and glacier that is generated from glacier melting. SWIGM includes melt originating from glacial ice, as well as outflow that may occur due to rain on a glacier, as any precipitation that falls on the snow or glacier surface is first added to the snow/glacier to account for its energy in the total energy content then melt outflow occurs if the energy content results in liquid water in excess of the liquid holding capacity.	m/hr
SWIR	Rainfall outflow	The part of outflow that is due to rain or snow that immediately melts. This only occurs on a non-glacier surface and when the surface snow water equivalent is 0. Precipitation that is rain, or that is snow that immediately melts due to a high temperature of the thermally active ground layer comprises this outflow.	m/hr

<b>Code</b>	<b>Name</b>	<b>Definition</b>	<b>Units</b>
SWISM	Snowmelt outflow	The part of outflow that is due to the melting of the seasonal snow pack. SWISM includes melt originating from the seasonal snow as well as outflow that may occur due to rain on a snowpack, as any precipitation that falls on the snow or glacier surface is first added to the snow/glacier to account for its energy in the total energy content then melt outflow occurs if the energy content results in liquid water in excess of the liquid holding capacity. If surface snow is present then melt outflow is generated from the surface snow. Glacier melt outflow is only generated when the surface snow water equivalent ablates to 0 and the substrate is glacier.	m/hr

**Appendix E****Coauthor approval letter**



Department of Civil and Environmental Engineering  
4110 Old Main Hill  
Logan, UT 84322-4110  
Telephone: (435) 797-2932  
Fax: (435) 797-1185

December 3, 2014

Dr. Adina Racoviteanu  
Laboratoire de Glaciologie et Géophysique de l'Environnement  
54, rue Molière, Domaine Universitaire, BP 96  
38402 Saint Martin d'Hères cedex  
France  
Email: adina.racoviteanu@lgge.obs.ujf-grenoble.fr

Dear Dr. Racoviteanu,

I am in the process of preparing my dissertation in the Civil and Environmental Engineering Department at Utah State University. I hope to complete my degree in November 2014.

I am requesting your permission to include the attached paper, of which you are coauthor, as a chapter in my dissertation. I will include acknowledgements to your contributions as indicated. Please advise me of any changes you require.

Please indicate your approval of this request by signing in the space provided, attaching any other form or instruction necessary to confirm permission. If you have any questions, please contact me.

Thank you,

Avirup Sen Gupta

I hereby give permission to Avirup Sen Gupta to use and reprint all of the material that I have contributed to Chapter 4 of this dissertation.

---

Adina Racoviteanu



## Utah State University

Department of Civil and Environmental Engineering  
4110 Old Main Hill  
Logan, UT 84322-4110  
Telephone: (435) 797-2932  
Fax: (435) 797-1185

December 3, 2014

Dr. Molly E. Brown  
Research Scientist  
NASA Goddard Space Flight Center  
Greenbelt, MD 20771  
Email: molly.e.brown@nasa.gov

Dear Dr. Brown,

I am in the process of preparing my dissertation in the Civil and Environmental Engineering Department at Utah State University. I hope to complete my degree in November 2014.

I am requesting your permission to include the attached paper, of which you are coauthor, as a chapter in my dissertation. I will include acknowledgements to your contributions as indicated. Please advise me of any changes you require.

Please indicate your approval of this request by signing in the space provided, attaching any other form or instruction necessary to confirm permission. If you have any questions, please contact me.

Thank you,

Avirup Sen Gupta

I hereby give permission to Avirup Sen Gupta to use and reprint all of the material that I have contributed to Chapter 2 and Chapter 4 of this dissertation.

---

Molly Brown



December 3, 2014

Paul Hummel  
Agricultural Systems Engineer & Office Manager  
AQUA TERRA Consultants  
150 E. Ponce de Leon Ave., Suite 355  
Decatur, GA 30030  
Email: prhummel@aquaterra.com

Dear Mr. Hummel,

I am in the process of preparing my dissertation in the Civil and Environmental Engineering Department at Utah State University. I hope to complete my degree in November 2014.

I am requesting your permission to include the attached paper, of which you are coauthor, as a chapter in my dissertation. I will include acknowledgements to your contributions as indicated. Please advise me of any changes you require.

Please indicate your approval of this request by signing in the space provided, attaching any other form or instruction necessary to confirm permission. If you have any questions, please contact me.

Thank you,

Avirup Sen Gupta

I hereby give permission to Avirup Sen Gupta to use and reprint all of the material that I have contributed to Chapter 2 of this dissertation.

---

Paul Hummel



## Utah State University

Department of Civil and Environmental Engineering  
4110 Old Main Hill  
Logan, UT 84322-4110  
Telephone: (435) 797-2932  
Fax: (435) 797-1185

December 3, 2014

Dr. Shahid Habib  
Chief, Office of Applied Sciences  
NASA Goddard Space Flight Center  
Greenbelt, MD 20771  
Email: shahid.habib-1@nasa.gov

Dear Dr. Habib,

I am in the process of preparing my dissertation in the Civil and Environmental Engineering Department at Utah State University. I hope to complete my degree in November 2014.

I am requesting your permission to include the attached paper, of which you are coauthor, as a chapter in my dissertation. I will include acknowledgements to your contributions as indicated. Please advise me of any changes you require.

Please indicate your approval of this request by signing in the space provided, attaching any other form or instruction necessary to confirm permission. If you have any questions, please contact me.

Thank you,

Avirup Sen Gupta

I hereby give permission to Avirup Sen Gupta to use and reprint all of the material that I have contributed to Chapter 2 and Chapter 4 of this dissertation.

---

Shahid Habib

## CURRICULUM VITAE

### **Avirup Sen Gupta**

30 USU Aggievillage Apt -K, Logan, Utah 84341, USA  
Office: (+1)435-797-3533, E-mail: avirup.sengupta@aggiemail.usu.edu

### **Career Overview**

My research interests lie in advancing hydrologic modeling, and information systems. In the area of hydrologic modeling, I improve existing methods for better representation of climate variability and prediction of surface water flows under climate change conditions. My research on information system makes the model easier to set up and use within the GIS software, so that the model can be placed in the hands of local scientists and decision-makers, for them to use in managing water resources.

### **Education**

**PhD Student** in Civil and Environmental Engineering, *Utah State University, USA*  
**Dissertation:** *Improving the physical processes and model integration functionality of an energy balance model for snow and glacier melt*  
**Academic advisor:** Prof. David G. Tarboton

**MS** in Civil Engineering, *University of Maine, USA*. August 2010  
**Thesis:** *Hydrologic Variability and Change in the New England Region, USA.*

**EIT/FE:** Successfully passed Fundamentals of Engineering (Civil) Exam in October 2009  
*Maine State Board of Registration for Professional Engineers (license no: 8568)*

**BS** in Civil Engineering, *Khulna University of Engineering and Technology, Bangladesh*, 2007

### **Employment**

---

#### **Aug 2010- Present Research Assistant, Utah State University**

- Enhanced the Utah Energy Balance (UEB) snowmelt model to include a representation for glacial ice and quantification water melt water from snow and glaciers.
- Developed software tools for downscaling globally available reanalysis climate data to drive hydrologic models for streamflow forecasting
- Wrote computer programs for enhancing the capability of UEB model to use gridded NASA climate model outputs and reanalysis data products provided in netCDF format.

- Developed a structured data model for UEB's input and output to facilitate its inclusion as a plugin component of the United States Environmental Protection Agency (EPA)'s BASINS modeling framework and enabling its interaction with other hydrologic models such as GeoSFM.
- Developed climate and radiation data downscaling methodologies
- Contributed in database development for NIDIS drought information using the CUAHSI Hydrologic Information System (HIS) tools.

**Dec 2013 - Mar 2014 E-Intern for the Securing Water for Food Grand Challenge Program, USAID**

I reviewed the proposals through a concept note criteria using an online platform. The proposals focused on (1) reusing wastewater, (2) capturing and storing water and (3) removing salinity of water supply in the developing countries.

**July 2011- August 2011 Visiting Fellow at Unidata, University Corporation for Atmospheric Research (UCAR)**

- Learnt Unidata software tools, such as: netCDF API, THREDDS, RAMADDA and IDV
- Examined and recommended approaches for HydroDesktop to access Unidata online data repository.
- Explored and recommended approaches to CUAHSI for using Unidata functionality to enhance the HIS cataloguing functionality that supports centralized search and data discovery.

**Sep 2008- July 2010 Graduate Research Assistant, University of Maine**

- Characterized long-term drought phenomena for the state of Maine
- Developed statistical framework/model for regional drought risk assessment.
- Performed hydrograph analysis and seasonal variability investigation related to baseflow.
- Performed uncertainty studies associated with precipitation prediction by Atmospheric General Circulation Models (AGCM).

**Computer Skills**

---

- Programming and Scripting Languages: Fortran 77/90/95, R, C#, C++, and Python
  - Database Management: MySQL and MS SQL Server.
  - Geo-spatial Data and Remote Sensing Analysis: ArcGIS and ERDAS IMAGINE
  - Hydrologic and Water Resources Simulation Models: SWAT (in R and BASINS), WEAP, WAS, HEC-HMS, HEC-RAS, UEB and GeoSFM.
  - Operating systems: Windows and Linux (Ubuntu)
  - Additional Skills: Microsoft Office Suite and Software version control tools (e.g. Mercurial)
-

## Publications

---

**Sen Gupta, A.**, S. Jain, and J.-S. Kim (2011), Past climate, future perspective: An exploratory analysis using climate proxies and drought risk assessment to inform water resources management and policy in Maine, USA. *Journal of Environmental Management*, 92, 941-947. doi:10.1016/j.jenvman.2010.10.054  
<http://www.sciencedirect.com/science/article/pii/S0301479710003865>

Brown, M. E., Racoviteanu, A. E., Tarboton, D. G., **Sen Gupta, A.**, Nigro, J., Policelli, F., Habib, S., Tokay, M., Shrestha, M., Bajracharya, S., Hummel, P., Gray, M., Duda, P., Zaitchik, B., Mahat, V., Artan, G., and Tokar, S.: An integrated modeling system for estimating glacier and snow melt driven streamflow from remote sensing and earth system data products in the Himalayas, *Journal of Hydrology*, 2014.  
<http://www.sciencedirect.com/science/article/pii/S002216941400732X>

**Sen Gupta, A.**, D. G. Tarboton, P. Hummel, M. E. Brown and S. Habib, (2014), Integration of an energy balance snowmelt model into an open source modeling framework. *Environmental Modelling & Software* (in review, submitted in July, 2014; submitted copy is available from the authors).

**Sen Gupta, A.**, D. G. Tarboton, A. Racoviteanu, M. E. Brown, S. Habib and F. Policelli, (2014), Estimating snow and glacier melt in a Himalayan watershed using an energy balance snow and glacier melt model. *Hydrological Processes* (in preparation; manuscript is available from the authors).

**Sen Gupta, A.**, D. G. Tarboton, (2014), A spatial downscaling tool for MERRA surface climate and radiation data. *Environmental Modelling & Software* (in preparation; manuscript is available from the authors).

## Software Manual

---

Nigro, J., **A. Sen Gupta**, G. A. Artan, M. Brown, D. Tarboton, S. Habib, M. Tokay, P. Hummel, M. Gray, P. Duda, A. Racoviteanu, K. Asante, J. Smith, S. Pervez, D. Entenmann, J. Verdin and J. Rowland, (2014), "HIMALA BASINS User's Manual,"  
[http://hspf.com/pub/HIMALA\\_BASINS/Documentation/HIMALA\\_BASINS\\_Users\\_Manual.pdf](http://hspf.com/pub/HIMALA_BASINS/Documentation/HIMALA_BASINS_Users_Manual.pdf)

## Conference Proceedings and Posters

---

**A. Sen Gupta**, and D. G. Tarboton, (2013), "Using the Utah Energy Balance Snow Melt Model to Quantify Snow and Glacier Melt in the Himalayan Region". *Proceedings of the*

Western Snow Conference 81st Annual Meeting. Adaptive Water Management in a Changing Climate, Jackson Hole, Wyoming, April 15-18, p. 103-114,  
<http://www.westernsnowconference.org/sites/westernsnowconference.org/PDFs/2013SenGupta.pdf>

Schreuders, K.A., D.G. Tarboton, J.S. Horsburgh, **A. Sen Gupta**, S. Reeder (2011), Enhancing Access to Drought Information Using the CUAHSI Hydrologic Information System, Abstract H43H-1325 presented at the 2011 AGU Fall Meeting, San Francisco, CA, 5-9 December. Abstract is available at  
<http://adsabs.harvard.edu/abs/2011AGUFM.H43H1325S>

Jain, S., Kim, J.-S., **A. Sen Gupta**, and Ricupero, K., (2009) Incorporating hydroclimatic variations and change into water allocation policy making and adaptation, Climate Change: Global Risks, Challenges and Decisions, Copenhagen, Denmark, 10-12 March. Abstract is available at [http://m.iopscience.iop.org/1755-1315/6/39/392023/pdf/1755-1315\\_6\\_39\\_392023.pdf](http://m.iopscience.iop.org/1755-1315/6/39/392023/pdf/1755-1315_6_39_392023.pdf)

## **Awards**

---

- *CUAHSI Pathfinder Graduate Student Fellowship* to Support Multi-site Research in Hydrology, 2010
- *Hydrologic Change Financial Award*, University of Maine, Summer 2010

## **Activities**

---

- Student member, *American Geophysical Union* (AGU)
- Manuscript reviewer, *Natural Hazards*, 2013
- Developer of the following hydro-climatic software tools
  - Glacier and grid capability for the Utah Energy Balance (UEB) Snowmelt model
  - MERRA Spatial Downscaling for Hydrology (MSDH)

## References

**Dr. David Tarboton**

Professor, Civil and Environmental Engineering Department,  
Utah State University, 4110 Old Main Hill,  
Logan, UT 84322-4110,  
Phone: (435) 797-3172  
Email: [david.tarboton@usu.edu](mailto:david.tarboton@usu.edu)

**Dr. Jeffery S. Horsburgh**

Assistant Professor, Civil and Environmental Engineering  
Utah Water Research Laboratory  
Utah State University  
8200 Old Main Hill  
Logan, UT 84322-8200  
Phone: (435) 797-2946  
Email: [jeff.horsburgh@usu.edu](mailto:jeff.horsburgh@usu.edu)

**Dr. David Rosenberg**

Associate Professor,  
Civil and Environmental Engineering Department,  
Utah State University  
Phone: (435) 797 8689  
Email: [david.rosenberg@usu.edu](mailto:david.rosenberg@usu.edu)

**Dr. Molly E. Brown**

Research Scientist  
Biospheric Sciences Laboratory, Code 618  
NASA Goddard Space Flight Center  
Greenbelt, MD 20771  
Phone: (301) 614-6616  
Email: [molly.e.brown@nasa.gov](mailto:molly.e.brown@nasa.gov)