Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature _____

Date _____

A VIRTUAL WATERSHED APPROACH FOR SNOW MODELLING

by

W. Joel Johansen

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in the Department of Geosciences

Idaho State University

Summer 2015

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of W. JOEL JOHANSEN find it satisfactory and recommend that it be accepted.

_____

Donna Delparte, PhD, Assistant Professor,
Department of Geosciences, Idaho State University
Major Advisor

_____

Dan Ames, PhD, Associate Professor
Department of Civil and Environmental Engineering, Brigham Young University
Committee Member

_____

Yolonda Youngs, PhD, Assistant Professor,
Department of History, Idaho State University
Graduate Faculty Representative

Dedication

For Ash,

You'll grow up in a world filled with virtual reality, but we'll help you know the joy of playing in the snow.

Acknowledgements

I would first like to express gratitude to my adviser, Dr. Donna Delparte, not only for her support over the past two years, but also for giving me the opportunity to receive such a unique education. That gratitude extends to the entire Geoscience Department at Idaho State University. I'm also grateful to Dr. Dan Ames and Dr. Yolonda Youngs for their willingness to be on my committee and for their valuable feedback. Finally, at a point when the appropriate words are painfully sought after, my profound love and unmeasurable gratitude are owed to my eternal companion and best friend Amanda. Thank you for dealing with the difficult task of supporting a husband in graduate school, while at the same time learning how to be a mother. Ash and I are blessed to have you.

Funding for this project was provided through Idaho State EPSCoR and the National Science Foundation under award #IIA-1329513.

Table of Contents

List of Figures

List of Tables

A VIRTUAL WATERSHED APPROACH

FOR SNOW MODELING

Thesis Abstract – Idaho State University (2015)

Understanding snowpack is a critical component to estimating fresh water supplies in the western United States. The iSNOBAL model is effective at characterizing the development and melting of a snowpack. As an initial step towards implementing the model within a virtual watershed platform, methods for simplifying and automating the creation of model inputs are described. First, parameter distribution methods were improved by using empirical Bayesian kriging (EBK), elevation gradients, and advanced simulations for several of the model input grids. Second, the process for creating the input grids was automated using the Python scripting language. Automation was also enhanced through the creation of a structured query language (SQL) database containing historical weather station data. Third, the gridding tools resulting from this work were published as web processing services. Cross-validation results show that EBK performs similarly to original detrended kriging methods, with EBK having the added benefit of being automated. Web processing services allow for the widespread use of these tools in desktop-based GIS software environments, web mapping applications, and virtual watershed platforms.

Chapter 1: Introduction

1.1 Background

Water, as a natural resource, has many significant roles throughout the world. It generates electricity, grows crops, aids in extracting other natural resources, and is a major component in several forms of recreation. In the United States in 2010, it was estimated that total fresh water extractions exceeded 300 trillion gallons per day, with 90% of that number going towards thermoelectric power, irrigation, and municipal use (Hutson *et al*., 2004).

Today, the availability—or lack thereof—of fresh water largely determines the development and expansion of an area. In the drier regions of the western United States, human settlements were established on condition of available irrigation water, as the success of agriculture indicated a sustainable future (Hansen *et al.*, 2014). The U.S. Department of Agriculture reported that in 2007, nearly 57 million acres of the country's cropland and pastureland were irrigated, with three-quarters of that land in the western states. They also reported that over four-fifths of the 91.2 million acre-feet of applied irrigation water in 2008 occurred in the same region (Schaible and Aillery, 2012).

Given this dependence on water for irrigation in the West, managing demands are difficult when considering the substantial variability in supply (Hansen *et al.*, 2014). Though Averyt *et al.* (2013) reported an ongoing stabilization in water demands, they

also noted an uncertain future as those demands will most likely evolve. Contributing to this uncertainty is the question of how a changing climate will impact fresh water supplies. Water yield in the semi-arid West is projected to decrease due to rising temperatures and declining precipitation (Brown *et al.*, 2013). This leads to concerns over several matters, including plant communities (Mathys *et al.*, 2014), wildfire frequencies (Chikamoto *et al.*, 2015), and severe droughts (Svoboda *et al.*, 2002; Schubert *et al.*, 2007).

Concerns can be mitigated, however, with a better understanding of the hydrologic cycle, especially as it applies to the montane watersheds that cover most of the western U.S. Given the fact that approximately 23% of annual freshwater demands are met by groundwater (Averyt *et al.*, 2013), and not ignoring the contributions of above-ground streamflow, one is led to examine the influence of seasonal snowpack on the overall hydrological process in a mountain watershed. Several studies have been conducted which explore the effects of a warming climate on snow and watershed dynamics, focusing on the western U.S. (Hamlet *et al.*, 2005; Mote *et al.*, 2005; Safeeq *et al.*, 2013; Godsey *et al.*, 2013; Lute *et al.*, 2015). In the most recent study listed, Lute and co-authors summarized that changes in snow metrics, including less precipitation falling as snow, earlier snowmelt, decreased late-season snow-water equivalent (SWE), and decreased late-season snow cover extent have "serious implications for water resources, agriculture, and ecosystems" (pg. 2). As the strain on freshwater supplies continues to grow, a more thorough understanding of mountain watershed resources is necessary (Tidwell *et al.*, 2014).

To assist watershed scientists in the region, three states have established the Western Consortium for Watershed Analysis, Visualization, and Exploration (WC-WAVE) ("Western Tri-State Consortium [Online]," 2015). As part of the group of researchers and students in Idaho, Nevada, and New Mexico, we are working in a highly collaborative environment to develop tools for the purpose of studying localized impacts of climate change on high-mountain watersheds. The consortium is comprised of three components: Watershed Science, Visualization and Data Cyberinfrastructure, and Workforce Development. These three groups are working together to explain the interactions between precipitation, snowpack, groundwater flow, and other properties within mountain catchments. To do this, we are building a Virtual Watershed Platform (VWP) to promote data exploration and analysis. The VWP will enable researchers to easily access and visualize several forms of data, including hydrologic model input and output. This visualization can take place in multiple settings, such as individual workstations (desktops), web-based environments, and advanced interactive 3D environments such as stereo projection, immersive CAVEs, and virtual reality.

Given the importance of snowpack, researchers in the WC-WAVE project chose to include the image snow cover energy- and mass-balance (iSNOBAL) model (Marks, Domingo, and Susong, 1999) within the VWP. The iSNOBAL model is included as part of the Image Processing Workbench (IPW) suite of software tools (Marks, Domingo, and Frew, 1999). The model is effective at predicting snowmelt and runoff and can be applied to a wide range of watershed sizes (1 to 2500 km^2). A single model run can encompass a week or an entire season. iSNOBAL has been used to study wind effects on

snow redistribution (Winstral and Marks, 2002), soil infiltration at the rain-snow transition zone (Kormos *et al.*, 2013), and soil evaporation trends with respect to precipitation (Wang *et al.*, 2013).

This thesis focuses on methods for creating input grids from weather station data for the iSNOBAL model. When run as part of the IPW, the model requires three different forms of input: an initial conditions image, precipitation images, and input forcing data images. Each of these three types of input is comprised of multiple spatial data layers. The initial conditions image consists of elevation, roughness length, total snow cover depth, average snow cover density, active snow layer temperature, average snow cover temperature, and percentage of liquid H_2O saturation of the snow cover. A single precipitation image is made up of total precipitation mass, percentage of precipitation as snow, density of the snow portion of precipitation, and average precipitation temperature (dew point temperature). Input forcing data images are composed of incoming thermal (long-wave) radiation, solar (short-wave) radiation, air temperature, vapor pressure, wind speed, and soil temperature. While the model requires a single initial conditions image, and precipitation images only for precipitation events (i.e. storms), it calls for separate input forcing data images for each individual time step. In the case of a model run over an entire year with one-hour time steps, this results in over 8500 images.

Following the generation of the above input data, iSNOBAL uses a two-layer snow model to create two output images. The first is a snow properties image, which contains layers for predicted depth of snow cover, snow density, and snow mass. The

second is an energy and mass flux image with layers for predicted evaporation, snowmelt, and runoff.

Coupling the capabilities of the iSNOBAL model with a virtual watershed framework will enhance the model's usability. To understand how this coupling will enhance model usability it is useful to better understand the VWP concept itself. The proposal for the WC-WAVE project states,

> *"[The virtual watershed] denotes a combination of data resources and computing activities and services that enable linking scientific modeling, visualization, and data management components for the purpose of enhanced analysis and exploration of real or hypothetical watersheds,"* (Goodwin *et al.*, 2013).

In a typical watershed study workflow, data is gathered from several sources, such as online repositories and personal databases. These data can consist of observational data from weather stations, digital elevation models, and GIS layers such as stream networks and watershed boundaries. Once the various forms of data are collected, preprocessing steps are performed to prepare the data for model ingestion. Preprocessing is usually required due to the fact that raw data is rarely in an appropriate format for a given model. These processes generally include some form of quality assessment and data reformatting. Following model runs, the output is visualized in one of several ways, usually depending on format. Formats can include geotiffs and Web Mapping Services which can be consumed for visualization in desktop GIS software and web applications or web maps.

Rather than requiring separate software components for each of these steps, researchers can find necessary data, prepare required model input, run multiple models, and visualize the results, all under the unifying umbrella of the VWP.

1.2 Statement of Purpose

Many hydrological modeling tools today do not take advantage of current computing and data resources (Humphrey *et al*., 2012), such as data access methods (Leonard and Duffy, 2013), data preprocessing methods (Ly *et al*., 2013), and easy-to-use web applications (Swain *et al*., 2015). The iSNOBAL model is a prime example of this problem. The process for creating input grids for the model is time-consuming. In addition, several of the methods for creating the grids are outdated, as the underlying IPW software was initially developed in 1990 (Frew, 1990) and last revised in 2002. Furthermore, its accessibility is limited since the IPW software is only compatible with Linux operating systems. By using newer computing technologies, and drawing on the ability to process and visualize results through web accessible computing tools, the iSNOBAL model can be more readily accessible by researchers and therefore more widely used in watershed studies.

This thesis describes initial steps that were taken to implement the iSNOBAL model within the VWP. Specifically, this work addresses the issues associated with the creation of input grids for the model. It is hypothesized that using the novel approach of empirical Bayesian kriging will improve input grids of air temperature, average

precipitation temperature, vapor pressure, precipitation mass, and initial snow depth. In addition, solar radiation and wind speed grids take advantage of modern software tools, and improvements are also made to soil temperature estimates.

Related to these improvements in methods are improvements in efficiency. The Marks *et al.* (1999) methods use a detrended kriging algorithm for several of the input grids, which requires a user to visually inspect and manually adjust semivariogram parameters for each time step. This time-consuming task can be automated using empirical Bayesian kriging methods. Further automation is achieved by using structured query language databases (SQL) and the Python scripting language. An SQL database allows for simple and efficient storage and retrieval of the source climate station data; and the Python scripting language automates data retrieval and the execution of gridding functions. By combining these two resources, the repetitive task of creating input grids for each time step can be made more efficient.

Finally, to address the issue of accessibility, the tools resulting from this work are made available in several formats. Using the Open Geospatial Consortium's (OGC) web processing service (WPS) standard (Michaelis and Ames, 2008), they can be published as a geoprocessing service. This enables their use in web mapping applications through RESTful uniform resource locators (URLs), and also in desktop-based GIS environments. The source code for the tools is also made available on a public repository, allowing researchers to modify and implement the tools according to their needs.

As additional steps are taken to fully implement the iSNOBAL model within the VWP, researchers will have a set of tools to assist them in the study of snowpack and

watershed dynamics. These tools can be useful in model sensitivity studies, exploring future scenarios, and a myriad of other applications.

1.3 Thesis Organization

The remainder of this thesis describes the work that was completed in two parts. Chapter 2 explains the initial development of the gridding tools using the Python scripting language and Esri's ArcServer geoprocessing services. The work reported in this chapter follows and improves upon the methods of Susong *et al.* (1999). Since the focus was primarily software development, these initial tools used pre-built data tables comprised of one hour's worth of weather station data, and included parameters such as air temperature, vapor pressure, solar radiation, precipitation depth, etc.

Chapter 3 presents the process of extending the functionality of the tools presented in Chapter 2 by using an SQL database of historical data. This chapter also presents an explanation of how the tools were published as geoprocessing services and how they are used in a web interface. In addition, Chapter 3 describes a process for validating the use of EBK methods, and the associated results of that validation process.

Chapter 4 presents the uses and advantages of the newly developed gridding tools, together with their integration into the VWP. Appendix A contains a table describing the data that was used in this work, Appendix B explains how to access the Python code through a public repository, Appendix C provides supplemental material

associated with the wind gridding tool, and Appendix D lists the steps necessary for building a custom SQL database compatible with the gridding tools.

Chapter 2: Initial development of gridding tools for creating iSNOBAL input

2.1 Introduction

Water resources were just as critical for civilizations historically—as evidenced by the discovery of early irrigation practices—as they are today. Water's significance can especially be seen in regions where it is more limited, such as the western United States. Here, the availability of irrigation water determined the success or failure of early settlements (Hansen *et al.*, 2014). Today, the sustainability and expansion of populations in the West rely heavily on successfully managing water as a resource (Pagano *et al.*, 2004). Successful management, however, first requires an understanding of the hydrological processes that govern fresh water supplies.

Mountain watersheds cover most of the western U.S. Within these high elevation catchments, snowpack plays a pivotal role in the year-round fresh water supply (Wood and Lettenmaier, 2008). It does this in two ways: first, through infiltration and groundwater recharge (Winograd *et al.*, 1998; Drexler *et al.*, 2013; Godsey *et al.*, 2013), and second, through above-ground surface flow (Mahanama *et al.*, 2012), both of which are related to the timing and magnitude of snowmelt. These three dynamics—groundwater recharge, surface flow, and snowmelt—are governed by several factors, including terrain, air temperature, soil moisture content and temperature, solar and thermal radiation, vegetation, and wind (see Schelker *et al.*, 2013; Kumar *et al.*, 2013;

Harpold *et al.*, 2014; Hinckley *et al.*, 2014). Recognizing that the above list is non-exhaustive, one gains an appreciation for the difficulty of predicting, and subsequently managing, water supplies in mountainous watersheds.

Several models have been developed for the purpose of predicting snowpack properties over the course of a season. SNOW17 (Anderson, 1973) was designed as part of the National Weather Service's River Forecasting System, and is still in use today (Franz *et al.*, 2008). The SNTHERM (Jordan, 1991) and SHAW (Flerchinger and Saxton, 1989) models accurately simulate snow properties at a point, while the Utah Energy Balance (UEB) model (Tarboton *et al.*, 1995) and the United States Geological Survey's Precipitation-Runoff Modeling System (PRMS; Leavesley *et al.*, 1983) can be applied as distributed simulations over small areas.

This chapter focuses on the image snow cover mass- and energy-balance model (iSNOBAL). iSNOBAL was developed by Marks *et al.* (1999) to address the limitations of existing models, which included complex, and in some cases, computationally expensive calculations, inabilities to be explicitly distributed over grids, and difficulties in obtaining adequate and necessary input data. iSNOBAL can accurately characterize the development and melting of a snowpack in various settings by representing the snow cover as a two-layer system. It can be spatially distributed over digital elevation model (DEM) grids ranging from 1 to 10,000 km^2 and can also be applied to various temporal ranges, such as a single week, or an entire season/year.

One of the advantages of iSNOBAL is its relatively simple list of input data. The model is initiated by an initial conditions image consisting of seven spatial data layers:

elevation, roughness length, total snow cover depth, average snow cover density, active snow layer temperature, average snow cover temperature, and percentage of liquid H_2O saturation of the snow cover. It is then driven by an input forcing data image for each time step comprised of six spatial data layers: incoming thermal (long-wave) radiation, air temperature, vapor pressure, wind speed, soil temperature, and net solar (short-wave) radiation. The model is updated during the run for each precipitation event (i.e. storm) with images that consist of four spatial data layers: total precipitation mass, percentage of precipitation mass that was snow, density of snow portion of the precipitation, and average precipitation temperature (dew point temperature). Most of these parameters can be acquired from weather stations found throughout a given basin.

Marks *et al.* (1999) introduced the iSNOBAL model, and followed the methods detailed by Susong *et al.* (1999) in creating the necessary input grids. Air temperature, precipitation mass, and initial snow mass grids were created using a detrended kriging algorithm which consisted of three general steps: first, subtracting linear parameter-elevation trends from observed values; second, performing ordinary kriging on the resulting residuals; and third, adding back to the residual grid the linear trend calculated in the first step. Since the input spatial data grids generally exhibit a changing mean from elevation effects—for example, higher temperatures at lower elevations and lower temperatures at higher elevations—and kriging assumes that the data is stationary, this process of detrending was necessary.

Since the introduction of the iSNOBAL model, interpolation methods have advanced and a more robust approach is now available, namely, empirical Bayesian kriging (EBK; Krivoruchko, 2012). While ordinary kriging methods require a user to visually inspect and adjust the estimated semivariogram parameters, EBK automates this procedure through Monte-Carlo-Markov-Chain techniques (Pilz and Spöck, 2008). When compared to other kriging methods, EBK is able to interpolate weakly nonstationary data over larger areas, at the same time requiring fewer data points (Krivoruchko, 2012; Deng *et al.*, 2013).

In a comparison of several interpolation methods, Cooper *et al.* (2015) found that EBK ranked among the highest in generating accurate water table elevation models. Mulcan *et al.* (2015) used EBK to identify acceptable locations for anchoring ocean current energy turbines by interpolating seafloor core samples. Other examples of its use include interpolating air pollution sources (Laurent *et al.*, 2014), soil organic content for validating spectroscopy studies (Deng *et al.*, 2013), and climate grids for mapping suitable snail habitats (Pedersen *et al.*, 2014).

iSNOBAL has limitations. As part of the IPW suite of software tools (Marks, Domingo, and Frew, 1999), the model only works on Linux operating systems, thus restricting its overall accessibility. In light of modern capabilities, researchers could profit from a simplified approach to the model by using the Open Geospatial Consortium's (OGC) web processing services (WPS) (Michaelis and Ames, 2008), such as those provided through Esri's geoprocessing services. Both services can be implemented in a desktop-based geographic information system (GIS) environment, or accessed

through simple web interfaces and RESTful uniform resource locators (URLs), allowing for more widespread accessibility and ease of use on current computing platforms.

Another limitation of iSNOBAL is the model's lack of efficiency regarding the creation of the required input grids. This is a manual, time-consuming process. For a model run over an entire year at one-hour time steps, a user would be required to visually inspect and adjust semivariograms for nearly 9,000 input images. Considering the advantages of EBK mentioned above (automation of estimating semivariogram parameters), the process of grid creation can potentially become more systematic. Furthermore, the repetitive nature of grid creation in general can be addressed by using the Python scripting language.

This research addresses the issues noted above through the creation of a set of gridding tools and geo- and web-processing services using the Python scripting language and Esri's ArcPy library. This allows for the implementation of more advanced methods, including EBK, to improve upon the Marks *et al.* (1999) and Susong *et al.* (1999) techniques. The developed tools can be implemented within an automated process to make the creation of input grids much less time consuming.

The remainder of the chapter is organized as follows: the methods section includes descriptions for the study area and the data tables used in testing and developing the gridding tools, as well as detailed explanations of the methods employed by each gridding tool; the results section describes the output grids created from running the tools on historical data from the Reynolds Creek Experimental Watershed. The section also outlines the advantages of employing the gridding tools. Finally, the

discussion section summarizes the findings of this paper and addresses specific needs that arose during the course of the study.

2.2 Methods

2.2.1 Study Area

The data used in this study came from the numerous weather stations within the Reynolds Creek Experimental Watershed (RCEW) in southwest Idaho. The RCEW is a mountainous rangeland catchment approximately 50 miles (80 km) southwest of Boise, Idaho. It covers nearly 239 km^2, with elevations ranging from 1100 m mean sea level (msl) to 2240 m msl. Dominant vegetation types include sagebrush and grasses at lower elevations, and aspen and fir in the higher elevations.

The RCEW was originally established back in the early 1960's, when authorities recognized a growing need for long term data sets related to fresh water supplies (National Research Council, 1999). Since that time, several weather and hydrologic recording stations have been collecting data on a continuous basis. **Error! Reference ource not found.** shows the locations of the various recording stations throughout the watershed, along with the watershed boundary, and a shaded relief map.

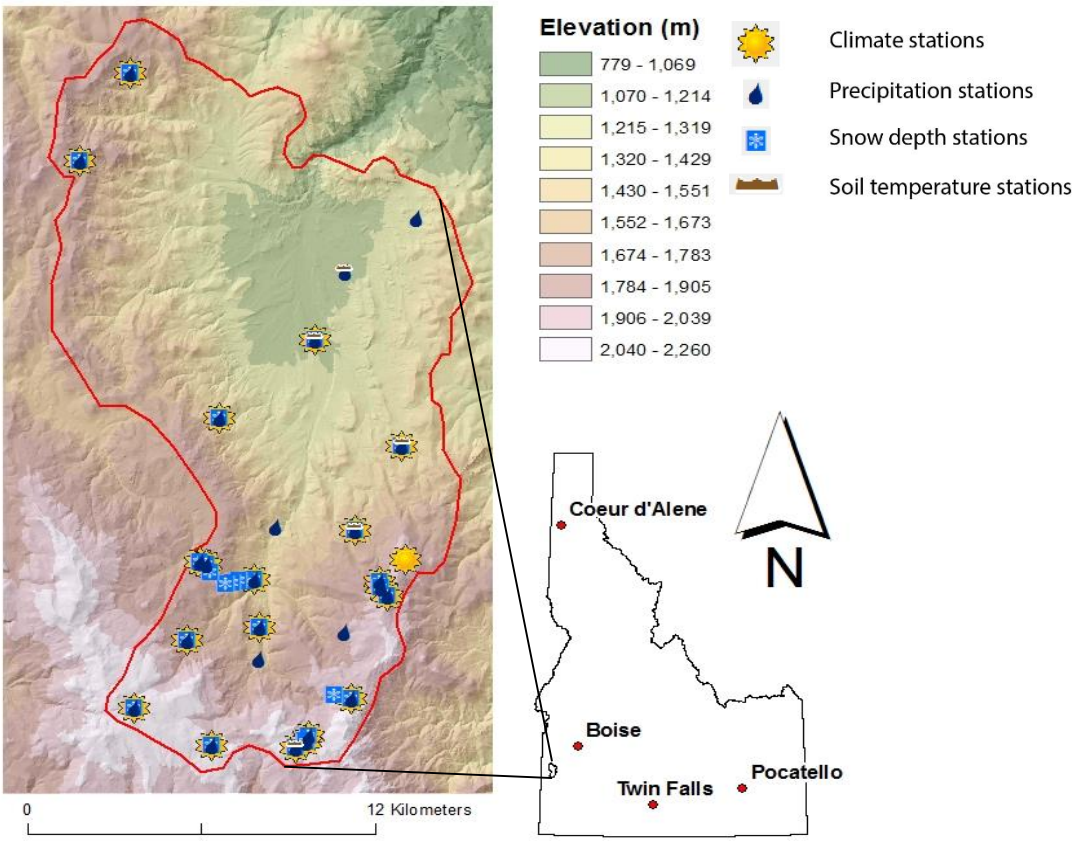


Figure 1. Diagram depicting the Reynolds Creek Experimental Watershed, including a shaded relief map, the watershed boundary, and locations of recording stations throughout the watershed.

2.2.2 Gridding Scripts

Table 1 lists 16 grids that are required for an iSNOBAL model run, subdivided into three different input types: initial conditions, precipitation, and input forcing data. The table also shows the Marks *et al.* 1999 methods, and the modified methods employed by the Python gridding scripts described in this chapter, with improved methodologies highlighted. The script names correspond to the different Python files that were written (see Appendix B).

The model requires 17 grids in all, as an elevation model is counted as part of the initial conditions image. A 10 m resolution DEM was created by blending a LiDAR derived model for the watershed with a National Elevation Dataset (NED) for the surrounding areas. This resulted in an elevation grid that was 190 km^2 in size and consisted of 1,119 rows and 1,688 columns, which was used as an input parameter for each of the gridding scripts.

Table 1. This table lists the 16 input grids for the iSNOBAL model, grouped by input type. The table also lists the Marks *et al.* (1999) methods and the modified methods for creating the grids, along with the associated gridding tool script name (note: Improved methedologies have been bolded).

Parameter	Marks *et al.*, 1999 Methods	Modified Methods	Script Name
Initial Conditions			
roughness length	constant	constant	Create Constant Grids
total snow cover depth	detrended kriging	**empirical Bayesian kriging**	Create Initial Snow Depth Grid
average snow cover density	elevation gradient	elevation gradient	Create Snow Properties Grids
active snow layer temperature	elevation gradient	elevation gradient	Create Snow Properties Grids
average snow cover temperature	elevation gradient	elevation gradient	Create Snow Properties Grids
% of liquid H_2O saturation	constant	constant	Create Constant Grids
Precipitation			
total precipitation mass	detrended kriging	**empirical Bayesian kriging**	Create Precipitation Mass Grid
percentage of precipitation mass that was snow	lookup table	lookup table	Create Dew Point Temperature Grid
density of snow portion of the precipitation	lookup table	lookup table	Create Dew Point Temperature Grid
average precipitation temperature	lookup table	**empirical Bayesian kriging**	Create Dew Point Temperature Grid
Input Forcing Data			
incoming thermal (long-wave) radiation	simulated	simulated	Create Thermal Radiation Grid
air temperature	detrended kriging	**empirical Bayesian kriging**	Create Air Temperature Grid
vapor pressure	elevation gradient	**empirical Bayesian kriging**	Create Vapor Pressure Grid
wind speed	elevation gradient	**simulated**	Create Wind Speed Grid
soil temperature	constant	**elevation gradient**	Create Soil Temperature Grid
net solar radiation	simulated	simulated	Create Solar Radiation Grid

In addition to the DEM, most of the tools also required one of four types of data tables: climate, precipitation, snow depth, and soil temperature. These four tables were built using historical data (see Appendix A) from March 1st, 2008, as shown in Tables 2 – 5. The process to build them consisted of downloading weather station data files from the FTP site and compiling all of the observed data for the given time step into a single table.

Table 2. Climate data table from 20 weather stations (designated by "Site Key") in the RCEW for the time step of March 1st, 2008 at 12:00 PM. Abbreviations denote: T_a = air temperature; H = humidity; ρ_v = vapor pressure; T_{dp} = dew point temperature; R_s = solar radiation; u = wind speed; u_{dir} = wind direction. "No-data" values are represented by "-999".

Site Key	T_a (°C)	H (%)	ρ_v (Pa)	T_{dp} (°C)	R_s (W/m^2)	u (mph)	u_{dir} (from north)
012	-0.5	0.88	514	-2.1	624	4.6	302
031	-2.6	0.81	400	-5	563	8.8	304
076	4.6	0.55	468	-3.2	600	7.2	282
095b	0.3	0.64	402	-5	705	7	276
124	-2.5	0.75	375	-5.8	704	12	288
124b	-0.8	0.83	475	-3	736	3	297
125	1.6	0.65	448	-3.7	253	3.6	266
127	-0.5	0.74	437	-4	605	9.7	282
128	-5.2	0.95	376	-5.8	-999	20.9	284
138d03	-3.4	0.81	373	-5.9	666	12.1	-999
138j10	-3.9	0.89	392	-5.3	679	8.8	-999
138L21	-4.5	0.82	344	-6.8	779	19	281
144	-1.3	0.72	397	-5.2	640	7.5	291
145	0.2	0.63	391	-5.3	628	4.2	261
163	-5.9	0.97	359	-6.3	471	8.6	302
166b	-5.2	0.84	331	-7.2	669	10	292
167	-3.6	0.93	418	-4.5	605	5.2	286
174	-3.8	0.95	424	-4.3	529	3.2	224
176	-4.6	0.86	359	-6.3	634	8.5	270
rmsp3	-4.8	0.99	404	-4.9	604	3.1	211

Table 3. Soil temperature data table from 5 stations (designated by "Site Key") in the RCEW for the time step of March 1st, 2008 at 12:00 PM. The parameter "st005" represents soil temperature at 5 cm below the ground surface. Elevations at each station were included in the table for linear regression purposes.

Site Key	st005	Elevation (m)
57	3.71	1186
76	2.9	1202

98	0.58	1410
127	0.11	1652
176	0.12	2095

Table 4. Snow depth data table from 28 weather stations (designated by "Site Key") in the RCEW for the time step of March 1st, 2008 at 12:00 PM.

Site Key	Snow Depth (cm)	Site Key	Snow Depth (cm)
012	44.1	166b	1.1
076	0.5	167	129.3
095b	0.7	167b	128.7
098c	0.2	174	98.9
124	1.5	176	70.2
124b	47.6	176b	146.8
125	18.3	bst	131.0
127	0.1	jdt1	58.9
138d03	1.4	jdt2	15.7
138j10	26.9	jdt3	67.7
138L21	0.0	jdt4	73.2
144	51.4	jdt5	23.7
145	40.3	rmsp3	151.7
163	116.2	spt	146.8

Table 5. Precipitation data table from 25 weather stations (designated by "Site Key") in the RCEW for the time step of March 1st, 2008 at 8:00 AM. Abbreviations denote: ppts = shielded; pptu = unshielded; ppta = dual gage wind corrected. All measurements are in mm and “no-data” values are represented by “-999”.

Site Key	ppts	pptu	ppta	Site Key	ppts	pptu	ppta
012	4.1	2.7	4.7	138L21	0	0	0
031	2.6	1.4	4.5	144	1.5	1.1	2.3
049	0.4	0.5	0.7	145	0.9	0.7	1.4
057	0.4	0.4	0.6	147	0.4	0.2	1.1
076	0.3	0.2	0.4	155	1.6	1.4	1.7
095b	1	0.6	1.6	163	1.4	0.4	2.2
098c	0.2	0.2	0.9	166b	0.5	0.2	0.7
116c	0.9	0.5	0.7	167	0.8	0.6	1.4
124	0.6	0.5	1.6	174	1.7	1.2	2.8
124b	2.6	-999	-999	176	0.8	0.6	2.3
125	1.5	1.5	1.9	138d03	0.4	0.1	0.3
127	0.6	0.3	0.8	rmsp3	1.3	0.7	2
138j10	0.1	0.1	0.2				

The sections below group the gridding tools into five categories based on their modified methods of creation. Total snow cover depth, total precipitation mass, average precipitation temperature (dew point temperature), air temperature, and vapor pressure all used empirical Bayesian kriging (EBK) methods to interpolate weather station data into the necessary grids. Average snow cover density, active snow layer temperature, average snow cover temperature, and soil temperature grids were created using an elevation gradient. Grids for thermal radiation, solar radiation, and wind speed were simulated over a DEM. Lookup tables were used to determine the percentage of precipitation mass that was snow, and the density of the snow portion of the precipitation based on dew point temperatures. And roughness length and percentage of liquid H_2O saturation grids were specified as constants.

2.2.2.1 Empirical Bayesian Kriging

For the grids that employ EBK, the first step is to extract elevations from the DEM for each of the weather stations. Next, the script is designed to average parameter values if the data tables contain multiple rows for each weather station (i.e. if a model is set to run on three-hour time steps). If any "no-data" values exist in the data table (see Table 5, for example), they are removed prior to averaging. Next, any row in the data table associated with a weather station that exists outside the domain of the DEM is removed. This makes it possible to perform ordinary least squares (OLS) regression using the parameter as the dependent variable, and elevation as the explanatory variable. The slope and intercept from the OLS regression are then used to predict

parameter values at weather stations that do not have observed values. Finally, EBK is performed on the combined predicted and observed parameter values to create distributed parameter grids with the same domain and spatial resolution as the DEM. The EBK function uses 100 simulated semivariograms, a smooth-circular search neighborhood with a 10,000 m radius, and a detrended-Whittle semivariogram model type.

Modifications were made to the *Create Initial Snow Depth* and *Create Precipitation Mass* gridding scripts, as it is possible for the interpolation to predict negative values. A conditional statement is used, and any predicted grid-cell values less than zero are set equal to zero.

2.2.2.2 Elevation Gradients

The *Create Snow Properties* script uses linear regression to predict values based on elevation. Linear regression equations are obtained from one of two sources- users can either specify their own interpolation points, or, default values will be used, as specified by Susong *et al.* (1999). Figure 2 shows an example dialog box where a user can input specific interpolation points as parameter-elevation pairs. Using the values for *Lower Layer Snow Temperature* from Figure 2 as an illustration, the linear regression equation takes the form of

$$\hat{y} = -0.002(h) + 4.0 \qquad [1]$$

where $\hat{y}$ is the predicted lower layer snow temperature value, and h is elevation. This equation is then applied to the DEM to produce the lower layer snow temperature grid.

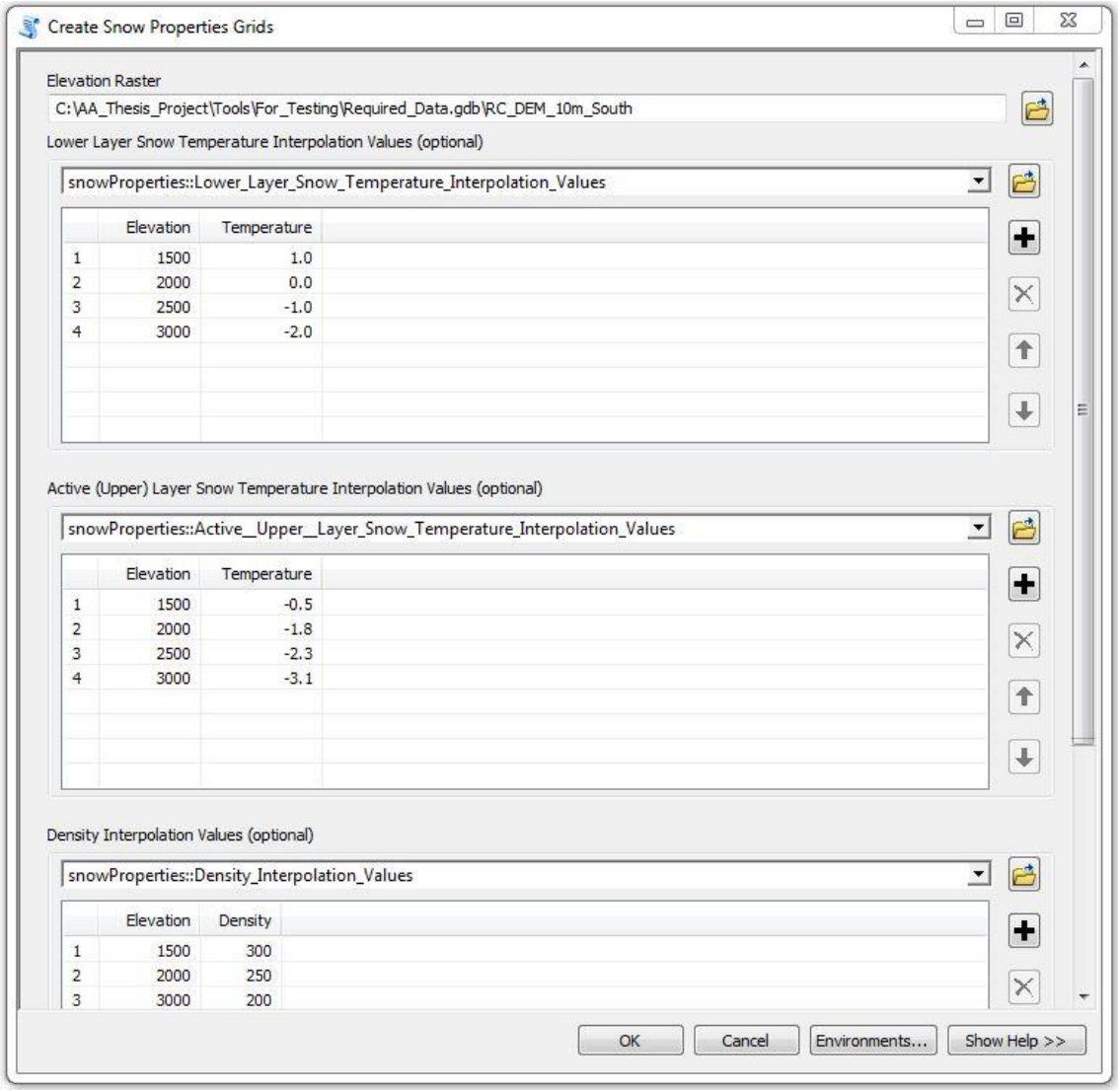


Figure 2. Dialog box showing user specified interpolation values for use in creating snow properties grids.

If the user chooses not to provide interpolation points, the script uses default values as follows: for snow density- 350 kg/m^3 at 1500 m, 300 kg/m^3 at 2500 m, and 250 kg/m^3 at 4000 m; for active snow layer temperature- −1.0 °C at 1500 m, −2.0 °C at 2500m, and −3.0 °C at 4000 m; for lower layer snow temperature- 0.0 °C at 1700 m, −1.0 °C at 2500 m, and −2.0 °C at 4000 m.

Once the active (upper) and lower layer snow temperature grids are created, they are averaged together to generate the average snow cover temperature grid.

The same general idea described above is used in the *Create Soil Temperature* script. Rather than creating a constant grid of 0˚ C, as was the convention, and since soil temperature data is available (see Table 3), the script uses ordinary least squares regression to calculate an equation for the best fit line, and uses the slope and y-intercept of that line to predict soil temperature values based on elevation.

2.2.2.3 Simulated Parameters

2.2.2.3.1 Solar Radiation

The ArcPy library contains functionality for simulating solar radiation values over a DEM. These services are used in the *Create Solar Radiation* gridding tool. Values are estimated using a 200 cell "sky size" (resolution for the viewshed, sky map, and sun map grids) for a single-hour time step. Simulated values are then corrected for cloud conditions in the following manner: first, simulated values are extracted to a table that contains observed values from weather stations (see R_s in Table 2), then stations containing "no-data" for observed values are removed from the table and, finally, ratios are calculated at each weather station by dividing simulated values by observed values. These ratios are then averaged across all weather stations and the entire simulated grid is multiplied by this averaged ratio.

2.2.2.3.2 Thermal Radiation

The approach used to create thermal radiation grids follows the methods of Marks and Dozier (1979). This requires the *Create Thermal Radiation* script to collect user input reference values for air pressure, air temperature, elevation, and surface air temperature. It also takes inputs of air temperature and vapor pressure grids that are created using the EBK methods detailed above. In addition, the script requires a view factor grid that is created following the instructions of Dozier and Outcalt (1979). The initial step in this process uses the open source Whitebox software (Lindsay, 2014) to calculate horizon angle grids at 10° increments. These grids are then divided by 36 to obtain an average horizon angle grid (H), which grid is used in equation 2 to compute a thermal view factor grid (Vf).

$$Vf = cos^2(H) \quad [2]$$

After all of the necessary input is created or specified, the script begins by converting any temperature parameters from Celsius to Kelvin. It then corrects the air temperature and vapor pressure grids from near surface to sea level equivalents, which corrected grids are used, along with the air pressure at a given elevation, to calculate effective atmospheric emissivity. Finally, equation 3 is used to calculate incoming long-wave radiation (I_i) for each grid cell of the DEM

$$I_i = \left(\epsilon_a \sigma {T_a}^4\right) Vf + \left(\epsilon_s \sigma {T_s}^4\right)(1 - Vf) \quad [3]$$

where ϵ_a is effective atmospheric emissivity, σ is the Stefan-Boltzman constant, T_a is near-surface air temperature, ϵ_s is surface emissivity, and T_s is surface temperature.

The methods above simulate thermal radiation for clear-sky conditions. The same approach used to correct solar radiation simulations for cloud conditions can be used in this instance. However, thermal radiation was not part of the observed values from weather station data.

2.2.2.3.3 Wind Speed

The *Create Wind Speed* script uses the open source WindNinja software to simulate measured wind speeds over the DEM (Forthofer *et al*., 2009). WindNinja was developed to assist fire managers in predicting spatially varying wind fields in complex terrains. It takes as input a DEM, a date and time, the dominant vegetation type of the area, a comma separated values file of measured wind speeds, and a configuration file. Examples of the latter two files that were used in developing the script can be seen in Appendix C.

After specifying the above input, the Python script uses the WindNinja command line interface (CLI) to run the simulation. The final output is an ASCII type grid of simulated wind speeds.

2.2.2.4 Lookup Tables

The same script that is used to create the dew point temperature grid is used to create the precipitation properties grids. This is due to the fact that precipitation properties (percentage of precipitation as snow, density of snow portion of

precipitation) are dependent on dew point temperature. This relationship is shown in Table 6.

Table 6. Lookup table used in the creation of snow properties grids. Precipitation temperature can also be referred to as dew point temperature. Table from Marks *et al.*, 1999.

Precipitation Temperature (°C)	Percent Snow	Snow Density (kg/m^3)
$T < -5$	100	75
$-5 \leq T < -3$	100	100
$-3 \leq T < -1.5$	100	150
$-1.5 \leq T < -0.5$	100	175
$-0.5 \leq T < 0$	75	200
$0 \leq T < 0.5$	25	250
$0.5 \leq T$	0	0

Following the creation and output of the dew point temperature grid, the *Create Dew Point Temperature* script uses two conditional statements, based on the values in Table 6, to create the grids for precipitation properties.

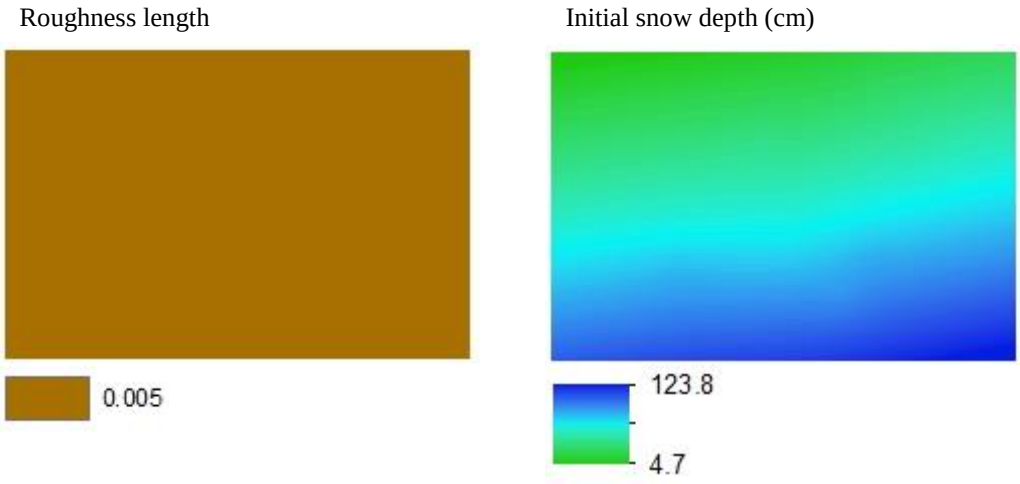
2.2.2.5 Constants

The *Create Constant Grids* script was designed with the same functionality as the *Create Snow Properties* script. That is, users can specify their own values for the roughness length and liquid H_2O saturation constants, or, if no values are specified, default values are used as recommended by Marks *et al.* (1999): 0.005 and 0.2, respectively.

2.3 Results

The Python scripts described above were tested using historical data from the RCEW, the source of which can be found in Appendix A. The data consisted of a single one-hour time step from March 1st, 2008 (see Tables 2 – 5).

Figure 3 shows the results of running the Python gridding scripts to create grids for the initial conditions image. The non-constant grids displayed variability that was anticipated. Snow temperatures were lower in the higher elevations than in the lower elevations; snow densities were also lower where the temperatures were cooler; and snow depths were greater in the higher elevations.



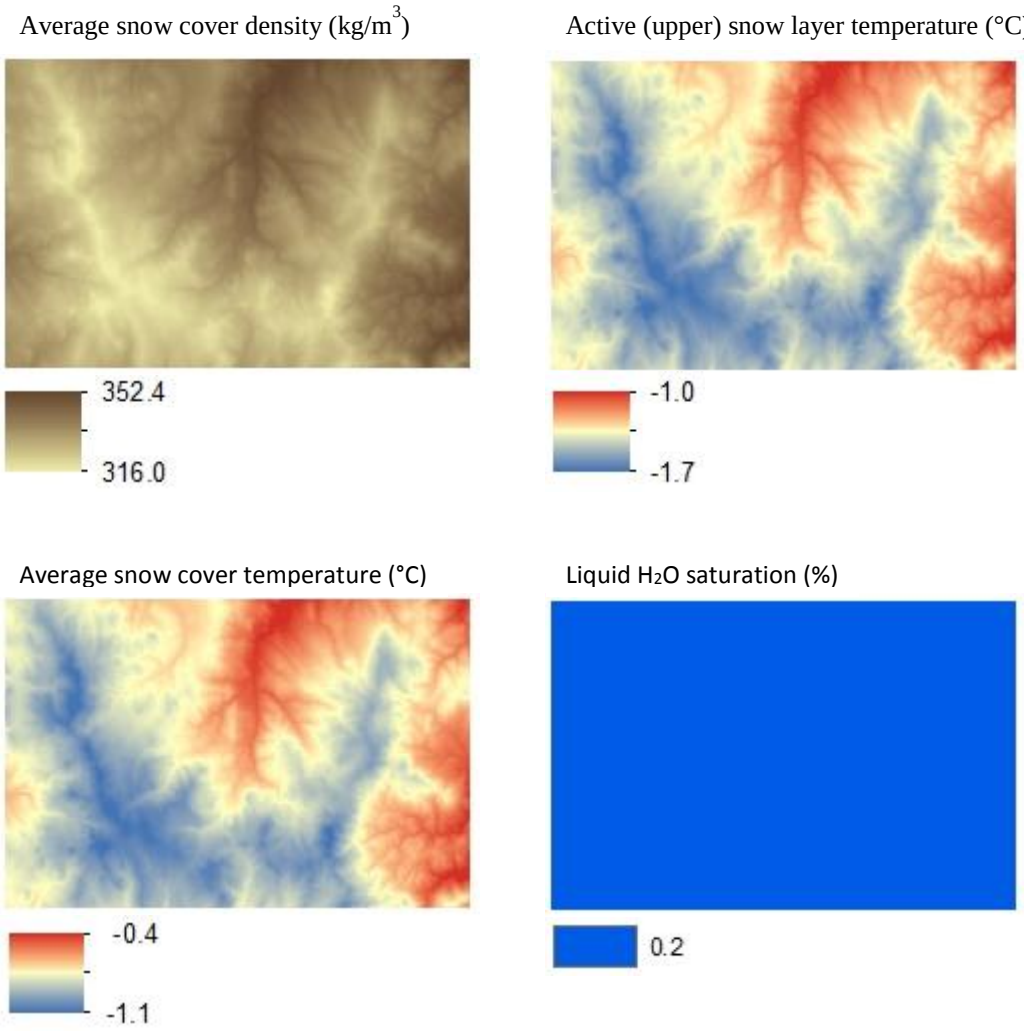


Figure 3. Initial conditions grids created using Python gridding scripts.

Precipitation grids also showed expected variability, as shown in Figure 4. Note how EBK predicted higher precipitation mass values in the west compared to the east. This was in accordance with local weather patterns, as most of the winter and early spring storms move over the RCEW from west to east (Hanson *et al.*, 2001). Thus, more precipitation was deposited on the western leeward slopes of the watershed, with less precipitation being available for the eastern slopes. Dew point temperatures also increased in the northern, lower elevations. Finally, the lack of variability in the *Portion*

of precipitation as snow and *Snow precipitation density* grids can be explained by the low precipitation temperatures associated with the time step (see Table 6).

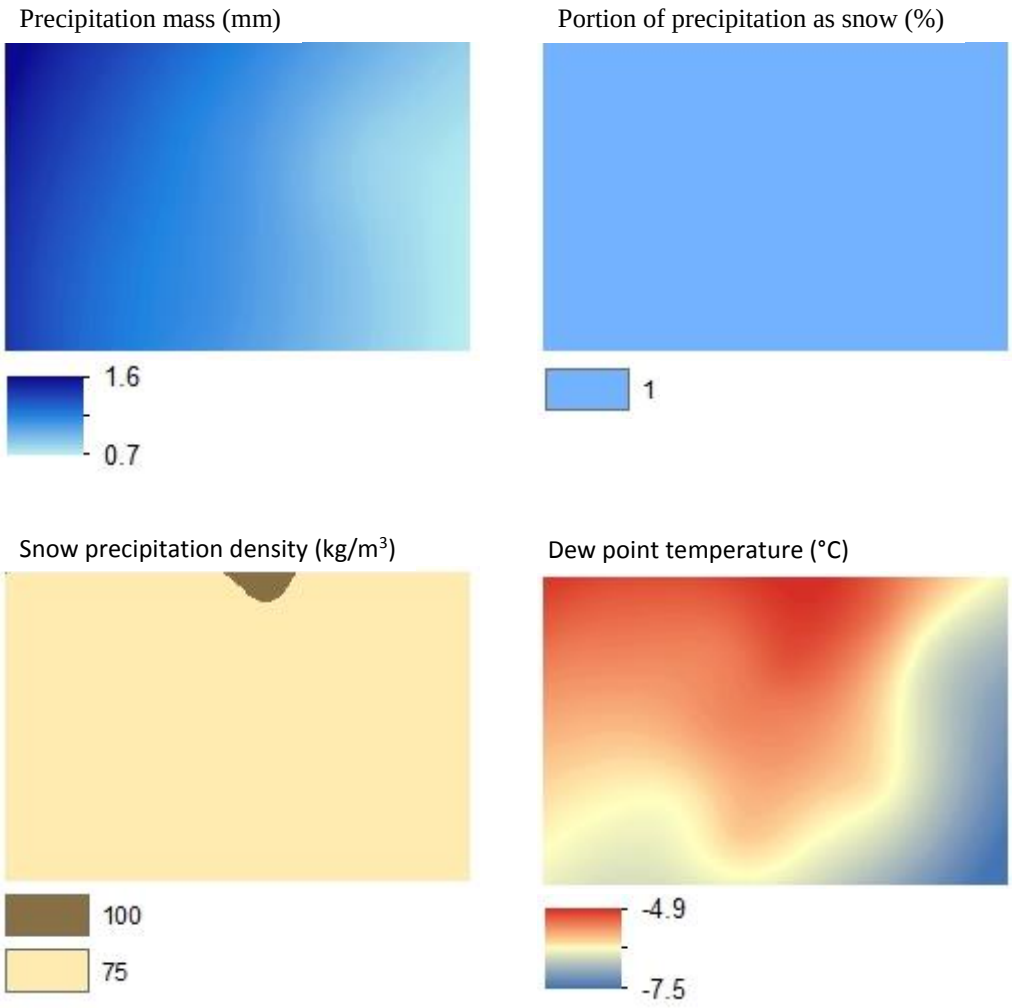
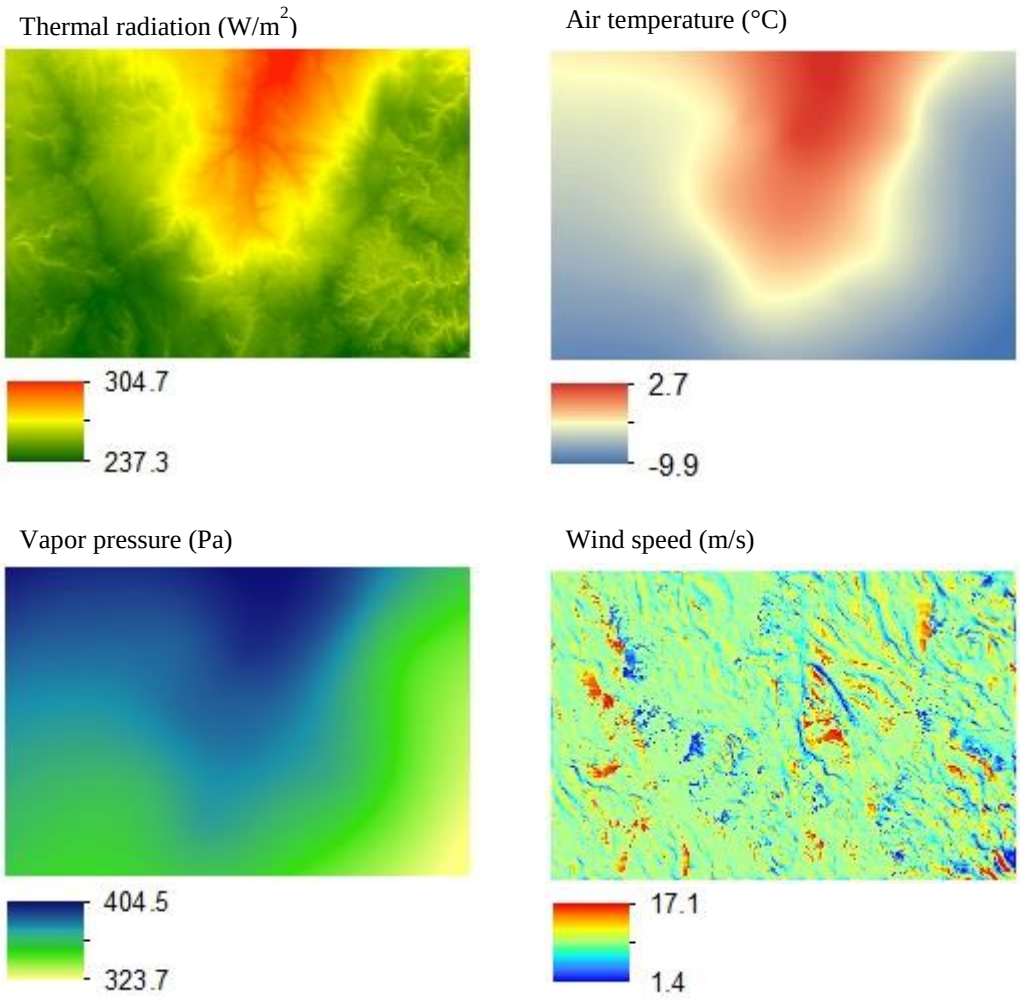


Figure 4. Precipitation grids created using the Python gridding scripts.

The input forcing data grids depicted in Figure 5 followed the pattern of expected variability as well. Air and soil temperatures were higher in the lower elevations, as were vapor pressures. Since the sources for thermal radiation include the atmosphere and surrounding terrain (Marks and Dozier, 1979), values were higher in the warmer valleys of the south when compared to the cooler ridges at higher

elevations. WindNinja accurately predicted higher wind speeds along ridges and lower wind speeds in valleys. Lastly, solar radiation values were higher for the south facing slopes than the north facing slopes, which agreed with early spring conditions in the northern hemisphere.



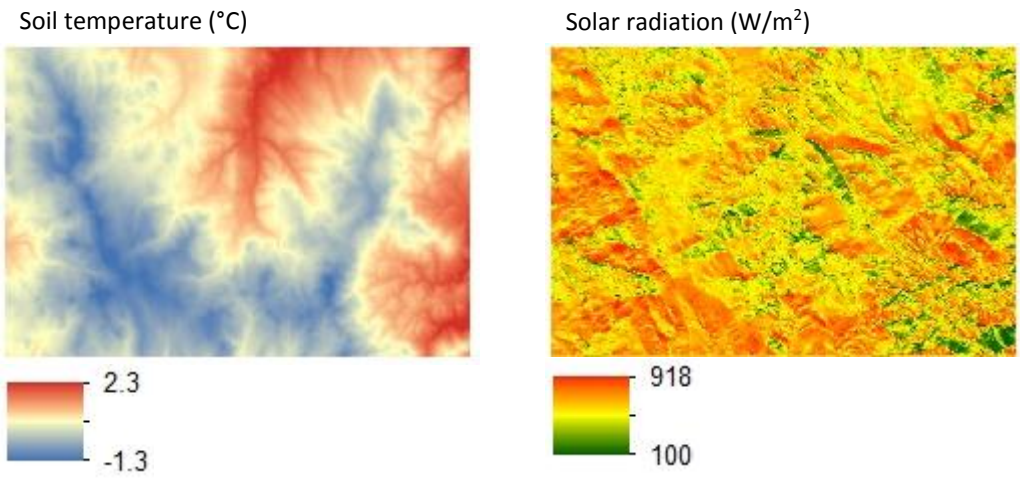


Figure 5. Input forcing data grids creating using the Python gridding tools.

The grids above were created by running the associated tools as a geoprocessing services from Esri's desktop software. The tools were published to a virtual machine containing 48 GB of RAM and 6 processing cores. In total, the 16 grids took approximately 1.5 hours to process, with the majority of the time being spent on grids that used EBK methods.

2.4 Discussion

The Python scripts developed in conjunction with this paper were created for the purpose of improving the iSNOBAL model, thus providing watershed scientists and researchers with an advanced set of tools. In creating these tools, the iSNOBAL model benefited in several ways. First, interpolation and parameter distribution methods were improved as several grids now use the more robust process of EBK. In addition, soil

temperature is no longer specified as a constant, but uses observed values to assist in predicting temperatures based on elevation. Other grids (snow properties and constants) allow for user input, which is valuable for two reasons: one, grids can now be created that are specific to location, researcher (study), or time-period and two, model sensitivity analyses can be performed with greater ease.

A second benefit of the tools is their widespread availability. The tools can be imported into Esri's ArcGIS desktop application or accessed as geoprocessing services using the same software. Furthermore, the tools can be executed as web processing services (WPS) via RESTful uniform resource locators (URLs). This means that anyone with an internet connection, regardless of operating system, can take advantage of the tools by incorporating them into web mapping applications and desktop GIS software.

The last objective of this work was to make the practice of input grid creation for the iSNOBAL model less hands-on and time-consuming. By using the more robust EBK methods, which automates the otherwise manual process of estimating semivariogram parameters, several of the grids can now be created in a much more efficient manner. Moreover, by using the Python scripting language, in conjunction with geoprocessing services, these tools can be implemented in an automated process.

As the tools take approximately 1.5 hours to create 16 input grids, it is important to note that not all 16 grids will be created for every time step. The 6 initial conditions image grids are created once per model run, and the 4 precipitation grids are created only during storm events. Of the 3 types of input grids, only the forcing data grids are created for every time step. This means that 5 to 6 grids will generally be created for

each time step since solar radiation grids are only created during daylight hours. This reduces the amount of computing time from 1.5 hours for all 16 grids to around 45 minutes for sunlight hours and 35 minutes for nighttime hours. Computing time can also be shortened by reducing the number of simulated semivariograms in the EBK function. By dropping the number from 100 to 50, grids that use EBK methods take roughly half as long to be created.

The following chapter describes improvements for implementing the scripts described above within the VWP. These improvements include building an SQL database to connect the tools with the full temporal range of available data, rather than the static one-hour data tables used above. Chapter 3 also discusses the development and use of a web interface to access and run the tools. Finally, cross-validation is performed to explore whether or not EBK methods are acceptable.

Chapter 3: Extending the functionality of the gridding tools with an SQL database and web interface

3.1 Introduction

The work described in this chapter builds upon the preceding work (described in Chapter 2) in three important ways. First, the tool's functionality is extended to the full temporal range of historical data available in the Reynolds Creek Experimental Watershed (RCEW) by using a structured query language (SQL) database. Second, to further address the concept of accessibility, a simple web interface is described for running the tools. This is done as a primary step towards implementing the tools within a virtual watershed platform (VWP). And third, cross-validation is performed to validate the use of the more robust empirical Bayesian kriging (EBK) methods. These additions to the work done in Chapter 2 are performed in order to provide a more complete set of tools for the virtual watershed platform (VWP) introduced in Chapter 1. Researchers studying watershed hydrological processes will benefit from these tools by having a simplified, automated, and improved means of running the iSNOBAL model.

In mentioning the improvements above, it is important to understand why they are needed. This in turn requires a discussion on two central aspects in most research workflows: data management and data integration within models. First consider the estimation that more scientific data will be generated in the next decade than has been

produced in all of human history (Horsburgh *et al.*, 2009). Furthermore, the data that has been collected, and that which will be collected, are heterogeneous in almost every characteristic, from the organization doing the collecting, to how it is described, organized, stored, and accessed (see Peckham and Goodall, 2013). From that list, the two elements relevant to this discussion are how data are stored and how it is accessed. The observation data from RCEW are stored as comma-separated-values files and are accessed through an anonymous file transfer protocol (FTP) network. While these methods of storage and access might be simple, they are somewhat outdated and limiting by nature. SQL databases address these issues by providing a simple interface for storing and retrieving data in an efficient manner (Jolly *et al.*, 2005). This efficiency is crucial when considering extremely large datasets, such as the RCEW, which collects several observational variables on an hourly basis. Similar studies have used SQL databases to store climate data for the purpose of monitoring micro-climates with agriculture applications (Ghobakhlou *et al.*, 2009).

The second aspect of a research workflow requiring discussion is that of data integration within models. Models have proven useful in nearly every field imaginable (Peckham *et al.*, 2012). This has certainly been the case with hydrological processes (Ly *et al.*, 2013). Unfortunately, the models themselves have not kept pace with advances in computer and data resources (Humphrey *et al*., 2012; Leonard and Duffy, 2013). One resource that is being underutilized is geoprocessing services offered through the web.

The remainder of this chapter describes the creation of an SQL database from historical observation data from the RCEW to improve data integration associated with

the gridding tools. It also details how the tools are made more widely accessible by implementing them in a web interface. Finally EBK methods are explored as an acceptable interpolation means through a relative comparison to detrended kriging methods using leave-one-out cross-validation.

3.2 Methods

3.2.1 SQL Database

The initial step in building the SQL database consisted of downloading the data files from the anonymous FTP site (see Appendix A) and storing them on a local hard drive. The downloaded files were divided into four directories based on type: climate, precipitation, snow depth, and soil temperature. Next, four empty tables were initiated within the SQL database consistent with the four different types just mentioned. Tables 7 – 10 show how these four database tables were set up prior to transferring observational data, and Figure 6 shows the corresponding database diagram.

Table 7. Specifications of the climate table initiated in the SQL database.

Parameter	Column Name	Data Type
Station name	Site_Key	varchar
Date and time of observation	date_time	datetime
Water year	wy	smallint
Water day	wd	smallint
Year	year	smallint
Month	month	smallint
Day	day	smallint

Hour	hour	smallint
Minute	minute	smallint
Air temperature	tmp3	double
Relative humidity	hum3	double
Vapor pressure	vap3	double
Dew point temperature	dpt3	double
Wind speed	wnd3sa	double
Wind direction	wnd3d	double

Table 8. Specifications of the precipitation table initiated in the SQL database.

Parameter	Column Name	Data Type
Station name	Site_Key	varchar
Date and time of observation	date_time	datetime
Shielded precipitation	ppts	double
Unshielded precipitation	pptu	double
Dual gage average precipitation	ppta	double

Table 9. Specifications of the snow depth table initiated in the SQL database.

Parameter	Column Name	Data Type
Station name	Site_Key	varchar
Date and time of observation	date_time	datetime
Water year	wy	smallint
Water day	wd	smallint
Year	year	smallint
Month	month	smallint
Day	day	smallint
Hour	hour	smallint
Minute	minute	smallint
Snow depth	snowdepth	double

Table 10. Specifications of the soil temperature table initiated in the SQL database.

Parameter	Column Name	Data Type
Station name	Site_Key	varchar
Date and time of observation	date_time	datetime
Soil temperature at 2.5 cm	stm002_5	double
Soil temperature at 5 cm	stm005	double

Soil temperature at 10 cm	stm010	double
Soil temperature at 15 cm	stm015	double
Soil temperature at 20 cm	stm020	double
Soil temperature at 30 cm	stm030	double
Soil temperature at 40 cm	stm040	double
Soil temperature at 50 cm	stm050	double
Soil temperature at 55 cm	stm055	double
Soil temperature at 60 cm	stm060	double
Soil temperature at 70 cm	stm070	double
Soil temperature at 90 cm	stm090	double
Soil temperature at 120 cm	stm120	double
Soil temperature at 180 cm	stm180	double

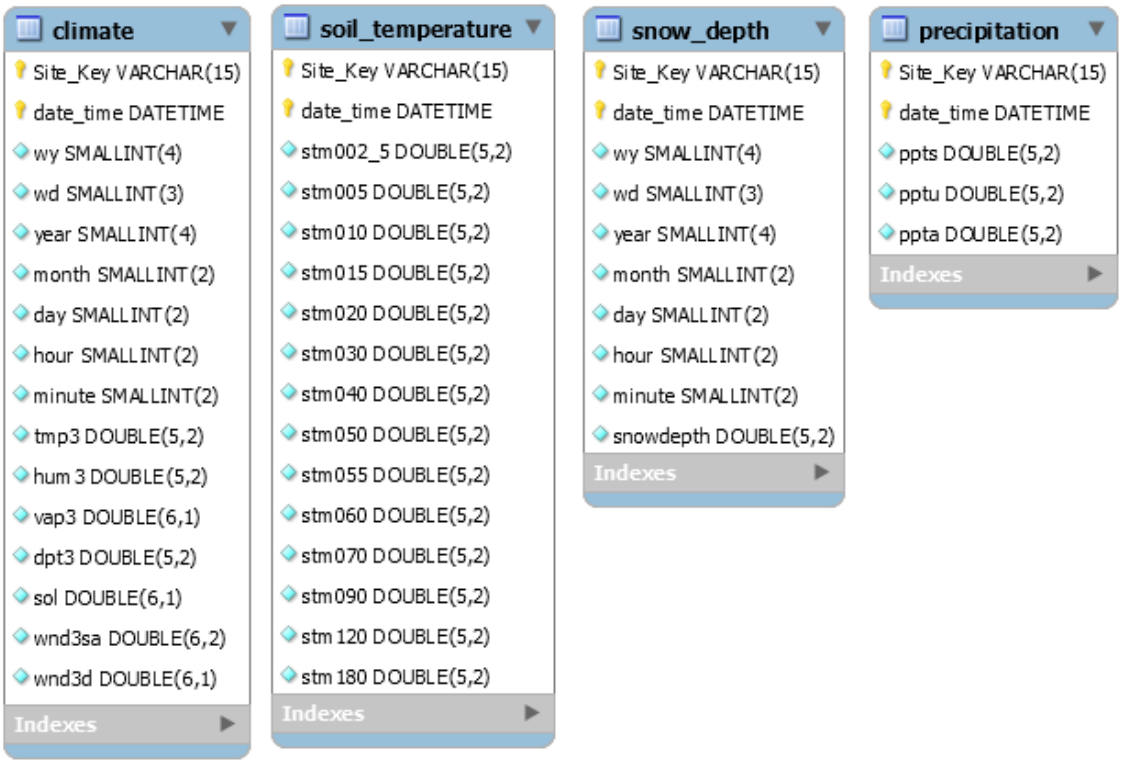


Figure 6. Database diagram showing the four tables within the SQL database and their associated columns.

Four corresponding Python scripts (see Appendix B) were written to transfer the data from the files to the database tables. As an example, the *Create Climate Table* script would start by listing the climate data files within the respective directory. Then,

after establishing a connection with the SQL database, the script would iterate over one data file at a time. Once the file was opened for reading, the script would iterate over the entire file one line at a time. This was done so that information could be extracted from the header of each data file, including the name of the weather station the file represented, and the list of parameters that the data file contained. Once these were obtained, the script continued by executing *INSERT* statements on each line of observation data. To make the process more efficient, rather than committing these *INSERT* statements to the database one line at a time, the commit was executed at the end of each data file. After the last data file was committed to the database, the connection to the database was closed, and the process was repeated for the remaining tables.

3.2.2 Geoprocessing Service and Web Interface

The Python scripts described in Chapter 2 were designed to create one grid per script (or, in some cases, multiple grids if they were related, such as initial snow properties). The approach taken in this chapter was to combine the functionality of all the previous gridding scripts into one master script. This master script was intended to simplify the publication of geoprocessing services by consolidating all of the functionality into one service, rather than having to publish and maintain multiple tools.

Input for this new tool included values for start and end date/time, a one or three hour time step, the type of kriging method, and which grids were to be created. An example of the tool dialog from the ArcMap desktop software is shown in Figure 7.

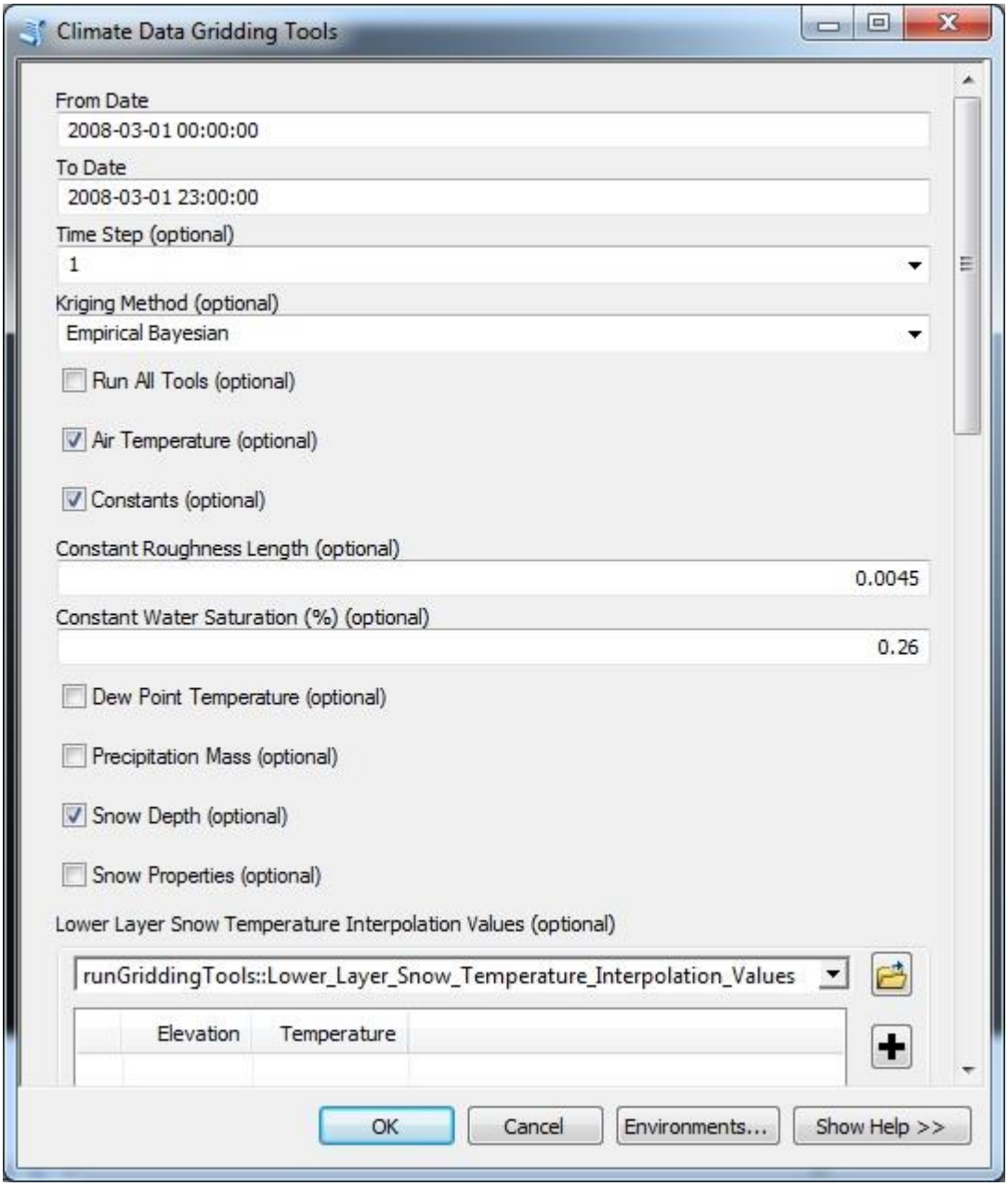


Figure 7. Example dialog box from running the master gridding tool in the ArcMap desktop software.

Given the above input, the tool was able to connect to the SQL database and create the desired grids by looping over the time range specified. Each iteration of the loop consisted of three general steps. First, the SQL database was queried using a *SELECT* statement, and observational values were selected based on the current time step. Next, results from the *SELECT* query were appended to parameter lists, which lists were then used in building data tables specific to each time step. Finally, these data tables were used as input for individual gridding functions within the script.

Once inside the separate gridding functions, grids were created according to the methods described in Chapter 2. However, one change is important to note. In the methods detailed in Chapter 2, for grids that used empirical Bayesian kriging, linear regression was used to predict parameters at stations that did not have observed values. This was done prior to the kriging interpolation. This method was removed from the current version of the tool, as it was believed to add unnecessary bias and error.

The completed script was published as a geoprocessing service on Esri's ArcServer on a server machine containing 48 GB of RAM and 6 processing cores. In publishing the service, the tool was made available through web processing services (WPS), and as a RESTful service. These can in turn be ingested into any platform, including web interfaces and web maps, desktop GIS software, and virtual watershed platforms.

3.2.3 Cross-Validation

To assess whether EBK methods produced acceptable interpolated grids, a simple leave-one-out cross-validation process was employed. First, multiple subsets of data were created by removing a single weather station and its associated observed parameter value. These subsets were then used in the gridding tool to generate predicted grids. Next, predicted values were extracted from the grids in the location of each missing weather station, and errors were calculated by subtracting observed values from predicted values.

This cross-validation process was performed for all 5 iSNOBAL input grids that used the EBK methods. Errors were calculated for air temperature, dew point temperature, and vapor pressure for four time steps across an entire year, once each in January, April, July, and October. Precipitation errors were calculated for two time steps- once for a winter storm in December, and again for a summer storm in July. Snow depth errors were also calculated for two time steps- once in December, and once in March.

To determine if EBK produced satisfactory results, the same cross-validation process was repeated on grids that were created using detrended kriging algorithms (see Susong *et al.*, 1999), and the two sets of errors were compared.

3.3 Results

The process described above for creating the SQL database resulted in a comprehensive dataset for the RCEW. Table 11 shows specific details related to each of the 4 tables within the SQL database, including the number of data files associated with each data type, the size of the SQL table (number of observations), and the date and time of the first and last observations.

Table 11. Description of RCEW historical data and the resulting SQL tables.

Data type	# of data files	Size of SQL table	First observation	Last observation
Climate	21	2,360,982	1981-06-18 11:00:00	2009-09-30 23:00:00
Precipitation	25	9,120,000	1962-01-01 01:00:00	2014-12-31 23:00:00
Snow Depth	29	3,568,752	1996-10-11 10:00:00	2008-10-01 00:00:00
Soil Temperature	5	1,154,088	1977-12-22 16:00:00	2014-10-01 00:00:00

Once the SQL database was created locally, it was exported to the same server machine that housed the geoprocessing service for running the gridding tool. This was done in order to connect the two resources in an efficient and reliable manner. Upon running the tool, the process of retrieving the necessary data for a single time step takes a matter of seconds.

To make the execution of the tool as simple as possible, a web interface was designed to collect user input, run the geoprocessing service, and return the resulting grids. A screenshot of that interface is shown in Figure 8. After specifying a beginning and ending date/time, and selecting a time step and the grids to be created, a user can run the process, and the subsequent grids are returned as downloadable .tif files. The

execution is carried out using the RESTful URL associated with the geoprocessing service. When run in this manner, a test case took approximately 1.5 hours to create 6 initial conditions grids, 4 precipitation grids, and 6 input forcing data grids.

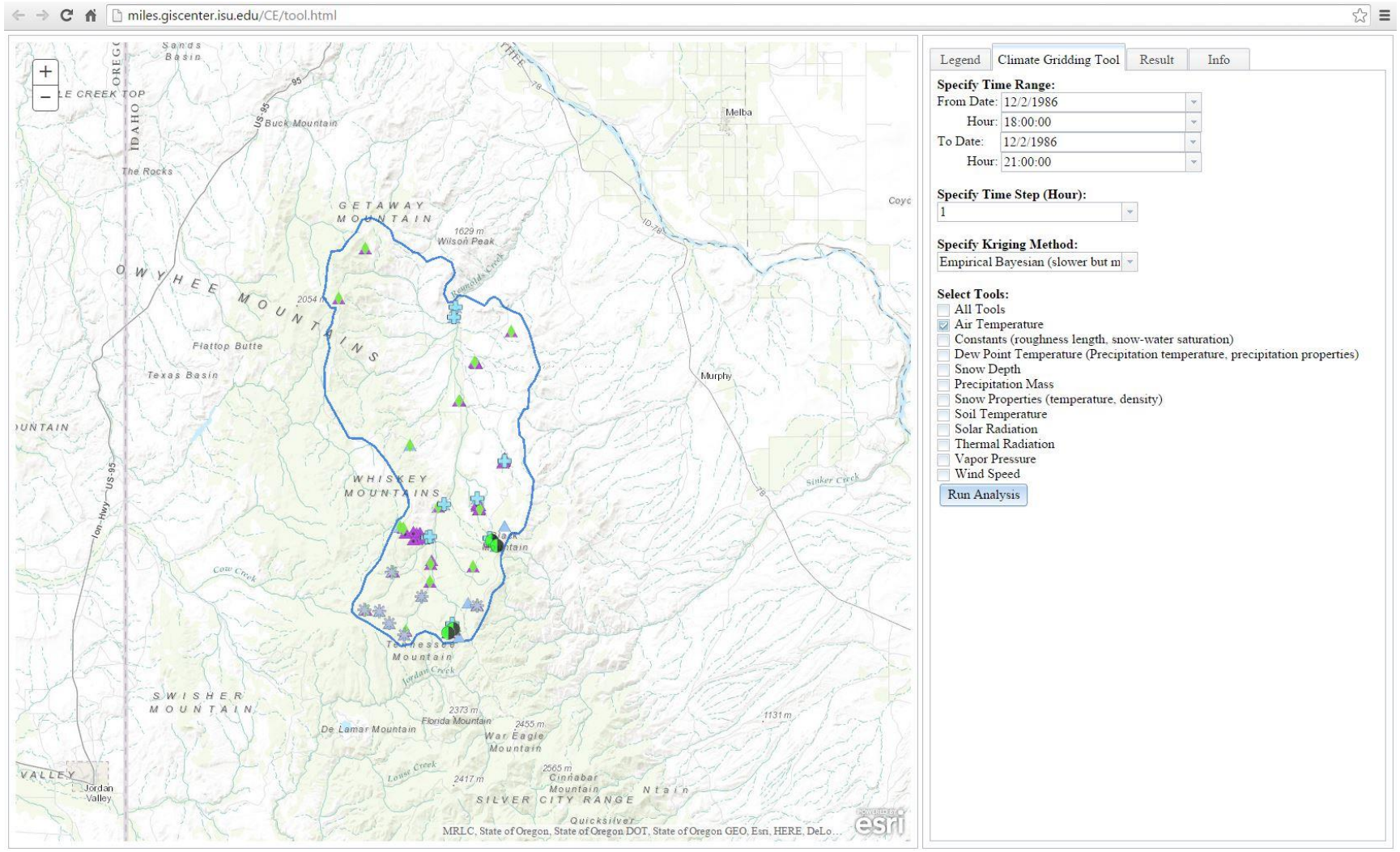


Figure 8. Screenshot of the web interface for running the geoprocessing service for creating iSNOBAL input grids.

The majority of processing time required to create the 16 input grids for a single time step was in association with the 5 grids that used EBK methods. To determine whether EBK produced appropriate results, leave-one-out cross-validation was performed on grids of air temperature, dew point temperature, vapor pressure, precipitation mass, and snow depth for multiple time steps. RMSE values were calculated between observed and predicted grids that were created using both empirical Bayesian kriging (EBK) methods, and detrended kriging algorithm (DKA) methods.

Table 12 shows the RMSE values for the three climate variables (air temperature, dew point temperature, and vapor pressure). In general, DKA methods performed slightly better than EBK methods, with EBK equaling or outperforming DKA in 3 of the 12 cases.

Table 12. Cross-validation results for air temperature, dew point temperature, and vapor pressure from 4 time steps. Numbers represent RMSE between observed and predicted values, with side-by-side comparison of empirical Bayesian kriging (EBK) and detrended kriging algorithm (DKA) methods.

	Air Temperature (˚C)		Dew Point Temperature (˚C)		Vapor Pressure (Pa)	
	EBK	DKA	EBK	DKA	EBK	DKA
January	1.8	1.9	2.5	2.1	36.3	37.5
April	1.9	0.9	0.7	0.6	11.5	10.6
July	1.3	0.6	0.8	0.8	52.8	50.4
October	1.0	0.4	0.9	0.4	36.2	19.9

EBK methods performed very well in predicting precipitation mass, as seen by the similar RMSE values in Table 13. For both the winter and summer storms, there was only a 0.1 mm difference between the EBK and DKA RMSE values.

Table 13. Cross-validation results for precipitation mass for a winter and summer storm. Numbers represent RMSE between observed and predicted values, with side-by-side comparison of empirical Bayesian kriging (EBK) and detrended kriging algorithm (DKA) methods.

	Precipitation Mass (mm)	
	EBK	DKA
Winter	1.6	1.7
Summer	1.0	0.9

Finally, even though Table 14 shows EBK once again equaling or outperforming DKA methods in estimating snow depth, the errors indicated by the results for both methods were unusually high. This was most likely due to the fact that neither method took into account aspect.

Table 14. Cross-validation results for snow depth for winter and early-spring time steps. Numbers represent RMSE between observed and predicted values, with side-by-side comparison of empirical Bayesian kriging (EBK) and detrended kriging algorithm (DKA) methods.

	Snow Depth (cm)	
	EBK	DKA
December	16.2	16.7
March	42.5	42.5

3.4 Discussion

The improvements discussed in this chapter build upon the tools described in Chapter 2. The first improvement involved building an SQL database to automate the process of data retrieval. Where the initial development of the tools dealt with static, single time step tables built by hand, the SQL database allows for efficient querying of the full temporal range of historical data from the RCEW. In addition, the SQL database can be used independent of the geoprocessing tool to identify data gaps, summarize

specific periods of data, etc. And while the database described herein contains only that data which was available at the time of creation, the process for adding more data as it is made available is simple and straightforward. Along these same lines, the tools can easily be modified to query custom databases, or databases from watersheds other than RCEW, to extend their functionality further.

The second improvement introduced in this chapter was the development of a web interface for the purpose of running the gridding tools. This was done using a Javascript API for ArcGIS Server. The simple interface collects user input, executes the gridding functions through a REST service, and returns the resulting grids as .tif files that can be downloaded for further use. This was done to not only extend the accessibility of the tool, but to make its execution more simple and direct.

With these two developments in mind, especially the ability to execute the tools from a web interface, it is important to remember that these are the initial steps towards implementing the iSNOBAL model in its entirety within the VWP. Further steps will be required to automate the process of uploading the created grids to the VWP so that they can be accessed by a model run from the platform itself.

This chapter also included a process for judging the relative accuracy of EBK interpolation methods when compared to DKA methods. The general finding was that EBK produced similar cross-validation results to DKA, especially for predicted precipitation mass and snow depth grids. However, cross-validation results for interpolated snow depth grids showed unusually high RMSE values for both empirical

Bayesian and detrended kriging methods. This was most likely due to the fact that neither of the two methods took into account aspect. In their study, Susong *et al.* (1999) created initial snow depth grids in a 3-step process. First, the same DKA methods were applied to snow-water-equivalent (SWE) observations to create a snow mass grid. Second, to account for differences in solar radiation between the two aspects, grid cells that fell on north facing slopes were multiplied by 1.2 and grid cells that fell on south facing slopes were multiplied by 0.8. And third, this modified snow mass grid was divided by a snow density grid to obtain a grid representing initial snow depths. Rather than interpolating SWE values and dividing by density, the current practice of the gridding tool interpolates observed snow depth values directly.

The results of the cross-validation process established EBK as an acceptable means for creating several of the input grids for iSNOBAL. An added benefit of EBK over DKA methods is the aspect of automation. Where DKA requires a user to estimate semivariogram parameters for each grid that employs the method, and then to repeat the process for every single time step, EBK eliminates this need by finding the best fit values systematically.

Chapter 4: Conclusion

The iSNOBAL model is an effective means of characterizing the development and melting of a snowpack in mountainous watersheds. However, several limitations exist in association with the model. This thesis describes work that was done to address these limitations. The product of this work is a unique set of tools which automate and simplify the creation of input grids for iSNOBAL by using the Python scripting language, an SQL database, and geoprocessing services.

The original methods of Marks *et al.* (1999) used a detrending kriging algorithm to produce several of the input grids for iSNOBAL. It was hypothesized that by using a novel approach in the form of empirical Bayesian kriging (EBK), the process for creating these grids could be made more efficient, and the accuracy of the grids themselves would improve. Initial cross-validation results showed that there was little improvement in accuracy between detrended kriging and EBK methods. However, cross-validation results were similar between the two methods, with EBK having the advantage of automating the estimation of semivariogram parameters using a Monte-Carlo methodology, whereas the detrended kriging approach requires a time-consuming process of adjusting the parameters for each grid. In order to take full advantage of both approaches, future work will adapt these two methods to work in tandem. This will be done by using the original detrending methods to produce a set of stationary residuals, using EBK to distribute the residuals, and then adding back the elevation trend.

Improvements were made to other parameter distribution methods as well. Marks *et al.* suggested using a constant 0 °C grid for soil temperature. The methods described in this thesis use elevation gradients calculated from observed soil temperature values to provide more realistic estimations. In addition, thermal radiation and wind speed grids are created using advanced simulations. Cross-validation will be performed in a future work to determine if these methods provide improved accuracy.

In an effort to more fully automate the process of grid creation, an SQL database was created containing historical data from the Reynolds Creek Experimental Watershed. The use of an SQL database is an improvement over current practices—which include storing the data as comma-separated-values files on an FTP network—by providing a simple and efficient technique for data storage and retrieval. The value of building an SQL database does not end with its integration in the gridding tools, however, as the database can be used as an independent resource for effective data analysis, including the identification of data gaps. Furthermore, the procedures for updating the database and creating additional databases are simplified through the use of SQL (see Appendix D), thus making it easy to run the gridding tools on current data and data from other watersheds.

A main objective of the work presented in this thesis was to make the process of creating input grids for the iSNOBAL model more efficient. The gridding tools described herein are an improvement to the IPW's command line interface methods, and provide a simple means for automating grid creation. However, work still needs to be done to improve computing efficiencies. The largest amount of computing time is spent on

parameters that use EBK, as well as solar radiation and wind speed simulations. Computing time can be reduced by specifying fewer simulated semivariograms in the EBK function, as well as running the WindNinja command line interface with more processors.

The final improvement made to the iSNOBAL model relates to its accessibility. Currently, the model and the associated gridding functionalities are confined to the IPW software and the Linux operating system. Once it is fully integrated within the VWP, iSNOBAL will be accessible to any researcher with an internet connection, regardless of operating system. This is made possible by using geoprocessing services and RESTful URLs. The gridding tools—in the form of geoprocessing services—along with the associated SQL database, can be consumed by platforms such as HydroShare (Tarboton *et al.*, 2014) and the WC-WAVE VWP, desktop GIS software such as ArcGIS, and web mapping applications.

In an effort to make the iSNOBAL model easier to run, all of these improvements were made by using computing and data resources commonly accessible by researchers through online tools or through GIS desktop software. The resulting gridding tools can assist researchers in exploring the effects of a changing climate on snowpack and other watershed dynamics by enhancing data sharing, modeling, and analysis. This in turn will lead to improved management practices in a future where the supply and demand of water resources are uncertain.

References

Anderson, E. A., 1973, National Weather Service River Forecast System - Snow Accumulation and Ablation Model: NOAA Technical Memorandum NWS HYDRO-17, 87 p.

Averyt, K., Meldrum, J., Caldwell, P., Sun, G., McNulty, S., Huber-Lee, a, and Madden, N., 2013, Sectoral contributions to surface water stress in the coterminous United States: Environmental Research Letters, v. 8, p. 035046, doi: 10.1088/1748-9326/8/3/035046.

Brown, T.C., Foti, R., and Ramirez, J. a., 2013, Projected freshwater withdrawals in the United States under a changing climate: Water Resources Research, v. 49, p. 1259–1276, doi: 10.1002/wrcr.20076.

Chikamoto, Y., Timmermann, A., Stevenson, S., DiNezio, P., and Langford, S., 2015, Decadal predictability of soil water, vegetation, and wildfire frequency over North America: Climate Dynamics, doi: 10.1007/s00382-015-2469-5.

Cooper, H., Zhang, C., and Selch, D., 2015, Incorporating uncertainty of groundwater modeling in sea level rise assessment: a case study in South Florida: Climatic Change, v. 129, p. 281–294.

Deng, F., Minasny, B., Knadel, M., McBratney, A., Heckrath, G., and Greve, M.H., 2013, Using Vis-NIR Spectroscopy for Monitoring Temporal Changes in Soil Organic Carbon: Soil Science, v. 178, p. 389–399, doi: 10.1097/SS.0000000000000002.

Dozier, J., and Outcalt, S.I., 1979, An approach toward energy balance simulation over rugged terrain: Geographical Analysis, v. 11, p. 65–85.

Drexler, J.Z., Knifong, D., Tuil, J., Flint, L.E., and Flint, A.L., 2013, Fens as whole-ecosystem gauges of groundwater recharge under climate change: Journal of Hydrology, v. 481, p. 22–34, doi: 10.1016/j.jhydrol.2012.11.056.

Flerchinger, G.N., and Saxton, K.E., 1989, Snow-Residue-Soil System I . Theory and Development: Soil and Water Division ASAE, v. 32, p. 565–567.

Forthofer, J., Shannon, K., and Butler, B., 2009, Simulating diurnally driven slope winds with WindNinja, *in* Proceedings of 8th Symposium on Fire and Forest Meteorological Society, p. 13.

Franz, K.J., Hogue, T.S., and Sorooshian, S., 2008, Operational snow modeling: Addressing the challenges of an energy balance model for National Weather

Service forecasts: Journal of Hydrology, v. 360, p. 48–66, doi: 10.1016/j.jhydrol.2008.07.013.

Frew, J.E., 1990, The Image Processing Workbench: University of California Santa Barbara, 382 p.

Ghobakhlou, a, Shanmuganthan, S., and Sallis, P., 2009, Wireless Sensor Networks for Climate Data Management Systems: 18th World IMACS / MODSIM Congress, p. 959–965.

Godsey, S.E., Kirchner, J.W., and Tague, C.L., 2013, Effects of changes in winter snowpacks on summer low flows: Case studies in the Sierra Nevada, California, USA: Hydrological Processes, v. 5064, p. 5048–5064, doi: 10.1002/hyp.9943.

Goodwin, P., Delparte, D., Flores, A., Sheneman, L., and Glenn, N., 2013, Collaborative Research: Western Consortium for Watershed Analysis, Visualization, and Exploration (WC-WAVE), p. 1–20.

Hamlet, A.F., Mote, P.W., Clark, M.P., and Lettenmaier, D.P., 2005, Effects of temperature and precipitation variability on snowpack trends in the sestern United States: Journal of Climate, v. 18, p. 4545–4561, doi: 10.1175/JCLI3538.1.

Hansen, Z.K., Lowe, S.E., and Xu, W., 2014, Long-term impacts of major water storage facilities on agriculture and the natural environment: Evidence from Idaho (U.S.): Ecological Economics, v. 100, p. 106–118, doi: 10.1016/j.ecolecon.2014.01.015.

Hanson, C.L., Marks, D., and Van Vactor, S.S., 2001, Long-term climate database, Reynolds Creek Experimental Watershed, Idaho, United States: Water Resources Research, v. 37, p. 2839–2841, doi: 10.1029/2001WR000417.

Harpold, A. a., Molotch, N.P., Musselman, K.N., Bales, R.C., Kirchner, P.B., Litvak, M., and Brooks, P.D., 2014, Soil moisture response to snowmelt timing in mixed-conifer subalpine forests: Hydrological Processes, p. 2782–2798, doi: 10.1002/hyp.10400.

Hinckley, E.L.S., Ebel, B. a., Barnes, R.T., Anderson, R.S., Williams, M.W., and Anderson, S.P., 2014, Aspect control of water movement on hillslopes near the rain-snow transition of the Colorado Front Range: Hydrological Processes, v. 28, p. 74–85, doi: 10.1002/hyp.9549.

Horsburgh, J.S., Tarboton, D.G., Piasecki, M., Maidment, D.R., Zaslavsky, I., Valentine, D., and Whitenack, T., 2009, An integrated system for publishing environmental observations data: Environmental Modelling and Software, v. 24, p. 879–888, doi: 10.1016/j.envsoft.2009.01.002.

Humphrey, M., Beekwilder, N., Goodall, J.L., and Ercan, M.B., 2012, Calibration of watershed models using cloud computing, *in* 2012 IEEE 8th International Conference on E-Science, p. 1–8.

Hutson, S.S., Barber, N.L., Kenny, J.F., Linsey, K.S., Lumia, D.S., and Maupin, M.A., 2004, Estimated use of water in the United States in 2000: *U.S. Geological Survey Circular No. 1268*.

Jolly, W.M., Graham, J.M., Michaelis, A., Nemani, R., and Running, S.W., 2005, A flexible, integrated system for generating meteorological surfaces derived from point sources across multiple geographic scales: Environmental Modelling and Software, v. 20, p. 873–882, doi: 10.1016/j.envsoft.2004.05.003.

Jordan, R., 1991, A One-Dimensional Temperature Model for a Snow Cover: Technical Documentation for SNTHERM.89: Unbc.Ca, p. 49.

Kormos, P.R., Marks, D., Williams, C.J., Marshall, H.P., Aishlin, P., Chandler, D.G., and McNamara, J.P., 2013, Soil, snow, weather, and sub-surface storage data from a mountain catchment in the rain–snow transition zone: Earth System Science Data Discussions, v. 6, p. 811–835, doi: 10.5194/essdd-6-811-2013.

Krivoruchko, K., 2012, Empirical Bayesian Kriging: ESRI Press, v. Fall 2012, p. 6–10.

Kumar, M., Marks, D., Dozier, J., Reba, M., and Winstral, A., 2013, Evaluation of distributed hydrologic impacts of temperature-index and energy-based snow models: Advances in Water Resources, v. 56, p. 77–89, doi: 10.1016/j.advwatres.2013.03.006.

Laurent, O., Hu, J., Li, L., Cockburn, M., Escobedo, L., Kleeman, M.J., and Wu, J., 2014, Sources and contents of air pollution affecting term low birth weight in Los Angeles County, California, 2001-2008: Environmental Research, doi: 10.1016/j.envres.2014.05.003.

Leavesley, G.H., Lichty, R.W., Troutman, B.M., and Saindon, L.G., 1983, Precipitation-Runoff Modeling System: User´s Manual: U.S. Geological Survey, Water-Resources Investigations Report 83-4238, p. 207.

Leonard, L., and Duffy, C.J., 2013, Essential terrestrial variable data workflows for distributed water resources modeling: Environmental Modelling and Software, v. 50, p. 85–96, doi: 10.1016/j.envsoft.2013.09.003.

Lindsay, J.B., 2014, The Whitebox Geospatial Analysis Tools project and open-access GIS, *in* Proceedings of the GIS research UK 22nd annual conference, p. 16–18.

Lute, a. C., Abatzoglou, J.T., and Hegewisch, K.C., 2015, Projected changes in snowfall extremes and interannual variability of snowfall in the western United States (Accepted Article): Water Resources Research, v. 51, p. 960–972, doi: 10.1002/2014WR016267.

Ly, S., Charles, C., and Degré, A., 2013, Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review.: Biotechnol. Agron. Soc. Environ., v. 17, p. 392–406.

Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D., and Reichle, R., 2012, Soil Moisture, Snow, and Seasonal Streamflow Forecasts in the United States: Journal of Hydrometeorology, v. 13, p. 189–203, doi: 10.1175/JHM-D-11-046.1.

Marks, D., Domingo, J., and Frew, J., 1999, Software tools for hydro-climatic modeling and analysis: Image Processing Workbench, ARS - USGS Version 2: ARS Technical Bulletin (Electronic Document: http://www.nwrc.ars.usda.gov/ipw), Electronic Document: http://www.nwrc.ars.usda.gov/.

Marks, D., Domingo, J., and Susong, D., 1999, A spatially distributed energy balance snowmelt model for application in mountain basins: Hydrological Processes, v. 13, p. 1935–1959.

Marks, D., and Dozier, J., 1979, A clear-sky longwave radiation model for remote alpine areas: Archiv fur Meteorologie, Geophysik und Bioklimatologie Serie B, v. 27, p. 159–187, doi: 10.1007/BF02243741.

Mathys, A., Coops, N.C., and Waring, R.H., 2014, Soil water availability effects on the distribution of 20 tree species in western North America: Forest Ecology and Management, v. 313, p. 144–152, doi: 10.1016/j.foreco.2013.11.005.

Michaelis, C.D., and Ames, D.P., 2008, Evaluation and Implementation of the OGC Web Processing Service for Use in Client-Side GIS: GeoInformatica, v. 13, p. 109–120, doi: 10.1007/s10707-008-0048-1.

Mote, P.W., Hamlet, A.F., Clark, M.P., and Lettenmaier, D.P., 2005, Declining mountain snowpack in western north America: Bulletin of the American Meteorological Society, v. 86, p. 39–49, doi: 10.1175/BAMS-86-1-39.

Mulcan, A., Mitsova, D., Hindle, T., Hanson, H., and Coley, C., 2015, Marine Benthic Habitats and Seabed Suitability Mapping for Potential Ocean Current Energy Siting Offshore Southeast Florida: Journal of Marine Science and Engineering, v. 3, p. 276–298, doi: 10.3390/jmse3020276.

National Research Council, 1999, New Strategies for America's Watersheds: Committee on Watershed Management, National Research Council, p. 328 S.

Pagano, T.C., Garen, D., and Sorooshian, S., 2004, Evaluation of Official Western U. S. Seasonal Water Supply Outlooks , 1922 – 2002: Journal of Hydrometeorology, v. 5, p. 896–909, doi: 10.1175/1525-7541(2004)005<0896:EOOWUS>2.0.CO;2.

Peckham, S.D., and Goodall, J.L., 2013, Driving plug-and-play models with data from web services: A demonstration of interoperability between CSDMS and CUAHSI-HIS: Computers and Geosciences, v. 53, p. 154–161, doi: 10.1016/j.cageo.2012.04.019.

Pedersen, U.B., Midzi, N., Mduluza, T., Soko, W., Vennervald, B.J., Mukaratirwa, S., and Kristensen, T.K., 2014, Modelling spatial distribution of snails transmitting parasitic worms with importance to human and animal health and analysis of distributional changes in relation to climate: v. 8, p. 335–343.

Pilz, J., and Spöck, G., 2008, Why do we need and how should we implement Bayesian kriging methods: Stochastic Environmental Research and Risk Assessment, v. 22, p. 621–632, doi: 10.1007/s00477-007-0165-7.

Safeeq, M., Grant, G.E., Lewis, S.L., and Tague, C.L., 2013, Coupling snowpack and groundwater dynamics to interpret historical streamflow trends in the western United States: Hydrological Processes, v. 27, p. 655–668, doi: 10.1002/hyp.9628.

Schaible, G.D., and Aillery, M.P., 2012, US irrigated agriculture: water management and conservation: Agricultural Resources and Environmental Indicators, 2012 edition, p. 29.

Schelker, J., Kuglerová, L., Eklöf, K., Bishop, K., and Laudon, H., 2013, Hydrological effects of clear-cutting in a boreal forest - Snowpack dynamics, snowmelt and streamflow responses: Journal of Hydrology, v. 484, p. 105–114, doi: 10.1016/j.jhydrol.2013.01.015.

Schubert, S., Koster, R., Hoerling, M., Seager, R., Lettenmaier, D., Kumar, A., and Gutzler, D., 2007, Predicting drought on seasonal-to-decadal time scales: Bulletin of the American Meteorological Society, v. 88, p. 1625–1630, doi: 10.1175/BAMS-88-10-1625.

Susong, D., Marks, D., and Garen, D., 1999, Methods for developing time-series climate surfaces to drive topographically distributed energy- and water-balance models: Hydrological Processes, v. 13, p. 2003–2021.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., and Stephens, S., 2002, The Drought Monitor: Bulletin of the American Meteorological Society, v. 83, p. 1181–1190.

Swain, N.R., Latu, K., Christensen, S.D., Jones, N.L., Nelson, E.J., Ames, D.P., and Williams, G.P., 2015, A review of open source software solutions for developing water resources web applications: Environmental Modelling & Software, v. 67, p. 108–117, doi: 10.1016/j.envsoft.2015.01.014.

Tarboton, D.G., Idaszak, R., Horsburgh, J.S., Heard, J., Ames, D., Goodball, J.L., Merwade, V., Couch, A., Arrigo, J., Hooper, R., Valentine, D., and Maidment, D.R., 2014, HydroShare: Advancing Collaboration through Hydrologic Data and Model Sharing: International Environmental Modelling and Software Society (iEMSs) 7th International Congress on Environmental Modelling and Software, doi: 978-88-9035-744-2.

Tarboton, D.G., Jackson, T.H., Liu, J.Z., Neale, C.M.U., Cooley, K.R., and McDonnell, J.J., 1995, A grid based distributed hydrologic model: Testing against data from Reynolds Creek Experimental Watershed: Proceedings of the American Meterological Society Conference on Hydrology, p. 77–90.

Tidwell, V.C., Moreland, B.D., Zemlick, K.M., Roberts, B.L., Passell, H.D., Jensen, D., Forsgren, C., Sehlke, G., Cook, M. a, King, C.W., and Larsen, S., 2014, Mapping water availability, projected use and cost in the western United States: Environmental Research Letters, v. 9, p. 064009, doi: 10.1088/1748-9326/9/6/064009.

Wang, R., Kumar, M., and Marks, D., 2013, Anomalous trend in soil evaporation in a semi-arid, snow-dominated watershed: Advances in Water Resources, v. 57, p. 32–40, doi: 10.1016/j.advwatres.2013.03.004.

Western Tri-State Consortium [Online], 2015, Westernconsortium.org.

Winograd, I.J., Riggs, A.C., and Coplen, T.B., 1998, The relative contributions of summer and cool-season precipitation to groundwater recharge, Spring Mountains, Nevada, USA: Hydrogeology Journal, v. 6, p. 77–93, doi: 10.1007/s100400050135.

Winstral, A., and Marks, D., 2002, Simulating wind fields and snow redistribution using terrain-based parameters to model snow accumulation and melt over a semi-arid mountain catchment: Hydrological Processes, v. 16, p. 3585–3603, doi: 10.1002/hyp.1238.

Wood, A.W., and Lettenmaier, D.P., 2008, An ensemble approach for attribution of hydrologic prediction uncertainty: Geophysical Research Letters, v. 35, p. 1–5, doi: 10.1029/2008GL034648.

Appendix A: Data Description and Source

The following table lists the climate station data that were used in this thesis.

Data Type	Parameter	Units	# of Stations	Temporal Resolution	Years of Record
Climate	air temperature vapor pressure dew point temperature solar radiation wind speed wind direction	°C Pa °C °C W/m^2 m/s ° from N	21	hourly	1981—2009
Precipitation	shielded unshielded averaged	mm mm mm	25	hourly	1962—2014
Snow	snow depth	cm	29	quarter-hourly	1996—2008
Soil Temperature	soil temperature at 5 cm depth	°C	5	hourly	1977—2014

Northwest Watershed Research Center, Reynolds Creek Experimental Watershed, Public-use Datafiles and Documentation, ftp://ftp.nwrc.ars.usda.gov/publicdatabase/, accessed Nov. 2014.

Appendix B: Python Code Repository

The gridding tools mentioned throughout this thesis that convert weather station data to distributed grids are derived from several Python scripts. These scripts are made publicly available under the MIT free software license, and can be found at:

https://github.com/delparte/WCWAVE

Appendix C: WindNinja Supplemental Files

The tables in this appendix represent supplemental files that are used by the WindNinja software to create estimated wind speed grids. The first table represents an example weather station CSV containing observed wind speeds and other parameters, along with station information. This CSV is used for initialization in the WindNinja simulations. The second table represents how a WindNinja configuration file is formatted and lists several variables pertaining to a WindNinja simulation run.

Example WindNinja weather station CSV file for a one-hour time step from March 1st, 2008 at 12:00 PM.

Station_Name	Coord_Sys (PROJCS,GEOGCS)	Datum (WGS84,NAD83,NAD27)	Lat/YCoord	Lon/XCoord	Height	Height_Units (meters,feet)	Speed	Speed_Units (mph,kph,mps)
012	PROJCS	NAD83	4793787.279	513952.8347	3	meters	4.025	mps
031	PROJCS	NAD83	4790459.234	512180.8464	3	meters	7.55	mps
076	PROJCS	NAD83	4783622.941	520288.0621	3	meters	6.35	mps
095b	PROJCS	NAD83	4780655.111	516986.352	3	meters	6.425	mps
124	PROJCS	NAD83	4775179.724	516317.994	3	meters	12	mps
124b	PROJCS	NAD83	4775115.638	516544.4316	3	meters	3	mps
125	PROJCS	NAD83	4774528.367	518189.0407	3	meters	3.6	mps
127	PROJCS	NAD83	4776395.437	521668.1306	3	meters	8.525	mps
128	PROJCS	NAD83	4775272.174	523390.85	3	meters	18.325	mps
138d03	PROJCS	NAD83	4774415.125	522515.3166	3	meters	12.1	mps
138j10	PROJCS	NAD83	4774199.4	522563.7815	3	meters	8.8	mps
138L21	PROJCS	NAD83	4773910.564	522800.5003	3	meters	19	mps
144	PROJCS	NAD83	4772188.618	515872.1455	3	meters	7.5	mps
145	PROJCS	NAD83	4772697.028	518399.46	3	meters	4.2	mps
163	PROJCS	NAD83	4769628.697	514056.6037	3	meters	8.6	mps
166b	PROJCS	NAD83	4768561.291	520063.2138	3	meters	10	mps
167	PROJCS	NAD83	4769981.191	521523.3537	3	meters	5.2	mps
174	PROJCS	NAD83	4768221.621	516735.8817	3	meters	3.2	mps
176	PROJCS	NAD83	4768127.662	519612.6749	3	meters	8.5	mps
rmsp3	PROJCS	NAD83	4768321.947	519977.3673	3	meters	2.9	mps

Direction (degrees)	Temperature	Temperature_Units (F,C)	Cloud_Cover (%)	Radius_of_Influence	Radius_of_Influence_Units (miles,feet,meters,km)
304	0.25	C	0	15	miles
306.75	-2.075	C	0	15	miles
287.5	5.175	C	0	15	miles
281.25	0.6	C	0	15	miles
288	-2.5	C	0	15	miles
297	-0.8	C	0	15	miles
266	1.6	C	0	15	miles
285.25	0.175	C	0	15	miles
286.25	-4.75	C	0	15	miles
275	-3.4	C	0	15	miles
275	-3.9	C	0	15	miles
281	-4.5	C	0	15	miles
291	-1.3	C	0	15	miles
261	0.2	C	0	15	miles
302	-5.9	C	0	15	miles
292	-5.2	C	0	15	miles
286	-3.6	C	0	15	miles
224	-3.8	C	0	15	miles
270	-4.6	C	0	15	miles
231.5	-4.85	C	0	-1	miles

initialization_method	=	pointInitialization
num_threads	=	4
elevation_file	=	elevation_raster
match_points	=	true
year	=	strYear,
month	=	strMonth,
day	=	strDay,
hour	=	strHour,
minute	=	strMinute,
mesh_resolution	=	output_cell_size
vegetation	=	brush
time_zone	=	America/Boise
diurnal_winds	=	true
write_goog_output	=	false
write_shapefile_output	=	false
write_farsite_atm	=	false
write_ascii_output	=	true
units_mesh_resolution	=	m
units_output_wind_height	=	m
output_speed_units	=	mph
output_wind_height	=	3
wx_station_filename	=	station_file

Appendix D: Considerations for applying methods to additional databases

The SQL database that was built in conjunction with this thesis, which contains historical weather station data from the Reynolds Creek FTP, was created using the MySQL database platform. In order to apply the gridding tools described in this thesis to new or custom databases, the following conventions must be strictly adhered to:

- Observational data must be separated into comma-separated values (CSV) files based on recording (climate, soil, etc.) station (one file per station)
- CSV files must be named with the following convention:

  [data type]_[station name].csv

  Example- “precipitation_124.csv”
- Each CSV file must contain an appropriate header in the first row

  For climate/weather (depending on parameters observed at each station):

  “date_time, ta, rh, ea, td, si, ws, wd”

  For precipitation:

  “date_time, ppts, pptu, ppta”

  For snow depth:

  “date_time, zs”

  For soil temperature (depending on measurement depths at each station):

  “date_time, st005, st010, st020,…”
- The “date_time” column must be in the following format:

  YYYY-MM-DD HH:MM
- “No-data” must be specified by “-999”