

AN ABSTRACT OF THE THESIS OF

Matthew Guy Cooper for the degree of Master of Science in Water Resources Science presented on May 1, 2015.

Title: Modeling Elevation-Dependent Snow Sensitivity to Climate Warming in the Data Sparse Eastern Oregon Cascades.

Abstract approved: _____

Anne W. Nolin

In the mountains of the Western US, shifts in the timing and magnitude of snow water equivalent (SWE) over the past century are well documented and attributed to climate warming, but the magnitude of sensitivity depends on elevation. We modeled the spatial distribution of SWE and its sensitivity to climate warming in the 1500 km² Upper Deschutes River Basin, Oregon, with a spatially distributed snowpack energy balance model forced by a gridded meteorological dataset. The forcing data, gridded at a spatial scale of 1/16°, were downscaled to a 100 m spatial-scale digital elevation model using two sets of temperature lapse rates, with and without bias-correction applied prior to downscaling. The bias-correction method adjusted the spatial patterns of temperature and precipitation in the 1/16° gridded data to match 30 arcsecond Parameter Regressions on Independent Slopes Model (PRISM) climate data. During production, the 1/16° temperature data was adjusted for the effect of elevation using a spatially uniform and temporally constant 6.5°C km⁻¹ lapse rate, whereas PRISM adjusts temperature for the effect of elevation using spatially and temporally variable lapse rates that are computed directly from regional weather station data. Thus, bias-correction implicitly adjusted the lapse rates in the 1/16° gridded data to match measured lapse rates. To test the effect of this implicit adjustment of the input data lapse rates vs. adjusting the lapse rates during downscaling, the 30 arcsecond bias-corrected data and 1/16° original data were each downscaled with 1) a spatially uniform and temporally constant 6.5°C km⁻¹ lapse rate, and 2) with monthly varying lapse rates computed from PRISM. Precipitation was

downscaled with the same method for each case. This procedure produced four sets of 100 m spatial scale data used as input to the snow model. Model parameters that control empirical estimates of incoming irradiance and the partitioning of precipitation into rain and snow were estimated independently with each dataset to optimize the agreement between modeled and observed SWE. We then modeled the sensitivity (percent change) of basin SWE in response to +2°C and +4°C warming with each of the four downscaled datasets and their respective optimized parameters. Pre-calibration, modeled SWE for the historical climate period differed depending on bias correction and choice of downscaling lapse rates. Post-calibration, modeled SWE for the historical climate period did not differ depending on choice of lapse rates but substantial differences emerged between modeled SWE with the original and bias-corrected forcing data. Inter-forcing dataset differences in modeled SWE during the historical period were largely controlled by differences in estimates of longwave irradiance and temperature between datasets. For the warming scenarios, the SWE sensitivity differed significantly at all elevations between the bias-corrected and original data, but (as in the post-calibration historical period) did not depend on choice of lapse rates. At low to mid elevations, climate change impacts on snow were largely controlled by temperature-driven shifts from snowfall to rainfall, while at high elevations, precipitation variability controlled SWE sensitivity. With just a 2°C increase in temperature, peak snow accumulation occurred 20-30 days earlier and was 20-60% smaller, the length of the snow covered season decreased up to 50 days, and winter rainfall increased by 20-60%. With a 4°C increase, the shifts in timing were roughly doubled and the declines in snow and snowfall increased up to 80%. A 10% increase in precipitation had a negligible impact on basin-integrated declines, indicating that future precipitation variability has little chance of offsetting regional climate warming impacts on snow in the Oregon Cascades. These results highlight the challenges of modeling SWE in data sparse regions, the importance of bias correcting gridded meteorological forcing datasets for hydrologic modeling applications in regions of complex topography, and the strong temperature dependence of snow in the Oregon Cascades.

©Copyright by Matthew Guy Cooper
May 1, 2015
All Rights Reserved

Modeling Elevation-Dependent Snow Sensitivity to Climate Warming in the Data Sparse
Eastern Oregon Cascades

by
Matthew Guy Cooper

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented May 1, 2015
Commencement June 2015

Master of Science thesis of Matthew Guy Cooper presented on May 1, 2015

APPROVED:

Major Professor, representing Water Resources Science

Director of the Water Resources Graduate Program

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Matthew Guy Cooper, Author

ACKNOWLEDGEMENTS

I express sincere appreciation to my adviser, Professor Anne Nolin, who has provided constant support and inspiration to me since I arrived at Oregon State University. Anne has provided me with every resource necessary for success, has always encouraged me to think big and go after important questions, but has also watched as I delved deeper and deeper into the depths of my project without stopping me from exploring my interests. This combination of support, encouragement, and freedom has made me a very happy graduate student, and I will forever be thankful. Anne presides over a group of my fellow graduate students in the Mountain Hydroclimatology Research Group, to whom all I owe a sincere thank you for the feedback, encouragement, collaboration, and shared learning.

I'd like to thank my committee members for their guidance. In particular, Dr. Mohammad Safeeq has been a critical source of methodological guidance throughout this project. Small suggestions from Safeeq often turned into large questions in my mind that shaped my thinking and approach to my project.

This thesis would not be possible without the hard work of Dr. Glen E. Liston, Senior Research Scientist at the Cooperative Institute for Research in the Atmosphere, who authored the numerical snow model I used for my research. I am grateful to Glen, and to all of the other great scientists in the field of snow hydrology that cut fresh tracks and post-holed the way before me.

Finally, I wish to dedicate this thesis to the memory of my mother, Karen Honnell Cooper, who taught me that service is the highest calling in life, the pursuit of knowledge the second, and going to parties the third.

TABLE OF CONTENTS

1. Introduction.....	1
1.1. Significance and motivation	1
1.2. Objectives and roadmap.....	4
1.3. Literature review	5
2. Study region description	10
3. Methods.....	12
3.1. Approach.....	12
3.2. Model forcing data and bias correction	13
3.3. Regional lapse rate computation and downscaling.....	16
3.4. Model description	18
3.5. Parameter estimation and model calibration.....	20
3.6. Evaluating modeled SWE during the reference period.....	22
3.7. Scenario development.....	23
4. Results and discussion	24
4.1. Impact of bias correction on estimated temperature and precipitation	24
4.2. Regional lapse rate variability and downscaled temperature and precipitation.....	28
4.3. Baseline SnowModel runs with original and bias-corrected data.....	35

TABLE OF CONTENTS (Continued)

4.4. Model parameter estimation	40
4.5. Model calibration, validation, and evaluation.....	46
4.6. Model evaluation and reference period simulations across elevation gradients	50
4.7. Climate change impacts on snow in the study region.....	57
4.7.1. Interdataset comparison of climate change impacts	57
4.7.2. A best estimate of climate change impacts to snow in the study region.....	61
5. Summary	68
5.1. Summary of research findings	68
5.2. Summary of research implications.....	71
6. Conclusion	74
Bibliography	75
Appendices.....	84

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
Figure 2.1: Map of the study region showing the locations of meteorological station networks, watershed boundaries for the Upper Deschutes and McKenzie River Basins, and the model domains for each basin.	1
Figure 3.1: Conceptual model of the experimental approach. Three scales of topographic resolution are shown: the 1/16° (~6 km) L13 digital elevation model, the 30 arcsecond (~800 m) PRISM digital elevation model, and the 100 m digital elevation model used for the snow model. The 6 km gridded climate data is first bias-corrected with 800 m PRISM data, then downscaled to the high resolution 100 m DEM. Bias correction improves the input data and parameter estimation provides an estimate of the model sensitivity to physical processes that lack direct observations in the study region.....	13
Figure 4.1: The mean-error (bias) between the L13 climate data and PRISM (solid black lines) and between L13 and station measurements (x's, linear fit shown as dashed line) for minimum temperature (top row), maximum temperature (middle row), and precipitation (bottom row), and for each season, plotted vs. elevation. Temperature and precipitation were averaged across all days in the respective season during the study period, and gridded data were binned and averaged across 100 m elevation bands.....	25
Figure 4.2: Monthly minimum and maximum temperature and precipitation from the L13 and PRISM climate data compared to SNOTEL station measurements for December, January, February, and March grouped. The L13 data is shown in purple, PRISM in orange. Results for other seasons were similar but the largest error was during winter.....	27
Figure 4.3: The distribution of daily minimum air temperature for the L13 and BC data compared to station measurements of minimum air temperature during winter.	28
Figure 4.4: Top: Minimum, maximum, and average daily temperature lapse rates computed from the station data in the study region; bottom: correlation coefficients for daily temperature vs. elevation.	29
Figure 4.5: Distributions of minimum, maximum, and average monthly temperature lapse rates computed from the bias-corrected temperature data and station temperature data in the study region.	30
Figure 4.6: Downscaled surfaces of mean daily air temperature for the December – March period for a) L13 data downscaled with a constant 6.5°C km ⁻¹ lapse rate, b) L13 data downscaled with PRISM lapse rates c) BC data downscaled with a constant 6.5°C km ⁻¹ lapse rate, and d) BC data downscaled with PRISM lapse rates	32
Figure 4.7: Downscaled mean daily air temperature compared to measurements from each meteorological station in the study region. Sisters COOP station was the only station used as input to the L13 interpolation algorithm; all stations were used as input to PRISM.	33

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure 4.8: Downscaled mean daily precipitation compared to measurements from each meteorological station in the study region. Sisters COOP station was the only station used as input to the L13 interpolation algorithm; all stations were used as input to PRISM.	34
Figure 4.9: Baseline simulations of SWE with each dataset. Timeseries of measured SWE from Three Creeks, Hogg Pass, and McKenzie SNOTEL, and corresponding modeled timeseries, are averaged to create one representative timeseries.	36
Figure 4.10: Mean values of simulated meteorological variables vs. elevation with the uncalibrated model. Hydrologic fluxes are cumulative sums on April 1; irradiance, temperature, and humidity are daily averages for the Dec. 1 – Apr. 1 period, binned and averaged into 100 m elevation bands across the 1989-2011 period.	37
Figure 4.11: The function used by Snow Model to estimate cloud fraction from 700 mb relative humidity.	39
Figure 4.12: The distribution of average daily air temperature ($^{\circ}\text{C}$) on days where snow accumulated in the study region from SNOTEL records. The expected value is very close to 0°C with the majority of precipitation events between 0 and -2°C	41
Figure 4.13: Cumulative posterior and uniform prior distributions for each calibration parameter. The TS, AL-MELT, and CF parameters show strong convergence to expected values, demonstrating the strong control on model performance that these parameters exert.	43
Figure 4.14: Calibrated simulations of SWE with each dataset. Timeseries of measured SWE from Three Creeks, Hogg Pass, and McKenzie SNOTEL, and corresponding modeled timeseries, were averaged to create one representative timeseries.	47
Figure 4.15: Each year in the study period classified by mean winter air temperature and April 1 cumulative precipitation at Hogg Pass and Three Creeks SNOTEL. Vertical red and horizontal blue bars are the interquartile range for each distribution.	49
Figure 4.16: Mean daily-simulated SWE compared to measurements at each of the three automated SNOTEL sites (top row) for the entire 23-year reference period. Manual SWE measurements from Snowcourses and field measurement sites bottom row.	51
Figure 4.17: Interannual model performance as evaluated with the traditional NSE, RMSE, and ME metrics of model evaluation. The bias-corrected data is more consistent between sites, with the lowest RMSE and most neutral ME.	54

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure 4.18: Mean values of simulated meteorological variables vs. elevation with the calibrated model. Hydrologic fluxes are cumulative sums on April 1; irradiance, temperature, and humidity are daily averages for the Dec. 1 – Apr. 1 period, binned and averaged into 100 m elevation bands across the 1989-2011 period.....	55
Figure 4.19: Maps of peak SWE with each of the four downscaled forcing datasets. From left to right: L13-6.5, L13-PLR, BC-6.5, and BC-PLR. Peak SWE is the date on which basin-wide volumetric SWE is greatest. Shown here are maps of the mean peak SWE for the 1989-2009 period.	56
Figure 4.20: Climate change impacts on mean SWE in the study region for the +2°C (top row), +2°C with +10% precipitation (second from top row), +4°C (third from top), and +4°C with +10% precipitation (bottom row) scenarios, simulated with each forcing dataset.	57
Figure 4.21: The percent change of mean SWE, SWE-melt, snowfall, and incoming longwave irradiance with the L13-PLR and BC-PLR datasets, for the +2°C and +4°C warming scenarios.	60
Figure 4.22: Climate change impacts on mean SWE in the study region, estimated with the bias-corrected data. Shown are the percent change in peak SWE and rainfall, and the number of day's difference for the date of peak SWE and the length of the snow-covered season.	62
Figure 4.23: April 1 SWE vs. elevation (upper left), mean winter air temperature vs. elevation (upper right), the percent change in April 1 SWE vs. elevation (lower left), and the percent change in April 1 SWE vs. mean winter air temperature (lower right).	63
Figure 4.24: Simulated daily SWE at each SNOTEL location. Top left: Three Creeks SNOTEL; top right: Hogg Pass SNOTEL; bottom left: McKenzie SNOTEL; and bottom right: basin-integrated.....	66
Figure 4.25: Absolute change in peak SWE in the Upper Deschutes River Basin with a 2°C warming.	67
Figure 5.1: Mean measured precipitation (solid blue) and SWE (solid black) during the period 1989-2011, and for water year 2014 (top, dashed) and 2015 (bottom, dashed), compared to mean modeled SWE for a 2°C (top, dashed red) and 4°C (bottom, dashed red) warmer climate.....	72

LIST OF TABLES

<u>Table</u>	<u>Page</u>
Table 2.1. Observational network stations and associated data used for this analysis.	12
Table 3.1. The snow-accumulation period climate, and descriptions for the original, bias-corrected, and downscaled meteorological surface datasets used for model forcing and intercomparison.....	17
Table 4.1 Model parameter values used for the uncalibrated runs, the minimum and maximum values of the uniform prior distributions, and the optimal estimates used for the final simulations.....	46
Table 4.2 Modeled SWE evaluation for the uncalibrated/calibrated model during the calibration (2000 – 2011) and validation (1989 – 1999) period for each dataset.	50
Table 4.3. Objective evaluation of each dataset at each SNOTEL location. The non-traditional metrics of performance correspond to the metrics used for the climate change sensitivity analysis.....	52
Table 4.4 Estimated basin-mean climate change impacts on SWE in the study region	64

LIST OF APPENDICES

<u>Appendix</u>	<u>Page</u>
Appendix A: Simulated SWE on April 1.....	85
Appendix B: Simulated SWE on March 28, the date of peak basin-integrated SWE.	86
Appendix C: Simulated SWE on April 1 with a 2°C warming.....	87
Appendix D: Simulated SWE on March 14, the date of peak basin-integrated SWE with a 2°C warming.....	88
Appendix E: Simulated SWE on April 1 with a 2°C warming and +10% precipitation.	89
Appendix F: Simulated SWE on March 14, the date of peak basin-integrated SWE with a 2oC warming and +10% precipitation.....	90
Appendix G: Simulated SWE on April 1 with a 4°C warming.....	91
Appendix H: Simulated SWE on February 20, the date of peak basin-integrated SWE with a 4°C warming.....	92
Appendix I: Simulated SWE on April 1 with a 4°C warming and +10% precipitation.	93
Appendix J: Simulated SWE on February 20, the date of peak basin-integrated SWE with a 4°C warming and +10% precipitation.	94
Appendix K: Simulated absolute change in SWE on April 1 with a 2°C warming.....	95
Appendix L: Simulated percent change in SWE on April 1 with a 2°C warming.....	96
Appendix M: Simulated absolute change in SWE on April 1 with a 2°C warming and +10% precipitation.	97
Appendix N: Simulated percent change in SWE on April 1 with a 2°C warming and +10% precipitation.	98
Appendix O: Simulated absolute change in SWE on April 1 with a 4°C warming.....	99
Appendix P: Simulated percent change in SWE on April 1 with a 4°C warming.....	100
Appendix Q: Simulated absolute change in SWE on April 1 with a 4°C warming and +10% precipitation.	101
Appendix R: Simulated percent change in SWE on April 1 with a 4°C warming and +10% precipitation.	102
Appendix S: Simulated absolute change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +2oC scenario.	103
Appendix T: Simulated percent change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +2°C scenario.	104
Appendix U: Simulated absolute change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +2°C with +10% precipitation scenario.	105

LIST OF APPENDICES (Continued)

<u>Appendix</u>	<u>Page</u>
Appendix V: Simulated percent change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +2°C with +10% precipitation scenario.	106
Appendix W: Simulated absolute change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +4°C scenario.	107
Appendix X: Simulated percent change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +4°C scenario.	108
Appendix Y: Simulated absolute change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +4°C with +10% precipitation scenario.	109
Appendix Z: Simulated percent change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +4°C with +10% precipitation scenario.	110

1. Introduction

1.1. Significance and motivation

In Mediterranean climates such as the Pacific Northwest (PNW) region of the United States, strong seasonality in precipitation leads to a reliance on winter snowfall to provide summer water supply for agriculture, municipalities, and ecosystems. This reliance on annual snow accumulation underscores the significance of the gradual decline of the mountain snowpack in the PNW during the last century and associated changes in the region's hydrologic regime. These changes include a 30% decline of average peak snow accumulation [Mote *et al.*, 2005], earlier snowmelt and spring runoff [Stewart *et al.*, 2005; Fritze *et al.*, 2011], shifts from snowfall to rainfall [Knowles *et al.* 2006], and decreased summer flows [Jefferson *et al.*, 2008; Luce *et al.*, 2013; Safeeq *et al.*, 2013a].

These changes are attributed to anthropogenic warming [Barnett *et al.*, 2008; Pierce *et al.*, 2008; Hidalgo *et al.*, 2009], and are not explained by decadal or interannual climate variability [Hamlet *et al.*, 2005; Mote, 2006]. Given this attribution and current trajectories, there will likely be continued shifts in the region's hydrologic regime toward less snowfall, more rainfall, and earlier snowmelt [Elsner *et al.*, 2010]. Snow in the Oregon Cascades is particularly vulnerable because most winter precipitation falls close to the rain-snow transition temperature [Nolin and Daly, 2006; Klos *et al.*, 2014] and because mid-winter melt is strongly correlated with peak snow-water-equivalent (SWE) [Mote *et al.*, 2005]. Regionally, warming temperature is driving these changes in SWE [Mote, 2003; Hamlet *et al.*, 2005; Mote *et al.*, 2005]. However, the temperature sensitivity of SWE depends on elevation due to the variation of temperature with altitude [Knowles and Cayan, 2004; Hamlet *et al.*, 2005; Howat and Tulaczyk, 2005; Mote, 2006; Sproles *et al.*, 2013].

These indicators of regional climate change have largely been observed in long-term datasets from weather stations or snowpack measurement locations that represent single points on the landscape [e.g. Mote, 2003; Mote *et al.*, 2005; Knowles *et al.* 2006]. When taken together, the trends provide a robust regional signal, but may not provide an estimate of the full variability of snow sensitivity to climate warming across complex topography. Estimating the spatial distribution of snow sensitivity to climate warming

across complex landscapes is important for several reasons including 1) the sensitivity is temperature-threshold dependent [Nolin and Daly, 2006] and patterns of surface temperature warming may depend on patterns of surface topography [Daly *et al.*, 2010; Pepin *et al.*, 2011]; 2) water management strategies benefit from sub-basin estimates of snow storage to improve predictions of climate change impacts on the magnitude and timing of snowmelt-driven streamflow [Knowles and Cayan, 2004]; and 3) snow regulates soil moisture and soil temperature which regulate biogeochemical cycles in snow covered ecosystems [Brooks and Williams, 1999], which are not adequately sampled by sparse point measurements [Randin *et al.*, 2014].

These needs motivate the use of spatially-distributed models to estimate snow sensitivity to climate warming. Because the physics of snow accumulation and ablation are relatively well understood there are abundant process-based numerical models available for such estimates [e.g. Liston and Elder, 2006a; Bavay *et al.*, 2009; Sproles *et al.*, 2013]. However, a key challenge for research in this area is the accurate representation of spatially-varying fields of temperature and precipitation used as model input, especially when incoming irradiance is estimated from the temperature and precipitation fields [Mizukami *et al.*, 2013; Wayand *et al.*, 2013; Elsner *et al.*, 2014]. This requirement is particularly challenging in data-sparse regions where ground-based monitoring networks lack sufficient spatial coverage to capture the true variability of climatic and hydrologic processes [Bales *et al.*, 2006; Nolin, 2012], and even more so in mountain landscapes where strong gradients in moisture and energy interact with topography and vegetation to create complex hydrologic patterns [Mahrt, 2006; Daly *et al.*, 2007].

In lieu of station data, many gridded model-forcing data products are available that facilitate fast and easy model implementation [e.g. Maurer *et al.*, 2002; Hamlet and Lettenmaier, 2005; Xia *et al.*, 2012; Abatzoglou, 2013; Livneh *et al.*, 2013]. These datasets provide quality-controlled model forcing data at sub-daily timescales and spatial resolutions of $1/16^\circ$ for any region in the continental US (many global datasets exist as well), making them particularly useful in data sparse regions. Generally, they are created by interpolating irregular observations of environmental variables such as temperature

and precipitation onto a regular grid using a horizontal interpolator and (in the case of temperature) combined with a spatially uniform and temporally constant lapse rate that adjusts for the assumed linear decrease of air temperature with increasing elevation. Downwelling shortwave and longwave irradiance and surface humidity are often derived from the temperature and precipitation fields via empirical relationships, for example Variable Infiltration Capacity (VIC) model-derived 2-m humidity and downwelling irradiance fields [Livneh *et al.*, 2013]. Thus, the climate-elevation relationship, which dictates temperature and precipitation regimes, is embedded in the production of these datasets [Mizukami *et al.*, 2013].

Gridded model-forcing datasets have been widely used as input to process-based hydrologic models to estimate climate warming impacts to regional hydrology [Christensen and Lettenmaier, 2007; Elsner *et al.*, 2014] but have generally been applied to very large geographical areas with coarse spatial resolution (>6 km grid-spacing). Most studies have focused on changes in regional mean streamflow that were, in fact, largely driven by changes in snow accumulation and ablation in headwater catchments [Elsner *et al.*, 2010]. Very few studies have explicitly focused on spatial estimation of snow sensitivity to climate warming in areas where spatial and temporal climate patterns are complex, nor have they used topographic data at a scale that accurately represents the complex topography of mountain environments.

For high spatial-resolution hydrologic modeling in such regions, specification of the surface temperature lapse rate is a common methodological choice. Lapse rates are used to distribute point-based input data or to downscale gridded input data to the topographic model boundary conditions. Whether using station or gridded data, lapse rates are used to extrapolate observed or gridded temperature to elevations where direct measurements do not exist. Errors in lapse rates can lead to large errors in estimates of the rain-snow transition line [Lundquist *et al.*, 2008b], the equilibrium line altitude of glacierized catchments [Immerzeel *et al.*, 2014], and can propagate bias into derived fields such as empirical estimates of incoming irradiance [Mizukami *et al.*, 2013; Lapo *et al.*, 2015]. It follows that estimates of elevation-dependent snow sensitivity to climate warming in

complex topography may depend on the choice of temperature lapse rates used in spatial temperature interpolation and extrapolation.

Thus, gridded meteorological data provide useful model input and diagnostic information to the hydrologic modeling community, yet we have an insufficient understanding of how simplifying assumptions used to generate the data, such as the choice to prescribe a constant temperature lapse rate and gridding resolution, impact model-based hydrologic assessments that utilize gridded data such as the effect of climate change on snowmelt-driven streamflow [Minder *et al.*, 2010; Elsner *et al.*, 2014]. The suitability of downscaling these datasets to high-resolution topography for watershed-scale hydrologic modeling where lapse rates are known to be highly variable in space and time is similarly uncertain.

1.2. Objectives and roadmap

The first objective of this study was to bias-correct and downscale a gridded temperature and precipitation dataset [Livneh *et al.*, 2013] onto a high resolution topographic dataset for use as input to a distributed snow accumulation and energy balance snowmelt model [Liston and Elder, 2006a]. The second objective of this study was to estimate the historical snowpack in the study region using the downscaled data as input to the snow model. The final, and overarching objective, was to estimate the sensitivity of the study region snowpack to increased temperature, with and without increased precipitation. Within this context, I sought to understand how the estimated sensitivity depended on elevation within the watershed. Further, I tested how the estimated historical snowpack, and the estimated sensitivity to climate warming, differed depending on bias-correction technique and the choice of lapse rates used to downscale the gridded climate data to the high-resolution topography.

In the following section, I summarize recent research examining gridded climate datasets used for hydrologic modeling, seasonal lapse rate variability in regions of complex topography, and implications of this variability for distributed hydrologic modeling. Section 2 describes the study region. Section 3 describes the methods I used to evaluate the gridded data and quantify seasonal lapse rate variability, the bias correction algorithm, model calibration and validation, and a description of the climate change

scenarios. Section 4 presents the bias-corrected and downscaled datasets, the model sensitivity to calibration parameters, the model performance with each downscaled dataset, and estimated sensitivity to climate warming with each downscaled dataset. Section 5 summarizes the contributions of this investigation to modeling efforts in data-sparse regions, implications of climate change for snowmelt-driven hydrology in this region, and a discussion of limitations and potential sources of error. Section 6 concludes the paper.

1.3. Literature review

Several recent studies have evaluated the sensitivity of hydrologic models to biases in gridded model-forcing datasets in regions of complex topography. These studies have consistently found elevation-dependent biases in gridded temperature and precipitation data that propagated into modeled estimates of incoming irradiance, snowmelt, and streamflow [Materia *et al.*, 2009; Mizukami *et al.*, 2013; Wayand *et al.*, 2013; Elsner *et al.*, 2014]. However, the magnitude and direction of interdataset differences, and errors relative to observations, differed in complex ways.

Output from the Weather Research and Forecasting (WRF) atmospheric model was compared to gridded forcing data empirically estimated from station observations in the Sierra Nevada [Wayand *et al.*, 2013], and each dataset was used to force the Distributed Hydrology and Soil Vegetation Model [Wigmosta *et al.*, 1994]. WRF-estimated shortwave and net-irradiance were consistently less biased on cloudy days, but performance was similar to empirical methods on clear-sky days. However, shortwave was empirically estimated (using methods of Thornton and Running, [1999]) from temperature that was extrapolated from a single low elevation station using a spatially uniform and temporally constant $6.5^{\circ}\text{C km}^{-1}$ temperature lapse rate. Temperature observations indicated that $6.5^{\circ}\text{C km}^{-1}$ was a good fit on an annual timescale, but results for winter or cloudy days were not presented. Despite large differences in the estimated shortwave irradiance, interdataset modeled-snowmelt differences were most sensitive to differences in estimated longwave irradiance, especially at low elevations. Errors relative to observations appeared to strongly depend on these differences, but the timing of the effect depended on elevation and forest cover. The empirical longwave estimates (using

methods of *Dilley and O'brien*, [1998]) produced melt rates smaller than observed melt rates at a high elevation site during spring, but the same forcing produced mid-winter melt at a low elevation site that was not observed. Further, interdataset differences in snowmelt were much smaller in forested than open sites due to the dampening effect of the canopy on atmospheric longwave differences. The majority of the basin was forested, thus interdataset differences in streamflow were strongly dependent on, and positively correlated with, interdataset differences in precipitation, despite the strong sensitivity of snowmelt to longwave irradiance.

In other cases, shortwave irradiance controlled interdataset snowmelt differences, and precipitation was negatively correlated with streamflow bias [*Mizukami et al.*, 2013]. In the upper Colorado River Basin, modeled SWE and peak runoff were both underestimated despite high bias in the precipitation and low bias in the temperature fields of the *Maurer et al.* [2002] climate data (hereafter referred to as M02) compared to the *Xia et al.*, [2012] climate data, when each were used to force the VIC model [*Mizukami et al.*, 2013]. The counterintuitive result was explained by the diurnal temperature range (DTR) difference between the two datasets that lead to overestimated incoming shortwave irradiance and hence snowmelt from the *Xia et al.*, [2012] data. Unlike the results from the Sierra [*Wayand et al.*, 2013], in Colorado the high elevation regions of the study region were more sensitive to the irradiance differences because the DTR bias increased with elevation. While not directly quantified, it was assumed that the temporally constant $6.5^{\circ}\text{C km}^{-1}$ temperature lapse rate used to produce the M02 data was responsible for the elevation-dependent DTR bias. Further, it is highly likely that differences in climate (warm Maritime vs. cold Continental) played an important role in the differential sensitivity of snowmelt and streamflow to sources of irradiance bias in the gridded climate data in each of these regions. Clearly, interactions between errors in precipitation, temperature, and irradiance estimates can propagate through the modeling chain in unexpected ways, can correlate with model boundary conditions such as elevation, and can differ depending on regional climate. The practical implication of these results is that the choice of meteorological forcing data, and the associated methods

used to produce that data, can be as important as intra-dataset variability for hydrologic model-based assessments [Elsner *et al.*, 2014].

Investigations of temperature lapse rate variability in regions of complex topography have consistently found strong seasonal cycles that deviate from $6.5^{\circ}\text{C km}^{-1}$, yet this value is often used in the gridding algorithm of datasets such as those cited above. Seasonally and spatially variable lapse rates have been observed in mountainous regions of Northern Italy [Rolland, 2003], the Washington Cascades [Minder *et al.*, 2010], southwestern Idaho [Blandford *et al.*, 2008], British Columbia [Stahl *et al.*, 2006], and the Nepalese Himalaya [Immerzeel *et al.*, 2014]. These observations suggest that average temperature lapse rates rarely or never exceed $6.5^{\circ}\text{C km}^{-1}$, are shallower (lesser decrease in temperature with altitude) than this value during winter (and the monsoon season in Nepal), and reach a maximum in late summer that tends to fluctuate closely around $6.5^{\circ}\text{C km}^{-1}$.

Minimum and maximum temperature lapse rates have distinct seasonal cycles that may or may not be in phase (e.g. out of phase in the PNW, in-phase in Italy and British Columbia) and are very rarely close in magnitude to each other or to $6.5^{\circ}\text{C km}^{-1}$. Maximum temperature lapse rates are steeper (greater decrease in temperature with altitude) than minimum temperature lapse rates with rare exception [e.g. Safeeq *et al.*, 2013b], have larger amplitudes of variability, and tend to exceed $6.5^{\circ}\text{C km}^{-1}$, except during winter when they are typically shallower than $6.5^{\circ}\text{C km}^{-1}$. Minimum temperature lapse rates tend to fluctuate between 0 and $4^{\circ}\text{C km}^{-1}$, rarely exceeding $6.5^{\circ}\text{C km}^{-1}$, but consistently have a weak linear relationship with elevation regardless of season [Mahrt, 2006; Stahl *et al.*, 2006; Daly *et al.*, 2007; Blandford *et al.*, 2008; Lundquist *et al.*, 2008a]. In general, the magnitude of seasonal variability is consistent among locations, but the absolute magnitudes of lapse rates vary regionally, for example in windward vs. leeward montane climates [Minder *et al.*, 2010] and at spatial scales larger than 1° [Rolland, 2003; Blandford *et al.*, 2008]. These observations clearly suggest that seasonal and spatial variability in lapse rate values should be addressed for watershed-scale hydrologic modeling applications.

Despite the lack of consistent seasonal or geographic mean values for minimum, maximum, or average temperature lapse rates, one clear result emerges: $6.5^{\circ}\text{C km}^{-1}$ is a poor estimate of the true surface temperature lapse rate during winter, especially if it is applied concurrently to minimum and maximum temperature. Further, wintertime lapse rates have the weakest linear relationship and are consistently shallower than $6.5^{\circ}\text{C km}^{-1}$. This suggests that where linearity is assumed, a constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate used to distribute temperature will lead to cold bias at high elevations during winter.

Despite this somewhat obvious conclusion, the effect of observed vs. assumed lapse rate variability on environmental modeling applications has received limited attention. Results from previous studies indicate large model sensitivity to choice of prescribed lapse rates when applied in mountainous regions dominated by steep topography and a rain/snow/ice transition. For example, prescribing observed lapse rates computed from local weather stations to distribute temperature in a glacierized catchment in the Nepalese Himalaya produced up to 400% more simulated runoff than with a constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate. The observed lapse rates were shallower than $6.5^{\circ}\text{C km}^{-1}$, thus the high elevation portions of the catchment were simulated warmer and snow accumulation decreased on the glacier, which increased icemelt and runoff regardless of season. Distributing temperature with the observed lapse rate values accurately simulated the equilibrium line altitude of the glacierized catchment [Immerzeel *et al.*, 2014].

In contrast, seasonally-dependent changes in modeled runoff were found for the seasonal-snow dominated Washington Cascades, where observed lapse rates varied from $3.9 - 5.2^{\circ}\text{C km}^{-1}$ depending on month. Modeled runoff was highly sensitive to the choice of a 6.5, 5, or $4^{\circ}\text{C km}^{-1}$ mean annual lapse rate but the relative difference depended on time of year [Minder *et al.*, 2010]. Model-simulated mid-winter and spring runoff was larger, snowmelt onset occurred one month earlier, and summer runoff was significantly reduced when the shallower $4^{\circ}\text{C km}^{-1}$ lapse rate was used to distribute temperature compared to a $6.5^{\circ}\text{C km}^{-1}$ lapse rate, due to the warmer estimated temperatures at high elevations. When the effect of prescribing observed and seasonally varying lapse rates for either side of the Cascades was compared, higher model sensitivity was observed on the leeward than windward side because lapse rate variability was much larger on the

leeward side of the Cascades. However, an intentionally simple modeling strategy was used, with no calibration of the SNOW-17 hydrologic model. This strategy isolated the impact of the lapse rate parameterization but it remains unclear 1) whether the model sensitivity to lapse rates would extend to applications where calibration is pursued; 2) which of the internal model processes were impacted by the lapse rate parameterization; and 3) how the lapse rate parameterization would impact model estimates of snowmelt sensitivity to climate warming with a more complex energy balance snowmelt model.

Finally, parameter sensitivity to choice of lapse rates was tested with the RHESSys eco-hydrologic model [Tague and Band, 2004] in the Western Oregon Cascades. Prescribing local station-derived lapse rates significantly reduced the number of parameter sets capable of simulating streamflow with high skill compared to a constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate. Interpolating temperature with the gridded 800 m PRISM product, which computes lapse rates directly from the station data, produced similar results. Further, simulated values of net primary productivity were up to four times lower and peak SWE up to four times higher when using the prescribed $6.5^{\circ}\text{C km}^{-1}$ vs. the 800 m PRISM product [Garcia et al., 2013]. In each of the aforementioned cases, observed lapse rates were shallower than $6.5^{\circ}\text{C km}^{-1}$ during winter, and simulated SWE was overestimated when $6.5^{\circ}\text{C km}^{-1}$ was used to extrapolate temperature across elevations. However, the relative impact of this difference depended on geographic location, time of year, the process being studied, and on the modeling strategy.

Given that 1/6 of the world's population relies on water resources derived from snow and ice-melt [Barnett et al., 2005], it is critical that we understand the sensitivity of hydrologic models to the temperature-elevation relationship, especially in a warming world. Here, we synthesize and build upon these lines of work [Minder et al., 2010; Garcia et al., 2013; Mizukami et al., 2013; Elsner et al., 2014; Immerzeel et al., 2014] by systematically evaluating the sensitivity of a distributed snowmelt energy balance model to the choice of lapse rates used to downscale gridded climate data to high resolution topography. To our knowledge no study has directly quantified regional lapse rate variability to diagnose errors in gridded temperature datasets and empirically estimated meteorological surfaces, demonstrated the method we use to correct these errors, nor

evaluated the sensitivity of a snowmelt energy balance model to choice of lapse rates used for downscaling. The effect of downscaling with observed regional-specific lapse rates vs. an assumed $6.5^{\circ}\text{C km}^{-1}$ lapse rate for estimating elevation-dependent snow sensitivity to climate warming has similarly not been presented.

2. Study region description

The Oregon Cascades have experienced the largest climate warming-driven declines in April 1 SWE in the United States over the last half century [Mote *et al.*, 2005] and modeling studies predict significant declines with a 2°C warming [Sproles *et al.*, 2013]. Our modeling domain is located in the central Oregon Cascades and covers the Metolius River and Shitike Creek watersheds, which comprise the headwaters of the Deschutes River Basin. The modeling domain spans from Mt. Jefferson southward to North Sister, the second and fourth tallest peaks in Oregon, respectively. We refer to these watersheds as the Upper Deschutes River Basin (UDRB) (Figure 2.1).

The UDRB is a 1500 km^2 watershed located east (lee) of the crest of the Cascades Mountains. Immediately west of the UDRB is the 3041 km^2 windward McKenzie River Basin (MRB), home to the H.J. Andrews Experimental Forest and Long Term Ecological Research Station. Elevation in the UDRB spans from 3200 m at the summit of Mt. Jefferson to 600 m in the lower reaches of the Metolius River canyon. Topography in the basin is characteristic of a volcanic plateau segmented by large volcanic peaks and deeply dissected river canyons, typical of the Oregon Cascades. Data from the Three Creeks Snow Telemetry (SNOTEL) station, characteristic of the climate on the east side of the Cascades, indicates annual average precipitation of $\sim 800\text{ mm}$ (compared to $\sim 2000\text{ mm}$ at McKenzie SNOTEL, a representative west side station), average wet day humidity of 73% (west side: 82%) and average winter wet day temperature of 4.4°C (west side: 3.8°C). Precipitation generally increases with elevation and temperature generally decreases, although inversions, cold air drainage, and continental air blockage are common, and the rain shadow effect complicates simple (e.g. linear) descriptions of the precipitation-elevation relationship [Daly *et al.*, 2008].

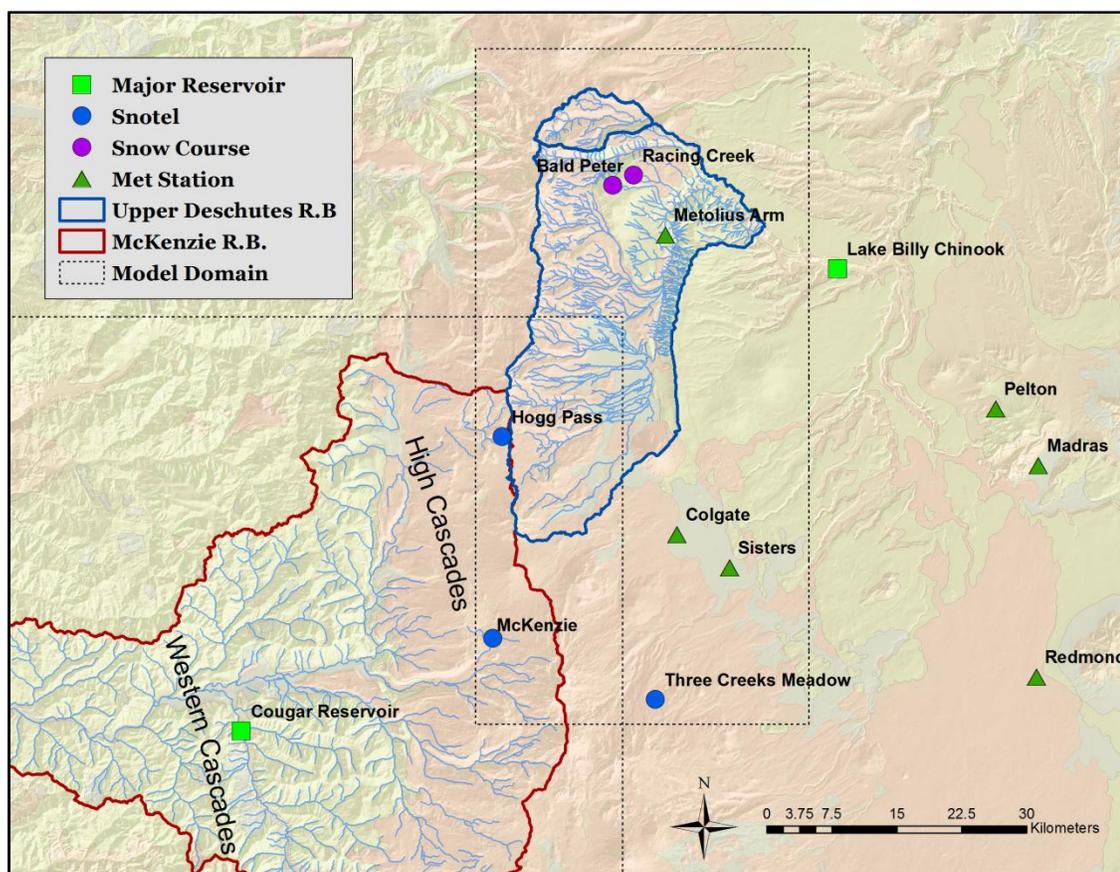


Figure 2.1: Map of the study region showing the locations of meteorological station networks, watershed boundaries for the Upper Deschutes and McKenzie River Basins, and the model domains for each basin.

Snow-water storage provides a critical reservoir for summer water supply here, as is common throughout the western US and Mediterranean climates worldwide. In the UDRB, snow provides water for hydropower, agriculture, municipalities, and anadromous fish. To facilitate management of water resources in the region, a variety of environmental observing networks are maintained by government agencies (Table 2.1). The Natural Resources Conservation Service SNOTEL and Snowcourses network monitors snowpack conditions at five locations within the modeling domain. Temperature, precipitation, and SWE are measured on an hourly schedule and quality-controlled data is available on a daily timestep. The U.S. Forest Service Remote Automated Weather Station (RAWS) network provides a suite of meteorological measurements to support land management and research applications, including two stations in the study region. The National Weather Service Cooperative Observer

Program (COOP) station network provides daily observations of maximum and minimum temperature, snowfall, and 24-hr precipitation totals at four locations in the study region.

These station networks are primarily intended for public land and resource management support and can be challenging to use as input to distributed environmental models due to poor geographic coverage [Sproles *et al.* 2013], data gaps, and inconsistent quality assurance methods, hence the proliferation of gridded model forcing data products. In this study, we use quality-controlled observations of temperature, precipitation, and SWE from these station networks to evaluate the gridded meteorological forcing data product we use as input to the snow model, to compute regional temperature lapse rates, and to evaluate model output (Table 2.1).

Table 2.1. Observational network stations and associated data used for this analysis.

Station Name	Elevation (m)	Network	Variables	Analysis
Three Creeks	1702	SNOTEL	T, P, SWE	Model evaluation, lapse rates
McKenzie	1504	SNOTEL	T, P, SWE	Model evaluation
Hogg Pass	1476	SNOTEL	T, P, SWE	Model evaluation, lapse rates
Bald Peter	1709	Snowcourse	SWE	Model evaluation
Racing Creek	1573	Snowcourse	SWE	Model evaluation
Metolius Arm	1050	RAWS	T, P	Model evaluation, lapse rates
Colgate	1000	RAWS	T, P	Model evaluation, lapse rates
Sisters	970	COOP	T, P	Model evaluation, lapse rates
Redmond	933	COOP	T, P	lapse rates
Madras	680	COOP	T, P	lapse rates
Pelton	430	COOP	T, P	lapse rates

3. Methods

3.1. Approach

The methodological approach is summarized as following: 1) bias-correct a gridded meteorological dataset; 2) downscale the original and bias-corrected data to a high resolution digital elevation model (DEM); in the case of temperature, adjust for elevation two ways - first using a constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate and second, lapse rates computed from PRISM climate data in the study region; 3) apply the downscaled data as input to a

process-based, spatially distributed snow accumulation and energy balance snowmelt model; 4) calibrate the model independently with each dataset; 5) compare interdataset differences in estimates of incoming irradiance, precipitation, temperature, snowfall, snowmelt, SWE, and humidity across elevation gradients pre-and post calibration, and 6) compare interdataset differences in modeled snow sensitivity to warming temperature with the original and bias-corrected data downscaled with each set of lapse rates. The experimental approach is presented conceptually in Figure 3.1.

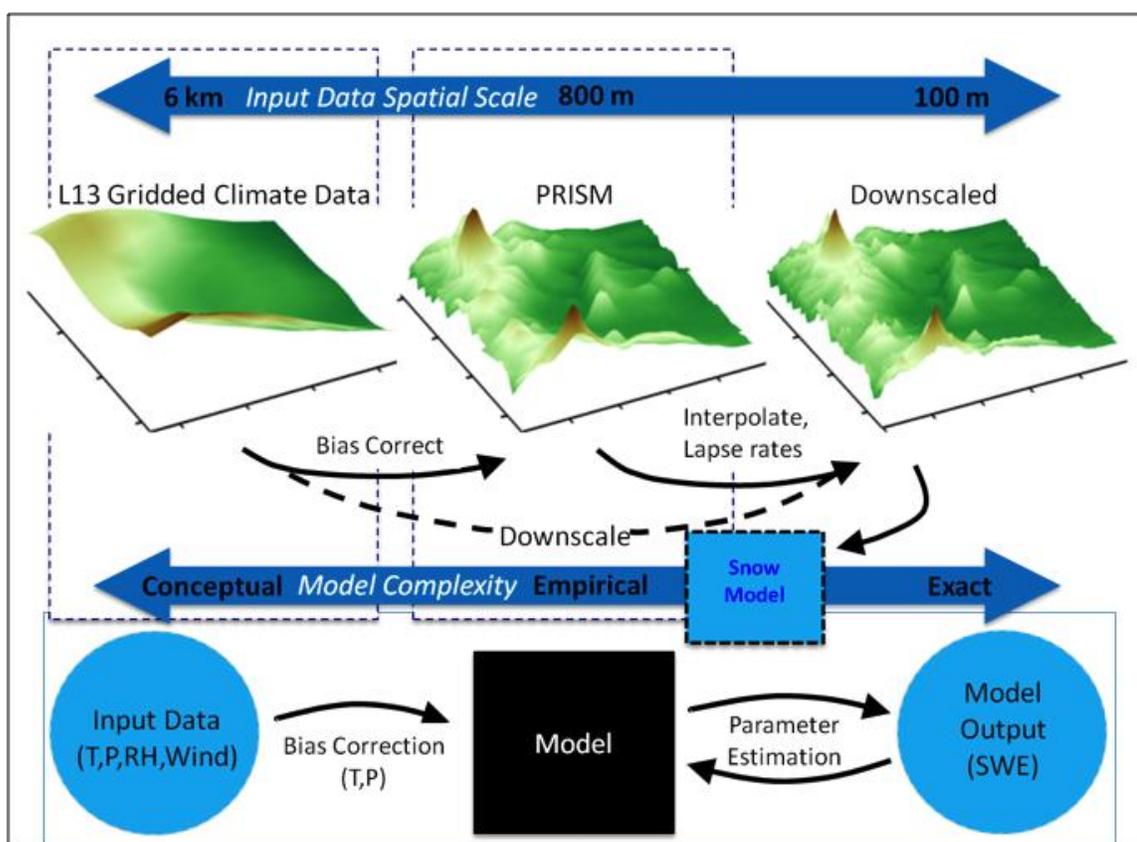


Figure 3.1: Conceptual model of the experimental approach. Three scales of topographic resolution are shown: the $1/16^\circ$ (~6 km) L13 digital elevation model, the 30 arcsecond (~800 m) PRISM digital elevation model, and the 100 m digital elevation model used for the snow model. The 6 km gridded climate data is first bias-corrected with 800 m PRISM data, then downscaled to the high resolution 100 m DEM. Bias correction improves the input data and parameter estimation provides an estimate of the model sensitivity to physical processes that lack direct observations in the study region.

3.2. Model forcing data and bias correction

In this study we used SnowModel [Liston and Elder, 2006a], a spatially-distributed snow accumulation and energy balance snowmelt model, to simulate the daily snowpack in the UDRB for the period 1989 – 2011. SnowModel includes MicroMet [Liston and

Elder, 2006b]), a micro-meteorological interpolation model that is used to interpolate the input forcing data to the elevation and land-cover boundary condition datasets. SnowModel is designed to use gridded input data for the meteorological forcings needed to solve the model equations, for example output from a regional atmospheric model, and uses MicroMet to interpolate the gridded data to the resolution of the boundary condition datasets. In basins where observational data is limited, a variant of this option is to use one of several gridded meteorological data products that have been interpolated from weather station observations. We used the $1/16^\circ$ (~6 km) temperature and precipitation forcing data described by *Livneh et al.* [2013], hereafter referred to as L13. We downloaded version 1.2 of the L13 data from the University of Washington Land Surface Hydrology Research Group ftp site <ftp://ftp.hydro.washington.edu/pub/blivneh/CONUS>. The L13 data is an updated version of the *Maurer et al.* [2002] data and was created by interpolating observations of daily minimum and maximum temperature and precipitation from the COOP network using the Synographic Mapping System [*Shepard*, 1968] on a 3-hourly timestep (aggregated to daily) at $1/16^\circ$ spatial resolution. In the L13 interpolation procedure, daily precipitation was scaled so that monthly mean precipitation matched the 1961 – 1990 PRISM monthly precipitation climatology and interpolated temperature was adjusted for elevation using a spatially uniform and temporally constant $6.5^\circ\text{C km}^{-1}$ lapse rate [*Livneh et al.* 2013]. A full description of the gridding algorithm is provided by *Widmann and Bretherton* [2000], and validation and description of the updated dataset by *Livneh et al.* [2013].

The snowpack model also required inputs of wind speed, wind direction, and relative humidity. We computed daily mean wind speed and direction from the zonal and meridional wind components of the U.S. National Centers for Environmental Prediction and National Center for Atmospheric Research reanalysis data product [*Kalnay et al.*, 1996]. Relative humidity was downloaded from the L13 data portal. The L13 data was used to force the VIC macro-scale hydrologic model [*Liang et al.*, 1994] and the modeled relative humidity was provided at the same timestep and spatial resolution as the L13 forcing data.

For bias-correction we used the PRISM LT71m monthly timeseries of temperature and precipitation provided at 30 arcsecond (~800 m) spatial resolution. PRISM is the official climate dataset for the U.S. Department of Agriculture, and is designed to incorporate state of the art knowledge of how physiographic variables affect the distribution of precipitation and temperature across complex topography. For example, the PRISM model includes parameters that differentiate windward and leeward mountain slopes, proximity to coastlines, the effect of local temperature inversions and cold air pools, and several measures of terrain complexity [Daly *et al.*, 2008]. Much of the validation of the PRISM model was done in the Oregon Cascades, where these processes exert strong controls on local climate [Daly *et al.*, 2007]. PRISM uses all available station observations, including SNOTEL, the highest elevation climate data available in the US, and an encoded knowledge base to create a “best” estimate of temperature and precipitation at any given location on a monthly timestep. While the temporal inconsistency of this approach may limit the use of PRISM to estimate long term trends [Guentchev *et al.*, 2010], it has been successfully used to create spatial datasets of input data for modeling experiments that test the sensitivity of a hydrologic system to perturbations or quantify the spatial distribution of hydrologic fluxes [Daly *et al.*, 2007, 2010; Abatzoglou, 2013; Klos *et al.*, 2014]. Most importantly for our application, PRISM is unique among climate gridding algorithms in its ability to temporally and spatially adjust the climate-elevation regression function (lapse rate) it uses to adjust temperature and precipitation for the effect of elevation [Daly *et al.*, 2008]. A full description of the PRISM model is provided at <http://prism.oregonstate.edu>, and the LT71m dataset by Daly *et al.* [2008].

Prior to bias-correction, the 6 km L13 temperature, precipitation, and humidity data were downscaled to the 800 m PRISM DEM with MicroMet using a constant $6.5^{\circ}\text{C km}^{-1}$ temperature lapse rate and the default precipitation scale factors of MicroMet. The downscaled daily L13 temperature was aggregated to monthly averages and precipitation to monthly totals for spatial and temporal consistency with the PRISM 800 m data. The wind fields were interpolated to the PRISM DEM using bilinear interpolation.

The bias-correction algorithm then superimposed the daily anomalies from the monthly average values of the L13 data onto the monthly averages from PRISM using the delta and ratio methods [Watanabe *et al.*, 2012]:

$$T_{d,BC} = T_{d,L13} + (T_{m,PRISM} - T_{m,L13})$$

$$PPT_{d,BC} = PPT_{d,L13} \times \frac{PPT_{m,PRISM}}{PPT_{m,L13}}$$

where the subscripts d and m denote daily and monthly, and BC is the bias-corrected L13 data. The method assumed that the daily temperature and precipitation bias was equivalent to the monthly bias, and was essentially equivalent to the method of Climatologically-Aided-Interpolation [Daly, 2006]. This same procedure was used to produce the Abatzoglou [2013] temperature and precipitation data except we used L13 as opposed to the National Land Data Assimilation Phase 2 data [Mitchell *et al.*, 2004]. As noted above, the same correction procedure was applied to the L13 precipitation data during production, but the 4 km 1961-1990 PRISM monthly precipitation climatology was used as a correction dataset. Temperature, however, was not scaled to match PRISM and instead the $6.5^{\circ}\text{C km}^{-1}$ lapse rate was the only post-interpolation correction made. Our approach was thus unique because we used 800 m PRISM monthly timeseries that corresponded to our modeling time period, as opposed to 4 km climatological mean values for 1961 - 1990, and we corrected temperature in addition to precipitation. Thus, for every month in the data period we computed a unique bias-correction scaling factor and applied that value to the L13 data for every day in the respective month. We used the ratio method for precipitation to avoid negative precipitation values and preserve dry grid cells and delta for temperature [Widmann and Bretherton, 2000]. The wind and humidity fields were not adjusted.

3.3. Regional lapse rate computation and downscaling

We used standard least-squares linear regression to compute daily and monthly temperature lapse rates using data from eight weather stations on the eastern side of the Cascade crest (Figure 2.1; Table 2.1). Minimum, maximum, and average temperature lapse rates were computed on daily and monthly timesteps as the slope of the least-squares linear regression equation with average daily or monthly temperature as the

response and elevation as the predictor variable. The mean monthly lapse rate was computed as the mean of the population of monthly lapse rates. We repeated this procedure with the PRISM data, restricted to the PRISM grid cells encompassed by our modeling domain. All lapse rate computations were made for the period 1989-2011.

The original and bias-corrected data were then each downscaled with MicroMet to the 100 m DEM that we used as the boundary grid for our model implementation. MicroMet is designed to use mean monthly lapse rates, and for each day in the model simulation, the daily lapse rate is computed by MicroMet using linear interpolation between monthly lapse rates centered on the mid-point of each month. The original and bias-corrected data were each downscaled twice, once with a constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate and once with the varying mean monthly temperature lapse rates computed from the regional PRISM data. The downscaling procedure thus produced four sets of 100 m spatial scale temperature data that was used as input to the model: the original L13 data downscaled with a constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate (L13-6.5) and with the PRISM lapse rates (L13-PLR), and the bias-corrected data downscaled each way (BC-6.5 and BC-PLR) (Table 3.1).

Table 3.1. The snow-accumulation period climate, and descriptions for the original, bias-corrected, and downscaled meteorological surface datasets used for model forcing and intercomparison.

Dataset Name	Lapse Rate Downscaling Method	Mean DJFM air temperature ($^{\circ}\text{C}$)	Mean Apr 1 cumulative precipitation (mm)	Spatial scale
L13	-	-1.41	470	$1/16^{\circ}$ (~6 km)
BC	-	0.25	595	800 m
L13-6.5	$6.5^{\circ}\text{C km}^{-1}$	-1.41	482	100 m
L13-PLR	PRISM	-1.40	482	100 m
BC-6.5	$6.5^{\circ}\text{C km}^{-1}$	0.28	592	100 m
BC-PLR	PRISM	0.26	592	100 m

The downscaled temperature and precipitation were evaluated against measurements of temperature and precipitation from weather stations in the study region. We applied a 3-day moving average to the weather station records of temperature and precipitation to account for the possibility of inconsistent reporting times between station networks (e.g. SNOTEL, COOP, and RAWS networks, *Chris Daly, personal communication*), and to corresponding gridded data. We computed the Nash-Sutcliffe efficiency (NSE), root-

mean-squared-error (RMSE), and mean error (ME) between the 3-day moving averaged observed and downscaled temperature and precipitation at each of the six stations in the modeling domain (Table 2.1) for each water year in the 1989-2011 period, and for the entire period.

Quality control (QC) procedures differed between station networks so a robust QC procedure was employed. We used out of range limits (LIM), rate of change limits (ROC), and no-change limits (NOC) to detect possibly erroneous values of temperature and precipitation [Meek and Hatfield, 1994]. For temperature, LIM thresholds were set to the historical observed minimum (-48°C) and maximum (49°C) values for the state of Oregon, and the maximum ROC threshold was set to 25°C . For precipitation the minimum LIM threshold was set to 0 mm and maximum LIM and ROC thresholds to 250 mm [Serreze *et al.*, 1999]. The temperature NOC threshold was set to 2 days, effectively excluding any consecutive constant values. Consecutive values for precipitation were not flagged [Serreze *et al.*, 1999]. Following these checks, daily temperature values exceeding 5 standard deviations from the mean monthly value were flagged. This procedure was repeated 2 additional times to account for the sensitivity of the mean and standard deviation to outliers. Daily values that exceeded the minimum and maximum LIM, ROC, and NOC thresholds or failed the outlier test were excluded from analysis, along with corresponding gridded and modeled data. Combined with previously missing data these checks excluded 7.2%, 8.5%, 9.7%, and 3.7% of all data for average, maximum, and minimum temperature, and precipitation, respectively. The L13 and PRISM data and daily SWE data from the SNOTEL network were provided in QC'd format and required no filtering. Where monthly values were calculated, we required that 90% of the data be complete for any station to enter the equation. All computations and evaluations were performed with the QC'd data.

3.4. Model description

SnowModel is a spatially-distributed snow accumulation and energy balance snowmelt model that uses a hierarchical modeling architecture to simulate the time evolution of snow in any region where snow occurs [Liston and Elder, 2006a]. SnowModel is an aggregate of four sub-models: MicroMet, EnBal, SnowPack, and SnowTran-3D.

MicroMet estimates values of temperature, precipitation, wind speed, wind direction, relative humidity, incoming shortwave and longwave irradiance, and surface pressure in each grid cell of a DEM, for each timestep. These interpolated fields are then passed to EnBal and SnowPack to compute the energy and mass balance of the simulated snowpack. SnowTran-3D is used to simulate interactions between snow and wind. We chose not to implement SnowTran-3D for our simulations.

MicroMet spatially interpolates the input forcing data in a two-step process. First, point or gridded observations of the first five variables, and empirical estimates of the latter three variables, are horizontally interpolated using a Barnes objective analysis scheme [Barnes, 1964; Koch *et al.*, 1983]. Physical sub-models based primarily on relationships with location, terrain, and vegetation are then applied to each variable to improve their realism at each point in time and space [Liston and Elder, 2006b]. In the case of temperature, a linear reduction in temperature with elevation is used to adjust the horizontally interpolated temperature surfaces to the elevation boundary conditions provided by the DEM. The model user specifies mean monthly temperature lapse rates and a linear interpolation centered on the midpoint of each month defines the daily temperature lapse rate. Precipitation is adjusted for elevation following the methods of Thornton *et al.*, [1997], with monthly-varying non-linear scaling factors that are interpolated to daily in the same way as temperature. Total precipitation is partitioned into solid and liquid components with a function that was modified to include mixed-phase precipitation for modeling in warm winter climates where large proportions of total winter precipitation accumulates when air temperature is near 0°C [Sproles *et al.* 2013]:

$$snowfall(t) = \begin{cases} ppt(t), & t_{air}(t) \leq TS \\ 0, & t_{air}(t) \geq TR \\ \frac{1}{TR - TS} \times (TR - t_{air}(t)) \times ppt(t) & TS < t_{air} < TR \end{cases}$$

where $snowfall(t)$ is the fraction of precipitation $ppt(t)$ defined as snow at time t , t_{air} the average air temperature, TS the temperature below which precipitation is defined to be all snow, and TR the temperature above which precipitation is defined to be all rain.

The slope of the linear equation that connects the lower (TS) and upper (TR) temperature thresholds scales the fraction of precipitation defined as snow between these thresholds.

The gridded solid and liquid precipitation, humidity, pressure, and incoming irradiance surfaces provide the input data to EnBal and SnowPack, which compute the mass and energy balance of the snowpack in each grid cell for every timestep. The energy balance is used to compute the energy available for ablation, while internal snowpack processes, forest-snow interactions, snow sublimation, and wind-snow interactions are computed by SnowPack and SnowTran-3d. A full description of each sub-model, governing equations, approximations, and assumptions can be found in *Liston*, [1995]; *Liston and Sturm*, [1998]; *Liston and Elder*, [2006a, 2006b]; *Liston et al.*, [2007]; and *Liston and Hiemstra*, [2008].

SnowModel has been successfully applied across a wide range of climates where snow occurs [e.g. *Strack et al.*, 2004; *Liston and Elder*, 2006a; *Mernild and Liston*, 2010; *Sproles et al.*, 2013; *Randin et al.*, 2014] and is well suited for this study as it is designed to work with irregularly spaced observations of model forcing data (e.g. from a network of meteorological stations), or with a gridded data product, and uses monthly lapse rates to interpolate the temperature data. As is common with many hydrologic modeling applications, we ran the model at a higher spatial resolution than the input forcing data, because the processes that control snow accumulation and ablation vary on scales of tens to hundreds of meters [*Winstral et al.*, 2014].

3.5. Parameter estimation and model calibration

Optimal model parameters were estimated independently with each of the four forcing datasets using Latin Hypercube Sampling (LHS) [*Iman*, 2008] from uniform prior distributions (N=1000) with NSE as a cost function. We used model-developer advice [*Glen Liston, personal communication*] and guidelines in *Sproles et al.*, [2013] to identify six parameters that control the primary model processes that control snow accumulation and ablation. The snow-rain transition temperatures (TS, TR) define the phase of precipitation in the model (as described in section 3.4) and are essentially empirical indexes that relate the 2-m air temperature to on-the-ground precipitation phase, and are highly uncertain [*Dai*, 2008]. Net shortwave irradiance is the primary determinant of

snowmelt in open-canopy clear-sky conditions and longwave in forested or cloudy conditions [Marks and Dozier, 1979; Male and Granger, 1981; Marks et al., 1992]. We selected four model parameters that exert strong control on these radiative fluxes. The gap-fraction (GF) varies from 0-1 and scales the incoming shortwave irradiance below the forest canopy, accounting for gaps in the canopy that let through direct solar-beam irradiance. A GF value of 0 will treat all below-canopy shortwave irradiance as the canopy-modified (diffused) top-of-canopy direct plus diffuse shortwave irradiance; a GF of 1 will treat the below-canopy irradiance as equivalent to the top-of-canopy direct plus diffuse shortwave irradiance. The cloud factor (CF) scales the model-computed increase in longwave irradiance and diffuse shortwave irradiance due to the presence of clouds. The model estimates cloud-cover and cloudy-sky incoming irradiance following the methods of *Iziomon et al.*, [2003] and *Walcek*, [1994]. There is considerable uncertainty in this estimation [Walcek, 1994; Liston and Elder, 2006b]. Finally, the albedo decay-rate parameters GRMELT and GRCOLD determine the rate at which snow-surface albedo decreases following a new snow event for melting snow (defined as snow surface temperature $> 0^{\circ}\text{C}$) and frozen snow, respectively. Details of modifications to the albedo decay routine in SnowModel made for the Oregon Cascades can be found in *Sproles et al.* [2013].

The calibration routine was designed to identify an optimal combination of one value for each parameter that yielded the highest NSE score when averaged across all snow measurements in the study region. Given the highly variable nature of the processes and associated parameters in space and time, wide limits for uniform prior distributions were defined based on physical considerations and field observations (Table 4.1). The a priori limits for TS and TR assumed that all precipitation fell as snow on days with average air temperature less than -1.5°C , and all precipitation fell as rain on days with average air temperature greater than 2.5°C . The a priori limits for GF and CF assumed that forest gap fraction varied between 10 and 80% and cloud fraction varied between 20 and 100%. A priori distribution limits for the albedo decay parameters were not available from the published literature and do not have a direct natural analog, thus computationally reasonable values that bounded the default values were selected. One thousand unique

sets of parameter combinations were randomly sampled from the prior distributions using LHS [Iman, 2008]. We chose $N=1000$ as a computationally feasible sample size that adequately sampled the parameter space given the highly efficient LHS strategy. Default parameter values used for the baseline uncalibrated model runs were selected from *Sproles et. al.* [2013].

The 1989 – 2011 study period was split into a calibration (2000 – 2011) and validation (1989 – 1999) period, and the model was solved for each of the 1000 parameter sets during the calibration period. The NSE goodness-of-fit metric was computed for the 12-year calibration period excluding days when the observations and the model both recorded zero SWE, using records of measured SWE from each of the three SNOTEL stations in the study region. In addition, we grouped the manual measurements of SWE at Bald Peter and Racing Creek Snowcourses and the Hogg Pass field measurement site and computed the NSE between this sample population and the corresponding 3-day moving-averaged modeled timeseries centered on the day of each observation. Parameter sets were ranked based on the average of the NSE at each SNOTEL location and the grouped manual measurements. This procedure was repeated with each of the four forcing datasets, and the highest ranked parameter set was selected for the final set of model simulations with each dataset. These parameters were then used to run the model during the 11-year validation period, and the NSE was computed for this period to validate model performance outside of the calibration period.

The calibration routine was employed for two reasons. First, we evaluated the impact of inter-dataset temperature and precipitation differences on optimal parameter estimates and model skill, thus identifying the sensitivity of internal model processes to input dataset differences and corresponding impacts on model skill. Second, by optimizing the parameters with each forcing dataset we were able to evaluate the sensitivity of the modeled climate warming impacts based on the data given a calibrated model, and not an interaction between the data and non-optimized parameters.

3.6. Evaluating modeled SWE during the reference period

In addition to the calibration and validation period evaluation, calibrated model simulations of SWE were evaluated for the entire 23-year study period (herein referred to

as the “reference” period). The modeled SWE was evaluated against daily automated measurements of SWE recorded at the SNOTEL stations, as well as for the population of manual SWE measurements at the Snowcourse sites. In addition to the traditional measures of NSE, RMSE, and ME, we further evaluated the modeled SWE with several non-traditional measures of model performance. We computed the difference between observed and modeled date and amount of peak SWE, the date of snow disappearance, and the length of the snow-covered season. The amount of peak SWE was computed as a percentage difference and the timing metrics as absolute differences (i.e. number of days difference). The four non-traditional metrics were chosen based on the objective of modeling specific characteristics of the snowpack that corresponded to metrics of change evaluated for the climate warming experiments, while the traditional metrics gave an indication of average model performance during reference period.

3.7. Scenario development

To evaluate the sensitivity of SWE in our study region to changes in mean climate, we developed four simple climate change scenarios. We used the delta method, wherein the historical (reference) records of temperature that were used to force the model for the reference period were scaled by a change factor. We added +2°C and +4°C to every daily temperature record in each of the four downscaled temperature datasets, with and without +10% precipitation added to each precipitation record. Respectively, these are referred to as T2, T2P10, T4, and T4P10 scenarios. This approach does not take into account future changes in variability, nor seasonal or sub-daily differences in the rate of warming or changes in future precipitation [Salathe *et al.*, 2007; Abatzoglou *et al.*, 2013]. The model was then run with the scaled forcing data and each set of calibrated model parameters. The reference period simulations were used as a baseline for comparison with the results of the climate change scenarios. For each set of scenario simulations, we computed the difference in amount and date of peak SWE, the date of snow disappearance, the length of the snow-covered season, and the percent increase in rainfall. These were each computed for every grid cell in the model domain and for every day in the simulation period, and then averaged across all years and binned into 100 m elevation bands.

Further, at each of the SNOTEL measurement locations, we computed changes in the mean daily SWE, and we repeated this for the daily basin-averaged mean SWE.

4. Results and discussion

4.1. Impact of bias correction on estimated temperature and precipitation

We first present a direct comparison between the L13 climate data, PRISM climate data, and meteorological station data in the study region, to establish the primary differences between pre- and post-bias-corrected datasets and observations. When gridded data values were compared to weather station data, the L13 data systematically underestimated measurements of both temperature and precipitation in the study region, as well as gridded estimates from the PRISM model (Figure 4.1). The temperature bias was clearly correlated with elevation, varied seasonally, and differed for minimum and maximum temperature. Most importantly for our application, the bias was largest for both temperature and precipitation during winter.

Precipitation differences between datasets were not as clearly correlated with elevation as temperature, however they were substantial. The largest difference between L13 and PRISM and between L13 and station measurements occurred in the 1400 – 1800 m elevation band (Figure 4.1) where 63% of the modeled peak snowpack accumulated (results presented in section 4.6.2). This is also the elevation range where the seasonal snowpack was estimated as most sensitive to climate warming in the geographically adjacent McKenzie River Basin [*Sproles et al.*, 2013]. In this elevation range, monthly total precipitation was underestimated during the study period by up to 10 cm during winter on average.

The precipitation bias was difficult to diagnose. Precipitation is not well correlated with elevation at the scale of this analysis and is more difficult to interpolate across complex topography than temperature [*Daly et al.*, 2008]. However, the bias may be related to the inconsistency between the 1961 – 1990 time period of the PRISM precipitation climatology used by the L13 interpolation procedure and the 1989 – 2011 time period of this study. This hypothesis does not necessarily imply that climate has changed. The difference may also be due to changes in the PRISM interpolation algorithm and the station networks used as input to the PRISM model. Further, we used

the monthly timeseries as opposed to the climatological mean values, which may have captured anomalous precipitation events that were smoothed by the climatological mean values. Regardless of the precise explanation, our results suggest that a “double” bias correction is recommended for modeling studies that use gridded forcing datasets based on climatology outside the modeling period. We corrected the precipitation bias, but a comprehensive diagnosis was not the objective of this study. For the remainder of this paper, we focus on the temperature bias, its relationship to regional lapse rate variability, and implications of the temperature bias for our modeling objectives.

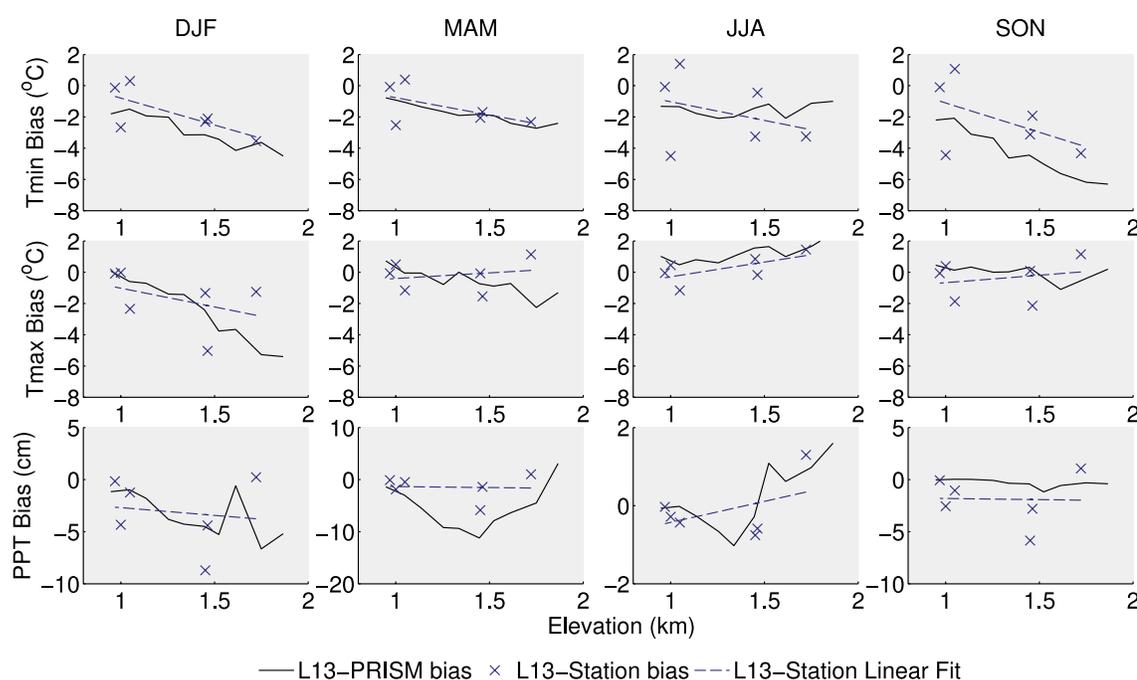


Figure 4.1: The mean-error (bias) between the L13 climate data and PRISM (solid black lines) and between L13 and station measurements (x's, linear fit shown as dashed line) for minimum temperature (top row), maximum temperature (middle row), and precipitation (bottom row), and for each season, plotted vs. elevation. Temperature and precipitation were averaged across all days in the respective season during the study period, and gridded data were binned and averaged across 100 m elevation bands.

During winter, the L13 temperature data underestimated measured air temperature at high elevation stations in the study region. The wintertime (DJF) daily average temperature bias averaged across the three SNOTEL stations in the study region was $-2.43^{\circ}\text{C d}^{-1}$ ($-1.36^{\circ}\text{C d}^{-1}$ for all stations grouped), which is similar to the $-3.0^{\circ}\text{C d}^{-1}$ bias between SNOTEL stations in the Colorado River Basin and the M02 data [Mizukami *et al.*, 2013]. A bias of this magnitude is particularly significant for model applications in

regions where wintertime precipitation occurs at temperatures close to 0°C, as it will lead to misidentification of precipitation phase. To demonstrate this, we estimated snowfall with the snow-rain partitioning function of SnowModel using the L13 daily temperature timeseries with the observed precipitation timeseries from the three SNOTEL stations in the study region. Estimated snowfall with the L13 temperature was greater by 22-46% on average than snowfall estimated with the observed temperature timeseries. In the context of a model this suggests that snowfall will be overestimated, but does not take into account propagation of the temperature bias into additional processes that depend on temperature, such as empirically estimated irradiance.

For both maximum and minimum temperature, the bias increased approximately linearly with elevation, suggesting a lapse rate bias. Minimum temperature was estimated colder than measurements during all seasons by the L13 data, whereas maximum temperature was estimated as warmer than measurements except during winter when it was also estimated as colder. The concurrent cold bias in both minimum and maximum temperature during winter lead to a corresponding cold bias in average temperature during winter, but during other seasons the opposite directions of the bias tended to compensate.

On average, L13 air temperature was between 0 and 4°C colder than both PRISM and station measurements (Figure 4.1), which is the same temperature range used in this study to estimate the sensitivity of the modeled snowpack to climate warming. The fact that the mean difference between the L13 air temperature and station measurements was as large or larger than the prescribed and projected near-term [Mote and Salathe, 2010] effect of climate change in this region suggests that this data may lead to biased estimates of climate warming impacts on snowmelt-driven hydrology. Further, the fact that the bias increased with elevation suggests that estimates of snow sensitivity to climate warming above the range of current monitoring networks may also be biased.

In general, PRISM data matched station measurements on daily and monthly timescales (Figures 4.1 and 4.2). When spatially averaged across elevation bands, PRISM was a close approximation to the station measurements (Figure 4.1). When compared directly to stations, PRISM provided a nearly equivalent estimate of monthly climate

conditions (Figure 4.2). However, these same stations are used as input to the PRISM model. This detail is discussed in the following section.

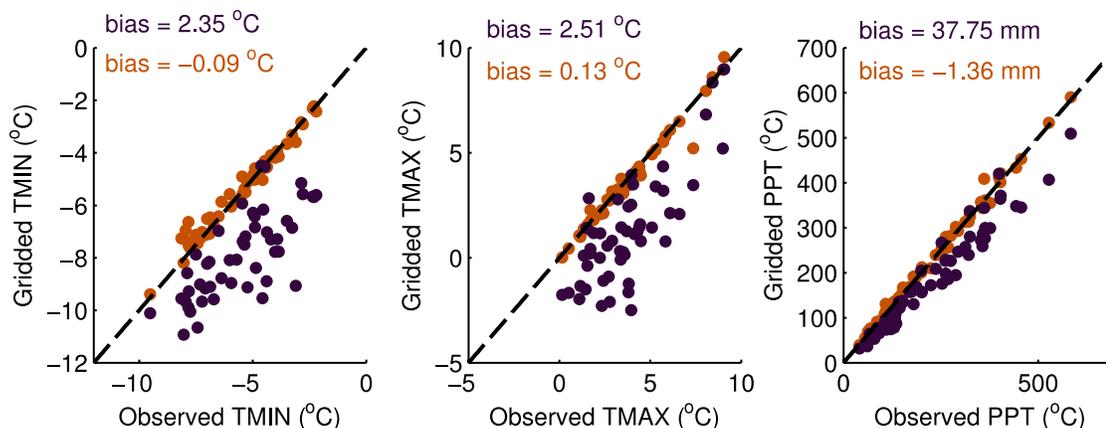


Figure 4.2: Monthly minimum and maximum temperature and precipitation from the L13 and PRISM climate data compared to SNOTEL station measurements for December, January, February, and March grouped. The L13 data is shown in purple, PRISM in orange. Results for other seasons were similar but the largest error was during winter.

Bias correction with PRISM effectively removed the systematic low bias in the L13 daily temperature data (Figure 4.3). The bias correction procedure we used scaled the daily data such that the mean of the L13 monthly temperature and precipitation distributions equaled the mean of the PRISM monthly distributions, which in turn were almost equivalent to the station distributions (Figure 4.2). However, bias-correction did not fully correct the daily distribution. Here, daily measurements from all stations were grouped for the entire study period and shown for winter when the bias was most significant, but results for individual stations and seasons were similar. When pre- and post-bias-corrected daily temperature and precipitation timeseries for the entire 23 year modeling period were compared directly to station measurements, the daily NSE, RMSE, and ME for all stations grouped was 0.58, $1.88^{\circ}\text{C d}^{-1}$, and $-1.36^{\circ}\text{C d}^{-1}$ for the uncorrected L13 data, and 0.83, $1.20^{\circ}\text{C d}^{-1}$, and $0.24^{\circ}\text{C d}^{-1}$ for the bias-corrected data. For precipitation, the daily NSE, RMSE, and ME for all stations grouped was 0.82, 2.48 mm d^{-1} , and -1.0 mm d^{-1} for the uncorrected data, and 0.89, 1.93 mm d^{-1} , and -0.13 mm d^{-1} for the bias-corrected data. Comparisons between downscaled pre- and post-bias-corrected data and station measurements are presented in the following section.

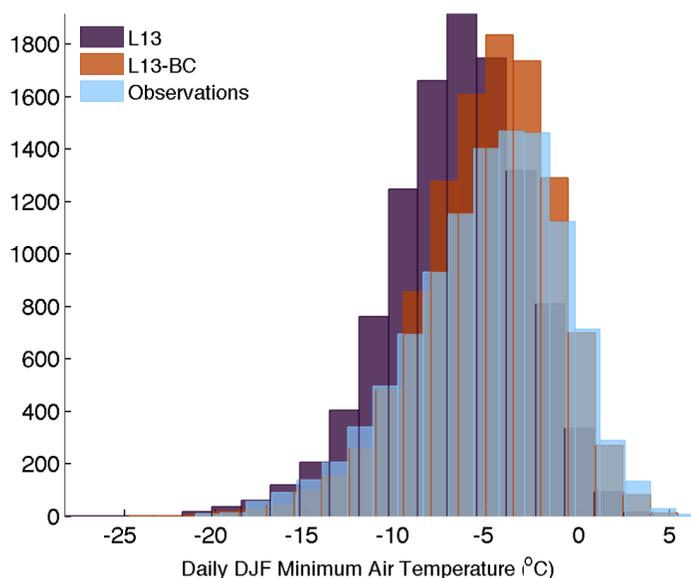


Figure 4.3: The distribution of daily minimum air temperature for the L13 and BC data compared to station measurements of minimum air temperature during winter.

4.2. Regional lapse rate variability and downscaled temperature and precipitation

Lapse rates computed from station observations clearly demonstrated that $6.5^{\circ}\text{C km}^{-1}$ was a poor estimate of surface temperature lapse rates for minimum and maximum temperature on daily (Figure 4.4) and monthly (Figure 4.5) timescales in this region, and largely explain the elevation-dependent temperature bias presented in the previous section. Average temperature lapse rates, however, were closely approximated by $6.5^{\circ}\text{C km}^{-1}$, except during winter when they were shallower than $6.5^{\circ}\text{C km}^{-1}$. All temperature lapse rates computed from stations followed a distinct seasonal cycle during the study period, reaching a minimum in winter and maximum in summer, but seasonal variability was larger for maximum and average temperature lapse rates. Minimum temperature lapse rates were shallower than $6.5^{\circ}\text{C km}^{-1}$ all year, most so during the wet season from September through January and least so in May and June. Maximum temperature lapse rates had the largest amplitude and approached the $9.8^{\circ}\text{C km}^{-1}$ dry adiabatic lapse rate on average much of the year, but were shallower than $6.5^{\circ}\text{C km}^{-1}$ in December and January. In general, the temperature-elevation relationship was not well described by a linear process for minimum, maximum, or average temperature during winter. However, a

linear process described average and maximum temperature lapse rates much better than minimum temperature (Figure 4.4).

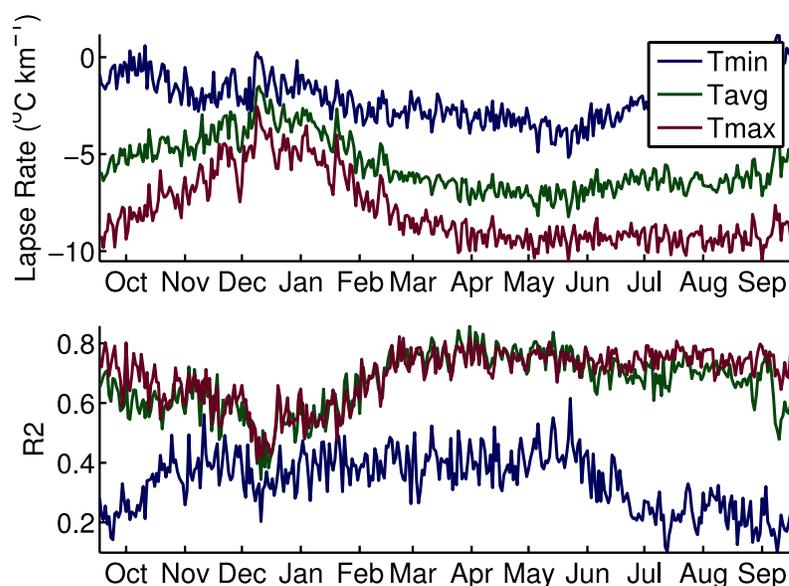


Figure 4.4: Top: Minimum, maximum, and average daily temperature lapse rates computed from the station data in the study region; bottom: correlation coefficients for daily temperature vs. elevation.

From a regional perspective, mean monthly average temperature lapse rates in the leeward UDRB basin were similar to those reported for the leeward Washington Cascades (max difference was $0.7^{\circ}\text{C km}^{-1}$ in February) [Minder *et al.*, 2010] and British Columbia (max difference was $1.1^{\circ}\text{C km}^{-1}$ in August) [Stahl *et al.*, 2006]. From March through August, $6.5^{\circ}\text{C km}^{-1}$ was a close approximation to observed monthly average temperature lapse rates in each of these regions, but from September through February lapse rates were shallower than $6.5^{\circ}\text{C km}^{-1}$, especially in December and January. While there was significant variability in the population of monthly lapse rates (Figure 4.5), mean monthly R^2 values for average temperature lapse rates ranged from 0.84 – 0.98.

This variability of lapse rates makes sense in light of the seasonality of moisture and airflow in this region. During winter, orographically-enhanced precipitation releases latent heat, warming the upper atmosphere and depressing the lapse rate. Cloud cover during winter suppresses daytime radiative cooling, weakening the reduction of temperature with elevation. Regionally, cold continental air from the east impinges on the Cascades depressing low elevation temperatures [Steenburgh *et al.*, 1997]. Inversions,

including near-surface cold air drainage and regional high pressure inversion caps, also complicate the linearity assumption [Daly *et al.*, 2010]. Each of these processes likely contributed to the wintertime depression of lapse rates in the study region.

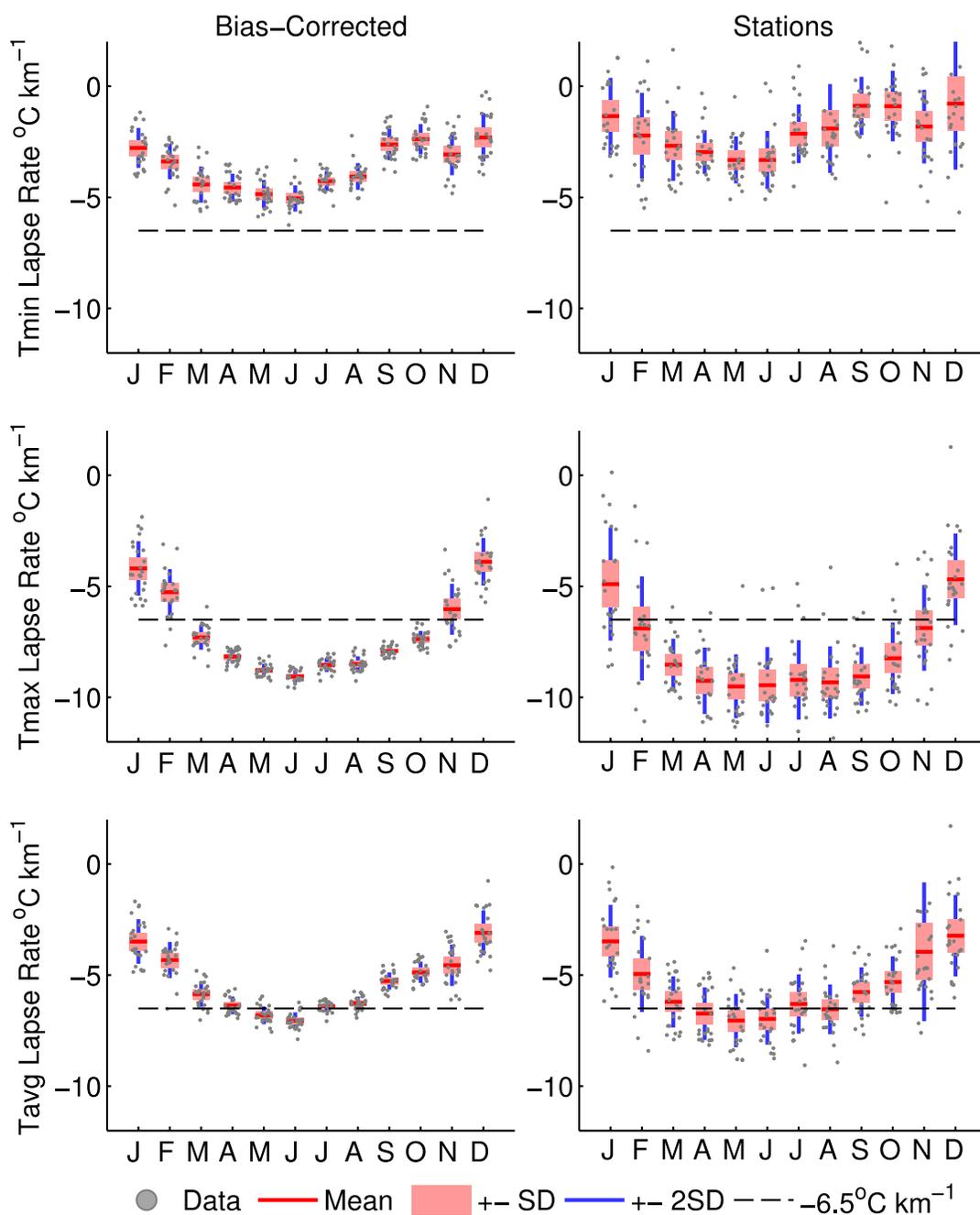


Figure 4.5: Distributions of minimum, maximum, and average monthly temperature lapse rates computed from the bias-corrected temperature data and station temperature data in the study region.

The impact of using a temporally constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate for spatial temperature interpolation in this region is evident when the L13 temperature bias (Figures 4.1 and 4.2) is revisited in light of the observed seasonal lapse rate variability (Figures 4.4 and 4.5). The L13 daily temperature was biased low most significantly in winter (DJF), when observed lapse rates were most shallow relative to $6.5^{\circ}\text{C km}^{-1}$. The minimum temperature bias increased approximately linearly with elevation during all seasons, which corresponded with the strong decrease of minimum temperature lapse rates relative to $6.5^{\circ}\text{C km}^{-1}$ during all seasons. Conversely, L13 maximum temperature was only biased low relative to stations and PRISM during winter, when observed maximum temperature lapse rates were shallower than $6.5^{\circ}\text{C km}^{-1}$. During summer and fall, observed maximum temperature lapse rates were steeper than $6.5^{\circ}\text{C km}^{-1}$, and the increase of maximum temperature bias vs. elevation reflected this (Figure 4.1). Average temperature lapse rates were closely approximated by $6.5^{\circ}\text{C km}^{-1}$ due to the balancing of the strongly shallow minimum temperature lapse rates and the steep maximum temperature lapse rates. However, during late fall and winter, observed average temperature lapse rates were greatly decreased relative to $6.5^{\circ}\text{C km}^{-1}$.

Lapse rates computed from the gridded PRISM data provided very close approximations to the station-derived lapse rates ($R^2 = 0.93$; Figure 4.5), however this was not surprising because the PRISM algorithm computes a unique lapse rate for every grid cell and every month from the station data directly, including the higher elevation SNOTEL sites. PRISM uses locally weighted physiographic parameters that account for the effects of topographic depressions and inversions in addition to the local climate-elevation regression function, which varies spatially and on monthly timesteps [Daly *et al.*, 2008]. For these reasons, PRISM was an effective bias-correction dataset, and lapse rates computed from the bias-corrected data closely followed the observed seasonal variability (Figure 4.5). Additionally, because PRISM provided a close approximation to the station-derived lapse rates, we used the mean monthly average temperature lapse rates computed from PRISM (shown as red lines in Figure 4.5) for downscaling the original L13 and bias-corrected BC temperature data to the 100 m DEM used for our SnowModel implementation. Along with the two datasets downscaled with a constant $6.5^{\circ}\text{C km}^{-1}$, the

resulting set of four downscaled temperature datasets are summarized in Table 3.1.

Hereafter, we note that “between” dataset comparisons will refer to comparisons between the original L13 and bias-corrected datasets (for example L13-6.5 vs. BC-6.5), “within” dataset comparisons will refer to comparisons that differ based on lapse-rate downscaling method but do not differ based on bias-correction (for example, L13-6.5 vs. L13-PLR), and “interdataset” will generally refer to comparisons between datasets as noted.

The effect of bias-correcting with PRISM climate data is evident when maps of the 100 m spatial scale temperature from each of the four downscaled datasets is compared (Figure 4.6). Interdataset differences in downscaled temperature were significant. In Figure 4.6, the mean temperature of each dataset is -1.42°C , -1.40°C , 0.28°C , and 0.26°C from left to right for the L13-6.5, L13-PLR, BC-6.5, and BC-PLR datasets, respectively. Thus, within datasets the effect of lapse rates on mean temperature was much smaller than the between-dataset effect of bias-correction. Further, the effect of topography on temperature patterns was more evident with the bias-corrected downscaled data, but within datasets, the $6.5^{\circ}\text{C km}^{-1}$ lapse rate resulted in stronger temperature gradients. Most striking was the stark contrast in mean winter air temperature in high elevation portions of the model domain between the L13 and BC data. Large portions of the model domain were estimated by the downscaled L13 data to have mean winter air temperature less than -4°C , whereas the same regions with the BC data were between -1 and -3°C (Figure 4.6).

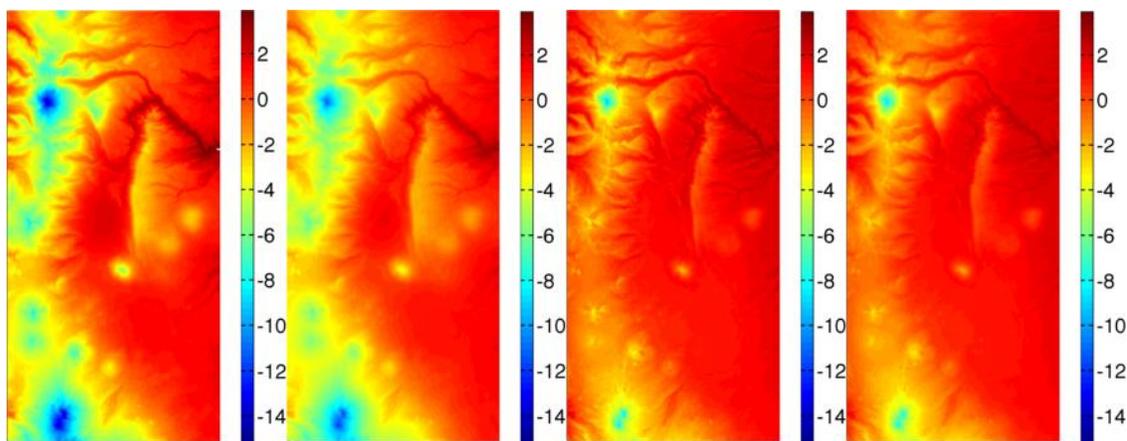


Figure 4.6: Downscaled surfaces of mean daily air temperature for the December – March period for a) L13 data downscaled with a constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate, b) L13 data downscaled with PRISM lapse rates c) BC data downscaled with a constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate, and d) BC data downscaled with PRISM lapse rates

When evaluated against measurements of temperature and precipitation from weather stations in the study region, the downscaled bias-corrected temperature matched observations with much higher accuracy than the uncorrected data, especially during winter (Figure 4.7). The mean NSE of the downscaled temperature ranged from 0.98 to 1.00 for the bias-corrected data, and from 0.93 to 1.00 for the uncorrected data. The mean NSE of the downscaled precipitation was much more variable and ranged from 0.86 to 0.97 for the bias-corrected data and from 0.74 to 0.90 for the uncorrected data. In particular, the effect of the wintertime cold bias in the L13 data was pronounced at the SNOTEL and RAWS measurement stations that were not used as input to the L13 interpolation. Further, the temperature data shown in Figure 4.7 was downscaled with PRISM lapse rates (results for temperature downscaled with the constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate not shown for clarity), indicating that using region-specific lapse rates for downscaling was not sufficient to offset the implicit lapse rate bias present in the original uncorrected climate data. Thus, bias correction was necessary to remove the cold bias in the L13 climate data prior to downscaling. Similarly, precipitation was largely underestimated and scattered on the daily timescale with the L13 data, whereas the BC data matched station observations with much higher accuracy (Figure 4.8).

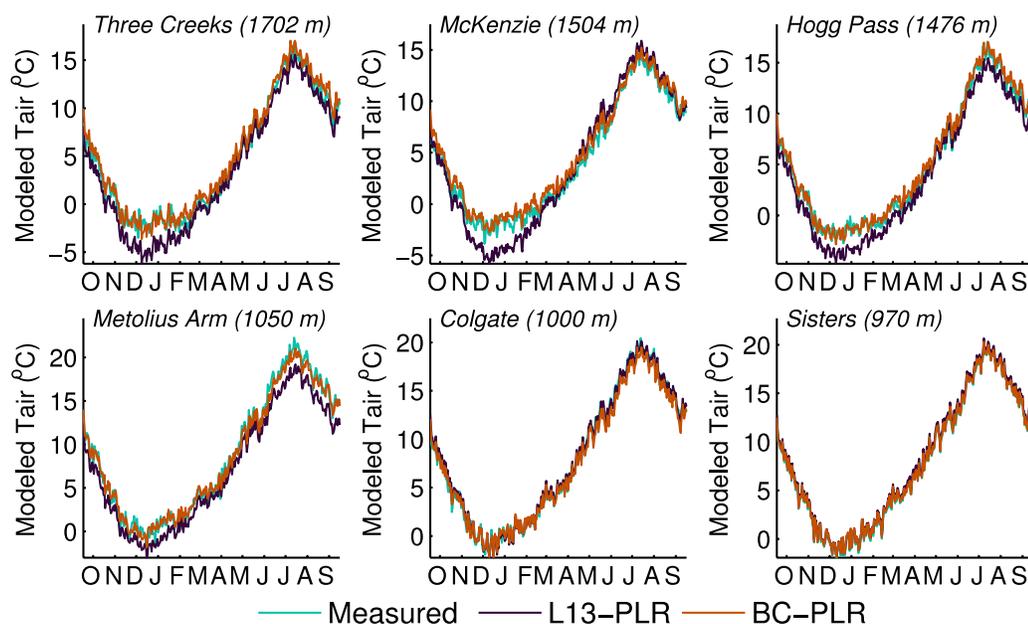


Figure 4.7: Downscaled mean daily air temperature compared to measurements from each meteorological station in the study region. Sisters COOP station was the only station used as input to the L13 interpolation algorithm; all stations were used as input to PRISM.

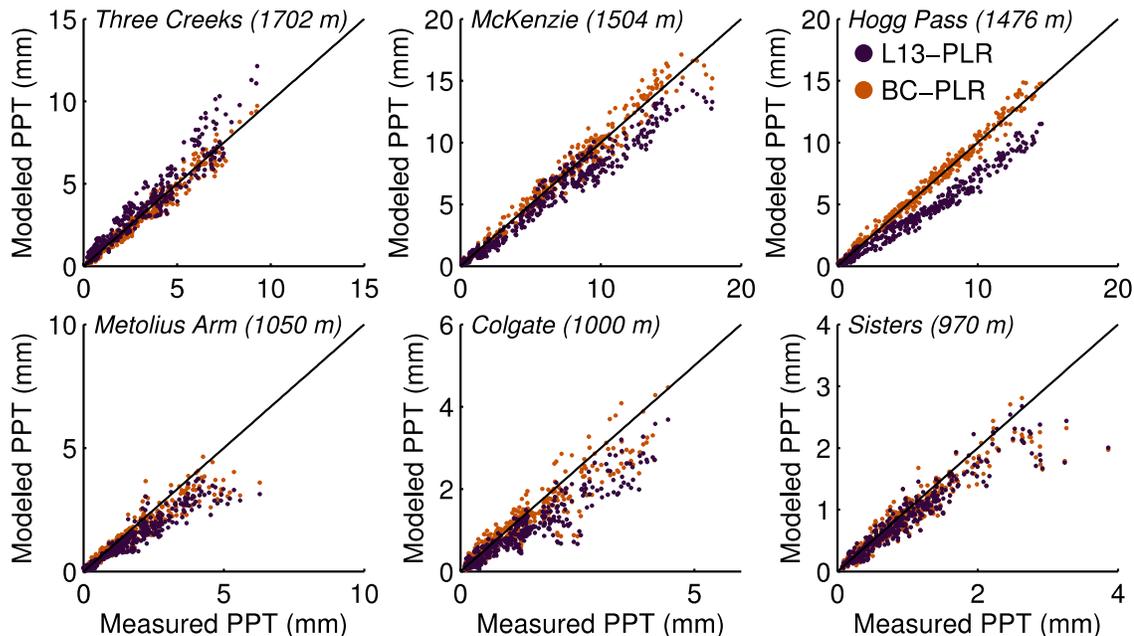


Figure 4.8: Downscaled mean daily precipitation compared to measurements from each meteorological station in the study region. Sisters COOP station was the only station used as input to the L13 interpolation algorithm; all stations were used as input to PRISM.

The fact that the bias-corrected temperature and precipitation data was in such close agreement with station measurements is not surprising because the same stations used for evaluation are included as input to the PRISM model, which we used as the correction dataset. We don't present these results as a validation of the PRISM model, or as evidence that our methods fundamentally improve our understanding of spatial interpolation of climate data. For example, we don't evaluate the ability of the model to simulate air temperature and precipitation at independent measurement locations. Our results are presented to demonstrate a transferrable methodology that can be used to generate improved daily meteorological forcing data for any region in the Continental US, including data-sparse regions. The bias-correction method combined the high-temporal resolution daily variability from the L13 data with the high-spatial resolution monthly average conditions from PRISM. The PRISM model is significantly more sophisticated with regards to the effects of physiography on spatial climate patterns and computes a unique temperature lapse rate for every grid cell from the station data directly, however it is only provided on monthly timescales. Thus, PRISM bias-correction provided a convenient and transferrable method to incorporate the additional

spatial variability, algorithmic complexity, and expert knowledge of the PRISM model into the daily variability from the L13 dataset.

4.3. Baseline SnowModel runs with original and bias-corrected data

Interdataset comparisons of simulated SWE with each of the four downscaled forcing datasets and the uncalibrated model revealed a wide range of model performance when evaluated against measurements of SWE at the SNOTEL stations in the study region (Figure 4.9). Simulated SWE with the original L13 data appeared qualitatively better than with the bias-corrected data in most years. This is surprising given that the L13 precipitation almost uniformly underestimated measured precipitation in the study region except for a handful of months, mostly in drier years. Further, the bias was largest in winter (-38mm m^{-1}) when the majority of precipitation in the study region occurs as snowfall. The presence of such large precipitation bias when aggregated to a monthly timestep indicated an overall mass balance bias and not a temporal misalignment of precipitation magnitude (Figure 4.2), which one might expect to be reflected in simulated SWE. However, monthly L13 winter temperature also uniformly underestimated measured temperature, which resulted in more simulated snowfall, all else equal. Thus, the cold bias appeared to compensate for the low precipitation bias.

However, given the known low bias in precipitation, the fact that SWE was simulated well indicated that total simulated precipitation was skewed toward snowfall. To further demonstrate the impact of the temperature bias relative to the precipitation bias, we simulated SWE with bias-corrected precipitation but left the temperature data uncorrected (L13 PPT-C, Figure 4.9). In all years, simulated SWE with bias-corrected precipitation was greater or equal to simulated SWE with the original, uncorrected data. Further, in many years, simulations with bias-corrected precipitation but uncorrected temperature performed best between datasets (e.g. 1989, 1990, 2000). This verified that the temperature and precipitation bias each influenced simulated SWE, and that low biased temperature resulted in greater simulated SWE when precipitation was corrected. However, it also verified that SWE-melt was overestimated during years when simulations matched observations. Given the known low temperature bias and associated overestimated snowfall with the uncorrected temperature data, it follows that SWE-melt

was overestimated during the accumulation period in order for the simulated SWE with the uncorrected temperature to match the observations. The source of the overestimated SWE-melt and its relationship to prescribed lapse rate values is explored in more detail throughout the remainder of this paper.

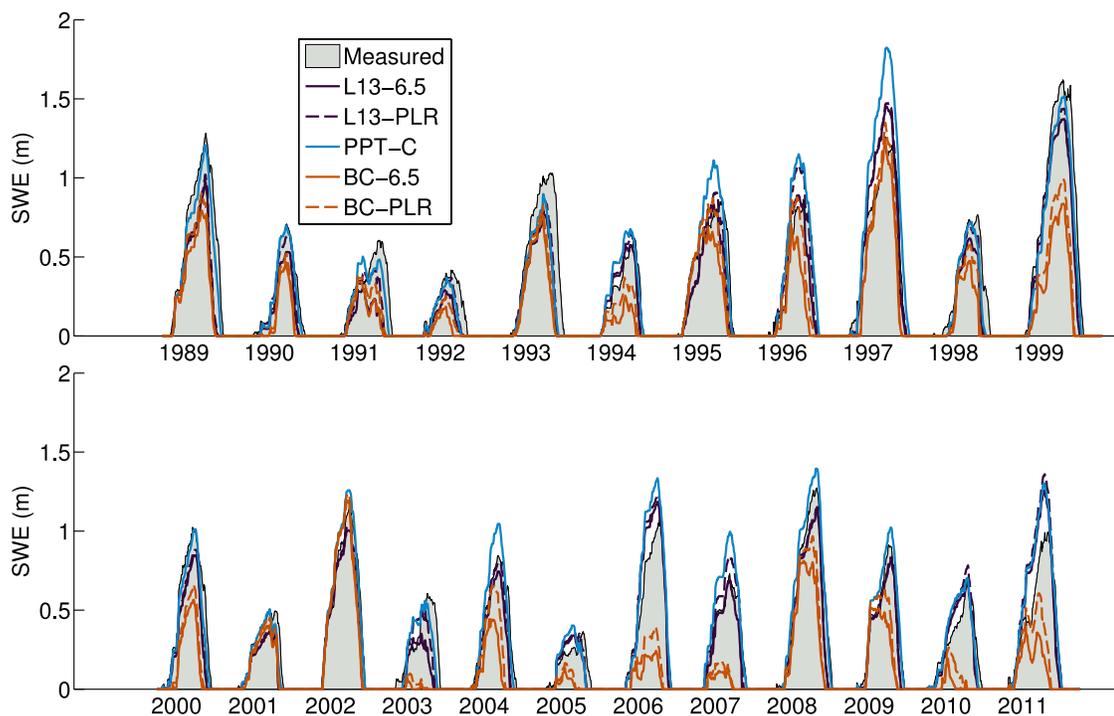


Figure 4.9: Baseline simulations of SWE with each dataset. Timeseries of measured SWE from Three Creeks, Hogg Pass, and McKenzie SNOTEL, and corresponding modeled timeseries, are averaged to create one representative timeseries.

Baseline (uncalibrated) SWE simulations with the bias-corrected input data were not improved relative to the original data (Figure 4.9). In all years, there were observed mid-winter accumulation events that were not simulated and simulated melt that was not observed. The most consistent patterns with the bias-corrected forcing data were underestimated peak SWE and early melt leading to early snow-disappearance, indicating errors in both the snow accumulation and melt-phase process representations of the model. The poor model performance with the bias-corrected forcing data suggested that model parameter or structural bias might explain the inability of the model to simulate the accumulation and ablation phases accurately.

Most surprisingly, within both the original and bias-corrected datasets, the warmer PRISM lapse rates resulted in more simulated SWE than simulations with the constant

$6.5^{\circ}\text{C km}^{-1}$ lapse rate (Figure 4.9, solid vs. dashed lines). Thus, the slightly warmer temperature data downscaled with PRISM lapse rates resulted in more SWE than the slightly colder temperature data downscaled with the constant lapse rate. This result is explained in Figure 4.10, where the mean simulated accumulation period (Dec. 1 – Apr. 1) SWE, SWE-melt, precipitation, snowfall, incoming long- and shortwave irradiance, air temperature, and relative humidity are shown for each dataset across the full elevation range of the model domain.

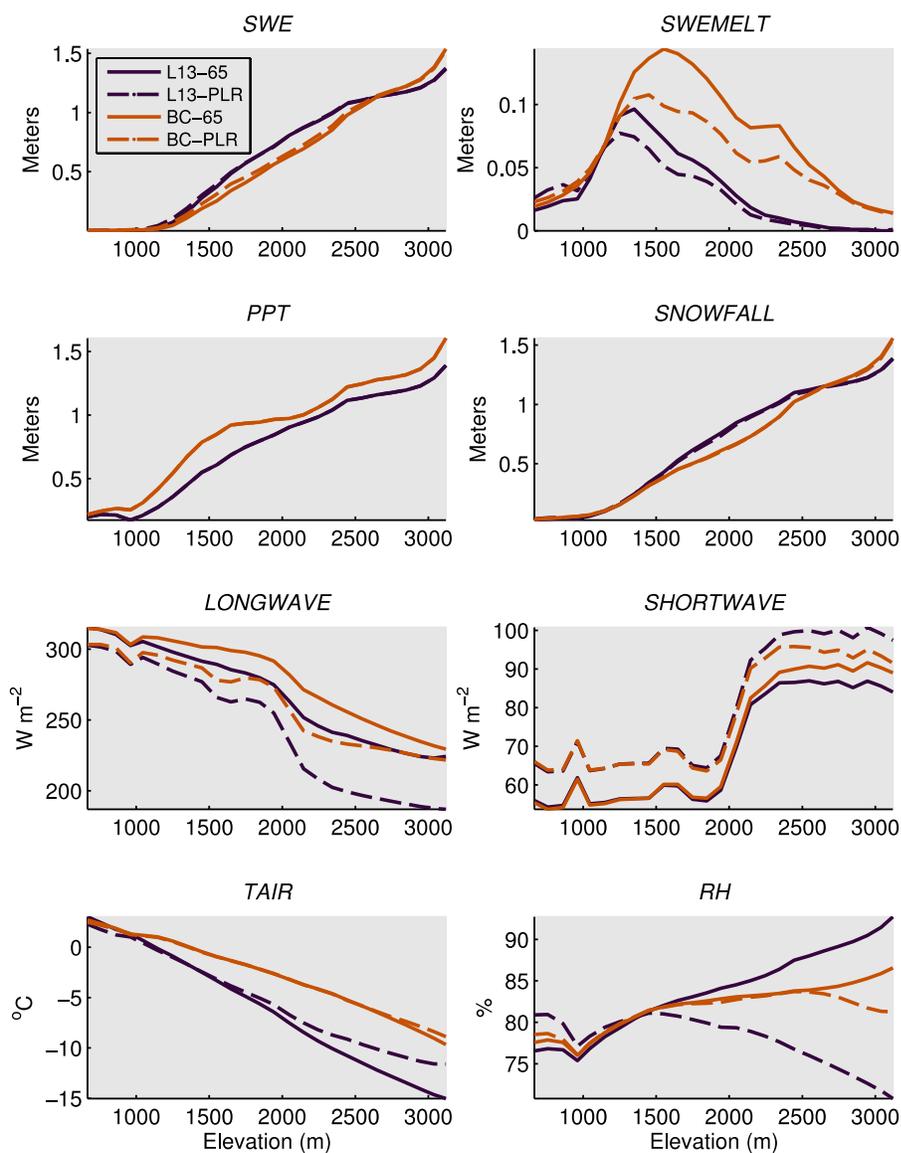


Figure 4.10: Mean values of simulated meteorological variables vs. elevation with the uncalibrated model. Hydrologic fluxes are cumulative sums on April 1; irradiance, temperature, and humidity are daily averages for the Dec. 1 – Apr. 1 period, binned and averaged into 100 m elevation bands across the 1989-2011 period.

Within the L13 and BC datasets, the colder, constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate resulted in more simulated incoming longwave irradiance and more mid-winter SWE-melt than the warmer PRISM lapse rates (Figure 4.10) (hereafter, “incoming longwave irradiance” and “incoming shortwave irradiance” are referred to as longwave irradiance and shortwave irradiance, respectively, as the emitted, reflected, and net components are not discussed). However, between datasets (lapse rates equal) the opposite effect was observed and simulated longwave irradiance was less with the colder L13 data than the warmer BC data. The latter effect is expected, but the counterintuitive former result is explained by two competing factors in the model. First, by definition the steeper $6.5^{\circ}\text{C km}^{-1}$ lapse rate will estimate slightly colder surface temperatures than the shallower PRISM lapse rates and subsequently less longwave irradiance derived from the surface temperature estimates. However, the model treats free atmosphere lapse rates the same as surface temperature lapse rates. Thus, the steeper $6.5^{\circ}\text{C km}^{-1}$ lapse rates will always estimate colder 700 mb tropospheric temperatures. These colder 700 mb temperatures trigger the cloud-fraction parameterization more frequently, which increases the estimated increase in longwave irradiance due to cloud cover. Ultimately, within datasets, the increased SWE-melt due to the increased longwave irradiance as a result of the steeper lapse rates outcompeted the effect of the colder surface temperatures on snow accumulation, and other processes that tend to increase simulated SWE with colder temperatures, in particular in the 1500 – 2000 m elevation range where the three SNOTEL stations are located (Figure 4.10).

To compute the cloud fraction, dew point temperature is calculated from the co-located temperature and humidity input data. The Barnes objective analysis scheme [Barnes, 1964; Koch *et al.*, 1983] then interpolates the dew point temperatures to the model grid. The gridded dew point temperatures and air temperatures are then adjusted to the 700 mb atmospheric level (approximated as 3000 m in the model) using monthly varying air and dew point temperature lapse rates. These are then converted to vapor pressure and 700 mb relative humidity, which is used to estimate the cloud fraction at the 700 mb level with the equation:

$$\%f = \left(78 + \frac{\Delta x}{15.5}\right) \times \left(\frac{P/P_s - 0.1}{0.6}\right) \times e^{\frac{(RH_{700}-1)}{(1-RH_e)}}$$

where RH_{700} is the relative humidity at the 700 mb level (0-1), Δx is the averaging area (Δx^2), P/P_s is the ratio of atmospheric to surface pressure, and $1 - RH_e$ is the “e-folding” critical relative humidity at which cloud cover is 100%:

$$1 - RH_e = 0.196 + \left(\frac{0.76 - \Delta x}{2834}\right) \times \left(1 - P/P_s\right)$$

With $\Delta x = 80$ km (the minimum averaging area from *Walcek*, [1994]) and $P/P_s = 0.7$ at 700 mb, the equation is equivalent to:

$$\%f = 0.8316 \times e^{\frac{(RH_{700}-1)}{0.4155}}$$

Thus, because RH increases as temperature decreases, the colder 700 mb temperatures simulated with $6.5^\circ\text{C km}^{-1}$ vs. the PRISM lapse rates tended to simulate higher relative humidity values at the 700 mb level and higher estimated cloud fractions (Figure 4.11) than the PRISM lapse rates, which increased the longwave irradiance, leading to increased SWE-melt.

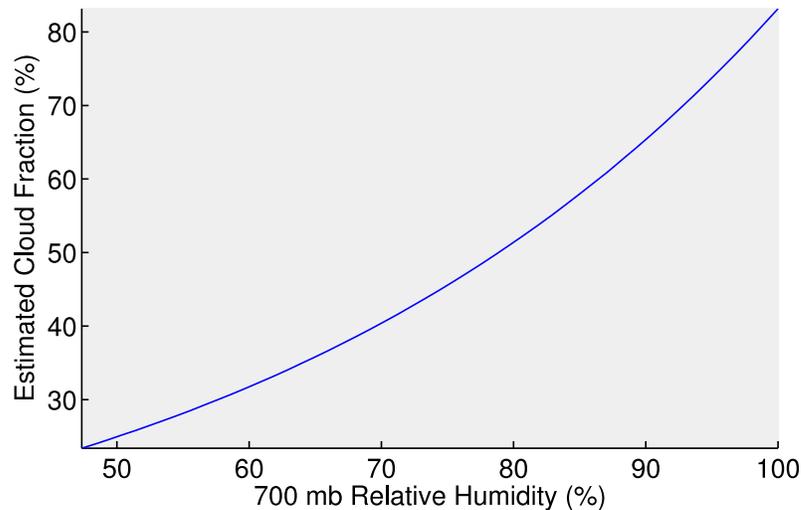


Figure 4.11: The function used by Snow Model to estimate cloud fraction from 700 mb relative humidity.

However, despite the apparent decrease of accumulated SWE due to increased longwave irradiance and subsequent SWE-melt that in turn were due to the colder lapse rates when comparing within the L13 and BC datasets, the same effect on SWE

accumulation was not observed when comparing between the colder L13 and warmer BC surface temperature datasets. These results suggest there is a tradeoff between the sensitivity of the model-estimated longwave irradiance to differences in absolute surface temperature, and to differences in prescribed lapse rates (from which 700 mb air temperature is derived) that influence the estimates of cloud fraction. This effect is such that within the BC dataset, small decreases in surface and 700 mb temperature due to steeper lapse rates lead to increased longwave irradiance, increased SWE-melt, and decreased SWE, but between datasets the L13 data was sufficiently colder than the BC data to overwhelm the cloud fraction effect on SWE-melt and accumulated SWE. These results lead to the conclusion that competing processes in the model which depend on temperature and lapse rates can lead to counterintuitive effects on modeled snow accumulation and ablation.

To summarize, the counterintuitive result was that within datasets, the steeper $6.5^{\circ}\text{C km}^{-1}$ lapse rate resulted in less simulated SWE than the shallower PRISM lapse rates (Figure 4.9) because they resulted in greater estimates of longwave irradiance (Figure 4.9 and 4.10), which increased SWE-melt during the accumulation period relative to the PRISM lapse rates. Between datasets, however, the colder surface temperature data from the L13-6.5 data did not result in less SWE than the warmer surface temperature data from the BC-PLR data, even though the longwave irradiance was almost identical on average (Figure 4.9). Further evidence for the sensitivity of modeled SWE to the cloud fraction parameterization is provided in the following two sections.

4.4. Model parameter estimation

Modeling snow in environments where mixed-phase precipitation and mid-winter melt is common is a major research challenge [Essery *et al.*, 2013]. In the relatively warm winters of the PNW, snow is isothermal for much of winter and thus relatively small energy inputs are required to produce melt [Lundquist *et al.*, 2013]. Further, accumulation is difficult to predict with simple static temperature partitioning thresholds because most winter precipitation falls close to the melting point (Figure 4.12). Diagnosing model structure or parameter bias is particularly difficult in data-sparse regions where measurements of irradiance and snowfall are unavailable. For these reasons, diagnosing

the exact physical mechanisms that explain the differences between modeled and measured quantities was not possible in this data sparse region. A common alternative approach is to calibrate a hydrologic model based on a desired outcome (e.g. match the observed mean monthly streamflow) and assume the model correctly represents physical processes [Kirchner, 2006]. We used this approach, but first focused on the validation of the temperature and precipitation inputs via bias-correction. We then employed the parameter optimization routine to determine the sensitivity of the model to different parameters given different input data combined with a given model structure.

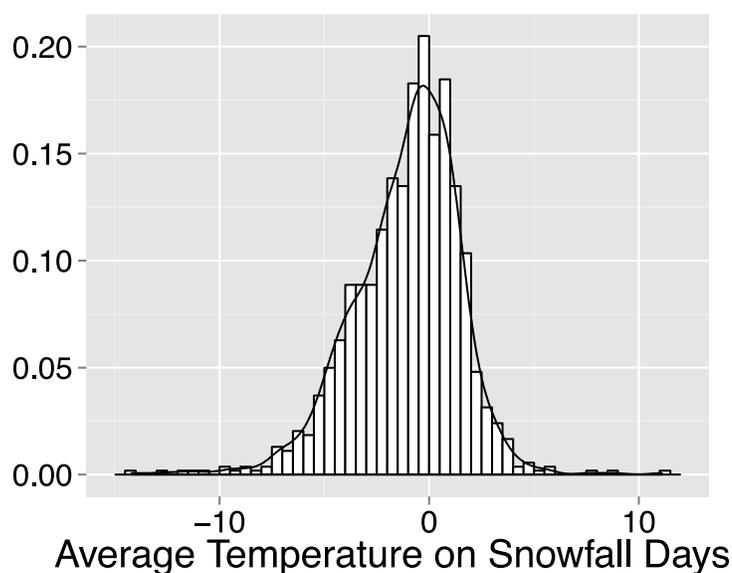


Figure 4.12: The distribution of average daily air temperature ($^{\circ}\text{C}$) on days where snow accumulated in the study region from SNOTEL records. The expected value is very close to 0°C with the majority of precipitation events between 0 and -2°C .

Cumulative posterior distributions for the top 100 ranked parameter values are compared between datasets and to the uniform prior distribution for each parameter in Figure 4.13. Parameters were ranked based on the average of the NSE score computed between the 11-year timeseries of modeled and measured SWE at the Three Creeks and Hogg Pass SNOTEL stations and the grouped population of manual SWE measurements from Racing Creek and Bald Peter Snowcourses and the Hogg Pass field measurements. The snow-threshold temperature (TS) showed strong convergence within each dataset, but converged to different values between datasets. For the bias-corrected L13 data, the majority (interquartile range) of parameter sets that modeled SWE accurately used a TS

value between -0.14 and -0.67°C , whereas with the original L13 data the majority of parameter sets in the posterior distribution ranged between -0.87 and -1.38°C . With a colder lower temperature threshold it is less likely that precipitation will be estimated as snowfall, thus convergence to a lower threshold offsets the cold bias effect on snow accumulation. Consequently, between datasets, the colder data converged to significantly colder values of TS. The same effect was evident to a smaller degree within datasets. For example, the L13 data downscaled with a constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate (L13-6.5) propagated the winter-time cold bias into the downscaling and produced colder estimates of surface temperature than when PRISM lapse rates were used to downscale (Figure 4.10), thus the TS converged to a colder temperature (Figure 4.13). However, this effect between downscaling procedures was not evident within the BC data, possibly indicating that the increased spatial resolution of the 800 m BC data vs. the 6 km L13 data limited the influence of the downscaling procedure on the estimate of 100 m temperature and the corresponding influence on the estimate of TS. While there is considerable uncertainty as to the physical relationship between daily average air temperature measured at a few weather stations in a region and the phase of on-the-ground precipitation, the range of values estimated with the BC data were closer to the theoretical 0°C isotherm where precipitation should occur as rainfall, and the posterior distribution appeared more reasonable for TS with the BC data than with the L13 data.

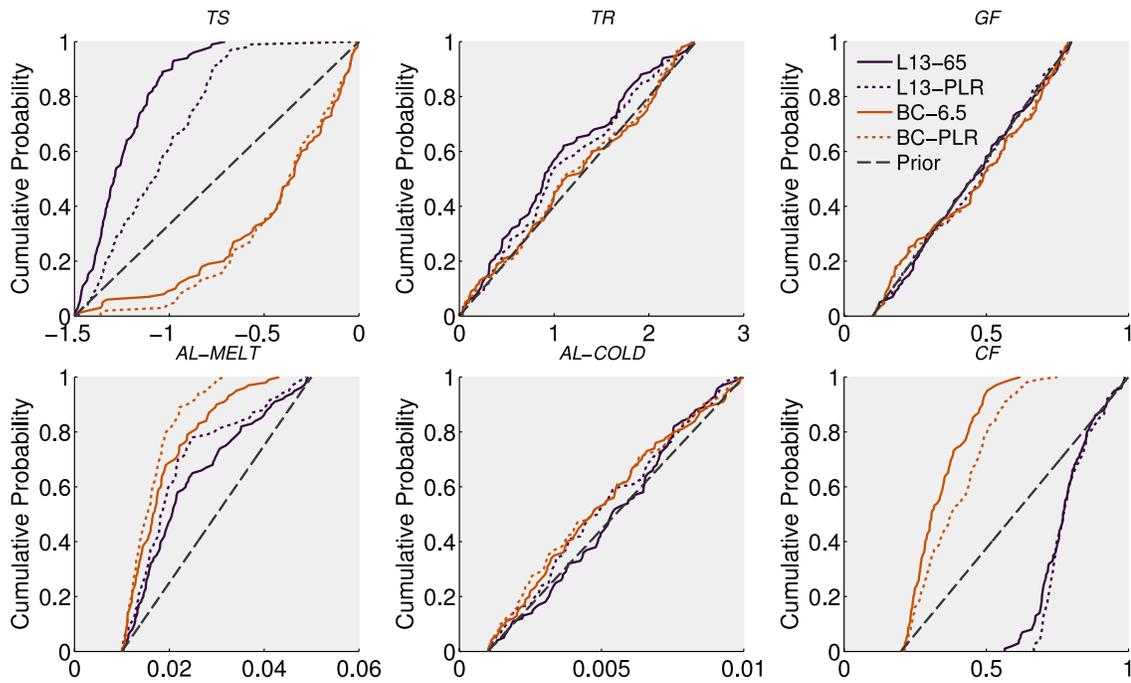


Figure 4.13: Cumulative posterior and uniform prior distributions for each calibration parameter. The TS, AL-MELT, and CF parameters show strong convergence to expected values, demonstrating the strong control on model performance that these parameters exert.

The melting-snow albedo decay gradient (AL-MELT) appeared to be less sensitive to the choice of forcing data, as the majority of posterior values ranged between 0.012 and 0.032 for each dataset. Some differences between datasets emerged, most notably between the L13-6.5 data and the BC-PLR data, which were the coldest and warmest data, respectively. Here, the colder data tended to simulate SWE accurately with larger values of AL-MELT, possibly to compensate for the decreased likelihood of melting conditions (air temperature $> 0^{\circ}\text{C}$) occurring with this data. A larger value of AL-MELT causes faster albedo decay, which increases the absorption of solar irradiance into the snowpack, thus increasing the internal energy of the snowpack. Conversely, the cold-snow albedo decay gradient (ALCOLD) did not converge toward any particular value in the prior distribution, and did not differ in any significant way between datasets, likely because it is only effective during non-melting conditions and hence only contributes directly to the snowpack energy balance during these periods. Further, albedo decay during non-melting conditions is parameterized as much slower (i.e. smaller gradients) than during melting conditions, therefore the value of AL-COLD has a smaller impact on the snowpack energy balance than AL-MELT. However, it should also be noted that

melting conditions are defined in SnowModel as conditions where air temperature is greater than 0°C, and are not defined based on the actual snow surface temperature.

The cloud-factor (CF) differed significantly between the L13 and BC data, and to some extent depended on lapse rates within each dataset, particularly the BC data. At first glance, this may be counterintuitive because the previous section showed that within datasets, differences due to longwave that resulted from choice of lapse rates led to large differences in modeled SWE and SWE-melt, whereas between datasets, differences due to longwave were not as important as the air temperature differences. Thus, one might expect parameter convergence for CF to differ strongly within datasets, to offset the influence of lapse rates on estimated longwave irradiance. However, a close examination reveals consistency between these results.

As discussed in the previous section, the cloud-factor scales the increase in longwave irradiance due to the model-estimated cloud fraction, which is a function of 700 mb relative humidity. Within datasets, differences in 700 mb relative humidity depend on lapse rates, and steeper (colder) lapse rates result in lower relative humidity and larger cloud fractions. Accordingly, within datasets, slightly colder lapse rates resulted in larger estimates of longwave irradiance, more SWE-melt, and less SWE, but between datasets the colder L13 surface temperature data resulted in smaller estimates of longwave, less SWE-melt, and more SWE, when lapse rates were equal. The key to disentangling the effect of the cloud fraction on simulated SWE is to compare the warmest and coldest data, which have almost identical estimates of longwave irradiance but very different SWE and SWE-melt (Figure 4.10 orange dashed line vs. purple solid). Here, the increase in longwave due to the colder lapse rates (L13-6.5 vs. L13-PLR) resulted in equivalent estimates of longwave as the warmest dataset (BC-PLR), but simulated SWE-melt was much larger and simulated SWE much smaller with the warmest data, indicating that the L13 data was sufficiently cold enough to offset the longwave effect. Via parameter estimation, the sensitivity of the model to these key differences was confirmed. The CF tended toward values between 0.71 and 0.85 for the L13 data, and to 0.26 and 0.49 for the BC data, indicating that the CF tended toward larger values, which preserved more of the estimated longwave irradiance, for the colder L13 data than with the warmer BC data to

compensate for inhibited melting conditions throughout the winter with the colder L13 data. However, within datasets, the opposite effect occurred; the slightly colder temperature lapse rates resulted in slightly smaller CF estimates, which scaled back the increase in estimated longwave due to the colder lapse rates. This significant behavioral difference demonstrates the strong influence of the CF parameter on model performance, and its dual dependence on absolute surface temperature of the input forcing data and prescribed lapse rates. However, the lack of direct measurements of longwave irradiance in our study region is a key limitation of our ability to draw any firm conclusions about the realism of the values for this parameter and its impact on simulated longwave irradiance. The importance of this uncertainty is highlighted by the fact that this limitation applies to the vast majority of distributed hydrologic model simulation experiments. In either case, the difference between the range of estimated values of CF for the L13 data and the BC data lead to very large differences in simulated longwave irradiance. This difference played a key role in interdataset differences in modeled SWE during the reference historical climate period and to a lesser extent during the climate change scenario periods, discussed in the following sections.

Finally, the rain-threshold temperature (TR) and the forest gap-fraction (GF) lacked convergence toward any particular value, although the optimal GF estimate used for all final simulations was very similar between datasets (Table 4.1). This result was somewhat surprising, given the importance of net shortwave radiation for snowpack energy balance. Larger values of GF increase the transmission of direct-beam solar irradiance to the snowpack beneath the forest canopy. However, this parameter only influences snow simulations at locations defined as forest by the land-cover boundary condition dataset. In our modeling domain, the McKenzie SNOTEL and both Snowcourse locations are defined as forested. However, McKenzie SNOTEL was not included in the ranking algorithm, and the Snowcourses were grouped, thus the influence of forest processes on model evaluation was limited in this experiment. Further, GF determination is a source of considerable uncertainty when using SNOTEL data as evaluation for grid-based models where land cover is defined uniformly for each grid cell. SNOTEL stations are typically sited in clearings to avoid the influence of the forest

canopy on automated measurements. However, the stations are often surrounded by dense forest that influence the local climate of the measurement site and it is therefore unclear to what extent forest-snow interactions influence the automated SWE measurements.

Table 4.1 Model parameter values used for the uncalibrated runs, the minimum and maximum values of the uniform prior distributions, and the optimal estimates used for the final simulations.

Parameter	Prior	Un-calibrated	Optimal Estimate			
			L13-6.5	L13-PLR	BC-6.5	BC-PLR
Lower temperature threshold	-1.5 - 0°C	-2	-1.17	-0.71	-0.66	-0.66
Upper temperature threshold	0 – 2.5°C	2	1.88	0.45	0.99	0.99
Cold albedo decay gradient	0.001 – 0.01	0.008	0.0097	0.0011	0.0061	0.0061
Warm albedo decay gradient	0.01 – 0.05	0.018	0.022	0.013	0.011	0.011
Cloud factor	0.2 – 1.0	1.0	0.54	0.79	0.29	0.29
Gap fraction	0.1 – 0.8	0.4	0.12	0.13	0.15	0.15

4.5. Model calibration, validation, and evaluation

Differences in simulated SWE between datasets remained post-calibration, but differences were much smaller than with the uncalibrated model (Figure 4.14). The largest differences post-calibration were between the original and bias corrected datasets and did not tend to depend on lapse rates, shown by the negligible difference between dashed and solid lines in Figure 4.14. In the case of the bias-corrected data, this is not surprising because the same parameter set was ranked highest for both the BC-6.5 and BC-PLR datasets. Thus, the only difference between the calibrated simulations with these two datasets was the lapse rates, which resulted in very slight differences in simulated temperature, shortwave irradiance, and humidity, but essentially no difference in simulated SWE. For the L13 datasets, the slightly warmer L13-PLR data resulted in slightly more SWE and later melt in all but two years, which is consistent with the lapse rate-cloud fraction parameterization effect observed in the pre-calibration simulations. However, these differences were negligible compared to the large differences between the L13 and BC datasets. As noted for the pre-calibration simulations, this indicates that

bias-correction alone explained much more of the interdataset differences than lapse rates. However, unlike the pre-calibration simulations, post-calibration model sensitivity to lapse rates was larger (albeit small in absolute terms) for the coarse spatial-resolution L13 data than the bias-corrected data.

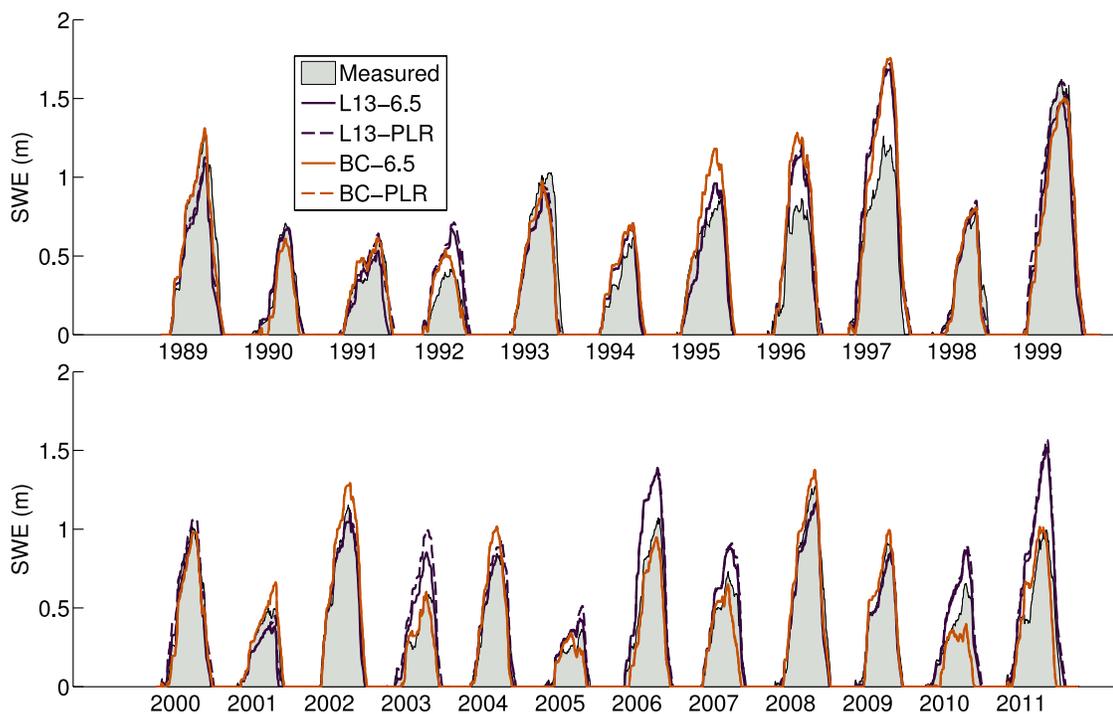


Figure 4.14: Calibrated simulations of SWE with each dataset. Timeseries of measured SWE from Three Creeks, Hogg Pass, and McKenzie SNOTEL, and corresponding modeled timeseries, were averaged to create one representative timeseries.

When compared to measured SWE at the three SNOTEL sites in the study region, simulated SWE with each dataset was very similar. All datasets performed worse during the validation period than the calibration period, but objective evaluation during the validation period was dominated by poor model performance during 1995, 1996, and 1997 (Figure 4.14). During February 1996, a large rain on snow event melted significant proportions of the Cascades snowpack in a matter of days [Marks *et al.*, 1998]. During this event, the sensible and latent heat fluxes accounted for 60-90% of the snowpack energy balance, thus it is not surprising that the model was unable to simulate these anomalous conditions and failed to capture the large melt event. Similarly, in 1997 there were anomalously large mid-winter melt events evident in the measurements that were not simulated by the model, indicating some weakness in the model's ability to simulate

anomalous mid-winter melt. Measurements of temperature and precipitation from the SNOTEL stations indicated that 1995 was wetter and colder than average, 1996 was very wet and slightly colder than average, and 1997 was very wet and much colder than average (Figure 4.15). Short periods of strong temperature inversions persisted during each of these winters, but similar inversions persisted during all winters in the study period. Similar climatic years, such as 1999 and 2008, were simulated well post-calibration. As noted previously, the exact mechanisms that explain differences between model simulations and observations were impossible to identify in this data-sparse region. Perhaps the simplest explanation is that in 1995, 1996, and 1997, the accumulation phase was well simulated with each dataset prior to calibration, whereas in the majority of years during the calibration period accumulation was underestimated prior to calibration. The two parameters that showed the strongest impact on model performance (TS and CF), both converged to values that increased SWE accumulation. For example, the snow threshold temperature converged to a less-negative value (-0.66 vs. uncalibrated -2.0), which by definition produced a larger quantity of accumulated solid precipitation, and the CF converged to values less than 1.0, which reduced the longwave irradiance. Thus, in years when the accumulation phase was well simulated prior to calibration, the post-calibration simulations tended to overestimate accumulation.

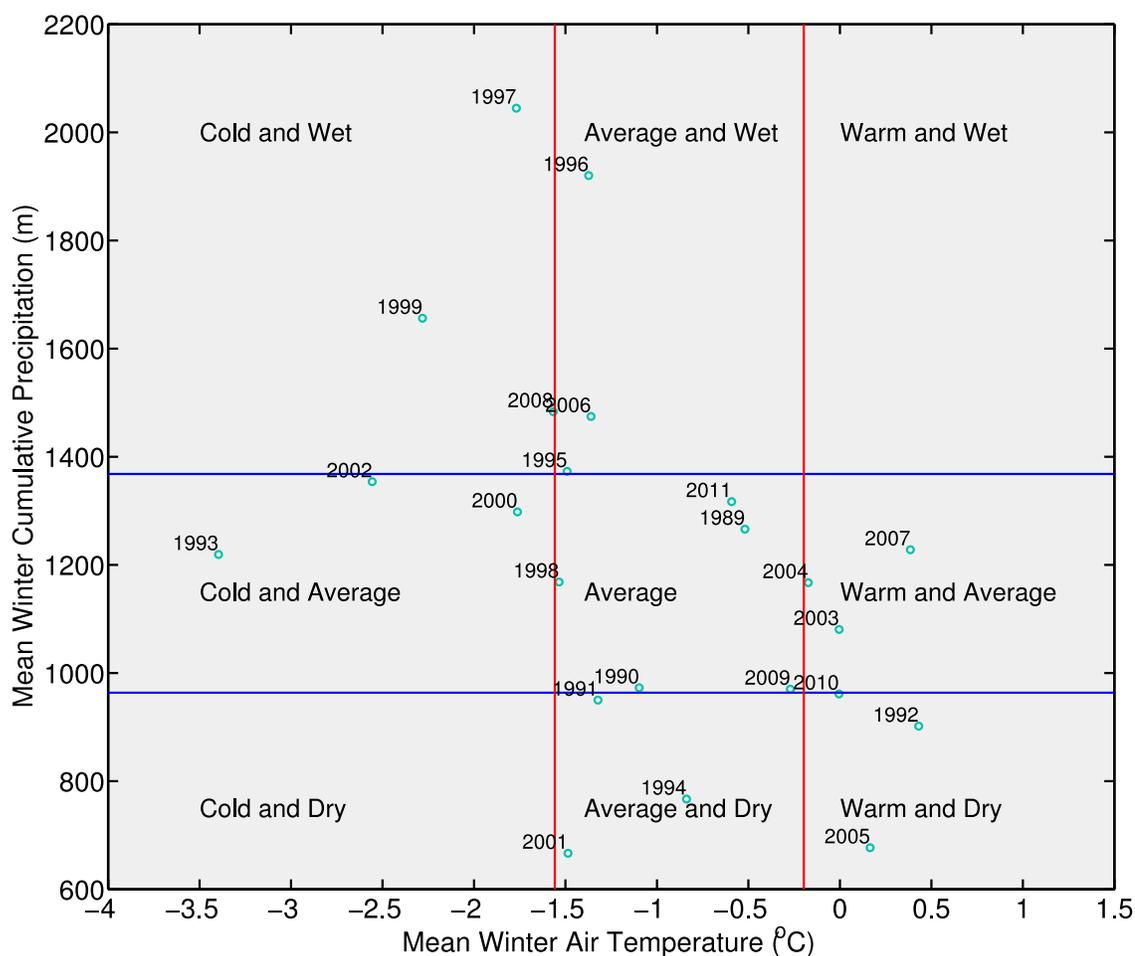


Figure 4.15: Each year in the study period classified by mean winter air temperature and April 1 cumulative precipitation at Hogg Pass and Three Creeks SNOTEL. Vertical red and horizontal blue bars are the interquartile range for each distribution.

In general, model performance was better during the calibration period than the validation period, but all datasets performed with reasonable skill (~ 0.70 NSE) during the validation period (Table 4.2). Objectively, the bias-corrected data performed better than the original L13 data during the calibration period, but the original L13 data performed better during the validation period. However, both of these differences were small (e.g. NSE = 0.86 for L13-PLR during the calibration period vs. 0.91 for BC-PLR, and NSE = 0.71 for L13-PLR during the validation period vs. 0.68 for BC-PLR). A much larger difference between datasets was observed for the difference between pre- and post-calibration model performance. Here, the improvement due to calibration was much larger with the bias-corrected data than the original data (Table 4.2). This effect was due to the strong convergence of the CF and TS parameters with the warmer bias-corrected

data that corrected the effects of underestimated snow precipitation and increased longwave irradiance on snow accumulation. Nonetheless, performance was very similar between all datasets, and some notable limitations to model performance were identified. In particular, the RMSE was large with all datasets, but the general characteristics of the regional snowpack were simulated well.

It should be noted that the calibration routine was by no means exhaustive and was designed to vary six parameters of a high dimensional, complex physical model. The full sensitivity of the model to all possible parameters was not the focus of this investigation. Additional performance gains and convergence between datasets may have been possible with a comprehensive calibration experiment. Further evaluation of model performance for the entire 23-year study period and implications of this performance for the study objectives are considered in the following section.

Table 4.2 Modeled SWE evaluation for the uncalibrated/calibrated model during the calibration (2000 – 2011) and validation (1989 – 1999) period for each dataset.

Dataset	NSE (-)		RMSE (cm d ⁻¹)		ME (cm d ⁻¹)	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
L13-6.5	0.81/0.88	0.57/0.71	16/13	22/18	-8.6/1.6	-12.2/1.6
L13-PLR	0.85/0.86	0.68/0.71	14/14	19/18	-4.7/6.4	-7.0/5.1
BC-6.5	-0.07/0.91	0.23/0.69	39/11	29/19	-28.4/-2.4	-21.4/5.4
BC-PLR	0.19/0.91	0.45/0.68	34/11	24/19	-23.7/-2.1	-15.8/5.6

4.6. Model evaluation and reference period simulations across elevation gradients

In addition to evaluating the model during the calibration and validation periods using the NSE metric computed for each period timeseries, we further evaluated interannual and mean model performance for the entire 23-year reference period at each measurement location. The calibration and validation periods provided an indication of the impact of parameter estimation on model skill, the transferability of the estimated parameters from the calibration to the validation time period, and key model process sensitivity. The model evaluation presented in this section focuses on model performance at each measurement site for the entire reference period to provide an indication of the spatial variability of interannual and mean model performance. This evaluation is presented because the climate impact assessment is based on the mean model simulations at each site and for basin-wide mean model simulations during the reference period.

In addition to computing the NSE, RMSE and ME for each year and for the mean simulations, we also evaluated the ability of the model to simulate the date and amount of peak SWE, the date of snow disappearance, and the length of the snow-covered season. These metrics correspond to the metrics we used to evaluate the impact of climate warming on the snowpack, thus improving our confidence in the ability of the model to simulate these characteristics of the snowpack.

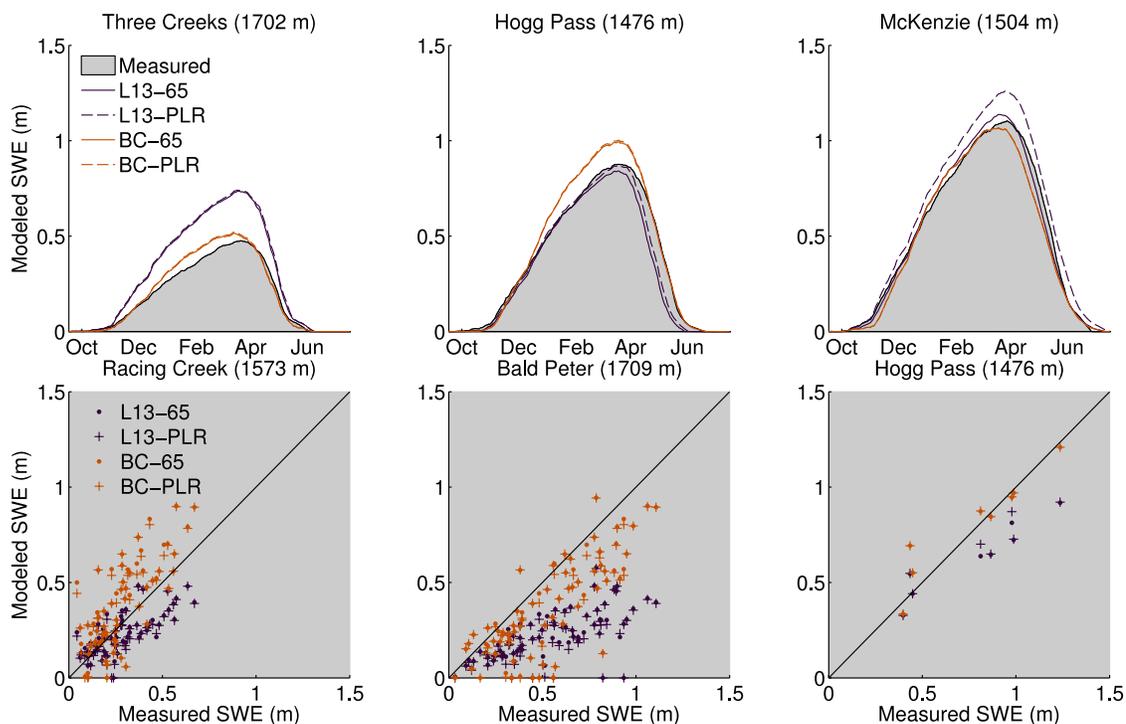


Figure 4.16: Mean daily-simulated SWE compared to measurements at each of the three automated SNOTEL sites (top row) for the entire 23-year reference period. Manual SWE measurements from Snowcourses and field measurement sites bottom row.

While objective model evaluation during the calibration and validation periods was similar between datasets when metrics were averaged across sites (Table 4.2), mean model performance at individual sites during the reference period differed between datasets (Figure 4.16; Table 4.3). Model performance was more consistent between sites with the bias-corrected data, which tended to overestimate SWE during the accumulation phase, although this was largely due to the anomalous overestimates during the winters of 1995, 1996, and 1997. The spring snowmelt period, in particular the date of snow disappearance, was simulated with very high accuracy at all sites with the bias-corrected data. Results from the uncorrected L13 data, on the other hand, were inconsistent

between sites. The date and amount of peak SWE and the date of snow disappearance at Three Creeks SNOTEL were overestimated, but at Hogg Pass SNOTEL each of these metrics were underestimated. At McKenzie SNOTEL, the results differed within the uncorrected data depending on the choice of lapse rates, where the warmer lapse rates resulted in more peak SWE and larger errors in the date of snow disappearance. This is likely a result of the higher CF value selected for the final simulations for that dataset. The McKenzie SNOTEL station is on the western side of the Cascades crest where a much wetter winter climate persists and, as noted previously, model performance at McKenzie was not included in the calibration ranking routine. However, we present the model results at McKenzie because it is within the modeling domain and provides an additional source of information on model performance.

Table 4.3. Objective evaluation of each dataset at each SNOTEL location. The non-traditional metrics of performance correspond to the metrics used for the climate change sensitivity analysis

Station Name	Forcing Data	NSE (-)	RMSE (cm)	ME (m)	Peak SWE (% dif)	Date of Peak SWE (days)	Date of Snow Disappear. (days)	Length Snow Season (days)
Three-Creeks	L13-6.5	0.40	13	9	55	-1	5	5
	L13-PLR	0.35	14	10	56	-3	4	4
	BC-6.5	0.97	3	1	8.5	-7	3	-6
	BC-PLR	0.97	3	1	10	-7	3	-6
Hogg-Pass	L13-6.5	0.95	7	-4	-4	0	-19	-31
	L13-PLR	0.97	5	-2	-0.7	0	-14	-26
	BC-6.5	0.97	6	3	13	0	4	-6
	BC-PLR	0.96	6	3	14	0	4	-6
McKenzie	L13-6.5	0.99	3	0	3	-8	13	0
	L13-PLR	0.96	8	6	14	-1	20	7
	BC-6.5	0.98	5	-2	-3	-9	14	1
	BC-PLR	0.98	5	-2	-3	-9	14	1

Interannual model performance demonstrated that the model was able to simulate the mean snowpack at each location with the bias-corrected data with reasonable accuracy, but there was a wide range of model performance between years (Figure 4.17). Model performance was most variable at the Three Creeks SNOTEL station, which is the station that also showed the largest variability between datasets for mean model performance. At this station, the improvement due to bias-correction was most pronounced, in terms of

NSE, RMSE, and ME. Noteworthy is that this is the only station in the region that is truly on the eastern side of the Cascades crest. In terms of estimating the sensitivity of the snow in the leeward side of the Cascades, the significant improvement due to bias correction was especially important at this location.

Within our modeling domain, the only sources of fully independent model validation data were the manual SWE field measurement locations. These included two NRCS-operated Snowcourses (Racing Creek and Bald Peter, Table 2.1), and a small number of direct measurements of SWE that were collected for research purposes near the Hogg Pass SNOTEL site during the winters of 2008 and 2009 [Sproles *et al.* 2013]. It is notable that the latter measurements show near-perfect agreement with simulated SWE with the bias-corrected data but were underestimated by the uncorrected L13 data (Figure 4.16). Further, automated measurements on these dates at the Hogg Pass SNOTEL station recorded much less SWE than both the model and the manual field measurements. This is consistent with the mean model performance at Hogg Pass SNOTEL, where SWE was overestimated by the bias-corrected data and underestimated by the original data. These direct field measurements were very limited in number, but provided the most reliable measurements. While it is impossible to draw any definite conclusions, these results suggest that the automated SNOTEL measurements at Hogg Pass may under-represent the quantity of SWE in the surrounding terrain where the field measurements were taken. Locational accuracy is a significant limitation in grid-based model experiments, where a single grid cell represents a large spatial footprint even at the unusually high spatial resolution of this experiment. Thus, it is essential to recognize that grid-cell simulations compared to point-based manual measurements are inherently limited by this scale inconsistency. Along these lines, simulations compared to point-based measurements at Snowcourse locations showed considerable scatter between datasets, although performance was better with the bias-corrected data. At Racing Creek Snowcourse, the bias-corrected data tended to overestimate SWE relative to observations, whereas at Bald Peter Snowcourse SWE was slightly underestimated. At all field measurement locations, the uncorrected data underestimated measured SWE.

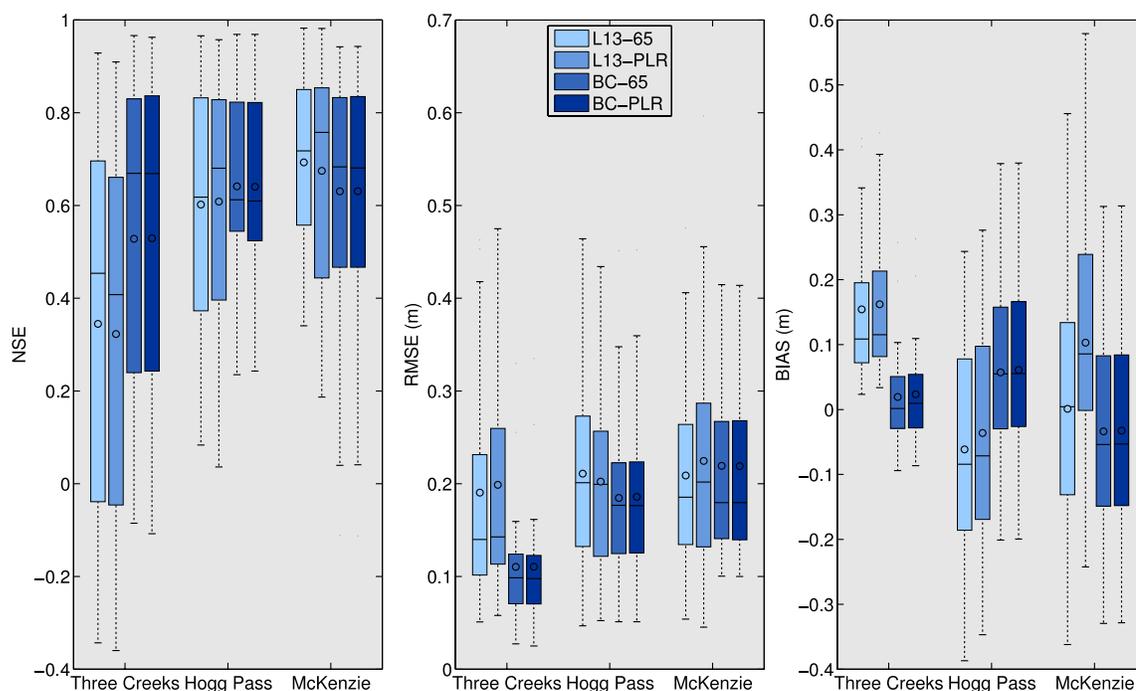


Figure 4.17: Interannual model performance as evaluated with the traditional NSE, RMSE, and ME metrics of model evaluation. The bias-corrected data is more consistent between sites, with the lowest RMSE and most neutral ME.

In addition to interdataset differences based on objective evaluation metrics compared at the measurement locations, pre- and post-calibration differences between datasets compared across the full elevation range of the study region were significant (Figure 4.18). Chief among these was the strong convergence of the estimated longwave irradiance with each dataset post-calibration, which contrasted sharply with the strong divergence of this estimated flux with the uncalibrated model (Figure 4.10 vs. Figure 4.18). This result is consistent with the hypothesis that longwave irradiance controlled interdataset differences pre-calibration. Simulated SWE converged as well, although at elevations between 1700 and 2300 m the bias-corrected data simulated less cumulative SWE. This trend was reversed at elevations above 2300 m, where it is likely that the increase in precipitation in the bias-corrected data resulted in more accumulated SWE where temperatures were well below freezing during winter. Similar to the uncalibrated model, SWE-melt was greater with the warmer bias-corrected data than the uncorrected data at all elevations, and snowfall was greater with the colder uncorrected data at all but the highest elevations where the bias-corrected precipitation was greater as well. In general, the largest post-calibration impact on mean modeled states and fluxes across

elevation gradients was the convergence of the longwave irradiance estimates between datasets.

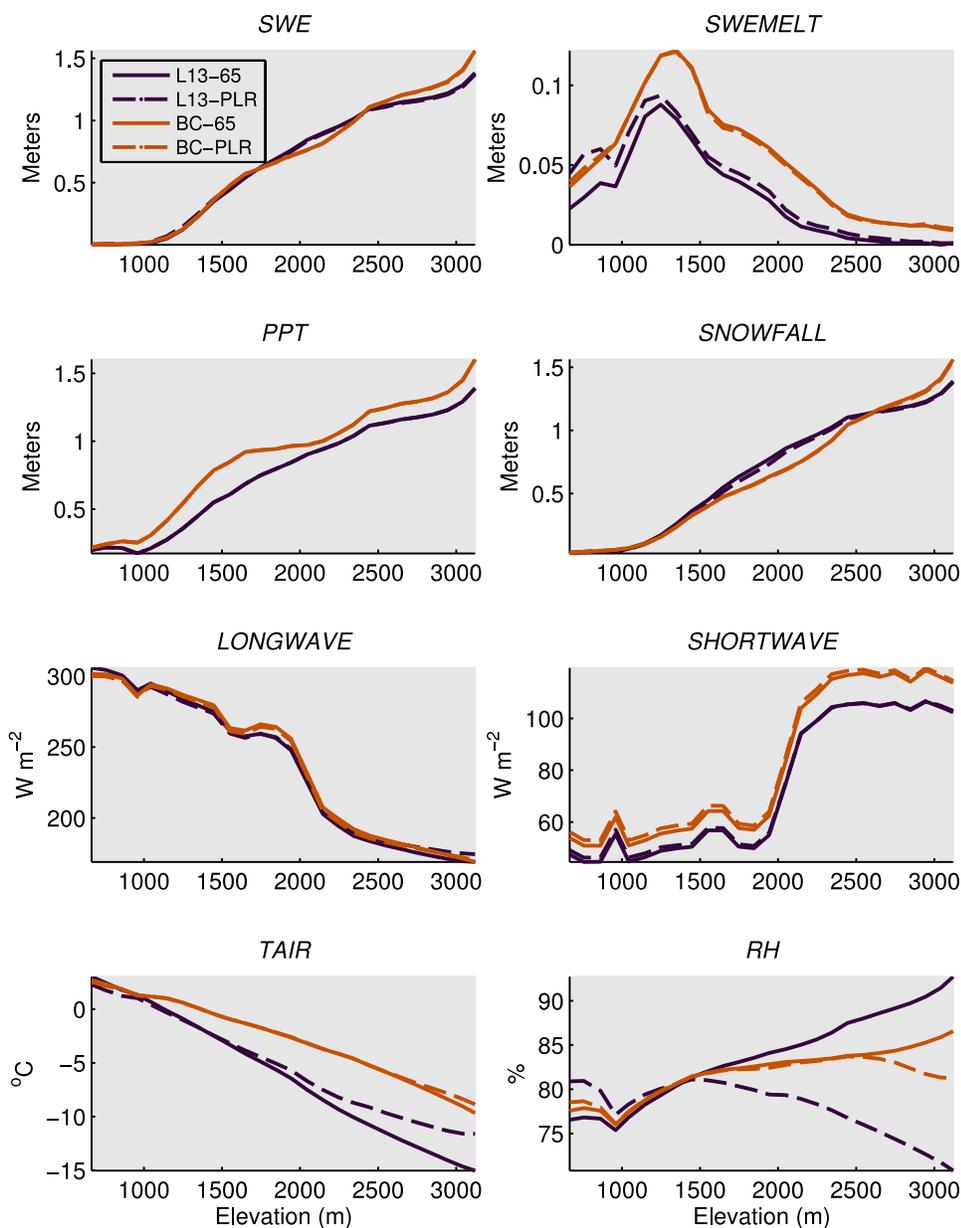


Figure 4.18: Mean values of simulated meteorological variables vs. elevation with the calibrated model. Hydrologic fluxes are cumulative sums on April 1; irradiance, temperature, and humidity are daily averages for the Dec. 1 – Apr. 1 period, binned and averaged into 100 m elevation bands across the 1989-2011 period.

In general, mean model performance during the reference period considered across all sites was better and more consistent with the bias-corrected data (Table 4.3). Further, spatial SWE variability appeared more realistic with the bias-corrected data (Figure

4.19). An increase of spatial gradients due to the $6.5^{\circ}\text{C km}^{-1}$ lapse rate was evident within datasets, and an increase of spatial detail due to bias correction was evident between datasets. This increase in spatial detail with the bias-corrected forcing data was the result of the increased spatial resolution of the PRISM climate data used for bias-correction. These results demonstrate that PRISM bias-correction improved the meteorological forcing data and resulted in better estimates of temperature, precipitation, and simulations of SWE, but model calibration was essential to reduce model parameter bias. As a whole, these results demonstrate that significant differences in modeled estimates of irradiance, snowfall, snowmelt, and SWE resulted from the use of different forcing datasets. These differences derived from complex interactions between the model boundary conditions (elevation), spatial scale (coarse vs. fine), process representations (cloud cover), model parameters (cloud factor), and characteristics of the input data (bias). In the following section, we compare how the differences in meteorological forcing data and modeled SWE during the historical climate period lead to differences in estimates of SWE sensitivity to changes in climate between datasets. We then present our best estimate of climate change impacts on snow hydrology in the study region, using the bias-corrected input data downscaled with the PRISM lapse rates, and with the optimal parameters estimated for that dataset.

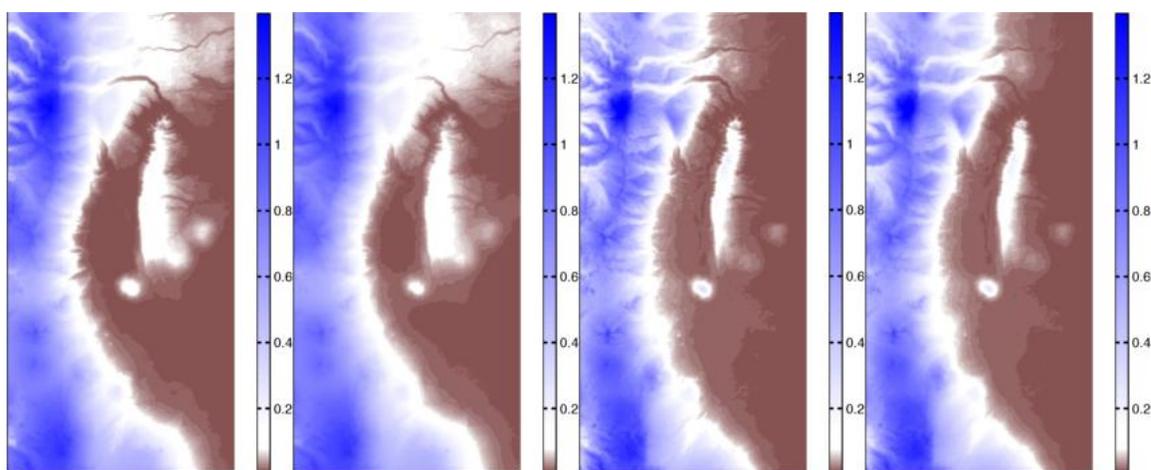


Figure 4.19: Maps of peak SWE with each of the four downscaled forcing datasets. From left to right: L13-6.5, L13-PLR, BC-6.5, and BC-PLR. Peak SWE is the date on which basin-wide volumetric SWE is greatest. Shown here are maps of the mean peak SWE for the 1989-2009 period.

4.7. Climate change impacts on snow in the study region

The first portion of this section describes the interdataset differences in estimated SWE sensitivity to warmer temperature with and without increased precipitation. The comparison is provided as an illustration of the impact that differences in forcing data can have on estimates of climate warming impacts derived from model-based assessments. However, we don't present an intercomparison of equally plausible hypotheses. Our best estimate for climate change impacts in the region is provided in the second portion of this section, which describes the estimated impacts with the bias-corrected forcing dataset.

4.7.1. Interdataset comparison of climate change impacts

In the previous sections, it was shown that estimating parameters independently with each dataset removed the interdataset differences in incoming longwave irradiance that lead to large interdataset differences in modeled SWE and SWE-melt prior to calibration. Thus, during the historical reference period, the post-calibration interdataset differences were largely due to the surface temperature and precipitation data differences. Under the conditions of $+2^{\circ}\text{C}$ and $+4^{\circ}\text{C}$ warming, these differences in surface temperature resulted in very large differences in the estimate of climate change impacts on SWE in the study region (Figure 4.20).

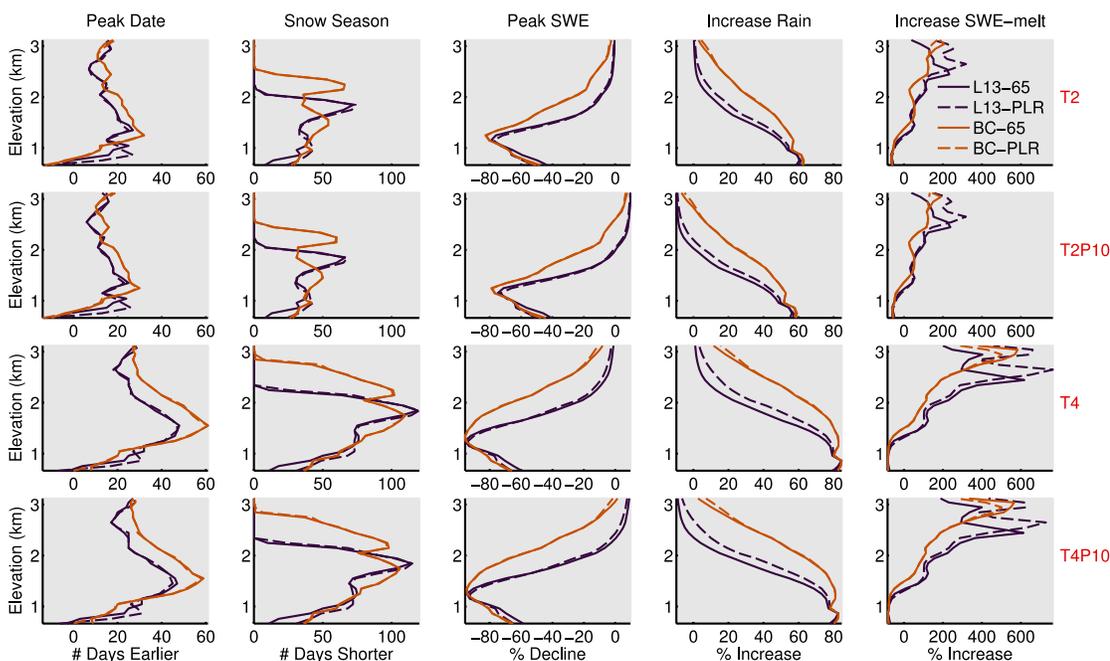


Figure 4.20: Climate change impacts on mean SWE in the study region for the $+2^{\circ}\text{C}$ (top row), $+2^{\circ}\text{C}$ with $+10\%$ precipitation (second from top row), $+4^{\circ}\text{C}$ (third from top), and $+4^{\circ}\text{C}$ with $+10\%$ precipitation (bottom row) scenarios, simulated with each forcing dataset.

Several metrics were used to evaluate the sensitivity of the modeled SWE to changes in climate. All metrics were computed as differences between the reference period mean simulated quantity and the climate change scenario mean simulated quantity. Previous studies have focused on changes in basin-wide April 1 SWE because of its importance for water managers as an indication of spring and summer streamflow [e.g. *Mote, 2006*]. Here, we focused on changes in the timing and amount of modeled peak SWE, which was defined as the date on which the basin-wide total volumetric SWE was greatest. This approach was used to understand how the peak amount of SWE in the basin responded to changes in climate regardless of a specific day of year such as April 1 that may be a biased estimate of peak SWE [*Montoya et al., 2014*]. As such, the change in the date of peak SWE was computed as the difference between the date of basin-wide volumetric peak SWE during the reference period and the date of basin-wide volumetric peak SWE during the climate change scenario period. The length of the snow season was computed as the number of days between the first and last day in which SWE was simulated to occur in a grid cell. All of the metrics were computed for each grid cell in the model domain and then binned into 100 m increasing elevation increments.

Climate change impacts on SWE were estimated much more conservatively with the uncorrected forcing data than with the bias-corrected data. These differences were much greater for the hydrologic quantities of percent change in peak SWE and percent increase in rainfall than for the change of timing metrics. However, the differences in estimates were elevation-dependent. The estimated sensitivity was roughly equivalent between datasets below approximately 1200 m with a 2°C warming. Above this elevation, the effect of estimating impacts with the uncorrected forcing data was approximately equivalent to shifting the elevation of the estimated impacts with the bias-corrected data upward by 200-300 m. This difference is very simple to explain. The sensitivity of SWE in the study region is largely controlled by temperature-driven shifts from snowfall to rainfall that decrease snow accumulation. Below 1200 m, estimates of mean winter air temperature with both the uncorrected and bias-corrected data were above 0°C (Figure 4.18). Above 1200 m, estimates of air temperature between datasets diverged and the uncorrected temperature data were much colder. Similarly, within the uncorrected data,

the data downscaled with a $6.5^{\circ}\text{C km}^{-1}$ lapse rate was colder above 1500 m. These differences in air temperature are reflected in the estimates of sensitivity between datasets (Figure 4.20), which diverge between datasets at roughly the same elevations where air temperature diverges.

While the differences in estimated sensitivity were largely explained by the differences in surface air temperature between datasets, there were also differences in the simulated increase in longwave irradiance between datasets that lead to slightly different estimates of sensitivity across the elevation range in the study region. Figure 4.21 shows in greater detail the four key variables that differed between datasets. At all elevations above ~1200 m, the pattern of SWE sensitivity roughly corresponded to snowfall sensitivity, with the largest declines in the 1500-m elevation range. At high elevation regions in the study region, though, increased mid-winter SWE-melt corresponded to increased longwave irradiance. However, as in the reference period climate, the relative impact of the increased longwave was not the same between datasets, and the colder L13 surface temperature dataset estimated smaller increases in SWE-melt with corresponding increases in incoming longwave than the warmer BC surface temperature dataset.

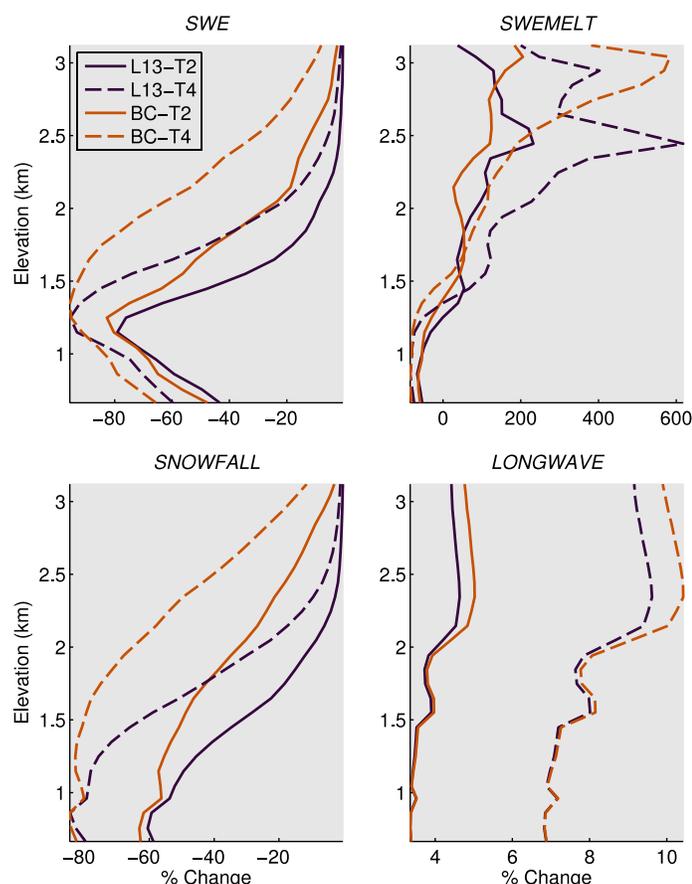


Figure 4.21: The percent change of mean SWE, SWE-melt, snowfall, and incoming longwave irradiance with the L13-PLR and BC-PLR datasets, for the +2°C and +4°C warming scenarios.

While the percent differences in SWE-melt between datasets were large and the percent increase in mid-winter SWE-melt was also very large in the high elevation regions of the watershed, the absolute increase in SWE-melt was very small. This is reflected in the very small sensitivity of SWE in the same elevation range as the increased SWE-melt. Thus, post-calibration, very large interdataset differences in surface temperature and lapse rates between and within datasets led to small differences in estimates of incoming longwave, and the impact of these differences on estimated climate impacts was negligible. When the reference period pre-and-post calibration period differences in SWE, SWE-melt, and longwave irradiance are considered, these results confirm the importance of the parameter estimation routine and its ability to remove the impact of the interdataset differences in temperature and lapse rates on estimated longwave irradiance. Thus, comparisons between datasets under the climate warming scenarios are explained by the interdataset surface temperature differences. Accordingly, the percent decline in

SWE under a 2°C warming with the bias-corrected data was very similar to the percent decline in SWE under a 4°C warming with the uncorrected data, suggesting that interdataset differences due to temperature were almost as large as the impact of climate change itself.

This comparison was included as an illustration of the impact that the choice of forcing data, and the methods used to downscale the forcing data, had on estimates of sensitivity, and was not meant to be a comparison of equally plausible hypotheses. In the previous sections of this paper, we showed that via calibration, each dataset was able to model SWE in the study region with high accuracy. In particular, calibration removed the effect of the interdataset differences in estimated longwave irradiance that explained interdataset differences in model performance prior to calibration. Post-calibration, model performance with the uncorrected forcing data was good, but the model was calibrated to the observed SWE at the expense of accurately estimating air temperature and snowfall vs. rainfall. By all accounts, the bias-corrected data downscaled with PRISM lapse rates was the most consistent dataset and is the dataset we consider to be the best estimate of the spatial distribution of temperature, precipitation, and modeled SWE in the study region. In the following section, we focus on the modeled climate change impacts to SWE in the study region using this dataset alone.

4.7.2. A best estimate of climate change impacts to snow in the study region

In the Upper Deschutes River Basin, the mountain snowpack is highly sensitive to a 2°C increase in temperature. Our results indicate that the relative peak SWE sensitivity (percent decline from a 2°C increase in temperature relative to historic conditions) was greatest at elevations between 1000 – 2000 m (Figure 4.22). Mean annual SWE accumulation declined by 80% in the rain-snow transition elevation of 1150 m, but sensitivity decreased to less than 10% above 2050 m. The sharp decline in sensitivity above 2000 m corresponded to the elevation at which mean winter air temperature in a 2°C warmer climate remained below -1.0°C (Figure 4.23).

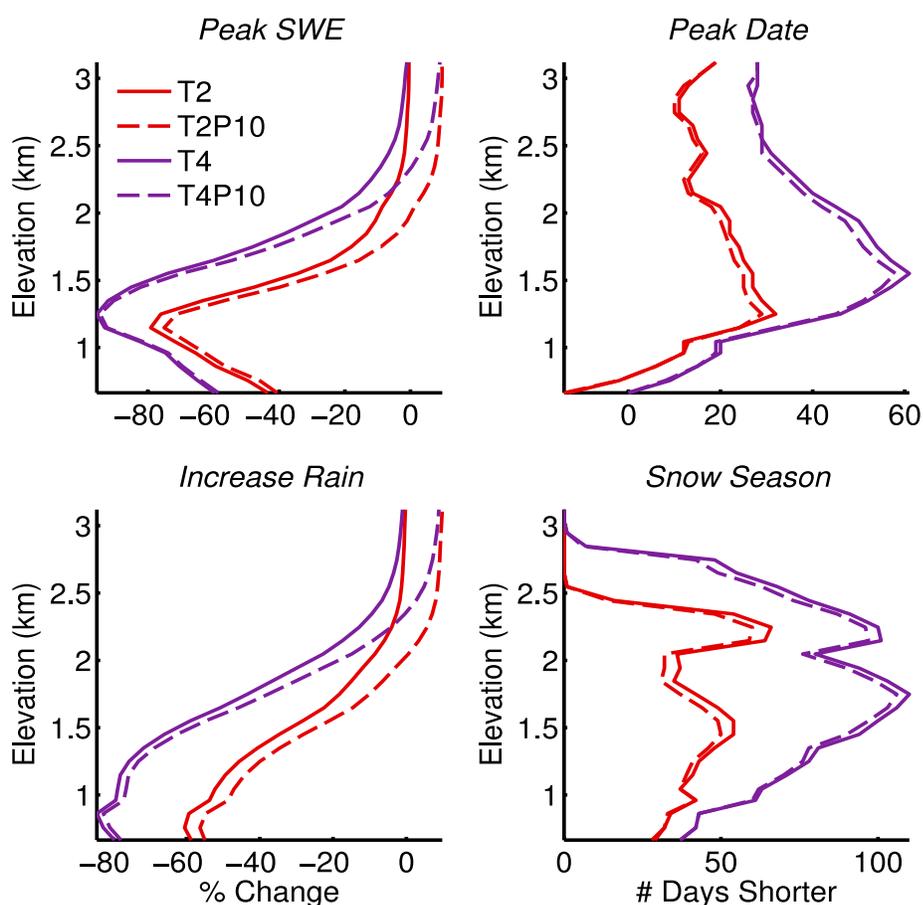


Figure 4.22: Climate change impacts on mean SWE in the study region, estimated with the bias-corrected data. Shown are the percent change in peak SWE and rainfall, and the number of day's difference for the date of peak SWE and the length of the snow-covered season.

In this analysis, elevation was essentially a proxy for winter temperature. From a management perspective, it is of interest to know where on the landscape the largest declines in SWE may occur, but in fact the declines largely depend on mean winter air temperature (Figure 4.23). The dependence of SWE sensitivity on mean winter air temperature was very strong, and was driven by the corresponding shift from snowfall to rainfall. Across most of the elevation range, the sensitivity of SWE corresponded very closely with the sensitivity of snowfall (the percent increase in rainfall). In the rain-snow transition elevations below 1000 m, the sensitivity of snowfall was higher, but at the mid to high elevations where the majority of SWE accumulated, SWE sensitivity was higher than snowfall sensitivity. This indicates that SWE sensitivity was driven primarily by shifts from snowfall to rainfall but also by increased ablation in this elevation range.

Above ~1750 m, the sensitivity of SWE and snowfall were essentially identical. (Figure 4.22)

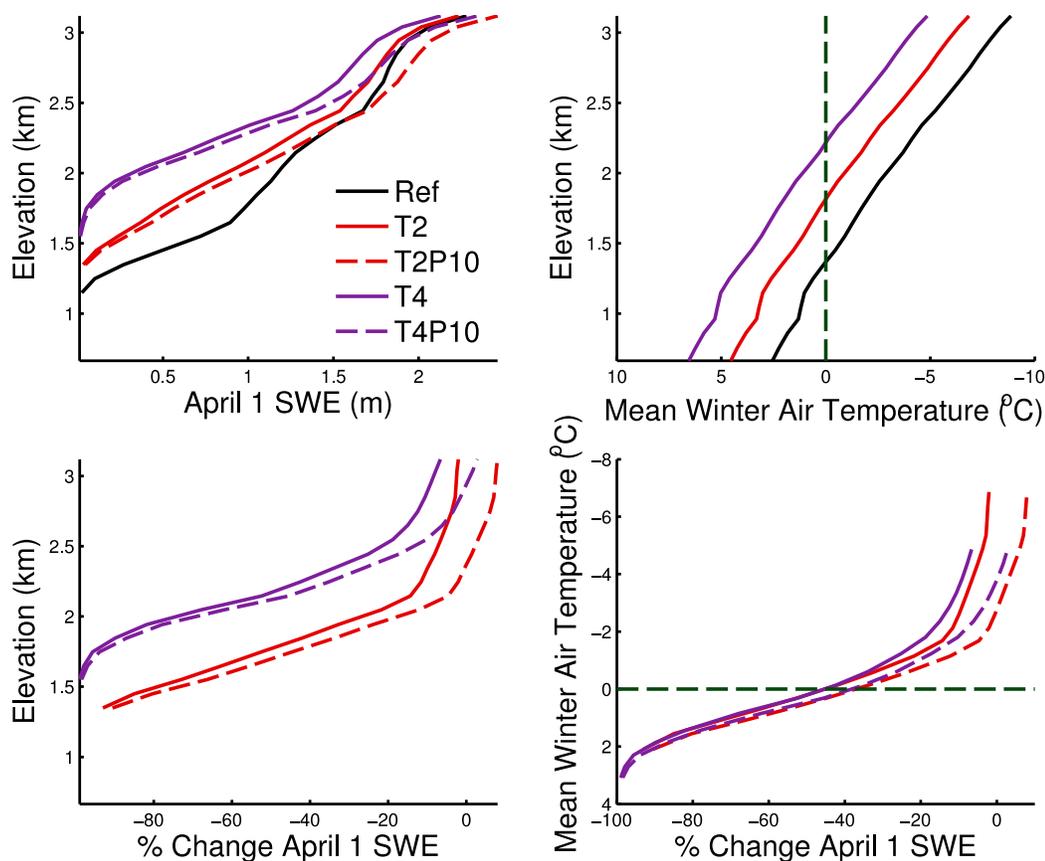


Figure 4.23: April 1 SWE vs. elevation (upper left), mean winter air temperature vs. elevation (upper right), the percent change in April 1 SWE vs. elevation (lower left), and the percent change in April 1 SWE vs. mean winter air temperature (lower right).

Basin-wide, the percent change in peak SWE with a 2°C increase in temperature was estimated to be -38%, and with a 4°C increase in temperature was -51%. Increases in precipitation made very little difference when sensitivity was integrated across the entire basin (Table 4.4). Volumetrically, these declines equate to 0.26 km³ on the date of peak SWE with a 2°C increase in temperature (Figure 4.25), and 0.44 km³ with a 4°C increase. These absolute declines in basin wide volumetric peak SWE are equivalent to 40% and 66% of the holding capacity of Green Peter Dam, the largest reservoir in the region, respectively.

With a 4°C increase in temperature, sensitivity was greater across all elevations but the increase in sensitivity from a 2°C to a 4°C warming was smaller than the initial 2°C

warming. Peak sensitivity shifted upward to 1250 m where 96% of the seasonal snowpack was lost with a 4°C warming, but as with a 2°C warming, the sensitivity steadily declined above that elevation. Between ~1200 m and 2000 m, the effect of a 4°C warming was essentially equivalent to shifting the effect of a 2°C warming upward by 300 m. In the context of lapse rates, 300 m corresponded to a decline of roughly 2°C in mean annual air temperature when the annual lapse rate in our region was applied.

The effect of increased precipitation on sensitivity was significant when considered at specific elevations in the region. With a 2°C warming in combination with a 10% increase in precipitation, the magnitude of SWE sensitivity decreased in a roughly linear fashion by approximately 3-10% between the lowest elevations and 2000 m. Above 2000 m, the offsetting effect of increased precipitation leveled off. Under a 2°C warming scenario, total winter SWE accumulation actually increased above 2050 m, and at the highest elevations above 3000 m in the basin, total winter SWE increased by 10% with a 2°C warming. With a 4°C warming, the offsetting effect of precipitation shifted upward by ~300 m and smaller percent declines were estimated across all elevations. However, when integrated across the entire basin, the offsetting effect was minimal (Table 4.4) because the majority of the snowpack accumulated in the 1700 – 2000 m elevation range where the sensitivity of SWE was large and the offsetting effect of precipitation was small. With a 4°C warming and 10% increase in precipitation, the estimate of basin-wide percent change in SWE was essentially equivalent to the 4°C warming alone. Volumetrically, the difference was 0.0125 km³.

Table 4.4 Estimated basin-mean climate change impacts on SWE in the study region

Scenario	Peak SWE Difference		Time of SWE difference		
	Absolute (km ³)	Percent (%)	Date of Peak SWE	Date of Snow Disappear.	Length of Snow Season
T2	0.26	-38	-15	-24	-31
T2P10	0.23	-36	-15	-22	-29
T4	0.44	-51	-36	-51	-66
T4P10	0.42	-50	-36	-49	-63

In addition to shifts in volumetric SWE across elevations, significant shifts in the timing of basin-wide SWE and modeled timeseries of SWE at the SNOTEL monitoring

locations were estimated for a 2°C and 4°C warming (Table 4.4, Figure 4.24). The date of basin-wide peak SWE shifted earlier by 15 days with a 2°C warming, and by 36 days earlier with a 4°C warming. Increased precipitation had no impact on shifts in the date of peak SWE, and when integrated across the basin, a 10% increase in precipitation had negligible impact on the amount of peak SWE. The estimated date of peak SWE in the basin for the historical period was March 28, which for practical purposes is equivalent to April 1, the date used by the NRCS to estimate peak SWE for streamflow forecasting purposes. Thus, our estimates of differences in the amount and timing of peak SWE can roughly be considered estimates of the differences for April 1. With a 2°C warming, the date of peak SWE shifted to March 13, and with a 4°C warming the date shifted to February 20. These shifts have significant implications for water management in this region, especially when considered in tandem with the shifts toward earlier date of snow melt (the first date after the date of peak SWE that the estimated SWE is zero). In our study region, we estimated that the date of snowmelt occurred 24 days earlier with a 2°C increase, and up to 51 days earlier with a 4°C increase. Again, increased precipitation had a negligible impact on these shifts in timing. Further, the general characteristics of these impacts were very similar at each SNOTEL station in the region.

Finally, we consider shifts in the length of the snow-covered season. The shift in the snow covered season is directly related to the shift in timing of snow melt and is important hydrologically, but also serves as an indicator of the ecological impact of the estimated changes in SWE. Snow serves as an important moisture subsidy for ecosystems, and the elevational distribution of snow has been correlated with forest productivity [Trujillo *et al.*, 2012]. As seasonal snow cover shifts higher in elevation, it is possible that this will create increased habitat for species adapted to warmer environments and impact forest productivity. Snow insulates the ground during winter, which significantly affects biogeochemical cycles in seasonally snow-covered environments [Brooks and Williams, 1999]. Finally, earlier snowmelt and shorter snow season duration result in earlier soil moisture wetting and, across broad scales, there is evidence that declines in snowpack across the West are increasing the frequency of forest fires [Westerling *et al.*, 2006]. With just a 2°C warming, the length of the snow covered

season decreased by >50 days in the forest-alpine transition zone between 2000 and 2500 m. With a 4°C warming the snow covered season was >100 days shorter and the elevation of the greatest declines shifted downward and expanded to the 1500 – 2500 m range. Again, increased precipitation had a negligible effect on these impacts.

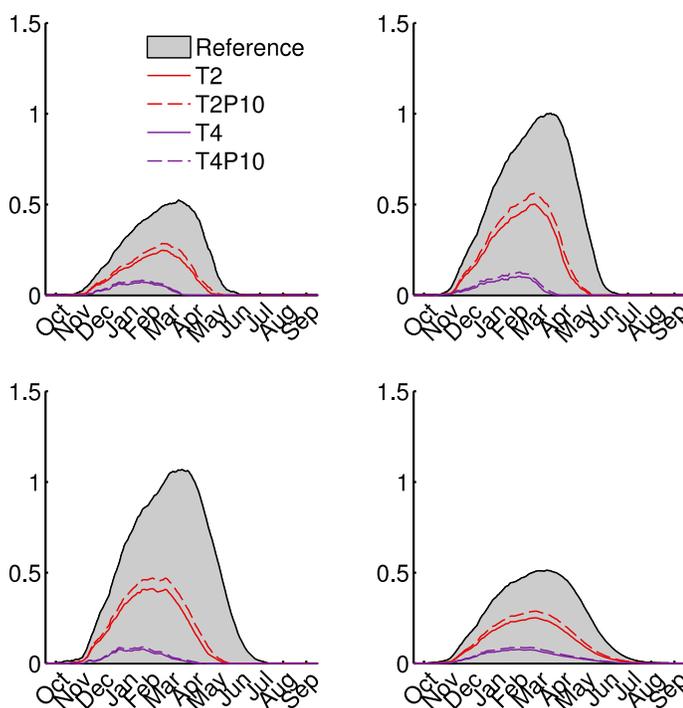


Figure 4.24: Simulated daily SWE at each SNOTEL location. Top left: Three Creeks SNOTEL; top right: Hogg Pass SNOTEL; bottom left: McKenzie SNOTEL; and bottom right: basin-integrated.

These results confirm that the sensitivity of SWE accumulation in the Oregon Cascades is largely driven by shifts from snowfall to rainfall during the accumulation period, and that the sensitivity is largest between 1200 m and 2000 m where mean winter air temperature is close to 0°C. However, above ~2500 m, SWE appears insensitive to climate warming, and SWE actually increased in this elevation range with increased precipitation. Thus, precipitation variability controlled SWE sensitivity above the freezing level. Nonetheless, when basin-integrated, a 10% increase in precipitation had a negligible impact on changes in SWE accumulation with a 2°C and 4°C warming. Climate warming shifted peak SWE accumulation earlier into winter, shortened the length of the snow-covered season, and decreased the basin-wide volumetric SWE by up to 38% with just a 2°C warming. Our best estimate of climate change impacts to snow in

this region is an indicator that warming temperatures will continue to drive the gradual decline of the mountain snowpack in the Oregon Cascades, and that this decline will have significant implications for water resources and ecosystem management in this region.

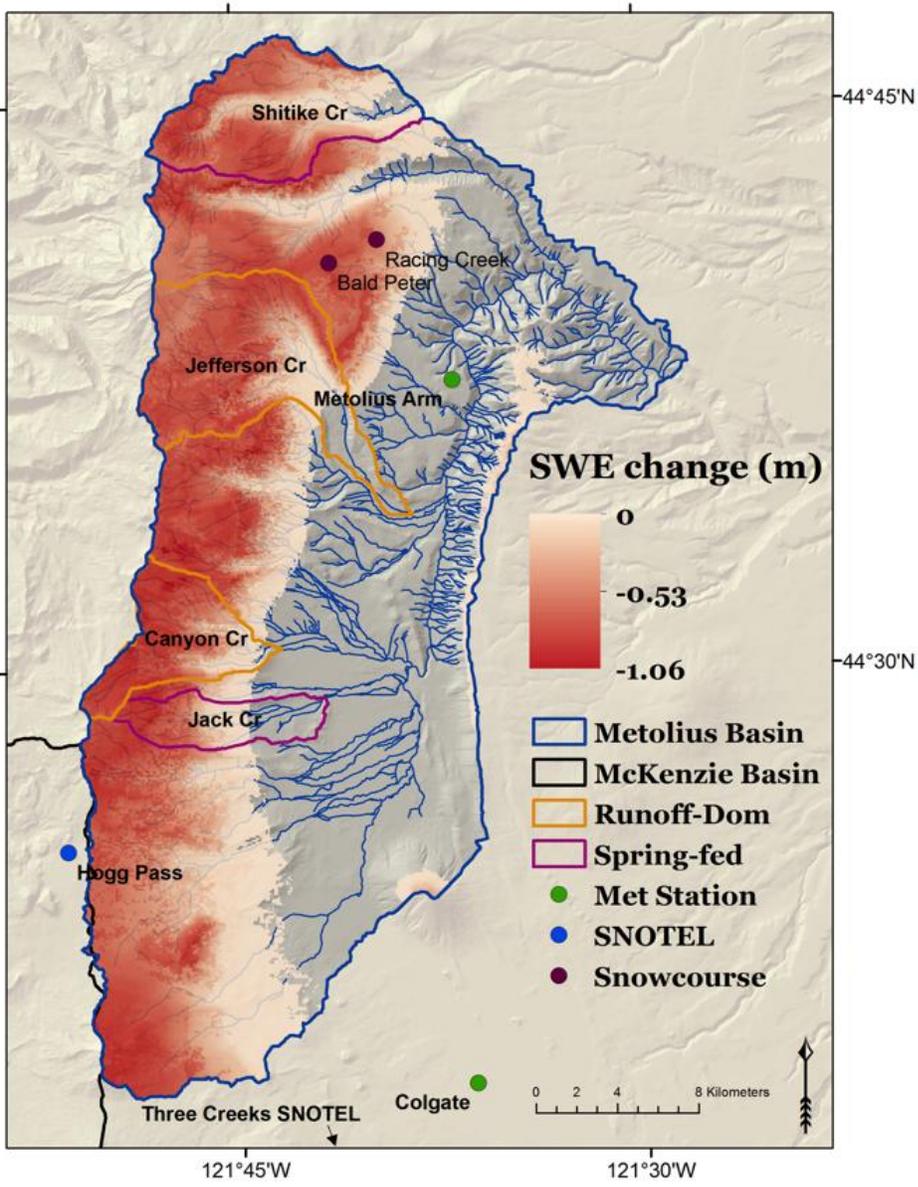


Figure 4.25: Absolute change in peak SWE in the Upper Deschutes River Basin with a 2°C warming.

5. Summary

5.1. Summary of research findings

When evaluated against observations from weather stations in the study region, the widely used L13 gridded meteorological forcing dataset was found to have consistent and significant low bias in the gridded temperature and precipitation, largest so in winter. In the case of temperature, the bias was found to correlate linearly with elevation. To diagnose this bias, lapse rates computed from observations were compared to those computed from PRISM data covering the same spatial extent and time period. The seasonal temperature bias in the gridded data was explained by the use of a temporally constant $6.5^{\circ}\text{C km}^{-1}$ lapse rate that overestimated the decline of temperature with increasing elevation in winter compared to measured temperature in the study region. The implication of this finding was that high elevation regions of the study region that lacked direct temperature measurements were likely estimated to be much colder by the gridded data than in reality. It was hypothesized that modeled snow sensitivity to climate warming derived from this data would be estimated more conservatively than when bias-corrected with PRISM climate data, which was confirmed.

The lapse rates computed from observations agreed well with conclusions presented by previous studies in the Pacific Northwest [e.g. *Blandford et al.*, 2008; *Minder et al.*, 2010]. In particular, lapse rates in the study region were shallower than the $6.5^{\circ}\text{C km}^{-1}$ lapse rate, especially so in winter, and reached a maximum in late summer. Several limitations to the lapse rate computations should be noted. First, there were very few stations available and these stations likely did not capture the full variability of topographic and land cover variability in the study region. Second, we included three stations (Redmond, Madras, and Pelton COOP) in our calculations that are located east of the modeling domain on the broad high-plateau landscape that is characteristic of eastern Oregon. Mixing these stations with the montane stations in the modeling domain may have biased our calculation of the environmental lapse rate in the study watershed. Further, the poor linear fit for minimum temperature lapse rates (and all lapse rates during winter) suggests that other factors controlled the distribution of cold temperatures in the study region. The linearity assumption fails to account for competing processes that control temperature in complex topography, ranging from purely advective mixing of air

along the steep slopes of the Cascades to radiative cooling of pooled air within convex valleys. While beyond the scope of this investigation, additional research is needed to develop efficient modeling tools for non-linear spatial temperature estimation [e.g. *Lundquist et al.*, 2008a], as well as fundamental process understanding [*Pepin et al.*, 1999, 2011].

PRISM climate data was used to correct the bias. The bias correction technique significantly improved the performance of the distributed snowpack energy balance model relative to forcing the model with the uncorrected data, after model parameters were estimated independently with each dataset. The parameter estimation experiment also revealed a key source of model sensitivity. For a given input temperature dataset, the model was highly sensitive to the choice of temperature lapse rates used to distribute the temperature across the elevation boundary conditions. However, the sensitivity was due to a complex interaction between the model-estimated cloud fraction and the prescribed lapse rates. Counterintuitively, colder lapse rates resulted in less simulated SWE, because the colder lapse rates simulated higher relative humidity values aloft which triggered the model-computed increase in longwave irradiance due to estimated cloud fraction. This key model sensitivity was previously unrecognized.

Further, the model parameter estimation experiment successfully explained interdataset differences in simulated SWE, and also resulted in improved model performance. While the bias-corrected data produced a higher goodness of fit overall, there appeared to be enough flexibility in the model calibration routine to adequately model the basin snowpack with each dataset, regardless of bias. However, despite the appearance of improved model behavior post-calibration, the partitioning of solid and liquid precipitation was significantly different between datasets. This implied that snow was simulated with high accuracy at the expense of rainfall and/or water available for runoff. The implication of this for streamflow modeling applications is a logical next step for research in this area.

Here it should be noted that the observational data necessary to truly evaluate the performance of the model and diagnose the sources of model error were not available in the study region. Modeling in data sparse regions will continue to be a challenge for the

environmental modeling community. We present our analysis here as a case study for some of the issues that arise when trying to utilize gridded forcing datasets to model land surface hydrology in data sparse regions.

Finally, we found that SWE in the study region is highly sensitive to increased temperature. With just a 2°C warming, the timing of peak basin-integrated SWE shifted 15 days earlier, peak basin-integrated SWE declined by 38%, and the date of snow melt and the length of the snow season decreased by 24 and 31 days, respectively. With a 4°C warming, the change in timing metrics were roughly doubled and peak SWE declined by 50%. The results confirmed the strong temperature dependence of snow in the Oregon Cascades, which in turn is dependent on elevation. In the study region, the most dramatic declines in SWE were estimated for elevations between 1200 and 2000 m.

The results reported here are limited by several important factors including the small geographic coverage, the use of one single hydrologic model [*Vano et al.*, 2012], and a single model calibration scheme [*Clark et al.*, 2011]. Chief among the limitations to our approach was the lack of high quality measurements of incoming and outgoing radiative fluxes necessary to quantify the energy balance of the simulated snowpack, which is necessary to comprehensively diagnose differences in ablation between datasets and how those differences relate to the bias correction and downscaling procedures. The calibration (parameter estimation) experiment provided some insight into the sensitivity of the model to forcing data and how that sensitivity depended on the modeled radiative fluxes, but again, the lack of measured radiative fluxes prevented an evaluation of the realism of the calibrated fluxes. Further, the calibration scheme was relatively simplistic. In particular, we did not evaluate the sensitivity of the model calibration scheme to input data climate. Given that the calibrated model was used to estimate the sensitivity of the snowpack to changes in climate, it would be a logical next step to test the sensitivity of the model to differences in climate. For example, one could calibrate the model during cold years and warm years separately to test how optimal parameter estimates depend on the “climate” of the forcing data.

5.2. Summary of research implications

The results of this study highlight the need to expand and improve upon existing ground based climate and hydrologic monitoring networks. Chief among these needs is an expansion of high elevation monitoring locations and fuller characterization of landscape-scale climatic variability, for example valley vs. ridge locations and forested vs. open locations. Land-cover type in the study region was estimated to be 69% coniferous forest, yet there were no long-term records of forest meteorology within the study region. At 1702 m, Three Creeks SNOTEL is the highest elevation snowpack monitoring location in the study region, yet 42% of the historical peak April 1 SWE accumulated above this station. With a 2°C and 4°C warming, the snowline shifted upward such that 54% and 72% of peak SWE (which shifted to March 14 and Feb 20), and 61% and 91% of April 1 SWE accumulated above this monitoring location, indicating that the vast majority of SWE will likely accumulate where there are no historical measurements. Further, with a 2°C and 4°C warming, this location was estimated to lose 53% and 87% of peak SWE, and 63% and 99% of April 1 SWE, respectively. Lower elevation monitoring locations in the region were even more vulnerable to warming. With a 2°C and 4°C warming, April 1 SWE at the Racing Creek (1573 m) Snowcourse declined by 80% and 100%, effectively rendering this measurement location obsolete.

While a 4°C warming may seem a distant future, the record warm spell of 2014-2015 may provide a glimpse of future snowpack conditions in the Oregon Cascades. At the Three Creeks SNOTEL station, 2014 and 2015 were 2.4°C and 4.4°C warmer than the study period mean temperature, while April 1 precipitation was 100% and 105% of normal, respectively. These unique meteorological conditions provide analogs for the +2°C and +4°C warming scenarios presented in this study. Comparison between observed SWE during this record warm spell and the mean modeled response of SWE to a +2°C and +4°C warming at this location strongly suggests that our model results accurately capture the sensitivity of SWE to increased temperature – and provide a cautionary glimpse of future snowpack conditions at our current monitoring locations (Figure 5.1).

However, it is important to note that the long-term regional warming trend has, and will, continue to be superimposed upon regional variability that causes the PNW to

experience oscillating warm and cold winters, therefore a comparison of this sort is complicated by mean vs. variable response. Nonetheless, the frequency and magnitude of warm winters may increase in concert with the long term warming trend [Diffenbaugh *et al.*, 2015], and the warming trend may already be exacerbating natural variability such that when warmer winters coincide with dry winters, conditions such as those experienced in the Western U.S. during 2014 and 2015 become more common [Mao *et al.*, 2015].

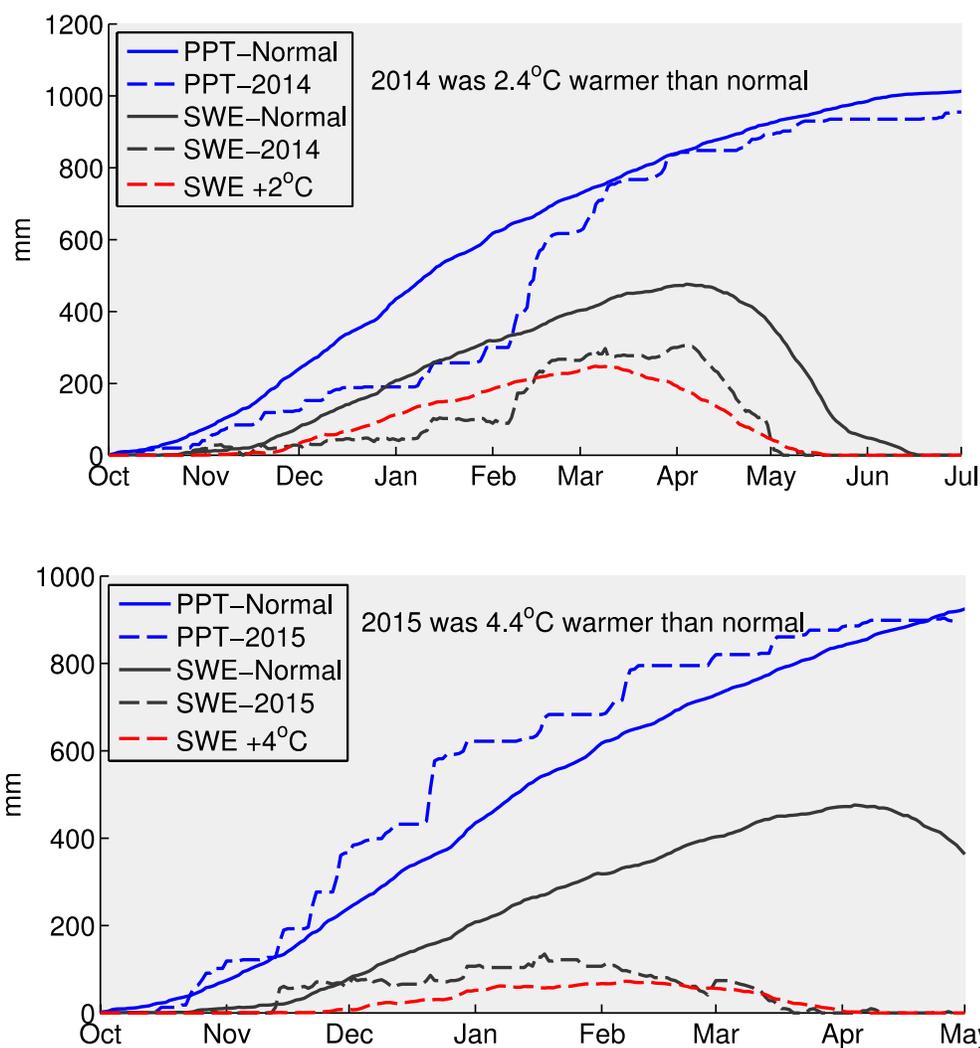


Figure 5.1: Mean measured precipitation (solid blue) and SWE (solid black) during the period 1989-2011, and for water year 2014 (top, dashed) and 2015 (bottom, dashed), compared to mean modeled SWE for a 2°C (top, dashed red) and 4°C (bottom, dashed red) warmer climate

The ability to characterize these losses at any given location or scale on the landscape is one of the key strengths of spatially distributed models such as SnowModel. Via the micrometeorological submodel MicroMet, the model forcing data was allowed to vary in space and time. Coupled with the spatially varying land cover and elevation boundary conditions (which in turn give rise to spatially varying topographic variables such as slope and aspect), the physical realism of the model was greatly increased. The spatially distributed, i.e. gridded, model output is therefore compatible with spatial statistical modeling and Geographic Information System (GIS) software. This compatibility permits the model output to be incorporated into data-driven management decisions. For example, maps of SWE sensitivity can be incorporated into a GIS system and specific locations on the landscape that currently lack direct measurements, but are estimated to be highly sensitive to change, can be targeted for future measurements. Conversely, if the goal is to maintain continuous information about snow conditions in a warmer climate, locations that are relatively insensitive to climate warming can be identified and targeted for future station installations. Perhaps more exciting is the possibility of coupling the gridded model output with spatial statistical models that determine the optimal sampling locations for capturing the underlying landscape variability that controls hydrologic processes, and using these to guide “intelligent” future monitoring networks.

However, while the modeling methodology used for this study provided a robust estimate of snow sensitivity to climate warming as a function of location, the ability to confidently predict the effect of climate change on snow and related resources at any given location and point in time at the scale of management decisions is still limited by the lack of sufficient ground-based data. For example, accurate results required bias-correction with PRISM data, which is entirely dependent on ground-based measurements. The PRISM model was able to improve our methods, but independent validation of the meteorological forcing data was limited by the lack of independent weather station data in the study region. Moreover, the full scope of measurements needed to critically evaluate process-based (e.g. energy balance) models is categorically unavailable in all but the most well instrumented research basins. At a minimum, monitoring networks should

expand to include routine net radiation measurements so that model error can be confidently evaluated.

In summary, a robust and integrated effort to design and implement ground-based data collection campaigns is ever more necessary to plan for future water needs in a warming climate. Stakeholder engagement that guides the planning and implementation process will be key to the success of these efforts. Models are useful, but only to the extent that they provide information that is useful to data users. Here, the importance of providing visualization tools to stakeholders and managers should be emphasized [Sproles, 2012]. Our results are meant to be useful, illustrative, and diagnostic, but are not meant to be predictive.

6. Conclusion

The need for spatially distributed hydrologic understanding and prediction is growing as climate change, land use change, and disturbance place increasing pressure on our water resources. Spatially distributed hydrologic models can help fill gaps in data sparse regions and provide diagnostic and prognostic skill when properly calibrated and validated. However, these models are highly sensitive to the input temperature and precipitation forcing data. It's critical that we understand how methodological decisions impact model simulations in data sparse regions, especially considering that 1/6th of the world's population relies on snow and ice-melt for their water resources [Barnett *et al.*, 2005], and very few of those regions contain the necessary data for comprehensive model assessment [Bales *et al.*, 2006].

Thus, while it's clear that regional climate warming is driving the gradual decline of the mountain snowpack in the PNW, uncertainty as to how these trends might emerge at the watershed scale persists. A key source of this uncertainty is the lack of long term and spatially representative hydrologic and meteorological data. The results of this research serve as an illustration of the challenges related to distributed hydrologic modeling in data sparse regions, and provide an example of a transferrable methodological approach that can be used to develop model-based assessments of climate change impacts to water resources in data sparse regions.

Bibliography

- Abatzoglou, J. T. (2013), Development of gridded surface meteorological data for ecological applications and modelling, *Int. J. Climatol.*, 33(1), 121–131, doi:10.1002/joc.3413.
- Abatzoglou, J. T., D. E. Rupp, and P. W. Mote (2013), Seasonal Climate Variability and Change in the Pacific Northwest of the United States, *J. Clim.*, 27(5), 2125–2142, doi:10.1175/JCLI-D-13-00218.1.
- Bales, R. C., N. P. Molotch, T. H. Painter, M. D. Dettinger, R. Rice, and J. Dozier (2006), Mountain hydrology of the western United States, *Water Resour. Res.*, 42(8), doi:10.1029/2005WR004387.
- Barnes, S. L. (1964), A Technique for Maximizing Details in Numerical Weather Map Analysis, *J. Appl. Meteorol.*, 3(4), 396–409, doi:10.1175/1520-0450(1964)003<0396:ATFMDI>2.0.CO;2.
- Barnett, T. P., J. C. Adam, and D. P. Lettenmaier (2005), Potential impacts of a warming climate on water availability in snow-dominated regions, *Nature*, 438(7066), 303–309, doi:10.1038/nature04141.
- Barnett, T. P. et al. (2008), Human-Induced Changes in the Hydrology of the Western United States, *Science*, 319(5866), 1080–1083, doi:10.1126/science.1152538.
- Bavay, M., M. Lehning, T. Jonas, and H. Löwe (2009), Simulations of future snow cover and discharge in Alpine headwater catchments, *Hydrol. Process.*, 23(1), 95–108, doi:10.1002/hyp.7195.
- Blandford, T. R., K. S. Humes, B. J. Harshburger, B. C. Moore, V. P. Walden, and H. Ye (2008), Seasonal and Synoptic Variations in Near-Surface Air Temperature Lapse Rates in a Mountainous Basin, *J. Appl. Meteorol. Climatol.*, 47(1), 249–261, doi:10.1175/2007JAMC1565.1.
- Brooks, P. D., and M. W. Williams (1999), Snowpack controls on nitrogen cycling and export in seasonally snow-covered catchments, *Hydrol. Process.*, 13(14-15), 2177–2190, doi:10.1002/(SICI)1099-1085(199910)13:14/15<2177::AID-HYP850>3.0.CO;2-V.
- Christensen, N. S., and D. P. Lettenmaier (2007), A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River Basin, *Hydrol. Earth Syst. Sci. Discuss.*, 11(4), 1417–1434.
- Clark, M. P., D. Kavetski, and F. Fenicia (2011), Pursuing the method of multiple working hypotheses for hydrological modeling, *Water Resour. Res.*, 47(9), W09301, doi:10.1029/2010WR009827.

- Dai, A. (2008), Temperature and pressure dependence of the rain-snow phase transition over land and ocean, *Geophys. Res. Lett.*, 35(12), n/a–n/a, doi:10.1029/2008GL033295.
- Daly, C. (2006), Guidelines for assessing the suitability of spatial climate data sets, *Int. J. Climatol.*, 26(6), 707–721, doi:10.1002/joc.1322.
- Daly, C., J. W. Smith, J. I. Smith, and R. B. McKane (2007), High-Resolution Spatial Modeling of Daily Weather Elements for a Catchment in the Oregon Cascade Mountains, United States, *J. Appl. Meteorol. Climatol.*, 46(10), 1565–1586, doi:10.1175/JAM2548.1.
- Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris (2008), Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, *Int. J. Climatol.*, 28(15), 2031–2064, doi:10.1002/joc.1688.
- Daly, C., D. R. Conklin, and M. H. Unsworth (2010), Local atmospheric decoupling in complex topography alters climate change impacts, *Int. J. Climatol.*, 30(12), 1857–1864, doi:10.1002/joc.2007.
- Diffenbaugh, N. S., D. L. Swain, and D. Touma (2015), Anthropogenic warming has increased drought risk in California, *Proc. Natl. Acad. Sci.*, 112(13), 3931–3936, doi:10.1073/pnas.1422385112.
- Dilley, A. C., and D. M. O'Brien (1998), Estimating downward clear sky long-wave irradiance at the surface from screen temperature and precipitable water, *Q. J. R. Meteorol. Soc.*, 124(549), 1391–1401, doi:10.1002/qj.49712454903.
- Elsner, M., L. Cuo, N. Voisin, J. Deems, A. Hamlet, J. Vano, K. B. Mickelson, S.-Y. Lee, and D. Lettenmaier (2010), Implications of 21st century climate change for the hydrology of Washington State, *Clim. Change*, 102(1-2), 225–260, doi:10.1007/s10584-010-9855-0.
- Elsner, M. M., S. Gangopadhyay, T. Pruitt, L. D. Brekke, N. Mizukami, and M. P. Clark (2014), How Does the Choice of Distributed Meteorological Data Affect Hydrologic Model Calibration and Streamflow Simulations?, *J. Hydrometeorol.*, 15(4), 1384–1403, doi:10.1175/JHM-D-13-083.1.
- Essery, R., S. Morin, Y. Lejeune, and C. B. Ménard (2013), A comparison of 1701 snow models using observations from an alpine site, *Adv. Water Resour.*, 55, 131–148, doi:10.1016/j.advwatres.2012.07.013.
- Fritze, H., I. T. Stewart, and E. Pebesma (2011), Shifts in Western North American Snowmelt Runoff Regimes for the Recent Warm Decades, *J. Hydrometeorol.*, 12(5), 989–1006, doi:10.1175/2011JHM1360.1.

- Garcia, E. S., C. L. Tague, and J. S. Choate (2013), Method of spatial temperature estimation influences ecohydrologic modeling in the Western Oregon cascades, *Water Resour. Res.*, n/a–n/a, doi:10.1002/wrcr.20140.
- Guentchev, G., J. J. Barsugli, and J. Eischeid (2010), Homogeneity of Gridded Precipitation Datasets for the Colorado River Basin, *J. Appl. Meteorol. Climatol.*, 49(12), 2404–2415, doi:10.1175/2010JAMC2484.1.
- Hamlet, A. F., and D. P. Lettenmaier (2005), Production of Temporally Consistent Gridded Precipitation and Temperature Fields for the Continental United States*, *J. Hydrometeorol.*, 6(3), 330–336, doi:10.1175/JHM420.1.
- Hamlet, A. F., P. W. Mote, M. P. Clark, and D. P. Lettenmaier (2005), Effects of Temperature and Precipitation Variability on Snowpack Trends in the Western United States*, *J. Clim.*, 18(21), 4545–4561, doi:10.1175/JCLI3538.1.
- Hidalgo, H. G. et al. (2009), Detection and Attribution of Streamflow Timing Changes to Climate Change in the Western United States, *J. Clim.*, 22(13), 3838–3855, doi:10.1175/2009JCLI2470.1.
- Howat, I. M., and S. Tulaczyk (2005), Climate sensitivity of spring snowpack in the Sierra Nevada, *J. Geophys. Res.*, 110(F4), doi:10.1029/2005JF000356.
- Iman, R. L. (2008), Latin Hypercube Sampling, in *Encyclopedia of Quantitative Risk Analysis and Assessment*, John Wiley & Sons, Ltd.
- Immerzeel, W. W., L. Petersen, S. Ragetti, and F. Pellicciotti (2014), The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas, *Water Resour. Res.*, n/a–n/a, doi:10.1002/2013WR014506.
- Iziomon, M. G., H. Mayer, and A. Matzarakis (2003), Downward atmospheric longwave irradiance under clear and cloudy skies: Measurement and parameterization, *J. Atmospheric Sol.-Terr. Phys.*, 65(10), 1107–1116, doi:10.1016/j.jastp.2003.07.007.
- Jefferson, A., A. Nolin, S. Lewis, and C. Tague (2008), Hydrogeologic controls on streamflow sensitivity to climate variation, *Hydrol. Process.*, 22(22), 4371–4385, doi:10.1002/hyp.7041.
- Kalnay, E. et al. (1996), The NCEP/NCAR 40-Year Reanalysis Project, *Bull. Am. Meteorol. Soc.*, 77(3), 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.
- Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resour. Res.*, 42(3), n/a–n/a, doi:10.1029/2005WR004362.

- Klos, P. Z., T. E. Link, and J. T. Abatzoglou (2014), Extent of the rain-snow transition zone in the western U.S. under historic and projected climate, *Geophys. Res. Lett.*, 2014GL060500, doi:10.1002/2014GL060500.
- Knowles, N., and D. Cayan (2004), Elevational Dependence of Projected Hydrologic Changes in the San Francisco Estuary and Watershed, *Clim. Change*, 62(1-3), 319–336, doi:10.1023/B:CLIM.0000013696.14308.b9.
- Knowles, N., M. D. Dettinger, and D. R. Cayan (2006), Trends in Snowfall versus Rainfall in the Western United States, *J. Clim.*, 19(18), 4545–4559, doi:10.1175/JCLI3850.1.
- Koch, S. E., M. desJardins, and P. J. Kocin (1983), An Interactive Barnes Objective Map Analysis Scheme for Use with Satellite and Conventional Data, *J. Clim. Appl. Meteorol.*, 22(9), 1487–1503, doi:10.1175/1520-0450(1983)022<1487:AIBOMA>2.0.CO;2.
- Lapo, K. E., L. M. Hinkelman, M. S. Raleigh, and J. D. Lundquist (2015), Impact of errors in the downwelling irradiances on simulations of snow water equivalent, snow surface temperature, and the snow energy balance, *Water Resour. Res.*, n/a–n/a, doi:10.1002/2014WR016259.
- Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *J. Geophys. Res. Atmospheres*, 99(D7), 14415–14428, doi:10.1029/94JD00483.
- Liston, G. E. (1995), Local Advection of Momentum, Heat, and Moisture during the Melt of Patchy Snow Covers, *J. Appl. Meteorol.*, 34(7), 1705–1715, doi:10.1175/1520-0450-34.7.1705.
- Liston, G. E., and K. Elder (2006a), A Distributed Snow-Evolution Modeling System (SnowModel), *J. Hydrometeorol.*, 7(6), 1259–1276, doi:10.1175/JHM548.1.
- Liston, G. E., and K. Elder (2006b), A Meteorological Distribution System for High-Resolution Terrestrial Modeling (MicroMet), *J. Hydrometeorol.*, 7(2), 217–234, doi:10.1175/JHM486.1.
- Liston, G. E., and C. A. Hiemstra (2008), A Simple Data Assimilation System for Complex Snow Distributions (SnowAssim), *J. Hydrometeorol.*, 9(5), 989–1004, doi:10.1175/2008JHM871.1.
- Liston, G. E., and M. Sturm (1998), A snow-transport model for complex terrain, *J. Glaciol.*, 44(148), 498–516.
- Liston, G. E., R. B. Haehnel, M. Sturm, C. A. Hiemstra, S. Berezovskaya, and R. D. Tabler (2007), Simulating complex snow distributions in windy environments

- using SnowTran-3D, *J. Glaciol.*, 53(181), 241–256, doi:10.3189/172756507782202865.
- Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. M. Andreadis, E. P. Maurer, and D. P. Lettenmaier (2013), A long-term hydrologically based data set of land surface fluxes and states for the conterminous U.S.: Update and extensions, *J. Clim.*, 130712132648002, doi:10.1175/JCLI-D-12-00508.1.
- Luce, C. H., J. T. Abatzoglou, and Z. A. Holden (2013), The Missing Mountain Water: Slower Westerlies Decrease Orographic Enhancement in the Pacific Northwest USA, *Science*, 342(6164), 1360–1364, doi:10.1126/science.1242335.
- Lundquist, J. D., N. Pepin, and C. Rochford (2008a), Automated algorithm for mapping regions of cold-air pooling in complex terrain, *J. Geophys. Res. Atmospheres*, 113(D22), n/a–n/a, doi:10.1029/2008JD009879.
- Lundquist, J. D., P. J. Neiman, B. Martner, A. B. White, D. J. Gottas, and F. M. Ralph (2008b), Rain versus Snow in the Sierra Nevada, California: Comparing Doppler Profiling Radar and Surface Observations of Melting Level, *J. Hydrometeorol.*, 9(2), 194–211, doi:10.1175/2007JHM853.1.
- Lundquist, J. D., S. E. Dickerson-Lange, J. A. Lutz, and N. C. Cristea (2013), Lower forest density enhances snow retention in regions with warmer winters: A global framework developed from plot-scale observations and modeling, *Water Resour. Res.*, 49(10), 6356–6370, doi:10.1002/wrcr.20504.
- Mahrt, L. (2006), Variation of Surface Air Temperature in Complex Terrain, *J. Appl. Meteorol. Climatol.*, 45(11), 1481–1493, doi:10.1175/JAM2419.1.
- Male, D. H., and R. J. Granger (1981), Snow surface energy exchange, *Water Resour. Res.*, 17(3), 609–627, doi:10.1029/WR017i003p00609.
- Mao, Y., B. Nijssen, and D. P. Lettenmaier (2015), Is climate change implicated in the 2013–2014 California drought? A hydrologic perspective, *Geophys. Res. Lett.*, 42(8), 2015GL063456, doi:10.1002/2015GL063456.
- Marks, D., and D. J. Dozier (1979), A clear-sky longwave radiation model for remote alpine areas, *Arch. Für Meteorol. Geophys. Bioklimatol. Ser. B*, 27(2-3), 159–187, doi:10.1007/BF02243741.
- Marks, D., J. Dozier, and R. E. Davis (1992), Climate and energy exchange at the snow surface in the Alpine Region of the Sierra Nevada: 1. Meteorological measurements and monitoring, *Water Resour. Res.*, 28(11), 3029–3042, doi:10.1029/92WR01482.
- Marks, D., J. Kimball, D. Tingey, and T. Link (1998), The sensitivity of snowmelt processes to climate conditions and forest cover during rain-on-snow: a case study

of the 1996 Pacific Northwest flood, *Hydrol. Process.*, 12(10-11), 1569–1587, doi:10.1002/(SICI)1099-1085(199808/09)12:10/11<1569::AID-HYP682>3.0.CO;2-L.

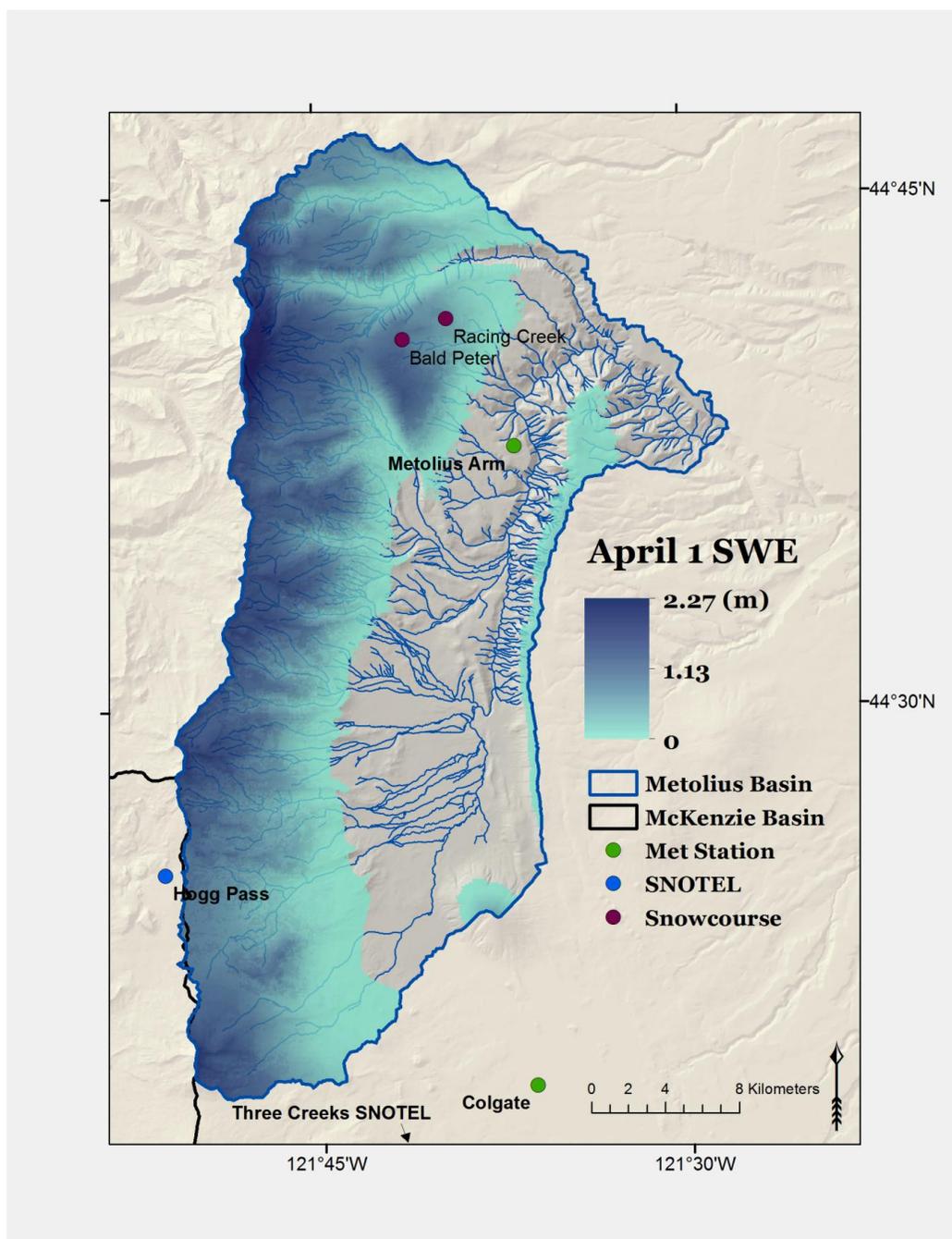
- Materia, S., P. A. Dirmeyer, Z. Guo, A. Alessandri, and A. Navarra (2009), The Sensitivity of Simulated River Discharge to Land Surface Representation and Meteorological Forcings, *J. Hydrometeorol.*, 11(2), 334–351, doi:10.1175/2009JHM1162.1.
- Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States*, *J. Clim.*, 15(22), 3237–3251, doi:10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2.
- Meek, D. W., and J. L. Hatfield (1994), Data quality checking for single station meteorological databases, *New Vis. New Perspect. Agric. Meteorol. Model.*, 69(1–2), 85–109, doi:10.1016/0168-1923(94)90083-3.
- Mernild, S. H., and G. E. Liston (2010), The Influence of Air Temperature Inversions on Snowmelt and Glacier Mass Balance Simulations, Ammassalik Island, Southeast Greenland, *J. Appl. Meteorol. Climatol.*, 49(1), 47–67, doi:10.1175/2009JAMC2065.1.
- Minder, J. R., P. W. Mote, and J. D. Lundquist (2010), Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains, *J. Geophys. Res. Atmospheres*, 115(D14), n/a–n/a, doi:10.1029/2009JD013493.
- Mitchell, K. E. et al. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res. Atmospheres*, 109(D7), D07S90, doi:10.1029/2003JD003823.
- Mizukami, N., M. P. Clark, A. G. Slater, L. D. Brekke, M. M. Elsner, J. R. Arnold, and S. Gangopadhyay (2013), Hydrologic Implications of Different Large-Scale Meteorological Model Forcing Datasets in Mountainous Regions, *J. Hydrometeorol.*, 15(1), 474–488, doi:10.1175/JHM-D-13-036.1.
- Montoya, E. L., J. Dozier, and W. Meiring (2014), Biases of April 1 snow water equivalent records in the Sierra Nevada and their associations with large-scale climate indices, *Geophys. Res. Lett.*, 41(16), 5912–5918, doi:10.1002/2014GL060588.
- Mote, P. W. (2003), Trends in snow water equivalent in the Pacific Northwest and their climatic causes, *Geophys. Res. Lett.*, 30(12), doi:10.1029/2003GL017258.
- Mote, P. W. (2006), Climate-Driven Variability and Trends in Mountain Snowpack in Western North America*, *J. Clim.*, 19(23), 6209–6220, doi:10.1175/JCLI3971.1.

- Mote, P. W., and E. P. S. Jr (2010), Future climate in the Pacific Northwest, *Clim. Change*, 102(1-2), 29–50, doi:10.1007/s10584-010-9848-z.
- Mote, P. W., A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier (2005), DECLINING MOUNTAIN SNOWPACK IN WESTERN NORTH AMERICA*, *Bull. Am. Meteorol. Soc.*, 86(1), 39–49, doi:10.1175/BAMS-86-1-39.
- Nolin, A. W. (2012), Perspectives on Climate Change, Mountain Hydrology, and Water Resources in the Oregon Cascades, USA, *Mt. Res. Dev.*, 32(S1), S35–S46, doi:10.1659/MRD-JOURNAL-D-11-00038.S1.
- Nolin, A. W., and C. Daly (2006), Mapping “At Risk” Snow in the Pacific Northwest, *J. Hydrometeorol.*, 7(5), 1164–1171, doi:10.1175/JHM543.1.
- Pepin, N., D. Benham, and K. Taylor (1999), Modeling Lapse Rates in the Maritime Uplands of Northern England: Implications for Climate Change, *Arct. Antarct. Alp. Res.*, 31(2), 151–164, doi:10.2307/1552603.
- Pepin, N. C., C. Daly, and J. Lundquist (2011), The influence of surface versus free-air decoupling on temperature trend patterns in the western United States, *J. Geophys. Res. Atmospheres*, 116(D10), n/a–n/a, doi:10.1029/2010JD014769.
- Pierce, D. W. et al. (2008), Attribution of Declining Western U.S. Snowpack to Human Effects, *J. Clim.*, 21(23), 6425–6444, doi:10.1175/2008JCLI2405.1.
- Randin, C. F., J.-P. Dedieu, M. Zappa, L. Long, and S. Dullinger (2014), Validation of and comparison between a semidistributed rainfall–runoff hydrological model (PREVAH) and a spatially distributed snow-evolution model (SnowModel) for snow cover prediction in mountain ecosystems, *Ecohydrology*, n/a–n/a, doi:10.1002/eco.1570.
- Rolland, C. (2003), Spatial and Seasonal Variations of Air Temperature Lapse Rates in Alpine Regions, *J. Clim.*, 16(7), 1032–1046, doi:10.1175/1520-0442(2003)016<1032:SASVOA>2.0.CO;2.
- Safeeq, M., G. E. Grant, S. L. Lewis, and C. L. Tague (2013a), Coupling snowpack and groundwater dynamics to interpret historical streamflow trends in the western United States, *Hydrol. Process.*, n/a–n/a, doi:10.1002/hyp.9628.
- Safeeq, M., A. Mair, and A. Fares (2013b), Temporal and spatial trends in air temperature on the Island of Oahu, Hawaii, *Int. J. Climatol.*, 33(13), 2816–2835, doi:10.1002/joc.3629.
- Salathe, E. P., P. W. Mote, and M. W. Wiley (2007), Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology in the United States pacific northwest, *Int. J. Climatol.*, 27(12), 1611–1621, doi:10.1002/joc.1540.

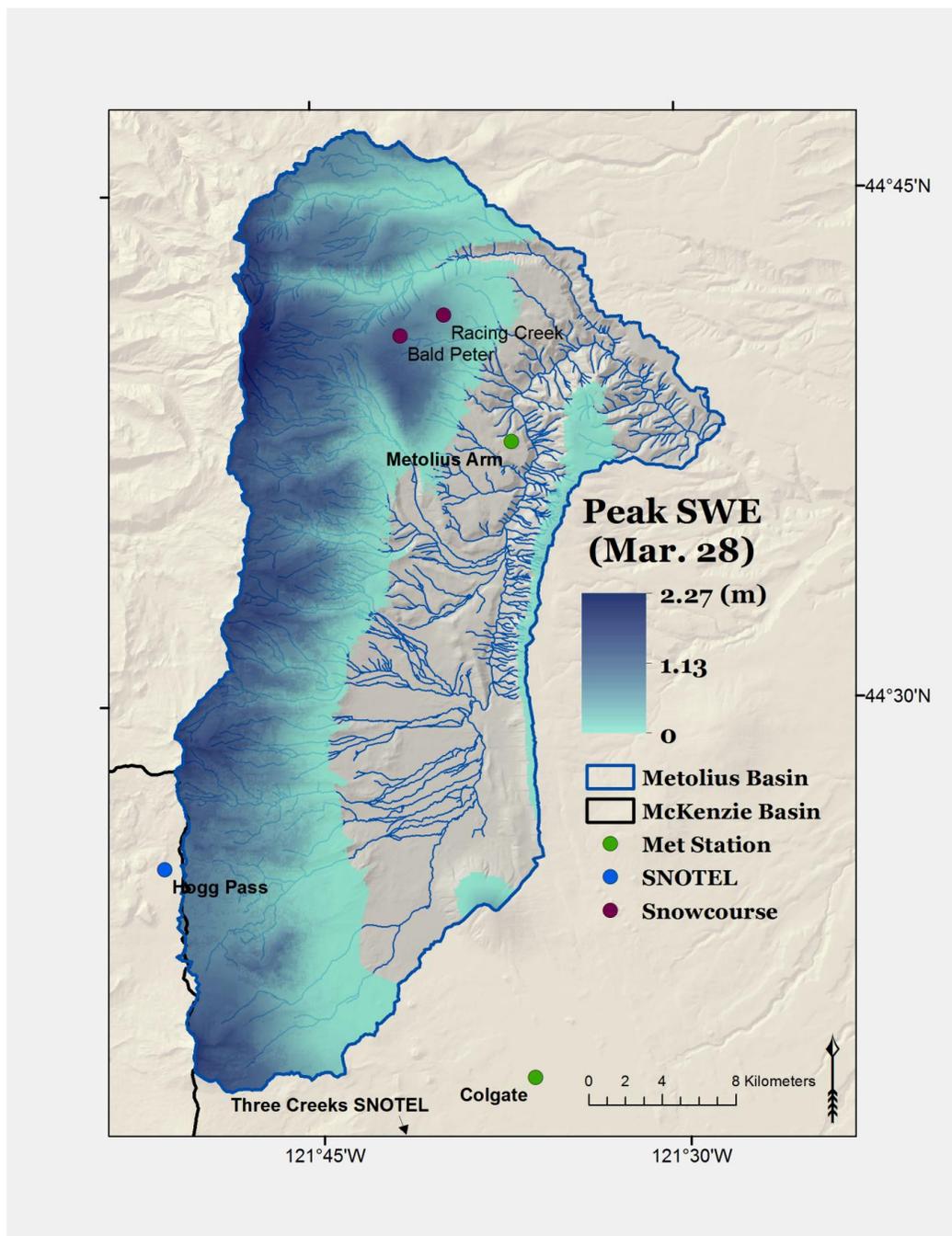
- Serreze, M. C., M. P. Clark, R. L. Armstrong, D. A. McGinnis, and R. S. Pulwarty (1999), Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, *Water Resour. Res.*, 35(7), 2145–2160, doi:10.1029/1999WR900090.
- Shepard, D. (1968), A Two-dimensional Interpolation Function for Irregularly-spaced Data, in *Proceedings of the 1968 23rd ACM National Conference*, pp. 517–524, ACM, New York, NY, USA.
- Sproles, E. A. (2012), Climate change impacts on mountain snowpack presented in a knowledge to action framework,
- Sproles, E. A., A. W. Nolin, K. Rittger, and T. H. Painter (2013), Climate change impacts on maritime mountain snowpack in the Oregon Cascades, *Hydrol Earth Syst Sci*, 17(7), 2581–2597, doi:10.5194/hess-17-2581-2013.
- Stahl, K., R. D. Moore, J. A. Floyer, M. G. Asplin, and I. G. McKendry (2006), Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density, *Agric. For. Meteorol.*, 139(3–4), 224–236, doi:10.1016/j.agrformet.2006.07.004.
- Steenburgh, W. J., C. F. Mass, and S. A. Ferguson (1997), The Influence of Terrain-Induced Circulations on Wintertime Temperature and Snow Level in the Washington Cascades, *Weather Forecast.*, 12(2), 208–227, doi:10.1175/1520-0434(1997)012<0208:TIOTIC>2.0.CO;2.
- Stewart, I. T., D. R. Cayan, and M. D. Dettinger (2005), Changes toward Earlier Streamflow Timing across Western North America, *J. Clim.*, 18(8), 1136–1155, doi:10.1175/JCLI3321.1.
- Strack, J. E., G. E. Liston, and R. A. Pielke (2004), Modeling Snow Depth for Improved Simulation of Snow–Vegetation–Atmosphere Interactions, *J. Hydrometeorol.*, 5(5), 723–734, doi:10.1175/1525-7541(2004)005<0723:MSDFIS>2.0.CO;2.
- Tague, C. L., and L. E. Band (2004), RHESys: Regional Hydro-Ecologic Simulation System—An Object-Oriented Approach to Spatially Distributed Modeling of Carbon, Water, and Nutrient Cycling, *Earth Interact.*, 8(19), 1–42, doi:10.1175/1087-3562(2004)8<1:RRHSSO>2.0.CO;2.
- Thornton, P. E., and S. W. Running (1999), An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation, *Agric. For. Meteorol.*, 93(4), 211–228, doi:10.1016/S0168-1923(98)00126-9.
- Thornton, P. E., S. W. Running, and M. A. White (1997), Generating surfaces of daily meteorological variables over large regions of complex terrain, *J. Hydrol.*, 190(3–4), 214–251, doi:10.1016/S0022-1694(96)03128-9.

- Trujillo, E., N. P. Molotch, M. L. Goulden, A. E. Kelly, and R. C. Bales (2012), Elevation-dependent influence of snow accumulation on forest greening, *Nat. Geosci.*, 5(10), 705–709, doi:10.1038/ngeo1571.
- Vano, J. A., T. Das, and D. P. Lettenmaier (2012), Hydrologic Sensitivities of Colorado River Runoff to Changes in Precipitation and Temperature*, *J. Hydrometeorol.*, 13(3), 932–949, doi:10.1175/JHM-D-11-069.1.
- Walcek, C. J. (1994), Cloud Cover and Its Relationship to Relative Humidity during a Springtime Midlatitude Cyclone, *Mon. Weather Rev.*, 122(6), 1021–1035, doi:10.1175/1520-0493(1994)122<1021:CCAIRT>2.0.CO;2.
- Watanabe, S., S. Kanae, S. Seto, P. J.-F. Yeh, Y. Hirabayashi, and T. Oki (2012), Intercomparison of bias-correction methods for monthly temperature and precipitation simulated by multiple climate models, *J. Geophys. Res. Atmospheres*, 117(D23), n/a–n/a, doi:10.1029/2012JD018192.
- Wayand, N. E., A. F. Hamlet, M. Hughes, S. I. Feld, and J. D. Lundquist (2013), Intercomparison of Meteorological Forcing Data from Empirical and Mesoscale Model Sources in the North Fork American River Basin in Northern Sierra Nevada, California*, *J. Hydrometeorol.*, 14(3), 677–699, doi:10.1175/JHM-D-12-0102.1.
- Westerling, A. L., H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam (2006), Warming and Earlier Spring Increase Western U.S. Forest Wildfire Activity, *Science*, 313(5789), 940–943, doi:10.1126/science.1128834.
- Widmann, M., and C. S. Bretherton (2000), Validation of Mesoscale Precipitation in the NCEP Reanalysis Using a New Gridcell Dataset for the Northwestern United States, *J. Clim.*, 13(11), 1936–1950, doi:10.1175/1520-0442(2000)013<1936:VOMPIT>2.0.CO;2.
- Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier (1994), A distributed hydrology-vegetation model for complex terrain, *Water Resour. Res.*, 30(6), 1665–1679, doi:10.1029/94WR00436.
- Winstral, A., D. Marks, and R. Gurney (2014), Assessing the Sensitivities of a Distributed Snow Model to Forcing Data Resolution, *J. Hydrometeorol.*, 15(4), 1366–1383, doi:10.1175/JHM-D-13-0169.1.
- Xia, Y. et al. (2012), Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, *J. Geophys. Res. Atmospheres*, 117(D3), D03109, doi:10.1029/2011JD016048.

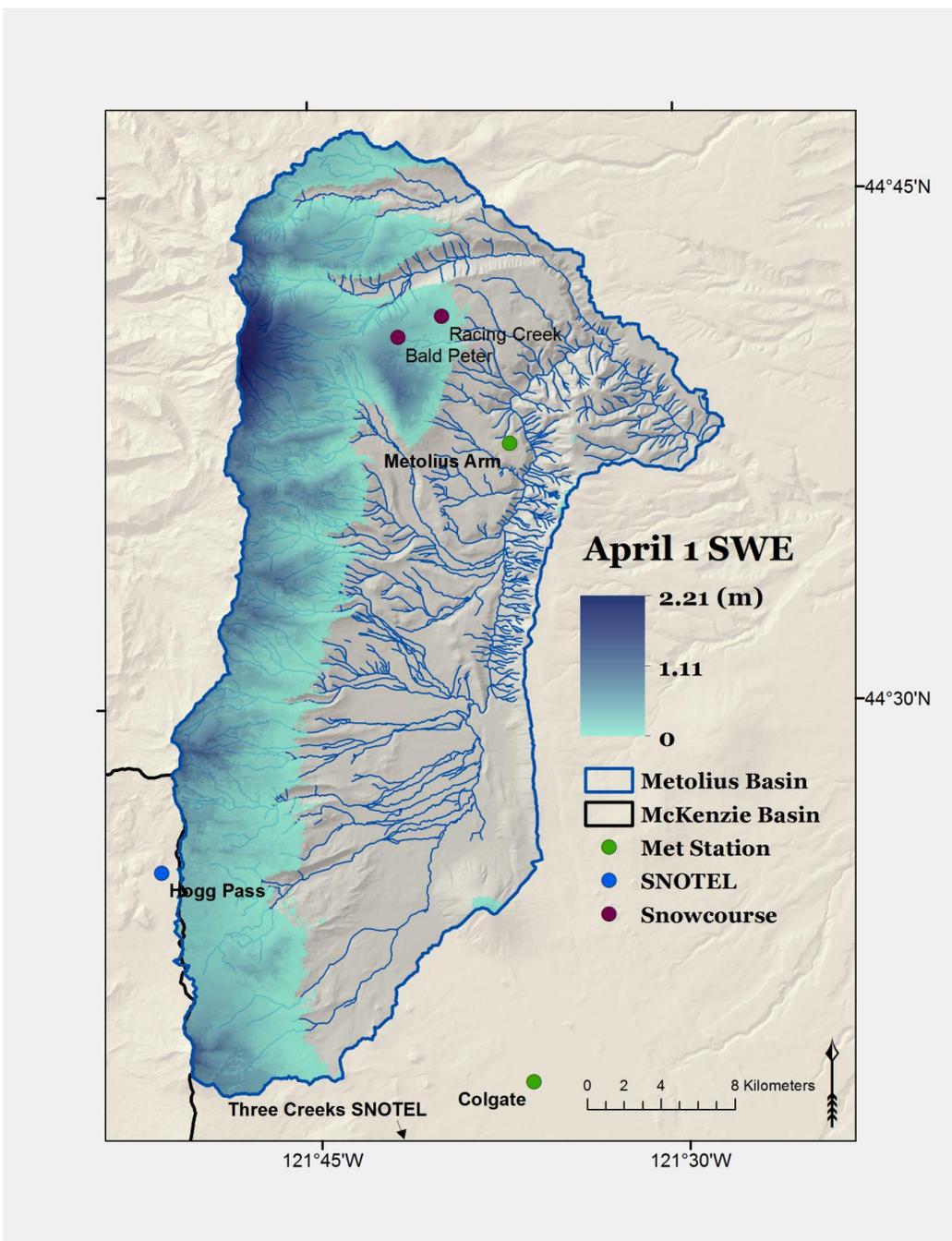
Appendices



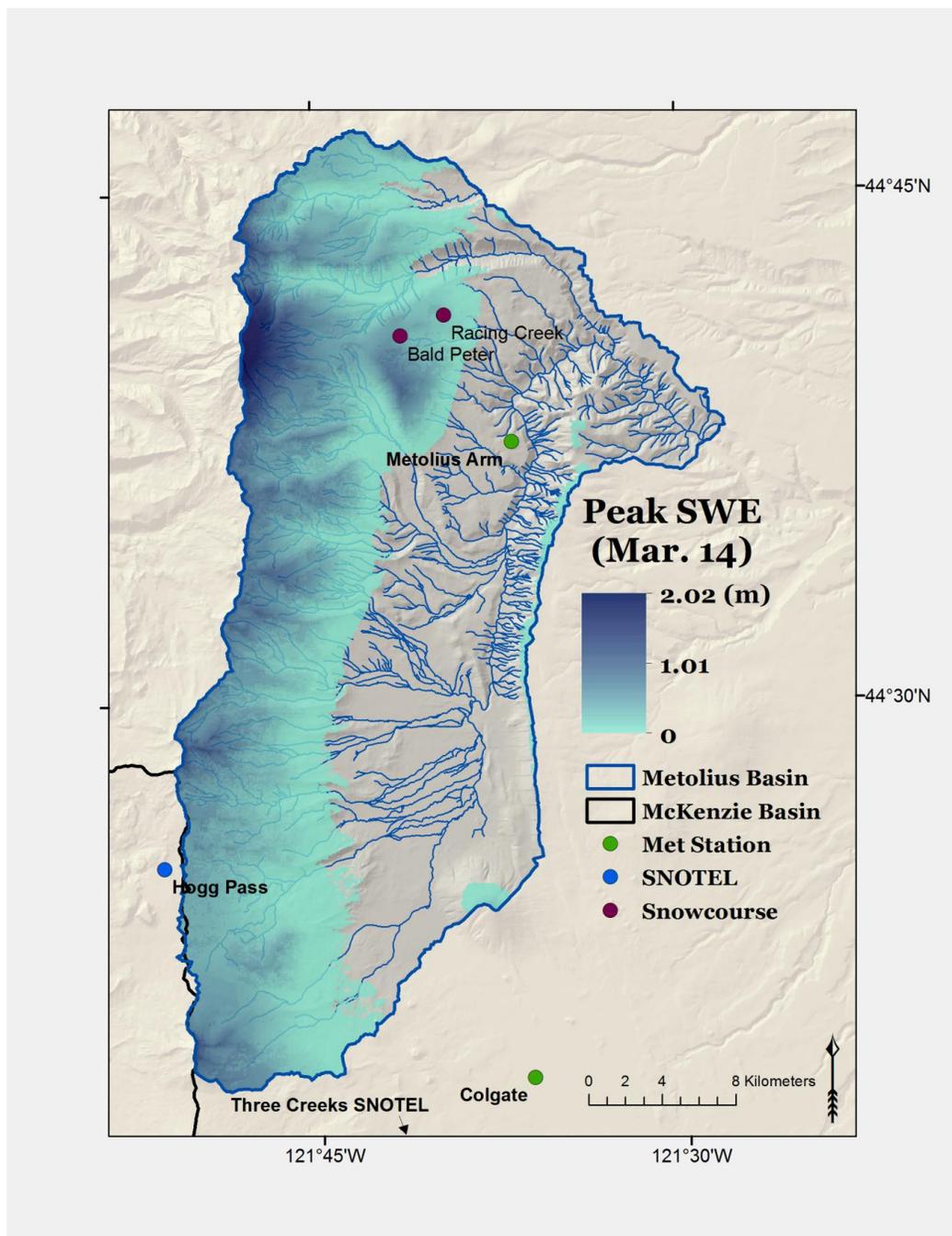
Appendix A: Simulated SWE on April 1.



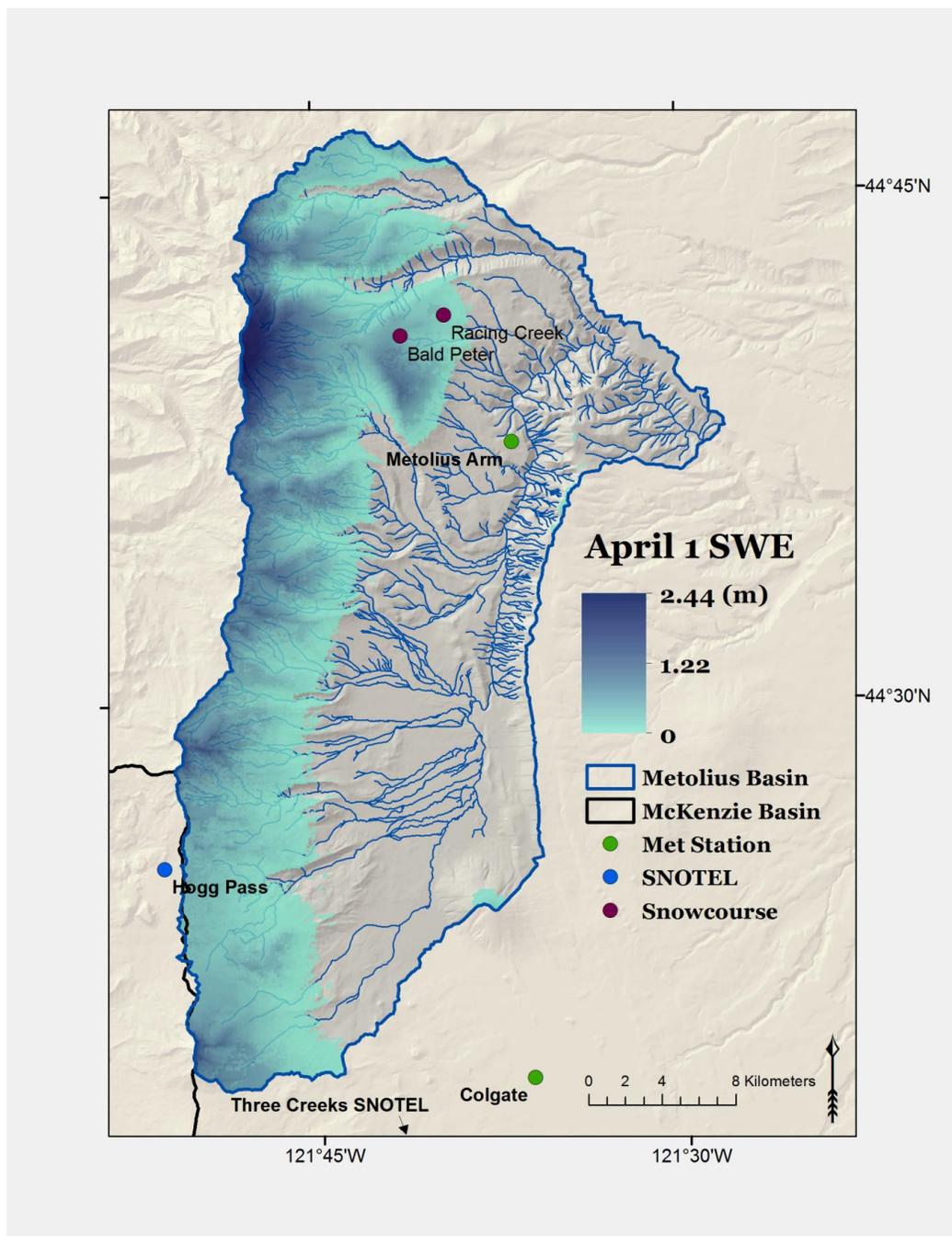
Appendix B: Simulated SWE on March 28, the date of peak basin-integrated SWE.



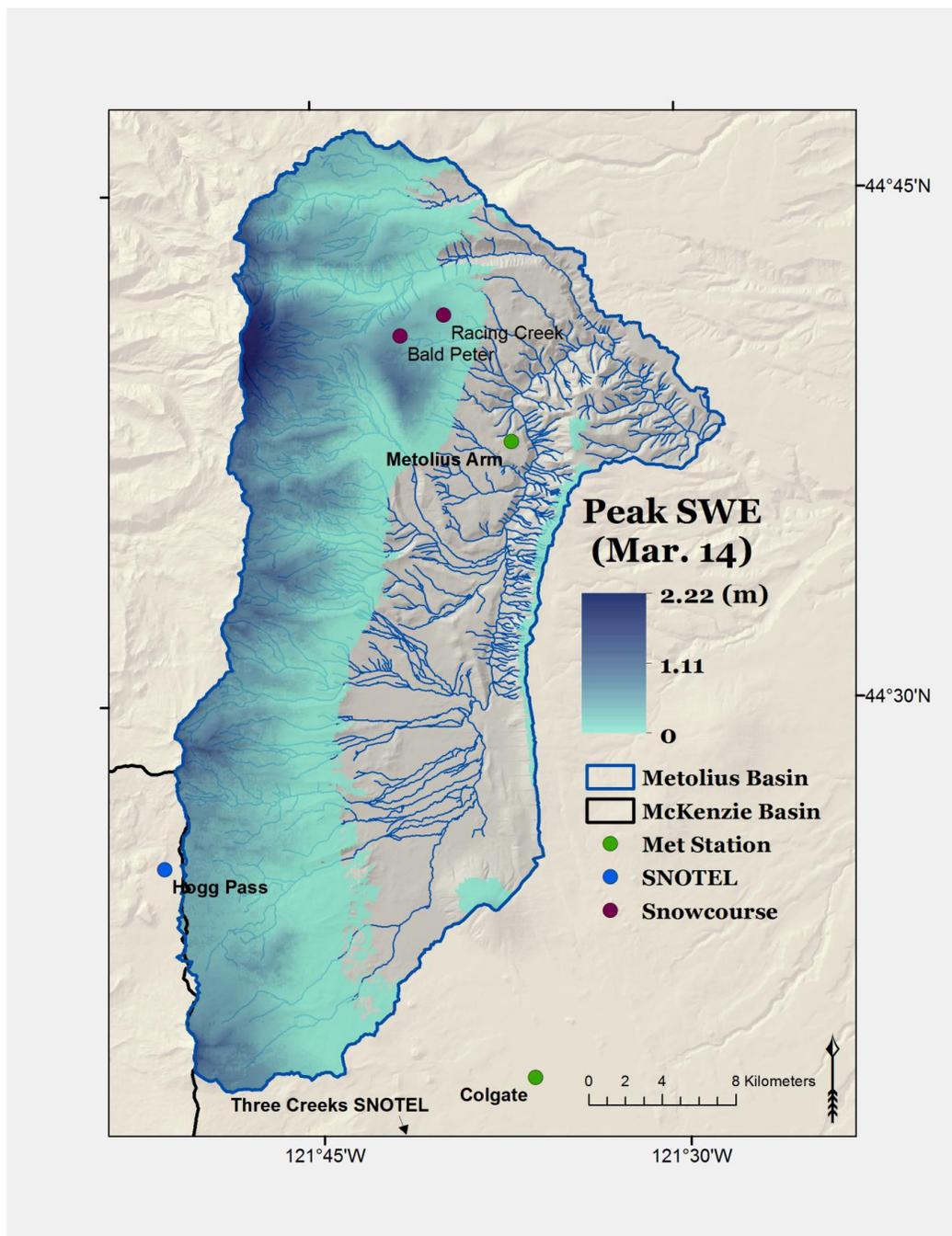
Appendix C: Simulated SWE on April 1 with a 2°C warming.



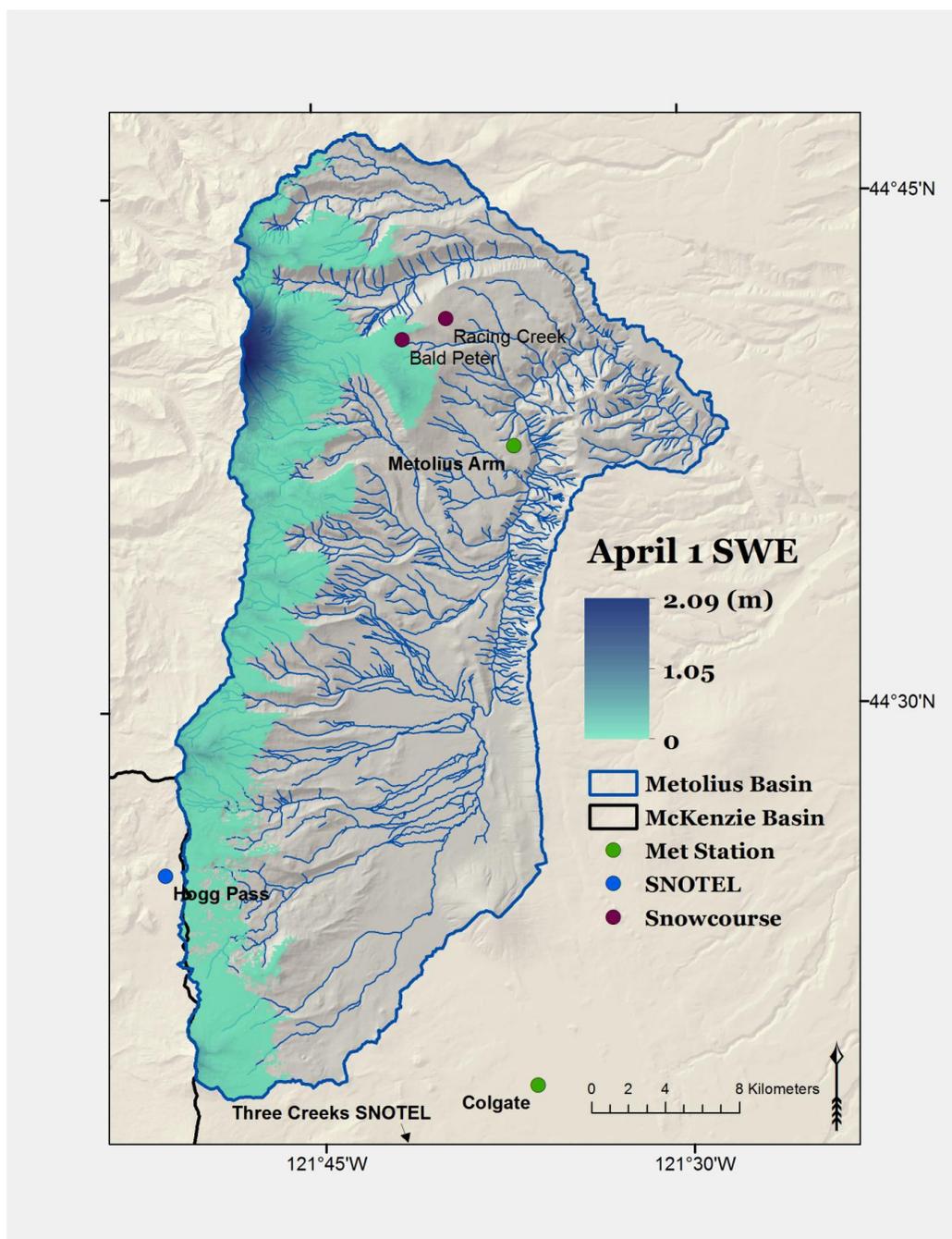
Appendix D: Simulated SWE on March 14, the date of peak basin-integrated SWE with a 2°C warming.



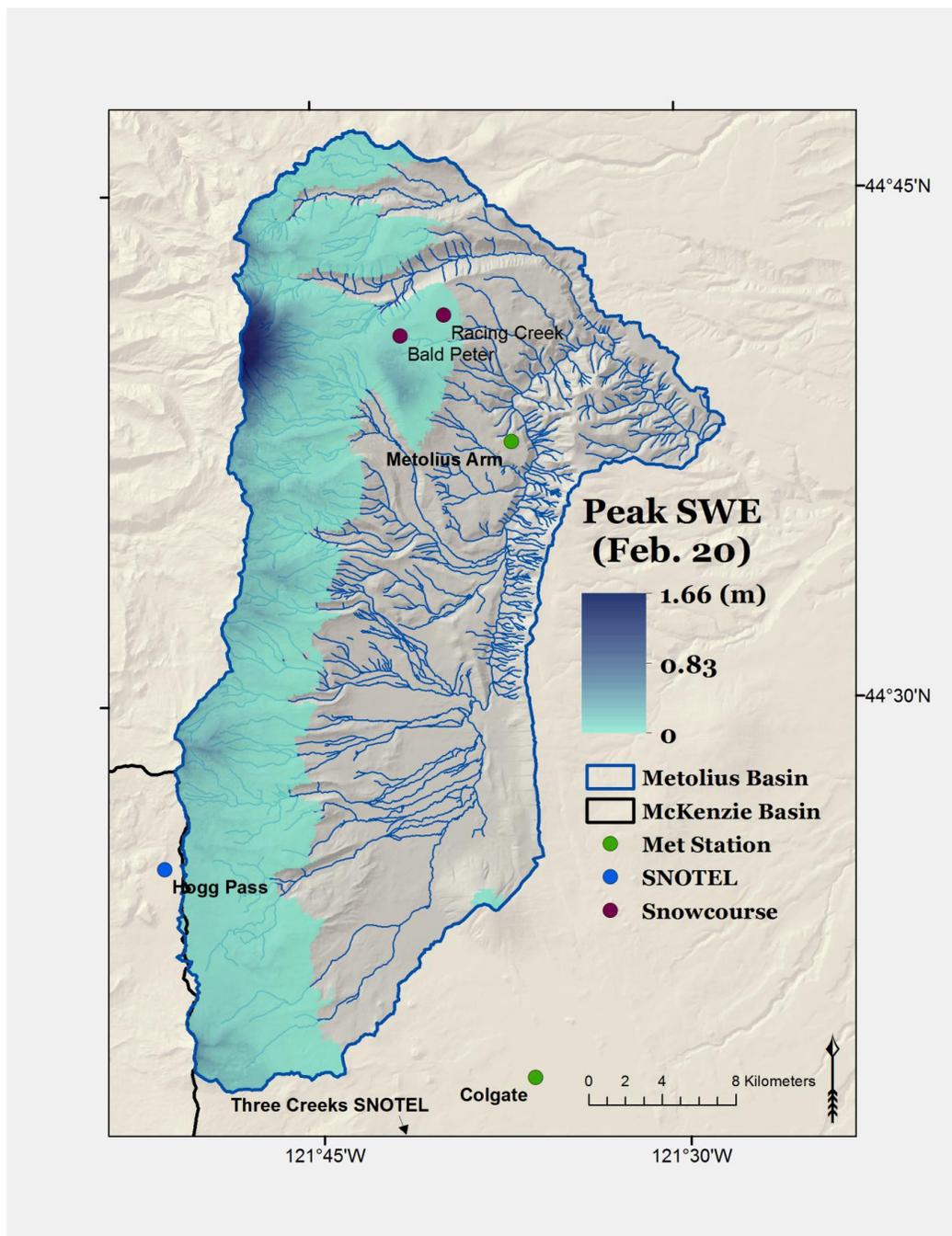
Appendix E: Simulated SWE on April 1 with a 2°C warming and +10% precipitation.



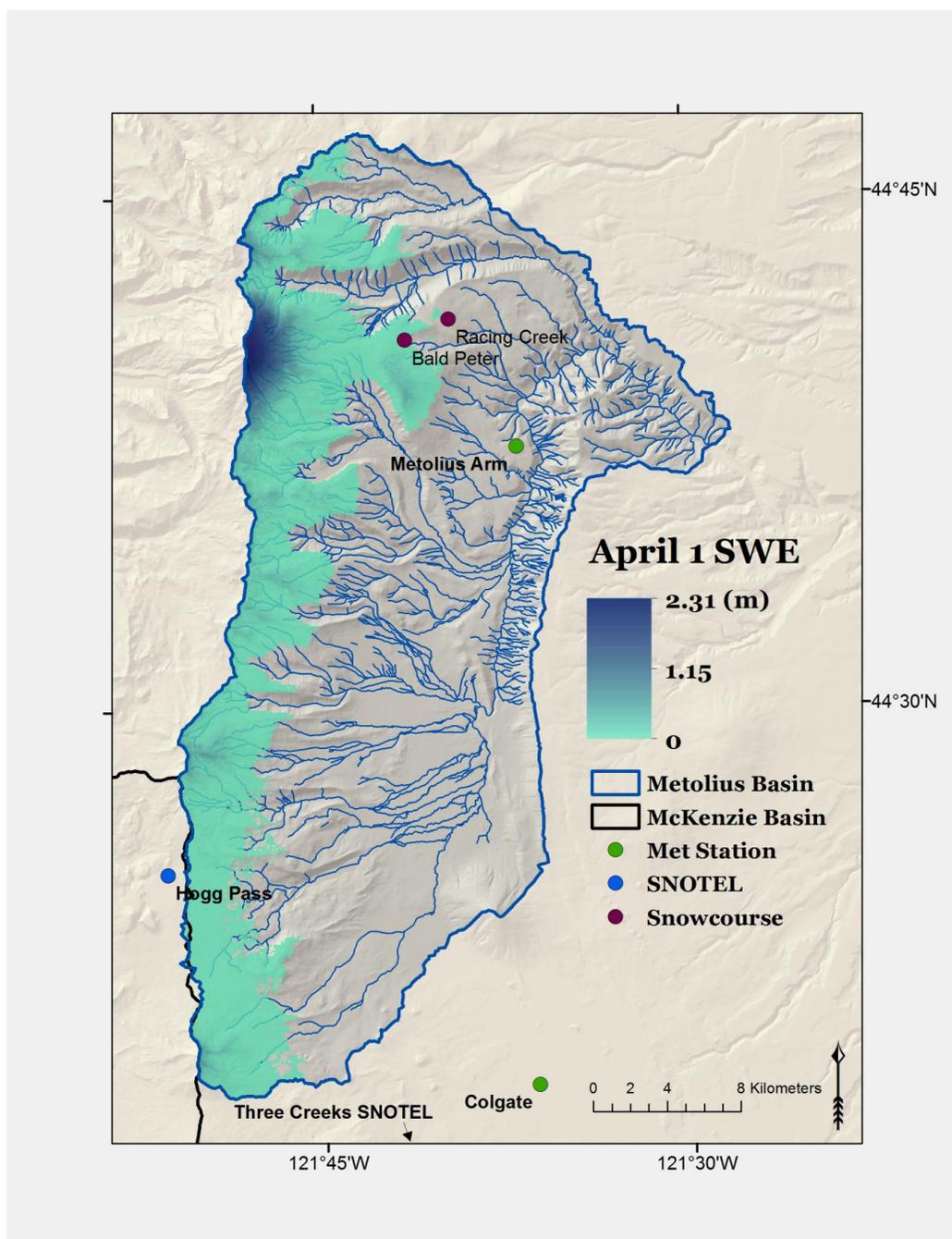
Appendix F: Simulated SWE on March 14, the date of peak basin-integrated SWE with a 2oC warming and +10% precipitation.



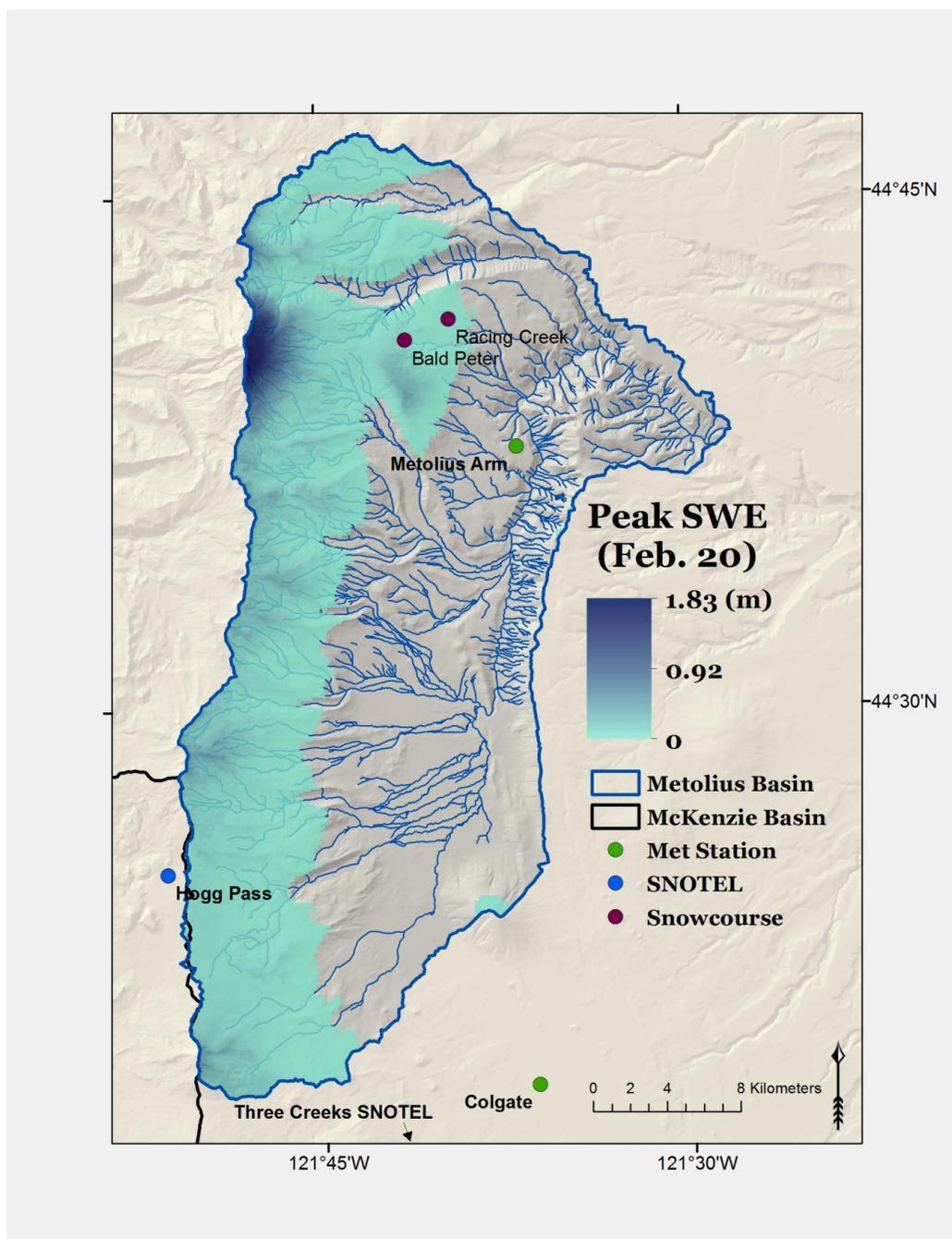
Appendix G: Simulated SWE on April 1 with a 4°C warming.



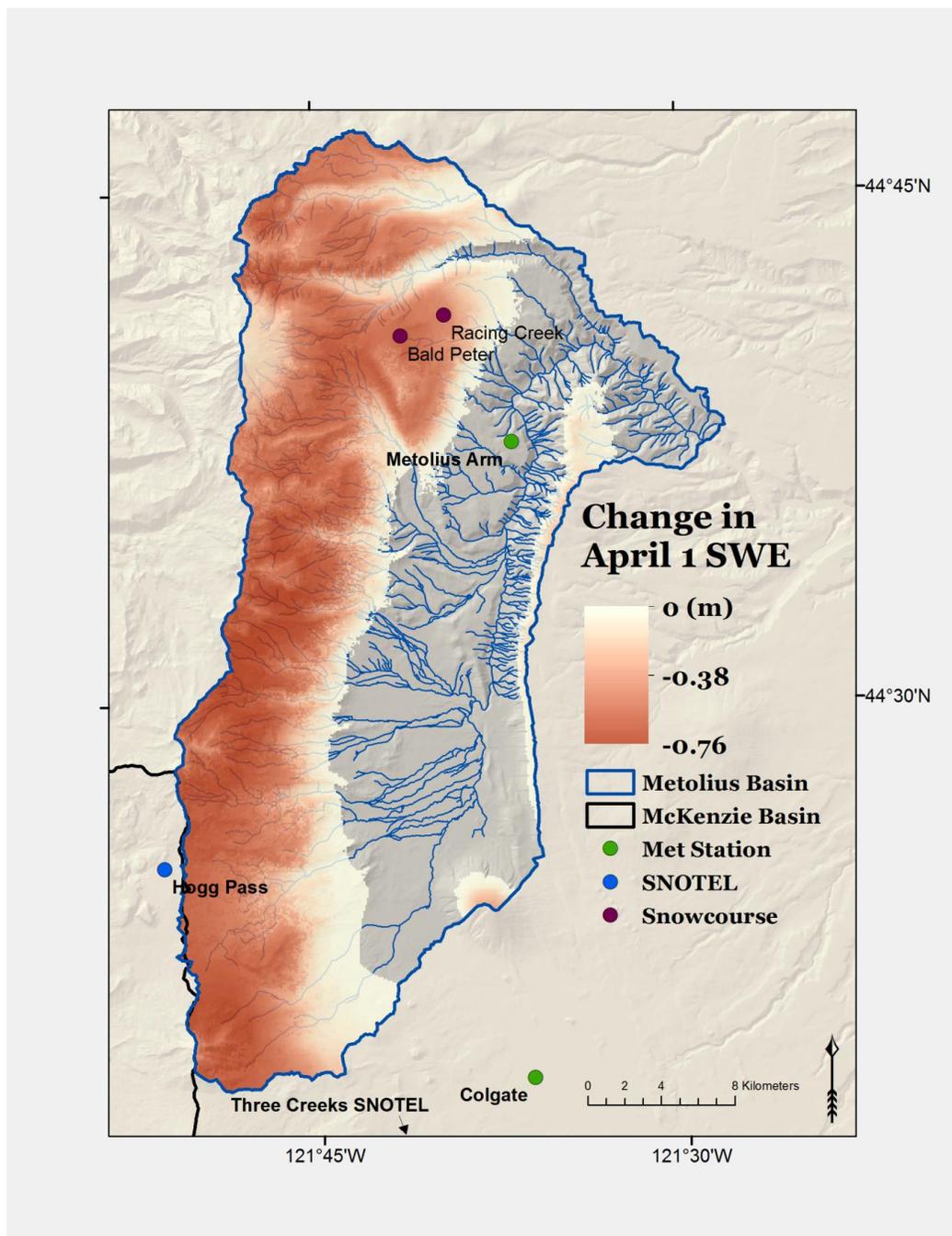
Appendix H: Simulated SWE on February 20, the date of peak basin-integrated SWE with a 4°C warming.



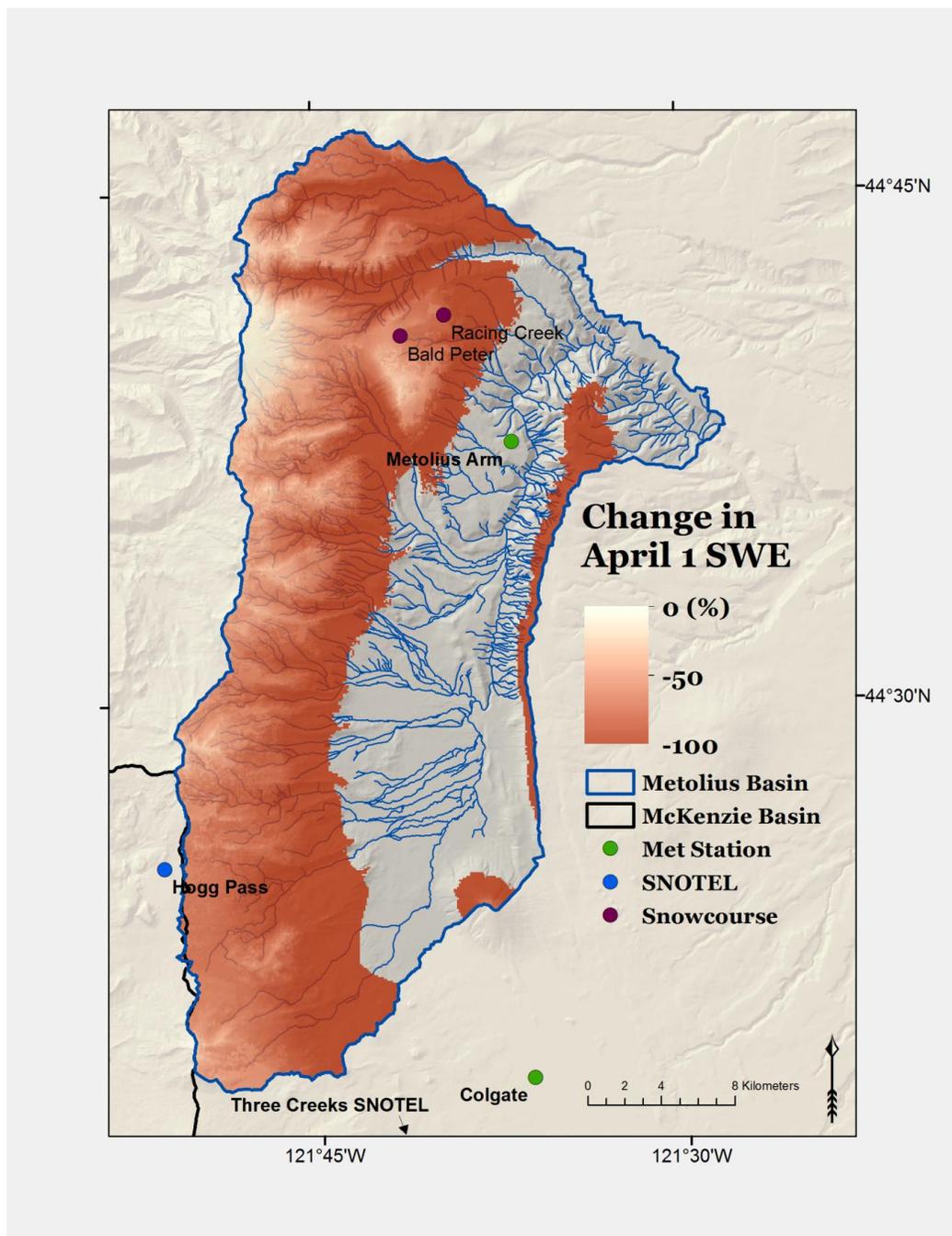
Appendix I: Simulated SWE on April 1 with a 4°C warming and +10% precipitation.



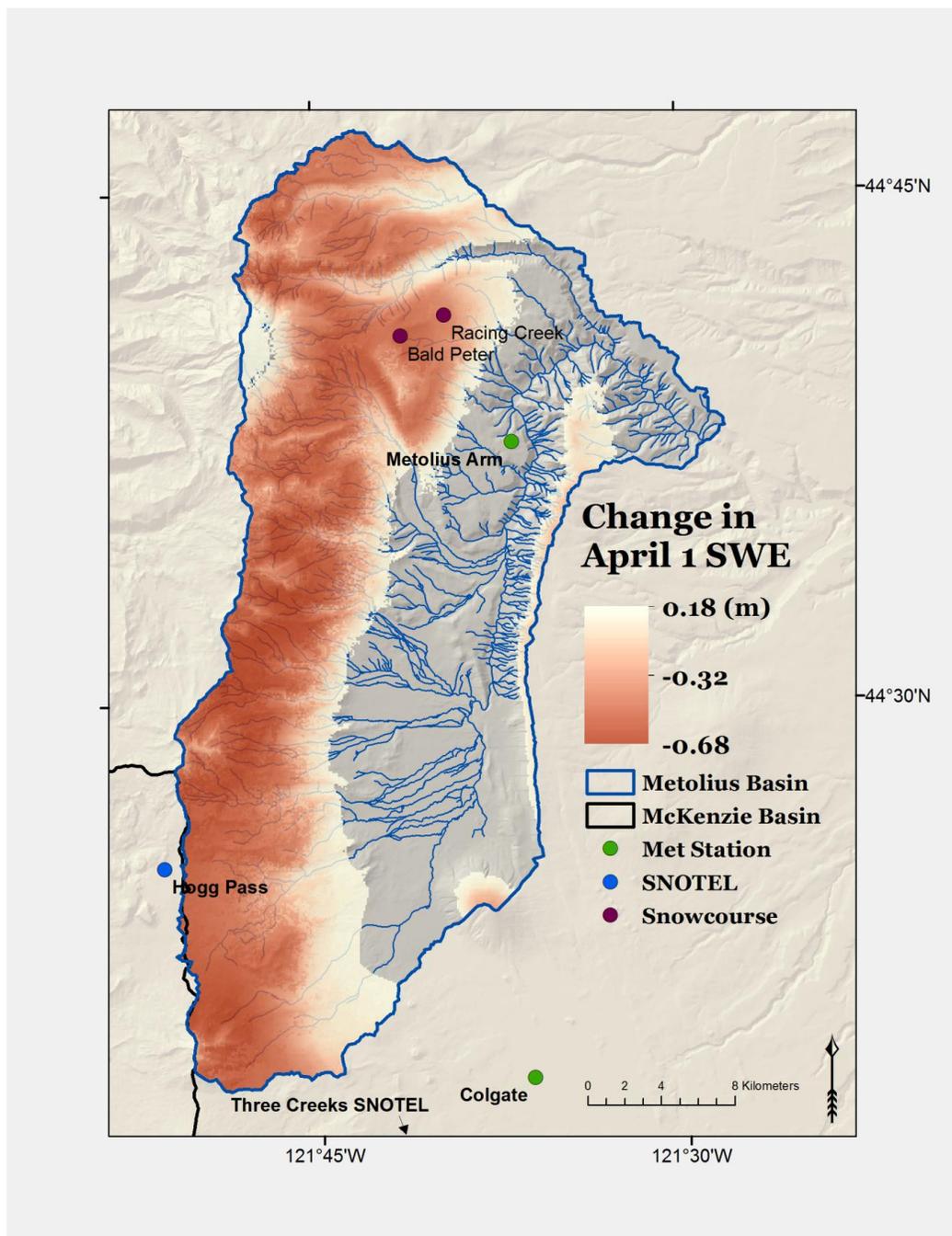
Appendix J: Simulated SWE on February 20, the date of peak basin-integrated SWE with a 4°C warming and +10% precipitation.



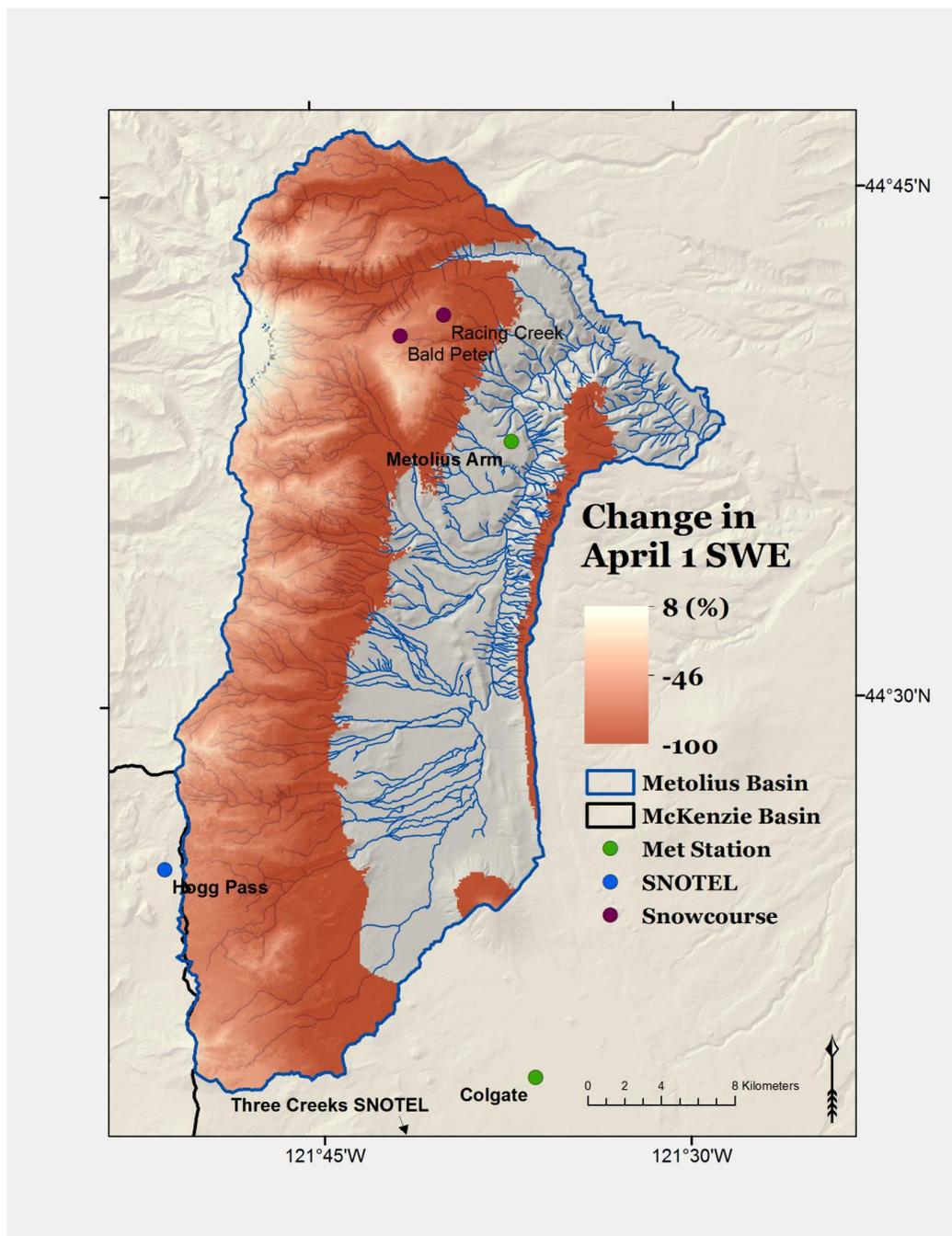
Appendix K: Simulated absolute change in SWE on April 1 with a 2°C warming.



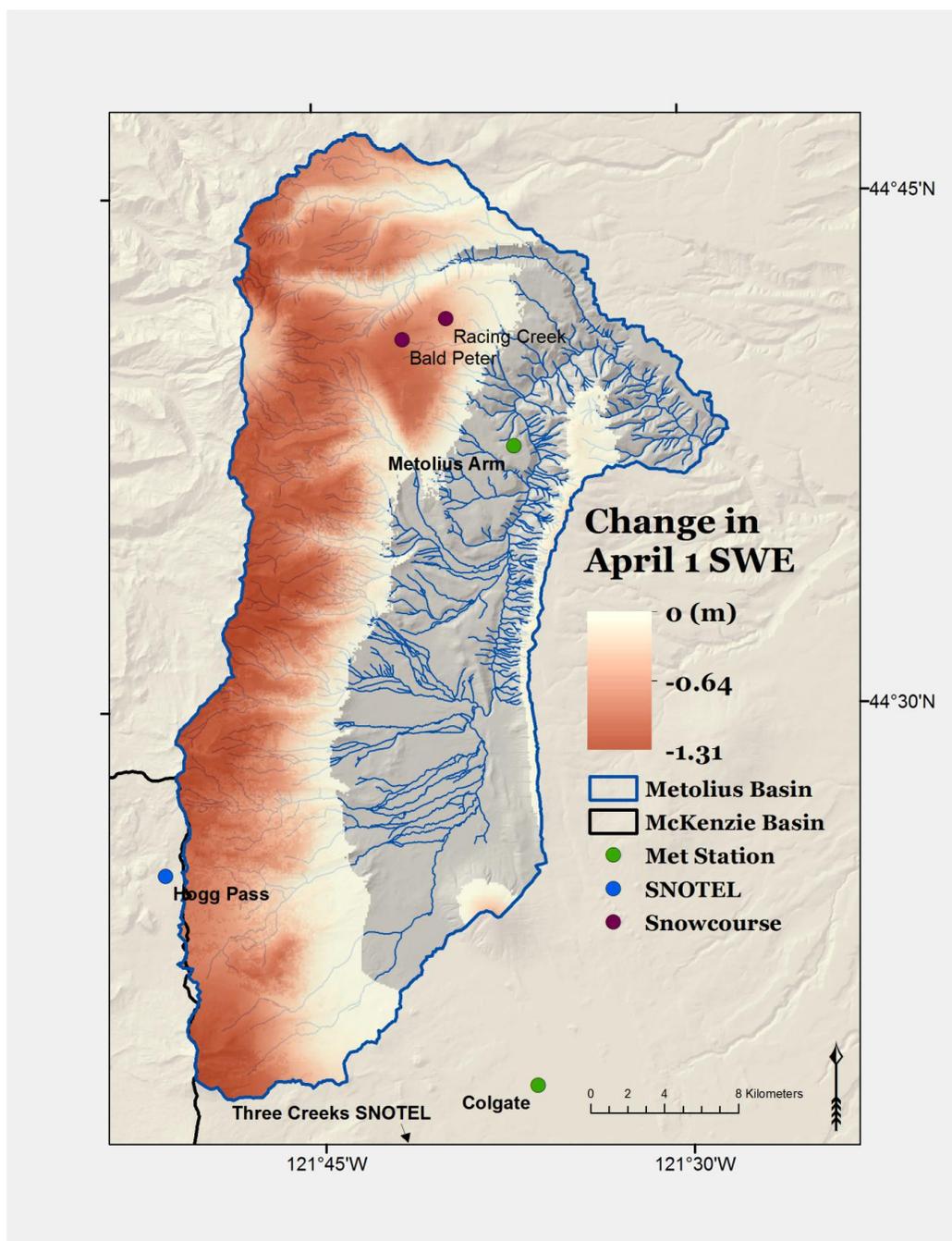
Appendix L: Simulated percent change in SWE on April 1 with a 2°C warming.



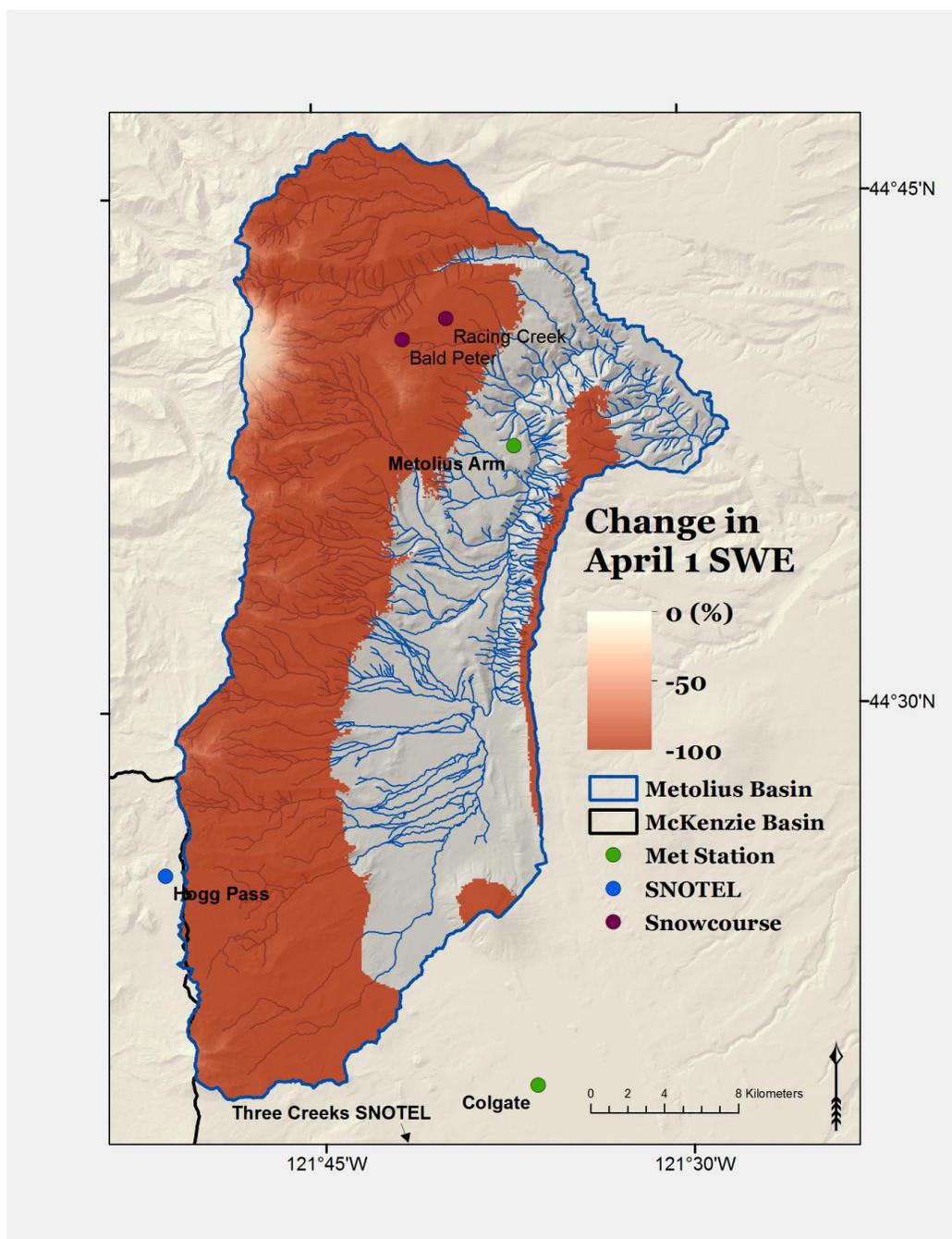
Appendix M: Simulated absolute change in SWE on April 1 with a 2°C warming and +10% precipitation.



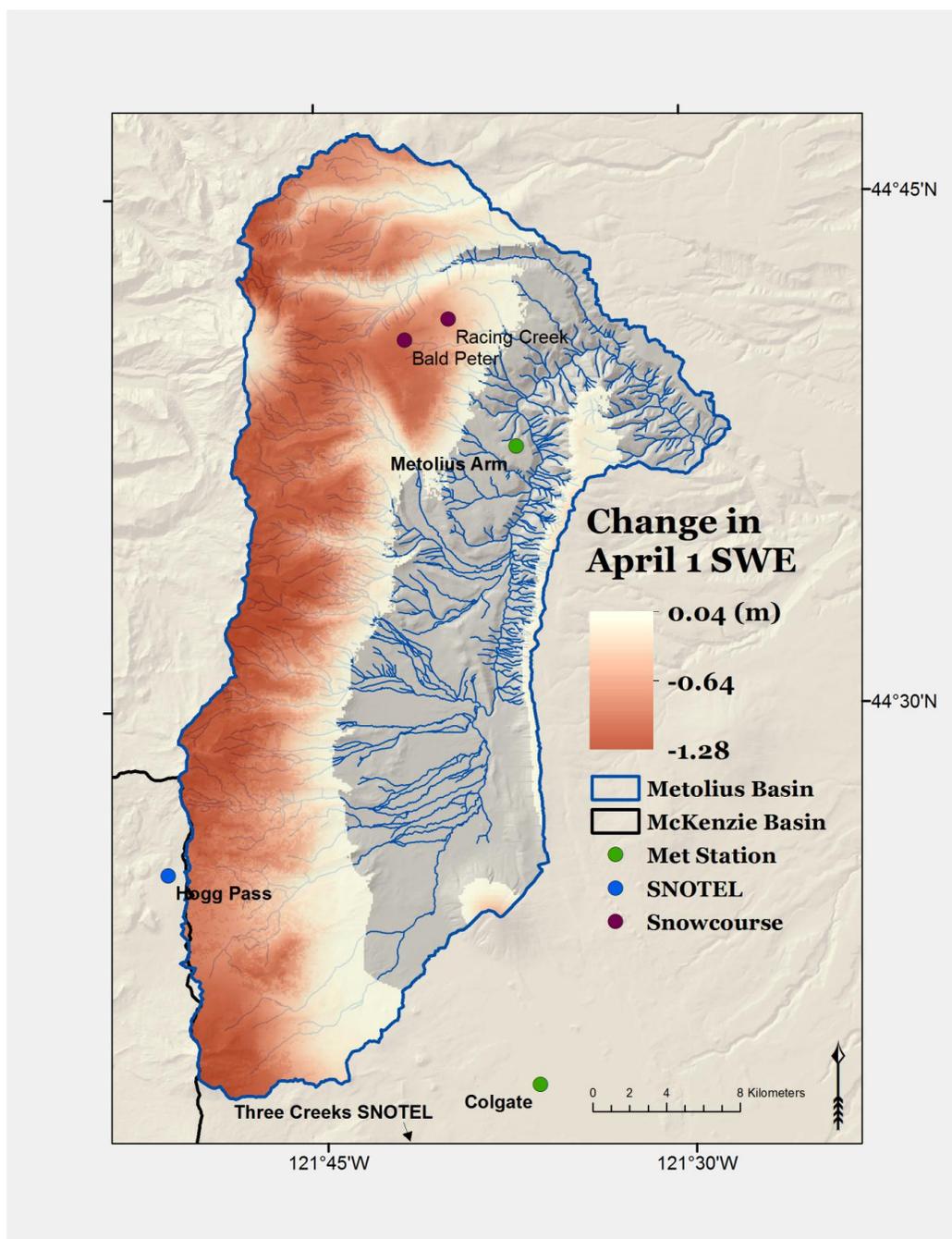
Appendix N: Simulated percent change in SWE on April 1 with a 2°C warming and +10% precipitation.



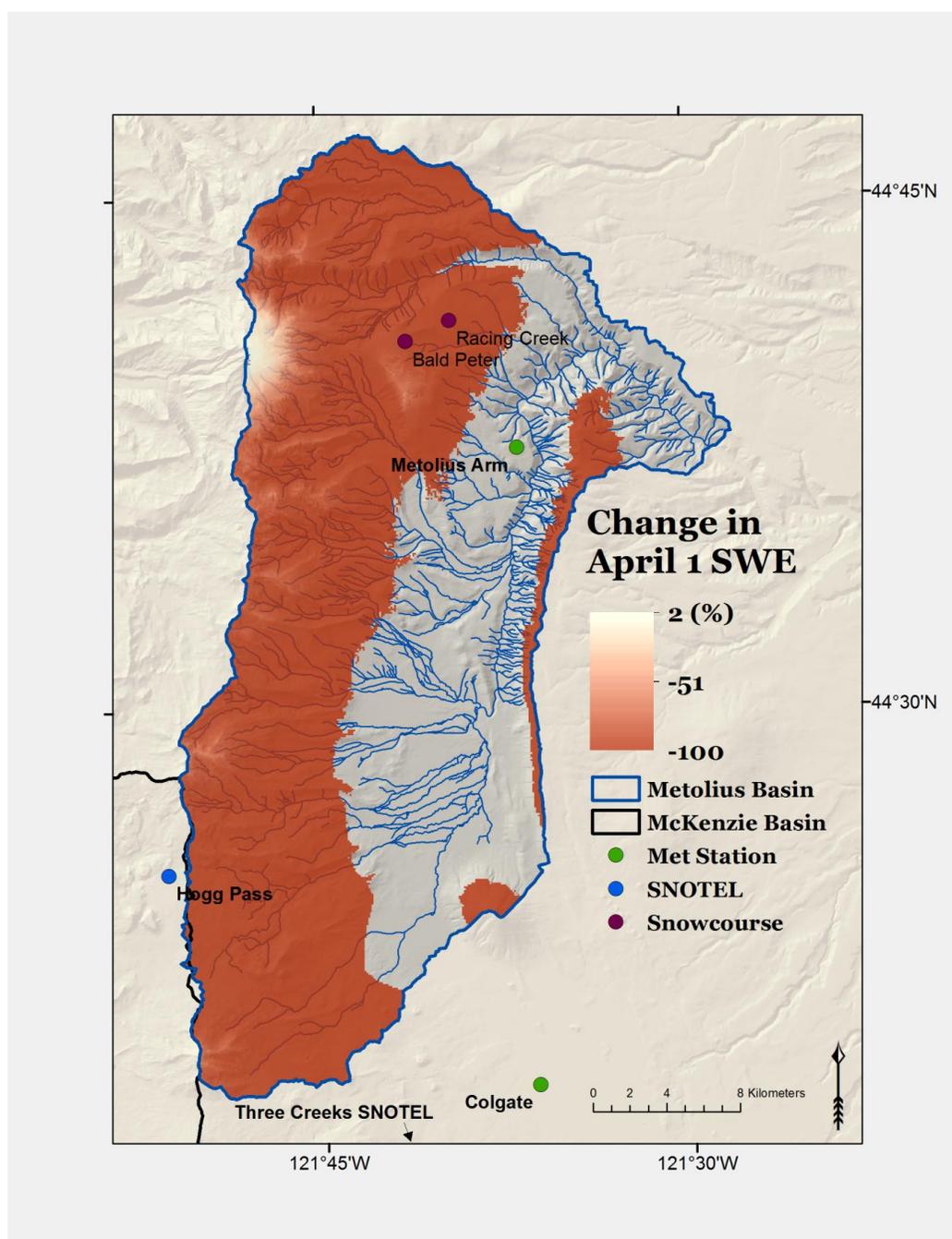
Appendix O: Simulated absolute change in SWE on April 1 with a 4°C warming.



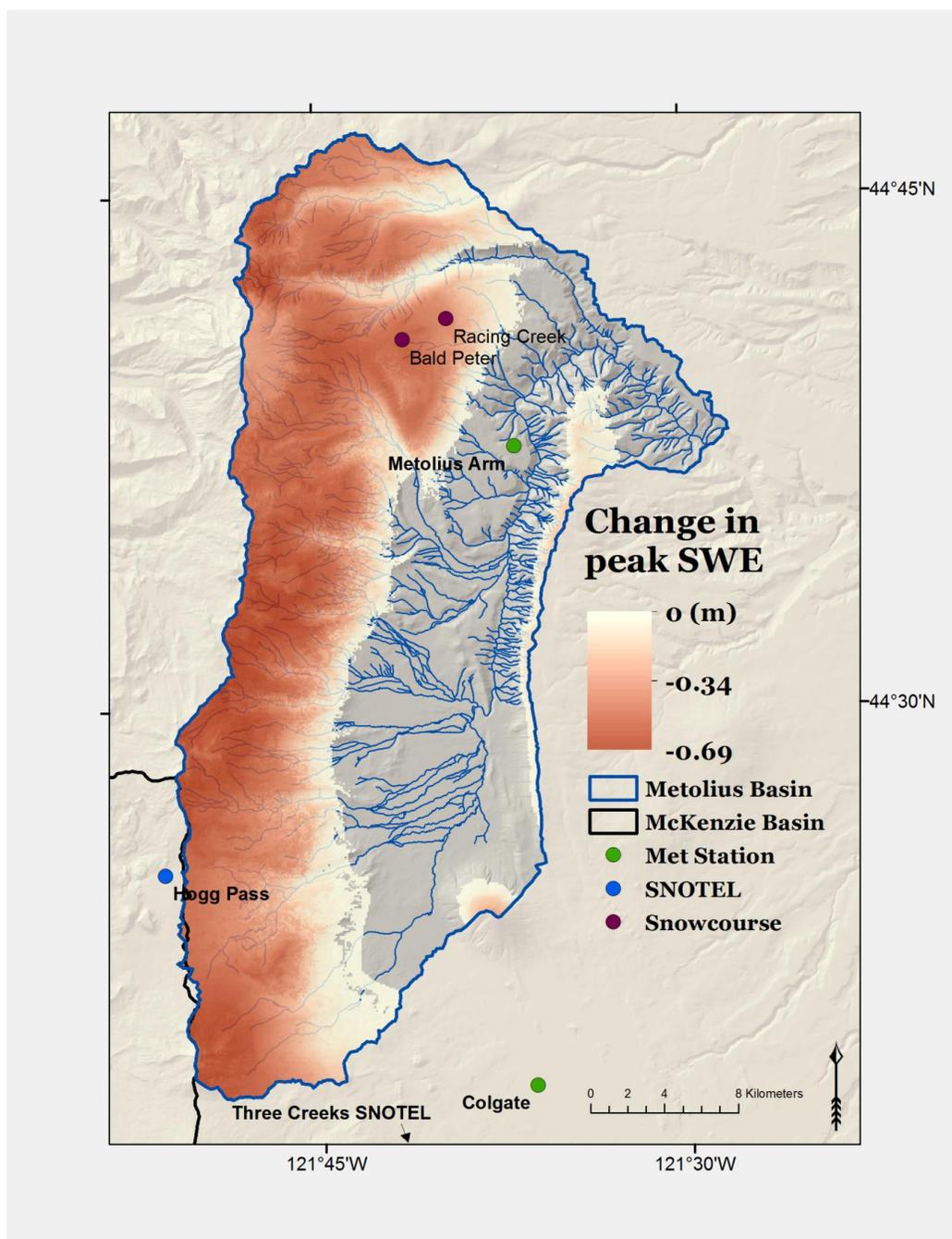
Appendix P: Simulated percent change in SWE on April 1 with a 4°C warming.



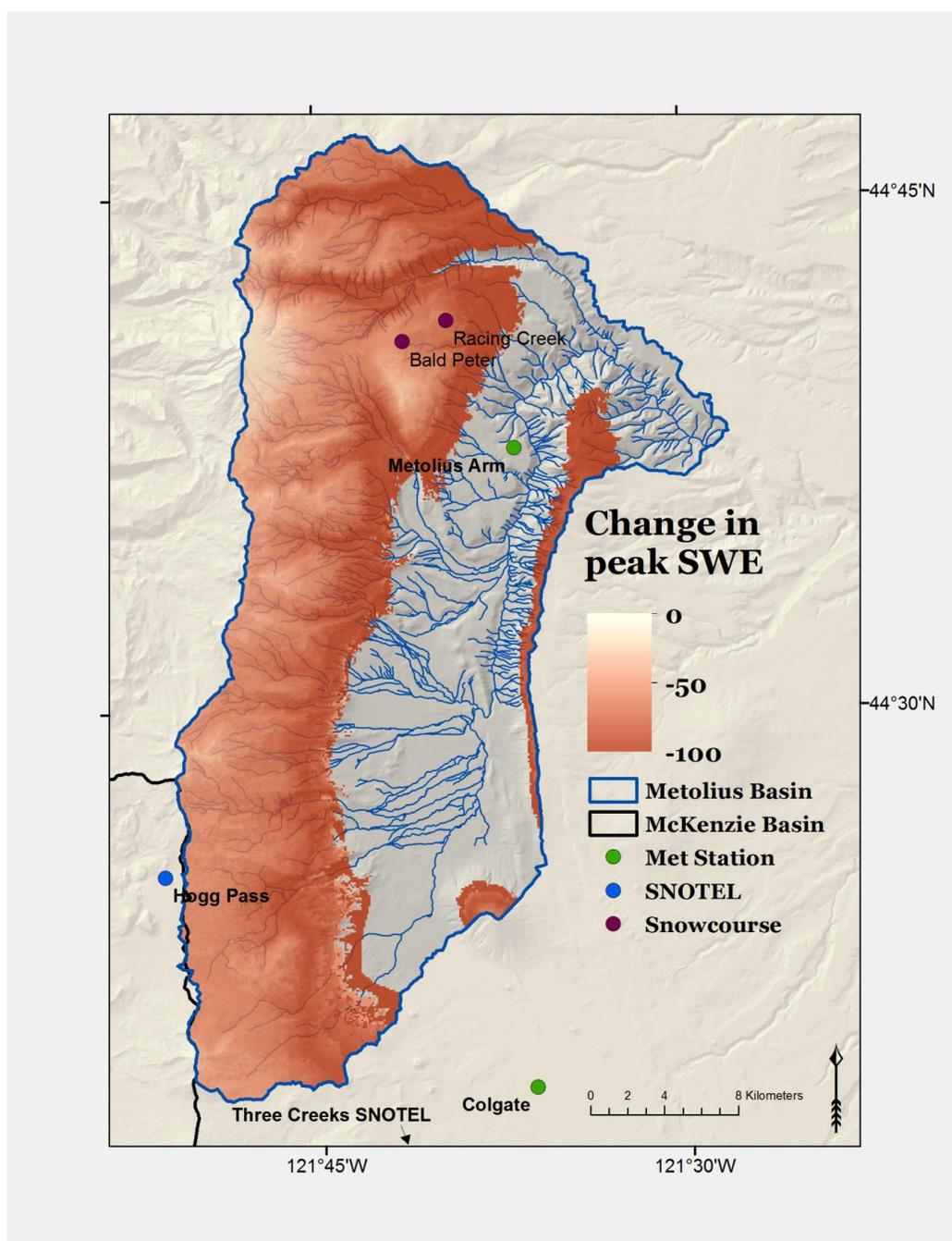
Appendix Q: Simulated absolute change in SWE on April 1 with a 4°C warming and +10% precipitation.



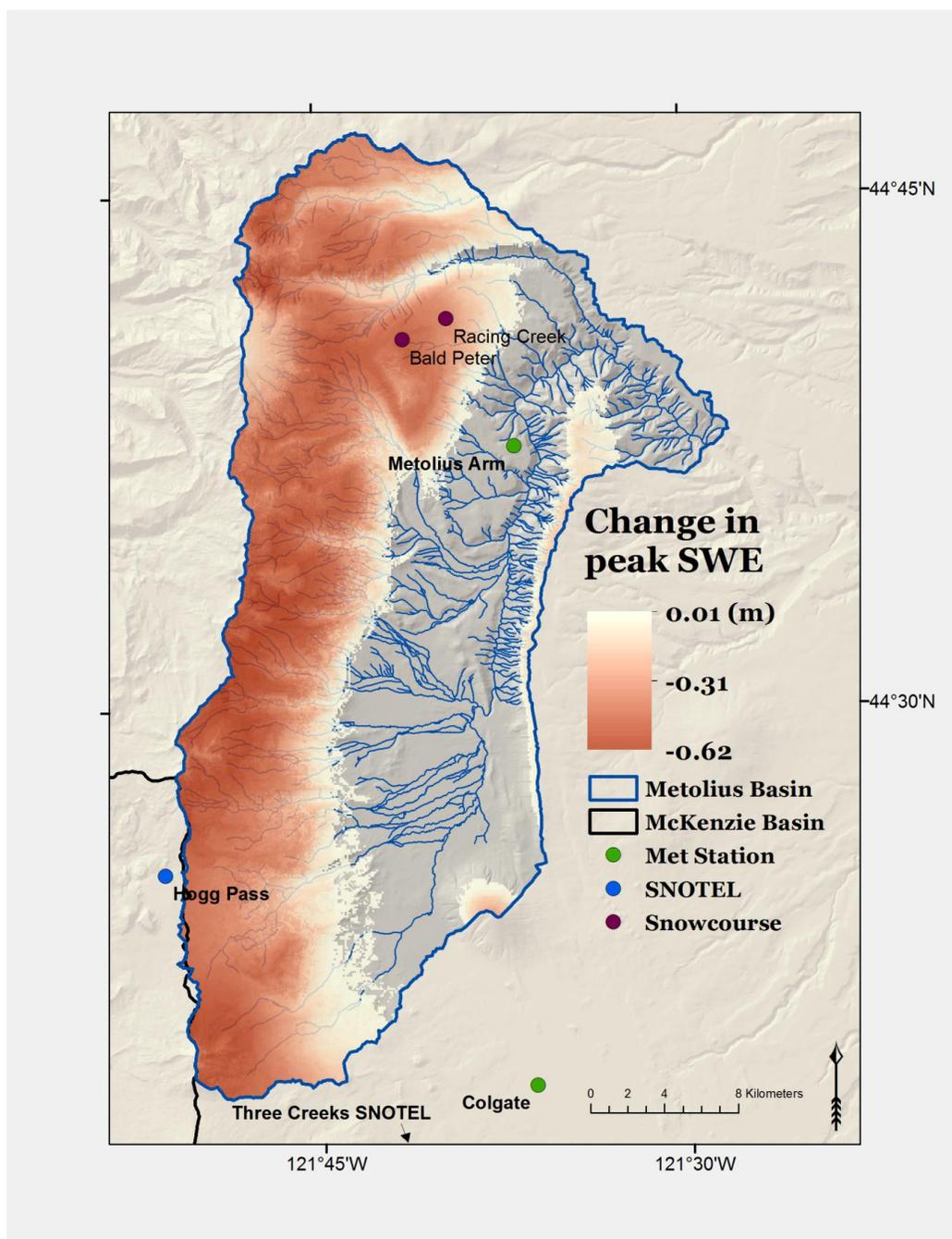
Appendix R: Simulated percent change in SWE on April 1 with a 4°C warming and +10% precipitation.



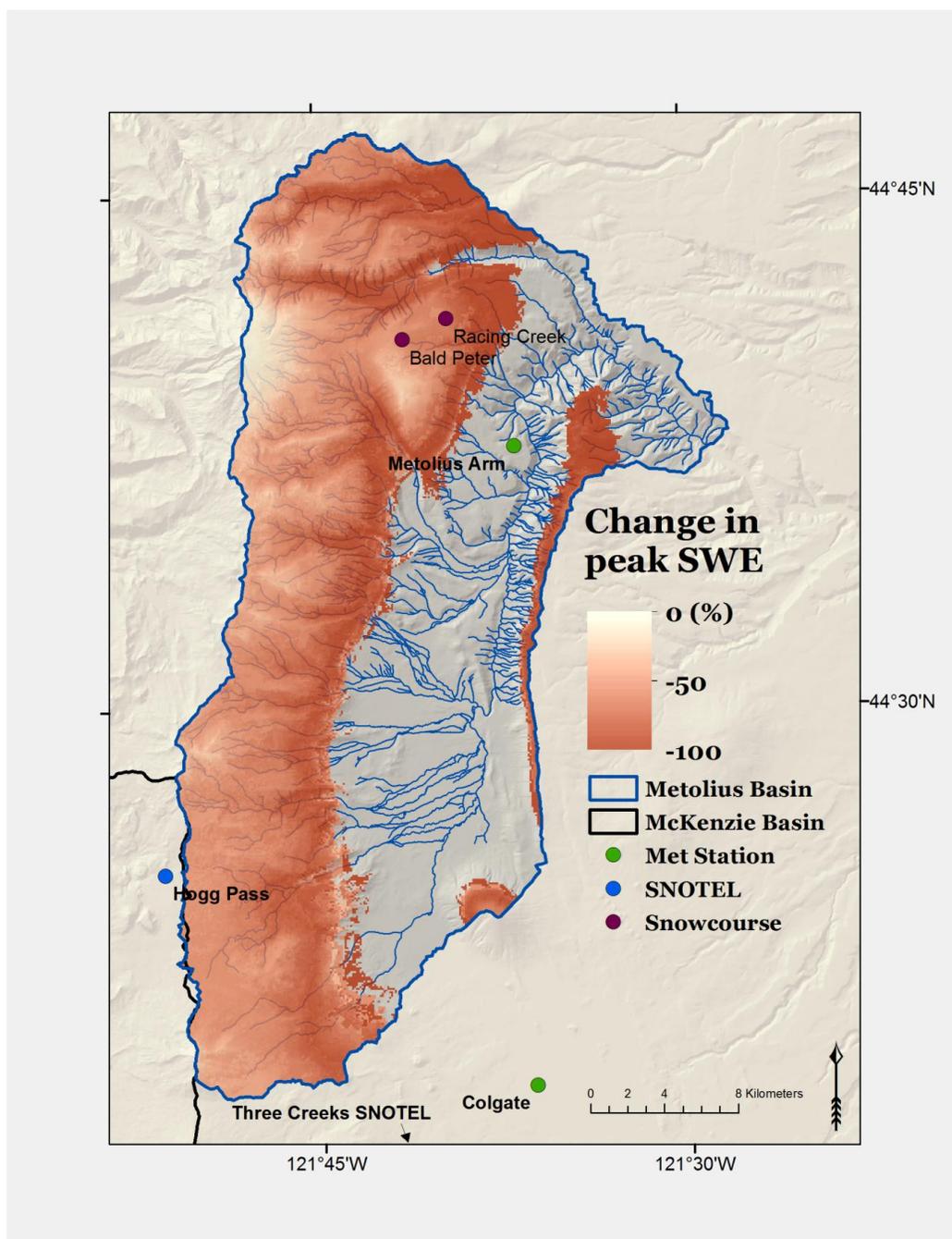
Appendix S: Simulated absolute change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +20C scenario.



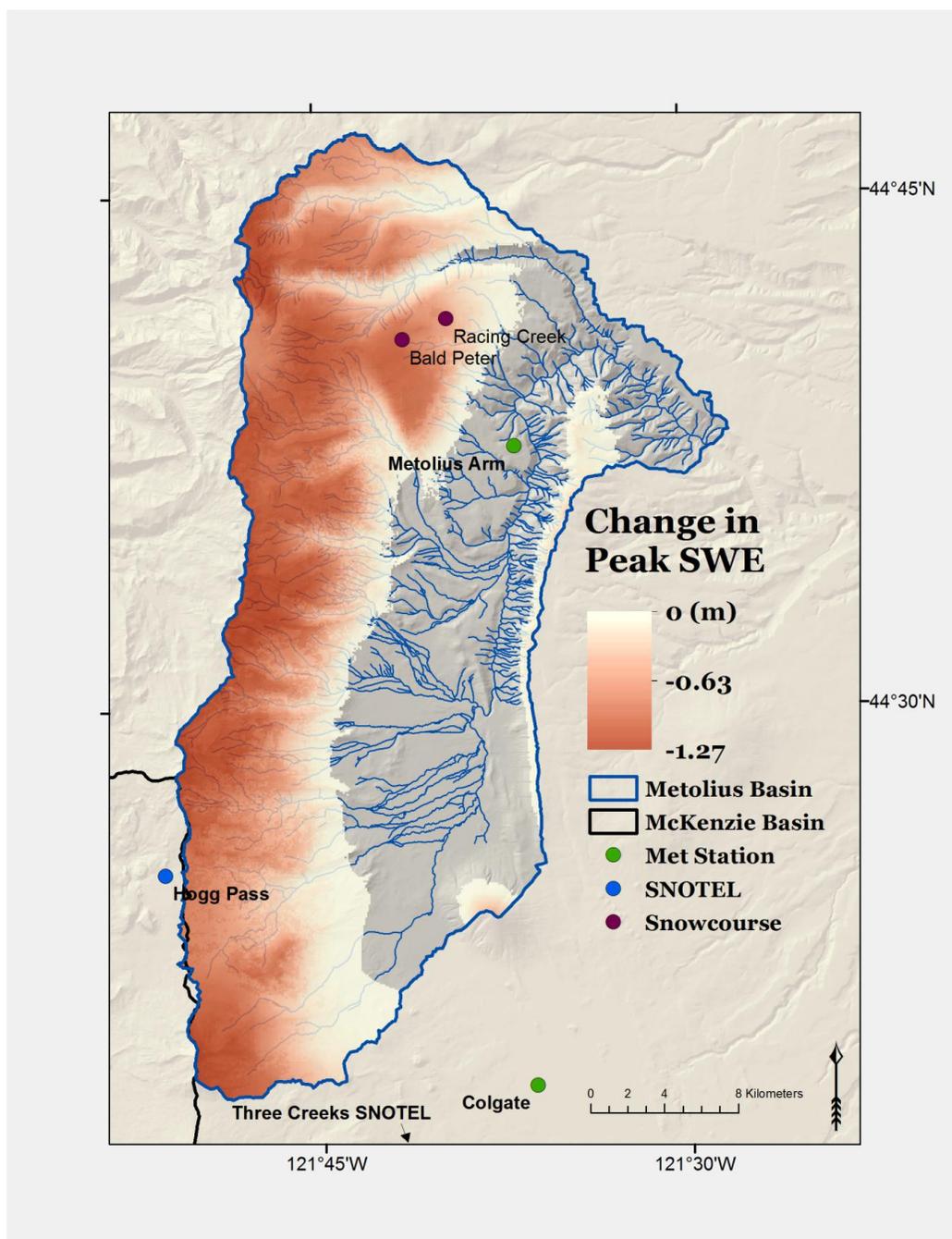
Appendix T: Simulated percent change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +2°C scenario.



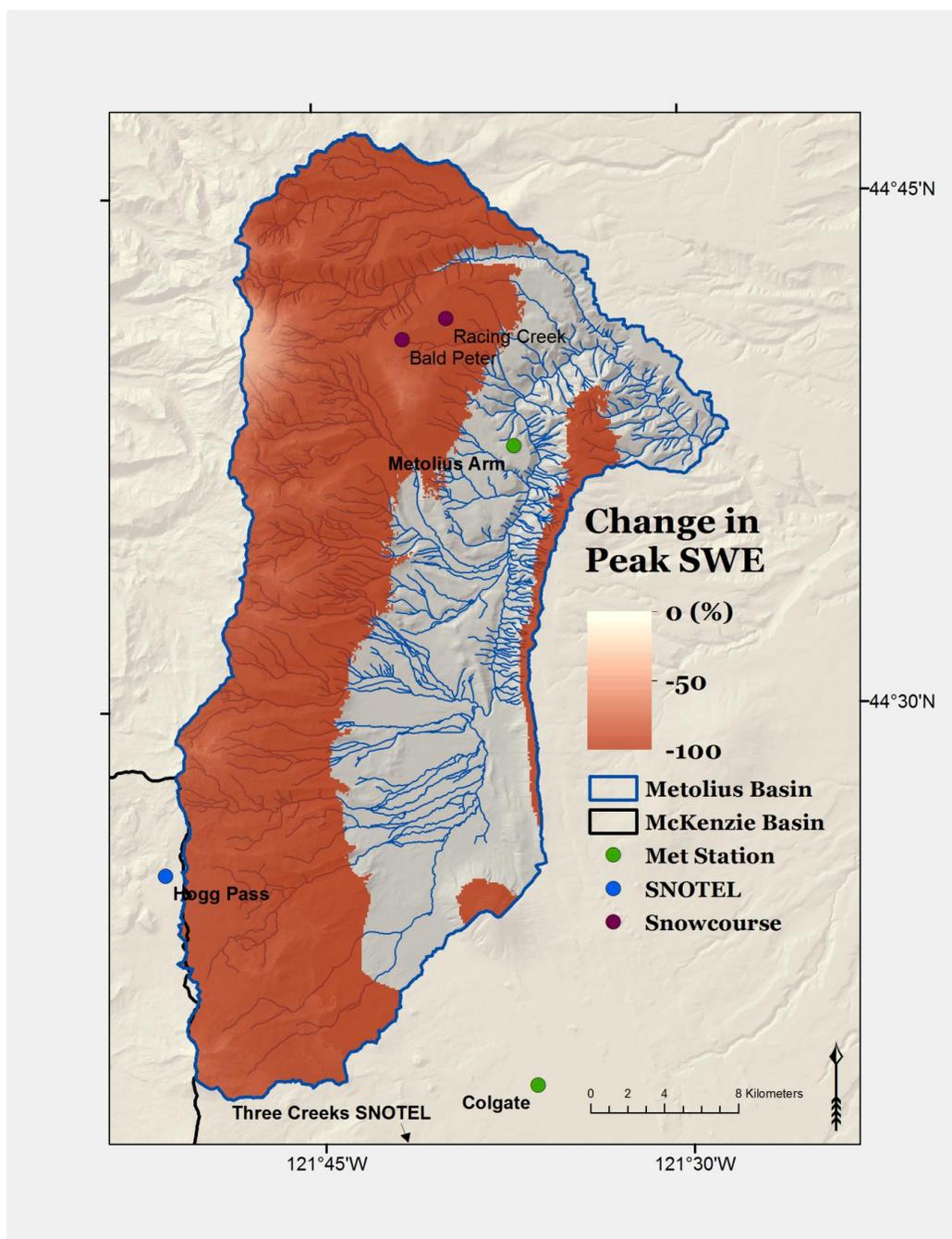
Appendix U: Simulated absolute change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +2°C with +10% precipitation scenario.



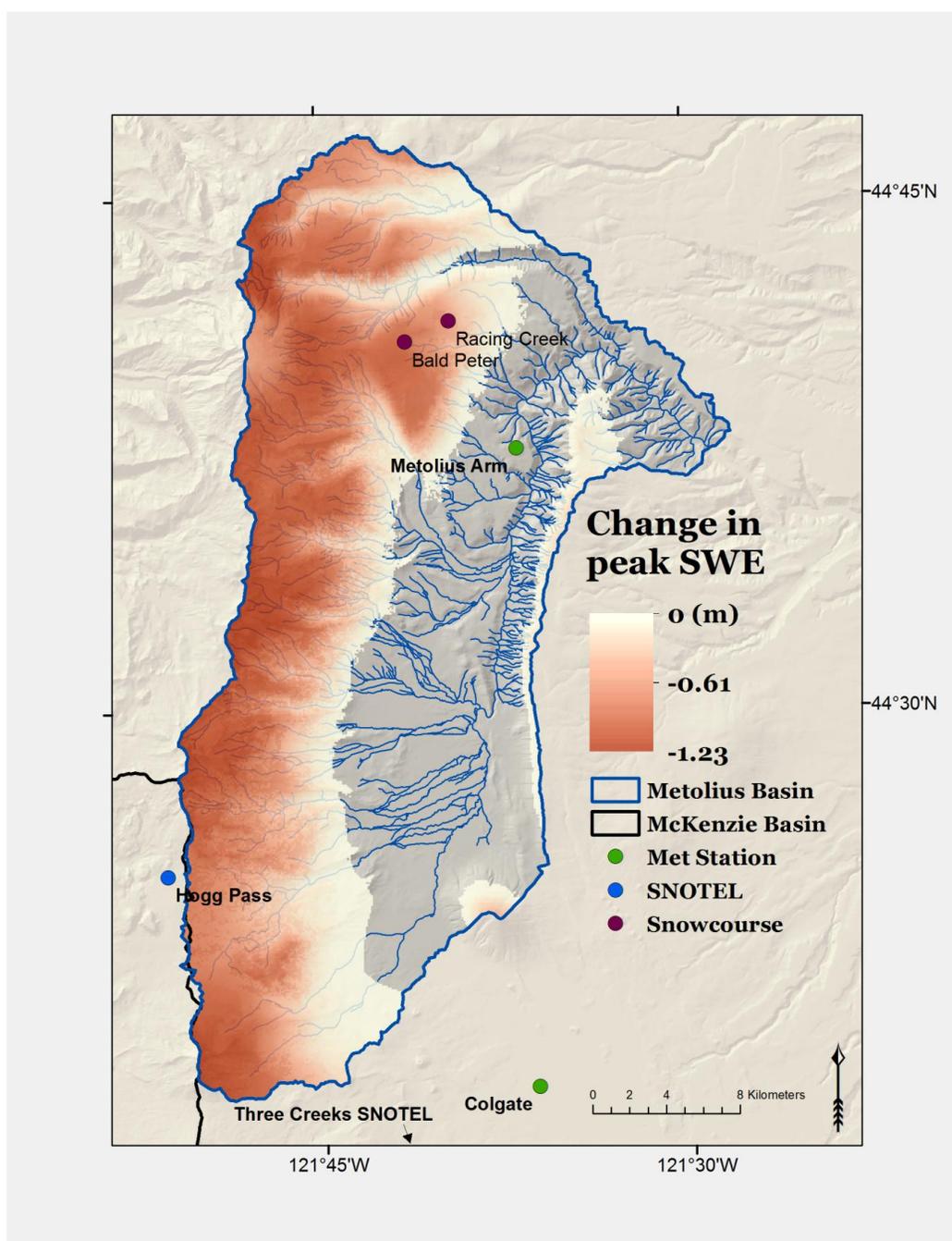
Appendix V: Simulated percent change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +2°C with +10% precipitation scenario.



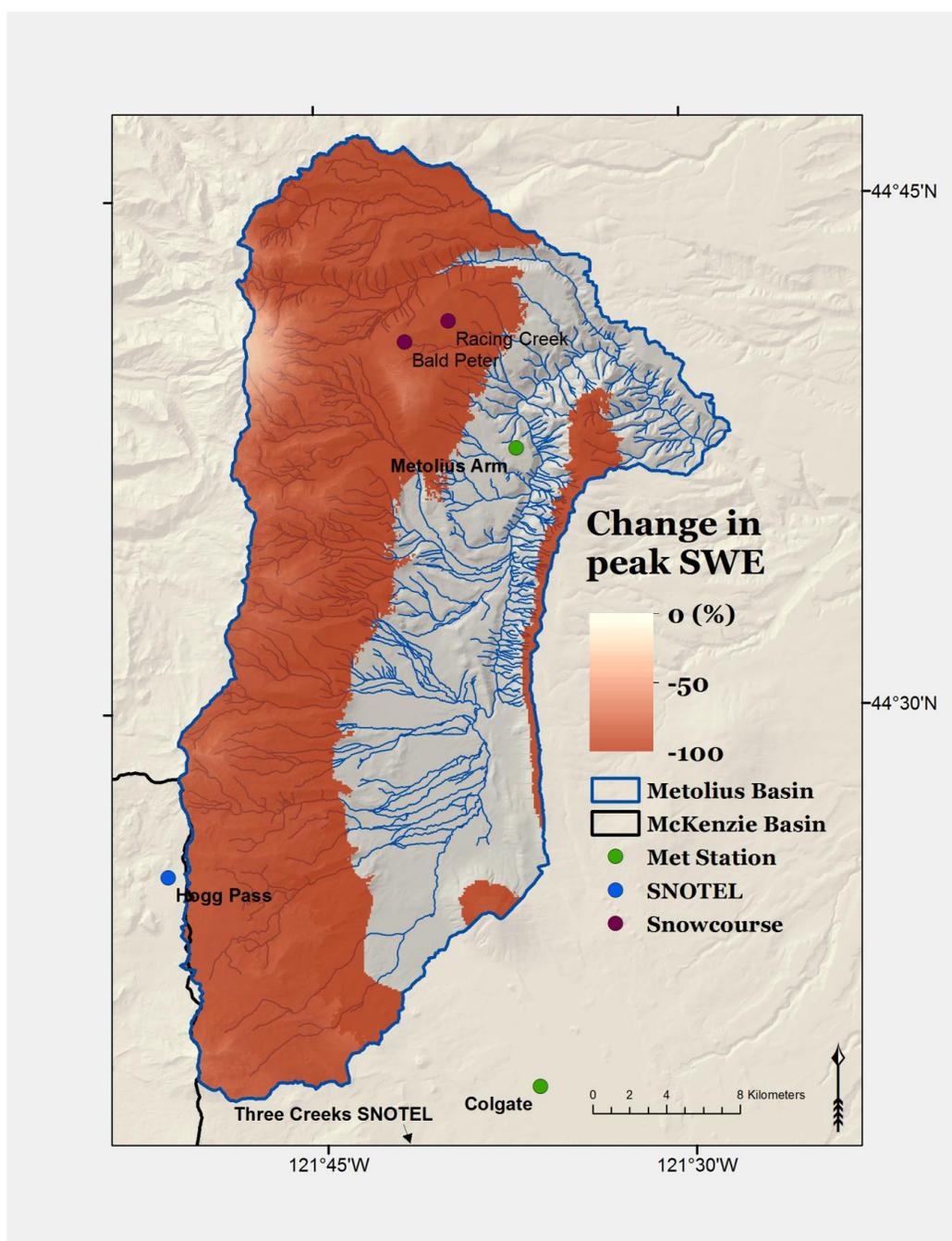
Appendix W: Simulated absolute change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +4°C scenario.



Appendix X: Simulated percent change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +4°C scenario.



Appendix Y: Simulated absolute change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +4°C with +10% precipitation scenario.



Appendix Z: Simulated percent change in peak SWE on the date of peak SWE for the reference period and the date of peak SWE for the +4°C with +10% precipitation scenario.

