Quantification of uncertainties in snow accumulation, snowmelt, and snow disappearance dates

Mark S. Raleigh

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Jessica D. Lundquist, Chair

Janneke Hille Ris Lambers

James Lutz

Program Authorized to Offer Degree:

Civil and Environmental Engineering

University of Washington

Abstract

Quantification of uncertainties in snow accumulation, snowmelt, and snow disappearance dates

Mark S. Raleigh

Chair of the Supervisory Committee:
Associate Professor Jessica D. Lundquist
Civil and Environmental Engineering

Seasonal mountain snowpack holds hydrologic and ecologic significance worldwide. However, observation networks in complex terrain are typically sparse and provide minimal information about prevailing conditions. Snow patterns and processes in this data sparse environment can be characterized with numerical models and satellite-based remote sensing, and thus it is essential to understand their reliability. This research quantifies model and remote sensing uncertainties in snow accumulation, snowmelt, and snow disappearance as revealed through comparisons with unique ground-based measurements.

The relationship between snow accumulation uncertainty and model configuration is assessed through a controlled experiment at 154 snow pillow sites in the western United States. To simulate snow water equivalent (SWE), the National Weather Service SNOW-17 model is tested as (1) a traditional "forward" model based primarily on precipitation, (2) a reconstruction model based on total snowmelt before the snow disappearance date, and (3) a combination of (1) and (2). For peak SWE estimation, the reliability of the parent models was indistinguishable,

while the combined model was most reliable. A sensitivity analysis demonstrated that the parent models had opposite sensitivities to temperature that tended to cancel in the combined model. Uncertainty in model forcing and parameters significantly controlled model accuracy.

Uncertainty in remotely sensed snow cover and snow disappearance in forested areas is enhanced by canopy obstruction but has been ill-quantified due to the lack of sub-canopy observations. To better quantify this uncertainty, dense networks of near-surface temperature sensors were installed at four study areas ($\leq$1 km^2) with varying forest cover in the Sierra Nevada, California. Snow presence at each sensor was detected during periods when temperature was damped, which resulted from snow cover insulation. This methodology was verified using time-lapse analysis and high resolution (15m) remote sensing, and then used to test daily 500 m canopy-adjusted MODIS snow cover data. Relative to the ground sensors, MODIS underestimated snow cover by 10-20% in meadows and 10-40% in forests, and showed snow disappearing 12 to 30 days too early in the forested sites. These errors were not detected with operational snow sensors, which have seen frequent use in MODIS validation studies.

The link between model forcing and snow model uncertainty is assessed in two studies using measurements at well-instrumented weather stations in different snow climates. First, representation of snow surface temperature (T_s) with temperature and humidity is examined because T_s tracks variations in the snowmelt energy balance. At all sites analyzed, the dew point temperature (T_d) represented T_s with lower bias than the dry and wet-bulb temperatures. The potential usefulness of this approximation is demonstrated in a case study where detection of model bias is achieved by comparing daily T_d and modeled T_s. Second, the impact of forcing data availability and empirical data estimation is addressed to understand which types of data most impact physically-based snow modeling and need improved representation. An experiment

is conducted at four well-instrumented sites with a series of hypothetical weather stations to determine which measurements (beyond temperature and precipitation) most impact snow model behavior. Radiative forcings had the largest impact on model behavior, but these are typically the least often measured.

DEDICATION

To Morgan, for her never-ending support, love, and encouragement.

"Mountains should be climbed with as little effort as possible and without desire.
The reality of your own nature should determine the speed.
If you become restless, speed up.
If you become winded, slow down.
You climb the mountain in an equilibrium between restlessness and exhaustion.

Then when you are no longer thinking ahead,
each footstep isn't just a means to an end but a unique event in itself.
This leaf has jagged edges.
This rock looks loose.
From *this* place the snow is less visible, even though closer.
These are things that you should notice anyway.

To live only for some future goal is shallow.
It's the sides of the mountain that sustain life, not the top.
Here's where things grow."

-- Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values*

ACKNOWLEDGEMENTS

I am indebted to my academic advisor Dr. Jessica Lundquist for her guidance, enthusiasm, and patience while I discovered and developed my research abilities and style during my graduate education. She has been a refreshing example of how one can be both passionate and scientific about a personal interest. She has also shown me there are always new questions to be asked and always more to be learned, and this is usually done best by getting out of the office and into the field. Thank you for giving me the opportunity to research mountain hydrology with you.

I would also like to thank my entire committee (Dr. Jessica Lundquist, Dr. Janneke Hille Ris Lambers, Dr. Jim Lutz, and Dr. Tom Ackerman) for dedicating time, energy, and interest in reviewing my research and for providing critical feedback from a diversity of perspectives.

Thank you to the students and researchers in the Mountain Hydrology Research Group at the University of Washington (UW) for reading many paper drafts, offering ideas for improvements, and for bringing tasty treats to our weekly meetings. I would like to specifically thank my fellow graduate students who did field work in the Sierra Nevada with me, including Courtney Moore, Brian Henn, and Nic Wayand. You were good company on our long hikes, and were good sports about the summer heat, winter snow, and the ever present threat of marauding black bears. Thanks also to Susan Dickerson-Lange for providing on-demand feedback of figures and text.

I would like to thank my past and ongoing collaborators outside of the UW Mountain Hydrology Group. I have had the pleasure of working with many individuals from other UW departments and from institutions outside of UW. There are far too many names to list here.

My graduate education would have been impossible without support from multiple organizations, and I am grateful for their support. Funding for my final year was provided by the Hydro Research Foundation (Department of Energy). I was also supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program – Grant NNX09AO22H. My education was partially funded by the Joint Institute for the Study of the Atmosphere and Ocean under NOAA Cooperative Agreement NA10OAR4320148, Contribution No. 7436, in association with the NOAA Hydrometeorological Testbed Project. I would also like to acknowledge support from a UW Valle Scholarship, a UW Nece Fellowship, a United States Society on Dams Scholarship, a CH2M HILL EWB-USA Scholarship, an AWWA Scholarship (Ameron International) and a fellowship from the AWRA-WA section.

I would like to acknowledge my family for setting the stage of this accomplishment. Thank you to my grandparents for settling in Colorado, where my fascination with mountains and snow began. Thanks to my parents for emphasizing the importance of quality education and for making many sacrifices along the way to ensure I had access to it. Thanks also to my siblings for sharing a love of snow and mountains. Thanks to the Forrey family for their curiosity in my work and for providing me with a home in the beautiful Pacific Northwest. Finally, I would like to thank my wife Morgan for patiently managing the inconveniences of me being in graduate school, including the late nights, the conference/field travel, and seemingly never ending stack of papers to read at home. I cannot thank her enough for her love and her confidence in me.

At the time of writing, two of the four chapters have been published in peer-refereed journals, and I would like to acknowledge *Water Resources Research* (John Wiley and Sons) and *Remote Sensing of Environment* (Elsevier, Inc.) for granting permission to reproduce those papers in this dissertation. The formal acknowledgements are:

Reprinted from *Water Resources Research*, *48*(1), Raleigh, M. S., and J. D. Lundquist, Comparing and combining SWE estimates from the SNOW-17 model using PRISM and SWE reconstruction, 1-16, 2012, with permission from Wiley.

Reprinted from *Remote Sensing of Environment*, *128*, Raleigh, M. S., K. Rittger, C. E. Moore, B. Henn, J. A. Lutz, and J. D. Lundquist, Ground-based testing of MODIS fractional snow cover in subalpine meadows and forests of the Sierra Nevada, 44-57, 2013, with permission from Elsevier.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

The "water towers of the world" [*Viviroli et al.*, 2003] are mountainous regions where orographically enhanced snowfall [*Roe*, 2005] and the resulting seasonal snowpack permit storage of vast quantities of water relative to the lowlands [*Viviroli et al.*, 2007]. Of interest are the three "life-cycle" stages of a seasonal snowpack, that is, (1) accumulation, (2) melt, and (3) disappearance, which influence downstream water resources (e.g., hydropower, water supply), watershed hydrology, and ecology. About one-sixth of the global population lives where mountain snowpack supplements reservoir storage and where snowmelt is the dominant contributor to annual streamflow [*Barnett et al.*, 2005]. Melting and disappearance of the snowpack during also marks the beginning of the growing season for many local ecological communities, which may impact species distribution [e.g., ribbon forests, *Billings*, 1969], diversity [*Litaor et al.*, 2008], and productivity [*Trujillo et al.*, 2012]. Quantifying uncertainty in tools that represent snow patterns and processes is therefore vital for understanding how the "water towers" work and for projecting the range of hydrologic responses to changes in climate and land cover. However, uncertainties in snow distributions and melt drivers are not often well understood because of high spatial variability in snow [*Scipión et al.*, 2013] and because ground-based observations are sparse in complex terrain [*Bales et al.*, 2006; *Lundquist et al.*, 2003].

The goal of this dissertation is to provide new insights into the utility of models and remote sensing to capture the magnitude of snow accumulation and melt and the timing of snow disappearance through quantification of errors. With advances in high-performance computing and the advent of satellite-borne remote sensing, the last three decades have witnessed continual refinements in modeling techniques and remote sensing of seasonal snow characteristics across different spatial and temporal scales [*Nolin*, 2010]. These tools have augmented the sparse

ground observations to provide more complete spatial and temporal coverage, and permitted predictions of future conditions. *Dozier* [2011] suggested that the snow research community now has the ability to exploit "the fourth paradigm" [*Hey et al.*, 2009], where large datasets from models and remote sensing are leveraged to address science questions. This direction is worth pursuing, but advances in ground-based observations have not kept pace with those in models and remote sensing [*Bales et al.*, 2006]. Therefore, the potential emerges for large model and remote sensing datasets to be analyzed with only a vague sense of the embedded uncertainties.

Throughout this research, individual processes (i.e., snow accumulation, snowmelt, and disappearance) from models or remote sensing are isolated and compared to ground-based observations, which are assumed the most likely approximation of reality. However, it is critical to note that no single method – ground observations, models, and remote sensing – lacks uncertainty, and that these imperfect tools can only be compared to each other (Figure 1.1), which is often complicated by differences in scale [*Blöschl*, 1999]. Several possible sources of uncertainty are imbedded in the ground-based observations, ranging from measurement precision [*van den Broeke et al.*, 2004] to environmental errors [*Huwald et al.*, 2009; *Johnson and Marks*, 2004; *Sieck et al.*, 2007] to spatial representativeness [e.g., *Molotch and Bales*, 2006]. Uncertainty in the ground observations is considered via model sensitivity analysis and by sampling large populations of ground observations whenever possible.

A central insight guiding the research is that accumulation, melt, and disappearance of seasonal snow are strongly linked, and knowledge of two of these components permits estimation of the third [*Liston*, 1999]. Consequently, uncertainties in one or two of the components will propagate into estimation of the third. For example, using a traditional "forward" running snow model, uncertainty in modeled snow disappearance is the result of

uncertainties in snow accumulation and snowmelt [*Shamir and Georgakakos*, 2006]. Likewise, snow disappearance timing information and calculated snowmelt can yield retrospective estimates of maximum snow accumulation, in what has been called "snow water equivalent (SWE) reconstruction" [*Molotch and Bales*, 2005; *Rice et al.*, 2011]. However, work presented in this dissertation and elsewhere [*Slater et al.*, 2013] indicate that the uncertainty in SWE reconstruction, particularly the model forcing and parameters, must be carefully considered.

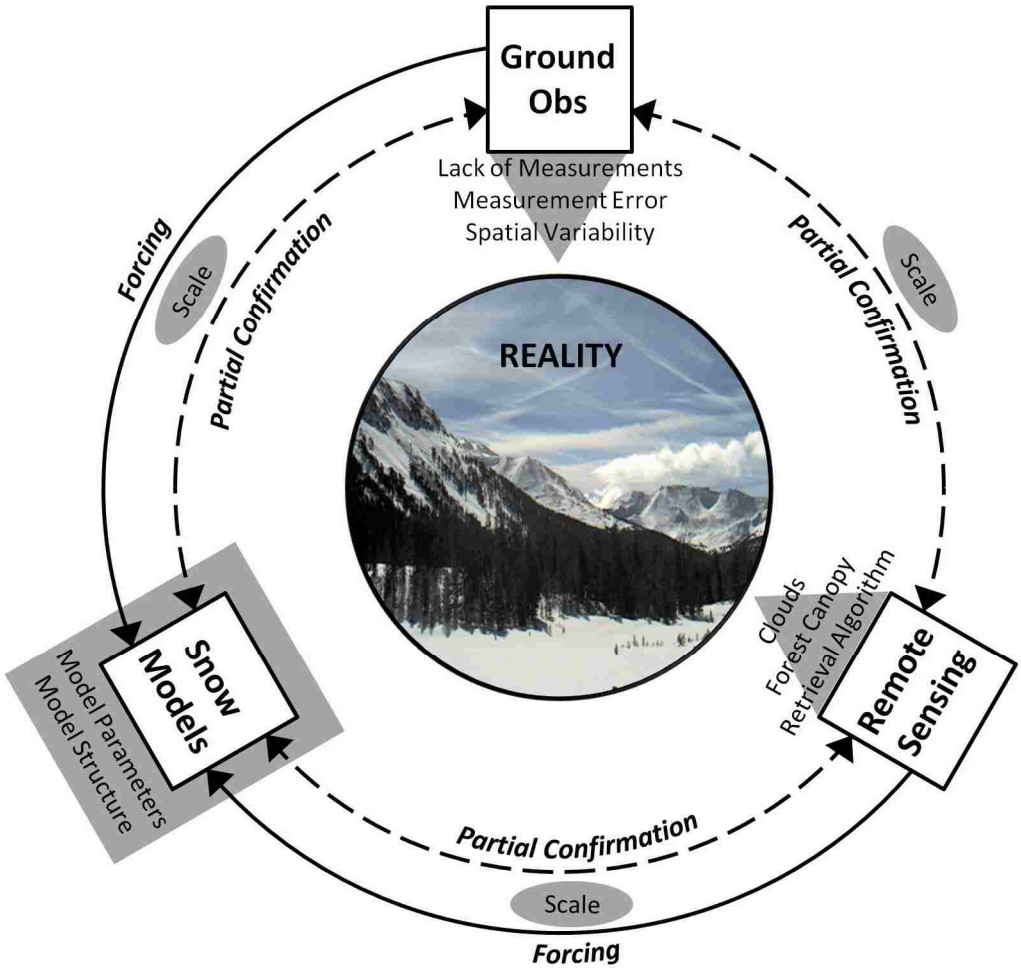

Figure 1.1 Representation of reality with ground observations, snow models, and remote sensing. Major sources of uncertainty are shown in the shaded areas. The gray triangles indicate that those methods sense reality, while the gray box indicates that models have no direct connection to reality. Opportunities for partial confirmation [*Oreskes et al.*, 1994] between methods and models are indicated. The photograph is at the Tioga Pass entrance to Yosemite National Park (courtesy of University of California San Diego).

A thorough understanding of the uncertainties in modeled and remotely sensed snow data is therefore critical to the successful use of these data. Through the course of this dissertation, a variety of unique approaches are employed in order to achieve this overarching theme. Where the data are available, the impact of snow climate and forest canopy on the uncertainty is considered. Four research questions motivated this research:

1. Which configuration (i.e., forward, reconstruction, or a combined forward-reconstruction approach) of a snow model yields the most robust estimates of SWE and precipitation, given uncertainties in model forcing data and parameters?

2. What is the accuracy of snow disappearance timing derived from MODIS fractional snow covered area at fine spatial scales (~1-2 km^2) across sites with varying forest density and topography?

3. Can standard height temperature and humidity approximate snow surface temperature, so as to allow detection of bias in a snowmelt model?

4. Given the scarcity of meteorological stations in mountains areas, which meteorological forcings are most critical to measure for physically-based modeling of snow in different climates?

The following chapters seek to address these four questions.

Chapter 2 [*Raleigh and Lundquist*, 2012] tests how different configurations of the same snow model impact snow accumulation simulations to address Question 1. This chapter primarily concerns uncertainties in snow accumulation using data from a network of 154 snow pillow sites in the western U.S. for model forcing and evaluation. SWE is simulated with three different configurations of SNOW-17 [*Anderson*, 1976], the operational snowmelt model of the National Weather Service River Forecast Center. Snow disappearance and air temperature (the primary forcing for snowmelt calculation) are assumed to be known, such that uncertainties in precipitation and maximum SWE accumulation can be quantified. Sensitivity to forcing data and model (i.e., structure, calibration, parameters) uncertainty are also explored and discussed.

Chapter 3 [*Raleigh et al.*, 2013c] develops a ground-based method for sampling snow covered area with networks of distributed near-surface temperature sensors, and then uses this approach to quantify errors in remotely sensed snow cover and snow disappearance in forest and meadow sites to address Question 2. The investigation takes place during water years 2010 and 2011 in the Sierra Nevada, California (USA). Fractional snow covered area data from the physically-based MODIS Snow Covered Area and Grain Size [MODSCAG, *Dozier and Frew*, 2009; *Dozier et al.*, 2008; *Painter et al.*, 2009] are corrected with a commonly employed forest canopy correction, and then compared with the unique ground-based datasets from dense networks of temperature sensors. This study provides unique quantification of forest effects on remote sensing accuracy, and reports and discusses errors in remotely sensed snow cover and snow disappearance that cannot be quantified with existing observational networks.

Chapter 4 [*Raleigh et al.*, *submitted* 2013a] addresses Question 3 through an investigation of how standard-height temperatures (i.e., dry-bulb, wet-bulb, and dew point) compare to measured snow surface temperature (T_s) at seven different study sites. The motivation of this study is to understand which standard temperature provides the most representative approximation (i.e., least bias) of T_s, which is strongly linked to the surface energy balance of the snowpack. The study is therefore related to the uncertainty in snowmelt, as snowmelt is governed by energy exchanges at the snow surface. After assessing how the standard temperatures compare to T_s and how these relationships change with atmospheric conditions and climate, the study demonstrates the potential utility of improved surface temperature approximation for detecting energy bias in snowmelt modeling.

Chapter 5 [*Raleigh et al.*, *in prep* 2013b] addresses Question 4 by evaluating the response of a physically-based snow model to data uncertainty and availability scenarios commonly

encountered. The study's goal is to identify the next "best" sensor to install at a weather station that already measures temperature and precipitation, and to identify which forcings may be adequately approximated with common methods. A series of data availability scenarios are generated (i.e., "hypothetical weather stations") and used to force a physically-based snow model. The data availability scenarios employ common empirical methods for estimating missing forcings. The sophisticated SNTHERM model [*Jordan*, 1991] is used to simulate snowpack at four sites from contrasting snow climates. This experiment provides insights into the importance of forcing data, and aims to provide guidance for future research.

When viewed from the context of global climate change, the research is particularly relevant. Whether estimating historic distributions of snow or projecting future changes in snowpack, a ubiquitous challenge is the uncertainty in model forcing. Modeling studies [e.g., *Elsner et al.*, 2010] suggest that seasonal snowpack in regions with mild winter temperatures, such as the Pacific Northwest (USA), is particularly sensitive to increasing temperature due to climate change. These changes in mountain hydrology have direct implications for irrigated agriculture [*Vano et al.*, 2010], hydropower [*Hamlet et al.*, 2002], aquatic habitat [*Cristea and Burges*, 2009], and the winter recreation industry [*Nolin and Daly*, 2006]. Ecological communities are also expected to shift or expire with earlier snow disappearance timing, although the ecological response may be complex because the spatial variability of snow results in microclimates that may create a buffer for some species [*Ford et al.*, 2013]. Because numerical models are used to inform natural resource managers and decision makers of quantitative changes in future snowpack and snowmelt in the "water towers", it is imperative that model uncertainties during the historic period are assessed thoroughly from a diversity of approaches.

Chapter 2 Comparing and combining SWE estimates from the SNOW-17 model using PRISM and SWE reconstruction

This chapter has been published in its current form in *Water Resources Research* [*Raleigh and Lundquist*, 2012]. Permission has been granted for reproduction in this dissertation. Sections, figures, tables, and equations have been renamed here, and some citations (e.g. papers in press) may have been updated.

Abstract

Snow models such as SNOW-17 may estimate past snow water equivalent (SWE) using either a forward configuration based on spatial extrapolation of measured precipitation, such as with the Parameter Regressions on Independent Slopes Model (PRISM), or a reconstruction configuration based on snow disappearance timing and back-calculated snowmelt. However, little guidance exists as to which configuration is preferable. Because the two approaches theoretically have opposite sensitivities to model forcing, combining (averaging) their SWE estimates may be advantageous. Using 154 snow pillow sites located in maritime mountains of the western United States, we compared forward, reconstruction, and combined configurations of a simplified SNOW-17. We evaluated model errors in (1) annual precipitation and (2) peak SWE, and (3) mean SWE errors during the accumulation and ablation seasons. We also conducted a separate analysis to assess the sensitivity of peak SWE to biased forcing data and parameters. The forward model had the greatest precipitation accuracy, while the combined model had the greatest accuracy in peak SWE and SWE during the accumulation and ablation seasons. In determining peak SWE, the forward and reconstruction models demonstrated opposite sensitivities to errors in air temperature and model parameters, and the combined model minimized errors due to temperature bias and parameter uncertainty. In basins with precipitation gauges, we recommend PRISM for precipitation estimation and the combined model for SWE estimation. In areas with high precipitation uncertainty, reconstruction is more viable. Accurate model parameters dramatically improved reconstruction, so more work is needed to advance parameter estimation techniques in complex terrain.

2.1 Introduction

Snow hydrologists often ask the fundamental questions, "What is the snow water equivalent (SWE) at an ungauged location in a mountainous basin, and how does it change in time?" These questions are important to hydrologists because understanding spatial distributions of SWE is essential for constructing depletion curves [*Homan et al.*, 2010; *Luce et al.*, 1999] which may be used to forecast seasonal runoff [*Rango and Martinec*, 1982]. The questions are also important for understanding seasonal snowpack interactions with fine-scale ecology. The

magnitude of peak SWE (and snow depth) is positively related to snowpack persistence [*Liston*, 1999], and these variables impact ecology, such as vegetation distributions [*Barbour et al.*, 1991], tree growth [*Littell et al.*, 2008], and wildlife habitat [*Millar and Westfall*, 2010].

The answers to the above questions remain elusive because of extreme spatial variability in snow properties [*Elder et al.*, 1989] and observational networks are sparse in many mountainous basins worldwide [e.g., California, USA: *Lundquist et al.*, 2003; British Columbia, Canada: *M Miles and Associates*, 2003; New Zealand: *Weingartner and Pearson*, 2001]. Furthermore, observation stations are typically located in flat clearings [*Farnes*, 1967], which often report systematically higher SWE than the surrounding area [*Grünewald and Lehning*, 2011; *Lee et al.*, 2005; *Molotch and Bales*, 2005]. Thus, the available observations are not likely to accurately represent the true spatial distribution of SWE, especially in mountainous basins with complex terrain and heterogeneous vegetation [*Blöschl*, 1999]. Researchers must either collect more snow data (e.g., field surveys, remote sensing) or use models forced by other information (e.g., terrain characteristics, precipitation, air temperature) to estimate SWE.

To expand SWE observations, intensive ground-based field surveys [e.g., *Cline et al.*, 2003; *Molotch and Bales*, 2005] have been conducted in relatively small areas over discontinuous time periods. These surveys are uncommon because they rely on intense manual labor operating in challenging terrain. Observations of SWE or snow depth have also been provided by remote sensing, such as laser scanning technology [*Grünewald et al.*, 2010; *Prokop et al.*, 2008] or scanning microwave radiometers [e.g., *Dahe et al.*, 2006]. Despite the promise of these remote sensing instruments, limitations remain. Laser scanning technology is not routinely employed in most basins and microwave radiometers observe SWE in large footprints (e.g., 0.5°), which are too coarse to resolve SWE variability in many basins. Passive remote sensing

instruments (e.g., USGS/NASA Landsat, NASA MODIS, and ESA MERIS) provide global observations of SCA but do not observe SWE or snow depth.

Because of these challenges, researchers are left to model SWE. To interpolate and extrapolate SWE observations, studies around the northern hemisphere have utilized many techniques, such as multivariate statistical methods [e.g., *Anderton et al.*, 2004], probabilistic approaches[e.g., *Skaugen*, 2007], masked interpolation methods [e.g., *Fassnacht et al.*, 2003], global interpolation models [*López-Moreno and Nogués-Bravo*, 2006], and regression-trees based on terrain characteristics [e.g., *Elder et al.*, 1998]. Methods that rely on snow observations alone may not accurately model SWE in areas above the highest observations [*Rice et al.*, 2011], so a deterministic snow model [e.g., *Anderson*, 1976] that simulates SWE time series based on meteorological data (e.g., air temperature, precipitation) may be preferred. Two different configurations of the same snow model may be employed to simulate SWE with time.

In the first configuration, off-line estimates of precipitation are input into the snow model, which partitions the precipitation at each time step into rain and snow (typically with a threshold air temperature), and stores snowfall accumulation as SWE. The model reduces SWE when environmental conditions favor snowmelt. This precipitation-driven approach is common, and we refer to it herein as the "forward" model (Figure 2.1a). If a gauge network exists nearby, the off-line estimates of precipitation can be estimated with analytic mapping models (such as the Parameter-elevation Regressions on Independent Slopes Model, PRISM [*Daly et al.*, 1994]), multivariate regression [*Marquínez et al.*, 2003], kriging[*Garen and Marks*, 2005], inverse-distance weighting [*Gemmer et al.*, 2004], or truncated Gaussian filters [*Thornton et al.*, 1997]. When few precipitation gauges are available in a mountainous basin, which is not uncommon,

uncertainty increases in precipitation inputs [e.g., *Tsintikidis et al.*, 2002] and this uncertainty will propagate through the forward model [*Shamir and Georgakakos*, 2006].

With the second configuration, SWE is reconstructed by summing modeled snowmelt backwards in time starting from the date of snow disappearance, in order to estimate how much SWE must have existed before snowmelt commenced (Figure 2.1b). Herein, we refer to this method as "reconstruction" [*Molotch and Bales*, 2005], though it has also been called the "depletion method" [*Cline et al.*, 1998; *Rice et al.*, 2011]. Several members of the snow science community (Appendix A) have selected reconstruction over forward modeling for various reasons. First, reconstruction does not require precipitation observations to estimate peak SWE in years when no snow accumulates after peak SWE. At a minimum, reconstruction requires air temperature to calculate snowmelt, and air temperature may be more reliably estimated than precipitation [*Ninyerola et al.*, 2000]. Secondly, reconstruction incorporates observed snow disappearance timing (Figure 2.1b), which provides additional information because it is correlated with peak SWE and ablation season melt rates [*Liston*, 1999]. Thirdly, snow disappearance timing observations are available for most basins worldwide from SCA products derived from passive remote sensing imagery (e.g., Landsat, MODIS, and MERIS). Thus, reconstruction is possible in most basins worldwide, regardless of the availability of precipitation observations.

When estimating past SWE in locations that have nearby precipitation gauges, one must decide to use either a forward model or reconstruction, but there has been little research to guide this decision. Studies have independently examined the accuracy and sensitivity of forward models [*He et al.*, 2011a, 2011b; *Shamir and Georgakakos*, 2006] and reconstruction [*Rice et al.*, 2011; *Slater et al.*, 2013], but no study has examined these together. The critical premise of

reconstruction is that snowmelt can be modeled more accurately than snowfall with forward modeling, but this hypothesis has not yet been tested. Additionally, if these models have opposite sensitivities to model inputs or parameters (see Section 2.2), then a "combined model" that averages SWE from forward and reconstruction models may reduce the likelihood of SWE errors. No prior study has examined the hypothetical benefits of this type of combined model.

The purpose of this study is to compare the accuracy and sensitivity of three configurations (forward, reconstruction, and combined) of the same snow model, in order to provide guidance for selecting a model configuration to estimate SWE and precipitation. Two specific questions are addressed: (1) Which model configuration (forward, reconstruction, combined) is likely to produce the most accurate estimates of (a) annual precipitation, (b) peak SWE, and (c) SWE during the accumulation and ablation seasons? (2) How sensitive are the model configurations to biases in model data and parameters when estimating SWE?

To answer these questions, we employ a simplified version of SNOW-17 [*Anderson*, 1976] to calculate snow accumulation and ablation. We select a temperature-index model because it only requires air temperature and precipitation data; an energy-balance approach requires additional data (e.g., radiation, wind, humidity) which are not widely available in most mountainous basins. We use air temperature, precipitation, and SWE data from 154 snow pillows in the western U.S. to calibrate SNOW-17 and test the three model configurations. Although these flat, clearing sites may not be representative of zonal or basin SWE, we use them because they readily provide a large pool of SWE, air temperature, and precipitation data that allow testing and comparison between the model configurations. In contrast, observations are rarely available along sloped terrain [*Pomeroy et al.*, 2003].

In testing each model configuration, we only assume that air temperature and snow disappearance timing data are known; the SWE data at each snow pillow site are used only for cross-validation and do not inform model calibration or the mass balance simulation at that site. For the forward model, we estimate precipitation using a PRISM precipitation map because PRISM has been shown to produce smaller errors and less bias than other techniques [*Daly et al.*, 1994]. The methodology was developed here to reflect the likely data available to a researcher who needs to estimate past SWE at ungauged locations.

2.2 Forward and reconstruction models: theory, limitations, and opportunities

Daily SWE time series with a forward snow model (Figure 2.1a) can be generalized as:

$$\overrightarrow{SWE_n} = \sum_{t=1}^{n} A_t - \sum_{t=1}^{n} M_t \tag{2.1}$$

where n is the n^{th} day of the water year (1 Oct – 30 Sept), A_t is estimated daily snowfall accumulation, M_t is estimated daily snowmelt, and t is the timestep (days). At a minimum, forward models require precipitation and air temperature data, as well as model parameters for rain-snow threshold temperatures, snowmelt temperatures, and snowmelt factors.

In contrast, a SWE time series from a reconstruction model (Figure 2.1b) is:

$$\overleftarrow{SWE_n} = \sum_{t=d}^{n} M_t - \sum_{t=d}^{n} A_t \tag{2.2}$$

where d is the date of snow disappearance (also known as the snow depletion or snow-free date, or the final day of seasonal snow cover), and n is the day of interest, which, by definition, must be before date d. The reconstruction model runs in reverse from date d to date n (Figure 2.1b). Because the model runs in reverse during this time domain, calculated snowmelt increases reconstructed SWE while snow accumulation decreases reconstructed SWE. At a minimum, reconstruction requires air temperature data and the snow disappearance date, as well as model

parameters for snowmelt temperatures and snowmelt factors. Precipitation data and rain-snow threshold parameters are required by reconstruction when snow storms occur after day *n* (to the right of t=n in Figure 2.1b). Many reconstruction studies have assumed no snowfall (i.e., A_t=0) after a specified date in March or April (see Appendix A). This assumption was not made in this study because total observed snow accumulation during the melt season was typically 10% of peak SWE at the study sites, and in extreme cases, exceeded 35% of peak SWE.

Forward models are intuitive because they operate in the same direction as time, and are practical because they can simulate past, current, or future conditions. Reconstruction models are retrospective and can only simulate past SWE because they first require the snow to disappear. Despite this limitation, reconstruction may yield the only viable estimate of past SWE when precipitation distributions are unknown because it does not require precipitation data. When using a temperature-index model, reconstruction will have comparable or improved accuracy relative to a forward model, given accurate air temperature data, low uncertainty in snow disappearance dates, and appropriate model parameters (see Section 2.4).

In practice, both models are impacted by errors in data and parameters, and these are likely to impact SWE estimation. At ungauged locations, air temperature is commonly estimated with a lapse rate, which may introduce error when not based on regional observations [*Minder et al.*, 2010]. *Slater et al.* [2013] reviewed a variety of studies that estimated air temperature and summarized that errors of 1°C were typical across mid-elevations, and errors exceeding 1.5°C were common when extrapolating data to high elevations. Precipitation biases of 10-50% may occur when windy conditions cause gauge undercatch [e.g., *Goodison et al.*, 1998], while additional errors may arise during interpolation or extrapolation of gauge data across a basin (e.g., PRISM or kriging). Snow cover from remote sensing may be inaccurate for various

reasons, such as cloud cover, missing scenes, concealment by forest canopy, or if the spatial scale of interest is smaller than the satellite instrument's footprint [*Dozier et al.*, 2008; *Slater et al.*, 2013]. For temperature-index models, uncertainty in model parameters at ungauged locations may be large [*He et al.*, 2011a].

In theory, forward and reconstruction models should exhibit varying sensitivities to model data and parameters when estimating peak SWE (Table 2.1). One key observation to note is that the models have opposite sensitivities to air temperature and to the model parameters, though the magnitudes of these sensitivities should vary. Additionally, the bias in model parameters depends on location and seasonal climate. For example, forward model estimates of peak SWE are only sensitive to the snowmelt parameters in years when melt occurs before peak SWE, while reconstruction is sensitive to rain-snow threshold parameters only when storms bring rain and snow after peak SWE. Bias in air temperature may impact the forward model through rain-snow partitioning [e.g., *Minder et al.*, 2010; *Moore and Owens*, 1984], and reconstruction through snowmelt rates [e.g., *Minder et al.*, 2010; *Richard and Gratton*, 2001]. *Minder et al.* [2010] found that air temperature estimates with a warm bias may result in less snowfall and greater snowmelt rates.

Therefore, forward and reconstruction models should have opposite sensitivities to accumulation and ablation processes, presenting an opportunity where averaging their peak SWE estimates may minimize the impact of biased data and/or parameters. We hypothesize that this type of combined model is more likely to produce smaller SWE errors than either forward or reconstruction models because averaging the opposing sensitivities will reduce the overall error.

2.3 Data

2.3.1 Observational sites and quality control

The study was conducted at 154 snow pillows (Figure 2.2a) in the western U.S. maritime ranges, including the Cascades of Washington and Oregon, the California Sierra Nevada, the Blue Mountains of Oregon, the Pacific Northwest Coastal Range, and the California Klamath Range. In these ranges, 85-95% of annual precipitation arrives between October and May [*Baker*, 1944]. In the subalpine areas of these ranges, 50-67% of annual precipitation is snowfall [*Serreze et al.*, 1999], while 90-100% of alpine precipitation falls as snow [*Kattelmann and Elder*, 1991; *Smith and Berg*, 1982]. Site elevations ranged from 685 m (snow transition zone) to 3475 m (alpine zone); 42% of the study snow pillows were located below 1600m.

All study sites in Washington, Oregon, and California (east of the Sierra Nevada crest) were drawn from the Natural Resources Conservation Service (NRCS) snowpack telemetry (SNOTEL) network (http://www.wcc.nrcs.usda.gov/snow/). All California sites west of the Sierra Nevada crest were selected from the California Department of Water Resources (CDWR) network (http://cdec.water.ca.gov/), managed by the California Cooperative Snow Surveys. Observations between water years 1996-2004 were considered in Washington and Oregon, and between water years 1996-1998 in California; selection of these years was arbitrary. Peak SWE observations ranged from 75mm to 2540 mm and annual precipitation ranged from 330mm to 5275mm over the study. All sites had daily observations of mean air temperature and SWE. Precipitation data were available at all analysis sites except at 26 of the CDWR sites. Eight additional CDWR precipitation gauges (Figure 2.2a, red stars) augmented the California data pool. At sites with precipitation data, undercatch correction was not attempted because wind speed data (not typically available) are required to correct undercatch for storage gauges [*Sevruk*, 1983], which are standard at SNOTEL sites.

Daily air temperature, precipitation, and SWE data were quality controlled following *Meek and Hatfield's* [1994] framework. Quality control flags were placed when values exceeded limits (Table 2.2), and flagged data were either accepted or rejected based on visual inspection. Temperature and precipitation time series from individual stations were also compared to observations at neighboring stations to find anomalies. A station-year (i.e., one water year of data at a single station) was discarded if there were four or more consecutive days of missing or flagged data of any one variable during the observed snow season. Station-years were also discarded if the snow disappearance date could not be determined from the SWE data. Data gaps of less than four days in the temperature and SWE data were filled with interpolation from data immediately before and after each gap, while no precipitation was assumed during these gaps.

After quality control, 388 station-years remained. 54 station-years (i.e., 18 stations, 3 water years each) were isolated for snow model calibration (see Section 2.4) and were not used in evaluation statistics. This left 334 station-years (136 snow pillow sites) of data available for the evaluation. 40 of these station-years were at the CDWR sites which lacked precipitation data.

2.3.2 PRISM data

Output data from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) [*Daly et al.*, 1994] were used to estimate precipitation for the forward model (see Section 2.5). PRISM was selected because it is used widely to map precipitation in hydrologic and ecological models. PRISM divides a digital elevation model into topographic facets based on slope orientation and coastal proximity. For each topographic facet in a region, PRISM develops elevation-based regressions with gauge observations, and estimates monthly and annual grids of precipitation based on those regressions. *Daly et al.* [1994] reported mean absolute

errors (MAE) of 10-17% in PRISM annual precipitation when using a 52 station network in Oregon.

To estimate the 1971-2000 climatology of annual precipitation in the conterminous United States, the PRISM climate group (www.prism.oregonstate.edu) created a 30-arcsec (800 m) "normals" product (Figure 2.2b) from observations at over 13,000 stations [*Daly et al.*, 1994, 2008]. This product incorporates historical precipitation data or snow course data from most major observational networks (e.g., SNOTEL, CDWR). Data from 114 of the 136 analysis sites were used to produce the normals product [*C. Daly*, personal communication, 2010]. An annually varying 2.5-arcmin (4 km) PRISM "analysis" product is also available; we compared the two PRISM products but found that the normals product yielded improved results in our study. Therefore, we only used the mean annual precipitation data from the 800m normals product (henceforth called PRISM).

2.4 Snow accumulation and melt model

The snow accumulation and melt model used in this study was SNOW-17 [*Anderson*, 1976]. SNOW-17 is a single layer, temperature-index snow model used operationally by the National Weather Service (NWS) for flood forecasting. SNOW-17 estimates SWE and outflow (snowmelt + rain) at each time step. Although many past reconstruction studies (Appendix A) used snow models that required net radiation to calculate snowmelt, our study sites generally lacked radiation observations. Accordingly, we selected SNOW-17 because it required data inputs (e.g., air temperature) that were available at our study sites, and because it has been shown to simulate snowmelt as well as energy balance methods in some studies [*Franz et al.*, 2008].

We simplified the SNOW-17 model for the sake of computational efficiency. Whereas the full NWS SNOW-17 has 10 model parameters, our simplified version has only five. In the

simplified version, the rain-on-snow melt, light rain melt, ground heat melt, liquid water holding capacity, and heat deficit components of the full NWS SNOW-17 were deactivated. If the heat deficit and liquid water holding routines were activated, reconstruction would have required an iterative solution [*Raleigh*, 2009] because these routines are dependent on snowpack conditions before that time step. With these routines deactivated, the simplified SNOW-17 could reconstruct SWE without multiple iterations. We repeated the analysis using iterations with the full NWS model, but found the results of the two versions were not significantly different at our study sites. Thus, only the results from the simplified SNOW-17 were included here.

The five parameters used in the simplified SNOW-17 are listed in Table 2.3. Whereas the full NWS model has a single threshold temperature to distinguish rain and snow, we used two threshold temperature parameters (T_{snow} and T_{rain}) because mixed rain-snow storms are not uncommon at many of our maritime sites. In conjunction with mean daily air temperature (T_t), these two parameters were used to estimate the snowfall fraction (f_t) of daily precipitation at each snow pillow. All daily precipitation was assumed snow when T_t was less than or equal to the T_{snow} parameter, all was rain when the T_t was greater than or equal to T_{rain} parameter, and the precipitation was a linear rain-snow mixture between those two parameters [*U.S. Army Corps of Engineers*, 1956]. Snowfall was accumulated in the modeled snowpack while rainfall and melt water passed through the snowpack without being stored.

The three remaining model parameters (*MBASE*, *MFMIN*, and *MFMAX*) were used to calculate snowmelt. Snowmelt (M_t) on day t was calculated as [*Anderson*, 1976]:

$$M_t = MF_t \cdot (T_t - MBASE) \tag{2.3}$$

where, MF_t is the daily-varying melt factor (mm °C^{-1} day^{-1}), T_t is mean daily air temperature (°C), and *MBASE* is the minimum air temperature (°C) for snowmelt (no melt when $T_t \leq MBASE$).

MF_t is varied daily with a sinusoidal curve to reflect seasonal changes in net solar radiation, such that MF_t equals *MFMIN* on 21 Dec and *MFMAX* on 21 Jun [*Anderson*, 1976].

All five model parameters required calibration. 18 snow pillows were designated as calibration sites and were selected on a regional basis (every 1° of latitude, with separate stations on the east and west slopes, Figure 2.2a). The 18 calibration sites were excluded from the rest of the study. Each site was independently calibrated using three water years of data. At each calibration site, an optimization algorithm was used to find the single values of T_{snow}, T_{rain} and for *MBASE, MFMIN, and MFMAX* that produced the lowest root mean squared error (RMSE) in snowfall accumulation and snowmelt, respectively. The resulting calibrated parameters (Table 2.3) were comparable to values reported in other SNOW-17 studies [e.g., *Franz et al.*, 2008; *He et al.*, 2011a; *Shamir and Georgakakos*, 2006]. The study results exhibited sensitivity to the model calibration. This is demonstrated in Section 2.6.4 and discussed further in Section 2.7.

2.5 Methods

Forward, reconstruction, and combined configurations of the simplified SNOW-17 model were used to estimate annual precipitation and SWE at each study site (i.e., point scale). All sites had snow pillows and most had precipitation gauges, which allowed evaluation of each model configuration at each study site; local observations of SWE and precipitation were not used as input into the model or the calibration. Figure 2.3 shows the assumptions made regarding data availability at the study site (hereafter referred to as Site X) when estimating precipitation (P_X) and snow water equivalent (SWE_X). These assumptions are described below.

1.) To assess the models' applicability in locations without precipitation gauges, we assumed each study site (Site X) lacked a precipitation gauge and thus required estimation based on observations (P_Y) at the closest gauge, Site Y. To estimate daily precipitation at Site X, a precipitation multiplier (S) was used to uniformly increase or decrease P_Y, to account for accumulation differences between locations, due to effects such as orographic enhancement of precipitation [*Roe*, 2005].

2.) Observations of air temperature and snow disappearance timing were assumed available at Site X. The annual snow disappearance date was provided by the snow pillow at Site X and was assumed the first date with SWE=0 after peak SWE. We assumed these two observations were available because point values may be observed easily with distributed temperature sensors in applications outside of this study [e.g., *Lundquist and Huggett*, 2008; *Lundquist and Lott*, 2008].

3.) We assumed that the five snow model parameters from the nearest calibration station (Section 2.4) could be transferred to Site X and were constant from year-to-year. This assumption is tested in Section 2.6.5 to evaluate the errors associated with transferring model parameters from regional calibration stations to study sites.

4.) We assumed sublimation, wind transfer, and avalanches were negligible at Site X and thus did not require simulation. Model simulations in western Idaho suggest that sublimation is a minor component of the mass balance, with an expected magnitude of 3% of peak SWE during wet years and 10% during dry years [*Reba et al.*, 2011b]. Snow

pillows (i.e. Site X) are typically located in flat clearings where mechanical redistributions are minimal or non-existent [*Farnes*, 1967].

While the daily snowfall fraction and potential snowmelt were the same for each model configuration, the key differences were the precipitation multiplier (S), the direction of the model simulation (Figure 2.1), and the starting point of the forward and reconstruction models. These differences impacted modeled precipitation and SWE, and are described further below.

2.5.1 Forward model

In the forward model, the PRISM mean annual precipitation map (Figure 2.2b) was used to estimate the mean precipitation ratio between Site X and Site Y (Figure 2.3). For example, if PRISM showed mean annual precipitation of 1500mm at the Site X pixel and 1000mm at the Site Y pixel, then the multiplier used ($S_{PRISM,XY}$) to map daily observations from Site Y to Site X would be 1.5. This common methodology is used in distributed models [e.g., *Shamir and Georgakakos*, 2006; *Smith et al.*, 2004] and mountain microclimate models [e.g., *Running et al.*, 1987].

Snowfall accumulation at Site X on day t was estimated with the forward model as:

$$(A_{X,t})_{forward} = (P_{Y,t} \times f_{X,t}) \times S_{PRISM,XY} \quad (2.4)$$

where $P_{Y,t}$ was observed daily precipitation (mm) at the nearest offsite gauge (Site Y, Figure 2.3); $f_{X,t}$ was the snowfall fraction of precipitation at Site X, based on the transferred T_{snow} and T_{rain} parameters (see section 2.4); $S_{PRISM,XY}$ was the PRISM precipitation multiplier, $P_{PRISM,X}$ / $P_{PRISM,Y}$. $S_{PRISM,XY}$ was constant between years, as it was the mean precipitation difference between sites.

Each year's precipitation (P_{ann}) at Site X was estimated with the PRISM multiplier as:

$$(P_{X,ann})_{forward} = P_{Y,ann} \times S_{PRISM,XY} \quad (2.5)$$

Combining equations (2.1), (2.3), and (2.4), the forward model estimated SWE as:

$$\overrightarrow{SWE_n} = \sum_{t=1}^{n} (P_{Y,t} \times f_{X,t}) \times S_{PRISM,XY} - \sum_{t=1}^{n} MF_t \times (T_{X,t} - MBASE) \quad (2.6)$$

2.5.2 Reconstruction model

We modeled snowfall accumulation (A) with the reconstruction model based on the mass balance across the snow season. Every year at Site X, total snowfall must equal total snowmelt (equation (2.3)) over the course of the snow season (neglecting other mass transfers):

$$S_{recon,XY} \times \sum_{t=c}^{d} f_{X,t} P_{Y,t} = \sum_{t=c}^{d} MF_t \times (T_{X,t} - MBASE) \quad (2.7)$$

where, $S_{recon,XY}$ is an annually and spatially varying multiplier that relates mass outflow (i.e., snowmelt) to mass inflow (i.e., unadjusted snow accumulation) at Site X, $t=c$ denotes the first day of continuous snow cover modeled at Site X, and d is the observed snow disappearance date.

During each year, $S_{recon,XY}$ was solved in equation (2.7) and then used to model snowfall accumulation at Site X on day t with the reconstruction model:

$$(A_{X,t})_{recon} = (P_{Y,t} \times f_{X,t}) \times S_{recon,XY} \quad (2.8)$$

This addition to the reconstruction model permitted SWE modeling across the entire snow season and allowed estimation of annual precipitation with SWE reconstruction:

$$(P_{X,ann})_{recon} = P_{Y,ann} \times S_{recon,XY} \quad (2.9)$$

Combining equations (2.2), (2.3), and (2.8), the SWE reconstruction method was finalized as:

$$\overleftarrow{SWE_n} = \sum_{t=n}^{d} MF_t \times (T_{X,t} - MBASE) - \sum_{t=n}^{d} (P_{Y,t} \times f_{X,t}) \times S_{recon,XY} \quad (2.10)$$

While S_{PRISM} could have been used in the SWE reconstruction model (Equation 2.10) in place of S_{recon}, we found that S_{PRISM} did not improve SWE reconstruction (no results shown). We include S_{recon} here to assess the accuracy of a basic approach of backing out precipitation from the snowpack mass balance (Equations 2.7 and 2.9).

2.5.3 Combined model

Annual precipitation and SWE were estimated with the combined model as:

$$\left(P_{X,ann}\right)_{combined} = \frac{\left(P_{X,ann}\right)_{forward} + \left(P_{X,ann}\right)_{recon}}{2} \tag{2.11}$$

$$\overleftrightarrow{SWE_n} = \frac{\overrightarrow{SWE_n} + \overleftarrow{SWE_n}}{2} \tag{2.12}$$

2.5.4 Sensitivity analysis

The study was designed to represent a best case scenario for the model configurations. In practice, it is common that air temperature at Site X (Figure 2.3) is not observed and must be estimated from data at other stations. Snow disappearance timing is most readily observed with remote sensing, which is also subject to various errors (see Section 2.2). The study was further idealized because the PRISM map was trained by past data at 114 of the 136 study sites, so PRISM accuracy was likely maximized. Because biases in model inputs and parameters should impact the models differently (Table 2.1), a sensitivity analysis (section 2.6.4) was conducted to quantify the impact of biases in model input data (air temperature, precipitation, and snow disappearance timing) and model parameters (rain/snow delineation, melt threshold temperature, melt factors) on peak SWE. This was accomplished by introducing independent, artificial biases in each data input and model parameter, and observing the changes in peak SWE accuracy.

2.6 Results

2.6.1 Annual precipitation

Estimates of annual precipitation were compared to the uncorrected, on-site precipitation observations. The most accurate estimates of annual precipitation were associated with the forward model (i.e., PRISM). While median errors (Table 2.4) in annual precipitation from the three models were similar, the reconstruction and combined models had a higher frequency of larger errors as seen in Figure 2.4a. The reconstruction and combined models had larger errors in precipitation because of errors associated with transferring the five model parameters (see Section 2.6.5). Figure 2.4b presents the results as cumulative probability distributions and demonstrates that the forward model generally yielded smaller errors than the other two models. The reconstruction approach of estimating annual precipitation (equation (2.9)) was twice as likely as the forward model to produce an annual precipitation error exceeding 10%, and nearly 12 times as likely to produce an error exceeding 50% (Figure 2.4b). Not surprisingly, the combined model results fell between the extremes of the forward and reconstruction models.

2.6.2 Peak SWE

Median peak SWE errors (Table 2.4) from the three models were not significantly different and had negative biases (Figure 2.4c). The negative bias for the forward model may be indicative of the median measurement error due to undercatch. The combined model was more likely to produce smaller peak SWE errors than either the forward or reconstruction models (Figures 4c and 4d). The forward and reconstruction models consistently demonstrated similar probabilities of absolute errors (Figure 2.4d); this indicated that neither approach was statistically preferable for modeling SWE at the study sites. The forward and reconstruction models were

each 1.2 times as likely as the combined model to produce a peak SWE error exceeding 10%, and 3.2 times as likely to produce an error exceeding 50% (Figure 2.4d).

2.6.3 SWE during the accumulation and ablations seasons

To understand the accuracy of SWE estimation during specific seasons, mean absolute errors (MAE) in modeled SWE were recorded during each accumulation and ablation season. Seasonal errors were assessed because peak SWE errors (Figure 2.4d) may not be useful for seasonal-specific applications, such as development of snow depletion curves for the ablation season. During the accumulation season, the forward and reconstruction models produced similar MAE (Figure 2.5 and Table 2.4) while the combined model produced lower MAE. During the ablation season, the forward model had significantly higher MAE, while the reconstruction and combined model had similar MAE (Figure 2.5). Forward model errors were greater during the ablation season because errors from the accumulation season were carried over to the ablation season. This caused major errors in estimated snow disappearance timing with the forward model; 65% of the forward simulations had at least a 7 day error in snow disappearance.

2.6.4 Sensitivity of results to model inputs and parameters

The sensitivity analysis (Figure 2.6) confirmed the expectations of Table 2.1. As seen in Figure 2.6a, the forward (reconstruction) model SWE error was negatively (positively) correlated with air temperature bias. The combined model was significantly less sensitive to air temperature errors because averaging overestimation and underestimation errors from the forward and reconstruction models resulted in median SWE errors closer to zero (Figure 2.6a). This implied that some SWE errors in the original analysis (no artificial bias) may have been the

result of errors in the observational air temperature data and/or the model calibration. As hypothesized, these errors tended to cancel out in the combined model.

The reconstruction model was insensitive to precipitation bias (Figure 2.6b) while forward model SWE errors were directly proportional to the introduced precipitation bias. Although our formulation of the SWE reconstruction method (equation (2.10)) included precipitation to allow modeling across the entire snow season, it was insensitive to precipitation bias. This was explained by the multiplier development. For example, a precipitation bias of -50% (i.e., $0.5P_Y$) would increase S_{recon} by a factor of 2 in equation (2.7), which would cancel the same -50% P_Y bias in equation (2.10). This also explained why errors in reconstructed SWE during the accumulation season were not large (Figure 2.5), despite having large annual precipitation errors (Figure 2.4a).

The forward model was independent of biases in snow disappearance timing by definition. When reconstructing peak SWE, there was an average additional error of 4.3% for every 1 day of snow disappearance date bias (Figure 2.6c). This error is similar in magnitude to the results of *Slater et al.* [2013], who reconstructed SWE with an idealized snow model at SNOTEL stations across the conterminous United States.

When estimating peak SWE, the forward and reconstruction models had opposite sensitivities to the model parameters (Figures 6d, 6e, and 6f), but the magnitudes of their sensitivities varied. The forward model was more sensitive to the rain-snow threshold temperatures (T_{snow} and T_{rain}) than the reconstruction model (Figure 2.6d) because the forward model peak SWE is dependent on snowfall accumulation (Equation 2.1). Likewise, the reconstruction model had greater sensitivity to bias in the snowmelt threshold temperature (*MBASE*) and the snowmelt factors (*MFMIN* and *MFMAX*), because reconstructed peak SWE

was primarily a function of calculated snowmelt (Equation 2.2). The reconstruction model's sensitivity to *MFMAX* (Figure 2.6f) was comparable to the forward model's sensitivity to precipitation bias (Figure 2.6b).

2.6.5 Parameter transfer accuracy

The assumption that model parameters could be transferred (Section 2.5) was checked by running the calibration optimization routine (Section 2.4) at each study snow pillow during all available water years, such that a unique set of five parameters was developed on-site for each station-year. These on-site parameters were considered as the best-case calibration for each station-year and eliminated any errors associated with transferring parameters from the regional calibration stations. Figure 2.7 displays the results of this analysis, and generally shows that reconstruction benefited the most from improved model parameters. For annual precipitation (Figures 7a and 7c), the reconstruction and combined models demonstrated increased accuracy with on-site parameters, while the forward model accuracy was unchanged because the forward model's estimates of annual precipitation were independent of the snow model parameters. Peak SWE accuracy dramatically increased for reconstruction when on-site parameters were used (Figures 7b and 7d). With the forward and combined models, SWE accuracy increased only slightly with on-site calibration (Figures 7b and 7d). Consequently, reconstruction became the most accurate peak SWE estimator when improved model parameters were available (Figure 2.7b).

2.7 Discussion

2.7.1 Summary of key findings

This study answered the two study questions: (1) When transferring model parameters (Figure 2.3), the PRISM-forced forward model generally estimated annual precipitation with the greatest accuracy, while the combined model typically estimated SWE with the greatest accuracy (Figures 4 and 5). (2) As expected for peak SWE estimation, the combined model was least sensitive to air temperature errors (Figure 2.6a) and reconstruction was least sensitive to precipitation errors (Figure 2.6b). The forward model was most sensitive to the rain-snow parameters (Figure 2.6d), while the reconstruction model was most sensitive to the melt threshold temperature (Figure 2.6e) and the melt factors (Figure 2.6f); the combined model was less sensitive in those cases. In practice, uncertainty will exist in all model parameters and data, and the results here suggest that the combined model may yield the lowest overall sensitivity to uncertainty.

When transferring model parameters, the likelihood of errors in peak SWE was nearly identical for the forward and reconstruction models (Figure 2.4d), and thus the accuracy of estimating snowfall during the accumulation season was comparable to the accuracy of estimating snowmelt during the ablation season. While this result did not support the implicit premise of reconstruction, that snowmelt can be estimated more accurately than snowfall, it implies that both approaches may be equally viable, given reasonable input data (Figure 2.3). With similar distributions of peak SWE errors, the combined model improves accuracy at locations where one model (forward or reconstruction) overestimates peak SWE and the other underestimates peak SWE. This overestimation-underestimation situation, which is characteristic of errors in air temperature and model parameters (Figure 2.6a, 6d-f), occurred in 58% of the station-years (n=334). The combined model improved SWE estimation (Figure 2.4c) in 62% of those cases. By construction, the combined model could never produce the largest

error at any given station-year because it was the average of the other two model estimates. This guarantees that the combined model will always estimate peak SWE more accurately than at least one of the two models, a useful feature when uncertainty exists in the forcing data and parameters of both forward and reconstruction models. Thus, the combined model will estimate SWE distributions more reliably than both the forward and reconstruction models only if those models have similar overall accuracy, which is the case here (Figures 4c and 4d).

Because the forward and reconstruction models produced overestimation and underestimation errors at various station-years, it was evident that errors in observed air temperature and/or transferred model parameters must have corrupted the models' estimates of peak SWE. This was demonstrated with the forward model, which had high accuracy in annual precipitation in the original analysis (Figure 2.4a), and lower accuracy when estimating peak SWE (Figure 2.4c). The forward model underestimation of peak SWE may be partially attributed to gauge undercatch, as the -8.9% median error was in the range of the -4.8% to -9.5% mean undercatch errors reported in the study area [*Serreze et al.*, 1999]. Errors in air temperature and model parameters limited the accuracy of reconstruction as well, as the accuracy of the reconstruction model dramatically increased with on-site calibration (Figure 2.7d), which corrected errors in model parameters and compensated for bias (if any) in observed temperature.

Given improved (on-site) calibration parameters, the initial premise of reconstruction was supported, as the ablation season was simulated more accurately than the accumulation season (Figure 2.7b). Because the forward model gained little improvement in peak SWE accuracy with on-site calibration, we found that the rain-snow partitioning parameters (T_{rain} and T_{snow}) were difficult to calibrate with accuracy, but placed a major control on model performance. Optimal parameters for snowmelt (*MBASE*, *MFMIN*, *MFMAX*) improved reconstructions of peak SWE

(Figure 2.7d), but in practice, the problem of deriving these optimal parameters remains [*He et al.*, 2011a]. Until research demonstrates improved parameter transferability and estimation, errors in the model parameters are likely to be comparable to those in the original analysis, where the combined model produced more accurate peak SWE estimates (Figures 4b and 4d).

Parameter transferability presents a significant obstacle to temperature-index models, and energy balance models are often advocated as an alternative. However, energy balance models may have large uncertainties in data inputs and also have multiple parameters which must be estimated or transferred. A complete energy balance model might have 25 or more terms of potential uncertainty [see Table 7 of *Marks and Dozier*, 1992], whereas the simplified SNOW-17 reconstruction model had a total of seven terms of uncertainty (air temperature, snow disappearance timing, and the five model parameters). Even if a simple energy balance model is employed, uncertainty in the radiative terms alone (on the order of 10-40 W m^{-2}) may exceed the data uncertainty for a model like SNOW-17 [*Slater et al.*, 2013]. Most energy balance models simulate the required inputs (e.g., radiation, humidity) through empirical relationships [e.g., *Waichler and Wigmosta*, 2003] that are defined by parameters. These parameterized empirical equations must be transferred as well, and therefore parameter transfer is an inescapable issue for all types of snow models.

2.7.2 Guidelines for model selection

The purpose of this study was to evaluate the models' accuracy and sensitivity, in order to understand which should be employed in practice. The guidelines derived from the study's results are summarized below and depend on data availability and the user's objectives.

1. If a reliable precipitation gauge network exists near the study basin, then PRISM (or a comparable data interpolation method) should be used to estimate precipitation (Figures 4a and

4b). However, the combined model should be employed to estimate peak SWE (Figures 4c and 4d), especially in the context of data and parameter uncertainty (Figure 2.6). Both the combined and reconstruction models are preferred over the forward model when developing snow depletion curves that relate SCA and SWE in the basin during the snowmelt season (Figure 2.5).

2. If a basin with a precipitation gauge network has high uncertainty in remotely sensed snow disappearance timing (e.g., cloudy conditions during the snowmelt season or infrequent sampling), then the forward model should be considered to estimate peak SWE (Figure 2.6c).

3. If the precipitation gauge network is sparse or non-existent, then SWE reconstruction should be employed because it is likely to produce similar errors as a forward model driven by PRISM (Figures 4c and 4d). A crude estimate of spatially-distributed precipitation (Figures 4a and 4b) may be backed out using the reconstruction method with equations (2.7) and (2.9) if at least one precipitation gauge exists in the area.

4. In all cases, the accuracy of SWE estimation will vary with the spatial scale of interest. For example, when estimating peak SWE across all station-years (n=334), the median bias was relatively small for all three models (Table 2.4). This implies that when aggregating SWE estimates over a large spatial scale (e.g., a basin), any of the three models might have skill in estimating mean areal SWE, but at any one specific point location (e.g., an ecological study site), there is a high probability of producing an error that exceeds the median bias. The median bias was less than 10% for all three models, but 65% of the combined model simulations and 77% of the forward and reconstruction simulations exceeded 10% error (Figure 2.4d).

2.7.3 Representativeness of the results

The accuracy and sensitivity of model configurations were evaluated here at flat clearings located at mid-elevations in maritime mountain ranges, so the results are most applicable to sites

with similar characteristics and climate. These sites can be found worldwide (e.g., New Zealand, Japan, eastern Russia, northwest Europe, Chile, and southwestern British Columbia). Modeling SWE in these other areas may inevitably demand inclusion of other (scale-dependent) processes, discussed further below, which did not require representation at the study sites.

Forested and sloped regions present additional uncertainties in all three models. The forward model may need to represent canopy interception and sublimation dynamics, and this could introduce additional uncertainty in SWE estimation. Likewise, the forest canopy may introduce uncertainty in the date of snow disappearance, which increases reconstruction uncertainty. The forest canopy also places a major control on snowmelt dynamics by reducing incident shortwave radiation and reducing turbulent energy transfer, and the representation of these processes in a model requires the estimation of many additional model parameters [e.g., *Storck*, 2000]. Terrain aspect also controls snowmelt energy, although to a lesser extent than forest canopy [*Coughlan and Running*, 1997]. Additional work is needed to develop techniques for estimating model parameters at sloped and forested sites [*Rutter et al.*, 2009].

Due to the study's location (maritime mountains of the western U.S.), 10% of the SWE simulations were at sites located below 1000m, and nearly 50% were at sites below 1600m. Thus, the results presented here are most representative of sites in or just above the snow transition zone, where winter air temperatures are mild, and mixed rain/snow storms and mid-winter snowmelt events are common [*Marks et al.*, 1998]. In this zone, the forward model's sensitivities to air temperature and rain-snow threshold parameters are high, which provides more incentive to use the combined model. At higher elevations, alpine regions, and colder continental climates, all winter precipitation falls as snow. In these locations the forward model may be less sensitive to air temperature, and estimates of snowfall will be impacted more by

precipitation extrapolation errors. However, wind transfer and sublimation have larger impacts on the mass balance in these regions [*Marks and Dozier*, 1992] and may require representation.

Because the analysis focused on snow pillow sites, this study showed results under an ideal scenario for data inputs (Figure 2.4) and model parameters (Figure 2.7). In reality, the data inputs, precipitation multipliers, and model parameters will have increased uncertainty when estimating SWE at ungauged locations. With heightened uncertainty in all model forcings (Figure 2.6), the results show that the combined model will reduce the magnitude of bias in peak SWE because of compensating errors (Table 2.1).

2.8 Conclusions

When estimating precipitation and SWE at study locations (assuming only air temperature and snow disappearance timing are known), the selection of a model configuration (i.e., forward, reconstruction, or combined) depends on the density and quality of the precipitation gauge network, the uncertainty of the model inputs and parameters, and the user's objectives. Precipitation-based studies should be guided by PRISM or similar off-line precipitation modeling (no snow model necessary). Ablation-specific studies [snowmelt depletion curves, e.g., *Homan et al.*, 2010; *Lee et al.*, 2005] should utilize either reconstruction or a combined model (Figure 2.5).

A snow model yields different estimates of peak SWE depending on the model configuration (forwards or reconstruction), partly because each configuration is uniquely sensitive to errors in model data and parameters (Figure 2.6). The quality of the forward and reconstruction estimates cannot be known when estimating peak SWE at an ungauged point location (e.g., an ecological plot) because the magnitude and sign of the errors in the data and parameters are unknowable. However, the simple averaging of the forward and reconstruction

estimates of peak SWE in the combined model may yield the most reliable estimate at point locations because the contrasting sensitivities of the models tend to minimize overall sensitivity to errors in data and parameters. Based on our evaluation at snow pillow locations in the maritime region of the mountainous western U.S., forward and reconstruction configurations have comparable accuracy, and so we recommend the combined configuration in this region.

When estimating SWE over larger domains (e.g., zonal areas of a basin) for subsequent streamflow analysis, the three configurations may yield similar SWE estimates when averaged across the basin, but may produce contrasting SWE estimates in each zone. Similar mean SWE estimates will not change the seasonal flow volume, but the distribution of SWE across the zones will impact streamflow timing. The combined model acts to eliminate large SWE errors (Figure 2.4c) in these zones and therefore may improve estimation of streamflow timing.

Improving snow model parameterization and transfer remains a challenging research endeavor, but is nevertheless important because model parameterization places a fundamental control on model accuracy, as demonstrated in this study. Accurate snow model parameters may dramatically improve estimates of annual precipitation and SWE with reconstruction (Figure 2.7).

Because the study was restricted to flat clearings in the maritime zone, additional investigation is needed to compare the model configurations in areas with varying slope, aspect, and forest cover and in different snow regimes and climates. If snow models are to represent fine-scale spatial variations in SWE in these areas, they must accurately model the associated accumulation and ablation processes. Snow and meteorological data are not routinely collected along slopes and under forest canopies, but are nevertheless required for testing model performance in these environments.

2.9 Tables

Table 2.1 Theoretical sensitivity of forward and reconstruction models to biases in model inputs and parameters when estimating peak SWE. We assume that mid-winter melt may occur before peak SWE (impacting the forward model) while some snow accumulation may occur after peak SWE (impacting the reconstruction model).

Parameter	Forward	Reconstruction
Model Inputs		
Air Temperature	Strongly Negative	Strongly Positive
Precipitation	Strongly Positive	Negative
Snow Disappearance Date	Independent	Strongly Positive
Model Calibration Parameters		
Rain-snow transition temperature(s) [1]	Strongly Positive	Negative
Snowmelt threshold temperature [2]	Positive	Strongly Negative
Snowmelt factors (degree day factors) [3]	Negative	Strongly Positive

[1] The rain-snow threshold temperatures in this study are Train and Tsnow
[2] The snowmelt threshold temperature in this study is MBASE
[3] The minimum and maximum snowmelt factors in this study are MFMIN and MFMAX, respectively

Table 2.2 Data thresholds used to flag potentially erroneous data in the quality control process. All flagged values were visually inspected to make quality control decision. The rate-of-change (ROC) limit was used to detect jumps in the data series while the no-observed-change (NOC) limit was used to detect constant data.

Parameter	Min. Limit	Max. Limit	ROC Limit	NOC Limit
Air Temperature	-25°C	40°C	10°C day-1	3 days
Precipitation	0mm	150mm day-1	150mm day-1	3 days
SWE	0mm	3000mm	300mm day-1	10 days

Table 2.3 Descriptions and ranges of the calibrated snow model parameters. Parameters were optimized at the 18 snow pillow sites (Figure 2.2a) that were isolated for regional calibration.

Parameter	Description	Range
Tsnow	All precipitation is snow when Tair ≤ Tsnow	-5 to 0°C
Train	All precipitation is rain when Tair ≥ Train	1.5 to 5°C
MBASE	Minimum temperature required for snowmelt	0 to 0.6°C
MFMIN	Minimum temperature-index snowmelt factor	0.36 to 3.6 mm °C-1 day-1
MFMAX	Maximum temperature-index snowmelt factor	2.8 to 6.8 mm °C-1 day-1

Table 2.4 Summary statistics of errors (%) in annual precipitation (n=294) and peak SWE (n=334), and MAE (%) in seasonal SWE with the forward, reconstruction and combined models. SWE errors are relative to observed peak SWE. All error distributions were non-normal, and thus the non-parametric statistics are of prime interest. Parametric statistics are shown for reference only. Seasonal MAE was taken during the observed accumulation and ablation seasons. In the non-parametric statistics, the Interquartile Range (IQR) is the difference between the 75th and 25th percentiles.

	Non-Parametric				Parametric		
	Median	25th	75th	IQR	Mean	Std. Dev. (σ)	Variance (σ2)
Annual Precipitation							
Forward	-0.4	-8.8	8.9	17.1	1.4	17.1	290
Reconstruction	1.4	-23.0	46.5	69.5	22.4	74.5	5550
Combined	2.4	-11.3	22.2	33.5	11.9	38.7	1500
Peak SWE							
Forward	-8.9	-31.6	12.6	44.2	-5.1	45.0	2030
Reconstruction	-6.0	-25.8	17.5	43.3	-0.3	40.5	1640
Combined	-7.6	-18.6	6.5	25.1	-3.3	27.2	740
MAE Accumulation SWE							
Forward	13.8	7.1	23.5	16.4	19.4	19.8	390
Reconstruction	13.9	7.5	23.4	15.9	19.1	16.7	280
Combined	9.6	6.4	16.0	9.6	13.1	12.0	140
MAE Ablation SWE							
Forward	31.1	15.1	45.6	30.5	35.7	31.9	1020
Reconstruction	16.0	9.8	26.1	16.3	20.3	17.1	290
Combined	14.5	9.9	22.4	12.5	18.6	17.4	300

2.10 Figures

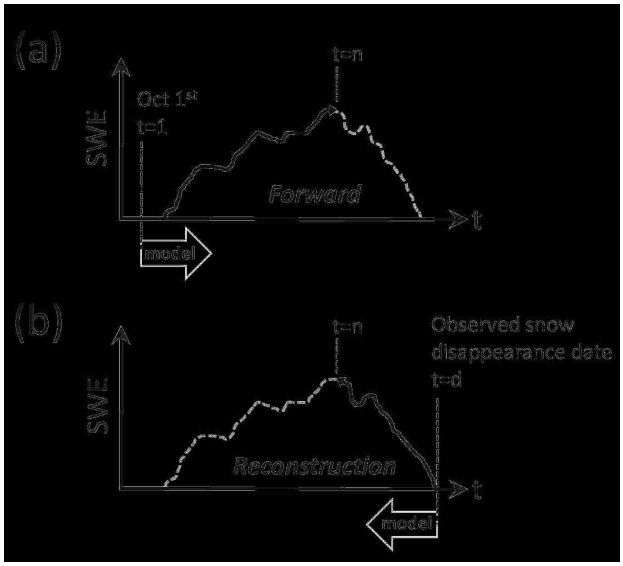

Figure 2.1 Conceptual schematic of SWE simulations from (a) forward and (b) reconstruction configurations of the same snow accumulation and melt model. SWE in (a) is simulated forwards in time from October 1st, while SWE in (b) is simulated backwards in time from the observed snow disappearance date. In this example, peak SWE occurs at t=n.

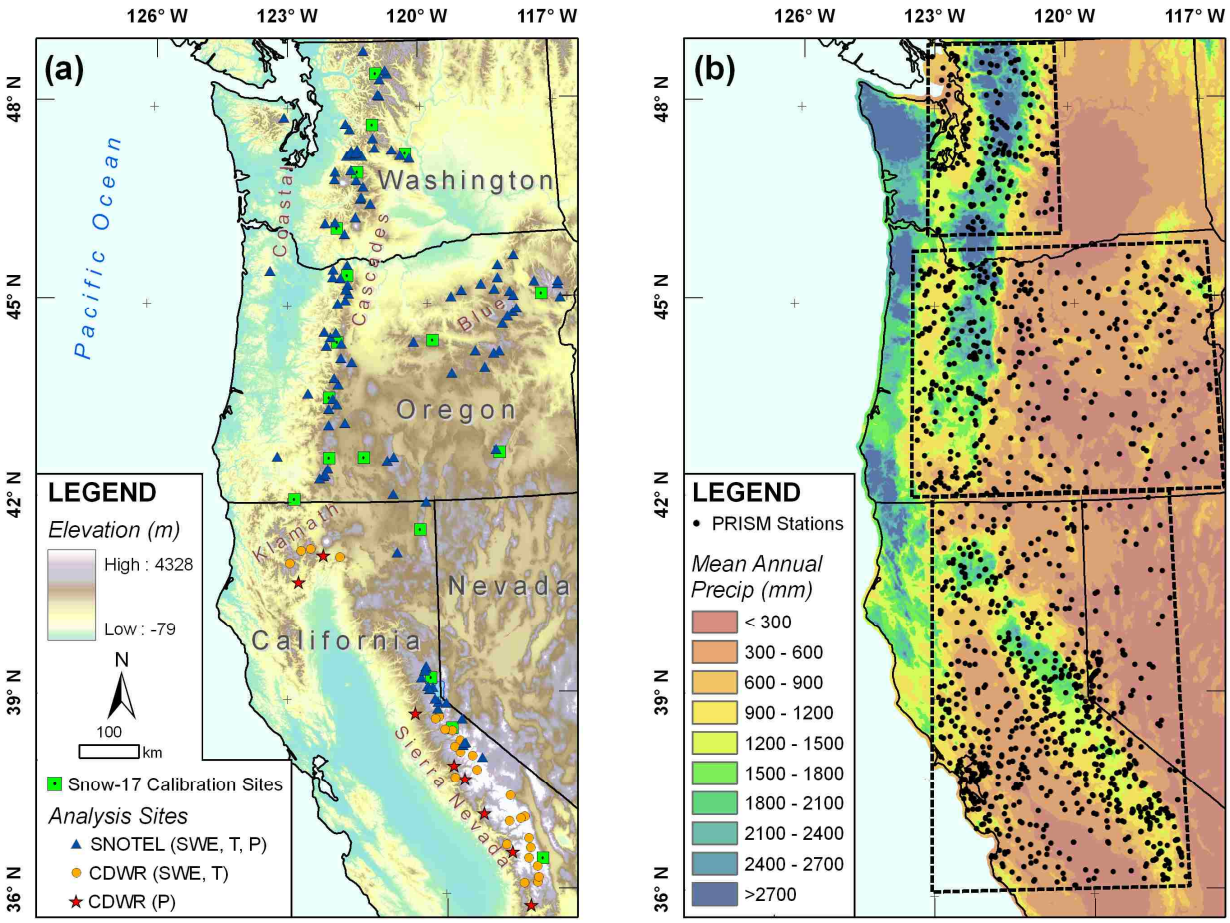

Figure 2.2 (a) Analysis sites and model calibration sites in Washington, Oregon, and California shown on top of a 1km digital elevation model. Analysis sites are from the SNOTEL and CDWR networks in California. (b) The PRISM 1971-2000 mean annual precipitation map with PRISM network stations revealed in three zones (dashed boxes). The PRISM map was used to derive the precipitation multiplier to estimate precipitation in the forward model.

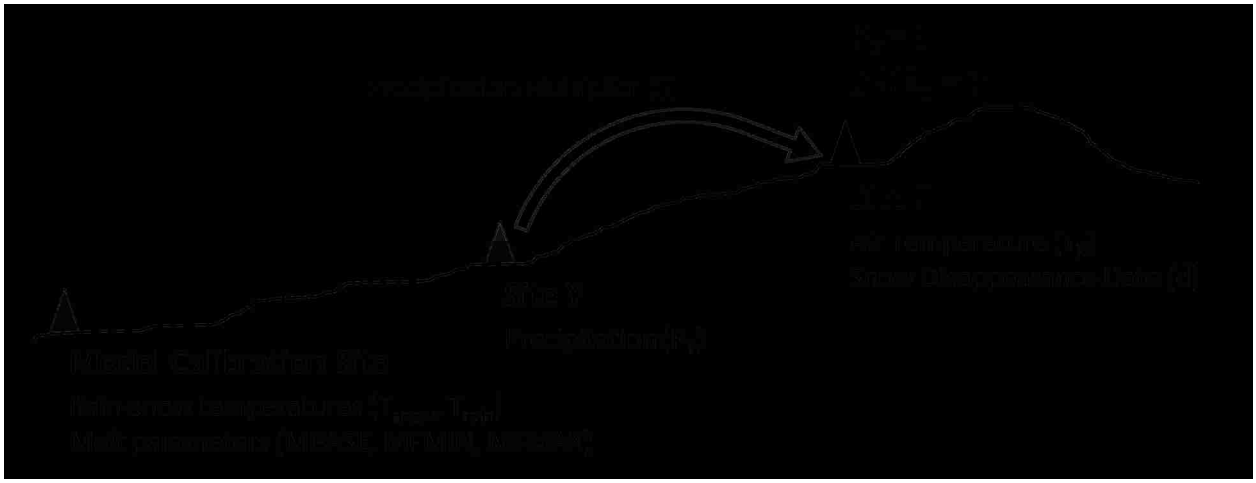

Figure 2.3 Conceptual schematic of available observations assumed when estimating annual precipitation and SWE at a study site (Site X). The snow accumulation and melt model was calibrated at the nearest regional calibration site (e.g., a snow pillow site) and parameters were transferred to Site X. Precipitation observations were taken from the nearest gauge (Site Y), which was not always located at the Calibration Site. To estimate precipitation at Site X, observations from Site Y were uniformly increased or decreased by a precipitation multiplier, S. Local observations of air temperature and snow disappearance at Site X were assumed available. Note: the relative elevation between sites was not always consistent (e.g., Site X was sometimes lower in elevation than Site Y and/or the Calibration Site).

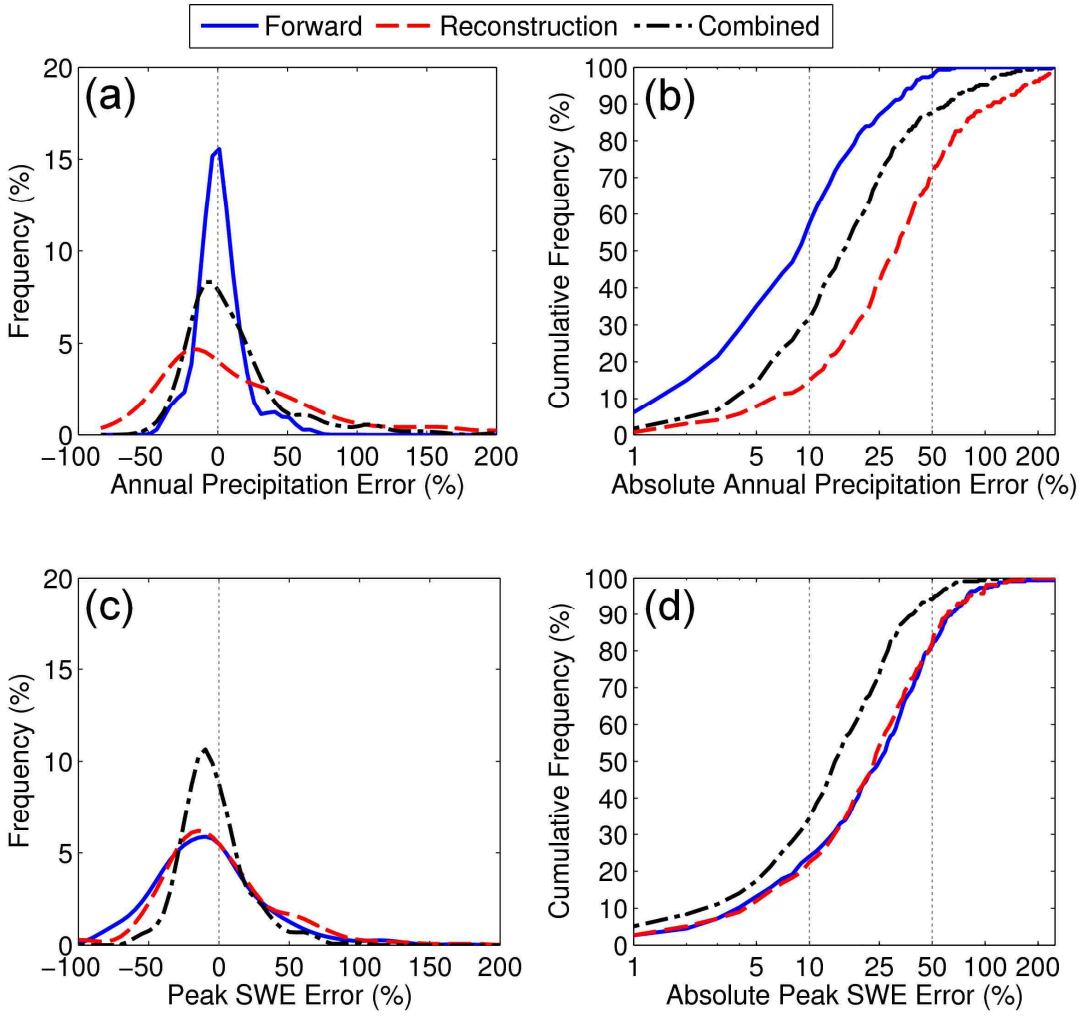

Figure 2.4 Smoothed error histograms for (a) annual precipitation (n=294) and (c) peak SWE (n=334) from the three model configurations. Histograms in (a) and (c) were smoothed across bins of 5% error in precipitation or SWE using kernel density estimation. Absolute errors from (a) and (c) are shown in terms of cumulative frequency in (b) and (d), respectively, and are plotted against a logarithmic scale. The snow model used regional calibration parameters transferred from the 18 calibration sites.

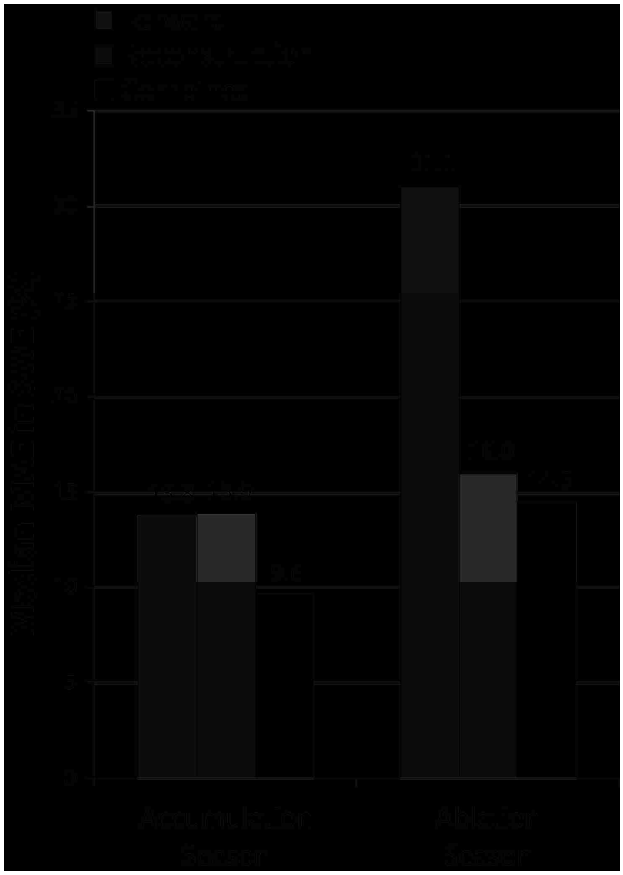

Figure 2.5 Median values of the mean absolute error (MAE) in modeled SWE (with regional calibration) during the accumulation and ablation seasons. For each station-year (n=334), the observed snow season was taken as the longest continuous period of observed SWE greater than 0, and the timing of observed peak SWE divided the snow season into the accumulation and ablation seasons. Seasonal MAE was normalized to observed peak SWE for each station-year.

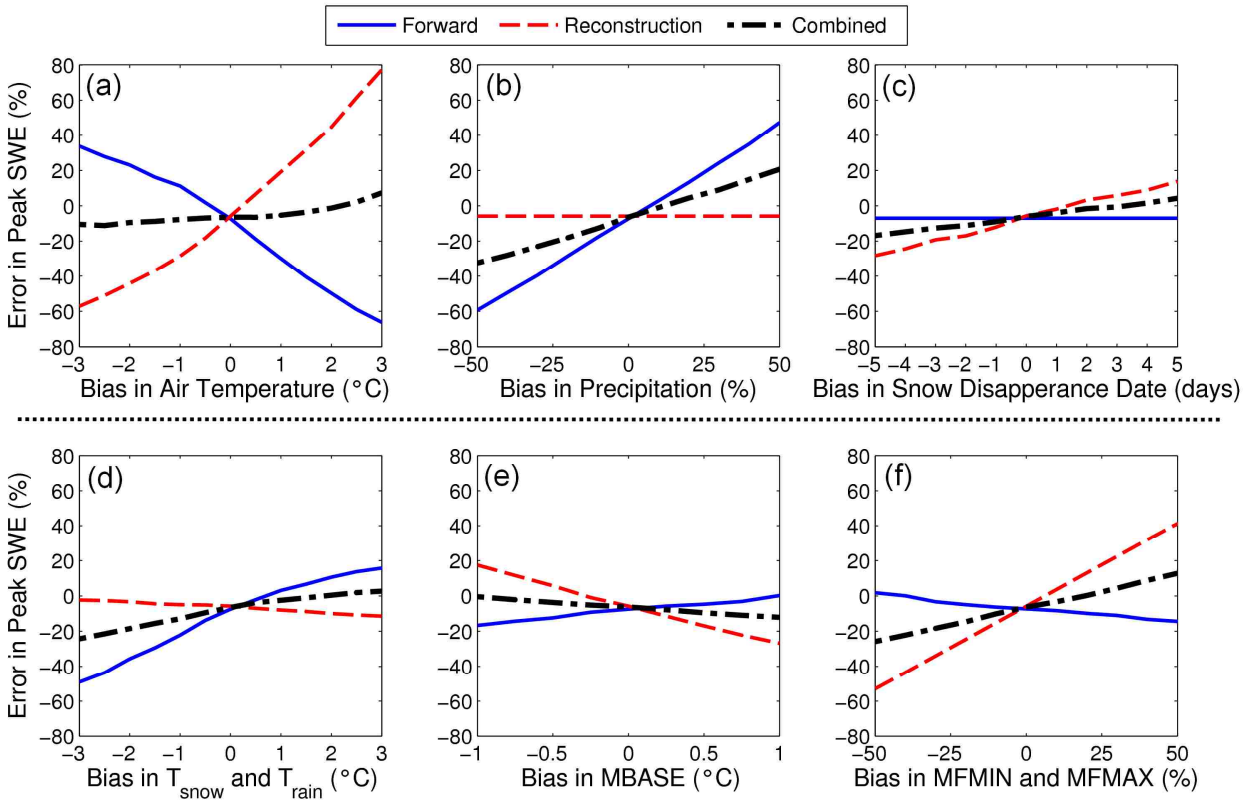

Figure 2.6 Sensitivity of median peak SWE errors (n=334) to artificial biases in model inputs (a-c) and calibrated model parameters (d-f) with the three models. Distributions about the medians were not presented for clarity; the original distribution (Figure 2.4c) was qualitatively preserved about the median. Cases with two parameters (i.e., d and f) were shifted uniformly with each bias. Initial calibration parameters were based on regional calibration.

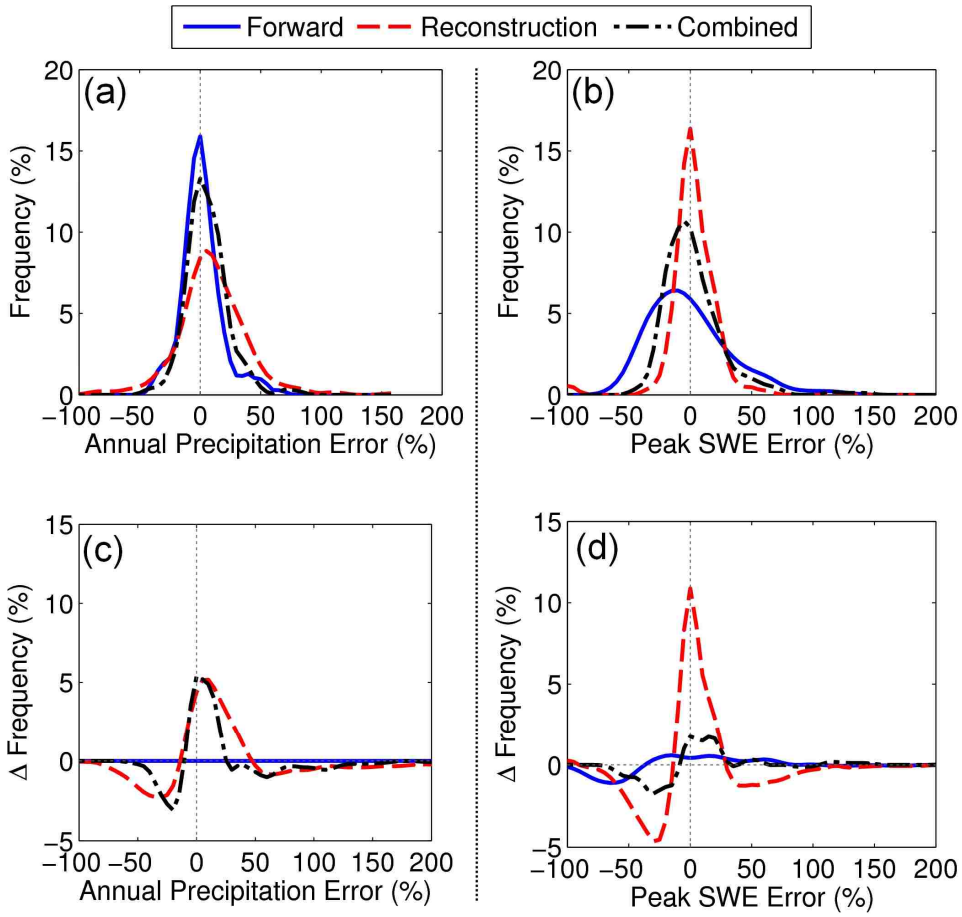

Figure 2.7 Smoothed error histograms for (a) annual precipitation (n=294) and (b) peak SWE (n=334) from the three models when using on-site (instead of regional) model calibration during each station-year. Changes in error frequency with calibrated on-site parameters (relative to regional parameters) are shown for (c) annual precipitation and (d) peak SWE.

2.11 Appendix A

Table 2.5 SWE and snow depth reconstruction studies to date.

Reference	Application [Location][1]	Melt Estimation Method	SCA/Snow Disappearance Instrument	Initial SWE	Included snowfall accum.?	Direct SWE Validation	Regression Tree Comparison
Bagchi [1979]	Snow depth [H]	Degree-Day approach	Landsat	1 Apr	No	--	--
Martinec and Rango [1981]	SWE distribution [RM], [SA]	Degree-Day approach	Landsat	1 Apr or 1 May	Yes	+4.8% error (n=1)	--
Cline et al. [1998]	SWE distribution [SN]	Energy Balance Approach	Landsat TM	21 Apr	No	+6% error (n=4)	69-78% within 1σ
Molotch et al. [2004]	Albedo comparison and snowmelt [SN]	Degree-day with net radiation	AVIRIS	6 Apr	No	--	64% of grids within 1σ
Molotch and Bales [2005]	Point/grid SWE scales [RM]	Same as *Molotch et al.* [2004]	Landsat ETM	1 Mar	No	--	8% mean error
Molotch and Bales [2006]	Albedo comparison and snowmelt [SN]	Same as *Molotch et al.* [2004]	AVIRIS	6 Apr	No	--	1% mean error
Molotch and Margulis [2008]	Reconstruction with different instruments [RM]	Same as *Molotch et al.* [2004]	Landsat ETM, MODIS, AVHRR	1 Mar	No	--	23% MAE (Landsat ETM)
Durand et al. [2008]	Reconstruction with multiple instruments [RM]	Same as *Molotch et al.* [2004]	Landsat ETM, MODIS	1 Mar	No	--	55-60% max. error
Molotch [2009]	Large scale SWE reconstruction [RM]	Same as *Molotch et al.* [2004]	Landsat ETM	1 Mar	No	--	23% MAE, & 37-55% within 1 σ
Homan et al. [2010]	Snow depletion curves [A], [I]	Energy Balance Approach	MODIS	2 Mar, 12 May, 15 May	No	< 50% max. error	--
Rice et al. [2011]	SWE distributions and melt runoff [SN]	Degree-Day approach	MODIS	Unspecified	No	15-18% MAE	--
Slater et al. [2011]	Uncertainty analysis [WU]	Degree-Day approach	Snow Pillows	Entire Year	Yes	90% within 10% error	--

[1] Study Locations: [A] = Alaska; [H] = Himalayas; [I] = Idaho; [RM] = Rocky Mountains; [SA] = Swiss Alps; [SN] = Sierra Nevada; [WU] = Western USA

Chapter 3 Ground-based testing of MODIS fractional snow cover in subalpine meadows and forests of the Sierra Nevada

This chapter has been published in its current form in *Remote Sensing of Environment* [*Raleigh et al.*, 2013c]. Permission has been granted for reproduction in this dissertation. Sections, figures, tables, and equations have been renamed here, and some citations (e.g. papers in press) may have been updated.

Abstract

The Moderate Resolution Imaging Spectroradiometer (MODIS) is used widely for mapping snow cover in climate and hydrologic systems, but its accuracy is reduced in forests due to canopy obstruction. Prior validation datasets cannot quantify MODIS errors in forests, because finer-resolution passive sensors (e.g., Landsat) encounter the same canopy errors, and operational ground-based networks sample snow in clearings where snow dynamics differ from those in the forest. To assess MODIS accuracy relative to forest cover, we applied a common canopy adjustment to daily 500 m fractional snow-covered area (f_{SCA}) from the physically-based MODIS Snow-Covered Area and Grain size (MODSCAG) algorithm, and tested it at subalpine meadow and forest sites (0.25 km^2 – 1 km^2) in the Sierra Nevada, California during two snow seasons. 37 to 89 sensors monitored hourly ground temperature at these sites. Damped diurnal variations provided a signal for snow presence due to the insulating properties of snow, yielding daily ground-based f_{SCA} at each site. Ground-based f_{SCA} values were validated in a canopy-free area of a meadow site using time-lapse imagery and 15 m snow maps from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). Ground-based f_{SCA} had high correlation (R^2=0.98) with time-lapse data and was within 0.05 of ASTER f_{SCA}. Comparisons between MODSCAG and ground-based f_{SCA} revealed that an underestimation bias remained in the canopy-adjusted MODSCAG f_{SCA}, ranging from -0.09 to -0.22 at the meadow sites and from -0.09 to -0.37 at the forest sites. Improved canopy adjustment methods are needed for MODIS f_{SCA}.

3.1 Introduction

Seasonal snow cover is a critical component of the energy and water budgets of mountainous watersheds. The high albedo and low thermal conductivity of snow reduce energy absorbed by the land surface, while snowpack stores water during the winter and releases it in the spring as snowmelt. Spatial mapping of snow cover with the NASA Moderate Resolution Imaging Spectroradiometer (MODIS) [*Hall et al.*, 2002] is convenient because of its frequent (daily) observations of snow at moderate spatial resolution (500 m). MODIS has been used to evaluate the spatial distribution of snow cover in models [*Shamir and Georgakakos*, 2006], to

improve streamflow forecasting for reservoir operations [*McGuire et al.*, 2006], to monitor climate change in areas with few snow observations [*Bormann et al.*, 2012], and to reconstruct spatial distributions of snow water equivalent (SWE) [*Homan et al.*, 2010; *Rice et al.*, 2011; *Rittger et al.*, 2011].

However, numerous factors limit the availability and accuracy of MODIS imagery, including cloud cover, large sensor view zenith angles (>30°), and the proportion and density of forest cover [*Dozier et al.*, 2008; *Hall et al.*, 1998; *Nolin*, 2010]. Forest cover is defined here as the fraction of the land surface obscured by tree canopy when viewed at nadir. Optical remote sensing is less accurate in forests as trees cast shadows [*Kane et al.*, 2008; *Vikhamar and Solberg*, 2003], and the canopy conceals the surface where snow may exist [*Liu et al.*, 2004, 2008]. Except in cases of abrupt disturbances (e.g., fire, wind storms, beetle outbreak, or timber harvest), coniferous forests change over annual to decadal time scales, and thus the canopy is the most persistent obstacle to remote sensing of snow in forested, temperate areas. Forests are extensive, covering 40% of the North American snow zone (Klein et al., 1998), and as much as 50% of the Sierra Nevada snow zone [*Richards*, 1959]. Quantifying MODIS errors in forests is critical for applications that rely on remotely sensed snow cover and snow disappearance timing [*Raleigh and Lundquist*, 2012].

Canopy or vegetation adjustments are typically made to binary snow presence and fractional snow-covered area (f_{SCA}) MODIS products in forested areas. Klein et al. (1998) found that the normalized difference vegetation index (NDVI) improved mapping of MODIS binary snow cover with the normalized difference snow index (NDSI) in forested areas. For f_{SCA} retrievals, sub-canopy snow cover has been assumed equivalent to snow cover in the viewable gap [*Liu et al.*, 2004], which adjusts (i.e., increases) the pixel f_{SCA} [*Durand et al.*, 2008; *Molotch*

and Margulis, 2008; *Rittger et al.*, 2011]. The accuracy of canopy adjustment methods for MODIS f_{SCA} is of prime interest, as viewable (i.e., canopy-free) f_{SCA} mapping with the physically-based MODIS Snow-Covered Area and Grain size algorithm (MODSCAG, Painter et al., 2009) has higher accuracy than NDSI-based empirical methods when using Landsat Enhanced Thematic Mapper Plus (ETM+) as validation [*Rittger et al.*, 2012].

While prior studies have acknowledged the limitations of remote sensing of snow in forested regions [*Hall et al.*, 1998, 2000, 2001; *Klein et al.*, 1998; *Liang et al.*, 2008; *Liu et al.*, 2004; *Nolin*, 2010; *Simic et al.*, 2004; *Vikhamar and Solberg*, 2003], few studies have quantified the impact of increasing forest cover on MODIS snow mapping accuracy and the effectiveness of f_{SCA} canopy adjustments. Comparisons with Landsat TM in Alaska have indicated that the original MODIS binary snow mapping algorithm [*Hall et al.*, 1995] has 96% accuracy in areas with < 50% forest cover and 71% accuracy in areas with > 50% forest cover [*Hall et al.*, 1998]. However, MODIS errors in forests cannot be reliably assessed with higher resolution sensors such as Landsat, because the Landsat sensor's line-of-sight is also obstructed by forest canopy and is susceptible to snow mapping errors from forest self-shadowing effects [*Kane et al.*, 2008]. Satellite intercomparison studies can yield uncertain conclusions because the highest resolution sensor is assumed the most accurate [*Hall et al.*, 2000], even though all passive instruments have common limitations (e.g., forest canopy). Therefore, ground-based observations provide an independent and robust approach to validate MODIS and quantify snow cover mapping errors in forested areas.

Most MODIS ground-based validation studies have used daily snow observation networks (e.g., SNOTEL stations in the United States) spread across regional and continental scales [*Brubaker et al.*, 2005; *Dong and Peters-Lidard*, 2010; *Klein and Barnett*, 2003; *Maurer*

et al., 2003; *Parajka and Blöschl*, 2008; *Pu et al.*, 2007; *Simic et al.*, 2004; *Tekeli et al.*, 2005; *Zhou et al.*, 2005]. These networks provide insufficient information about the accuracy of MODIS or canopy adjustment methods in forested areas because measurement stations are typically located in clearings, which exhibit different snow accumulation and melt dynamics relative to forests [*Varhola et al.*, 2010a]. Additionally, the spatial density of observational networks is relatively sparse, with an average of 1 station per 100 km^2 in the most dense network [*Parajka and Blöschl*, 2008]. MODIS pixels rarely encompass more than one ground-based snow sensor.

Intensive field surveys yield more dense spatial information to evaluate snow mapping in forests and the effectiveness of canopy correction methods. For example, the NASA Cold Land Processes Field Experiment (CLPX) [*Cline et al.*, 2003] measured snow properties in subalpine forests and alpine areas in Colorado that have allowed evaluations of microwave remote sensing and examination of MODIS view angle effects in forests [*Liu et al.*, 2008; *Xin et al.*, 2012]. CLPX teams sampled snow depth at 500 locations in nine 1 km^2 intensive study areas during 7 to 9 day periods near peak snow accumulation (late February and late March) in 2002 and 2003 [*Elder et al.*, 2009]. While the CLPX density (500 measurements per 1 km^2) is ideal, the timing of the observation periods (near peak accumulation) limits our understanding of potential snow cover mapping errors in the forest. The largest errors in remotely sensed snow cover are likely to occur late in the melt season when snow may persist longer in clearings (e.g., Storck et al., 2002) or in forests (e.g., Pomeroy & Granger, 1997). Field crews are rarely able to measure snow disappearance timing [*Jost et al.*, 2007] because it requires frequent (e.g., daily) surveys through the melt season.

Shallowly-buried temperature sensors provide a robust, ground-based approach to monitor snow presence under the forest canopy and in clearings through an entire snow season [*Lundquist and Lott*, 2008]. In temperate regions, diurnal fluctuations in near-surface soil temperatures are significantly reduced or absent when snow is present [*Lundquist and Lott*, 2008; *Tyler et al.*, 2008], allowing inference of snow presence during damped temperature cycles at each sensor. By using a network of such sensors, f_{SCA} can be inferred and used to test MODIS-derived f_{SCA} and canopy adjustment methods in ways that higher spatial resolution imagery (e.g., Landsat) or typical ground-based monitoring stations (e.g., SNOTEL) cannot.

The purpose of this study is (1) to test the seasonal accuracy of f_{SCA} from canopy-adjusted MODSCAG through comparison with dense networks of daily ground-based observations at four sites in the California Sierra Nevada with varying forest cover, and (2) to validate the ground-based methodology. We validated our ground-based f_{SCA} during one snowmelt season at a meadow site using observations of daily snow depletion from time-lapse photography and high resolution (15 m) snow maps from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) [*Yamaguchi et al.*, 1998].

3.2 Study sites and years

We selected four study sites (0.25 km^2 to 1.0 km^2) in the California Sierra Nevada (Figure 3.1a, Table 3.1) across a range of forest cover. These included subalpine meadow sites at Tuolumne Meadows (TUM) and Dana Meadows (DAN), and forest sites at the Onion Creek Experimental Forest (ONN) and the Yosemite Forest Dynamics Plot (YFDP). All four sites have a Mediterranean climate, with the majority of annual precipitation falling between October and May, typically as snowfall [*Baker*, 1944; *Serreze et al.*, 1999].

TUM and DAN are located in the headwaters of the Tuolumne River Basin in Yosemite National Park. TUM is located in a west-draining valley. A forested north-facing slope is found to the south, while a mixture of forest and rock outcroppings along a south-facing slope is found to the north. DAN is located 3 km west of Mt. Dana (3981 m) in a south-draining valley. Both meadow sites are flat, with mean slopes of 1° in TUM and 5° in DAN. Lodgepole pine (*Pinus contorta*) is the dominant tree species in these areas. Mean forest cover in the area surrounding each site is 23% at TUM and 32% at DAN (Table 3.1). The California Department of Water Resources (CDWR) measures SWE with snow pillows near each of these sites. SWE data were available at both sites through most of the study period, except after mid-February 2011 at TUM.

ONN is situated southwest of Donner Summit in the headwaters of the North Fork of the American River Basin. The study area spans 125 m of relief. Slopes are primarily southwest facing and average 15°. The forest cover at ONN averages 65% and is primarily mixed-conifer forest, with communities of red fir (*Abies magnifica*), white fir (*Abies concolor*), Jeffrey pine (*Pinus jeffreyi*), and incense-cedar (*Calocedrus decurrens*), with discontinuous forest cover in montane chaparral and meadow. A prominent thicket of chaparral is found along the northwestern edge of the ONN study site, while a 0.2 km^2 dry meadow is located to the south. In this meadow, the National Oceanic and Atmospheric Administration (NOAA), through their Hydrometeorological Testbed (HMT, Ralph et al., 2005), maintains a meteorological station that monitors snow depth. This station provided reliable data until March 2011, when heavy snow accumulations buried and damaged the mast arms of the tower. After March 2011, the snow depth data were suspect.

YFDP (http://www.yfdp.org) is located in the Tuolumne River Basin near Crane Flat in Yosemite National Park [*Lutz et al.*, 2012]. The site is predominantly north facing, with 18°

mean slopes, and a 115 m elevation span. An old-growth forest comprised primarily of sugar pine (*Pinus lambertiana*) and white fir (*Abies concolor*) characterizes the site. Of the four study sites, YFDP has the highest mean forest cover (79%), but also the lowest mean elevation (1860 m).

The study spanned water years (i.e., 1 Oct – 30 Sept) 2010 and 2011, which exhibited contrasting snow conditions according to monthly snow course measurements from the California Cooperative Snow Surveys (CCSS). During water year (WY) 2010, near average or above average snow conditions were found on 1 April, and above average conditions were found on 1 May due to additional snow storms and/or low melt rates throughout April (Figure 3.1b). Water year 2011 featured anomalously high snow accumulations on both 1 April and 1 May (Figure 3.1c). The CCSS surveys are made in clearings, where winter snow accumulations tend to be greater than under forest canopies. To measure the magnitude of this difference at ONN, we conducted a snow survey on 17 April 2011 at 47 points in the meadow and 51 points in the adjoining forest. Meadow SWE averaged 470 mm higher than forest SWE (Figure 3.1c). The CCSS 1 April snow course measurements in ONN matched our 17 April meadow survey.

3.3 Methods

3.3.1 Ground-based f_{SCA}

In temperate snow zones such as the subalpine regions of the Sierra Nevada, diurnal fluctuations in sub-surface ground temperature (T_g) are damped or absent when snow is present because the low thermal conductivity of snow causes it to insulate the ground [*Lundquist and Lott*, 2008; *Tyler et al.*, 2008]. Thus, by measuring hourly T_g with shallowly-buried temperature sensors, daily snow presence can be inferred during periods with a reduced diurnal cycle in T_g (Figure 3.2). Maxim (San Jose, California) Thermochron iButtons (model DS1922L) and Onset

(Cape Cod, Massachusetts) HOBO Pendant data-loggers were deployed at TUM, DAN, and ONN to measure T_g every hour from August 2009 through July 2010, and from September 2010 through August 2011. Sensors measured hourly T_g at YFDP from November 2010 to July 2011. 92% of all sensors deployed in the study were iButtons, while the remaining 8% were HOBO Pendants. Sensors were buried 2 cm to 10 cm under the surface (Figure 3.2a), following the methods of Lundquist & Lott (2008).

A network of temperature sensors sampled T_g across each study site at regular spatial intervals (Figure 3.3). During WY 2010, paired sensors were located within 5 m to 10 m of each other, and all sets of paired sensors were located 100 m apart from each other on a quasi-regular grid (Figure 3.3a, 3.3c, 3.3e). During WY 2011, the networks at TUM, DAN, and ONN were expanded to cover a larger area (Figure 3.3b, 3.3d, 3.3f). Sensors were no longer paired at each location and were spaced every 100 m at TUM and ONN and every 100 m to 200 m at DAN. Sensors at YFDP (Figure 3.3g) were only deployed in WY 2011, and were spaced at 40 m intervals along two parallel transects. Sensors at TUM, DAN, and ONN were geolocated with a handheld GPS unit while sensors at YFDP were installed at study points that were surveyed with a total station.

After retrieving the sensors, all hourly T_g time series (Figure 3.2b) were converted to daily time series of binary snow presence (i.e., 0=snow-free, 1=snow) with the following simple algorithm. To identify snow presence above a sensor, the algorithm required that the diurnal range in T_g did not exceed 1.0° C over a period of 48 hours (Figure 3.2c). The resulting daily snow presence time series (Figure 3.2d) were visually checked against the original hourly T_g. Additional quality control was conducted by comparing the timing of snowfall events (as observed at nearby snow pillows or as calculated based on air temperature and precipitation) to

the snow time series at the study sites. This eliminated spurious snow detection during cold, snow-free periods in the autumn with low diurnal temperature variations. Averaged across the sites, spurious increases in f_{SCA} occurred on 10 days during WY 2010 and 3 days during WY 2011. When spurious snow presence was detected, we reclassified the affected sensors as snow-free.

To derive daily f_{SCA} time series at each site (hereafter called "ground f_{SCA}"), the number of sensors reporting snow presence each day were summed and divided by the total number of sensors at that site (Table 3.1). We checked the confidence of ground f_{SCA} through a Monte-Carlo type test at each site, where 10 sensors were randomly sampled to produce a unique f_{SCA} time series. This was repeated 100 times and the standard deviation of the 100 f_{SCA} time series was computed during each day of the ablation season. Averaged across the ablation season, the standard deviation of ground f_{SCA} ranged from 0.045 to 0.092, suggesting that the ground networks adequately sampled the snow cover dynamics of each study area.

3.3.2 Validation of ground f_{SCA}

To test ground f_{SCA}, we used two independent observations of snow cover depletion at TUM during the spring and summer of 2010. These included a time-lapse camera, which was used to check the timing and rate of snow cover depletion, and three high-resolution (15 m) images from ASTER in a canopy-free area, which were used to further assess f_{SCA} accuracy.

3.3.2.1 Time-lapse analysis

An east-facing time-lapse camera took a photograph of Tuolumne Meadows every four hours in May and June 2010 and recorded the depletion of snowpack from full cover to snow-free conditions. The camera was located approximately 2 km west of the TUM ground sensors

and focused on the western area of the meadow, which we assumed had similar snow depletion timing and rates to the meadow as a whole. The camera, a Sony Cybershot (model DSC-W55) with 7.2 megapixel resolution, was placed in a protective casing and fixed to a tree for stability.

Numerous studies have employed time-lapse photography to assess patterns of snow presence, to quantify spatial and temporal components of snow cover depletion at small scales, and to measure snow depth (see Parajka et al., 2012 for a review). We employed a novel approach for detecting the temporal depletion of snow by mapping snow cover in each scene based on pixel brightness and then using singular value decomposition (SVD) (see Wall et al., 2003) to extract the temporal depletion information. SVD is the basis for Principal Component Analysis (PCA), which is essentially equivalent to Empirical Orthogonal Function (EOF) analysis. SVD reduces a complex system into its principal modes of variability; in other words, it finds the signals that explain the most variance of the data set in space and time.

Daily images (e.g., Figure 3.4a-4c) taken between 10AM and 2PM PST were retained to coincide with the overpass of Terra (10:30 AM equatorial crossing), to minimize shadows in the image, and to ensure relatively consistent lighting from day-to-day. Cloudy images were removed because they introduced noise in the analysis by reducing the lighting in each scene. We converted each image from RGB to the 0 to 255 range (0=black and 255=white) in order to map snow cover based on grayscale brightness. We confined the analysis to an area in the western extent of the meadow (i.e., closer to the camera) to reduce distortion of pixels, as pixels farther from the camera encompassed more land surface area than pixels closer to the camera.

A snow mapping algorithm was used and demonstrated to be consistent with visually identified snow cover in each original RGB image. To map snow in each grayscale image, we constructed a histogram and classified snow-covered pixels by finding pixel brightness values

greater than 155 (~60% brightness), which corresponds to the lower limit of snow albedo at visible wavelengths for shallow snow with a large grain radius [*Wiscombe and Warren*, 1980]. The algorithm translated each grayscale image into a binary snow map. The binary snow map of each image was reordered into a column vector, and a matrix (**M**) was constructed from all vectors, such that rows corresponded to spatial position and columns corresponded to time.

SVD was then used to derive the temporal component of snow depletion. **M** was input into the SVD.m routine in MATLAB (Version 7.9), which output the spatial and temporal modes, ordered by the proportion of variance explained. The first mode explained 34% of the variance in the time-lapse sequence, and the non-zero spatial weights indicated that this mode represented snow cover (Figure 3.4d). Thus, the first temporal mode (Figure 3.4e) described how snow cover depleted through time. To infer a time series of f_{SCA}, the absolute value of the first temporal mode was scaled to the 0 to 1 range, hereafter called "time-lapse f_{SCA}" (Figure 3.4e).

3.3.2.2 ASTER

Cloud-free ASTER images of TUM were obtained on 25 April, 5 June, and 30 July in 2010 to derive high resolution snow maps. ASTER visible and near infrared (VNIR) bands in the green-yellow (0.520 – 0.600 μm), red (0.630 – 0.690 μm), and near-infrared (0.780 – 0.860 μm) wavelengths were acquired. Supervised mapping of snow cover was implemented based on the methods of Vogel (2002), which were developed for the 15 m panchromatic band (0.52 – 0.92 μm) of Landsat 7 and found to have comparable performance to the NDSI approach. The raw digital numbers of each ASTER channel were first converted to radiance and reflectance based on NASA (2001). The VNIR reflectance values in each pixel were then averaged into a single panchromatic value. Pixels with panchromatic reflectance > 40% were mapped as snow,

consistent with the NDSI threshold [*Dozier*, 1984]. This threshold provided the best visual match to false-color images of snow cover (Figure 3.5). ASTER f_{SCA} was acquired in a 465 m square box coincident with the study area, and sized to exclude adjoining forested areas.

The number of snow-covered ASTER pixels divided by the total number of ASTER pixels (961) in the area (red box, Figure 3.5b) was taken as ASTER f_{SCA}. We carefully considered the accuracy of this approach, as fine resolution maps of binary snow cover may have bias when aggregated to a coarse f_{SCA} value [*Rittger et al.*, 2012]. As noted by Rittger et al. (2012), binary methods may underestimate snow cover at low fractions and overestimate snow cover at higher fractions. However, we did not find evidence that a major bias in ASTER f_{SCA} existed at TUM in 2010.

3.3.3 MODIS snow-covered area and grain size (MODSCAG)

The physically-based MODSCAG algorithm uses spectral mixture analysis on a pixel-by-pixel basis to derive gridded 500 m daily f_{SCA}. Spectral mixture analysis finds the best linear combination of land surface endmembers (e.g., snow, soil, rock, vegetation, lake ice) that matches MODIS surface reflectance from the Terra MOD09GA product [*Painter et al.*, 2009]:

$$R_{S,\lambda} = \sum_i f_i R_{\lambda,i} + \varepsilon_\lambda \tag{3.1}$$

where $R_{S,\lambda}$ is the pixel-averaged surface reflectance from MOD09GA at wavelength λ, f_i is the fraction of endmember i in the pixel, $R_{\lambda,i}$ is surface reflectance for endmember i at wavelength λ, and ε_λ is the residual error for all endmembers. Wavelength-dependent surface reflectance of non-snow endmembers (e.g., vegetation, soil) are acquired from a library of observations acquired in the field or in a laboratory. Snow reflectance is estimated using a hemispherical directional reflectance factor and a discrete ordinates radiative transfer model. Using the

approach in Equation 3.1, MODSCAG examines permutations of two or more endmembers and selects the model with the smallest error (relative to MOD09GA reflectance) and the fewest endmembers. If this combination of endmembers includes snow, then the daily f_{SCA} is computed as the fraction of the snow endmember, normalized by the fraction of photometric shade (e.g., due to terrain or vegetation shading) in the pixel. The lower detection limit is f_{SCA}=0.15 [*Painter et al.*, 2009]. For a complete description of the MODSCAG algorithm, the interested reader is directed to the model development paper of Painter et al. (2009). MODSCAG snow cover data are available through the NASA JPL Snow Data System Portal (http://snow.jpl.nasa.gov/).

After implementation of MODSCAG at all pixels and all daily scenes, cloudy and noisy pixels are filtered, producing gaps. Noisy pixels usually occur when one or more spectral bands have high frequency dropouts to zero reflectance, which cannot be used to estimate the snow cover properties of a pixel [*Dozier et al.*, 2008]. Therefore, noisy pixels require removal and estimation through interpolation. MODSCAG scenes were interpolated in time using a 16-day smoothing spline on a pixel-by-pixel basis following Dozier et al. (2008) and Dozier & Frew (2009). The implemented spline (csaps.m in MATLAB) was a weighted combination of a least-squares fit and a cubic spline. The best fit changed depending on the temporal spacing between available data, and therefore varied spatially as a result of non-uniformities in cloud cover and noisy data. Estimates from the spline were additionally weighted according to the cosine of the sensor view zenith angle and the view-angle-dependent pixel area, such that near-nadir views had the greatest weights.

Filtering and smoothing of the MODSCAG scenes produced a temporally continuous product of gridded daily f_{SCA} across the Sierra Nevada. Because the geolocation accuracy of MODIS (± 1.5 pixels) and the gridding procedure may introduce artifacts into a pixel-scale

validation [*Tan et al.*, 2006], we aggregated (i.e., averaged) the MODSCAG f_{SCA} in a 3x3 (i.e., 1500 m x 1500 m) pixel window encompassing each site to ensure collocation with the area sampled by the ground-based observations [*Xin et al.*, 2012].

3.3.4 MODSCAG canopy adjustments

MODSCAG f_{SCA} is created based on the land surface that is viewable by MODIS, which underestimates snow cover in forests due to effects such as canopy obstruction [*Rittger et al.*, 2012]. In forested areas, only the land surface in forest clearings, canopy gaps, and between canopy gaps is visible (Liu et al., 2008). These viewable areas provide the basis for estimating f_{SCA} with MODSCAG, which currently lacks a native canopy adjustment. We adjusted the MODSCAG f_{SCA} time series by the viewable gap fraction (VGF), which can be provided by a geometric optical model (Liu et al., 2004) or a satellite-derived product [*Durand et al.*, 2008; *Molotch and Margulis*, 2008; *Rittger et al.*, 2012]:

$$f_{SCA,adjusted} = \frac{f_{SCA,obs}}{1 - f_{can}} \tag{3.2}$$

where $f_{SCA,adjusted}$ is MODSCAG f_{SCA} adjusted for forest canopy, $f_{SCA,obs}$ is the f_{SCA} from the gridded daily 500 m MODSCAG product, $1\text{-}f_{can}$ is the viewable gap fraction, and f_{can} is the fractional forest cover of each grid cell. This adjustment increases f_{SCA} in areas with trees to account for the area hidden by the canopy and the greatest f_{SCA} adjustments occur in more dense forests. $f_{SCA,adjusted}$ is constrained to the [0, 1] interval. This canopy-adjusted MODSCAG f_{SCA} is hereafter referred to as "MODSCAG f_{SCA}".

For the value of f_{can} (Equation 3.2) at each site, we used the static (i.e., temporally constant) percent tree canopy from the 2001 National Land Cover Dataset (NLCD), which is derived from Landsat 5 and Landsat 7 data at 30 m resolution [*Homer et al.*, 2004; *USGS*, 2011].

This dataset is freely available and commonly used in snow research (e.g., Durand et al., 2008; Young et al., 2009). Changes in land cover at the sites were insignificant between 2001 and 2011.

There are several well-known limitations inherent in this canopy adjustment [*Rittger et al.*, 2012]. First, this correction assumes f_{SCA} under the forest canopy is equivalent to f_{SCA} in the viewable areas (e.g., meadows, clearings) [*Durand et al.*, 2008]. However, snow studies across different climate zones show accumulation and melt rates change based on forest cover and type of tree (see Varhola et al., 2010a). Viewable snow in a clearing is not likely to represent snow under the canopy, especially late in the melt season. Second, the simplifying assumption of constant f_{can} for MODIS may not be robust in areas with trees, as a greater area of each tree (e.g., leaves, branches, trunks) will be included as the view angle increases [*Hall et al.*, 1998; *Liu et al.*, 2008]. While the smoothing algorithm weights $f_{SCA,obs}$ based on view angle and pixel area (see Section 3.3.3), a weighting scheme is ineffective for static values of f_{can}, and therefore the static approach may not be fully sufficient. Third, a static adjustment is not robust when forest canopies are loaded with snow, as immediately following a precipitation event.

3.3.5 Evaluation metrics

We adopted the same binary and fractional metrics for evaluation as prior MODSCAG studies [*Painter et al.*, 2009; *Rittger et al.*, 2012] and other MODIS snow cover studies (e.g., Dong & Peters-Lidard, 2010). The binary metrics are first-order performance metrics that reveal the accuracy of MODSCAG in determining whether or not snow is present, regardless of the fractional value. Fractional metrics are used to assess actual f_{SCA} errors. MODSCAG f_{SCA} values below the MODSCAG detection limit (0.15) were set to 0 before calculating the binary

and fractional metrics. Ground f_{SCA} values below 0.15 were not changed in this manner, as this indicated the existence of patchy snow.

3.3.5.1 Binary metrics

During each day of the snow season at each study site, MODSCAG f_{SCA} was evaluated based on the agreement or disagreement of snow presence with the ground f_{SCA}. Days when snow was present (i.e., $f_{SCA} \geq 0.15$ for MODSCAG, $f_{SCA} \geq 0$ for ground) in both ground f_{SCA} and MODSCAG f_{SCA} were classified as a true positive (TP), while days when both reported snow-free conditions (i.e., $f_{SCA} < 0.15$ for MODSCAG, $f_{SCA} = 0$ for ground) were classified as a true negative (TN). A false positive (FP) indicated that MODSCAG identified snow cover not observed by the ground network (i.e., commission), and a false negative (FN) signified that MODSCAG missed snow cover that the ground network observed (i.e., omission). These daily values were input into three binary metrics to determine performance across the snow season:

$$Precision = \frac{TP}{TP + FP} \tag{3.3}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.4}$$

$$F = \frac{2TP}{2TP + FP + FN} \tag{3.5}$$

Precision tests for commission errors, *Recall* tests for omission errors, and the *F* score tests for both errors. All three binary metrics vary from 0 to 1, with 1 indicating perfect performance.

3.3.5.2 Fractional metrics

Direct comparisons of MODSCAG f_{SCA} and ground f_{SCA} were achieved through the use of mean bias (i.e., mean difference), median bias, and root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{N}(fSCA_{MODSCAG} - fSCA_{ground})^2} \quad (3.6)$$

where N is the number of snow-covered days, as observed at each ground network. Bias was taken as the difference between MODSCAG f_{SCA} and ground f_{SCA}, such that a positive (negative) bias indicated MODSCAG overestimated (underestimated) f_{SCA}.

3.4 Results

3.4.1 Validation of ground f_{SCA} at Tuolumne Meadows

Comparisons of ground f_{SCA} against time-lapse f_{SCA} and ASTER f_{SCA} in 2010 indicated that the ground f_{SCA} methodology was accurate. Time-lapse f_{SCA} and ground f_{SCA} matched each other in characterizing the timing and rate of snow cover depletion from late May to early June 2010 (Figure 3.6a). Both ground f_{SCA} and time-lapse f_{SCA} indicated that the majority of snow cover depletion occurred over a ten day period (29 May to 8 June). During this critical melt period, ground f_{SCA} had high correlation (R^2=0.98) with time-lapse f_{SCA}.

Ground f_{SCA} tracked ASTER f_{SCA} across the three available ASTER images (Figure 3.6a). Ground f_{SCA} on 25 April 2010 (full snow cover) and 30 July 2010 (no snow cover) matched ASTER f_{SCA}. A comparison on 5 June 2010 supported the accuracy of ground f_{SCA} during the critical melt out period, as ground f_{SCA} was 0.62 and ASTER f_{SCA} was 0.66 (Figure 3.5b). ASTER was valuable in that it provided additional f_{SCA} information in early June 2010, when MODSCAG snow cover disappeared abruptly (see section 3.4.2.1).

3.4.2 Time series comparisons

3.4.2.1 Tuolumne Meadows (TUM)

Across the 2010 and 2011 snow seasons, MODSCAG had high *Precision*, *Recall*, and *F* score values at TUM (Table 3.2). MODSCAG performance was better in 2011 than 2010

because omission errors (*Recall*) were present during the 2010 melt season, likely from cloud cover and view angle issues (described below). Omission errors also arose when MODSCAG either missed early season storms or the smoothing algorithm removed snow storms in October 2010 and 2011 (Figure 3.6).

The canopy adjustment reduced MODSCAG f_{SCA} bias at TUM by 13% to 15% (Figure 3.6), but canopy-adjusted MODSCAG still had a consistent negative bias at TUM during both years (Table 3.3). During the winter months, MODSCAG f_{SCA} oscillated between 0.60 and 0.95 in a pattern possibly introduced by the smoothing method (section 3.3.3); these multi-day f_{SCA} oscillations during the winter were a common feature of MODSCAG f_{SCA} at all four sites. At TUM, these oscillations did not consistently coincide with snowfall events, and therefore did not occur due to increased reflection from canopy interception.

A notable MODSCAG f_{SCA} error at TUM occurred in spring 2010 when MODSCAG f_{SCA} depleted rapidly and fell below the 0.15 threshold on 1 June, 7 days before the ground f_{SCA} and 8 days before the time-lapse f_{SCA} fell below 0.15 (Figure 3.6a). An examination of atmospheric transmissivity (calculated from insolation observations) and MODIS visible imagery indicated that cloudy conditions persisted on 1 June, 3 June, and 4 June (Figure 3.7a). Additionally, the Terra satellite was off-nadir (i.e., >30°) on 31 May (31°), 2 June (43°) and 6 June (49°). The limited availability of near-nadir view angles on clear days within the short snow cover depletion period (29 May – 8 June) likely caused MODSCAG f_{SCA} to decline rapidly on 1 June. Interestingly, SWE at the TUM snow pillow disappeared on 3 June 2010, two days after MODSCAG f_{SCA} fell below 0.15 (Figure 3.6a).

3.4.2.2 Dana Meadows (DAN)

MODSCAG had the fewest omission errors at DAN, as noted by the high *Recall* values. However, MODSCAG had a slight tendency to map snow cover that did not exist at DAN, as noted by the *Precision* (Table 3.2). These commission errors occurred in early October 2009 and 2010, and briefly in mid-October 2010 (Figure 3.8).

The canopy adjustment reduced MODSCAG f_{SCA} bias by 22% to 23% (Figure 3.8), but canopy-adjusted MODSCAG f_{SCA} still had an overall negative bias at DAN (Table 3.3). In both years, MODSCAG overestimated f_{SCA} during the early accumulation season (e.g., October) but underestimated f_{SCA} through the period of full snow cover (e.g., December through early June), with values typically fluctuating between 0.75 and 0.96. During the 2010 ablation season, MODSCAG f_{SCA} reasonably matched ground f_{SCA} (Figure 3.8a). During the 2011 ablation season, MODSCAG had a notable f_{SCA} overestimation error from 28 June to 30 June (Figure 3.7b, 8b), possibly introduced by cloud cover and off-nadir view angles. After this period, MODSCAG f_{SCA} reasonably tracked the ground f_{SCA} for the rest of the 2011 ablation season. Complete depletion of MODSCAG f_{SCA} (i.e., $f_{SCA}<0.15$) was only 1 to 2 days earlier than ground f_{SCA} depletion during the two years. During 2010, snow disappeared at the DAN snow pillow on 24 June, which was 1 day prior to MODSCAG and 2 days prior to the ground temperature network. In 2011, snow disappeared at the DAN snow pillow on 1 July, 7 days prior to MODSCAG and 9 days prior to the ground network.

3.4.2.3 Onion Creek Experimental Forest (ONN)

MODSCAG yielded more omission errors at ONN than at DAN, as suggested by lower *Recall* (Table 3.2). Omission errors occurred in November 2009, October 2010, and during the

second half of the melt season in both water years (Figure 3.9). The high *Precision* indicated commission errors were rare at ONN, which increased the *F* score.

The canopy adjustment had the greatest bias reduction at ONN, with bias reduced by 35% to 42% (Figure 3.9). However, canopy-adjusted MODSCAG f_{SCA} had a negative bias through both years at ONN (Table 3.3). The largest f_{SCA} errors occurred during the early accumulation and late ablation seasons (Figure 3.9). During the 2010 ablation season, MODSCAG f_{SCA} abruptly depleted and dropped below the 0.15 detection threshold on 1 June 2010 (Figure 3.9a). This date of complete snow depletion at ONN was 2 days prior to the date of snow disappearance at the NOAA HMT snow depth sensor, but 11 days prior to the ground temperature sensors. MODSCAG f_{SCA} depletion was more gradual during the 2011 ablation season (Figure 3.9b), but reported systematically lower f_{SCA} through this period and reached complete depletion 12 days prior to the ground sensors. During both ablation seasons, MODSCAG snow cover disappeared (i.e., $f_{SCA}<0.15$) once ground f_{SCA} approached the forest cover fraction (Figure 3.7c).

3.4.2.4 Yosemite Forest Dynamics Plot (YFDP)

MODSCAG had the lowest *Recall* and *F* score at the heavily forested YFDP (Table 3.2), indicating that snow omission errors were most common at this site. Omission errors were concentrated in May 2011 when the snow cover was melting (Figure 3.10). Sporadic periods with MODSCAG omission errors were also present during periods of partial snow cover disappearance in the middle of the snow season (e.g., early December, early February). Like ONN, commission errors were rare at the YFDP, as the *Precision* score was high.

While *Precision* was high at YFDP, errors in f_{SCA} were prevalent throughout the snow season (Figure 3.10), as MODSCAG f_{SCA} had a mean RMSE of 0.55 and a mean bias of -0.37

(Table 3.3). In other words, MODSCAG reported snow cover existed throughout most of the snow season, but MODSCAG f_{SCA} was generally too low. MODSCAG f_{SCA} was characterized by multiple cases with rapid increases to full snow cover and nearly equivalent drops in f_{SCA} within 1 to 2 weeks. These large fluctuations were not found in the ground f_{SCA} (Figure 3.10).

Comparing the MODSCAG f_{SCA} time series to snow accumulation data at the nearby Gin Flat snow pillow indicated that the MODSCAG f_{SCA} fluctuations often coincided with new snowfall events (Figure 3.10). This suggested either MODSCAG was viewing intercepted snow in the forest canopy, or storm clouds were being misclassified as snow cover, or some combination thereof due to the smoothing algorithm. However, not all f_{SCA} fluctuations coincided with new snow events at YFDP (e.g., single peak in early November, two peaks in mid-April).

During the ablation season of 2011, MODSCAG only mapped snow cover at YFDP when the ground f_{SCA} was greater than the forest cover fraction (Figure 3.7d), as at ONN. A snowfall event in mid-May 2011 brought ground f_{SCA} back to full cover and extended the snow season by 9 days. MODSCAG missed this snowfall event, as it reached complete depletion ($f_{SCA}<0.15$) prior to this storm (26 April) and remained below the 0.15 threshold through the summer. Ground f_{SCA} reached 0.15 on 26 May, 30 days after MODSCAG (Figure 3.10).

3.4.3 Seasonality of errors

We found MODSCAG errors varied relative to forest cover and time of year. We examined how errors changed across three periods: the early accumulation season (first day of $f_{SCA} \geq 0.15$ through 31 December), the winter (1 January through 31 March), and the ablation season (1 April to the final day of $f_{SCA} \geq 0.15$). At TUM, DAN, and ONN, MODSCAG errors tended to be most variable during the accumulation and ablation seasons (Figure 3.11a,c). Errors

at these three sites were generally less variable during the winter months (Figure 3.11b). At YFDP, MODSCAG tended to underestimate snow more severely and more consistently with time (Figure 3.11).

3.5 Discussion and conclusions

We demonstrated that networks of temperature sensors buried shallowly in the ground provide reliable values of daily f_{SCA}, which can be used to test canopy-adjusted MODSCAG f_{SCA} in forest locations not sampled by traditional methods (e.g. Landsat, SNOTEL sites). At our Sierra Nevada sites, we found that (1) the static canopy adjustment (Equation 3.2) reduced MODSCAG f_{SCA} bias (Figures 6, 8, 9, 10) but a consistent negative bias still remained (Figure 3.11), (2) the accuracy of canopy-adjusted MODSCAG f_{SCA} varied with forest cover, and (3) MODSCAG errors were usually most variable during the accumulation and ablation seasons. The results demonstrated the value of dense ground-based validations of remote sensing and underscored the need for improved canopy adjustments for MODIS f_{SCA}.

Canopy-adjusted MODSCAG f_{SCA} was systematically lower than ground f_{SCA} during the middle of winter (Figure 3.11b). This result was particularly surprising at the meadow sites (TUM and DAN), as we expected full snow cover at these flat, lightly forested, high-elevation locations, which had considerable snow accumulation during the two study years (Figure 3.1). We investigated several possible reasons for this. First, it was possible that the ground sensor sampling strategy (section 3.3.1) partially caused this difference, as sensors were not placed at potentially snow-free locations (e.g., rock outcroppings and open water). However, these features covered a small fraction (generally < 0.05) of the land cover at each site. Second, this difference may have resulted from a documented underestimation bias in the NLCD forest cover dataset, which has been reported at 9.7% across the USA [*Greenfield et al.*, 2009] and 23.4% in

the Sierra Nevada zone [*Nowak and Greenfield*, 2010]. However, by checking the NLCD forest cover against lidar-derived forest cover at TUM and YFDP, we found that the difference between NLCD forest cover and lidar forest cover was generally less than 5% at our sites. The mean f_{SCA} bias ranged from -0.13 to -0.25 during the winter at TUM and DAN (Figure 3.11b), and therefore neither the sampling strategy nor the possible NLCD bias provide complete explanations for the difference between MODSCAG f_{SCA} and ground f_{SCA}.

We hypothesize that the use of static forest cover data to adjust for canopy (Equation 3.2) was the primary cause of the MODSCAG f_{SCA} underestimation bias during the winter. Because canopy-adjusted MODSCAG f_{SCA} oscillated to values as high as 0.95 and 0.96 during the winter at TUM and DAN, respectively, this suggested that the canopy adjustment was effective over specific periods with favorable view angles and cloud conditions. MODIS band reflectance and viewable gap fraction (i.e., forest cover fraction) both change with view zenith angle, which changes daily for MODIS [*Liu et al.*, 2008; *Xin et al.*, 2012]. While MODSCAG accounts for the view angle-dependent changes in surface reflectance, the static canopy adjustment used here did not account for the variation of viewable gap fraction with view angle. Therefore, the static adjustment may not be appropriate for adjusting f_{SCA} from scanning sensors (e.g., MODIS). Concurrent estimates of fractional vegetation from MODSCAG that account for the changing view angle, larger pixel size, and different reflectance would likely provide superior snow maps. More work is needed to determine whether a view angle dependent forest cover fraction improves canopy adjustments.

Canopy-adjusted MODSCAG also exhibited a fundamental limitation in that no snow cover was mapped once ground f_{SCA} approached the forest cover fraction (Figures 7c, 7d). This problem cannot be corrected with the current adjustment method (Equation 3.2), as $f_{SCA,obs}$ was 0

in these cases, inevitably resulting in $f_{SCA,adjusted}$ of 0, regardless of the forest cover value (f_{can}). This remains an outstanding challenge for satellite remote sensing of snow in forested areas.

The prevalence of omission errors (Table 3.2) at the forest sites (ONN and YFDP) and the inability of MODSCAG to map snow below the forest cover fraction (Figures 7c, 7d) suggested snow persisted longer under the canopy than in the viewable gaps and clearings at these two sites. Snow lasted 12 to 30 days longer at the ground networks of the forest sites (Figures 9, 10) relative to MODSCAG snow cover. This difference in snow persistence in forests vs. clearings was consistent with other studies in the Sierra Nevada. Anderson (1956) found snow cover disappeared in a dense forest 16 days after snow in large forest openings near ONN, while Church (1914) noted snow persisted at least 7 to 10 days longer in pine and fir forests than treeless meadows near Lake Tahoe. These studies provide confidence in our observations, but we recognize that these results cannot be generalized for all forests, as forest characteristics (e.g., canopy structure, species, age) interact with snow in complex ways, resulting in variability of snow persistence [*Kittredge*, 1953]. Nevertheless, we note that errors in snow disappearance timing impact applications such as SWE reconstruction. An error of 12 days in point snow disappearance has a potential error of 50% in reconstructed peak SWE, assuming a mean SWE error of 4.3% per day of snow disappearance date bias [*Raleigh and Lundquist*, 2012]. Slater et al. (2012) show similar median SWE errors given ±10 days uncertainty in snow disappearance.

We also note that the operational snow depth and SWE sensors in clearings at our study areas did not reliably represent snow disappearance timing in the adjacent forests, and these sensors have had wide usage in prior MODIS validation studies. Taking f_{SCA}<0.15 for snow absence, the timing of snow disappearance in 2010 from MODSCAG was within 2 days of

observations at the operational snow sensors at TUM (Figure 3.6a) and ONN (Figure 3.9a). However, our ground temperature networks indicated that as much as 60% (ONN) to 84% (TUM) of the land was still snow-covered once snow disappeared at the operational sensors in 2010. In these cases, MODSCAG errors in snow disappearance were actually larger than the errors suggested by the operational data. We also observed cases when the operational snow sensors overestimated the MODSCAG errors in snow disappearance timing. For example, the DAN snow pillow data suggested MODSCAG had a 7 day error in snow disappearance date in 2011 (Figure 3.8b), but our ground-based sensors showed MODSCAG only had a 2 day error. Thus, the issue of snow sensor representativeness [*Rice and Bales*, 2008] is critically important for comparisons between a MODIS pixel and a single snow sensor in that pixel. Because most snow pillows were positioned to provide streamflow indices [*Farnes*, 1967] and not to represent the timing and duration of snow presence of an area, we conclude that a single snow depth or SWE sensor cannot validate a MODIS pixel with confidence, especially in forested pixels. This has direct implications for prior studies that have used single snow sensors to validate MODIS.

While MODSCAG is more accurate than empirical MODIS snow cover retrievals early in the accumulation season and late in the ablation season [*Rittger et al.*, 2012], our results show that MODSCAG errors are often most variable during these periods (Figure 3.11 a,c). This implies that intensive field surveys conducted near peak accumulation (e.g., CLPX) do not sample the largest MODIS snow cover errors in the seasonal snow zone. As suggested by the results at YFDP, large MODIS errors are expected to occur frequently in the transient snow zone where snow may accumulate and disappear multiple times in a single snow season. These difficulties support the use of distributed ground-based sensors to test MODIS snow cover, such as networks of ground temperature sensors (as in this study) or snow depth sensors [*Musselman*

et al., 2012; *Varhola et al.*, 2010b]. Lidar observations of snow cover (e.g., Deems et al., 2006) may also test MODIS errors in forests, but the tradeoffs between timing, frequency, and costs of lidar flights must be considered carefully.

Although our sample size (n=7 site years) is relatively small considering the large variability in forest cover across the globe, we have (1) demonstrated a new methodology for ground validation of MODIS and (2) identified errors in forests that cannot be detected with previously used validation techniques (e.g., comparisons with Landsat and SNOTEL). Thus our results provide a first quantification of forest effects on f_{SCA} errors and highlight the need for continued testing of MODSCAG and canopy adjustment methods over a more complete range of forest cover and environmental conditions. This in turn will benefit users of MODSCAG for distributed applications, such as SWE reconstruction and snow model testing.

3.6 Tables

Table 3.1 Characteristics of the four Sierra Nevada study sites, ordered by increasing forest cover.

	Tuolumne Meadows (TUM)	**Dana Meadows (DAN)**	**Onion Creek (ONN)**	**Yosemite Forest Dynamics Plot (YFDP)**
Mean Forest Cover [A], f_{can}	0.23	0.32	0.65	0.79
Latitude (N)	37° 52' 30"	37° 53' 58"	39° 16' 40"	37° 45' 59"
Longitude (W)	119° 21' 49"	119° 15' 20"	120° 21' 18"	119° 49' 9"
River Basin	Tuolumne	Tuolumne	American	Tuolumne
Mean Elevation (m)	2615	2985	1950	1860
Mean DJF Air Temp.[B] (°C)	-3.6	-4.1	-0.9	2.2
Mean Ann. Precip.[B] (mm)	830	970	1700	1060
WY 2010 Ground Sensors	45	45	89	--
WY 2011 Ground Sensors	47	52	75	37

[A] Average fractional forest cover of the 1500 m x 1500 m area encompassing each study site, based on 30 m 2001 National Land Cover Dataset forest canopy product [*Homer et al.*, 2004]
[B] Based on PRISM 1971-2000 monthly climate normals product at 800 m resolution [*Daly et al.*, 1994, 2008]

Table 3.2 Summary of binary metrics across each snow season at the four study sites.

Metric	Water Year	TUM	DAN	ONN	YFDP
Precision	2010	1.00	0.96	1.00	--
	2011	0.99	0.97	0.99	0.96
	Mean	1.00	0.96	1.00	0.96
Recall	2010	0.78	0.94	0.80	--
	2011	0.91	0.98	0.88	0.61
	Mean	0.84	0.96	0.84	0.61
F score	2010	0.87	0.95	0.89	--
	2011	0.95	0.97	0.93	0.75
	Mean	0.91	0.96	0.91	0.75

Table 3.3 Summary of fractional metrics across each snow season at the four study sites.

Metric	Water Year	TUM	DAN	ONN	YFDP
RMSE	2010	0.28	0.19	0.24	--
	2011	0.23	0.19	0.17	0.55
	Mean	0.25	0.19	0.21	0.55
Mean Bias	2010	-0.24	-0.11	-0.11	--
	2011	-0.19	-0.08	-0.07	-0.37
	Mean	-0.22	-0.09	-0.09	-0.37
Median Bias	2010	-0.25	-0.15	-0.03	--
	2011	-0.20	-0.12	-0.03	-0.41

3.7 Figures

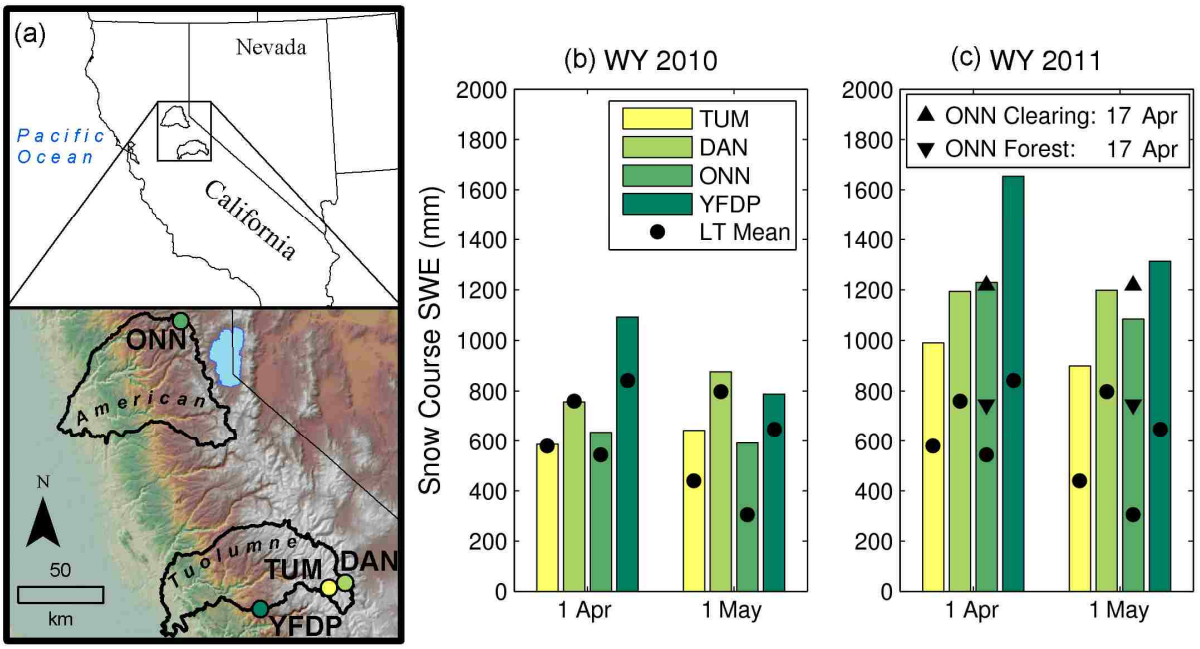

Figure 3.1 (a) Locations of Tuolumne Meadows (TUM), Dana Meadows (DAN), Onion Creek Experimental Forest (ONN), and Yosemite Forest Dynamics Plot (YFDP) in the Sierra Nevada, and snow course snow water equivalent (SWE) at or near the four study sites during water years (b) 2010 and (c) 2011. Snow courses are from the California Cooperative Snow Survey network, taken routinely every year on or near 1 April and 1 May. YFDP is represented by the Gin Flat snow course, 4 km east of YFDP and 300 m higher in elevation. Also shown are the long-term (LT) means at each snow course on 1 April (n=65 years) and 1 May SWE (n=30 years); only years with snow at all four snow courses were used to calculate the LT mean. An additional snow survey on 17 April 2011 was conducted at ONN to document the difference in SWE accumulation between a clearing and the adjoining forest.

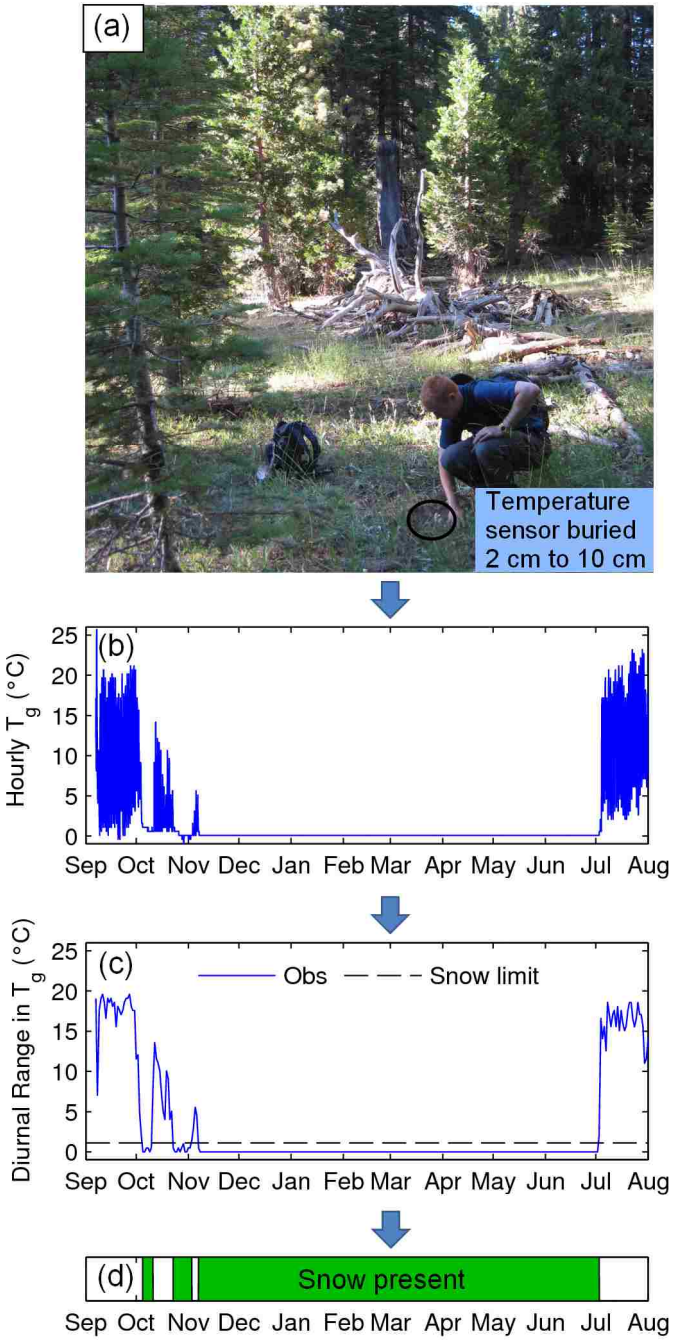

Figure 3.2 Ground-based method of determining snow presence at a point with a self-logging temperature sensor. Temperature sensors were (a) buried at a depth of 2 cm to 10 cm and recorded (b) hourly ground temperature (T_g) through the study year. (c) Diurnal temperature ranges were inspected and periods with diurnal temperatures below 1.0°C were (d) classified as snow-covered periods at that sensor.

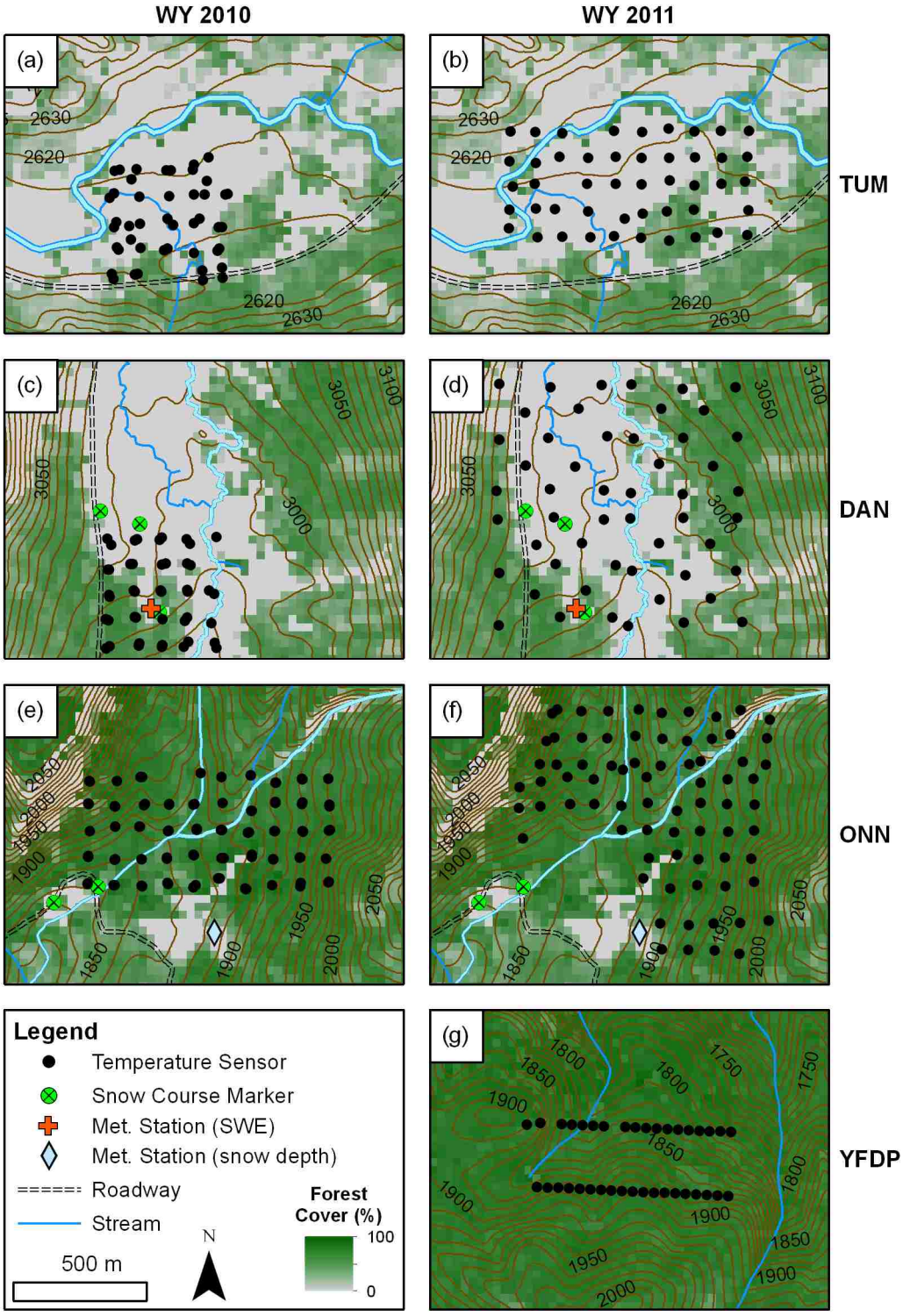

Figure 3.3 Locations of ground temperature sensors at the study sites during water years 2010 and 2011. Shown in order of increasing forest cover are (a-b) Tuolumne Meadows, (c-d) Dana Meadows, (e-f) Onion Creek, and (g) the Yosemite Forest Dynamics Plot. Most locations during WY 2010 (a, c, e) have two temperature sensors in close proximity (<10 m) but appear as a single dot. Contour interval is 5 m in (a) and (b), and 10 m in all other subplots.

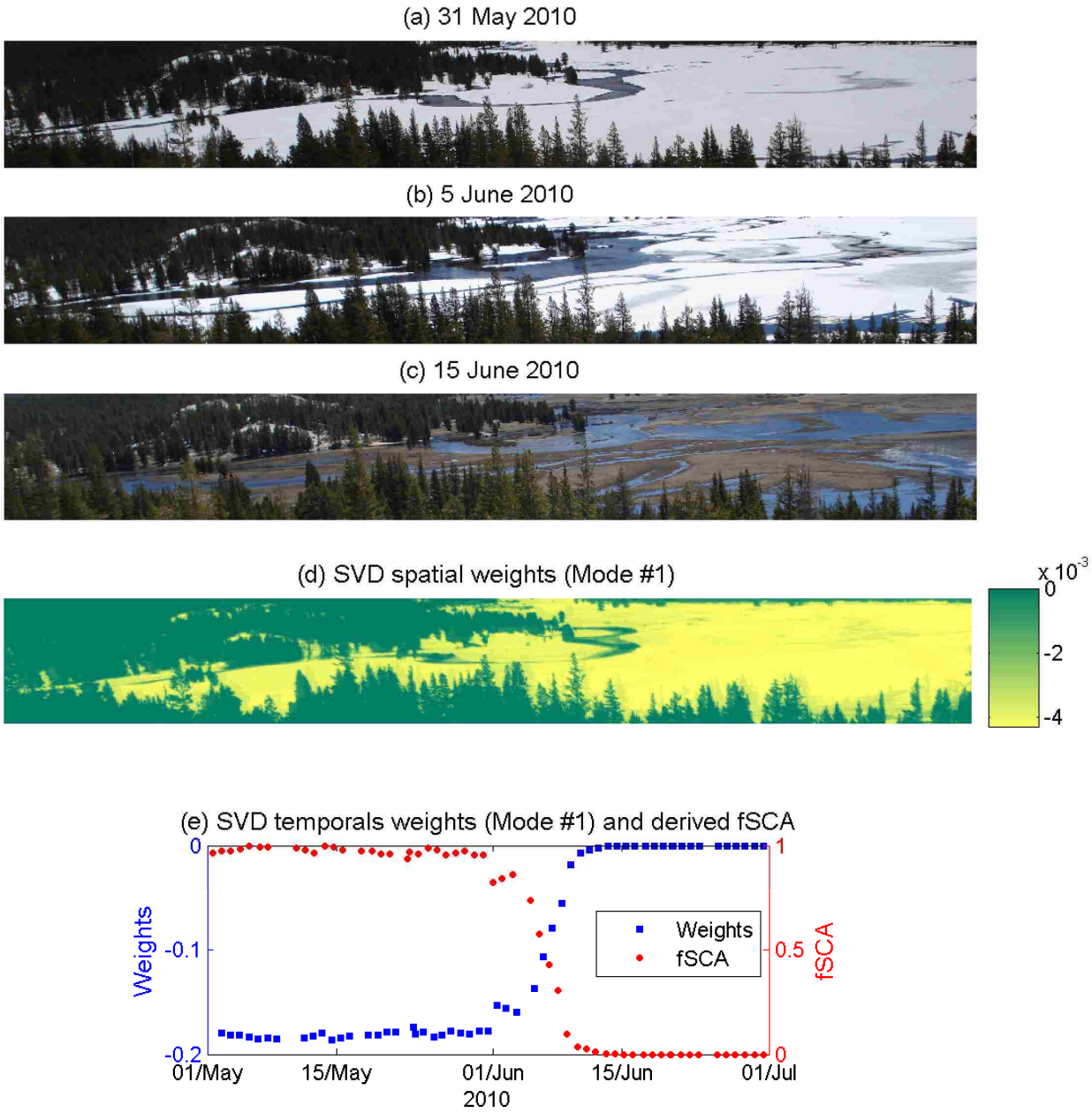

Figure 3.4 Tuolumne Meadows time-lapse photo analysis with singular value decomposition (SVD) during May-June 2010. Shown are sample RGB photographs taken to show progression of snow cover from (a) 31 May to (b) 5 June to (c) 15 June. Also shown are the (d) spatial and (e) temporal weights of the 1st SVD mode, which is interpreted as snow cover depletion. A fractional snow-covered area (f_{SCA}) time series was inferred from the temporal weights of the 1st SVD mode. The location of the time-lapse camera relative to the ground temperature sensors is indicated in Figure 3.5a.

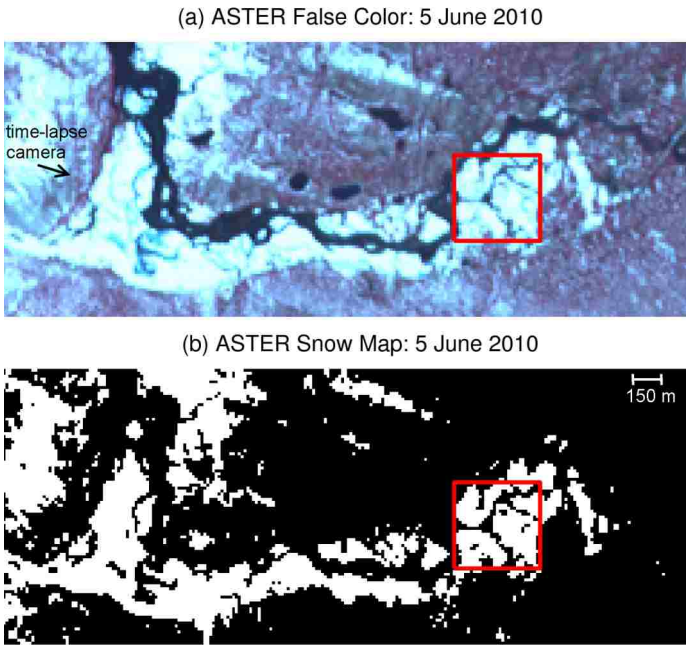

Figure 3.5 (a) ASTER nadir false color image (RGB 321) of Tuolumne Meadows on 5 June 2010, showing snow cover and melt water channels draining to the Tuolumne River. The approximate location and view direction of the time-lapse camera (Figure 3.4) are indicated. ASTER VNIR resolution is 15 m. (b) Mapped snow cover from the ASTER image on 5 June 2010 (white = snow cover, black = snow-free or unknown). The red box corresponds to the approximate location of the ground temperature sensors from Figure 3.3a. Fractional snow cover in the red box was 0.66 on this date.

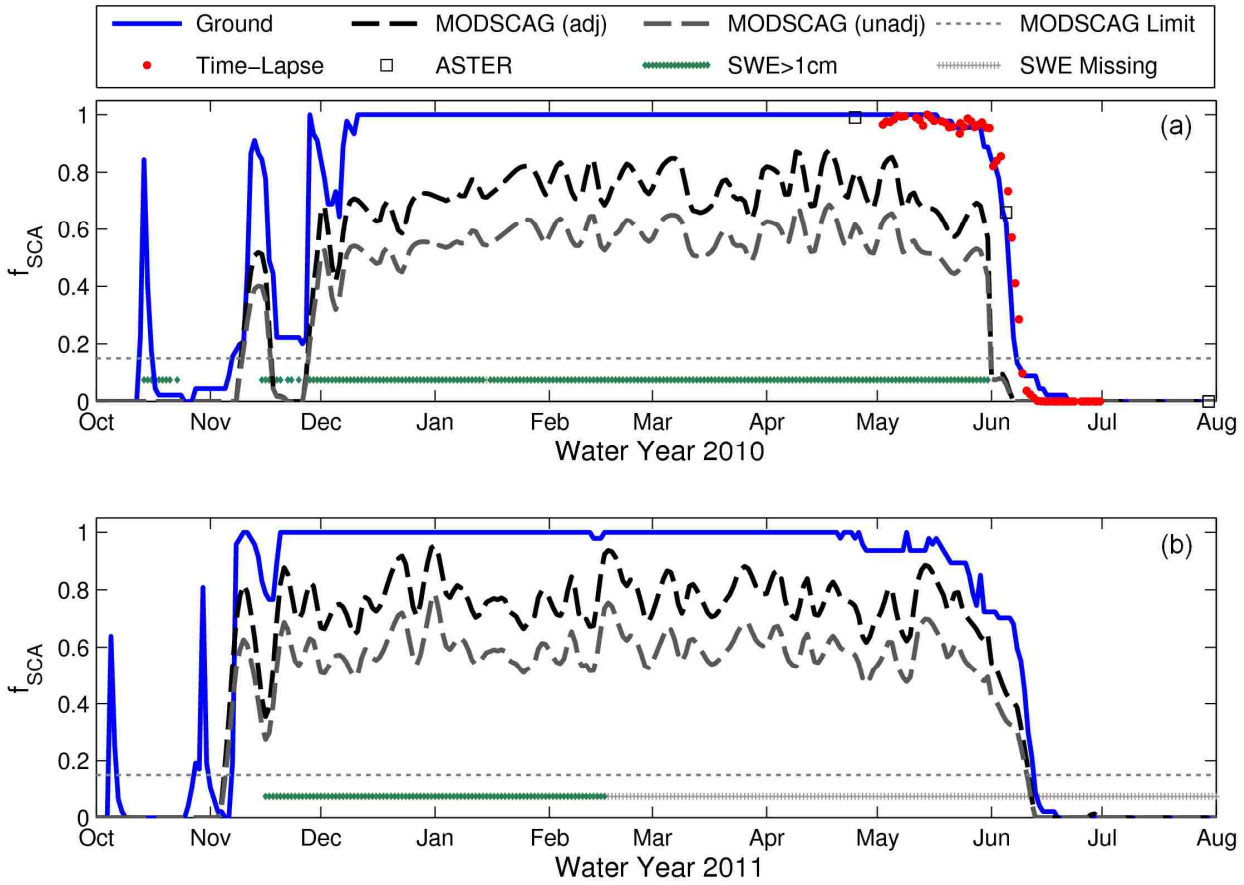

Figure 3.6 Fractional snow-covered area (f_{SCA}) at Tuolumne Meadows (TUM) during water years (a) 2010 and (b) 2011. Shown in both years are f_{SCA} from the ground temperature network, and MODSCAG f_{SCA} before and after the canopy adjustment. In May-June 2010, independent f_{SCA} data from a time-lapse camera and ASTER were included for validation. Also shown are periods when the TUM snow pillow reported snow presence (SWE>1cm) and when SWE data were missing.

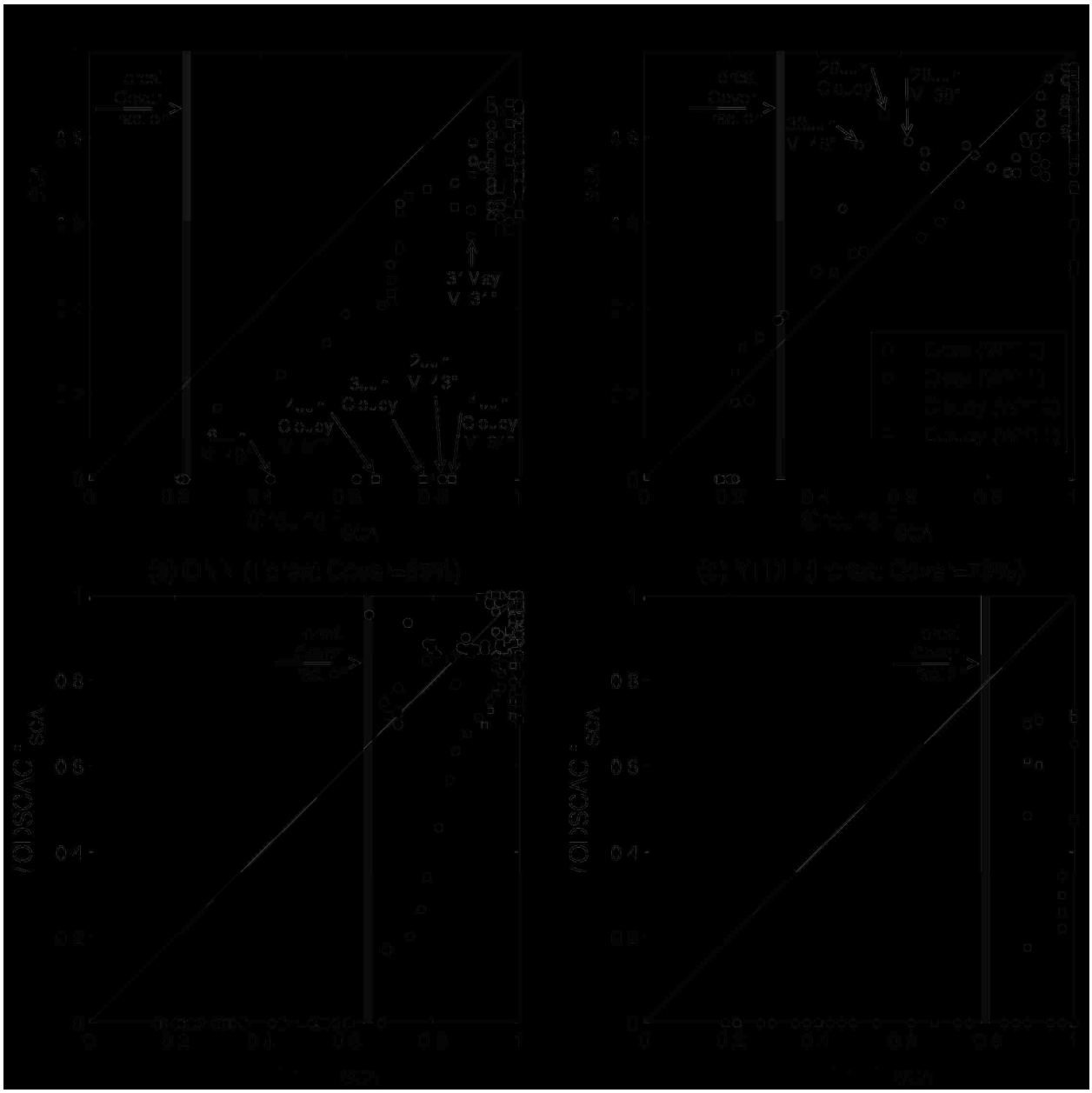

Figure 3.7 Ground fractional snow-covered area (f_{SCA}) vs. canopy-adjusted MODSCAG f_{SCA} at (a) Tuolumne Meadows, (b) Dana Meadows, (c) Onion Creek, and (d) the Yosemite Forest Dynamics Plot from 1 April to melt out during water years 2010 and 2011. Time generally progresses from the upper right corner to the lower left as the snow cover disappears. Points are classified based on whether cloudy or clear conditions prevailed, based on pyranometer data and MODIS visible imagery. Points during a one week period in WY 2010 at TUM (a) and during a three day period in WY 2011 at DAN (b) are labeled to indicate a combination of cloudy days and off-nadir view angles (V) that introduced errors in MODSCAG f_{SCA}. The NLCD forest cover fraction is also plotted, showing that MODSCAG f_{SCA} drops to 0 at the forest sites (c, d) as the ground f_{SCA} approaches the forest fraction.

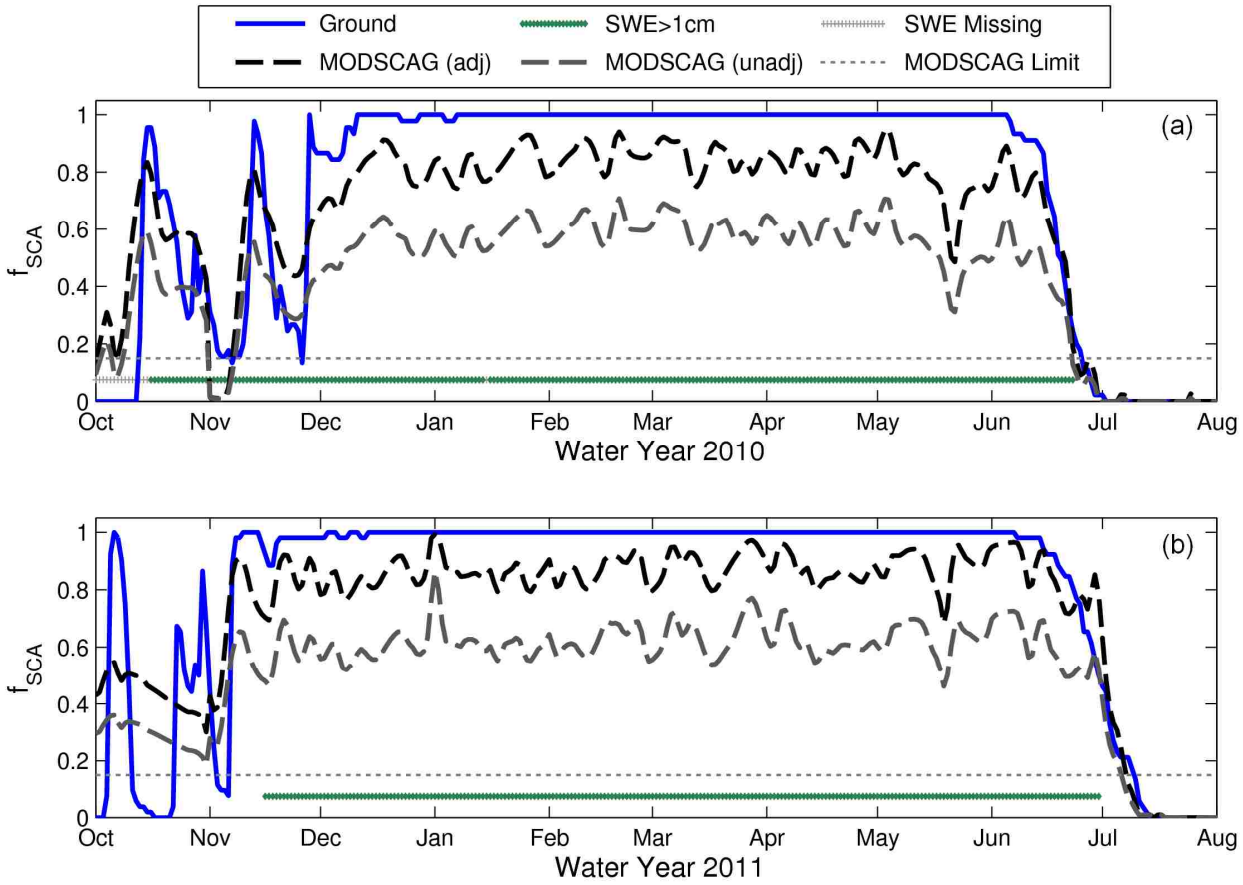

Figure 3.8 Fractional snow-covered area (f_{SCA}) at Dana Meadows (DAN) from the ground temperature network and MODSCAG during water years (a) 2010 and (b) 2011. MODSCAG f_{SCA} is shown before and after the canopy adjustment. Also shown are periods when the DAN snow pillow reported snow presence (SWE>1cm) and when SWE data were missing.

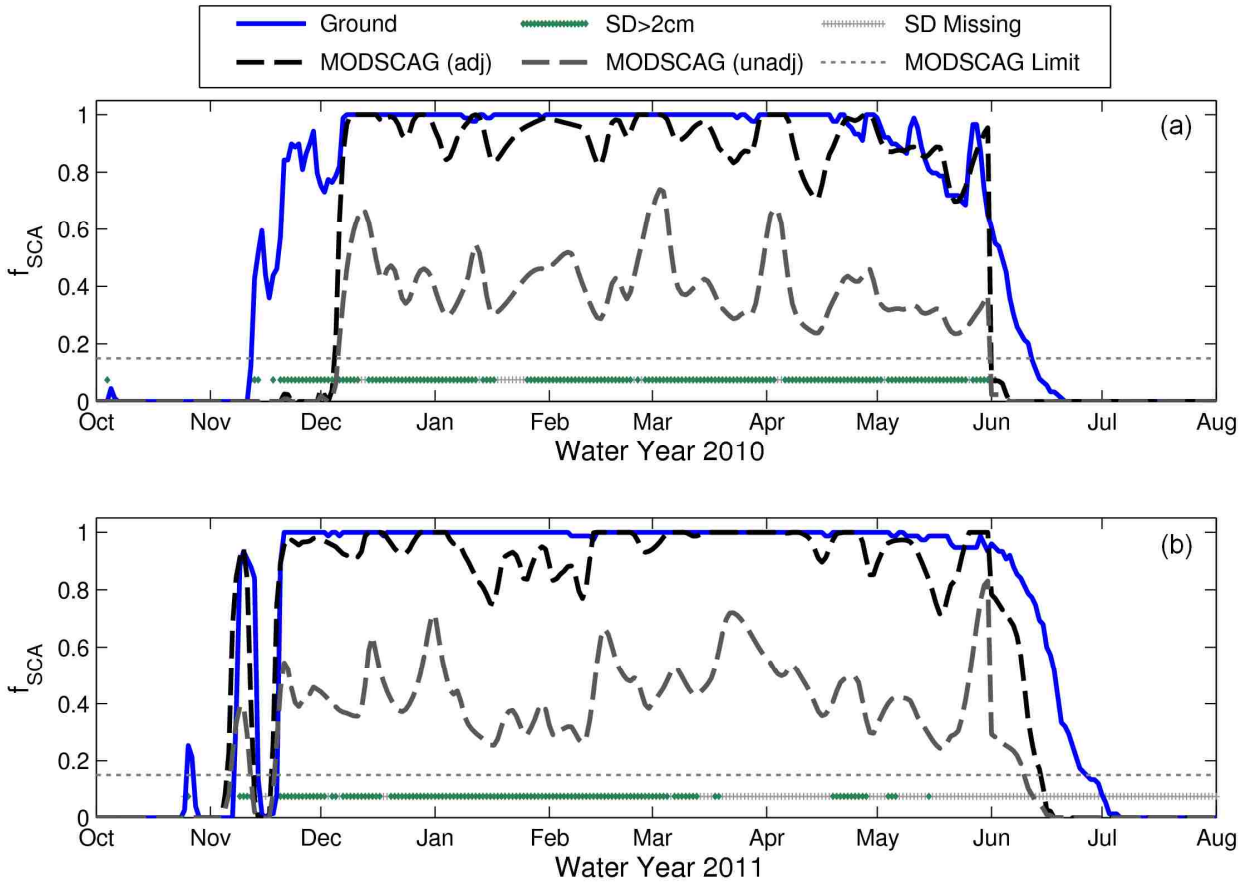

Figure 3.9 Same as Figure 3.8, except at the Onion Creek Experimental Forest (ONN), and with periods shown when snow depth (SD) exceeded 2 cm during 2010 and 2011. Most snow depth measurements were unavailable after March 2011 when the snow depth sensor arm was bent by heavy snow accumulation.

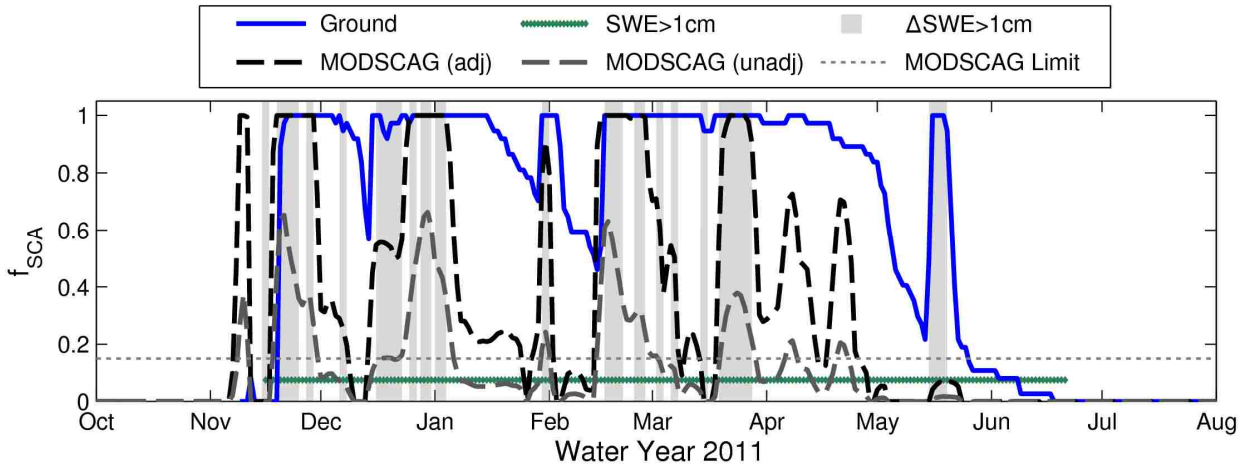

Figure 3.10 Same as Figure 3.8, except at the Yosemite Forest Dynamics Plot (YFDP) and during water year 2011 only. SWE values are taken from the Gin Flat snow pillow, 4 km from YFDP and 300 m higher in elevation. The shaded regions indicate periods when new SWE accumulation exceed 1 cm. Also shown are periods when the Gin Flat snow pillow reported snow presence (SWE>1cm). Note that the ground sensors did not begin recording data until 9 November 2010, so the first snow storm reported by MODSCAG was not evaluated.

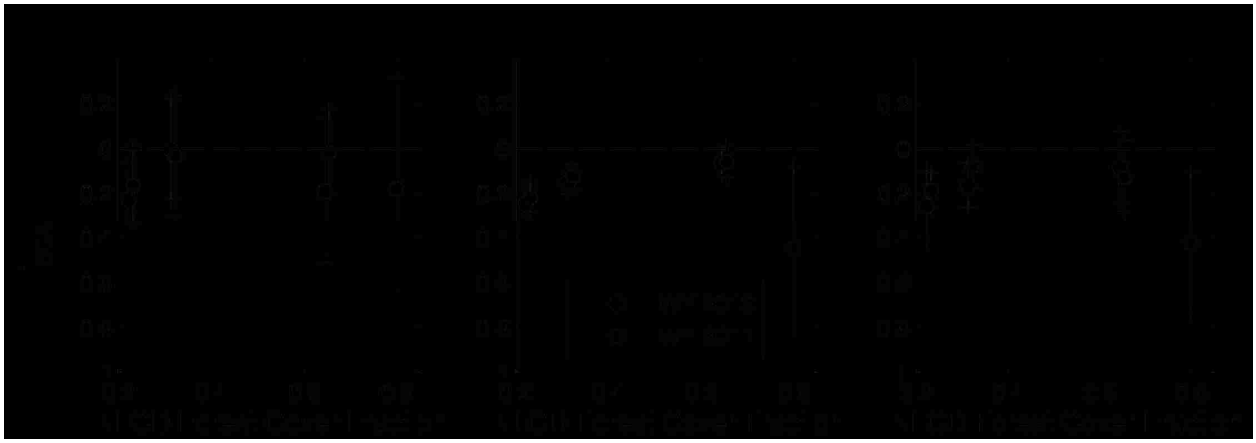

Figure 3.11 Daily errors in canopy-adjusted MODSCAG f_{SCA} during water years 2010 and 2011 versus mean NLCD forest cover during (a) the early accumulation season (Start-Dec), (b) winter (Jan-Mar), and (c) the ablation season (Apr-Disappearance). The starting date in (a) was the first day with $f_{SCA} \geq 0.15$, while the disappearance date in (c) was the last day when $f_{SCA} \geq 0.15$. Mean errors are denoted by a circle and whiskers indicate 1 standard deviation from the mean error. The markers were displaced horizontally so the two water years could be shown without overlap at each site. The four sites listed in increasing forest cover are TUM, DAN, ONN, and YFDP.

Chapter 4 Standard temperature and humidity approximate snow surface temperature: new possibilities for snow model calibration

This chapter was submitted in its current form to *Water Resources Research* [*Raleigh et al.*, *submitted* 2013a] in April 2013.

Abstract

Snow surface temperature (T_s) is an important variable in the snowmelt energy balance, land-atmosphere interactions, weak layer formation (avalanche risk), and winter recreation, but is rarely measured at observational networks. Reliable T_s datasets are needed to validate remote sensing and distributed modeling, in order to represent land-atmosphere feedbacks. Previous research demonstrated that the dew point temperature (T_d) close to the snow surface approximates T_s well because air is saturated immediately above snow. However, standard height (2 to 4 m) measurements of the saturation temperatures, T_d and wet-bulb temperature (T_w), are much more readily available. There is limited understanding of how these standard variables approximate T_s, and how the approximations vary with climate, seasonality, time of day, and atmospheric conditions (stability and radiation). We used sub-daily measurements from seven sites to test T_s approximations with standard height temperature and moisture. T_d produced the lowest bias (-2.2 °C to +2.6 °C) and root mean squared error (RMSE) when approximating T_s, but tended to underestimate daily extremes in T_s. For comparison, air temperature (T_a) was biased +3.2 °C to +6.8 °C. T_s biases increased with increasing frequency in nighttime stability and daytime clear sky conditions. We illustrate that mean daily T_d can be used to detect systematic input data bias in physically-based snowmelt modeling, a useful tool for model calibration in data sparse regions. Thus, improved understanding of T_d variations can advance understanding of T_s in space and time, providing a simple yet robust measure of surface feedback to the atmospheric energy budget.

4.1 Introduction

The surface (i.e., skin) temperature of snow (T_s) is a critical factor in the snow energy balance and in land-atmosphere interactions, modulating how much energy is used to warm or melt a snowpack and how much energy is returned to the atmosphere. Specifically, T_s controls outgoing longwave radiation and regulates the near-surface profiles of temperature and vapor pressure that influence sensible and latent heat transfer. Increases in T_s drive the growth of snow grains, reducing snow surface albedo and enhancing absorbed shortwave radiation [*Wiscombe and Warren*, 1980]. T_s is also a critical factor for slab avalanche formation, as large fluctuations in T_s can favor the development of unstable layers in a snowpack [*Armstrong and Armstrong*,

1987; *Birkeland*, 1998; *Birkeland et al.*, 1998]. Finally, T_s is an important variable for winter recreation (e.g., downhill skiing and cross-country skiing races), as it has a non-linear relationship with surface friction [e.g., *Wagner and Horel*, 2011].

Despite its importance, T_s is rarely measured at existing observational networks [*Bales et al.*, 2006] and must be acquired with remote sensing or modeling. Satellite-based remote sensing of surface emission at infrared wavelengths has yielded T_s datasets, but atmospheric emission of longwave radiation complicates this methodology [*Duguay*, 1993], and few studies have validated remotely sensed T_s, as noted by *Dozier and Painter* [2004]. In modeling applications, T_s has been estimated with (1) empirical relationships that track hourly [*Brubaker et al.*, 1996; *Marks et al.*, 1992] or mean daily [*Molotch*, 2009] air temperature, (2) conceptual approaches based on air temperature that incorporate radiative cooling effects [*Marsh and Pomeroy*, 1996; *Pohl et al.*, 2006], (3) longwave-based psychrometric formulations [e.g., *Ellis et al.*, 2010], and (4) physically-based approaches that solve for T_s in the surface energy balance using analytic [e.g., *Essery and Etchevers*, 2004; *Kondo and Yamazaki*, 1990] or iterative solutions [e.g., *Jordan*, 1991; *Tarboton and Luce*, 1996]. Due to the inherent scarcity of T_s observations, models generally lack validation of this important parameter. Instead, snow models are typically evaluated only against snow water equivalent (SWE) [*Essery and Etchevers*, 2004], a practice which neglects the contributing processes and limits process-based understanding [*Clark et al.*, 2011]. Depending on model selection, uncertainty in mid-winter T_s may be as large as 8 °C to 10 °C, [*Essery et al.*, 2013; *Slater et al.*, 2001], yielding up to 40 W m^{-2} of uncertainty in longwave emitted to the atmosphere, and signifying problems in the modeled surface energy balance.

In contrast to the above methods for estimating T_s, *Andreas* [1986] hypothesized that the dew point temperature (T_d) of air close to the snow surface approximates T_s. The physical reason

for this approximation is that snow cover is a saturated surface, such that the vapor pressure (e) of air close to the surface equals the saturation vapor pressure (e_{sat}). Air reaches saturation at T_d and e_{sat} is a function of T_s alone; thus, it can be seen that T_d close to the snow surface is expected be in equilibrium with T_s. Supporting this reasoning, *USACE* [1956] notes that "The vapor pressure has a strong tendency to remain close to that of the snow surface since the snowpack is both a sink and a source for vapor pressure greater or less than that of the snow. For air over a melting snowpack, the tendency is thus toward a vapor pressure of 6.11 millibars (the saturated vapor pressure at 32 °F)." While *Andreas* [1986] focused on T_d, there exists a second saturation temperature, the wet-bulb temperature (T_w) (see section 4.3), which is the temperature at which an air parcel becomes saturated reach through evaporative cooling. For an unsaturated air parcel, T_w is always greater than T_d. The relationship between T_w and T_s has seen little attention in the literature.

Andreas [1986] supported his hypothesis with theoretical analysis, and then demonstrated that T_s generally corresponded to T_d measurements 10 cm above the snow surface during January 1984 at a field site in Michigan. His theoretical analysis showed T_d is representative of T_s during periods with enhanced mixing (i.e., high wind speed) or when the near-surface vapor pressure gradient weakens (e.g., when e_{air} comes into equilibrium with $e_{sat}(T_s)$). Based on his field measurements, he found that approximation of T_s with T_d was accurate to ±1 °C. However, he found that T_s was higher than T_d during sunny periods and assumed that solar heating had biased the infrared pyrometer that measured T_s. *Andreas et al.* [2002] assessed the relationship between saturation temperatures and other components of the cryosphere, including polar sea ice. They found that air was close to saturation with respect to ice during all months, and these temperatures differed from ice surface temperatures by less than 1.6 °C.

To date, the results of *Andreas* [1986] have seen limited application in snow hydrology research, perhaps because T_d measurements are rarely taken close to the snow surface [*Marks et al.*, 1992]. Temperature sensors and hygrometers are typically installed at a standard height 2 m to 4 m above the ground surface at a climate station, with the height above the snow surface varying as the snow depth fluctuates with accumulation, wind scour/deposition, compaction, sublimation, and melt processes. Therefore, it remains unknown how saturation temperatures at standard height relate to T_s, how these relationships vary with climate and local conditions (i.e., boundary layer stability and radiation), and whether they improve T_s approximation over traditional methods. Furthermore, the relationship of T_s to saturation temperatures is uncertain during snowmelt periods, as the January experiment of [*Andreas*, 1986] was generally limited to T_s below -5 °C.

The purpose of this paper is to test representations of T_s with standard height measurements across a more complete range of seasonally-snow-covered environments, climates, and seasons. We specifically address four key questions: (1) How do measurements of T_s compare to standard height dry bulb (i.e., air) temperature (T_a), wet-bulb temperature (T_w), and dew point temperature (T_d)? (2) How well does standard height T_d represent T_s across different climates and seasons? (3) How well does T_d approximate T_s with variations in atmospheric conditions (i.e., radiation and stability)? We address the above questions using measurements at seven sites located across North America and Europe. To illustrate the relevance of these results to model calibration and validation, we also compared T_d measurements with those simulated by a physically-based snow model [SNTHERM; *Jordan*, 1991] for both cases of best-available forcing data and of biased radiative energy balance forcing.

4.2 Sites and Data

We selected study sites in both wet and dry climates that featured all surface observations required to test and compare the T_s approximations (Table 4.1, Figure 4.1). Météo-France provided an 18-year dataset from the Col de Porte site (CDP) in a subalpine maritime environment [*Morin et al.*, 2012]. The United States Army Corps of Engineers Cold Regions Research and Engineering Laboratory (CRREL) provided data at a low elevation, moist continental site (South Royalton Vermont, SRV) [*Peck and Fiori*, 1992]. The IP3 (Improved Processes and Parameterizations for Prediction in Cold Regions) Research Network (http://www.usask.ca/ip3/) provided data from stations in a wet alpine environment (Opabin, OPB) [*Hood and Hayashi*, 2010] and a subarctic bog in the zone of discontinuous permafrost (Scotty Bog, BOG) [*Williams et al.*, 2013]. The University of Calgary provided data from a cold grassland prairie (Spy Hill, SPY) [*Mohammed et al.*, 2013]. The Center for Snow and Avalanche Studies (http://www.snowstudies.org) provided data from paired weather stations in a dry continental climate (Colorado, USA), including a subalpine site (Swamp Angel Study Plot, SASP) and an alpine site (Senator Beck Study Plot, SBSP) [*CSAS*, 2012]. Although SASP and SBSP were located in the same climate, they provided insights into windy (SBSP) vs. sheltered (SASP) locations. Mean winter (DJF) temperatures at the sites ranged from -0.90 to -22 °C (Table 4.1).

The study period depended on data availability at each site, and ranged from one to 18 snow years (Table 4.1). All sites recorded 30-min or hourly measurements of air temperature, humidity, wind speed, snow depth, downwelling shortwave and longwave radiation, and either infrared-measured T_s or upwelling longwave radiation (i.e., from the downward-pointing pyrgeometer on a net radiometer). The meteorological measurements were made at a fixed

height above the ground at all sites except for CDP, where operators adjusted the instrument heights weekly to maintain a consistent height above the snow surface [*Morin et al.*, 2012].

Upwelling longwave radiation was converted to T_s using the Stefan-Boltzmann equation with snow thermal emissivity (ε) equal to 1, for consistency with the published dataset at CDP [*Morin et al.*, 2012]. However, we note that ε may be as low as 0.97 [*Kondo and Yamazawa*, 1986], in which case reflected longwave radiation must be considered when calculating T_s from upwelling longwave data. We checked the impact of ε selection at the four sites with pyrgeometers, and found that the mean difference between $T_s(\varepsilon=1)$ and $T_s(\varepsilon=0.97)$ was 0.2 °C, and therefore our selection of ε=1 did not significantly alter the results of the study. At CDP, T_s was measured directly with an infrared thermometer and indirectly with a downward-pointing pyrgeometer; only the colder T_s measurement was reported each hour in the dataset published by *Morin et al.* [2012]. Infrared measurements of T_s were available during two winters to check the pyrgeometer measurements at BOG, and we found that the mean difference in the two sensors was ±2 °C in the -30 °C to 0 °C range, but larger differences were present when the temperature was below -30 °C. This test provided confidence that the type of sensor (infrared sensor vs. pyrgeometer) did not significantly bias the results of the study. However, due to site-to-site differences in sensor field of view, we qualitatively considered the impact of sensor type when interpreting the results of the study. A full assessment of the relative uncertainties of these different sensors was outside the scope of the study. Based on manufacturer's specifications, both types of sensors had temperature-dependent uncertainties in measured T_s (Table 4.1).

Only T_s measurements in the -50 °C to 0 °C range were considered in the analysis. T_s observations exceeding 0 °C with snow cover present were manually set to the physical limit of 0 °C. Averaged across the sites, these cases represented 4.6% of observations, but 19% of

observations at CDP exceeded 0 °C (but all were less than 2 °C). These erroneous values may occur due to calibration errors, radiative heating errors, or when non-snow surfaces (e.g., vegetation, bare ground) are located in the field of view of the infrared thermometer or pyrgeometer. We further constrained the analysis to periods when measured snow depth exceeded 10 cm to minimize impacts from herbaceous vegetation protruding out of the snow. Debris on the snow surface may also alter the radiometric properties of the snow, and thus presents additional uncertainty in the T_s measurements. This was relevant at CDP where tree litter and atmospheric dust have been documented [*Etchevers et al.*, 2004], and at SASP and SBSP where dust deposition on the snow surface is common [*Painter et al.*, 2007].

4.3 Calculation of meteorological variables and conditions

4.3.1 Wet-bulb temperature (T_w)

The wet-bulb temperature (T_w) is the temperature that an air parcel would reach if cooled to saturation at constant pressure through evaporation of water into the parcel, where the parcel provides the latent heat for evaporation. When temperature is below the freezing point, the wet-bulb is sometimes referred to as the ice-bulb temperature. T_w can be measured in the field with a sling psychrometer, but these measurements were not made at the sites. In the absence of such measurements, T_w is commonly estimated with psychrometric charts or iterative numerical based on temperature, humidity, and atmospheric pressure (P_{atm}). To estimate T_w, we used the iterative solution described by *Iribarne and Godson* [1981], a common tool that is used both operationally (e.g., in the Advanced Weather Interactive Processing System of the US National Weather Service) and in research models (e.g., the Distributed Hydrology Soil Vegetation Model, [*Wigmosta et al.*, 1994]). P_{atm} was estimated at a constant value using an empirical relationship between elevation and pressure at each site, except at CDP and SASP where hourly P_{atm}

measurements were available. Using observations at CDP and SASP, we tested the assumption of temporally constant P_{atm} vs. hourly P_{atm} for estimating T_w, and found a maximum difference of 0.3 °C between the two approaches. Thus, assuming temporally constant P_{atm} introduced little error into T_w estimation.

4.3.2 Dew point temperature (T_d)

T_d is defined as the temperature that an air parcel would reach if cooled to saturation with respect to water at constant pressure without changes in moisture content. The frost point temperature is analogous to T_d, except saturation is considered with respect to an ice surface instead of water. For simplicity, we refer to both as T_d, but note that we are actually calculating frost point when $T_a \leq 0$ °C and the dew point when $T_a > 0$ °C. Here we assume that there is no freezing-point depression for liquid water, and that the water vapor and snow have no impurities. For an unsaturated air parcel, T_d is lower than the other standard temperatures, where $T_d < T_w < T_a$. For a saturated air parcel (i.e., e=e_{sat}), $T_d = T_w = T_a$. The difference between T_a and T_d is called the dew point depression, which is a measure of air saturation.

T_d can be measured with a variety of hygrometers (e.g., capacitive, resistive, chilled mirror), but often saturation state or moisture content is instead reported as relative humidity (*RH*) or specific humidity in published datasets. We therefore used a Magnus-Tetens approach [*Murray*, 1967] to calculate T_d (°C) as a function of T_a (°C) and fractional *RH* at each time step:

$$T_d = \frac{c\left[\ln(RH) + \frac{bT_a}{c+T_a}\right]}{b - \ln(RH) - \frac{bT_a}{c+T_a}} \tag{4.1}$$

Alduchov and Eskridge [1996] provide different coefficients for Equation 4.1 depending on whether saturation is taken with respect to water ($e_{sat,w}$) or with respect to ice ($e_{sat,i}$). When $T_a >$

0 °C we use the $e_{sat,w}$ coefficients b = 17.625 and c = 243.04 °C, and when $T_a \leq 0$ °C we use the $e_{sat,i}$ coefficients b = 22.587 and c = 273.86 °C. Relative to other common approximations, maximum relative error with this approach is 0.384% for calculating vapor pressure [*Alduchov and Eskridge*, 1996]. This maximum error in e_{sat} translates to a maximum error of 0.1°C in T_d. Thus, uncertainty in T_d arises primarily from uncertainties in measured T_a and RH.

4.3.3 Boundary layer stability

We hypothesize that atmospheric stability plays a role in the representation of T_s with the standard height temperatures (i.e., T_a, T_w, T_d). *Andreas* [1986] suggested this relationship by considering how increasing wind speed and sensible heat flux impact the difference between T_s with T_d, but atmospheric stability was not directly considered. We expect that the standard height T_d best represents T_s during unstable conditions when turbulent mixing reduces temperature stratification. During stable conditions, T_d at standard height should become decoupled from T_s.

To characterize boundary layer stability, we calculated the bulk Richardson number (Ri_b) [*Singh and Frevert*, 2002] at each time step:

$$Ri_b = \frac{gz(T_a - T_s)}{0.5(T_a + T_s)U(z)^2} \tag{4.2}$$

where g is gravitational acceleration (m s^{-2}), z is the height of the temperature sensor above the snow surface (m), temperature values are in Kelvin, and $U(z)$ is wind speed (m s^{-1}) at height z. Because wind speed was not measured at the same height as temperature at every study site (Table 4.1), we rescaled the measured wind speed (U_{obs}) to the height z at each time step assuming a power law wind profile:

$$U(z) = U_{obs}\left(\frac{z}{z_u}\right)^{\alpha} \tag{4.3}$$

where z_u is the height of the wind speed measurement (m) and α is 1/7, an average value assumed when the wind profile is unknown [*Peterson and Hennessey*, 1978].

The Richardson number compares buoyancy suppression of turbulence against turbulence generation to characterize stability conditions. Based on previous studies, a critical Richardson number in the 0.15 to 0.25 range separates unstable and stable conditions, with unstable (stable) conditions prevailing below (above) the critical value. We assume that the critical number is 0.20 at the study sites, but note that there is disagreement about the critical number over snow [for a discussion, see *Andreas*, 2002]. We note that many physically-based models [e.g., *Jordan*, 1991] implement a correction based on Ri_b when computing turbulent fluxes.

4.3.4 Radiative heating and clearness index

Because incoming radiation is the primary energy source for heating and melting a snowpack in many climates, it is important to consider how radiative heating impacts T_s and approximations thereof. Solar heating of instruments may also bias measurements of T_s [*Andreas*, 1986] and standard height temperature [*Huwald et al.*, 2009]. Incoming shortwave and longwave radiation vary with slope, elevation, time of day, time of year, forest cover, and cloud cover. For simplicity, we only consider how radiation changes in time and with sky conditions (i.e., clearness vs. cloudiness). To approximate sky conditions, we calculated a dimensionless clearness index (*CI*) at each time step:

$$CI = \frac{Q_{si,obs}}{Q_{si,pot}} \tag{4.4}$$

where $0 \leq CI \leq 1$, $Q_{si,obs}$ is measured incoming shortwave radiation (W m^{-2}) and $Q_{si,pot}$ is potential (i.e. clear sky) incoming shortwave radiation (W m^{-2}). $Q_{si,obs}$ was observed at all sites, however, we used the SBSP observations at SASP because of the tendency for the radiometer dome to become snow covered at SASP where it is less windy than SBSP. $Q_{si,pot}$ was calculated as a function of time of day, time of year, latitude, humidity, temperature, and elevation according to the equations of *Crawford and Duchon* [1999]. Equation 4.4 does not yield valid *CI* during the night, so we linearly interpolate between sunset and the following sunrise to approximate nighttime values of *CI*. High values of *CI* indicate sunny conditions during the day, and surface cooling at night. We assume clear sky conditions prevail when $CI > 0.5$ and cloudy conditions prevail when $CI < 0.5$. No adjustments were made for shading in $Q_{si,pot}$, however we acknowledge that topographic shading is a significant issue in complex terrain [*Duguay*, 1993]. Because *CI* was only used to classify cloudy vs. clear sky conditions, we expected minimal impact in the *CI* analysis from topographic shading. At sites and times of the year where topographic shading was important, this effect was similar to cloudy conditions near sunrise or sunset.

4.4 Physically-based snow modeling experiment

To illustrate how improved approximation of T_s might benefit snow modeling studies, we considered how biases in the energy balance of a physically-based snow model become manifested in T_s and whether T_d can be used to detect the bias. This was a relevant application, as data are rarely available to validate the energy balance or T_s in snowmelt modeling studies. We selected the 1-dimensional, multi-layer Snow Thermal Model (SNTHERM) [*Jordan*, 1991] for the physically-based simulation of T_s at the CDP site during water year 2006. SNTHERM was developed specifically for the prediction of T_s based on energy exchanges at the snow-atmosphere interface and is regarded as one of the more sophisticated and reliable snow models

available [*Etchevers et al.*, 2004]. SNTHERM simulates various snowpack processes, including snow accumulation, frost development, compaction, metamorphosis, grain growth, sublimation, and snowmelt. *Watson et al.* [2006] argued that SNTHERM is the "benchmark model" for physically-based simulation of snowmelt processes, though the model complexity has limited its application in distributed modeling problems.

SNTHERM simulates snowpack development by dividing new snowfall into horizontally infinite layers. The governing equations for energy and mass balance are applied between these layers, with the meteorological conditions applied at the upper (i.e., snow-atmosphere) boundary and steady state conditions assumed at the lower (i.e., snow-soil) boundary. An iterative numerical solution yields estimates of T_s, and layer specific states for thickness, density, temperature, and phase. Iterative solution of T_s from the energy balance is possible because multiple processes are functions of T_s, including outgoing longwave radiation, sensible heat, latent heat, and heat conduction into the snowpack [*Jordan*, 1991; *Liston and Elder*, 2006; *Tarboton and Luce*, 1996]. Physically-based, iterative approaches for estimating T_s have disadvantages in that they require accurate forcing data for the other components of the energy balance [*Pomeroy et al.*, 1998] and they can be computationally expensive to reach convergence [*Wigmosta et al.*, 1994]. Despite being physically-based, some parameters in these models require calibration (e.g., thermal conductivity) [*Essery and Etchevers*, 2004; *Tarboton and Luce*, 1996], which yields additional uncertainties in modeled T_s and the surface energy balance.

To understand the relationship between energy balance bias and T_s bias, we simulated snowpack with SNTHERM with a control simulation and with biased forcing data scenarios. For the control simulation, we used local observations (i.e., T_a, RH, wind speed, incoming shortwave radiation, outgoing shortwave radiation, and incoming longwave radiation) to drive

the model and compared modeled and observed values of T_s. This provided a benchmark for understanding the accuracy of modeled T_s under an ideal driving data scenario. We also compared T_d and T_s for reference. We then continued the experiment by introducing artificial biases in the incoming radiation data, ranging from -15% to +15%, and then modeled T_s with SNTHERM with these biased datasets. These scenarios with biased radiation resembled the reality of snowmelt modeling in a data sparse environment, as biases in model data and/or model structure can propagate to model outputs (e.g. T_s, SWE) in ways that are not easily detected. We then compared the T_s simulations generated with biased model data against both measured T_s and T_d to determine whether these yielded similar relationships. Finally, we compared the artificial energy balance biases against the mean difference between modeled T_s and T_d, to determine whether T_d could be used to detect bias in a snowmelt model. All temperature values were aggregated from hourly to daily mean values in this modeling experiment for simplicity.

4.5 Results and discussion

4.5.1 Comparing T_s *with standard height temperatures*

Sub-daily comparisons of the standard height temperatures (T_a, T_w, and T_d) and observed T_s are presented in Figures 2 and 3. We summarize the results with three performance metrics (i.e., R^2, bias, root mean squared error (RMSE)), assuming each has unique implications for different applications. The comparisons showed that despite reasonable correlation between T_a and T_s, T_a was consistently higher than T_s, and the magnitude of this bias was inconsistent between sites (Figure 4.2a-g, Figure 4.3b). Averaged across the sites, $T_a > T_s$ in 95% of measurements. For T_a, we also considered lagged correlations with T_s, but found that the greatest correlation was with no lag at four of the seven sites.

Like T_a, T_w also exhibited a consistent warm bias (Figure 4.3b), although this bias decreased with increasing T_s (Figure 4.2h-n). Despite this warm bias, T_w had the strongest correlation with T_s at all seven sites (Figure 4.3a). The correlation improved slightly at only two of the seven sites when a lag correlation was attempted (no figure.shown).

Of the three standard height temperature variables, T_d had the lowest absolute bias and RMSE when approximating T_s (Figures 3b, 3c), but underestimated T_s more frequently than T_a and T_w at warmer T_s (Figure 4.2o-u) and did not have stronger correlations with T_s than T_w had with T_s (Figure 4.3a). When approximating T_s with standard height T_d, the bias ranged from -2.2 °C to +2.6 °C. This error range was larger than the ±1 °C accuracy reported by *Andreas* [1986] when using 10 cm height T_d. However, at four of the study sites (SPY, SASP, OPB, BOG), the bias ranged from -1.0 °C to 0.0 °C, well within the range of *Andreas* [1986]. Thus, climate and environmental conditions may have played as large of a role in the bias as the measurement heights of T_d. These results were not linked to the type of instrument used to measure T_s.

In approximating daily maxima and minima of T_s, the saturation temperatures (i.e., T_w and T_d) provided improved representation over T_a (Figure 4.4). At all sites, T_a overestimated daily maxima and minima of T_s. At five of the seven sites, maximum daily T_s values were typically between T_w and T_d (Figure 4.4a). Both T_w and T_d overestimated minimum daily T_s at all sites, except at SBSP where T_d underestimated minimum daily T_s (Figure 4.4b). Daily minimum T_s was typically colder than standard height T_d, suggesting that stable conditions at night decoupled standard height T_d and T_s (see section 4.5c).

The results indicate that to first order, T_d represents average and minimum daily T_s with the least bias, and that the saturation temperatures (T_w and T_d) characterize maximum daily T_s more accurately than T_a. None of the standard height temperatures captured both the daily

maxima and minima of T_s because their diurnal ranges tended to be much lower than the T_s diurnal range (Table 4.2). Bias in maximum daily T_s might thus be reduced using both T_w and T_d as predictors, but development of such a method was outside the scope of this study.

These comparisons demonstrated the inherent difficulty in empirically estimating T_s based on T_a across different climates. Methods of estimating T_s based on a static offset from T_a attempt to correct the estimation bias, which ranged from +3.2 °C to +6.8 °C at the seven study sites. However, the bias cannot be known a priori, and can vary from published values even within a single climatic zone. For example, *Brubaker et al.* [1996] found that T_a was typically +2.5 °C higher than T_s at a site in Vermont, USA, but our data in Vermont (SRV) suggests the difference is +6.8 °C. Using the 2.5 °C offset, a bias of 4.3 °C would remain, exceeding the 2.6 °C bias found from assuming that T_d approximated T_s. We therefore expect methods that estimate T_s based on an offset from T_a have limited usefulness because their transferability is questionable.

4.5.2 Variation of results with climate and season

To characterize how the standard height approximations of T_s varied with climate, we tested correlations between daily RMSE statistics and six predictors (elevation, latitude, mean winter temperature, mean annual shortwave radiation, mean dew point depression, and mean wind speed). The only significant correlation between these predictors was between elevation and mean annual shortwave radiation ($p=0.036$) and between elevation and mean dew point depression ($p=0.023$). Mean winter temperature was not significantly correlated ($p=0.14$) with latitude, due to the variety of elevations and climates represented by the sites.

When testing correlations between the six predictors and RMSE, we only found significant relationships between maximum daily T_s RMSE and shortwave radiation and dew point depression, and between mean daily T_s RMSE and latitude (Figure 4.5). When

approximating maximum daily T_s with T_d , RMSE significantly increased (r=0.827, p=0.022) with increasing shortwave radiation (i.e., sunnier conditions,) and significantly increased (r=0.823, p=0.023) with increasing dew point depression (i.e., drier conditions). These predictors suggested that climatic conditions impacted the approximation of T_s with T_d. Specifically, sites with more solar radiation and drier conditions tended to have more variable differences between T_s and T_d. RMSE of mean daily T_s declined with increasing latitude, and this was statistically significant for T_a (p=0.022) and T_d (p=0.004), but not for T_w. The relationship between latitude and RMSE was strongest (r=-0.913) when using T_d to approximate mean daily T_s (Figure 4.5d).

We also examined temperature variations with respect to time of day and season by averaging across all hours with snow (i.e., depth > 10 cm) in January and April (Figure 4.6). During the nighttime hours, T_d generally provided the best approximation of T_s, although T_d was typically higher than T_s by several degrees at three sites (CDP, SRV, SASP) where stable conditions prevailed (see section 4.5c). This relationship did not appear linked to the instrument used to measure T_s, as T_s was measured with an infrared sensor at SASP, with a pyrgeometer at SRV, and with both types of instruments at CDP. During the daytime hours, T_d was generally lower than T_s at all sites except CDP and SRV. At CDP and SRV, radiative heating raised T_s close to T_d during midday, offsetting the nighttime positive bias of T_d. This effect was consistent in both winter (Figure 4.6a, c) and spring (Figure 4.6b, d). At the other sites, however, the daytime difference between T_s and T_d was more pronounced in April than in January (Figure 4.6), which provided additional evidence that radiative heating from increasing insolation caused T_s observations to rise rapidly. These results suggest that diurnal variations in stability and radiation, and seasonal variations in radiation impact how well standard height temperatures

represent T_s. *Andreas* [1986] conducted his experiment in January, but our multi-year results suggest that larger midday differences between T_s and T_d are found during times of the year with increased insolation. Stability and radiation are considered more in depth in the next section.

4.5.3 Impacts of stability and radiation

Because radiation interacts with boundary layer stability, it is difficult to completely distinguish their independent effects. We used a compositing approach to examine how each of these factors contributed to bias when approximating T_s with standard height T_d. To assess how stability impacts the approximation, we first calculated the difference between T_s and T_d and binned the results using the critical Ri_b as a classifier (Figure 4.7). Relative to unstable conditions, the difference between T_d and T_s tended to increase during stable conditions, such that T_s was generally lower than T_d. Such a cooling effect may arise at night due to surface cooling at night or during the day due to shallow stable layer formation over melting snow cover [*Halberstam and Schieldge*, 1981; *Mahrt and Vickers*, 2005]. Median $T_d - T_s$ was closest to 0 °C during unstable conditions at five of the sites with statistical significance (95% level, rank sum test).

Averaging the hourly bias at each site provided further clarification on how radiation and stability each affected approximations of T_s with T_d (Figure 4.8). Clear sky periods generally had $T_s < T_d$ at night and $T_s > T_d$ during the day (Figure 4.8a). Cloudy periods had similar differences between T_s and T_d at night, but smaller differences between T_s and T_d during the day. The midday underestimation bias was a persistent feature at most of the sites. Stable periods decoupled T_s and T_d at night, such that T_s was 3.0 °C lower than T_d when averaged across nighttime hours and across all sites (Figure 4.8b). In contrast, T_s measurements were typically higher than T_d by 0.5 °C or less during unstable periods at night. Regardless of stability

conditions, the effect of radiative heating caused $T_s > T_d$ during the daytime hours. When considering stability and radiation conditions together (Figure 4.8c), we found that cloudy and unstable periods had the lowest bias through the day. The midday bias was most reduced during cloudy and stable periods, but this was the result of offsetting effects from surface cooling (due to stable conditions) and surface heating (due to radiation). As noted previously, CDP exhibited this type of counteractive response due to stable and cloudy conditions.

The frequency of stable conditions and clear skies helped explain why approximating T_s with standard height T_d was more reliable at some sites and less reliable at others (Figure 4.9). Bias in approximating T_s with T_d at midday was not significantly related to the frequency of stable conditions (Figure 4.9a), but was significantly ($p=0.006$) related to the frequency of clear sky conditions (Figure 4.9b). At midnight, bias in the approximation was significantly related to both stability frequency (Figure 4.9c, $p=0.009$) and clear sky frequency (Figure 4.9d, $p=0.044$). However, we note there was uncertainty in clear conditions at night, as the clearness index was estimated from daytime measured solar radiation. The midnight bias was closer to 0 °C at sites where stable conditions were less frequent. In considering daily bias (Figure 4.3b), relationships with stability and clear sky frequencies were significant ($p<0.05$, Figure 4.9e-f). Figure 4.9e-f suggests that an optimal range of stability and clear sky conditions may exist for which T_d provides a reasonable approximation of T_s. However, low bias in daily average T_s does not imply low bias in extreme daily values or hourly values of T_s. For example, SASP had low bias in daily average T_s, but this was the result of compensating errors in midday underestimation bias and nighttime overestimation bias (Figure 4.6g).

Andreas [1986] originally found that T_d was lower than T_s on sunny days, but assumed that solar heating biased the infrared sensor that he used to measure T_s. Radiative heating

presents a source of uncertainty in the T_s measurements of our study and may explain some of the midday bias in approximating T_s with T_d. However, we also note that snow has been shown to exhibit remarkable fluctuations in temperature over sub-daily time scales in response to fluctuations in energy. In a controlled laboratory experiment, *Shea and Jamieson* [2011] documented that for a moderately dense snowpack (i.e., 270 kg m^{-3}), T_s increased 14 °C in response to 3 minutes of 175 W of infrared heating, and cooled back to the original temperature minutes after the heat source was removed. The field observations of *Wagner and Horel* [2011] also showed that T_s can exhibit large diurnal fluctuations, with temperature increasing from -17 °C to -0.3 °C over a six hour period. Our temperature data (Table 4.2) indicated that daily ranges of T_s were within these carefully measured values, and therefore we cannot attribute all of the midday bias to solar heating of the T_s sensors. With the available data, we could not distinguish T_s measurement errors from actual increases in T_s. However, simulations with the physically-based SNTHERM model indicated that these midday fluctuations were plausible, supporting the hypothesis that the observed increases in T_s were not the result of sensor heating (no figures shown).

4.5.4 Case study: comparisons with SNTHERM

The SNTHERM control simulation (i.e., all forcing data available) accurately represented T_s at CDP in water year 2006, with $R^2 = 0.97$, bias = +0.74 °C, and RMSE = 1.14 °C for mean daily T_s (Figure 4.10a). The control simulation tended to overestimate T_s mostly at colder temperatures but matched T_s better as the snow warmed to the melting point. These results demonstrated that the sophisticated snow model produced realistic simulations of T_s under an ideal data input scenario. For reference, the simple approximation of T_s with T_d yielded a similar bias in T_s (bias = +1.02 °C), although this T_d approximation had higher variability (R^2 =

0.75 and RMSE = 2.7 °C), especially at colder temperatures (Figure 4.10b). While the accuracy of the T_d approximation was lower than the snow model, it was accomplished with fewer data inputs (i.e., only temperature and humidity). The biases of both approaches were comparable to the highest accuracy (±0.7 °C) obtained with multiple snow models in the snow model intercomparison project [SnowMIP, *Etchevers et al.*, 2004]. The RMSE of the T_d approximation was also within the range (0.7-3 °C) for most models in SnowMIP.

The artificial biases introduced in the energy balance were directly correlated with the resulting bias in T_s, such that underestimation (overestimation) biases in the energy balance yielded underestimation (overestimation) biases in T_s (Figure 4.10c,e). Interestingly, similar biases were observed when comparing modeled T_s to observed T_s and T_d in the -10% energy bias case (Figure 4.10c-d), and when comparing modeled T_s to observed T_s and T_d in the +10% energy bias case (Figure 4.10e-f). In fact, the bias in the energy balance was strongly correlated ($p<0.001$) with the mean difference between modeled T_s and T_d (Figure 4.10g). Because T_d was shown to be a reasonable surrogate for T_s for daily averages (Figure 4.3b), the implication is that T_d has potential use as a diagnostic tool for detecting bias when simulating snowpack with a physically-based model. Specifically, our results suggest that large differences (i.e., exceeding 2 °C) between modeled T_s and T_d indicate the presence of bias in the energy balance (±10% or greater). Although this does not reveal the source of the bias (e.g., model structure, parameters, process representation, driving data), it provides a previously unrealized method for assessing bulk model performance. More work is needed to develop this methodology, but the results from this example show promise for model calibration and testing.

4.6 Summary and conclusions

In this study, we assessed how standard near-surface measurements of temperature and humidity (T_a, T_w, T_d) compared to measured values of T_s. We examined how climate, seasonality, time of day, boundary layer stability and radiation impacted T_s representation. We further compared T_s approximations based on T_d against T_s simulations from a physically-based, energy balance snowmelt model (SNTHERM), and demonstrated that T_d can be used to detect bias in the modeled snowpack energy balance . The goal was to provide a practical extension of the work of *Andreas* [1986] in order to understand how well standard height dew point temperature approximates T_s over a wider range of climates, seasons, and environmental conditions. The results demonstrated that the standard height saturation temperatures (T_w and T_d) represented T_s more accurately than T_a (Figures 2-4). Averaged across all samples, T_d yielded biases closer to 0 °C and the lowest RMSE. RMSE decreased with latitude and increased with shortwave radiation and dew point depression (Figures 5). However, T_d tended to be lower than maximum daily T_s and (at most sites) higher than minimum daily T_s (Figure 4.6). Biases were related to the frequency of stability and radiation conditions, and these effects varied with time of day and the frequency of those conditions (Figures 7-9). Results from a case study at CDP in water year 2006 showed that representing T_s with T_d is a potential tool for diagnosing forcing data bias in an energy balance model (Figure 4.10).

We therefore conclude that T_d at standard height is a reasonable first-order approximation of daily average T_s in many environments, and that it is preferred over approaches that track T_a. Results from our modeling experiment strongly indicate that comparing mean daily T_d to model estimates of mean daily T_s across the snow season has significant value for assessing model performance in mountainous areas, where validation data are rarely found at the location of interest. It was beyond the scope of this study to test this model diagnostic tool, and future work

should develop the method further. However, we expect that approximating T_s with T_d is most robust at locations where turbulent mixing occurs frequently (e.g., alpine areas, windy ridges) and in climates where insolation is low. Due to a lack of measurements in forested sites, we cannot comment on the accuracy of this approximation under forest canopies where turbulence and radiation dynamics are altered due to wind sheltering and canopy shading of solar radiation.

While T_d approximates T_s well for daily averages, we note that there are a few caveats. First, we note that errors were more variable for hourly values than for daily averages, and that none of the standard temperatures captured the daily extremes in T_s (Figure 4.4, Table 4.2). This will impact different applications and the user must consider whether accuracy in hourly T_s values is a priority for their particular application. Differences between T_s and T_d were largest at midday, especially at sites that experience increased radiative heating (Figure 4.9). This suggests that the vapor pressure of air at standard height is not in equilibrium with vapor pressure at the snow surface during sunny conditions, signifying a potential increase in the magnitude of latent heat flux. Future work should investigate how to improve upon these findings, such that the maxima and minima in T_s are captured. Second, we recognize that there is no vapor pressure gradient when T_s is assumed equal to T_d (i.e., $e_{air} - e_{surf} = 0$). Because radiation dominates the snowmelt energy balance in many climates, we do not expect this to be a major limitation for many applications. However, in settings where latent heat exchange is important, we caution the use of this simple approximation. Finally, we acknowledge that uncertainties in measured T_s may be large (Table 4.1), and our analysis assumed that T_s was measured accurately with the infrared sensors and pyrgeometers. Using paired sensors at BOG, we found reasonable correspondence between paired infrared sensor measurements and pyrgeometer measurements. Likewise, we did not find a substantial relationship between sensor type and T_s representation

across the sites (e.g., Figure 4.5, 9). Thus, instrument type did not significantly impact our results, but we recommend more research in how these instruments might influence interpretation of T_s. Sites that measure the energy balance rarely have both types of sensors, and it would be helpful to have more paired measurements to better understand uncertainties in measured T_s.

This study highlights the value in having improved temperature and humidity information in mountainous areas and provides motivation for expanded monitoring and improved understanding of humidity variations. This adds to the recent results of *Marks et al.* [2013], who found that the timing of $T_d = 0$ °C matched the measured timing of the precipitation phase transition (rain vs. snow) better than $T_w = 0$ °C and $T_a = 0$ °C during a winter storm in the Owyhee Mountains (Idaho, USA). *Harder and Pomeroy* [2013] found that wet-bulb temperature was a better predictor of precipitation phase than T_a in the Canadian Rockies. Thus, understanding spatial and temporal variations in both air temperature and humidity may yield improved representation of both T_s and event-scale precipitation phase across a watershed. Humidity variations can be monitored in a location of interest using networks of inexpensive thermometers and hygrometers [*Feld et al.*, 2013], which can be obtained at a fraction of the cost of an infrared thermometer or pyrgeometer and are easier to maintain. These inexpensive sensors can also yield reliable information about variations in temperature [*Lundquist and Cayan*, 2007] and snow covered area [e.g., *Raleigh et al.*, 2013c] within a basin to test modeling and remote sensing. More routine humidity observations in standard networks (e.g. NRCS SNOTEL in the US) would also advance knowledge of humidity variations.

Improved spatial and temporal information of T_d would be beneficial for a number of applications with remote sensing and distributed snow models that are related to T_s. For

validation purposes, remotely sensed T_s products from MODIS [*Wan and Li*, 1997] could be compared to spatial estimates of T_d, especially at night when T_d reasonably approximates T_s at many locations (Figure 4.6). Spatial distributions of T_d might also be used to downscale coarse scale T_s from MODIS to finer spatial resolutions. Finally, while we illustrated how T_d might serve as a diagnostic tool for checking daily averages of modeled T_s, we note that T_d might be used as a reasonable prognostic representation of T_s (Figure 4.10a-b) for distributed modeling of the daily snowmelt energy balance across a mountainous catchment.

4.7 Tables

Table 4.1 Site information. Sites are sorted from warmest to coldest mean winter temperatures

Characteristics							
Site Name	**Col de Porte**	**South Royalton**	**Spy Hill**	**Swamp Angel**	**Senator Beck**	**Opabin**	**Scotty Bog**
Acronym	CDP	SRV	SPY	SASP	SBSP	OPB	BOG
Site Climate	Subalpine Maritime	Continental Moist	Continental Prairie	Subalpine Continental	Alpine Continental	Moist Subalpine	Permafrost Wetland
Mean DJF T_a (°C)	-0.90	-5.5	-6.7	-8.5	-10	-10	-22
Elevation (m)	1325	150	1268	3371	3719	2230	283
Latitude (°N)	45.3	43.8	51.2	37.9	37.9	51.3	61.3
Country	France	USA	Canada	USA	USA	Canada	Canada
Study Period	1993-2011	1991-1993	2006-2011	2003-2012	2005-2012	2006-2009	2008-2009
Data resolution	1 hr	30 min	30 min	1 hr	1 hr	30 min	30 min
Sensors							
T_a	PT 100/3, PT 100/4 Wires	Unknown[a]	Vaisala HMP45	Campbell CS500-U	Campbell CS500-U	Vaisala HMP45	Vaisala HMP45
Humidity	Vaisala HMP35 Vaisala HMP45	Unknown[a]	Vaisala HMP45	Campbell CS500-U	Campbell CS500-U	Vaisala HMP45	Vaisala HMP45
U	Chauvin Arnoux Tavid 87	Unknown[a]	RM Young 05103	Campbell Scientific RM Young	Campbell Scientific RM Young	RM Young 05103	Met-One 014A
T_s or LW_{out} *(accuracy[b])*	Testo term Pyroterm (unknown), Heitronics KT15 (±1 °C)	Net radiometer, unknown model[a] (±3 °C)	Kipp and Zonen CNR1 net radiometer (±7 °C at 0 °C, ±6 °C at -35 °C)	AlpuG SnowSurf infrared thermometer (±0.5 °C at 0 °C, ±4 °C at -35 °C)	AlpuG SnowSurf infrared thermometer (±0.5 °C at 0 °C, ±4 °C at -35 °C)	Kipp and Zonen CNR1 net radiometer (±7 °C at 0 °C, ±6 °C at -35 °C)	Kipp and Zonen CNR1 net radiometer (±7 °C at 0 °C, ±6 °C at -35 °C)
Measurement heights (m) above ground							
T_a	1.5[c]	2.0	1.6	3.4	3.8	1.7	1.9
Humidity	1.5[c]	2.0	1.6	3.4	3.8	1.7	1.9
U	10	2.0	1.9	3.8	4.0	4.4	2.0
T_s or LW_{out}	1.2[c]	2.0	1.7	3.2	3.5	1.5	1.5

[a] Site was discontinued in 2005 and sensor records were unavailable [*Peck*, CRREL, 2013, personal communication]. Accuracy was taken from *Peck* [1994].
[b] Based on manufacturer's specifications. For the CNR1, accuracy may be better than this specification [*van den Broeke et al.*, 2004] perhaps as low as ±2 °C.
[c] Sensor heights were adjusted weekly to maintain consistent height above snow surface.

Table 4.2 Mean daily temperature ranges[a] (°C) at the study sites during periods with snow depth exceeding 10 cm.

	CDP	SRV	SPY	SASP	SBSP	OPB	BOG
T_s	**8.4**	**15**	**12**	**15**	**11**	**9.1**	**15**
T_a	6.2	14	9.2	12	7.8	6.4	14
T_w	4.6	11	7.8	8.8	6.2	5.1	12
T_d	4.9	9.9	7.5	9.1	10	6.3	11

[a] At each site, the largest temperature range is in bold.

4.8 Figures

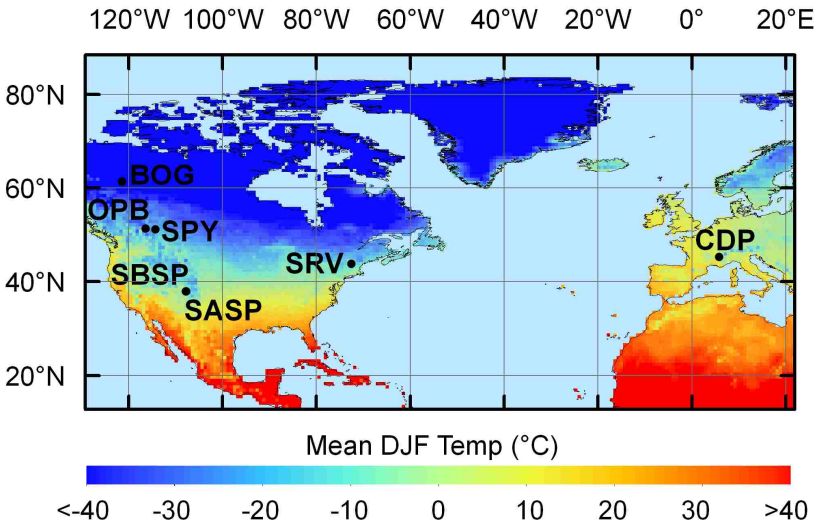

Figure 4.1 Sites used to evaluate snow surface temperature approximations. Acronyms are defined in Table 4.1. Background colors represent DJF air temperature at standard height averaged over 1900-2006 [*Legates and Willmott*, 1990].

Figure 4.2 Comparisons between observed sub-daily snow surface temperature (T_s) and standard height (a-g) air temperature, (h-n) wet-bulb temperature, and (o-u) dew point temperature at the seven study sites. The sites are organized from warmest (left) mean DJF temperatures to coldest (right). Comparisons are only shown during periods when snow depth exceeded 10 cm.

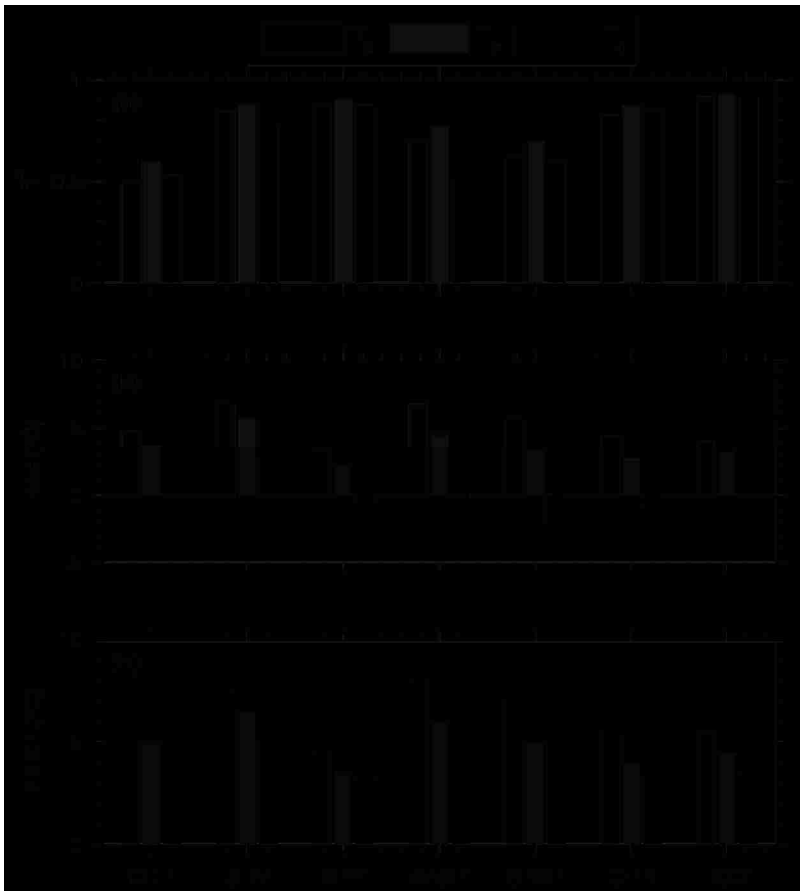

Figure 4.3 Summary statistics for approximating sub-daily snow surface temperature with air temperature, wet-bulb temperature, and dew point temperature at the seven study sites.

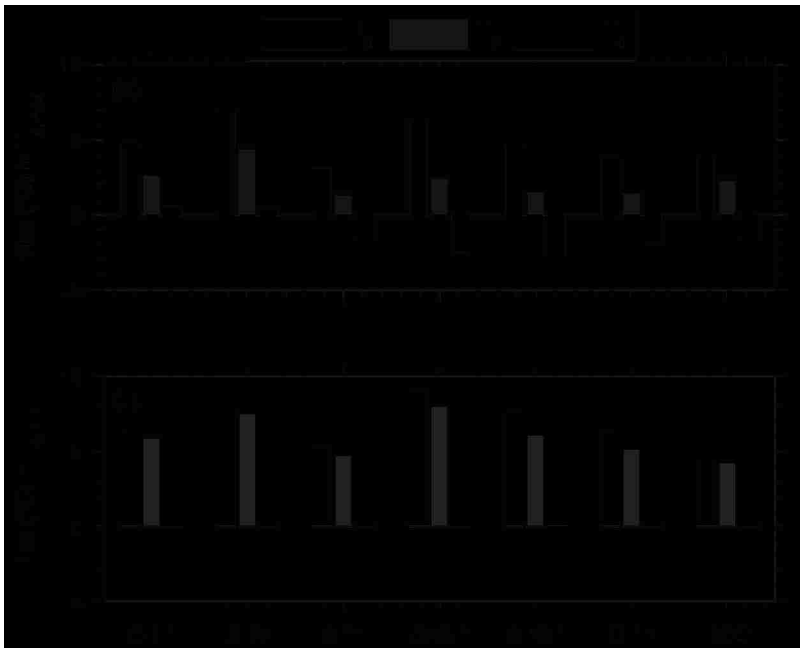

Figure 4.4 Bias in predicting (a) maximum daily surface temperature ($T_{s,max}$) and (b) minimum daily surface temperature ($T_{s,min}$) using air temperature (T_a), wet-bulb temperature (T_w), and dew point temperature (T_d) at the seven study sites. A positive bias indicates overprediction of T_s.

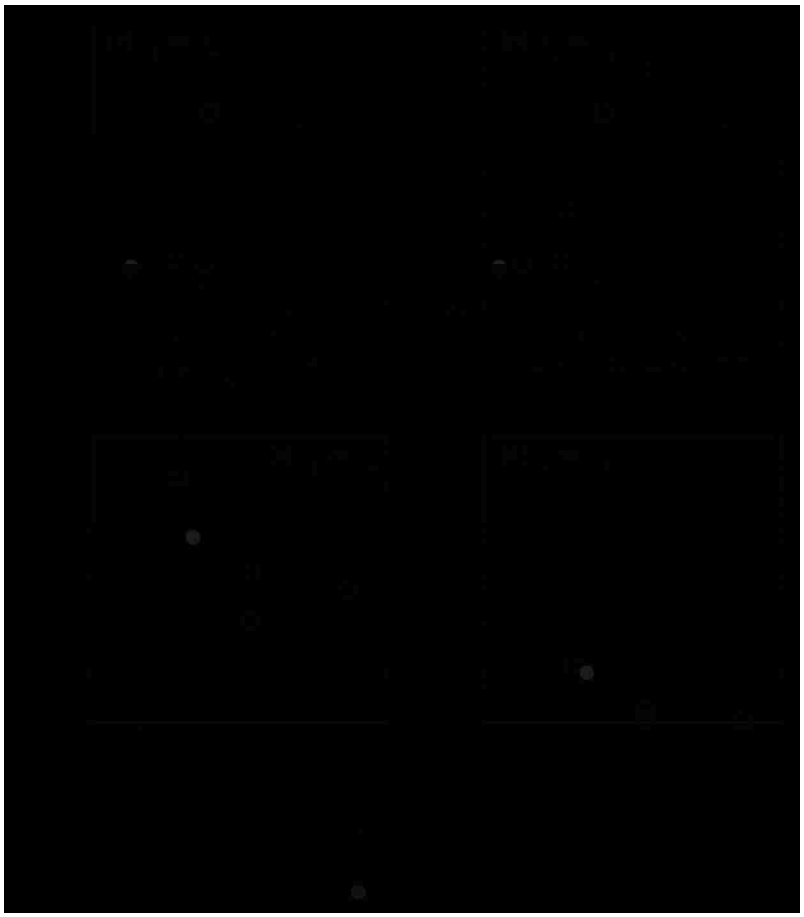

Figure 4.5 RMSE in simulating daily statistics of T_s vs. site characteristics. Shown are RMSE values when simulating maximum daily T_s with maximum daily T_d vs. (a) mean incoming shortwave radiation and (b) mean dew point depression (i.e., $T_a - T_d$). Also shown are RMSE values plotted against latitude when simulating average daily T_s with (c) average daily T_a and (d) average daily T_d. Symbols represent the type of sensor used to measure T_s. Only periods with snow depth exceeding 10 cm are included in these statistics.

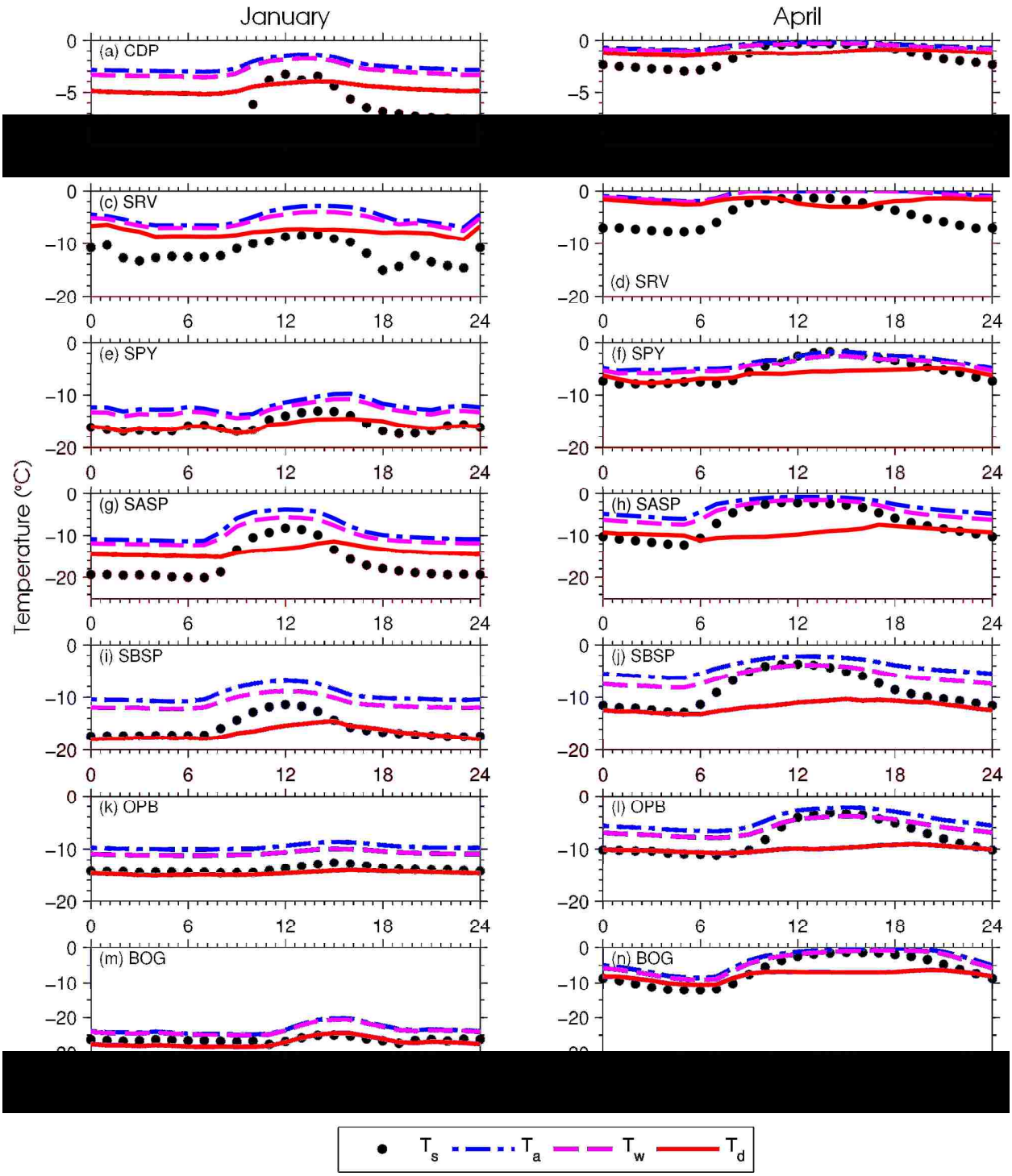

Figure 4.6 Comparisons of snow surface temperature, air temperature, wet-bulb temperature, and dew point temperature averaged hourly at the seven study sites during (left column) January and (right column) April. T_a, T_w, and T_d data are constrained to an upper limit of 0 °C for prediction of T_s.

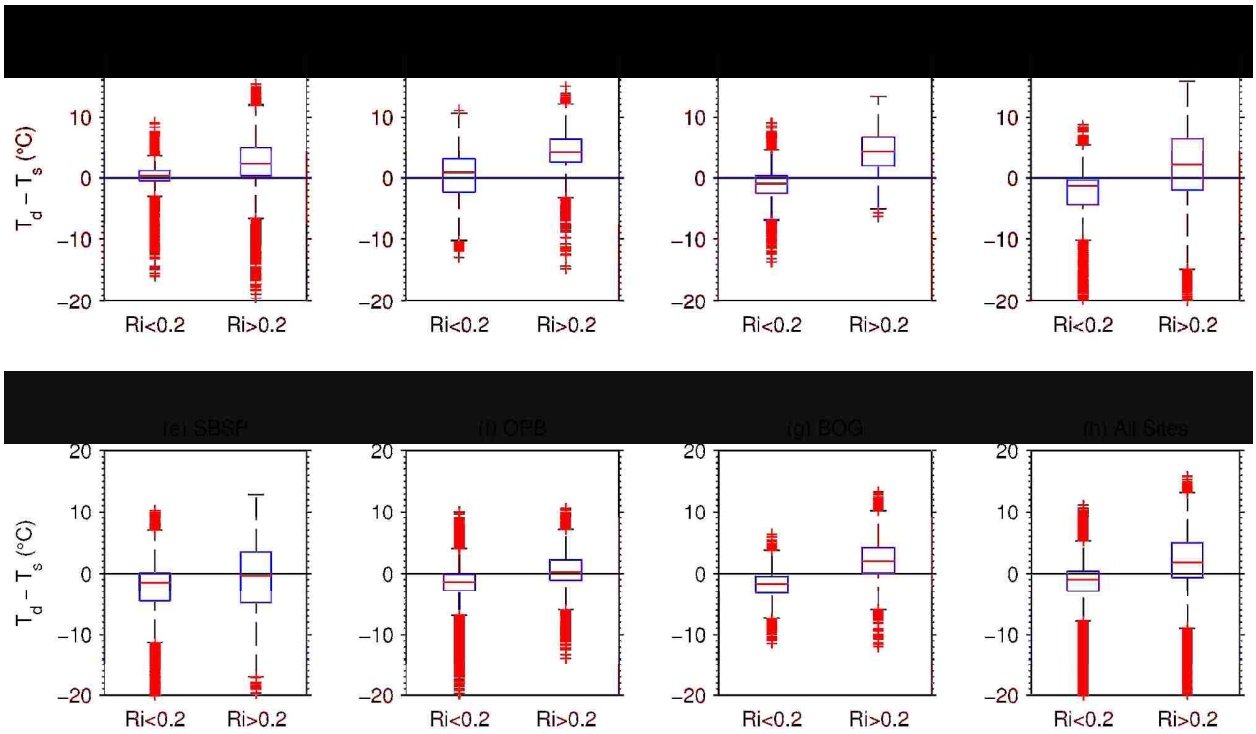

Figure 4.7 Box plots showing the difference between dew point temperature (T_d) and measured snow surface temperature (T_s) as a function of the bulk Richardson (Ri_b) number at the seven study sites. Hours are binned based on unstable (Ri_b <0.2) and stable (Ri_b >0.2) conditions. T_d data are constrained to an upper limit of 0 °C for prediction of T_s.

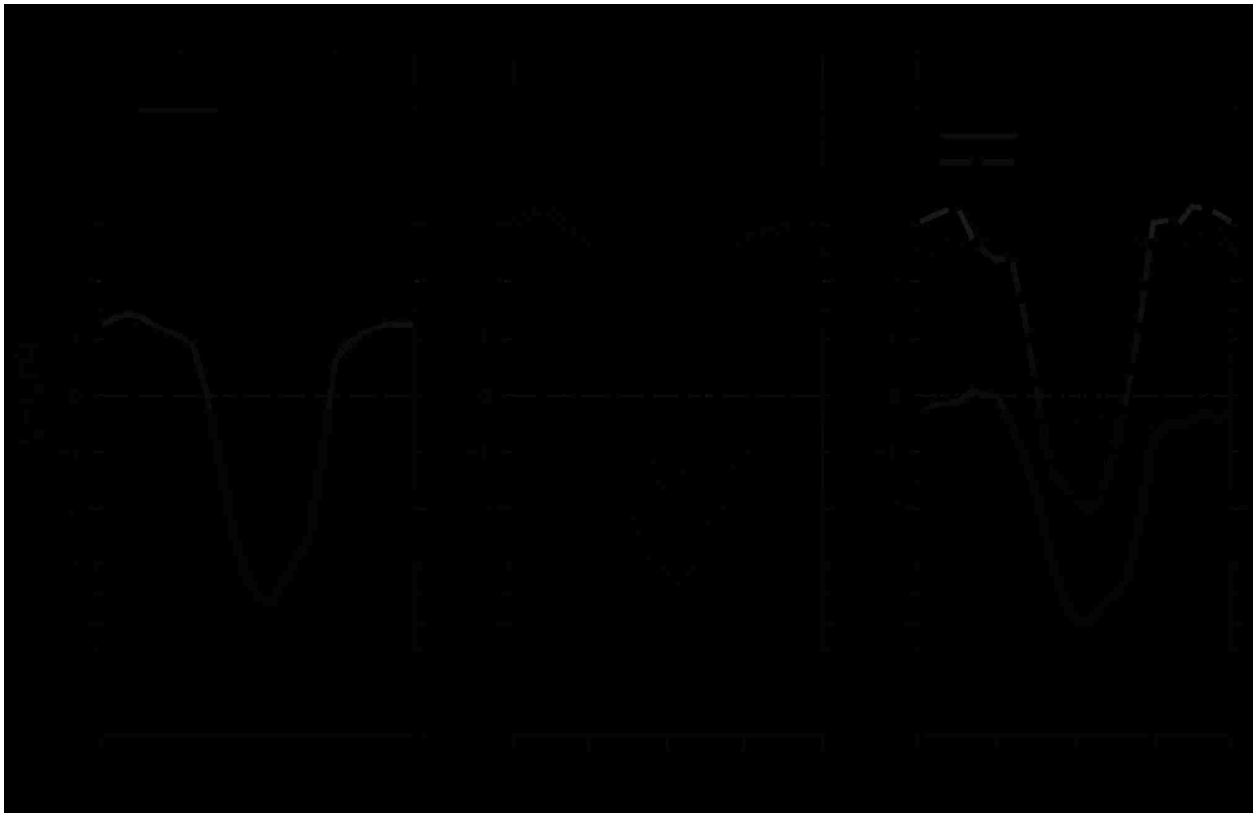

Figure 4.8 Difference between dew point temperature and measured snow surface temperature averaged across all seven sites during each hour of the day. Results are separated based on (a) radiation conditions (as defined by the clearness index, *CI*), (b) stability conditions (based on critical bulk Richardson number, Ri_b), and (c) both radiation and stability conditions. T_d data are constrained to an upper limit of 0 °C for prediction of T_s.

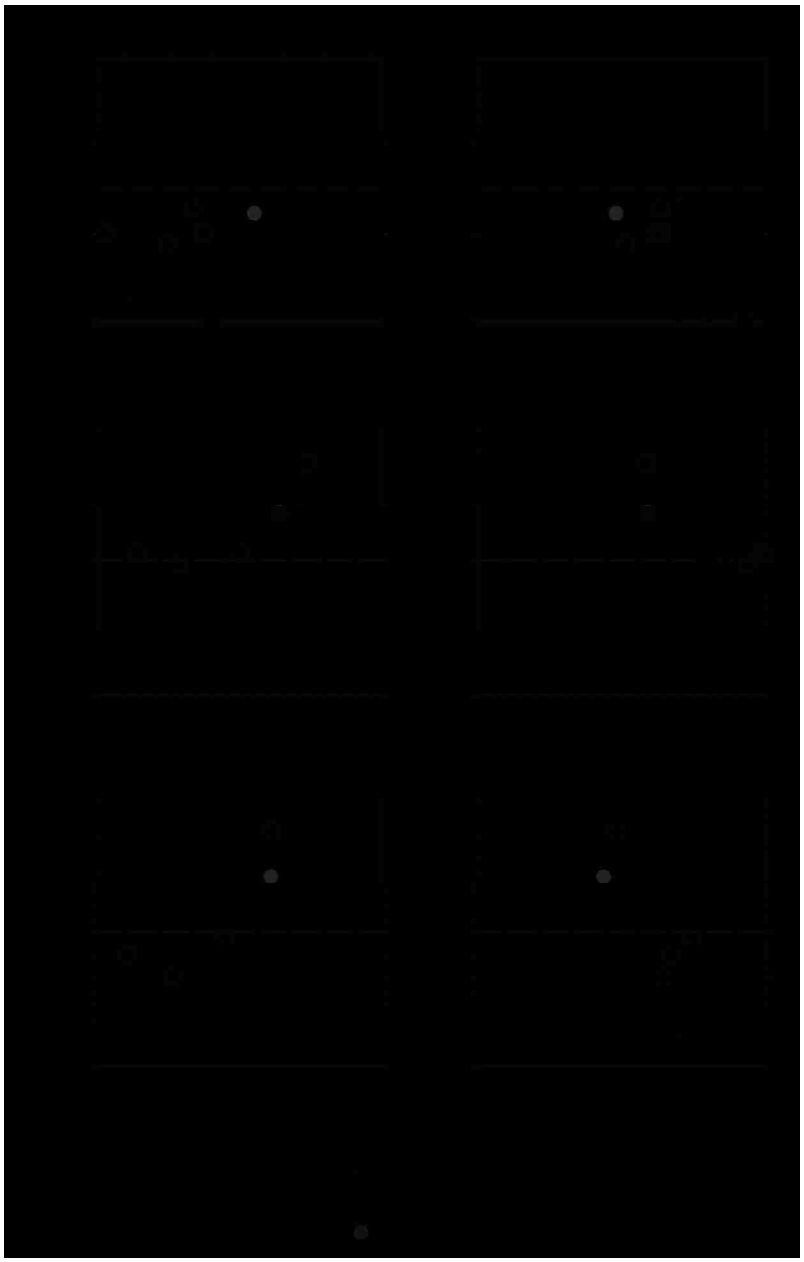

Figure 4.9 Frequency of stability conditions and clear sky conditions vs. mean bias when approximating T_s with T_d at noon, midnight, and daily. T_d data are constrained to an upper limit of 0 °C for prediction of T_s. Noon is taken as the average from 1100 to 1300 hrs, while midnight is taken as the average from 2300 to 0100 hrs.

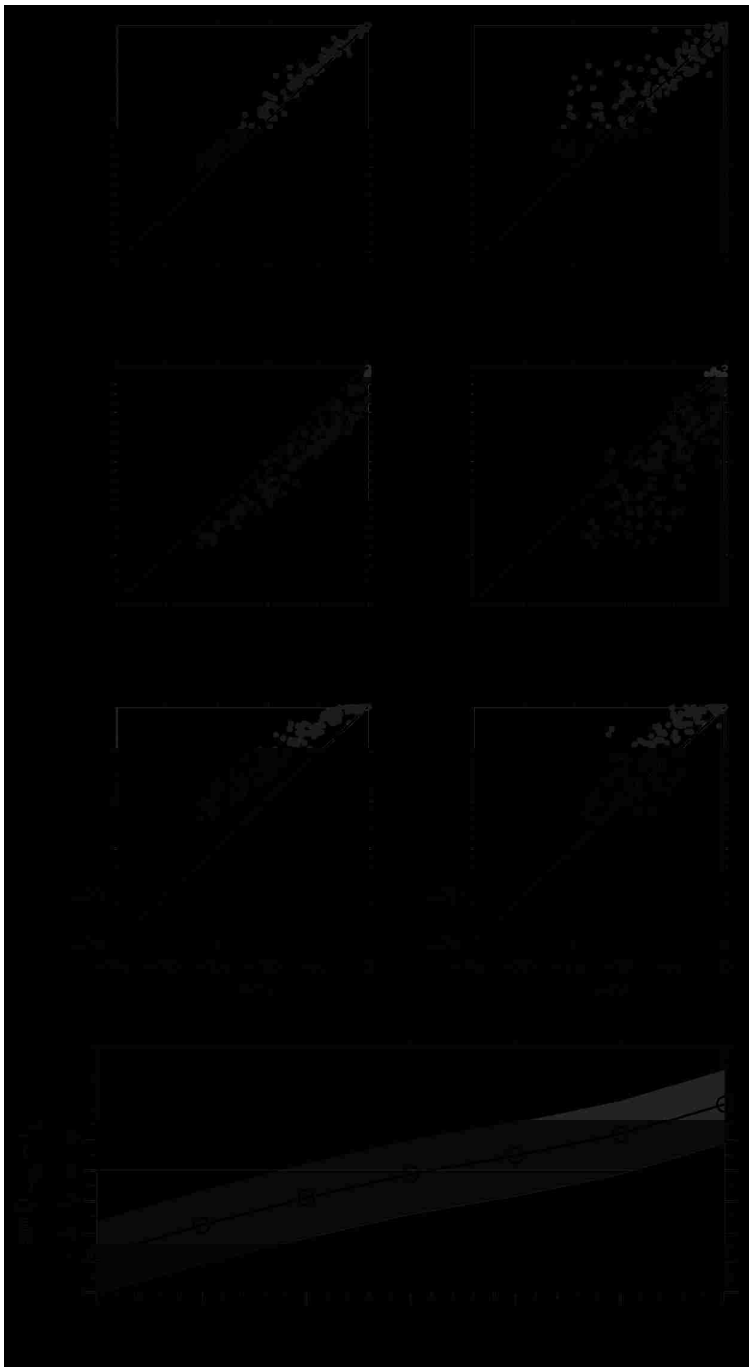

Figure 4.10 Modeling experiment results with SNTHERM at CDP in water year 2006. Shown are (a) the control model simulation (no bias introduced in model data), (b) the approximation of T_s with T_d, (c-d) model simulation with -10% radiation bias with comparisons to observed T_s with T_d, (e-f) model simulation with +10% radiation bias with comparisons to observed T_s with T_d, and (g) the relationship between energy balance bias (%) and the mean difference between modeled T_s with T_d. Only mean daily values are shown. The shading in (g) represents the -2.2 °C to +2.6 °C uncertainty range when approximating observed T_s with T_d, as found in this study.

Chapter 5 Impacts of forcing data uncertainty and availability on a physically-based snow model: the need for improving radiation representation

This chapter is in preparation [*Raleigh et al.*, *in prep* 2013b] for submission to *Water Resources Research*.

Abstract

Physically-based snow models can help provide understanding of seasonal snow processes in data sparse regions but require additional forcing data (e.g., wind, humidity, radiation) that are often unavailable and instead represented with empirical estimates. Forcing uncertainty propagates into model output in ways that cannot be readily understood without ground observations, and thus research is needed to understand which forcings require improved representation (e.g. through more frequent observation or robust approximation methods). Here we conduct an experiment with a sophisticated 1-D physical model (SNTHERM) at four data-rich sites in contrasting climates to compare the relative impacts of observed and empirically estimated forcings on model output across a range of data availability scenarios (n=128). The goal is to quantify which forcings impact model behavior the most, so as to prioritize which measurements (beyond temperature and precipitation) should be made at weather station in snow-dominated catchments. The results showed that model simulations diverged the most depending on the source of radiation data (observed vs. estimated). Therefore, there is a need for improved representation and observation of Q_{si}, Q_{so}, and Q_{li} for realistic simulations of snowpack.

5.1 Introduction

Seasonal snow cover serves as a major water supply [*Barnett et al.*, 2005; *Viviroli et al.*, 2003], acts as a natural reservoir for hydropower [*Madani and Lund*, 2009; *Winther and Hall*, 1999], impacts ecological activity [*Kudo*, 1991; *Trujillo et al.*, 2012], and alters weather [*Hawkins et al.*, 2002; *Jin and Miller*, 2007] and climate [*Qu and Hall*, 2006] through land-atmosphere feedbacks. Despite the importance of seasonal snow, weather stations remain scarce in cold regions [*Hijmans et al.*, 2005; *Lundquist et al.*, 2003; *Viviroli et al.*, 2011]. Extrapolation of available snow measurements is problematic because snow depth exhibits high spatial variability due to scale-variant topographic, vegetative, and wind effects [*Blöschl*, 1999; *Deems et al.*, 2006; *Scipión et al.*, 2013; *Sturm and Wagner*, 2010]. Thus, weather stations provide only a glimpse of a larger, more complex natural system. Given scarce observations and high spatial

variability, numerical models are often employed to predict and better understand snow distributions and processes. While physically-based models are useful for this purpose, the dearth of meteorological and snow observations limits confidence in modeled snow variables.

Comprehensive testing of physically-based snow models is often impractical because few stations measure all of the forcings required [*Bales et al.*, 2006]. These may include air temperature (T_{air}), precipitation (P), wind speed (U), relative humidity (RH), incoming shortwave radiation (Q_{si}), outgoing shortwave radiation (Q_{so}), and incoming longwave radiation (Q_{li}). Outgoing longwave radiation (Q_{lo}) may also serve as a forcing but is often solved as a model state variable in the energy balance [e.g., *Essery and Etchevers*, 2004; *Jordan*, 1991]. A survey of over 1300 stations that measure snow depth or snow water equivalent (SWE) in the western United States shows that T_{air} and P are most frequently measured (Figure 5.1). Radiation components are least frequently measured, even though net radiation dominates the snowmelt energy balance in many climates [*Male and Granger*, 1981; *Ohmura*, 2001; *U.S. Army Corps of Engineers*, 1956; *Zuzel and Cox*, 1975]. Thus, when testing a snow model at a typical station, only T_{air} and P are available to force the model, and the other forcings must be estimated. This practice injects additional uncertainty into the model and enhances the difficulty of testing modeled processes. This survey neglects measurements of sensible and latent heat fluxes, which are even rarer in cold regions but may be important for snowmelt in some locations [e.g., *Marks et al.*, 1998].

Given the current state of forcing data availability, the primary question driving this research is "Which additional forcings, beyond T_{air} and P, should be measured at a snow station to aid in the testing of physically-based snow models?" This goes beyond an investigation of the relative importance of forcings for snow modeling [*Zuzel and Cox*, 1975], as investments in

sensors to measure an important forcing (e.g., Q_{si}, Q_{li}) may not be warranted if that forcing can be estimated reasonably from other data (e.g., T_{air}, RH). Indeed, numerous empirical models have been developed for estimating forcings not measured in remote areas (see section 5.5), and these are regularly used in mountain climate models [MTCLIM; *Hungerford et al.*, 1989] and distributed hydrologic models [VIC, *Liang et al.*, 1994; DHSVM, *Wigmosta et al.*, 1994]. Yet there is incomplete knowledge about how uncertainties in these forcings propagate to snow model output, as 99% of stations do not measure the complete set of forcings (Figure 5.1). Priorities need to be established about which forcings need improved representation.

The purpose of this study is to assess how forcing data uncertainty, availability, and empirical synthesis impact the output of a physically-based snow model in different snow climates. The basic premise is that T_{air} and P are measured at a hypothetical station at hourly or daily resolution (e.g., historic SNOTEL observations are archived as daily), and we wish to determine which empirically-estimated forcings, compared to directly-measured forcings, yield the most significant changes in snow model output. By using data observed at four sites that measure all forcings required to run and evaluate a physically-based snow model, we are able to gain unique insights about model response to forcing uncertainty.

To address the research question, we assemble 128 different forcing data scenarios at each site to correspond to all possible weather stations configurations that could be used to test a model that requires seven forcings (T_{air}, P, U, RH, Q_{si}, Q_{so}, Q_{li}). When a hypothetical station lacks a specific forcing, we use a single method to estimate the missing forcing. Selection of this method is based on performance in prior studies. At each hypothetical station, we simulate snowpack to reveal the range of model responses to data availability and estimation. Finally, we

quantify the aggregate effect of forcing representation (i.e. observed vs. estimated) on the model output uncertainty to understand which forcings need improved representation or observation.

We simulate seasonal snowpack with the snow thermal model (SNTHERM) [*Jordan*, 1991]. SNTHERM was selected for several reasons: (1) it includes the most important process representations (e.g., snow density, storage and refreezing of liquid water) needed for consistent results [*Essery et al.*, 2013], (2) it is sophisticated and has been described as a "benchmark model" [*Watson et al.*, 2006], (3) it has been widely tested and has yielded reasonable simulations of snowpack properties [*Feng et al.*, 2008; *Jin et al.*, 1999; *Rutter et al.*, 2009], and (4) it attempts to be a generalized model that can achieve reasonable accuracy with "no calibration" [*Melloh*, 1999]. We assess how forcing uncertainty impacts modeled SWE, snow surface temperature (T_{surf}), snowmelt rates, and snow disappearance dates (SDD).

Because observed forcings are subject to uncertainties due to instrument accuracy and measurement errors [e.g., *Helgason and Pomeroy*, 2012; *Huwald et al.*, 2009; *Johnson and Marks*, 2004; *Melloh et al.*, 2004; *Rasmussen et al.*, 2012], we remain cautious in our interpretation of differences between measured and estimated forcings. While we make comparisons between these forcings, we carefully consider what may be concluded from model inference. Under this framework, we do not discount the seemingly bizarre possibility that some empirical forcings may yield more "realistic" snow simulations than their measured counterparts. When we find large differences in model output for measured vs. empirical forcing cases, we can only conclude that those forcings need improved understanding and further research, with attention to both possible measurement errors and to the robustness of empirical approaches.

5.2 Literature review

No prior study (to our knowledge) has used data-rich weather stations to examine the adequacy of empirical methods for estimating all model forcings in the context of modeling cold region processes. Responses of snow models to forcing data uncertainty have been examined in several studies [e.g., *Raleigh and Lundquist*, 2012; *Shamir and Georgakakos*, 2006; *Slater et al.*, 2013]. *Walter et al.*, [2005] suggested that a physical snow model could adequately simulate SWE using daily T_{air} and P, but did not provide guidance about the value of forcing data beyond these two forcings. In other recent studies, there has been considerable interest in comparing and evaluating different empirical methods [e.g., *Bohn et al.*, 2013; *Feld et al.*, 2013; *Flerchinger et al.*, 2009; *Shook and Pomeroy*, 2011] to understand comparative accuracy in different climates, but with little attention to how these various methods impact snow model output. No study has attempted to understand how uncertainty in all estimated forcings propagates through a snow model.

The most similar studies were conducted by *Waichler and Wigmosta* [2003] and *Schnorbus and Alila* [2004], where sub-daily forcing data scenarios were generated and used to force a distributed model in maritime basins. These studies provided useful insights into the response of a model to a suite of forcing data scenarios, but were confined to a single climate and only provided partial understanding of forcing uncertainty due to the lack of key measurements (e.g., SWE, Q_{so}, Q_{li}, and Q_{lo}). Here we aim to fill this research gap with a more comprehensive test of forcing uncertainty on snow model behavior across a variety of climates.

Model forcing is one of several sources of uncertainty when modeling a snowpack. Uncertainties in (1) model structure and parameterization and (2) evaluation data (e.g., SWE, snow surface temperature) also confound efforts to test the reliability of numerical models [*Clark and Vrugt*, 2006; *Essery et al.*, 2013; *Johnson and Marks*, 2004; *Kavetski et al.*, 2006; *Kuczera et*

al., 2006]. Considerable attention has been paid to model selection and intercomparison [e.g., *Boone and Etchevers*, 2001; *Essery et al.*, 2012; *Etchevers et al.*, 2004; *Feng et al.*, 2008; *Jin et al.*, 1999; *Koivusalo and Heikinheimo*, 1999; *Rutter et al.*, 2009; *Slater et al.*, 2001]. Whereas these comparison studies focus primarily on model uncertainty, we focus exclusively on forcing uncertainty and utilize a common model parameter set at all locations. For the evaluation data, we qualitatively consider uncertainty by showing the available snow survey and snow pit data at each site.

Alternative forcing estimation techniques have also been developed and implemented, such as geostatistical interpolations (e.g., kriging [*Jabot et al.*, 2012]), output from mesoscale weather models [*Wayand et al.*, 2013], and satellite-based estimates of forcings [*Forman and Margulis*, 2009; *Ma and Pinker*, 2012; *Pinker and Laszlo*, 1992]. While recognizing these other methods, we focus primarily on empirical approaches for forcing estimation due to their historical and sustained popularity in distributed hydrologic models.

5.3 Study sites and observed data

We selected study sites (Figure 5.2, Table 5.1) across a range of environments to assess how results may vary with climate. The sites ranged in elevation from 930 m to 3370 m, and included Imnavait Creek (IC) [*Kane et al.*, 1991; *Sturm and Wagner*, 2010] in Alaska (USA), Col de Porte (CDP) [*Morin et al.*, 2012] in the Rhône-Alpes (France), the Reynolds Mountain East (RME) sheltered site [*Reba et al.*, 2011a] in Idaho (USA), and the Swamp Angel Study Plot (SASP) [*Landry et al.*, 2013] from the Senator Beck Basin Study Area in Colorado (USA). All sites have sub-daily (i.e., hourly, 30-min, or 15-min) observations of T_{air}, P, U, RH, Q_{si}, and Q_{so}. All sites except IC recorded measurements of Q_{li} to force the model. At IC, Q_{li} was taken as the

residual from local measurements of net radiation (Q_{net}) and the other measured radiation components:

$$Q_{li} = Q_{net} - (Q_{si} - Q_{so} - Q_{lo}) \tag{5.1}$$

Sensor manufacturers and models are listed in Table 5.1.

We only retained Q_{so} observations in the 1100 to 1300 hr window, as pyranometers are less accurate in the early morning and late afternoon when the sun is lower in the sky and shadows are cast in the sensor field-of-view. Further, measuring Q_{so} under clear sky conditions is known to lead to non-negligible measurement errors due to the directional dependence of scattered radiation off rough or sloped surfaces [e.g., *Warren*, 1982]. The mid-day values were averaged each day, checked to ensure that calculated albedo fell within realistic limits, and then interpolated between days to get hourly values.

Model evaluation data included SWE and T_{surf}. SWE was measured locally with a snow pillow at all sites except SASP. At SASP, snow depth (SD) was measured hourly with an ultrasonic sensor. We converted measured SD to SWE using bulk snow density (ρ_b):

$$SWE = SD\frac{\rho_b}{\rho_w} \tag{5.2}$$

where ρ_w is the density of water (1 g cm^{-1}). To implement equation 5.2, we first calculated daily ρ_b from SWE and SD observations at three nearby NRCS SNOTEL sites (Red Mountain Pass, Idarado, and Mineral Creek), and then averaged these three values daily. Because spatial variability in snow density is much lower than spatial variability in snow depth, we assumed that mean density at these nearby sites was representative of density at SASP [e.g., *Sturm et al.*, 2010]. To further constrain density, we fit the SNOTEL-based daily ρ_b time series to match ρ_b observations from routine (weekly to bi-weekly) snow pit excavations at SASP. Thus, the SNOTEL density provided information on the temporal evolution of density, while the snow pit

density helped to control for any spatial differences in density. Snow pit SWE showed reasonable agreement with this methodology (see section 5.6). All sites had snow pit or snow survey data to help understand uncertainty in SWE evaluation data. To ensure quality SWE data, we aggregated SWE to daily values, checked for consistency between SWE and SD, and tested for errors in snow pillow SWE [*Johnson and Marks*, 2004]. We found three notable errors at the IC snow pillow that required correction based on SD, and none at the other sites.

T_{surf} was measured hourly with an infrared temperature sensor at SASP and CDP, but not at RME and IC. Thus, Q_{lo} measurements at RME and IC were used to calculate T_{surf} (°C) using the Stefan-Boltzmann equation, assuming snow surface emissivity of 1:

$$T_{surf} = \left(\frac{Q_{lo}}{\sigma} \right)^{1/4} - 273.15 \tag{5.3}$$

where σ is the Stefan-Boltzmann constant (5.67 x 10^{-8} W m^{-2} K^{-4}). We assumed snow emissivity was 1 for consistency with the published CDP T_{surf} dataset, which is based on observations from both an infrared temperature sensor and a pyrgeometer [*Morin et al.*, 2012]. Snow is nearly a black-body radiator [*Kondo and Yamazawa*, 1986], so this was a reasonable simplification.

At IC and CDP, all forcing and evaluation data were observed at the same location. At RME all forcings except Q_{so} and Q_{lo} were taken from the snow pillow site. These outgoing flux measurements were available at the snow pillow site, but we did not use them because they were measured on a 35 m tower, which presented significant field-of-view issues (e.g., trees in the sensor view). Thus, we used Q_{so} and Q_{lo} data from a nearby station (RME exposed site), where field-of-view issues were minimized because the measurements were made 1.5 m above ground [*Danny Marks, personal communication, 2013*]. The tradeoffs were the radiometer (1) casted a shadow on the snow when the solar zenith angle was large (e.g., at noon in spring and summer), and (2) was not located at the RME snow pillow. At SASP, all forcings were taken locally,

except radiation was used from a nearby exposed site (i.e., 1.3 km to the west, 340 m above SASP) where the radiometers were snow-covered less often.

At each study site, a single year was selected based on initial data completeness and quality. Table 5.2 shows the percentage of hours missing in the observations, and indicates that gaps were generally minimal. These small gaps required filling in order to obtain serially-complete (i.e., no missing values) datasets to drive the snow model. A hierarchy of spatial and temporal approaches was used to fill these gaps. The primary approach was to use regression analysis between nearby sites. IC, RME, and SASP all had paired sites in close proximity that had high correlations. When nearby sites were also missing data, temporal approaches were used. For short gaps (i.e., less than 24 hours), we used temporal interpolation [e.g., *Liston and Elder*, 2006], which *Henn et al.* [2013] demonstrated is the method of choice for filling short gaps in T_{air} at single stations. When gaps exceeded 24 hours (a case occurring in less than 0.5% of any record), the gap was filled using the monthly average observation at each hour. These techniques were employed to ensure data consistency at each site (i.e., missing data were filled based on observations). We assume these methods introduced negligible error into the analysis.

We found systematic differences between SWE and cumulative P at all sites, and this required rectification. When cumulative P was less than SWE, we assumed P undercatch bias was present [*Rasmussen et al.*, 2012; *Serreze et al.*, 1999; *Sieck et al.*, 2007]. This was found at all sites except CDP and was most pronounced at IC, where problems in measuring P are well documented [e.g., *Clagett*, 1988; *Kane et al.*, 1991; *Yang et al.*, 2000]. Because the premise of the study is that we are evaluating a snow model at a station with observations of T_{air}, P, and *SWE*, we argue that P undercatch could be detected and corrected with the available data. Thus, we corrected P based on observed SWE, an approach that has been used in other modeling

studies [e.g., *Sturm and Wagner*, 2010]. However, we note that undercatch "bias" in *P* is difficult to quantify with snow pillow data, as snow pillow SWE may be subject to effects from drifting snow [*Meyer et al.*, 2012]. Mean correction factors varied from 0.82 (i.e., we decreased *P*) at CDP to 2.73 (i.e., we increased *P*) at IC.

5.4 Snow model

SNTHERM [*Jordan*, 1991] is a physically-based, 1-D, multi-layer snow model that was developed for prediction of thermal characteristics of seasonal snow and frozen soil. The model accounts for snow accumulation, compaction, densification, metamorphosis, sublimation, liquid water retention, and melt processes for a snowpack on bare soil (i.e., no vegetation interactions). The model is based on mixture theory (i.e., snow and soil are porous matrices), allowing snow density and moisture content of layers to be simulated. The equations governing heat and moisture exchange are numerically solved with a finite control volume approach across horizontal layers of snow and soil. As snow accumulates in these layers (minimum and maximum thickness specified by the modeler) and the snowpack compacts with time, the 1-D grid compresses while maintaining the original finite element structure. Energy exchanges at the snow-atmosphere interface are simulated based on meteorological forcings and at the snow-soil interface based on simulated soil temperature profiles. Model accuracy is ensured by using a flexible time-step, such that shorter time-steps (e.g., 5 s) are invoked when convergence criteria are not met for a given layer. The end result is that SNTHERM provides a detailed treatment of processes within a snowpack, but at the expense of high computational intensity.

We ran SNTHERM at an hourly time step and provided the model with seven forcings: T_{air}, *P, U, RH*, Q_{si}, Q_{so}, and Q_{li}. Each type of forcing was either the observed data (section 5.3) or empirically estimated data (section 5.5), depending on the particular hypothetical weather

station. Although SNTHERM simulates snow grain growth and optical properties of the snow surface, we overrode this feature by providing Q_{so} data from observed data or from empirical estimates [*U.S. Army Corps of Engineers*, 1956] prepared offline. This was done to make the results more generally applicable to other studies, as many snow models do not provide detailed treatment of snow grain size and albedo.

5.5 Methods

5.5.1 Experiment design

We wanted to understand the range of model responses as a function of both weather station configuration and empirical forcing accuracy, with the explicit goal of identifying which forcings need improved representation (i.e., "which additional sensor should be installed at a station?"). For simplicity, we did not consider measurement accuracy for the observed forcings so that we could focus on differences between observed and estimated forcings.

We first assembled a series of 128 different data availability scenarios at each study site. This can be envisioned as the 128 different weather station configurations that are possible to test a snow model that requires seven forcings. The construction of these scenarios is conceptually shown in Figure 5.3a. We started with all seven forcings (i.e., T_{air}, P, U, RH, Q_{si}, Q_{so}, and Q_{li}) observed at hourly resolution (i.e., the maximal data scenario), and then incrementally replaced the observed forcing data with empirically estimated forcings (see below), until only daily T_{air} and P were available to estimate the missing forcings (i.e., the minimal data scenario). The minimal data scenario corresponded to the original premise of the study, where we assumed a typical snow station and wanted to find which forcings beyond daily T_{air} and P most need improved observation or representation. After constructing the 128 data scenarios, we simulated snowpack with SNTHERM at each site to generate 128 time series of SWE and T_{surf} (Figure 5.3a).

To understand the impact of observed vs. estimated forcings, we divided the 128 data scenarios into two groups of 64 "stations" for each forcing, one group where that forcing was measured and the other group where that forcing was not observed and was instead estimated empirically. Our null hypothesis was that mean model output for these two groups was similar because the observed and estimated forcings were not drastically different. We then computed and compared the aggregate statistics for SWE, T_{surf}, snowmelt, and SDD of these two groups to test and find cases when the model output was significantly ($p<0.05$) different based on a non-parametric Wilcoxon ranked sum test.

For each missing forcing, a single empirical estimation method was selected based (whenever possible) on prior comparison studies between alternative empirical methods and the popularity of usage (see section 5.5.2). The goal of this study was not to find the "best" empirical method, but to apply a single method that has been shown to produce reasonable results in other studies. We note that most of the empirical methods rely on other forcings, such that errors in the independent forcings may propagate into the dependent forcings (Figure 5.3b), an issue known to impact estimates of some forcings [e.g., Q_{li}, *Feld et al.*, 2013; *Wayand et al.*, 2013][*Feld et al.*, 2013; *Wayand et al.*, 2013]. Details about the selection and performance of the empirical methods are in section 5.5.2.

5.5.2 Selection and description of data estimation methods

For each forcing, a single estimation method was selected based (whenever possible) on prior comparison studies between alternative empirical methods and the popularity of usage (see below). We only selected one empirical method instead of multiple methods to reduce computational expense. Consideration of an additional empirical method for each forcing would yield $3^7 = 2187$ simulations and take 17 times more computation time than the original binary

experiment. Besides, the purpose of this study was not to find the "best" empirical method, but to apply a single method that has been shown to produce reasonable results in other studies. To shed light on the performance of these methods at the study sites, we statistically evaluated the empirical forcings against the observed forcings. We note that most of the empirical methods rely on other forcings, such that errors in the independent forcings may propagate into the dependent forcings (Figure 5.3b), an issue known to impact estimates of some forcings [e.g., Q_{li}, *Feld et al.*, 2013; *Wayand et al.*, 2013]. Details about the empirical methods are reported below, while evaluation and comparisons to other studies are found in Appendix B.

5.5.2.1 Air temperature (T_{air})

To supplement data scenarios (Figure 5.3) where hourly T_{air} was not observed, we used the semi-empirical model of *Cesaraccio et al.* [2001] to estimate T_{air} (°C) at hour t:

$$T_{air}(t) = \begin{cases} T_{\min} + DTR \times \sin\left[\left(\frac{t - t_{\min}}{t_{\max} - t_{\min}}\right)\frac{\pi}{2}\right] & t_{\min} < t < t_{\max} \\ T_{set} + (T_{\max} - T_{set}) \times \sin\left[\frac{\pi}{2} + \left(\frac{t - t_{\max}}{4}\right)\frac{\pi}{2}\right] & t_{\max} < t < t_{set} \\ T_{set} + \left(\frac{T_p - T_{set}}{\sqrt{t_p - t_{set}}}\right)\sqrt{t - t_{set}} & t_{set} < t < t_p \end{cases} \quad (5.4)$$

where T_{min} and T_{max} are the daily minimum and maximum temperatures (assumed available in our hypothetical experiment), DTR is the daily temperature range (i.e., T_{max} - T_{min}), t_{min} is the hour of T_{min} (assumed at sunrise), t_{max} is the hour of T_{max} (a model parameter), t_{set} is the sunset hour, T_p is the minimum air temperature of the next day, which occurs at hour t_p. This temperature model was selected because it yielded lower errors than other methods in the method intercomparison presented by *Cesaraccio et al.* [2001] and is a parsimonious model. Sunset and sunrise times were calculated based on solar geometry [*U.S. Naval Observatory*, 1990].

5.5.2.2 Precipitation (P)

In data scenarios where hourly *P* was not observed, the daily observed sum was divided evenly across all hours of the day. To reiterate, we made the assumption that daily *P* was available because the study was framed to understand which forcings beyond T_{air} and *P* could be better represented at weather stations for snow modeling studies. The daily *P* data were the observed data after rectification with the local SWE data (section 5.3). The uniform sub-daily disaggregation method has precedence in other works, including the widely-used *Maurer et al.* [2002] dataset and in the experiment of *Waichler and Wigmosta* [2003]. We expected that uniform hourly precipitation was only a bad assumption at warmer sites (e.g., CDP) that experienced transient rain-snow conditions.

5.5.2.3 Wind speed (U)

Wind speed (*U*) estimation remains crude in hydrologic research, and to our knowledge, few methods have been developed that represent scalar wind speed accurately at the hourly scale. Part of the difficulty in estimating wind speed empirically is that it exhibits low correlation with other surface forcings [*Parlange and Katz*, 2000]. Common approaches include extrapolating the nearest available observation in space or to assume a constant value [*Waichler and Wigmosta*, 2003; *Walter et al.*, 2005]. Many studies in the wind energy and atmospheric science communities have found that *U* follows a Weibull distribution [*Justus and Mikhail*, 1976; *Justus et al.*, 1978], and thus stochastic estimation of *U* is possible with random values from this probably distribution. However, doing so required knowledge of the distribution's shape and scale parameters, which are location-dependent. Based on our study's premise, these parameters were unknowable and therefore we did not use the Weibull distribution.

In lieu of established methods for estimating hourly wind speed, we used NCEP/NCAR reanalysis data [*Kalnay et al.*, 1996] in order to obtain a coarse resolution estimate of regional wind speed. NCEP reanalysis was the basis for the *Maurer et al.* [2002] dataset that has been used in many studies with the VIC model. We estimated hourly wind speed at the study sites based on the long-term mean 6-hourly dataset of scalar wind speed to capture typical diurnal and seasonal cycles in regional wind speed. Methods of downscaling reanalysis based on terrain features might reduce any bias [*Winstral et al.*, 2009]. However we elected not to incorporate these adjustments or other advanced techniques for modeling wind (e.g., GIS analysis, blowing snow models) because we did not have access to spatial snow data that would sufficiently help evaluate these methods. Nevertheless, we acknowledge the significance of these processes in our study regions [*Sturm and Wagner*, 2010; *Winstral et al.*, 2002].

5.5.2.4 Relative humidity (RH)

To estimate hourly *RH*, we implemented the approach of *Running et al.* [1987] because it is a commonly used empirical approach in areas without humidity observations for extrapolation. This method assumes that the daily dew point temperature (T_{dew}) equals T_{min}. Hourly values of *RH* are thus found by calculating the vapor pressure each hour as a function of T_{air} and comparing to the daily saturation vapor pressure. This approach effectively assumes *RH*=100% everyday when $T_{min} = T_{dew}$, and *RH* declines with as T_{air} increases through the day. A new correction to this empirical method has been developed [*Kimball et al.*, 1997] to improve RH estimation in arid regions. We applied the *Running et al.* [1987] approach instead of the *Kimball et al.* [1997] because (1) the methods do not yield substantially different errors in humidity during the snow season [*Feld et al.*, 2013], and (2) the *Running et al.* [1987] is easier to

implement, as it does not require simultaneous estimation with Q_{si} [*Bohn et al.*, 2013; *Thornton and Running*, 1999; *Thornton et al.*, 2000].

5.5.2.5 Incoming shortwave radiation (Q_{si})

We estimated incoming shortwave radiation using an empirical relationship between daily air temperature range (*DTR*) and atmospheric transmissivity. *DTR* is correlated with atmospheric transmissivity, such that cloudy (clear) days tend to have a smaller (larger) *DTR*. The approach we used was originally developed by *Bristow and Campbell* [1984] and later refined by *Thornton and Running* [1999] to eliminate the need for site specific calibration. Several studies have compared different approaches for estimating incoming shortwave radiation with empirical approaches or reanalysis datasets [*Abraha and Savage*, 2008; *Ball et al.*, 2004; *Shook and Pomeroy*, 2011; *Slater et al.*, 2013; *Yang and Koike*, 2005]. *Ball et al.* [2004] compared the *Thornton and Running* [1999] model to other empirical approaches and found that the different methods had similar accuracy, despite large differences in model complexity. We selected the *Thornton and Running* [1999] model because it is commonly used in mountain modeling [e.g., MTCLIM, *Hungerford et al.*, 1989] and land surface models such as VIC [e.g., *Bohn et al.*, 2013; *Liang et al.*, 1994]. As noted above, we did not use the simultaneous approach for estimating Q_{si} and vapor pressure and instead implemented the method in "stand-alone" mode [*Bohn et al.*, 2013].

5.5.2.6 Outgoing shortwave radiation (Q_{so})

We calculated hourly outgoing shortwave radiation (Q_{so}) from estimated albedo based on the *USACE* [1956] method because of its common usage in snow models [e.g., *Letsinger and Olyphant*, 2007; *Livneh et al.*, 2010; *Walter et al.*, 2005; *Wigmosta et al.*, 1994]. Albedo is

simulated as a function of time since last snowfall based on decay curves for the accumulation and ablation seasons. The time-decay in albedo is a representation of surface metamorphism processes (e.g., recrystallization) and litter. Modeled snowmelt can exhibit high sensitivity to the selected decay curve. While alternative methods exist for switching the curves [based on liquid water at the snow surface, *Andreadis et al.*, 2009; based on T_{surf}, *Livneh et al.*, 2010] the ideal method for switching between the curves is unclear. As a simplistic approach, we assumed that the accumulation season ran from 1 October to 1 March and that the ablation season ran after 1 March. Q_{so} was estimated as the product between albedo and Q_{si}.

5.5.2.7 Incoming longwave radiation (Q_{li})

Many methods exist for estimating incoming longwave radiation (Q_{li}), and some of these have been reviewed and compared in recent studies [*Flerchinger et al.*, 2009; *Herrero and Polo*, 2012; *Juszak and Pellicciotti*, 2013]. Based on the recommendations of *Flerchinger et al.* [2009] and *Juszak and Pellicciotti* [2013], we used the *Dilley and O'Brien* [1998] parameterization for clear-sky conditions and the *Unsworth and Monteith* [1975] parameterization for all-sky conditions. These studies reported that the two parameterizations had a mean bias of 1.1 W m^{-2} and RMSE ranging from 21 to 30 W m^{-2} when evaluated across a range of climates. We assumed there was no way to calibrate these methods in the context of our study premise (i.e., only T_{air} and P observation available at a remote site) and therefore applied the original calibration coefficients at our sites. However, we note that the errors in Q_{li} parameterizations increase when applied at uncalibrated sites [*Juszak and Pellicciotti*, 2013]. The sensitivity to calibration and the scarcity of longwave observations in mountainous areas (Figure 5.1) may explain why *Bohn et al.* [2013] found that empirical longwave bias was highest in mountains.

5.6 Results

We assessed how model output changed with respect to weather station configuration and empirical forcing synthesis (Figure 5.3a). The SNTHERM simulations of SWE and T_{surf} are shown in Figure 5.4. Simulations tended to cluster together at CDP and RME (Figure 5.4 c,e) during mid-winter periods when T_{surf} was brought to 0°C and surface melt occurred for some simulations but not others (Figure 5.4 d, f). The major differences in SWE often resulted from these events. *Wayand et al.* [2013] reported similar cases of mid-winter melting and subsequent offsetting of SWE at low elevations in a maritime basin, where warmer conditions and greater sensitivity to forcing errors prevail.

Daily average SWE from each of the two groups of weather stations (observed forcing vs. estimated forcing) are shown in Figure 5.5. All major differences in model output were related to radiation. For Q_{si}, this was most evident at IC, CDP, RME (Figure 5.5 c, h, m). These sites showed opposite responses to whether Q_{si} was observed or estimated, as SNTHERM SWE at IC was closer to observed SWE with the estimated Q_{si}, while SNTHERM SWE at CDP and RME was closer to observed SWE with observed Q_{si}. Based on model inference, this peculiar result at IC suggested that the pyranometer had a positive measurement bias, but we cannot conclusively determine this given the large spatial variability in the snow evaluation data. SWE differences due to Q_{so} emerged at all four sites. Q_{so} observations seemed to help the SNTHERM simulations at IC (Figure 5.5d) but hindered simulations at the other three sites (Figure 5.5 i, n, s). This analysis showed that the mid-winter melt events that caused "clustering" or results (Figure 5.5 c, e) were related to treatment of Q_{si} and Q_{so}. For example, a melt event in January at RME caused the observed and estimated groups to diverge in both the Q_{si} and Q_{so} experiments (Figure 5.5 m, n). Differences due to Q_{li} (observed vs. estimated) were also apparent, but less dramatic than the responses due to Q_{si} and Q_{so} (Figure 5.5 e, j, o, t). In some cases, the observed

forcing yielded SWE closer to the observations (e.g., IC, RME) while at CDP, use of the Q_{li} observations increased melt rates (Figure 5.5 j).

5.7 Discussion

The results of the analysis indicated that radiation placed a significant control on physically-based snow model output across multiple climate zones. Empirical estimates of radiation generally were different enough from observed values to cause significant changes in model output (Figure 5.5). This result had not been found in previous studies due to a lack of sites that measured the complete radiation balance.

The strong relationship between model behavior and radiation representation holds implications for distributed hydrology and land surface models that utilize empirical approaches for estimating Q_{si}. For example, the ranges in bias and mean absolute errors at the four study sites were smaller than what was reported by the global assessment of *Bohn et al.* [2013], yet these relatively smaller errors still yielded strong controls on the behavior of SNTHERM (Figure 5.5). Alternative approaches to empirical methods (e.g., reanalysis data) for estimating Q_{si} also have been shown to exhibit high errors [*Shook and Pomeroy*, 2011; *Slater et al.*, 2013], which provides additional motivation for future research to improve radiation data for modeling.

One surprising result was the consistent divergence of modeled SWE depending on whether observed or estimated Q_{so} forced SNTHERM. Results at CDP, RME, and SASP (Figure 5.5) provided compelling evidence that modeling snowpack with estimated Q_{so} may produce more realistic snowmelt rates than with observed Q_{so}. This outcome was likely a result of the difficulties of measuring albedo and Q_{so} accurately. Measurement issues may have been related to dust-on-snow at SASP [*Painter et al.*, 2010] and atmospheric dust deposition and tree-litter on the snow surface at CDP [*Etchevers et al.*, 2004]. Q_{so} measurements at RME were impacted by

shading issues in the spring [*Danny Marks, personal communication, 2013*], and were taken at an exposed site away from the snow pillow where snow tended to disappear two weeks earlier. Together, these issues in the observed data may have enhanced the SNTHERM melt rates due to lower albedo at RME. These results resonate with the findings of *Blöschl* [1991], who found that uncertainty in albedo was more important than uncertainty in T_{air} and argued that more research was needed to understand albedo variations in time and space. Indeed, further research is needed to assess the utility of Q_{so} observations for testing snow models and remote sensing.

We acknowledge that there were some limitations to the model, methods, sites, and data of this study. First, we note that wind interacts with P uncertainty and may be strongly linked to SWE accumulation at exposed sites (e.g., IC); our analysis could not account for these effects, which may be substantial in tundra and alpine areas [*Sturm et al.*, 2001]. Model sensitivity to wind cannot be tested completely with SNTHERM because it does not represent redistribution processes, which may significantly control snow accumulation patterns [*Scipión et al.*, 2013; *Winstral et al.*, 2013]. Second, snow distributions at IC were complicated when patches of snow-free tundra emerged during snowmelt, and the 1-D snow model could not account for these 2-D effects [*Liston*, 1995]. Third, despite being in contrasting climates, the four study sites were in radiation-dominated environments, which made estimation of turbulent fluxes less important. We can only speculate that the relative importance of forcings might change in environments where turbulent fluxes are more dominant. Finally, we did not address interannual variability in this study, as it was difficult to acquire measurements of all forcings for multiple years without extended gaps in observations (e.g., power or sensor failure). Given a longer study period, we expect that the relative importance of forcings would be linked to the annual precipitation climate (i.e., wet vs. dry). We hypothesize that wetter conditions would extend snowmelt into

the sunnier summer months, and that drier conditions would result in earlier snow disappearance and less energy contribution from shortwave radiation.

5.8 Conclusions

We conducted a modeling experiment at four well-instrumented sites in varying snow climates. This experiment was motivated by a need to understand which surface forcings, beyond temperature and precipitation, most impact output of a physically-based snow model and therefore need better representation (via more observations and/or improved estimation methods). We found that the radiative forcings, which are measured least frequently in the western United States (Figure 5.1), caused the greatest divergences in model behavior. This was not entirely surprising, given the widely-acknowledged importance of radiation in the snowmelt energy budget as cited in the literature [*Male and Granger*, 1981] and further confirmed here. However, because the estimated radiation forcings were reasonably close to the reported accuracy and precision found in the literature (see Appendix B), the empirical methods for estimating radiation are sufficiently different to alter model output with statistical significance.

Because we are unable to discriminate errors in measured radiation, model uncertainty, and snow evaluation data, we conclude from model inference that radiation needs to be a high priority for the snow hydrology research community. The results of this study therefore advocate radiation sensors as the "next best sensors" to install at existing weather stations in cold regions for the purpose of testing and running snow model experiments. This recommendation inevitably demands more detailed consideration of important factors not assessed here (e.g., sensor cost and access, maintenance, spatial variability of radiation, etc). Future research should also continue to investigate methods for estimating and testing radiative forcings and methods for detecting and correcting errors in radiation measurements.

While our conclusions focus on radiation, we note that we corrected for precipitation uncertainty in the controlled experiment. At locations that lack precipitation observations, we expect that precipitation plays a significant role in the modeled snow variables (e.g., SWE, SDD) and this will be comparable (at a minimum) to the effect of melt-related data [*Raleigh and Lundquist*, 2012]. Ongoing and future research is evaluating the comparative sensitivity of the model to all forcing inputs at locations that do not have any measurements. This work is being investigated with formal uncertainty/sensitive analysis methods [*Matott et al.*, 2009] that are designed to assess and quantify the relationship in uncertainty between model inputs and outputs.

Finally, a general conclusion that can be drawn from the results is that the particular configuration and location of a weather station will invariably impact a modeler's perception of model performance. *Kuczera et al.* [2006] notes that most modelers ignore forcing data errors and uncertainty, but the results of this study demonstrate that the source of the forcing data may significantly change model output. Given the pervasive threat of equifinality in hydrology [*Beven and Binley*, 1992], a modeler must be ever critical of their forcings, model, and evaluation data.

5.9 Appendix B: Performance of the empirical methods for estimating forcings

The empirical methods used to estimate the "missing" forcings were evaluated against observed values and summary statistics were computed. These statistics included mean bias (Table 5.4), mean absolute error (MAE, Table 5.5), and root mean squared error (RMSE, Table 5.6) to allow comparisons with the variety of summary statistics reported in the literature [e.g., *Bohn et al.*, 2013; *Flerchinger et al.*, 2009; *Shook and Pomeroy*, 2011]. We do not show statistics for P, as bias was forced to 0 in our experiments and MAE and RMSE were generally

small. We also only show statistics for the forcings computed from observed values, and not for the forcings computed from estimated forcings (Figure 5.3b) for simplicity. We were unable to find studies that provided statistics on U errors and Q_{so} errors, but report our evaluation of these forcings for reference.

Performance of the T_{air} estimation method was worse than prior studies (Table 5.6), as hourly RMSE ranged from 1.7°C (CDP) to 3.0°C (IC) and *Cesaraccio et al.* [2001] reported RMSE of 1.5°C when evaluated at five sites in California. We assumed that there some climatic factors were involved, as CDP and California had similar RMSE and were both maritime locations, whereas IC was in the Arctic Tundra. The dependency of the *Cesaraccio et al.* [2001] method on solar geometry was problematic at IC near the winter and summer solstices, when sunrise and sunset times were undefined.

The RH estimation method had a positive bias at all sites (Table 5.4), ranging from 0.6% to 11%. For comparison, *Bohn et al.* [2013] showed bias ranging from -5.5% to 6.6% using the similar method of *Kimball et al.* [1997]. Averaged across the sites, the monthly MAE (Table 5.5) of 10.2% fell within the 9.3% to 10.7% ranged reported by *Bohn et al.* [2013]. Vapor pressure was calculated from T_{air} and RH and statistics were computed to permit comparison with other studies.

The *Thornton and Running* [1999] method yielded a bias ranging from -23.7 W m^{-2} at IC to +27.8 W m^{-2} at CDP, with a bias of 1.3 W m^{-2} when averaged across the four sites (Table 5.4). *Bohn et al.* [2013] generally found a negative bias at their sites, ranging from -0.7 W m^{-2} to -47 W m^{-2}. Monthly MAE (Table 5.5) at the four sites ranged from 10.9 W m^{-2} to 25.8 W m^{-2}, whereas MAE ranged from 14.7 W m^{-2} to 51.6 W m^{-2} in the *Bohn et al.* [2013] evaluation. Thus,

the *Thornton and Running* [1999] method had comparatively better agreement with the data at our study sites.

Finally, the combined approach for computing Q_{li} [*Dilley and O'Brien*, 1998; *Unsworth and Monteith*, 1975] had mean bias (Table 5.4) ranging from -3.0 W m^{-2} to 15.7 W m^{-2}, with a study average of 6.1 W m^{-2}. Mean RMSE across the sites was 34.1 W m^{-2} for hourly data and 24.6 W m^{-2} for daily data (Table 5.6). For comparison, *Flerchinger et al.* [2009] reported a bias of 1.1 W m^{-2} and RMSE of 24.5 W m^{-2} for hourly and 14.9 W m^{-2} for daily values. While the performance of the Q_{li} method was worse here, we note that *Flerchinger et al.* [2009] calibrated the method at their study sites and thus this discrepancy was not surprising.

5.10 Tables

Table 5.1 Basic characteristics of the snow study sites, ordered by elevation

Site Name	Site ID	Elevation (m)	Study Period (Water Year)	Snow Cover Classification[a]	Sensors	Operator(s)
Imnavait Creek	IC	930	2011	Tundra	T_{air}: Vaisala HMP45C P: Campbell Scientific TE 525 U: Met One 014A RH: Vaisala HMP45C Q_{si}: Kipp & Zonen CMA 6 Q_{so}: Kipp & Zonen CMA 6	NRCS, CRREL, Ameriflux
Col de Porte	CDP	1330	2006	Mountain (maritime)	T_{air}: PT 100/4 wires P: PG2000, GEONOR U: Chauvin Arnoux Tavid 87 – non-heated RH: Vaisala HMP 45D Q_{si}: Kipp & Zonen CM14 Q_{so}: Kipp & Zonen CM14 Q_{li}: Eppley PIR	Météo-France
Reynolds Mountain East (sheltered site)	RME	2060	2007	Mountain (intermountain)	T_{air}: Vaisala HMP 45 P: Belfort Universal Gages U: Met One 013/023 RH: Vaisala HMP 45 Q_{si}: Eppley Precision Spectral Pyranometer Q_{so}: Kipp & Zonen CNR1 Q_{li}: Eppley PIR	Northwest Watershed Research Center, Agricultural Research Service
Swamp Angel Study Plot	SASP	3370	2008	Mountain (continental)	T_{air}: Vaisala CS500 P: ETI Noah II U: RM Young Wind Monitor 05103-5 RH: Vaisala CS500 Q_{si}: Kipp & Zonen CM21 Q_{so}: Kipp & Zonen CM21 Q_{li}: Kipp & Zonen CG-4	Center for Snow and Avalanche Studies

[a] Based on *Sturm et al.* [1995]. Three of the sites are considered "mountain" snow covers and it is therefore difficult to classify them with the *Sturm et al.* system. The parenthesis denotes the most approximate climate description of the mountain sites.

Table 5.2 Percent of hours with missing data when snow was present at three of the study sites[a]

Site	T_{air}	P	U	RH	Q_{si}	Q_{so}	Q_{li}
IC	0.2%	4.1%	0.2%	0.2%	0.2%	0.2%	0.2%
CDP	0%	0%	0%	0%	0%	0%	0%
SASP	0%	0%	0%	0%	<0.1%	1.3%	<0.1%

[a] Statistics at RME are not presented but *Reba et al.* [2011] notes that missing values "represent less than one half of one percent of the record"
[b] Statistics at CDP are based on Figure 4 of *Morin et al.* [2012]

Table 5.3 Mean bias statistics for estimation of meteorological forcings. Precipitation was excluded because it had no bias (by design).

		Sites				
Hourly Forcing	**Dependencies**	**IC**	**CDP**	**RME**	**SASP**	**Overall**
T_{air} (°C)	$f(T_{max}, T_{min})$	-0.3	0.0	-0.2	-0.2	-0.2
U (m s^{-1})	NCEP	-0.17	0.40	0.61	0.78	0.41
RH (fractional)	$f(T_{air})$	0.062	0.060	0.20	0.106	0.107
Q_{si} (W m^{-2})	$f(T_{air}, RH, P)$	-23.7	22.6	5.7	0.8	1.3
Q_{so} (W m^{-2})	$f(T_{air}, RH, P, Q_{si})$	13.4	37.6	13.1	27.8	23.0
Q_{li} (W m^{-2})	$f(T_{air}, RH, Q_{si})$	9.1	-3.0	15.7	2.5	6.1

Table 5.4 Same as Table 5.3, but for mean absolute error (MAE), and with statistics from different intervals for comparison with other studies. Q_{so} at the hourly scale was not shown because only daily observed values were retained at noon.

			Sites				
Hourly Forcing	**Dependencies**	**Interval**	**IC**	**CDP**	**RME**	**SASP**	**Overall**
T_{air} (°C)	$f(T_{max}, T_{min})$	Hourly	2.0	1.2	1.4	1.6	1.6
		Daily	1.1	0.4	0.7	0.7	0.7
		Monthly	0.7	0.2	0.2	0.3	0.4
U (m s^{-1})	NCEP	Hourly	1.76	0.99	1.08	1.08	1.23
		Daily	1.49	0.80	0.80	0.81	0.98
		Monthly	0.86	0.41	0.64	0.78	0.67
RH (fractional)	$f(T_{air})$	Hourly	0.155	0.123	0.247	0.181	0.177
		Daily	0.135	0.098	0.199	0.155	0.147
		Monthly	0.067	0.069	0.153	0.117	0.102
Q_{si} (W m^{-2})	$f(T_{air}, RH, P)$	Hourly	66.0	56.0	94.0	79.9	74.0
		Daily	31.9	39.0	26.4	36.2	33.4
		Monthly	24.0	25.8	10.9	19.0	19.9
Q_{so} (W m^{-2})	$f(T_{air}, RH, P, Q_{si})$	Daily	26.4	39.2	--	35.9	33.8
		Monthly	21.2	37.4	--	30.6	29.7
Q_{li} (W m^{-2})	$f(T_{air}, RH, Q_{si})$	Hourly	28.7	21.1	29.0	35.3	28.5
		Daily	20.5	13.1	20.5	28.7	20.7
		Monthly	13.9	4.7	14.5	16.7	12.5

Table 5.5 Same as Table 5.4, but for root mean squared error (RMSE).

			Sites				
Hourly Forcing	**Dependencies**	**Interval**	**IC**	**CDP**	**RME**	**SASP**	**Overall**
T_{air} (°C)	*f(T_{max}, T_{min})*	Hourly	3.0	1.7	2.1	2.3	2.3
		Daily	1.6	0.6	0.8	1.0	1.0
		Monthly	0.9	0.2	0.2	0.4	0.4
U (m s^{-1})	NCEP	Hourly	2.27	1.25	1.33	1.35	1.55
		Daily	1.91	1.00	1.02	0.92	1.21
		Monthly	1.06	0.48	0.79	0.81	0.79
RH (fractional)	*f(T_{air})*	Hourly	0.204	0.164	0.312	0.228	0.227
		Daily	0.180	0.136	0.290	0.192	0.200
		Monthly	0.086	0.075	0.229	0.137	0.132
Q_{si} (W m^{-2})	*f(T_{air}, RH, P)*	Hourly	117.4	106.4	150.5	136.4	127.7
		Daily	46.9	54.4	40.0	49.8	47.8
		Monthly	32.3	37.0	9.9	22.6	25.5
Q_{so} (W m^{-2})	*f(T_{air}, RH, P, Q_{si})*	Daily	40.3	56.9	31.7	47.4	44.1
		Monthly	32.1	52.4	23.5	37.4	36.4
Q_{li} (W m^{-2})	*f(T_{air}, RH, Q_{si})*	Hourly	35.6	24.6	35.3	40.8	34.1
		Daily	25.2	15.5	24.3	33.3	24.6
		Monthly	16.6	5.6	17.1	19.0	14.6

5.11 Figures

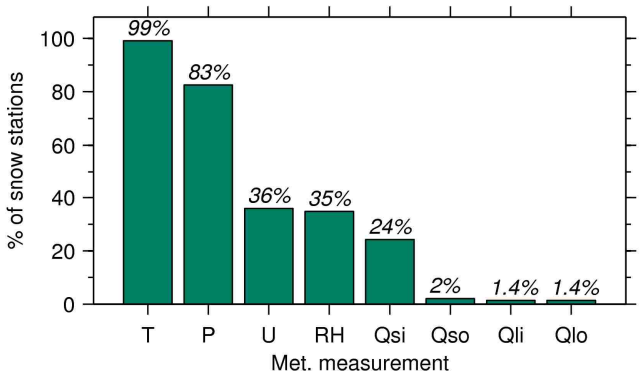

Figure 5.1 A survey of variables measured at weather stations (n=1318) in the western United States (i.e., NRCS SNOTEL domain) where either snow water equivalent (SWE) or snow depth (SD) are measured. Shown are the percentages of the snow stations measuring different meteorological variables. Measurement acronyms are defined in the text. Station operators include USDA (NRCS, USFS, Bureau of Reclamation), California Department of Water Resources, California Cooperative Snow Surveys, NOAA (NWS, COOP, HMT), Western Regional Climate Center, Desert Research Institute, US Army Corps of Engineers, US Army Cold Regions Research Laboratory, Fluxnet/Amerflux, Long Term Ecological Research sites, short-term research campaigns (e.g. NASA CLPX), NSF Critical Zone observatories, university research sites, state departments of transportation, airports (FAA), and avalanche centers. Lumped net radiation measurements are not tallied. Quality and serial completeness of the data are not reflected.

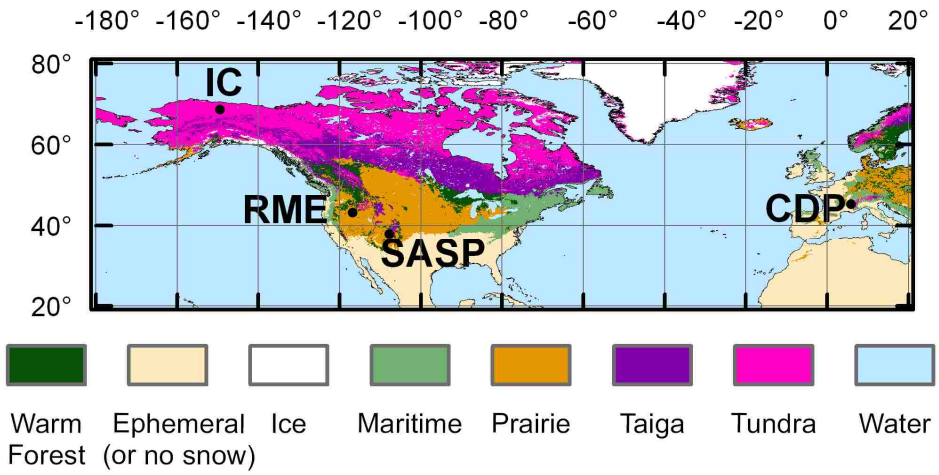

Figure 5.2 Locations of the study sites. The background shows an updated version of the *Sturm et al.* [1995] snow classification map [updated map courtesy of Glen Liston].

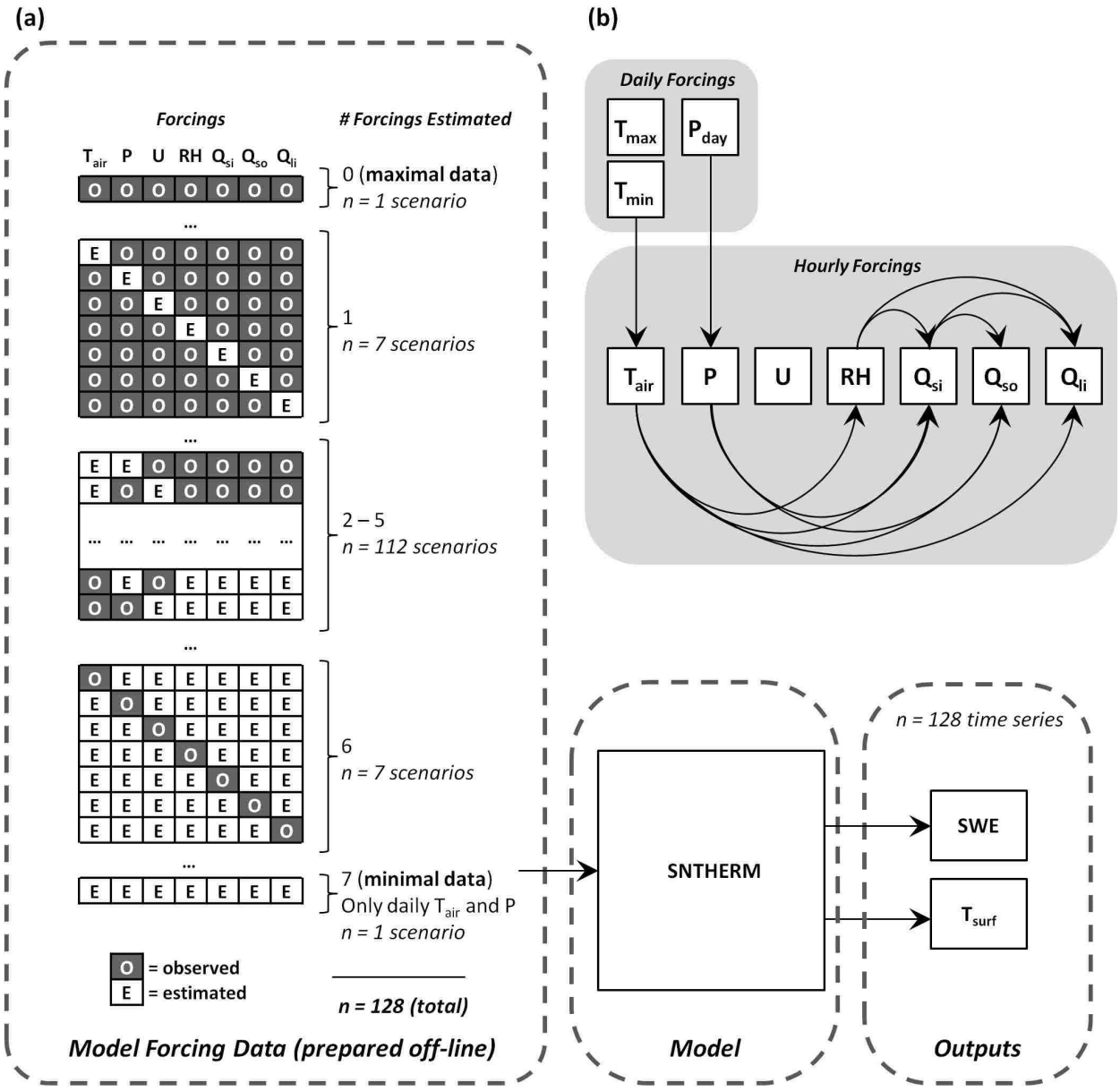

Figure 5.3 Conceptual workflow diagrams showing (a) the sequence of all combinations of observed forcing data used to create the data scenarios for snow model forcing and (b) the dependencies of estimated forcing data.

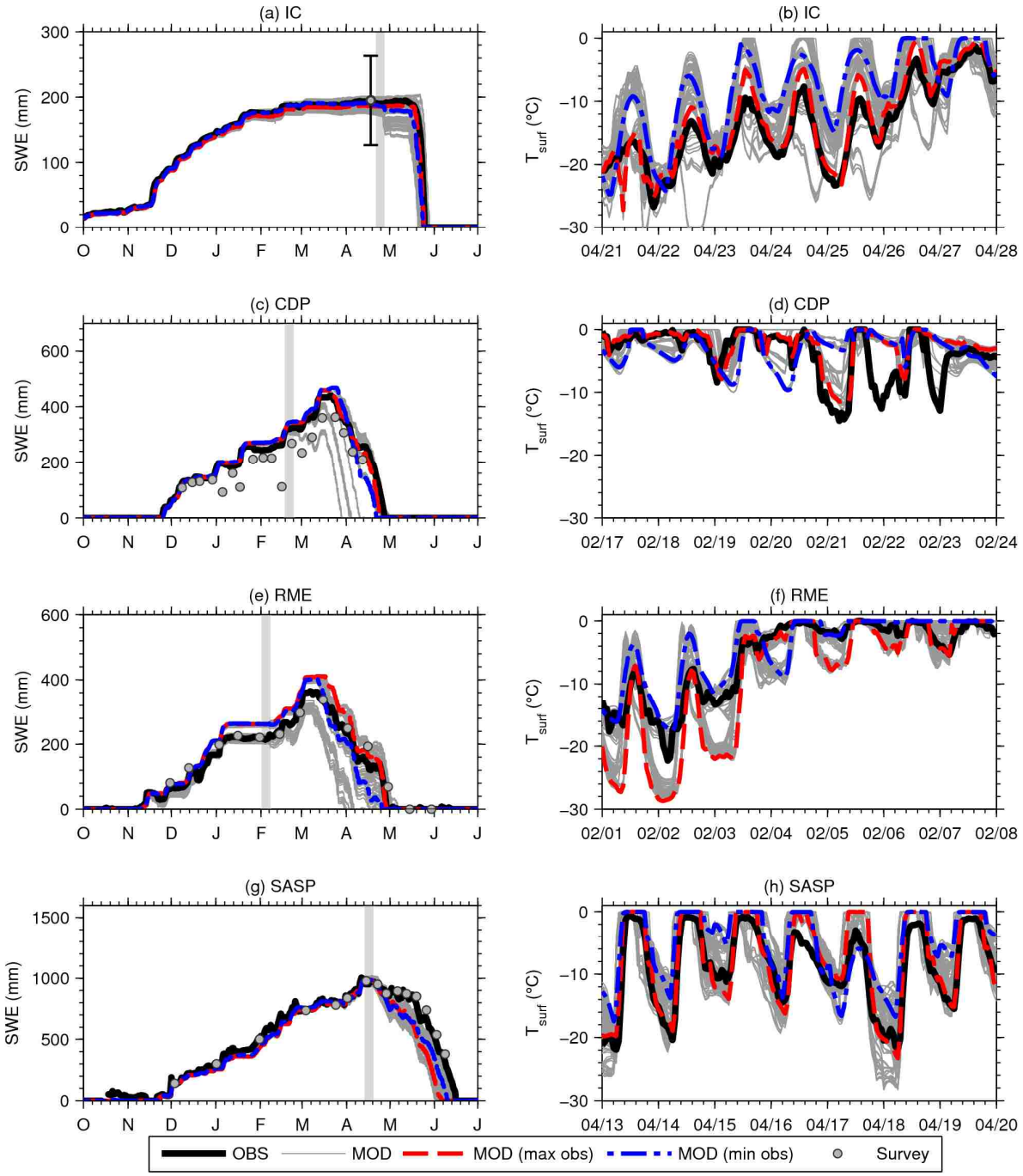

Figure 5.4 Response of SNTHERM SWE and snow surface temperature (T_{surf}) at the four study plots to uncertainty in model forcings based on n=128 data availability scenarios. Shown are the range of modeled observations (gray), the model simulation with all observed forcings (max obs), and the model simulation with no observed forcings beyond temperature and precipitation (min obs). Daily SWE evaluation data and snow survey/snow pit are also plotted. At IC, the error bars are +/- 1 std. deviation from the mean in a 1km x 1km spatial snow survey. The period of the T_{surf} plots is in the shaded region in the SWE plots.

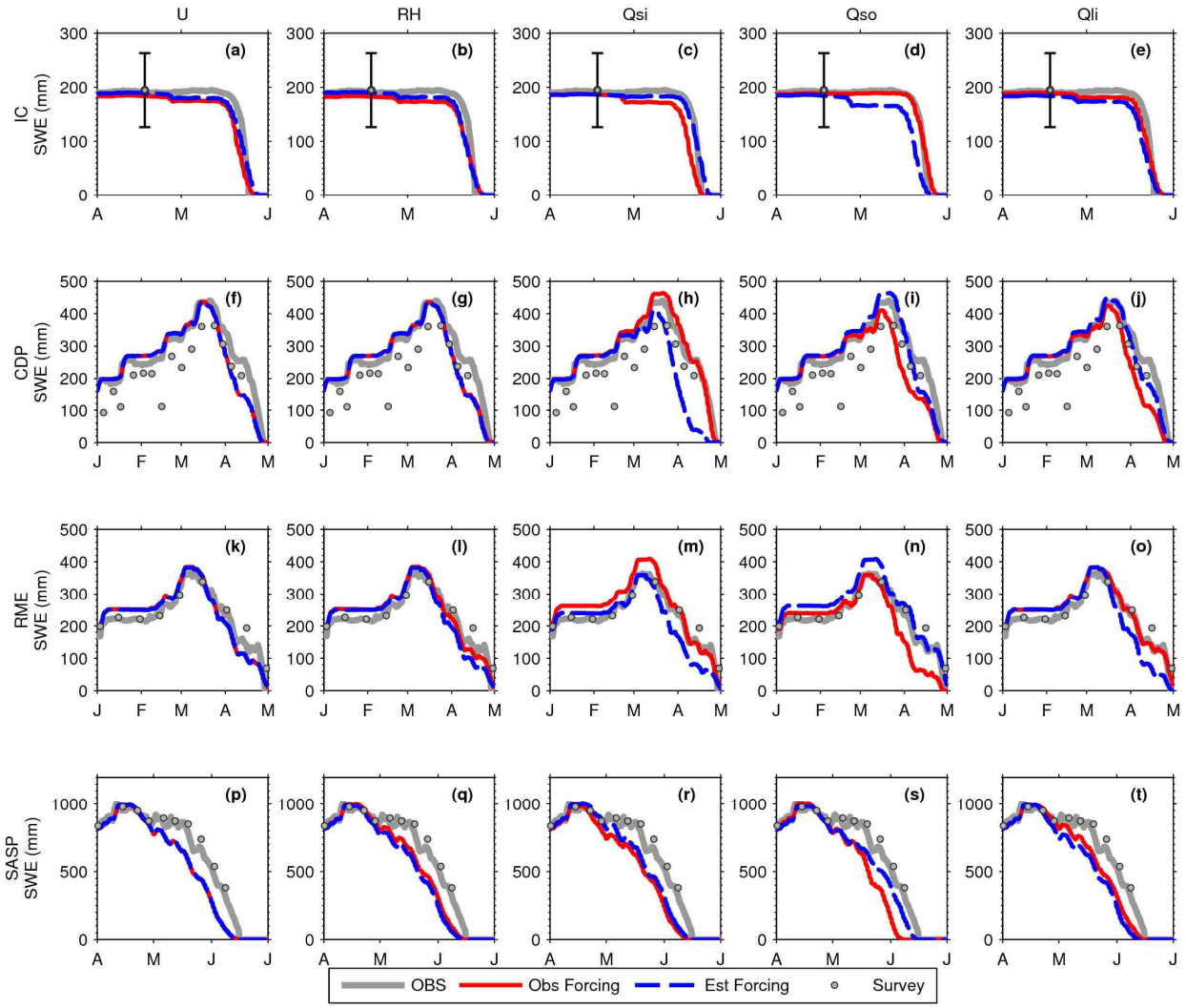

Figure 5.5 SWE simulations and evaluation data the four sites. Modeled SWE was averaged across 64 data scenarios based on whether each forcing (columns) was observed (red solid line) or estimated (blue dashed line).

Chapter 6 Conclusions

My research has sought to quantify uncertainties in numerical snow models and remote sensing, so as to understand the utility of these tools for research and management of water resources in snow-dominated catchments. Following the framework shown in Figure 1.1, the work has contributed to the field of snow hydrology through the comparisons and analysis of unique datasets and development of new methods for understanding snow states and processes. While comparing measurements, models, and remote sensing datasets is a useful approach for understanding relative uncertainties, it must be remembered that no single approach captures the complexity of the natural, real-world snowpack and that uncertainties of our tools must always be considered and quantified as rigorously as possible. The four chapters of this dissertation have attempted to quantify specific uncertainties by addressing different science questions.

In Chapter 2, I addressed the science question "*Which configuration of a snow model yields the most robust estimates of SWE and precipitation, given uncertainties in model forcing data and parameters*?" This chapter indicated that forward and reconstruction snow models may yield very similar error distributions for SWE, and supported the hypothesis that a combined model reduced SWE errors because of offsetting sensitivities to air temperature. More work is needed to understand how the advantages of the different model configurations may be best leveraged to reduce overall uncertainties in modeled SWE and precipitation, and whether more advanced methods of combing the model configurations is warranted. A combined, spatial estimate of SWE might be achieved with a weighting function based on proximity to precipitation/snow pillow sites and fractional forest cover so as to merge a gridded forward model with a satellite-based reconstruction model.

In assessing remote sensing uncertainty, I addressed the question, "*What is the accuracy of snow disappearance timing derived from MODIS fractional snow covered area at fine spatial scales (~1-2 km^2) across sites with varying forest density and topography*?" in Chapter 3. This chapter developed and tested an original ground-based approach for measuring fractional snow covered area, which was used to test MODIS in four sites with varying forest cover and elevation. The unique ground data showed that MODIS errors were on the order of 10-20% in meadows and 10-40% in forests, and showed snow disappearing 12 to 30 days too early in the forested sites. While it is generally understood that remote sensing uncertainty is larger in forests than in clearings, there was little quantification of this uncertainty prior to the study. Additionally, the results showed that a common vegetation correction method for snow covered area products was not robust in forested sites, particularly later in the ablation season. Future work should develop improved vegetation corrections for remotely sensed snow cover.

Chapter 4 explored the uncertainties in simple approximations of the snow surface temperature by addressing the question, "*Can standard height temperature and humidity approximate snow surface temperature, so as to allow detection of bias in a snowmelt model?*" Through comparing standard height temperature data with snow surface data at seven sites in different environments, it was found that the dew point temperature consistently had the lowest bias when approximating snow surface temperature. While there was more scatter for hourly values, the relationship between daily dew point and snow surface temperature was most robust. A modeling experiment demonstrated that the dew point can help diagnose systematic errors in the energy balance, which offers promise for evaluating snow models in remote locations. This study implied that more frequent monitoring of humidity at mountain weather stations and improved understanding of humidity fluctuations in space and time may be valuable for testing

energy balance snow models and remote sensing. Future work is needed to extend and refine these methods for evaluating modeled and remotely sensed snow surface temperature with humidity-based temperatures.

Chapter 5 addressed the question, "*Given the scarcity of meteorological stations in mountains areas, which meteorological forcings are most critical to measure for physically-based modeling of snow in different climates*?" By using data from four well-instrumented sites, I was able to conduct a controlled modeling experiment where 128 hypothetical weather stations were considered at each location, and the impact of data availability and empirical synthesis on model output was addressed. At all sites, the radiation components were the most critical measurements (assuming temperature and precipitation were already available). While net radiation is the dominant source of melt energy in many climates, this study showed that empirical surrogates are sufficiently different from observed values to cause significant divergences in model output. Future and ongoing work is considering how data uncertainty at unmonitored locations impacts snow model output (through formal sensitivity/uncertainty analysis) and how snow model selection interacts with forcing data uncertainty.

Taken as a whole, the research presented in this dissertation supports two ideas for advancing snow hydrology research in the future. First, the results imply that the associated uncertainties with model inputs and outputs must be carefully considered and reported. In the world of deterministic models, a single answer for a specific snow variable at each location in space and time is often given, but typically with little knowledge of the embedded uncertainties of that estimate. Explicitly considering uncertainties might be best achieved with routine uncertainty/sensitivity analysis or by treating the model forcing from a probabilistic perspective instead of a "single number" deterministic view. Second, the study supports the notion that

process-based snow models must be evaluated not only on the final product (i.e., SWE), but (possibly more importantly) the contributing processes as well [e.g., *Essery and Etchevers*, 2004]. The modeled snow surface temperature is one variable which has previously been difficult to test, but this research shows it is vital to do so, given its strong linkage with the surface energy balance. If applications of process-based models are to continue in the field, we must work to understand the uncertainties of these models and develop methods that test and evaluate their mechanics and contributing processes.

References

Abraha, M. G., and M. J. Savage (2008), Comparison of estimates of daily solar radiation from air temperature range for application in crop simulations, *Agricultural and Forest Meteorology*, *148*(3), 401–416, doi:10.1016/j.agrformet.2007.10.001.

Alduchov, O., and R. Eskridge (1996), Improved Magnus form approximation of saturation vapor pressure, *Journal of Applied Meteorology*, *35*, 601–609, doi:10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2.

Anderson, E. (1976), *A point energy and mass balance model of a snow cover*, NOAA Tech. Rep. NWS 19, Silver Spring, MD.

Anderson, H. (1956), Forest-cover effects on snowpack accumulation and melt, Central Sierra Snow Laboratory, *Trans. Amer. Geophys. Union*, *37*(3), 307–312.

Anderton, S. P., S. M. White, and B. Alvera (2004), Evaluation of spatial variability in snow water equivalent for a high mountain catchment, *Hydrological Processes*, *18*(3), 435–453, doi:10.1002/hyp.1319.

Andreadis, K. M., P. Storck, and D. P. Lettenmaier (2009), Modeling snow accumulation and ablation processes in forested environments, *Water Resources Research*, *45*(5), W05429, doi:10.1029/2008WR007042.

Andreas, E., P. Guest, P. Persson, C. Fairall, T. Horst, R. Moritz, and S. Semmer (2002), Near-surface water vapor over polar sea ice is always near ice saturation, *Journal of Geophysical Research*, *107*(C10), doi:10.1029/2000JC000411.

Andreas, E. L. (1986), A new method of measuring the snow-surface temperature, *Cold Regions Science and Technology*, *12*(2), 139–156, doi:10.1016/0165-232X(86)90029-7.

Andreas, E. L. (2002), Parameterizing Scalar Transfer over Snow and Ice: A Review, *Journal of Hydrometeorology*, *3*(4), 417–432, doi:10.1175/1525-7541(2002)003<0417:PSTOSA>2.0.CO;2.

Armstrong, R. L., and B. R. Armstrong (1987), Snow and avalanche climates of the western United States: a comparison of maritime, intermountain and continental conditions, *Avalanche Formation, Movement and Effects (Proc. of the Davos Symposium), IAHS*, *162*, 281–294.

Baker, F. S. (1944), Mountain climates of the western United States, *Ecological Monographs*, *14*(2), 223–254.

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 Resources Research*, *42*, W08432, doi:10.1029/2005WR004387.

Ball, R. A., L. C. Purcell, and S. K. Carey (2004), Evaluation of Solar Radiation Prediction Models in North America, , *397*, 391–397.

Barbour, M., N. Berg, T. Kittel, and M. Kunz (1991), Snowpack and the distribution of a major vegetation ecotone in the Sierra Nevada of California, *Journal of Biogeography*, *18*(2), 141–149.

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), 303–9, doi:10.1038/nature04141.

Beven, K., and A. Binley (1992), The future of distributed models: model calibration and uncertainty prediction, *Hydrological Processes*, *6*, 279–298.

Billings, W. (1969), Vegetational pattern near alpine timberline as affected by fire-snowdrift interactions, *Vegetatio*, *19*(1), 192–207.

Birkeland, K. W. (1998), Terminology and Predominant Processes Associated with the Formation of Weak Layers of Near-Surface Faceted Crystals in the Mountain Snowpack, *Arctic and Alpine Research*, *30*(2), 193–199, doi:10.2307/1552134.

Birkeland, K. W., R. F. Johnson, and D. S. Schmidt (1998), Near-Surface Faceted Crystals Formed by Diurnal Recrystallization: A Case Study of Weak Layer Formation in the Mountain Snowpack and Its Contribution to Snow Avalanches, *Arctic and Alpine Research*, *30*(2), 200–204, doi:10.2307/1552135.

Blöschl, G. (1991), The influence of uncertainty in air temperature and albedo on snowmelt., *Nord. Hydrol.*, *22*, 95–108.

Blöschl, G. (1999), Scaling issues in snow hydrology, *Hydrological Processes*, *13*, 2149–2175, doi:10.1002/(SICI)1099-1085(199910)13:14/15<2149::AID-HYP847>3.0.CO;2-8.

Bohn, T. J., B. Livneh, J. W. Oyler, S. W. Running, B. Nijssen, and D. P. Lettenmaier (2013), Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models, *Agricultural and Forest Meteorology*, *176*, 38–49, doi:10.1016/j.agrformet.2013.03.003.

Boone, A., and P. Etchevers (2001), An intercomparison of three snow schemes of varying complexity coupled to the same land surface model: Local-scale evaluation at an Alpine site, *Journal of Hydrometeorology*, *2*, 374–394.

Bormann, K. J., M. F. McCabe, and J. P. Evans (2012), Satellite based observations for seasonal snow cover detection and characterisation in Australia, *Remote Sensing of Environment*, *123*, 57–71, doi:10.1016/j.rse.2012.03.003.

Bristow, K. L., and G. S. Campbell (1984), On the relationship between incoming solar radiation and daily maximum and minimum temperature, *Agricultural and Forest Meteorology*, *31*(2), 159–166.

Van den Broeke, M., D. van As, C. Reijmer, and R. van de Wal (2004), Assessing and Improving the Quality of Unattended Radiation Observations in Antarctica, *Journal of Atmospheric and Oceanic Technology*, *21*(9), 1417–1431, doi:10.1175/1520-0426(2004)021<1417:AAITQO>2.0.CO;2.

Brubaker, K. L., R. T. Pinker, and E. Deviatova (2005), Evaluation and Comparison of MODIS and IMS Snow-Cover Estimates for the Continental United States Using Station Data, *Journal of Hydrometeorology*, *6*(6), 1002–1017, doi:10.1175/JHM447.1.

Brubaker, K., A. Rango, and W. Kustas (1996), Incorporating radiation inputs into the snowmelt runoff model, *Hydrological Processes*, *10*(10), 1329–1343, doi:10.1002/(SICI)1099-1085(199610)10:10<1329::AID-HYP464>3.0.CO;2-W.

Cesaraccio, C., D. Spano, P. Duce, and R. L. Snyder (2001), An improved model for determining degree-day values from daily temperature data, *International Journal of Biometeorology*, *24*, 161–169.

Church, J. E. (1914), Recent studies of snow in the United States, *Quarterly Journal of the Royal Meteorological Society*, *40*(169), 43–52.

Clagett, G. (1988), The Wyoming windshield—an evaluation after 12 years of use in Alaska, *Proceedings of the Western Snow Conference*, 113–123.

Clark, M. P., H. K. McMillan, D. B. G. Collins, D. Kavetski, and R. A. Woods (2011), Hydrological field data from a modeller's perspective: Part 2: process-based evaluation of model hypotheses, *Hydrological Processes*, *25*(4), 523–543, doi:10.1002/hyp.7902.

Clark, M. P., and J. a. Vrugt (2006), Unraveling uncertainties in hydrologic model calibration: Addressing the problem of compensatory parameters, *Geophysical Research Letters*, *33*(6), 1–5, doi:10.1029/2005GL025604.

Cline, D. et al. (2003), Overview of the NASA cold land processes field experiment (CLPX-2002), in *Microwave Remote Sensing of the Atmosphere and Environment III*, vol. 4894, edited by C. Kummerow, J. Jian, and S. Uratuka, pp. 361–372, Int. Society for Optical Engineering SPIE, Milpitas, CA.

Cline, D. W., R. C. Bales, and J. Dozier (1998), Estimating the spatial distribution of snow in mountain basins using remote sensing and energy balance modeling, *Water Resources Research*, *34*(5), 1275, doi:10.1029/97WR03755.

Coughlan, J., and S. Running (1997), Regional ecosystem simulation: A general model for simulating snow accumulation and melt in mountainous terrain, *Landscape Ecology*, *12*, 119–136.

Crawford, T., and C. Duchon (1999), An improved parameterization for estimating effective atmospheric emissivity for use in calculating daytime downwelling longwave radiation, *Journal of Applied Meteorology*, *38*, 474–480.

Cristea, N. C., and S. J. Burges (2009), An assessment of the current and future thermal regimes of three streams located in the Wenatchee River basin, Washington State: some implications for regional river basin systems, *Climatic Change*, *102*(3-4), 493–520, doi:10.1007/s10584-009-9700-5.

CSAS (2012), Archival Data from Senator Beck Basin Study Area, *Center for Snow and Avalanche Studies*. [online] Available from: http://snowstudies.org/data1.html

Dahe, Q., L. Shiyin, and L. Peiji (2006), Snow Cover Distribution, Variability, and Response to Climate Change in Western China, *Journal of Climate*, *19*(9), 1820–1833, doi:10.1175/JCLI3694.1.

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, *International Journal of Climatology*, *28*(15), 2031–2064, doi:10.1002/joc.1688.

Daly, C., R. P. Neilson, and D. L. Phillips (1994), A statistical–topographic model for mapping climatological precipitation over mountainous terrain, *J. Appl. Meteorol.*, *33*, 140–158.

Deems, J. S., S. R. Fassnacht, and K. J. Elder (2006), Fractal distribution of snow depth from LiDAR data, *Journal of Hydrometeorology*, *7*, 285–297.

Dilley, A., and D. O'Brien (1998), Estimating downward clear sky long-wave irradiance at the surface from screen temperature and precipitable water, *Quarterly Journal of the Royal Meteorological Society*, *124*(549), 1391–1401, doi:10.1256/smsqj.54902.

Dong, J., and C. Peters-Lidard (2010), On the Relationship Between Temperature and MODIS Snow Cover Retrieval Errors in the Western U.S., *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(1), 132–140, doi:10.1109/JSTARS.2009.2039698.

Dozier, J. (1984), Snow reflectance from Landsat-4 thematic mapper, *IEEE Transactions on Geoscience and Remote Sensing*, (3), 323–328.

Dozier, J. (2011), Mountain hydrology, snow color, and the fourth paradigm, *Eos, Transactions American Geophysical Union*, *92*(43), 373, doi:10.1029/2011EO430001.

Dozier, J., and J. Frew (2009), Computational provenance in hydrologic science: a snow mapping example., *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, *367*(1890), 1021–33, doi:10.1098/rsta.2008.0187.

Dozier, J., and T. H. Painter (2004), Multispectral and hypersepctral remote sensing of alpine snow properties, *Annual Review of Earth and Planetary Sciences*, *32*(1), 465–494, doi:10.1146/annurev.earth.32.101802.120404.

Dozier, J., T. H. Painter, K. Rittger, and J. Frew (2008), Time–space continuity of daily maps of fractional snow cover and albedo from MODIS, *Advances in Water Resources*, *31*(11), 1515–1526, doi:10.1016/j.advwatres.2008.08.011.

Duguay, C. (1993), Radiation modeling in mountainous terrain review and status, *Mountain Research and Development*, *13*(4), 339–357.

Durand, M., N. P. Molotch, and S. A. Margulis (2008), Merging complementary remote sensing datasets in the context of snow water equivalent reconstruction, *Remote Sensing of Environment*, *112*(3), 1212–1225, doi:10.1016/j.rse.2007.08.010.

Elder, K., D. Cline, G. E. Liston, and R. Armstrong (2009), NASA Cold Land Processes Experiment (CLPX 2002/03): Field Measurements of Snowpack Properties and Soil Moisture, *Journal of Hydrometeorology*, *10*(1), 320–329, doi:10.1175/2008JHM877.1.

Elder, K., J. Dozier, and J. Michaelsen (1989), Spatial and temporal variation of net snow accumulation in a small alpine watershed, Emerald Lake basin, Sierra Nevada, California, USA, *Annals of Glaciology*, *13*, 56–63.

Elder, K., W. Rosenthal, and R. E. Davis (1998), Estimating the spatial distribution of snow water equivalence in a montane watershed, *Hydrological Processes*, *12*(1011), 1793–1808, doi:10.1002/(SICI)1099-1085(199808/09)12:10/11<1793::AID-HYP695>3.3.CO;2-B.

Ellis, C. R., J. W. Pomeroy, T. Brown, and J. MacDonald (2010), Simulation of snow accumulation and melt in needleleaf forest environments, *Hydrology and Earth System Sciences*, *14*(6), 925–940, doi:10.5194/hess-14-925-2010.

Elsner, M. M., L. Cuo, N. Voisin, J. S. Deems, A. F. Hamlet, J. a. Vano, K. E. B. Mickelson, S.-Y. Lee, and D. P. Lettenmaier (2010), Implications of 21st century climate change for the hydrology of Washington State, *Climatic Change*, *102*(1-2), 225–260, doi:10.1007/s10584-010-9855-0.

Essery, R., and P. Etchevers (2004), Parameter sensitivity in simulations of snowmelt, *Journal of Geophysical Research*, *109*, 1–15, doi:10.1029/2004JD005036.

Essery, R., S. Morin, Y. Lejeune, and C. B Ménard (2013), A comparison of 1701 snow models using observations from an alpine site, *Advances in Water Resources*, *55*, 131–148, doi:10.1016/j.advwatres.2012.07.013.

Etchevers, P. et al. (2004), Validation of the energy budget of an alpine snowpack simulated by several snow models (SnowMIP project), *Annals of Glaciology*, *38*(1), 150–158, doi:10.3189/172756404781814825.

Farnes, P. (1967), Criteria for determining mountain snow pillow sites, in *Proc. 35th Western Snow Conf*, pp. 59–62, Boise, Idaho.

Fassnacht, S. R., K. A. Dressler, and R. C. Bales (2003), Snow water equivalent interpolation for the Colorado River Basin from snow telemetry (SNOTEL) data, *Water Resources Research*, *39*(8), doi:10.1029/2002WR001512.

Feld, S. I., N. C. Cristea, and J. D. Lundquist (2013), Representing atmospheric moisture content along mountain slopes: Examination using distributed sensors in the Sierra Nevada, California, *Water Resources Research*, *49*, doi:10.1002/wrcr.20318.

Feng, X., A. Sahoo, K. Arsenault, P. Houser, Y. Luo, and T. J. Troy (2008), The Impact of Snow Model Complexity at Three CLPX Sites, *Journal of Hydrometeorology*, *9*(6), 1464–1481, doi:10.1175/2008JHM860.1.

Flerchinger, G. N., W. Xaio, D. Marks, T. J. Sauer, and Q. Yu (2009), Comparison of algorithms for incoming atmospheric long-wave radiation, *Water Resources Research*, *45*(3), 1–13, doi:10.1029/2008WR007394.

Ford, K. R., A. K. Ettinger, J. D. Lundquist, M. S. Raleigh, and J. Hille Ris Lambers (2013), Spatial Heterogeneity in Ecologically Important Climate Variables at Coarse and Fine Scales in a High-Snow Mountain Landscape, edited by F. de Bello, *PLoS ONE*, *8*(6), e65008, doi:10.1371/journal.pone.0065008.

Forman, B. a., and S. a. Margulis (2009), High-resolution satellite-based cloud-coupled estimates of total downwelling surface radiation for hydrologic modelling applications, *Hydrology and Earth System Sciences Discussions*, *6*(2), 3041–3087, doi:10.5194/hessd-6-3041-2009.

Franz, K. J., T. S. Hogue, and S. Sorooshian (2008), Operational snow modeling: Addressing the challenges of an energy balance model for National Weather Service forecasts, *Journal of Hydrology*, *360*, 48–66, doi:10.1016/j.jhydrol.2008.07.013.

Garen, D. C., and D. Marks (2005), Spatially distributed energy balance snowmelt modelling in a mountainous river basin: estimation of meteorological inputs and verification of model results, *Journal of Hydrology*, *315*, 126–153, doi:10.1016/j.jhydrol.2005.03.026.

Gemmer, M., S. Becker, and T. Jiang (2004), Observed monthly precipitation trends in China 1951-2002, *Theoretical and Applied Climatology*, *77*(1-2), 39–45, doi:10.1007/s00704-003-0018-3.

Goodison, B., P. Louie, and D. Yang (1998), WMO solid precipitation measurement intercomparison: Final report, in *Instrum. Obs. Methods Rep. 67*, vol. 67, p. 211, World Meteorol. Organ., Geneva, Switzerland.

Greenfield, E., D. Nowak, and J. Walton (2009), Assessment of 2001 NLCD percent tree and impervious cover estimates, *Photogrammetric Engineering and Remote Sensing*, *75*(11), 1279–1286.

Grünewald, T., and M. Lehning (2011), Altitudinal dependency of snow amounts in two small alpine catchments: can catchment-wide snow amounts be estimated via single snow or precipitation stations?, *Annals of Glaciology*, *52*(58), 153–158.

Grünewald, T., M. Schirmer, R. Mott, and M. Lehning (2010), Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment, *The Cryosphere*, *4*(2), 215–225, doi:10.5194/tc-4-215-2010.

Halberstam, I., and J. Schieldge (1981), Anomalous Behavior of the Atmospheric Surface Layer over a Melting Snowpack., *Journal of Applied Meteorology*, *20*, 255 – 265.

Hall, D., J. Foster, A. Chang, C. Benson, and J. Y. L. Chien (1998), Determination of snow-covered area in different land covers in central Alaska, USA, from aircraft data--April 1995, *Annals of Glaciology*, *26*, 149–155.

Hall, D., J. Foster, and V. Salomonson (2001), Development of a technique to assess snow-cover mapping errors from space, *Geoscience and*, *39*(2), 432–438.

Hall, D. K., G. A. Riggs, V. V Salomonson, N. E. DiGirolamo, and K. J. Bayr (2002), MODIS snow-cover products, *Remote Sensing of Environment*, *83*(1-2), 181–194, doi:10.1016/S0034-4257(02)00095-0.

Hall, D. K., G. a. Riggs, and V. V. Salomonson (1995), Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data, *Remote Sensing of Environment*, *54*(2), 127–140, doi:10.1016/0034-4257(95)00137-P.

Hall, D., A. Tait, J. Foster, A. Chang, and M. Allen (2000), Intercomparison of satellite-derived snow-cover maps, *Annals of Glaciology*, *31*(1), 369–376.

Hamlet, A. F., D. Huppert, and D. P. Lettenmaier (2002), Economic Value of Long-Lead Streamflow Forecasts for Columbia River Hydropower, *Journal of Water Resources Planning and Management*, *128*(2), 91–101, doi:10.1061/(ASCE)0733-9496(2002)128:2(91).

Harder, P., and J. Pomeroy (2013), Estimating precipitation phase using a psychrometric energy balance method, *Hydrological Processes*, doi:10.1002/hyp.9799.

Hawkins, T., A. Ellis, J. Skindlov, and D. Reigle (2002), Intra-annual analysis of the North American snow cover-monsoon teleconnection: Seasonal forecasting utility, *Journal of Climate*, *15*, 1743–1753.

He, M., T. S. Hogue, K. J. Franz, S. a. Margulis, and J. a. Vrugt (2011a), Characterizing parameter sensitivity and uncertainty for a snow model across hydroclimatic regimes, *Advances in Water Resources*, *34*(1), 114–127, doi:10.1016/j.advwatres.2010.10.002.

He, M., T. S. Hogue, K. J. Franz, S. a. Margulis, and J. a. Vrugt (2011b), Corruption of parameter behavior and regionalization by model and forcing data errors: A Bayesian example using the SNOW17 model, *Water Resources Research*, *47*(7), 1–17, doi:10.1029/2010WR009753.

Helgason, W., and J. Pomeroy (2012), Problems Closing the Energy Balance over a Homogeneous Snow Cover during Midwinter, *Journal of Hydrometeorology*, *13*(2), 557–572, doi:10.1175/JHM-D-11-0135.1.

Henn, B., M. Raleigh, A. Fisher, and J. Lundquist (2013), A comparison of methods for filling gaps in hourly near-surface air temperature data, *Journal of Hydrometeorology, in press*.

Herrero, J., and M. J. Polo (2012), Parameterization of atmospheric longwave emissivity in a mountainous site for all sky conditions, *Hydrology and Earth System Sciences*, *16*(9), 3139–3147, doi:10.5194/hess-16-3139-2012.

Hey, A., S. Tansley, and K. Tolle (2009), *The fourth paradigm: data-intensive scientific discovery*, Microsoft Res., Redmond, WA.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis (2005), Very high resolution interpolated climate surfaces for global land areas, *International Journal of Climatology*, *25*(15), 1965–1978, doi:10.1002/joc.1276.

Homan, J. W., C. H. Luce, J. P. McNamara, and N. F. Glenn (2010), Improvement of distributed snowmelt energy balance modeling with MODIS-based NDSI-derived fractional snow-covered area data, *Hydrological Processes*, *25*(4), 650–660, doi:10.1002/hyp.7857.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan (2004), Development of a 2001 national land-cover database for the United States, *Photogrammetric Engineering and Remote Sensing*, *70*(7), 829–840.

Hood, J. L., and M. Hayashi (2010), Assessing the application of a laser rangefinder for determining snow depth in inaccessible alpine terrain, *Hydrology and Earth System Sciences Discussions*, *7*(1), 417–440, doi:10.5194/hessd-7-417-2010.

Hungerford, R., R. Nemani, S. Running, and J. Coughlan (1989), *MTCLIM: a mountain microclimate simulation model*, U.S. Forest Service Intermountain Research Station Research Paper INT-414, Ogden, UT.

Huwald, H., C. W. Higgins, M.-O. Boldi, E. Bou-Zeid, M. Lehning, and M. B. Parlange (2009), Albedo effect on radiative errors in air temperature measurements, *Water Resources Research*, *45*(8), 1–13, doi:10.1029/2008WR007600.

Iribarne, J., and W. Godson (1981), *Atmospheric Thermodynamics*, 2nd ed., D. Reidel, Dordrecht, Holland.

Jabot, E., I. Zin, T. Lebel, a. Gautheron, and C. Obled (2012), Spatial interpolation of sub-daily air temperatures for snow and hydrologic applications in mesoscale Alpine catchments, *Hydrological Processes*, *26*(17), 2618–2630, doi:10.1002/hyp.9423.

Jin, J., X. Gao, Z. Yang, R. Bales, S. Sorooshian, R. Dickinson, S. Sun, and G. Wu (1999), Comparative analyses of physically based snowmelt models for climate simulations, *Journal of Climate*, *12*, 2643–2657.

Jin, J., and N. L. Miller (2007), Analysis of the Impact of Snow on Daily Weather Variability in Mountainous Regions Using MM5, *Journal of Hydrometeorology*, *8*(2), 245–258, doi:10.1175/JHM565.1.

Johnson, J. B., and D. Marks (2004), The detection and correction of snow water equivalent pressure sensor errors, *Hydrological Processes*, *18*(18), 3513–3525, doi:10.1002/hyp.5795.

Jordan, R. (1991), A One-Dimensional Temperature Model for a Snow Cover: Technical Documentation for SNTHERM.89, p. 58, Special Report 91-16, US Army CRREL, Hanover, NH, USA.

Jost, G., M. Weiler, D. R. Gluns, and Y. Alila (2007), The influence of forest and topography on snow accumulation and melt at the watershed-scale, *Journal of Hydrology*, *347*(1-2), 101–115, doi:10.1016/j.jhydrol.2007.09.006.

Justus, C., W. Hargraves, A. Mikhail, and D. Graber (1978), Methods for estimating wind speed frequency distributions, *Journal of Applied Meteorology*, *17*, 350–353.

Justus, C., and A. Mikhail (1976), Height variation of wind speed and wind distributions statistics, *Geophysical Research Letters*, *3*(5), 261–264.

Juszak, I., and F. Pellicciotti (2013), A comparison of parameterizations of incoming longwave radiation over melting glaciers: Model robustness and seasonal variability, *Journal of Geophysical Research: Atmospheres*, *118*(February), n/a–n/a, doi:10.1002/jgrd.50277.

Kalnay, E., and Coauthors (1996), The NCEP/NCAR 40-year reanalysis project, *Bulletin of the American Meteorological Society*, *77*, 437–471, doi:http://dx.doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kane, D. L., L. D. Hinzman, C. S. Benson, and G. E. Liston (1991), Snow hydrology of a headwater Arctic basin: 1. Physical measurements and process studies, *Water Resources Research*, *27*(6), 1099–1109, doi:10.1029/91WR00262.

Kane, V. R., A. . Gillespie, R. McGaughey, J. A. Lutz, K. Ceder, and J. F. Franklin (2008), Interpretation and topographic compensation of conifer canopy self-shadowing, *Remote Sensing of Environment*, *112*(10), 3820–3832, doi:10.1016/j.rse.2008.06.001.

Kattelmann, R., and K. Elder (1991), Hydrologic characteristics and water balance of an alpine basin in the Sierra Nevada, *Water Resources Research*, *27*, 1553–1562.

Kavetski, D., G. Kuczera, and S. W. Franks (2006), Bayesian analysis of input uncertainty in hydrological modeling: 2. Application, *Water Resources Research*, *42*(3), n/a–n/a, doi:10.1029/2005WR004376.

Kimball, J., S. Running, and R. Nemani (1997), An improved method for estimating surface humidity from daily minimum temperature, *Agricultural and Forest Meteorology*, *85*, 87–98.

Kittredge, J. (1953), Influences of Forests on Snow in the Ponderosa, Sugar Pine, Fir Zone of the Central Sierra Nevada, *Hilgardia*, *22*, 1–96.

Klein, A., and A. Barnett (2003), Validation of daily MODIS snow cover maps of the Upper Rio Grande River Basin for the 2000–2001 snow year, *Remote Sensing of Environment*, *86*(2), 162–176, doi:10.1016/S0034-4257(03)00097-X.

Klein, A. G., D. K. Hall, and G. A. Riggs (1998), Improving snow cover mapping in forests through the use of a canopy reflectance model, *Hydrological Processes*, *12*, 1723–1744.

Koivusalo, H., and M. Heikinheimo (1999), Surface energy exchange over a boreal snowpack: comparison of two snow energy balance models, *Hydrological Processes*, *13*, 2395–2408.

Kondo, J., and T. Yamazaki (1990), A prediction model for snowmelt, snow surface temperature and freezing depth using a heat balance method, *Journal of Applied Meteorology*, *29*, 375–384.

Kondo, J., and H. Yamazawa (1986), Measurement of snow surface emissivity, *Boundary-Layer Meteorology*, *34*(4), 415–416.

Kuczera, G., D. Kavetski, S. Franks, and M. Thyer (2006), Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-

dependent parameters, *Journal of Hydrology*, *331*(1-2), 161–177, doi:10.1016/j.jhydrol.2006.05.010.

Kudo, G. (1991), Effects of snow-free period on the phenology of alpine plants inhabiting snow patches, *Arctic and Alpine Research*, *23*(4), 436–443.

Landry, C. ., K. A. Buck, M. P. Clark, and M. S. Raleigh (2013), Mountain system monitoring at Senator Beck Basin, San Juan Mountains, Colorado: a new integrative data source to develop and evaluate models of mountain processes, *Water Resources Research, in review*.

Lee, S., A. G. Klein, and T. M. Over (2005), A comparison of MODIS and NOHRSC snow-cover products for simulating streamflow using the Snowmelt Runoff Model, *Hydrological Processes*, *19*, 2951–2972, doi:10.1002/hyp.5810.

Legates, D., and C. Willmott (1990), Mean seasonal and spatial variability in global surface air temperature, *Theoretical and Applied Climatology*, *41*, 11–21.

Letsinger, S., and G. Olyphant (2007), Distributed energy-balance modeling of snow-cover evolution and melt in rugged terrain: Tobacco Root Mountains, Montana, USA, *Journal of Hydrology*, *336*, 48– 60, doi:10.1016/j.jhydrol.2006.12.012.

Liang, T., X. Huang, C. Wu, X. Liu, W. Li, Z. Guo, and J. Ren (2008), An application of MODIS data to snow cover monitoring in a pastoral area: A case study in Northern Xinjiang, China, *Remote Sensing of Environment*, *112*(4), 1514–1526, doi:10.1016/j.rse.2007.06.001.

Liang, X., D. Lettenmaier, E. Wood, and S. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *Journal of Geophysical Research*, *99*(D7), 14415–14428.

Liston, G. (1995), Local advection of momentum, heat, and moisture during the melt of patchy snow covers, *Journal of Applied Meteorology*, *34*, 1705–1715.

Liston, G. E. (1999), Interrelationships among Snow Distribution, Snowmelt, and Snow Cover Depletion: Implications for Atmospheric, Hydrologic, and Ecologic Modeling, *Journal of Applied Meteorology*, *38*(10), 1474–1487, doi:10.1175/1520-0450(1999)038<1474:IASDSA>2.0.CO;2.

Liston, G., and K. Elder (2006), A meteorological distribution system for high-resolution terrestrial modeling (MicroMet), *Journal of Hydrometeorology*, *7*, 217–234.

Litaor, M. I., M. Williams, and T. R. Seastedt (2008), Topographic controls on snow distribution, soil moisture, and species diversity of herbaceous alpine vegetation, Niwot Ridge, Colorado, *Journal of Geophysical Research*, *113*(G2), 1–10, doi:10.1029/2007JG000419.

Littell, J. S., D. L. Peterson, and M. Tjoelker (2008), Douglas-fir growth in mountain ecosystems: Water limits tree growth from stand to region, *Ecological Monographs*, *78*, 349–368, doi:10.1890/07-0712.1.

Liu, J., R. A. Melloh, C. E. Woodcock, R. E. Davis, and E. S. Ochs (2004), The effect of viewing geometry and topography on viewable gap fractions through forest canopies, *Hydrological Processes*, *18*(18), 3595–3607.

Liu, J., C. E. Woodcock, R. A. Melloh, R. E. Davis, C. McKenzie, and T. H. Painter (2008), Modeling the View Angle Dependence of Gap Fractions in Forest Canopies: Implications for Mapping Fractional Snow Cover Using Optical Remote Sensing, *Journal of Hydrometeorology*, *9*(5), 1005–1019, doi:10.1175/2008JHM866.1.

Livneh, B., Y. Xia, K. E. Mitchell, M. B. Ek, and D. P. Lettenmaier (2010), Noah LSM Snow Model Diagnostics and Enhancements, *Journal of Hydrometeorology*, *11*(3), 721–738, doi:10.1175/2009JHM1174.1.

López-Moreno, J. I., and D. Nogués-Bravo (2006), Interpolating local snow depth data: an evaluation of methods, *Hydrological Processes*, *20*(10), 2217–2232, doi:10.1002/hyp.6199.

Luce, C. H., D. G. Tarboton, and K. R. Cooley (1999), Sub-grid parameterization of snow distribution for an energy and mass balance snow cover model, *Hydrological Processes*, *13*, 1921–1933.

Lundquist, J. D., and D. R. Cayan (2007), Surface temperature patterns in complex terrain: Daily variations and long-term change in the central Sierra Nevada, California, *Journal of Geophysical Research*, *112*, D11124, doi:10.1029/2006JD007561.

Lundquist, J. D., D. R. Cayan, and M. D. Dettinger (2003), Meteorology and hydrology in Yosemite National Park: A sensor network application, *Proc. Second Int'l Workshop Information Processing in Sensor Networks (IPSN)*, 518–528.

Lundquist, J. D., and B. Huggett (2008), Evergreen trees as inexpensive radiation shields for temperature sensors, *Water Resources Research*, *44*, doi:10.1029/2008WR006979.

Lundquist, J. D., and F. Lott (2008), Using inexpensive temperature sensors to monitor the duration and heterogeneity of snow-covered areas, *Water Resources Research*, *44*, 8–13, doi:10.1029/2008WR007035.

Lutz, J. A., A. J. Larson, M. E. Swanson, and J. A. Freund (2012), Ecological Importance of Large-Diameter Trees in a Temperate Mixed-Conifer Forest, *PLoS ONE*, *7*(5), 1–15, doi:10.1371/journal.pone.0036131.

M Miles and Associates (2003), *British Columbia's climate-related observation networks: An adequacy review*, BC Ministry of Water, Land and Air Protection, Victoria, BC.

Ma, Y., and R. T. Pinker (2012), Modeling shortwave radiative fluxes from satellites, *Journal of Geophysical Research*, *117*(D23), D23202, doi:10.1029/2012JD018332.

Madani, K., and J. R. Lund (2009), Estimated impacts of climate warming on California's high-elevation hydropower, *Climatic Change*, *102*(3-4), 521–538, doi:10.1007/s10584-009-9750-8.

Mahrt, L., and D. Vickers (2005), Boundary-Layer Adjustment Over Small-Scale Changes of Surface Heat Flux, *Boundary-Layer Meteorology*, *116*(2), 313–330, doi:10.1007/s10546-004-1669-z.

Male, D., and R. Granger (1981), Snow surface energy exchange, *Water Resources Research*, *17*(3), 609–627.

Marks, D., and J. Dozier (1992), Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada: 2. Snow cover energy balance, *Water Resources Research*, *28*(11), 3043–3054.

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 Resources Research*, *28*(11), 3029–3042.

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, *Hydrological Processes*, *12*(10-11), 1569–1587, doi:10.1002/(SICI)1099-1085(199808/09)12:10/11<1569::AID-HYP682>3.0.CO;2-L.

Marks, D., A. Winstral, M. Reba, J. Pomeroy, and M. Kumar (2013), An evaluation of methods for determining during-storm precipitation phase and the rain/snow transition elevation at the surface in a mountain basin, *Advances in Water Resources*, *55*, 98–110, doi:10.1016/j.advwatres.2012.11.012.

Marquínez, J., J. Lastra, and P. García (2003), Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis, *Journal of Hydrology*, *270*, 1–11, doi:10.1016/S0022-1694(02)00110-5.

Marsh, P., and J. Pomeroy (1996), Meltwater fluxes at an arctic forest tundra site, *Hydrological Processes*, *10*, 1383–1400.

Matott, L. S., J. E. Babendreier, and S. T. Purucker (2009), Evaluating uncertainty in integrated environmental models: A review of concepts and tools, *Water Resources Research*, *45*(6), n/a–n/a, doi:10.1029/2008WR007301.

Maurer, E. P., J. D. Rhoads, R. O. Dubayah, and D. P. Lettenmaier (2003), Evaluation of the snow-covered area data product from MODIS, *Hydrological Processes*, *17*(1), 59–71, doi:10.1002/hyp.1193.

Maurer, E., A. Wood, J. Adam, D. Lettenmaier, and B. Nijssen (2002), A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States*, *Journal of Climate*, *15*(1993), 3237–3251.

McGuire, M., A. W. Wood, A. F. Hamlet, and D. P. Lettenmaier (2006), Use of Satellite Data for Streamflow and Reservoir Storage Forecasts in the Snake River Basin, *Journal of Water Resources Planning and Management*, *132*(2), 97, doi:10.1061/(ASCE)0733-9496(2006)132:2(97).

Meek, D., and J. Hatfield (1994), Data quality checking for single station meteorological databases, *Agricultural and Forest Meteorology*, *69*, 85–109.

Melloh, R. (1999), A synopsis and comparison of selected snowmelt algorithms, *US Army Corps of Engineers, CRREL Report 99-8*, 17.

Melloh, R. a., T. J. Hall, and R. Bailey (2004), Radiation data corrections for snow-covered sensors: are they needed for snowmelt modelling?, *Hydrological Processes*, *18*(6), 1113–1126, doi:10.1002/hyp.5510.

Meyer, J. D. D., J. Jin, and S.-Y. Wang (2012), Systematic Patterns of the Inconsistency between Snow Water Equivalent and Accumulated Precipitation as Reported by the Snowpack Telemetry Network, *Journal of Hydrometeorology*, *13*(6), 1970–1976, doi:10.1175/JHM-D-12-066.1.

Millar, C. I., and R. D. Westfall (2010), Distribution and Climatic Relationships of the American Pika (Ochotona princeps) in the Sierra Nevada and Western Great Basin, U.S.A.; Periglacial Landforms as Refugia in Warming Climates, *Arctic, Antarctic, and Alpine Research*, *42*(1), 76–88, doi:10.1657/1938-4246-42.1.76.

Minder, J. R., P. W. Mote, and J. D. Lundquist (2010), Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains, *Journal of Geophysical Research*, *115*, D14122, doi:10.1029/2009JD013493.

Mohammed, G. A., M. Hayashi, C. R. Farrow, and Y. Takano (2013), Improved characterization of frozen soil processes in the Versatile Soil Moisture Budget model, *Canadian Journal of Soil Science*, *93*(4), 511–531, doi:10.4141/cjss2012-005.

Molotch, N. P. (2009), Reconstructing snow water equivalent in the Rio Grande headwaters using remotely sensed snow cover data and a spatially distributed snowmelt model, *Hydrological Processes*, *23*(7), 1076–1089, doi:10.1002/hyp.7206.

Molotch, N. P., and R. C. Bales (2005), Scaling snow observations from the point to the grid element: Implications for observation network design, *Water Resources Research*, *41*, doi:10.1029/2005WR004229.

Molotch, N. P., and R. C. Bales (2006), SNOTEL representativeness in the Rio Grande headwaters on the basis of physiographics and remotely sensed snow cover persistence, *Hydrol. Processes*, *20*, 101002/hyp6128.

Molotch, N. P., and S. A. Margulis (2008), Estimating the distribution of snow water equivalent using remotely sensed snow cover data and a spatially distributed snowmelt model: A multi-resolution, multi-sensor comparison, *Advances in Water Resources*, *31*(11), 1503–1514, doi:10.1016/j.advwatres.2008.07.017.

Moore, R., and I. Owens (1984), Modelling alpine snow accumulation and ablation using daily climate observations, *J. Hydrol.(NZ)*, *23*(2), 73–83.

Morin, S., Y. Lejeune, B. Lesaffre, J.-M. Panel, D. Poncet, P. David, and M. Sudul (2012), An 18-yr long (1993–2011) snow and meteorological dataset from a mid-altitude mountain site (Col de Porte, France, 1325 m alt.) for driving and evaluating snowpack models, *Earth System Science Data*, *4*(1), 13–21, doi:10.5194/essd-4-13-2012.

Murray, F. (1967), On the computation of saturation vapor pressure, *Journal of Applied Meteorology*, *6*, 203–204.

Musselman, K. N., N. P. Molotch, S. A. Margulis, P. B. Kirchner, and R. C. Bales (2012), Influence of canopy structure and direct beam solar irradiance on snowmelt rates in a mixed conifer forest, *Agricultural and Forest Meteorology*, *161*, 46–56, doi:10.1016/j.agrformet.2012.03.011.

National Aeronautics and Space Administration (NASA) (2001), *Landsat 7 science data users' handbook*, Greenbelt, MD, USA.

Ninyerola, M., X. Pons, and J. M. Roure (2000), A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques, *International Journal of Climatology*, *20*(14), 1823–1841, doi:10.1002/1097-0088(20001130)20:14<1823::AID-JOC566>3.0.CO;2-B.

Nolin, A., and C. Daly (2006), Mapping "at risk" snow in the Pacific Northwest, *Journal of Hydrometeorology*, *7*, 1164–1171.

Nolin, A. W. (2010), Recent advances in remote sensing of seasonal snow, *Journal of Glaciology*, *56*(200), 1141–1150.

Nowak, D. J., and E. J. Greenfield (2010), Evaluating the national land cover database tree canopy and impervious cover estimates across the conterminous United States: a comparison with photo-interpreted estimates., *Environmental management*, *46*(3), 378–390, doi:10.1007/s00267-010-9536-9.

Ohmura, A. (2001), Physical Basis for the Temperature-Based Melt-Index Method, *Journal of Applied Meteorology*, *40*(4), 753–761, doi:10.1175/1520-0450(2001)040<0753:PBFTTB>2.0.CO;2.

Oreskes, N., K. Shrader-Frechette, and K. Belitz (1994), Verification, validation, and confirmation of numerical models in the earth sciences, *Science*, *263*, 641–646.

Painter, T. H., A. P. Barrett, C. C. Landry, J. C. Neff, M. P. Cassidy, C. R. Lawrence, K. E. McBride, and G. L. Farmer (2007), Impact of disturbed desert soils on duration of mountain snow cover, *Geophysical Research Letters*, *34*, L12502, doi:10.1029/2007GL030284.

Painter, T. H., J. S. Deems, J. Belnap, A. F. Hamlet, C. C. Landry, and B. Udall (2010), Response of Colorado River runoff to dust radiative forcing in snow., *Proceedings of the National Academy of Sciences of the United States of America*, *107*(40), 17125–30, doi:10.1073/pnas.0913139107.

Painter, T. H., K. Rittger, C. McKenzie, P. Slaughter, R. E. Davis, and J. Dozier (2009), Retrieval of subpixel snow covered area, grain size, and albedo from MODIS, *Remote Sensing of Environment*, *113*(4), 868–879, doi:10.1016/j.rse.2009.01.001.

Parajka, J., and G. Blöschl (2008), Spatio-temporal combination of MODIS images – potential for snow cover mapping, *Water Resources Research*, *44*(3), 1–13, doi:10.1029/2007WR006204.

Parajka, J., P. Haas, R. Kirnbauer, J. Jansa, and G. Blöschl (2012), Potential of time-lapse photography of snow for hydrological purposes at the small catchment scale, *Hydrological Processes*, doi:10.1002/hyp.8389.

Parlange, M., and R. Katz (2000), An extended version of the Richardson model for simulating daily weather variables, *Journal of Applied Meteorology*, *39*, 610–622.

Peck, L. (1994), Variation in Visual and Near-Infrared Contrast with a Snow Background, p. 27, US Army CRREL, Special Report 94-28.

Peck, L., and J. Fiori (1992), Soil temperatures at South Royalton, Vermont, USA, *NSIDC: National Snow and Ice Data Center*. [online] Available from: http://nsidc.org/data/ggd489.html

Peterson, E., and J. Hennessey (1978), On the use of power laws for estimates of wind power potential, *Journal of Applied Meteorology*, *17*, 390–394.

Pinker, R., and I. Laszlo (1992), Modeling surface solar irradiance for satellite applications on a global scale, *Journal of Applied Meteorology*, *31*, 194–211.

Pohl, S., P. Marsh, and G. Liston (2006), Spatial-temporal variability in turbulent fluxes during spring snowmelt, *Arctic, Antarctic, and Alpine Research*, *38*(1), 136–146.

Pomeroy, J., and R. Granger (1997), Sustainability of the western Canadian boreal forest under changing hydrological conditions. I. Snow accumulation and ablation, in *Sustainability of Water Resources under Increasing Uncertainty*, edited by D. Rosbjerg, B. Nour-Eddine, A. Gustard, Z. Kundzewics, and P. Rasmussen, pp. 237–242, IAHS Press, Wallingford.

Pomeroy, J., D. Gray, K. Shook, B. Toth, R. Essery, A. Pietroniro, and N. Hedstrom (1998), An evaluation of snow accumulation and ablation processes for land surface modelling, *Hydrological Processes*, *12*, 2339–2367.

Pomeroy, J., B. Toth, R. Granger, N. Hedstrom, and R. Essery (2003), Variation in surface energetics during snowmelt in a subarctic mountain catchment, *Journal of Hydrometeorology*, *4*(4), 702–719.

Prokop, A., M. Schirmer, M. Rub, M. Lehning, and M. Stocker (2008), A comparison of measurement methods: terrestrial laser scanning, tachymetry and snow probing for the determination of the spatial snow-depth distribution on slopes, *Annals of Glaciology*, *49*(1), 210–216, doi:10.3189/172756408787814726.

Pu, Z., L. Xu, and V. V. Salomonson (2007), MODIS/Terra observed seasonal variations of snow cover over the Tibetan Plateau, *Geophysical Research Letters*, *34*(6), 1–6, doi:10.1029/2007GL029262.

Qu, X., and A. Hall (2006), Assessing snow albedo feedback in simulated climate change, *Journal of Climate*, *19*, 2617–2630.

Raleigh, M. (2009), A statistical evaluation of a snow water equivalent reconstruction method using three snowmelt models at daily and hourly time steps, 148 pp., Dept. of Civil and Env. Eng., Seattle.

Raleigh, M. S., C. C. Landry, M. Hayashi, W. L. Quinton, and J. D. Lundquist (2013a), Standard temperature and humidity approximate snow surface temperature: new possibilities for snow model calibration, *Water Resources Research, in review*.

Raleigh, M. S., K. Lapo, and J. D. Lundquist (2013b), Impacts of forcing data uncertainty and availability on a physically-based snow model: the need for improving radiation representation, *Water Resources Research, in prep*.

Raleigh, M. S., and J. D. Lundquist (2012), Comparing and combining SWE estimates from the SNOW-17 model using PRISM and SWE reconstruction, *Water Resources Research*, *48*(1), 1–16, doi:10.1029/2011WR010542.

Raleigh, M. S., K. Rittger, C. E. Moore, B. Henn, J. A. Lutz, and J. D. Lundquist (2013c), Ground-based testing of MODIS fractional snow cover in subalpine meadows and forests of the Sierra Nevada, *Remote Sensing of Environment*, *128*, 44–57, doi:10.1016/j.rse.2012.09.016.

Ralph, F. M. et al. (2005), Improving Short-Term (0–48 h) Cool-Season Quantitative Precipitation Forecasting: Recommendations from a USWRP Workshop, *Bulletin of the American Meteorological Society*, *86*(11), 1619–1632, doi:10.1175/BAMS-86-11-1619.

Rango, A., and J. Martinec (1982), Snow accumulation derived from modified depletion curves of snow coverage, in *Hydrological Aspects of Alpine and High-Mountain Areas*, edited by J. Glen, pp. 83–90, International Assoc. Hydrological Sciences, Exeter, U.K.

Rasmussen, R. et al. (2012), How Well Are We Measuring Snow: The NOAA/FAA/NCAR Winter Precipitation Test Bed, *Bulletin of the American Meteorological Society*, *93*(6), 811–829, doi:10.1175/BAMS-D-11-00052.1.

Reba, M. L., D. Marks, M. Seyfried, A. Winstral, M. Kumar, and G. Flerchinger (2011a), A long-term data set for hydrologic modeling in a snow-dominated mountain catchment, *Water Resources Research*, *47*(7), W07702, doi:10.1029/2010WR010030.

Reba, M. L., D. Marks, A. Winstral, T. E. Link, and M. Kumar (2011b), Sensitivity of the snowcover energetics in a mountain basin to variations in climate, *Hydrological Processes*, *25*, doi:10.1002/hyp.8155.

Rice, R., and R. Bales (2008), Embedded sensor network design for spatial snowcover, in *Proc. 76th Western Snow Conference*, pp. 35–46, Hood River, Oregon.

Rice, R., R. C. Bales, T. H. Painter, and J. Dozier (2011), Snow water equivalent along elevation gradients in the Merced and Tuolumne River basins of the Sierra Nevada, *Water Resources Research*, *47*(8), 1–11, doi:10.1029/2010WR009278.

Richard, C., and D. Gratton (2001), The importance of the air temperature variable for the snowmelt runoff modelling using the SRM, *Hydrological Processes*, *15*, 3357–3370.

Richards, L. (1959), Forest densities, ground cover, and slopes in the snow zone of the Sierra Nevada west-side, *Technical Papers. Pacific Southwest Forest and Range Experiment Station*, *40*, 1–21.

Rittger, K., A. Kahl, and J. Dozier (2011), Topographic distribution of snow water equivalent in the Sierra Nevada, in *Proc. 79th Western Snow Conf.*, pp. 37–46, Stateline, Nevada.

Rittger, K., T. H. Painter, and J. Dozier (2012), Assessment of methods for mapping snow cover from MODIS, *Advances in Water Resources*, doi:10.1016/j.advwatres.2012.03.002.

Roe, G. (2005), Orographic precipitation, *Annu. Rev. Earth Planet. Sci.*, *33*, 645–671, doi:10.1146/annurev.earth.33.092203.122541.

Running, S. W., R. R. Nemani, and R. D. Hungerford (1987), Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest

evapotranspiration and photosynthesis, *Canadian Journal of Forest Research*, *17*(6), 472–483.

Rutter, N. et al. (2009), Evaluation of forest snow processes models (SnowMIP2), *Journal of Geophysical Research*, *114*, 1–18, doi:10.1029/2008JD011063.

Schnorbus, M., and Y. Alila (2004), Generation of an Hourly Meteorological Time Series for an Alpine Basin in British Columbia for Use in Numerical Hydrologic Modeling, *Journal of Hydrometeorology*, *5*(5), 862–882, doi:10.1175/1525-7541(2004)005<0862:GOAHMT>2.0.CO;2.

Scipión, D. E., R. Mott, M. Lehning, M. Schneebeli, and a. Berne (2013), Seasonal small-scale spatial variability in alpine snowfall and snow accumulation, *Water Resources Research*, *49*, n/a–n/a, doi:10.1002/wrcr.20135.

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 Resources Research*, *35*(7), 2145–2160, doi:10.1029/1999WR900090.

Sevruk, B. (1983), Correction of measured precipitation in the Alps using the water equivalent of new snow, *Nordic Hydrology*, *14*, 49–58.

Shamir, E., and K. P. Georgakakos (2006), Distributed snow accumulation and ablation modeling in the American River basin, *Advances in Water Resources*, *29*, 558–570, doi:10.1016/j.advwatres.2005.06.010.

Shea, C., and B. Jamieson (2011), Some fundamentals of handheld snow surface thermography, *The Cryosphere*, *5*(1), 55–66, doi:10.5194/tc-5-55-2011.

Shook, K., and J. Pomeroy (2011), Synthesis of incoming shortwave radiation for hydrological simulation, *Hydrology Research*, *42*(6), 433, doi:10.2166/nh.2011.074.

Sieck, L. C., S. J. Burges, and M. Steiner (2007), Challenges in obtaining reliable measurements of point rainfall, *Water Resources Research*, *43*, doi:10.1029/2005WR004519.

Simic, A., R. Fernandes, R. Brown, P. Romanov, and W. Park (2004), Validation of VEGETATION, MODIS, and GOES+ SSM/I snow-cover products over Canada based on surface snow depth observations, *Hydrological Processes*, *18*(6), 1089–1104, doi:10.1002/hyp.5509.

Singh, V., and D. Frevert (Eds.) (2002), *Mathematical Models of Small Watershed Hydrology and Applications*, Water Resources Publications, Highlands Ranch, Colorado.

Skaugen, T. (2007), Modelling the spatial variability of snow water equivalent at the catchment scale, *Hydrology and Earth System Sciences*, *11*(5), 1543–1550, doi:10.5194/hess-11-1543-2007.

Slater, A. G. et al. (2001), The Representation of Snow in Land Surface Schemes: Results from PILPS 2(d), *Journal of Hydrometeorology*, *2*(1), 7–25, doi:10.1175/1525-7541(2001)002<0007:TROSIL>2.0.CO;2.

Slater, A. G., A. P. Barrett, M. P. Clark, J. D. Lundquist, and M. S. Raleigh (2013), Uncertainty in seasonal snow reconstruction: Relative impacts of model forcing and image availability, *Advances in Water Resources*, *55*, 165–177, doi:10.1016/j.advwatres.2012.07.006.

Smith, J., and N. Berg (1982), The Sierra Ecology Project, vol. 3, p. 104, U.S. Department of the Interior, Washington, D.C.

Smith, M. B., D.-J. Seo, V. I. Koren, S. M. Reed, Z. Zhang, Q. Duan, F. Moreda, and S. Cong (2004), The distributed model intercomparison project (DMIP): motivation and experiment design, *Journal of Hydrology*, *298*, 4–26, doi:10.1016/j.jhydrol.2004.03.040.

Storck, P. (2000), Trees, snow and flooding: An investigation of forest canopy effects on snow accumulation and melt at the plot and watershed scales in the Pacific Northwest, in *Water Resourc. Series, Tech. Rep. 161*, p. 176, Dept. of Civil and Env. Eng., University of Washington.

Storck, P., D. Lettenmaier, and S. Bolton (2002), Measurement of snow interception and canopy effects on snow accumulation and melt in a mountainous maritime climate, Oregon, United States, *Water Resources Research*, *38*(11), 1–16.

Sturm, M., J. Holmgren, and G. Liston (1995), A Seasonal Snow Cover Classification System for Local to Global Applications, *Journal of Climate*, *8*, 1261–1283.

Sturm, M., J. McFadden, G. Liston, F. Chapin III, C. Racine, and J. Holmgren (2001), Snow-shrub interactions in Arctic tundra: A hypothesis with climatic implications, *Journal of Climate*, *14*, 336–344.

Sturm, M., B. Taras, G. E. Liston, C. Derksen, T. Jonas, and J. Lea (2010), Estimating Snow Water Equivalent Using Snow Depth Data and Climate Classes, *Journal of Hydrometeorology*, *11*(6), 1380–1394, doi:10.1175/2010JHM1202.1.

Sturm, M., and A. M. Wagner (2010), Using repeated patterns in snow distribution modeling: An Arctic example, *Water Resources Research*, *46*(12), 1–15, doi:10.1029/2010WR009434.

Tan, B., C. E. Woodcock, J. Hu, P. Zhang, M. Ozdogan, D. Huang, W. Yang, Y. Knyazikhin, and R. B. Myneni (2006), The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions, *Remote Sensing of Environment*, *105*(2), 98–114, doi:10.1016/j.rse.2006.06.008.

Tarboton, D., and C. Luce (1996), Utah Energy Balance Snow Accumulation and Melt Model (UEB), *Computer model technical description users guide, Utah Water Res. Lab. and USDA For. Serv. Intermt. Res. Station*, 64.

Tekeli, A. E., Z. Akyürek, A. Arda Şorman, A. Şensoy, and A. Ünal Şorman (2005), Using MODIS snow cover maps in modeling snowmelt runoff process in the eastern part of Turkey, *Remote Sensing of Environment*, *97*(2), 216–230, doi:10.1016/j.rse.2005.03.013.

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*, 211–228.

Thornton, P., H. Hasenauer, and M. White (2000), Simultaneous Estimation of Daily Solar Radiation and Humidity from observed Temperature and Precipitation: An Application over, *Complex Terrain in Austria', Agric. For. Meteorol.*, *104*, 255–271.

Thornton, P., S. Running, and M. White (1997), Generating surfaces of daily meteorological variables over large regions of complex terrain, *Journal of Hydrology*, *190*, 214–251.

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, *Nature Geoscience*, *5*(10), 705–709, doi:10.1038/ngeo1571.

Tsintikidis, D., K. P. Georgakakos, J. A. Sperfslage, D. E. Smith, and T. M. Carpenter (2002), Precipitation Uncertainty and Raingauge Network Design within Folsom Lake Watershed, *Journal of Hydrologic Engineering*, *7*(2), 175–184, doi:10.1061/(ASCE)1084-0699(2002)7:2(175).

Tyler, S. W., S. A. Burak, J. P. McNamara, A. Lamontagne, J. S. Selker, and J. Dozier (2008), Spatially distributed temperatures at the base of two mountain snowpacks measured with fiber-optic sensors, *Journal of Glaciology*, *54*(187), 673–679.

U.S. Army Corps of Engineers (1956), Snow hydrology, summary report of the snow investigations, *US Army Engineer Division, North Pacific*, 498.

U.S. Naval Observatory (1990), *Almanac for Computers*, Nautical Almanac Office, U.S. Naval Observatory, Washington, D.C.

Unsworth, M. H., and J. L. Monteith (1975), Long-wave radiation at the ground I. Angular distribution of incoming radiation, *Quarterly Journal of the Royal Meteorological Society*, *101*(427), 13–24, doi:10.1002/qj.49710142703.

USGS (2011), Seamless Data Warehouse, , (October). [online] Available from: http://seamless.usgs.gov/ (Accessed 20 October 2011)

Vano, J. a., M. J. Scott, N. Voisin, C. O. Stöckle, A. F. Hamlet, K. E. B. Mickelson, M. M. Elsner, and D. P. Lettenmaier (2010), Climate change impacts on water management and irrigated agriculture in the Yakima River Basin, Washington, USA, *Climatic Change*, *102*(1-2), 287–317, doi:10.1007/s10584-010-9856-z.

Varhola, A., N. C. Coops, M. Weiler, and R. D. Moore (2010a), Forest canopy effects on snow accumulation and ablation: An integrative review of empirical results, *Journal of Hydrology*, *392*(3-4), 219–233, doi:10.1016/j.jhydrol.2010.08.009.

Varhola, A., J. Wawerla, M. Weiler, N. C. Coops, D. Bewley, and Y. Alila (2010b), A New Low-Cost, Stand-Alone Sensor System for Snow Monitoring, *Journal of Atmospheric and Oceanic Technology*, *27*(12), 1973–1978, doi:10.1175/2010JTECHA1508.1.

Vikhamar, D., and R. Solberg (2003), Snow-cover mapping in forests by constrained linear spectral unmixing of MODIS data, *Remote Sensing of Environment*, *88*(3), 309–323, doi:10.1016/j.rse.2003.06.004.

Viviroli, D. et al. (2011), Climate change and mountain water resources: overview and recommendations for research, management and policy, *Hydrology and Earth System Sciences*, *15*(2), 471–504, doi:10.5194/hess-15-471-2011.

Viviroli, D., H. H. Dürr, B. Messerli, M. Meybeck, and R. Weingartner (2007), Mountains of the world, water towers for humanity: Typology, mapping, and global significance, *Water Resources Research*, *43*(7), 1–13, doi:10.1029/2006WR005653.

Viviroli, D., R. Weingartner, and B. Messerli (2003), Assessing the Hydrological Significance of the World's Mountains, *Mountain Research and Development*, *23*(1), 32–40, doi:10.1659/0276-4741(2003)023[0032:ATHSOT]2.0.CO;2.

Vogel, S. W. (2002), Usage of high-resolution Landsat 7 band 8 for single-band snow-cover classification, *Annals of Glaciology*, *34*(1), 53–57, doi:10.3189/172756402781818058.

Wagner, W., and J. Horel (2011), Observations and simulations of snow surface temperature on cross-country ski racing courses, *Cold Regions Science and Technology*, *66*(1), 1–11, doi:10.1016/j.coldregions.2010.12.003.

Waichler, S. R., and M. S. Wigmosta (2003), Development of Hourly Meteorological Values From Daily Data and Significance to Hydrological Modeling at H. J. Andrews Experimental Forest, *Journal of Hydrometeorology*, *4*(2), 251–263, doi:10.1175/1525-7541(2003)4<251:DOHMVF>2.0.CO;2.

Wall, M., A. Rechtsteiner, and L. M. Rocha (2003), Singular value decomposition and principal component analysis, in *A Practical Approach to Microarray Data Analysis*, edited by D. P. Berrar, W. Dubitzky, and M. Granzow, pp. 91–109, Kluwer, Norwell, MA.

Walter, T. M., E. S. Brooks, D. K. McCool, L. G. King, M. Molnau, and J. Boll (2005), Process-based snowmelt modeling: does it require more input data than temperature-index modeling?, *Journal of Hydrology*, *300*(1-4), 65–75, doi:10.1016/j.jhydrol.2004.05.002.

Wan, Z., and Z. Li (1997), A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data, *IEEE Transactions on Geoscience and Remote Sensing*, *35*(4), 980–996, doi:10.1109/36.602541.

Warren, S. (1982), Optical properties of snow, *Reviews of Geophysics and Space Physics*, *20*(1), 67–89.

Watson, F. G. R., W. B. Newman, J. C. Coughlan, and R. A. Garrott (2006), Testing a distributed snowpack simulation model against spatial observations, *Journal of Hydrology*, *328*(3-4), 453–466, doi:10.1016/j.jhydrol.2005.12.012.

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*, *Journal of Hydrometeorology*, *14*(3), 677–699, doi:10.1175/JHM-D-12-0102.1.

Weingartner, R., and C. Pearson (2001), A comparison of the hydrology of the Swiss Alps and the Southern Alps of New Zealand, *Mountain Research and Development*, *21*, 370–381, doi:10.1659/0276-4741(2001)021.

Wigmosta, M., L. Vail, and D. P. Lettenmaier (1994), A distributed hydrology-vegetation model for complex terrain, *Water Resources Research*, *30*(6), 1665–1679.

Williams, T. J., W. L. Quinton, and J. L. Baltzer (2013), Linear disturbances on discontinuous permafrost: implications for thaw-induced changes to land cover and drainage patterns, *Environmental Research Letters*, *8*(2), 025006, doi:10.1088/1748-9326/8/2/025006.

Winstral, A., K. Elder, and R. Davis (2002), Spatial snow modeling of wind-redistributed snow using terrain-based parameters, *Journal of Hydrometeorology*, *3*, 524–538.

Winstral, A., D. Marks, and R. Gurney (2009), An efficient method for distributing wind speeds over heterogeneous terrain, *Hydrological Processes*, *23*, 2526–2535, doi:10.1002/hyp.7141.

Winstral, A., D. Marks, and R. Gurney (2013), Simulating wind-affected snow accumulations at catchment to basin scales, *Advances in Water Resources*, *55*, 64–79, doi:10.1016/j.advwatres.2012.08.011.

Winther, J.-G., and D. Hall (1999), Satellite-derived snow coverage related to hydropower production in Norway: present and future, *International Journal of Remote Sensing*, *20*(15-16), 2991–3008.

Wiscombe, W., and S. Warren (1980), A model for the spectral albedo of snow. I: Pure snow, *Journal of the Atmospheric Sciences*, *37*, 2712 – 2733, doi:http://dx.doi.org/10.1175/1520-0469(1980)037<2712:AMFTSA>2.0.CO;2.

Xin, Q., C. E. Woodcock, J. Liu, B. Tan, R. a. Melloh, and R. E. Davis (2012), View angle effects on MODIS snow mapping in forests, *Remote Sensing of Environment*, *118*, 50–59, doi:10.1016/j.rse.2011.10.029.

Yamaguchi, Y., A. Kahle, H. Tsu, T. Kawakami, and M. Pniel (1998), Overview of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), *IEEE Transactions on Geoscience and Remote Sensing*, *36*(4), 1062–1071, doi:10.1109/36.700991.

Yang, D., D. Kane, L. Hinzman, B. Goodison, J. Metcalfe, P. Louie, G. Leavesley, D. Emerson, and C. Hanson (2000), An evaluation of the Wyoming gauge system for snowfall measurement, *Water Resources …*, *36*(9), 2665–2677.

Yang, K., and T. Koike (2005), A general model to estimate hourly and daily solar radiation for hydrological studies, *Water Resources Research*, *41*(10), n/a–n/a, doi:10.1029/2005WR003976.

Young, C. A., M. I. Escobar-Arias, M. Fernandes, B. Joyce, M. Kiparsky, J. F. Mount, V. K. Mehta, D. Purkey, J. H. Viers, and D. Yates (2009), Modeling the Hydrology of Climate Change in California's Sierra Nevada for Subwatershed Scale Adaptation, *JAWRA Journal of the American Water Resources Association*, *45*(6), 1409–1423, doi:10.1111/j.1752-1688.2009.00375.x.

Zhou, X., H. Xie, and J. M. H. Hendrickx (2005), Statistical evaluation of remotely sensed snow-cover products with constraints from streamflow and SNOTEL measurements, *Remote Sensing of Environment*, *94*(2), 214–231, doi:10.1016/j.rse.2004.10.007.

Zuzel, J., and L. Cox (1975), Relative importance of meteorological variables in snowmelt, *Water Resources Research*, *11*(I), 174–176.

CURRICULUM VITAE

Education

University of Washington, Seattle
Ph.D. in Civil & Environmental Engineering — *2013*

University of Washington, Seattle
Masters of Science in Civil Engineering — *2009*

Gonzaga University, Spokane
Bachelors of Science in Civil Engineering, Magna Cum Laude — *2005*

Employment Experience

University of Washington, Seattle, WA
Research Assistant — *2007 – 2013*
Instructor – CEE 345 Hydraulic Engineering — *2012*
Teaching Assistant – CEE 345 Hydraulic Engineering — *2009*
Grader – CEE 424 GIS for Civil Engineers — *2009 – 2012*

Merrick & Company, Aurora, CO
Design Engineer — *2005 – 2007*
Civil Engineering Intern — *2002 – 2004*

Fellowships and Scholarships

Hydro Research Foundation Fellowship — *2012 – 2013*

NASA Earth and Space Science Fellowship — *2009 – 2012*

Ronald E. Nece Fellowship, *University of Washington* — *2012*

US Society on Dams Grand Scholarship — *2012*

CH2M Hill Engineers Without Borders USA Scholarship — *2012*

AWWA – Ameron International Scholarship — *2010*

AWRA Washington Section Student Fellowship — *2008 – 2009*

Valle Scholarship, *University of Washington* — *2007 – 2008*

Regent's Scholarship, *Gonzaga University* — *2001 – 2005*

Awards

Dean's Award, Community of Innovators Award Ceremony — *2013*
College of Engineering, University of Washington

AGU Editor's Citation for Excellence in Refereeing — *2012*
Water Resources Research

Dr. James E. Church Memorial Award — *2011*
Best Student Paper at the Western Snow Conference

Professional Distinctions

Passed the Fundamentals of Engineering (FE) Exam *2004*

Peer-reviewed Publications

1. **Raleigh, M.S.**, Rittger, K., Moore, C.E., Henn, B., Lutz, J.A., and J.D. Lundquist (**2013**), Ground-based testing of MODIS fractional snow cover in subalpine meadows and forests of the Sierra Nevada, *Remote Sensing of Environment*, 128, 44-57, doi:10.1016/j.rse.2012.09.016.

2. **Raleigh, M.S.** and J.D. Lundquist (**2012**), Comparing and combining SWE estimates from the SNOW-17 model using PRISM and SWE reconstruction, *Water Resources Research*, *48*, W01506, doi:10.1029/2011WR010542.

3. Slater, A.G., Barrett, A.P., Clark, M.P., Lundquist, J.D., and **M.S. Raleigh** (**2013**), Uncertainty in seasonal snow reconstruction: relative impacts of model forcing and image availability, *Advances in Water Resources*, doi: 10.1016/j.advwatres.2012.07.006.

4. Henn, B., **Raleigh, M.S.**, Fisher, A., and J.D. Lundquist (**2013**), A comparison of methods for filling gaps in hourly near-surface air temperature data, *Journal of Hydrometeorology*, doi: 10.1175/JHM-D-12-027.1.

5. Ford, K.R., Ettinger, A.K., Lundquist, J.D., **Raleigh, M.S.**, and J.H.R. Lambers, Spatial heterogeneity in ecologically important climate variables at coarse and fine scales, *accepted by PLOS ONE*, April 2013.

Manuscripts in Review/Revision/Prep

Raleigh, M.S., Landry, C.C., Hayashi, M., Quinton, W.L., and J.D. Lundquist (2013), Standard temperature and humidity approximate snow surface temperature: new possibilities for snow model calibration, *submitted to Water Resources Research*, April 2013.

Raleigh, M.S., Lapo, K., and J.D. Lundquist (2013), Impacts of driving data uncertainty and availability on a physically-based snowmelt model (working title), *in prep*.

Landry, C. C., Buck, K.A., **M. S. Raleigh**, and M.P. Clark (2013), Mountain system monitoring at Senator Beck Basin, San Juan Mountains, Colorado: a new integrative data source to develop and evaluate models of mountain processes, *submitted to Water Resources Research,* July 2013.

Dickerson-Lange, S.E., Lutz, J.A., Martin, K.A., **Raleigh, M.S.**, and J.D. Lundquist (2013), Quantifying the variability of snow depth and presence across diverse forests in a maritime watershed, *in prep*.

Conference Presentations

Raleigh, M.S., and J.D. Lundquist, Modeling in the dark: how data scarcity and uncertainty impact snowmelt modeling, Davos Atmosphere and Cryosphere Assembly (DACA-13), Davos, Switzerland, July 8-12, **2013**.

Raleigh, M.S., and J.D. Lundquist, Different yet Similar: Studying Snow Cover Patterns in Wet and Dry Years Using Distributed Temperature Sensors in the Sierra Nevada, Yosemite Hydroclimate Meeting, Yosemite National Park, California, October 11-12, **2012**.

Raleigh, M.S., K. Rittger, and J.D. Lundquist, What lies beneath? Comparing MODIS fractional snow covered area against ground-based observations under forest canopies and in meadows of the Sierra Nevada, 79th Western Snow Conference, Stateline, Nevada, April 18 – 21, **2011**.

Raleigh, M.S. and J.D. Lundquist, An alternative approach to predicting snowfall across the Sierra Nevada, American Meteorological Society, 14th Conference on Mountain Meteorology, Squaw Valley, California, August 30 – September 3, **2010**.

Raleigh, M.S. and J.D. Lundquist, Calculating snowmelt backwards – using the date of snowpack disappearance to determine how much snow fell over a season, 2009 Geological Society of America Meeting, Portland, Oregon, October 18-21, **2009**.

Raleigh, M.S. and J.D. Lundquist, Rain-on-snow events in a warmer world: hydrologic implications for the Sierra Nevada, 2009 Graduate Climate Conference, Pack Forest Conference Center, Washington, April 17-18, **2009**.

Conference Poster Presentations

Raleigh, M.S., Rittger, K., and J.D. Lundquist, Buried Treasure: Using distributed ground temperature sensors to test remote sensing of fractional snow cover, American Geophysical Union Fall Meeting, San Francisco, California, December 3-7, **2012**.

Raleigh, M.S., Improving representation of high-elevation snowpack for summer water supply forecasting, United States Society on Dams, 2012 Annual Meeting and Conference, New Orleans, Louisiana, April 23 – 27, **2012**.

Raleigh, M.S. and J.D. Lundquist, Snowfall Accumulation in the Western United States: Comparing estimates from SWE reconstruction and PRISM, American Geophysical Union Fall Meeting, San Francisco, California, December 13-17, **2010**.

Raleigh, M.S. and J.D. Lundquist, A snow hydrologist's time machine: determining winter snow accumulation with springtime mass and energy exchanges at the snow-air interface, CUAHSI 2nd Biennial Colloquium, Boulder, Colorado, July 19-21, **2010**.

Raleigh, M.S. and J.D. Lundquist, Calculating snowmelt backwards – using the date of snowpack disappearance to determine how much snow fell over a season, American Geophysical Union Fall Meeting, San Francisco, California, December 14-18, **2009**.

Raleigh, M.S., F.C. Lott, and J.D. Lundquist, Evaluation of Precipitation Scaling Using the Observed Snow Cover Disappearance Date, American Geophysical Union Fall Meeting, San Francisco, California, December 15-19, **2008**.

Non-Peer Reviewed Publications

Raleigh, M., Modeling snow water equivalent: what is the ideal approach?, Mountain Views Newsletter, Consortium for Integrated Climate Research in Western Mountains (CIRMOUNT), Vol. 6, No. 1, 16-18, April 2012.

Raleigh, M., Not a Household Name, ASEE PRISM Magazine, January 2013.

Raleigh, M., Let's Elect the 1 Percent!, ASEE PRISM Magazine, October 2012.

Raleigh, M., Cheers!, ASEE PRISM Magazine, April 2012.

Raleigh, M., Teaching Without Borders, ASEE PRISM Magazine, January 2012.

Raleigh, M., It's Not Science Fiction, ASEE PRISM Magazine, October 2011.

Reviewer for Research Journals

Arctic, Antarctic, and Alpine Research
Cold Regions Science and Technology
Journal of Hydrology
Journal of Hydrometeorology
Remote Sensing of Environment
Water Resources Research

Professional Society Memberships

American Geophysical Union
American Meteorological Society
American Society for Engineering Education
American Water Resources Association
American Water Works Association
Engineers Without Borders
United States Society on Dams

Field Experience

Energy flux tower (Turbulent Fluxes, Radiation, Ground Flux) *2012 - 2013*
Snoqualmie Pass, Washington

NOAA Hydrometeorological Testbed Project *2007 – 2012*
American River Basin, California

Snow cover study plots *2009 – 2012*
Onion Creek Experimental Forest & Yosemite National Park, California

Snow and forest dynamics plot *2009 – 2012*
Cedar River Watershed, Washington

Leadership Experience

Engineers Without Borders, *University of Washington Chapter*

- President *2009 – 2011*
- Vice President *2008 – 2009*

American Water Resources Association, *University of Washington Chapter*

- President *2008 – 2009*
- Professional Liaison *2009 – 2010*
- Webmaster *2007 – 2013*

Tau Beta Pi Engineering Honor Society, *Gonzaga University Chapter*

- President *2005*
- Vice President *2004*

Service

Academic

- 2012 Civil engineering graduate student representative – Hydrologic Extremes Faculty Search Committee

Environmental

- 2012 Watershed restoration volunteer – Duwamish Alive!

Conference Chairing

- 2012 Graduate Climate Conference – Terrestrial Impacts Session

Conference Planning Committees

- 2010 Graduate Climate Conference (Pack Forest Conference Center, WA)
- 2009 American Water Resources Association National Conference (Seattle, WA)
- 2008 Engineers Without Borders International Conference (Seattle, WA)

Humanitarian

- 2008 Engineers Without Borders Roadway Improvement Project (Acasio, Bolivia) – Design and Implementation Team

Mentoring

- 2008 Future Cities Competition – Team Engineer (McClure Middle School, *Seattle)*

Tutoring

- 2006-2007 Math and Science Tutor – Colorado Association of Black Professional Engineers and Scientists
- 2005 Math and Physics Tutor – Tau Beta Pi