5-2022

Improving Flood Inundation and Streamflow Forecasts in Snowmelt Dominated Regions

Irene Garousi-Nejad
Utah State University

IMPROVING FLOOD INUNDATION AND STREAMFLOW FORECASTS IN SNOWMELT DOMINATED REGIONS

by

Irene Garousi-Nejad

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

of

DOCTOR OF PHILOSOPHY

in

Civil and Environmental Engineering

Approved:

David G. Tarboton, Sc.D. Major Professor	Jeffery S. Horsburgh, Ph.D. Committee Member
Alfonso Torres-Rua, Ph.D. Committee Member	S.-Y. Simon Wang, Ph.D. Committee Member
Robert R. Gillies, Ph.D. Committee Member	D. Richard Cutler, Ph.D. Interim Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2022

To my family,

I hope you think your sacrifices were worth it.

ABSTRACT

Improving Flood Inundation and Streamflow Forecasts

in Snowmelt Dominated Regions

by

Irene Garousi-Nejad, Ph.D.

Utah State University, 2022

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

Projected changes in temperature and precipitation for the 21st century are likely to have significant effects on the hydrology of the western U.S., including loss of natural water storage as snowpack, changes in runoff timing from summer to winter, and increasing intensity of floods and extreme low flows. Accurate predictions of streamflow and snowmelt quantity, timing, and spatial pattern are important for decision making in water-sensitive sectors such as emergency management during flood seasons to prepare for flooding events or during spring and summer seasons to supply water for agricultural, municipal, etc. demands. Faced with this necessity for prediction, the National Oceanic and Atmospheric Administration (NOAA) has embarked on transforming their U.S.-wide water prediction capabilities via the National Water Model (NWM), a 21st-century take on water forecasting. The NWM is a specific configuration of the WRF-Hydro community model. The research in this dissertation evaluated opportunities to improve process-based understanding of WRF-Hydro configured as the NWM in continental-scale hydrological flood and water supply modeling—through three main approaches. First, we

investigated improvements to the Continental-scale Flood Inundation Mapping (CFIM) approach using high-resolution hydrography data to resolve the overestimation of flooded areas where there are barriers such as roads. Secondly, we investigated the NWM snow outputs to examine sources of errors and biases in model outputs and identify areas where predictions from the NWM involving snow may be better or worse. Thirdly, we developed advancements to the representation and parameterization of snow process in the model centered on the rain snow separation. Results of this research showed that methods developed can improve the overestimation of flood extent areas in CFIM and help reduce the general under-estimation of the snow water equivalent and early melt issue in the NWM. Taken together, the results of these evaluations and improvements contribute to the knowledge needed for hydrological modelers to advance the creation of geographically specific, value-added, and tailored solutions that will enhance the capability for hydrologic response prediction where flood and water supply forecasts are essential for water management.

(236 pages)

PUBLIC ABSTRACT

Improving Flood Inundation and Streamflow Forecasts

in Snowmelt Dominated Regions

Irene Garousi-Nejad

Much effort has been dedicated to expanding hydrological forecasting capabilities and improving understanding of the continental-scale hydrological modeling used to predict future hydrologic conditions and quantify consequences of climate change. In 2016, the National Oceanic and Atmospheric Administration's (NOAA) Office of Water Prediction implemented the National Water Model (NWM) to provide nationally consistent, operational hydrologic forecasting capability across the continental U.S. The primary goal of this research was to develop hydrological tools that include modeling of flood inundation mapping and snowmelt contributions to river flow in snowmelt-dominated regions across the Western U.S. This dissertation first presents terrain analysis enhancements developed to reduce the overestimation of flooded areas, observed where barriers such as roads cross rivers, from the continental-scale flood inundation mapping method that uses NWM streamflow forecasts. Then, it reports on a systematic evaluation of the NWM snow outputs against observed snow water equivalent (SWE) and snow-covered area fraction (SCAF) at point locations across the Western U.S. This evaluation identified the potential causes responsible for discrepancies in the model snow outputs and suggests opportunities for future research directed towards model improvements. Then, it presents improvements to SWE modeling by quantifying the improvements when using better model inputs and implementing humidity information in separating

precipitation into rain and snow. These results inform understanding of continental-scale hydrologic processes and how they should be modeled.

ACKNOWLEDGMENTS

Many intelligent and generous people contributed to completing this dissertation that I sincerely want to thank

Dr. David Tarboton kindly agreed to have me join his research team and become my advisor in September 2016. His endless support, insightful wisdom, thoughtful advice, and inspiring encouragement made this journey fruitful and enjoyable. I am most grateful that I had the opportunity to study, research, and learn hydrology under his supervision and mentorship. David has always provided me with so many opportunities to explore new problem-solving approaches and gain courage and confidence by every experiment we designed in this research.

I would like to thank my committee members for being supportive of my research and the valuable feedback they provided me along the way. I appreciate your constructive comments during the internal review process of papers published from this work. I am also happy to acknowledge support from Dr. Shaowen Wang to attend the University Consortium for Geographic Information Science (UCGIS) summer school and work with Dr. Yan Liu on continental-scale flood inundation mapping. I appreciate that I received constructive comments and feedback from them regarding my work in Chapter 2. Also, I want to thank Dr. Arezoo RafieeiNasab and Dr. Anthony Castronova for their help and insightful inputs to the work presented in Chapter 3 and Chapter 4 of this dissertation.

I give sincere thanks to the faculty and staff in the Department of Civil and Environmental Engineering and Utah Water Research Laboratory at Utah State University for teaching me, training me, and supporting my research during the last five

years. Thanks for being my extended family over the past years and making my experience in Utah wonderful. Special thanks to Carri Richards for her technical help on the first draft of papers published from this work.

I would also like to express my immense gratitude for the help I received from the folks at the Consortium of Universities for the Advancement of Hydrologic Sciences, Inc. (CUAHSI), National Water Center, and National Center for Atmospheric Research. My work benefited from the generous support of CUAHSI which came in the form of a National Water Center Innovators Program: Summer Institute, Snow Measurement Field School, WRF-Hydro Workshop, and a Master Class on Advance Techniques in Watershed Science. All of these visits not only helped me develop new skills and perspectives to complete this dissertation, but it also provided me the opportunity to connect with the hydrologic community and foster my professional network.

Finally, to my family, relatives, and friends in U.S. and Iran for their cheerfulness when life by choice or by design was difficult. I especially want to thank my parents, Persheng and Jabar, and my sibling, Arvin, from the bottom of my heart for sending their unconditional love, support, and encouragement from about 7000 miles away.

I would like to give my wholehearted thanks to my husband—Mahyar. Mahyar has made considerable sacrifices and supported me in so many ways during the past years. Thanks for being there, and cheers to our new adventures.

Irene Garousi-Nejad

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

In western North America, winter and spring snowpack observations are frequently used to predict summer streamflow, often called water supply forecasting. It is projected that the western U.S. will see an average monthly reduction of about 30% in the extent of the land area within a wintertime snowfall regime, which will affect more than 60 million people who depend on snowmelt in these regions (Bales et al., 2006; Klos et al., 2014). Given the recent snowpack decline due to climate warming in the western U.S. and the uncertain impact on water resources, accurate water supply forecasts will become increasingly crucial as populations grow and demand more water, and as operational agencies have to manage water under global environmental change (Bhatti et al., 2016; Gergel et al., 2017; Li et al., 2017; Livneh & Badger, 2020; Mote, 2003; Mote et al., 2005; Regonda et al., 2005; Stewart et al., 2004, 2005).

Decades of model development, combined with advances in technology and software engineering, have gradually enabled snowmelt runoff models to evolve into continental-scale, high-resolution, and physically-based distributed models such as the National Oceanic and Atmospheric Administration's (NOAA) National Water Model (NWM) in the U.S. This evolution was driven in part by the need to shorten the time interval for streamflow forecasts; to accommodate the shift from simple temperature-index based to energy balance methods; and to enable predicting the effects of anthropogenic and environmental changes such as those caused by land-use change or climate change on large heterogeneous basins (DeWalle & Rango, 2008).

The main goals of the NWM are to provide forecast streamflow, produce spatially continuous, country-wide estimates of hydrologic states (soil moisture, snowpack, etc.), and implement a modeling architecture that permits rapid infusion of new technologies, data, and science. For instance, NWM forecast discharges are used with Height Above Nearest Drainage (HAND), which is calculated from a digital elevation model, to approximate reach-averaged hydraulic properties, estimate a synthetic rating curve, and map near real-time flood inundation from stage in the Continental-Scale Flood Inundation Mapping (CFIM) approach (Y. Liu et al., 2018; Zheng, Maidment, et al., 2018; Zheng, Tarboton, et al., 2018). The CFIM approach avoids the data intensive requirements for river cross-section-based hydraulic modeling; however, it has limitations that affect the accuracy of the estimated flood inundation extent under some conditions. For example, the CFIM approach suffers from the misalignment between the National Hydrography Dataset (NHDPlus medium resolution; McKay et al., 2012) catchments and DEM-derived catchments. DEM errors also manifest as flow paths appearing to go uphill or as artificial barriers in flow paths. Additionally, CFIM considers a fixed roughness parameter in Manning's equation to estimate the flood stage.

The NWM simulates snow processes (accumulation and ablation) using a specific configuration of the Noah Multi-parameterization (Noah-MP) land surface model where snow accumulation/ablation parameterizations are based on mass and energy balance in the snowpack. Noah-MP overcomes some limitations of simple temperature index-based models; however, studies show that Noah-MP has shortcomings attributed to incomplete or incorrect representation of some snow physics that yield inaccurate simulation of snow and ultimately streamflow estimates (Cuntz et al., 2016; Helbig et al., 2015; Magand et

al., 2014; Mendoza et al., 2015; Wang et al., 2019; Wrzesien et al., 2015). These limitations affect the accurate simulation of snow state variables, notably Snow Water Equivalent (SWE) estimates compared to in-situ measurements and satellite observations (Chen et al., 2014; C. Liu et al., 2017; Magand et al., 2014; Swenson & Lawrence, 2012).

Research that evaluates the NWM performance and enhances model output accuracy through more realistic inputs and physics process representations is essential, given that NWM is part of NOAA's water resources information system that provides water forecasts and products to protect life and property. For instance, Viterbo et al. (2020) evaluated the prediction of flooding in NWM streamflow forecasts. They found that errors were due to both meteorological input errors as well as hydrologic process representation. In another study, Lahmers et al. (2019) improved the performance of WRF-Hydro configured as NWM version 1.1 by implementing a conceptual channel infiltration function into the model architecture. They concluded that accounting for channel infiltration loss in the semi-arid western U.S. improves the streamflow behavior simulated when the model is forced with high-resolution precipitation input. However, we are not aware of a systematic and thorough evaluation of the NWM snow outputs.

In spite of existing efforts, further improvements are still necessary to identify causes of discrepancies and explore options for improving representation of physical processes in the NWM that affect the accuracy of water supply forecasts in the western U.S. The NWM represents a progressive leap toward meeting water challenges of the future. Over time, it is expected that the outputs from the NWM will continue to be used in forecasting and will deliver timely forecasts to serve the growing needs of stakeholders and the research community. The NWM is currently undergoing extensive validation and

verification to identify where scientific updates to the model can make the most improvement.

1.2 Objectives and Research Questions

The primary objective of this work was to develop hydrological tools that include modeling of flood inundation mapping and snowmelt contributions to river flow in snowmelt-dominated watersheds. The objectives begin with terrain analysis enhancements developed to improve the terrain-based flood inundation mapping using the NWM forecasts. Then, they progress through the examination of NWM performance in estimation of Snow Water Equivalent (SWE) and snow-covered area fraction (SCAF) to identify the model's limitations in different regions within the mountainous areas in the western U.S. and conclude by proposing improvements to overcome some of the limitations of the NWM in snow estimates. Lastly, they evaluate improvements to SWE modeling by quantifying the improvements when using better model inputs (precipitation and air temperature) and using humidity information in separating precipitation into rain and snow. Each objective is stated below with a list of research questions focused on the challenges faced by engineers and scientists in applying hydrologic modeling for flood and water supply forecasting in mountainous areas that were addressed.

Objective 1: Enhance the Height Above Nearest Drainage Flood Inundation Mapping Method through Terrain Analysis.

This objective follows on from the work of Y. Liu et al (2018), Zheng, Maidment, et al.(2018), and Zheng, Tarboton, et al.(2018) that developed and advanced the HAND approach for use on the U.S. NHDPlus stream network used by NWM. This work was

motivated by discrepancies between the modeled inundation and the inundation mapped from remote sensing indicating that the CFIM method performs poorly and that there are opportunities to improve the accuracy and precision of results. The CFIM approach suffers from misalignment between NHDPlus catchments and DEM-derived catchments. There are also DEM errors that manifest as flow paths appearing to go uphill, or as artificial barriers in flow paths. In addition, CFIM considers a fixed roughness parameter in Manning's equation to estimate the flood stage. Recognizing these limitations, this work explored the following questions:

- How can we improve the CFIM approach to better condition the DEM and ensure that elevation values do not increase moving downstream along hydrographic flow paths to avoid spurious inundation results?
- What are opportunities to enhance the precision with which flood inundation can be mapped using the HAND approach?

Objective 2: Assess the Strengths and Limitations in the NWM Snow Representation against In-Situ Measurements and Remote Sensing Products over the Western United States.

Across the western U.S., snow is observed at 808 snow telemetry (SNOTEL) sites that provide data intended to quantify snow and inform water supply forecasts. Preliminary illustrative comparisons of the NWM retrospective analysis results (NWM-R2) SWE to SNOTEL SWE indicate that SWE is well modeled at some locations while significantly different from observations at other locations. Accurate modeling of SWE is a necessary condition for accurate physically-based modeling of runoff. This motivated the need, addressed in this study, to systematically evaluate the performance of NWM-R2

simulations of SWE and SCAF against available SNOTEL measurements and the moderate resolution imaging spectroradiometer (MODIS) satellite imagery to look for causes of discrepancies between the modeled snow and observation. Also, evaluations of the NWM SWE and SCAF estimates are necessary to address whether the differences between the modeled snow and observations result from input errors (e.g., precipitation) or whether they are due to limitations in the snow parameterization. It is expected that a model operating at a 1 km spatial resolution will not perform optimally in regions where significant spatial variability occurs over tens of meters. However, it has not yet been studied whether the NWM snow estimates are consistently overestimated or underestimated compared to observations across the western U.S. and whether the model is statistically efficient in specific locations or deficient in other regions. This objective answered the following questions:

- How well does the NWM model simulate snowpack (in terms of SWE, SCAF, and snowmelt timing) compared to observations over the entire western U.S.?
- What are the potential causes of discrepancies in NWM-R2 SWE, SCAF, and snowmelt timing?
- Are these discrepancies associated with model input errors, the snow parameterization in the model, or measurement errors?

Objective 3: Evaluate Input Data and Rain Snow Separation Improvements to the National Water Model Simulation of Snow Water Equivalent.

In recent years, climate warming has induced a significant shift in the proportional amount of rainfall versus snowfall across regions reliant on mountain

snowpack. Continuing changes in the precipitation phase (rainfall, snowfall, or a mixture of both) are expected to alter snowpack dynamics, streamflow timing and amount, and frequency of rain-on-snow events; and thus present a new set of challenges for effective hydrologic modeling (Harpold et al., 2017; Musselman et al., 2018). Rainfall and Snowfall Separation (RSS) is one of the most sensitive parameterizations in simulating cold-region hydrological processes (Loth et al., 1993) and has a notable influence on the success of snowmelt models (Rutter et al., 2009). Despite advances in snowmelt modeling, most models, including the NWM's Noah-MP configuration, rely on empirical algorithms based on air temperature to separate precipitation into rain and snow. These ignore the effects of atmospheric humidity on exchanges of latent heat between a hydrometeor and atmosphere (Feiccabrino et al., 2015; Jennings et al., 2018). This has been reported to result in errors in SWE estimates (Harder & Pomeroy, 2014; Wang et al., 2019). While there has been significant prior work on RSS, the objective here was to evaluate the NWM snow model performance across a set of SNOTEL sites that are representative of various hydro-climatological conditions (in terms of rain-on-snow events) across the western U.S., and to indicate where model errors can be removed by using better inputs and a more physically accurate RSS method. Under this objective, the following questions were addressed:

- To what degree are discrepancies in NWM SWE and RSS predictions due to input errors and how much could they potentially be improved if inputs were better?
- How well does the NWM's RSS (rainfall and snowfall separation) parameterization work in comparison to SNOTEL observations?

- Do any other RSS parameterization methods yield more accurate snowfall compared to SNOTEL observations?
- Does incorporating a statistically better RSS scheme into the NWM translate into appreciable improvements in modeling of SWE?
- How do improvements in modeled SWE vary over sites grouped according to the percentage of precipitation events that are rain on snow?

1.3 Chapter Organization

Each of the objectives is addressed within one chapter of this dissertation as follows:

Chapter 2 addresses the first objective. It first presents an evaluation of the CFIM method over a section of the Bear River in Utah, USA, where record flooding due to rapid snowmelt occurred in 2017. The performance of CFIM in terms of its accuracy in representing flooded and non-flooded areas was evaluated by comparing the results with inundation observed by high-resolution Planet RapidEye Satellites. Then, this chapter evaluates a set of improvements developed to overcome some limitations and advance CFIM outcomes. These improvements include: (1) dispersing nodes that subdivide the stream reach into segments approximately uniformly along the reach to avoid the sometimes small and irregular-sized NHDPlus catchments that degrade synthetic rating curve estimation; (2) using high-resolution hydrography (i.e., 1:24,000 scale NHD high-resolution hydrography) to condition the DEM and breach DEM barriers, often due to roads; and (3) using a high-resolution (i.e., 1/9th arc-sec [3 m]) DEM that is available for this area. This work also suggests an approach to obtain a reach-specific Manning's n

from observed inundation and validated improvements for the flood of March 2019 in the Ocheyedan River, Iowa. This chapter shows that methods developed have the potential to improve CFIM.

Chapter 3 addresses the second objective and presents a comparison of the NWM reanalysis snow outputs for water years 2008-2018 to observed SWE and SCAF at SNOTEL sites across the western U.S. SWE was obtained from SNOTEL sites, while SCAF was obtained from MODIS observations at a nominal 500 m grid scale. Retrospective NWM results were at a 1000 m grid scale. This chapter used several metrics to evaluate the model performance. These included seasonal comparison, same-day comparison (modeled and observed SWE and SCAF were compared on the date of observed peak SWE), different-day comparison (modeled and observed peak SWE and SCAF were compared on the separate dates where peak SWE was modeled and observed), full and some snow cover comparisons of snow presence or absence, and comparison of the times of half melt from peak SWE. This work found that the differences between modeled and observed SWE were attributed to both model errors and errors in inputs, notably precipitation and temperature. Also, these differences were regional with generally better SWE and SCAF results in the Central Basin and Range and differences tending to become larger the further away regions are from this region.

Chapter 4 addresses the third objective. It used offline runs of the WRF-Hydro model with the NWM version 2.0 (NWM v2.0) configuration to evaluate the model performance in simulating snowfall and SWE with observations at a set of representative point-scale locations from SNOTEL sites across the western U.S. over 11 water years (2009-2019). It details the selection of the representative SNOTEL sites based on the

long-term averages of the percentage of rain-on-snow (ROS) events. Numerical experiments were conducted to investigate whether and quantify how much NWM snow outputs can be improved by using better inputs (observed precipitation and bias-corrected air temperature) and an RSS method that incorporates humidity information when separating precipitation into rain and snow. Similar to Chapter 3, several metrics were used to quantify the degree of improvement by model scenarios versus SNOTEL observations. These metrics included same-day comparison, different-day comparison, and half melt from peak SWE date comparison. Findings indicated that modifications to the NWM inputs and the Noah-MP's rain snow separation parameterization can improve the general performance of the NWM snow outputs.

1.4 Contribution

This work was driven by the need for better understanding and improving the modeling of flood inundation mapping and snowmelt contributions to streamflow in snowmelt-dominated regions within the western U.S. Results of this dissertation ultimately open the door for the creation of geographically specific, value-added, and tailored solutions that will close the gap between science and the people who need water resources information to protect and support them.

The main contribution of Chapter 2 (Objective 1) is the flow direction conditioning approach, a new DEM terrain analysis method that was developed using high-resolution hydrography data to alter, or condition a DEM, so that elevation values do not increase moving downstream along hydrographic flow paths. This removes artificial barriers in the DEM due to infrastructure such as road crossings, producing an

important improvement in the calculation of HAND and mapping of inundated flooding. High-resolution hydrography mapped at 1:24,000 scale, such as used in this work, is available from the U.S. National Hydrography Dataset for the entire continental U.S., and this approach could be applied similar to Liu et al. (2018) at a continental U.S. scale to compute HAND and associated channel hydraulic properties. While developed and tested for U.S.-based case studies, this approach can be applied globally to improve terrain-based flood forecasting methods where high-resolution hydrography information/data is available.

The contribution of Chapter 3 (Objective 2) that systematically evaluates the performance of the NWM retrospective simulations is that it identifies areas where predictions from the NWM involving snow may be better or worse, and suggests opportunities for research directed towards the NWM model snow component improvements, and ultimately runoff and water supply forecasts in snowmelt-dominated regions. While U.S.-based, the NWM was built using the WRF-Hydro modeling framework that has been applied worldwide, and the lessons learned from this comparison across the U.S. have an application to the representation of snow processes in continental-scale models throughout the world.

The contributions from Chapter 4 (Objective 3) are: (1) the NWM snowfall and SWE can be improved by using more accurate meteorological inputs (particularly precipitation and air temperature) and (2) the parameterization of snow processes can be improved by using a better RSS scheme in the code. While stating that better inputs lead to better model performance is not new, this work quantifies how much that performance improvement could be, which is important in considering where to invest time and effort

in improving the NWM overall. The misrepresentation of precipitation separation into rain and snow within hydrological models leads to cascading effects on hydrological simulations (Harpold et al., 2017), which ultimately propagates into spring snowmelt runoff (Mizukami et al., 2013). Overall, examination of physically-based hydrological models such as NWM through testing different theories, experiments, and modeling parameterizations improves the confidence in predictions in ungauged catchments (Hrachowitz et al., 2013), and this improvement is contingent on how well the model represents physical processes.

Overall, this research that evaluates the NWM performance and enhances model output accuracy through better inputs and physical process representations contributes to the work of the hydrology community that uses NWM outputs, given that the NWM is part of NOAA's water resources information system that provides water forecasts and products to protect life and property. This dissertation also contributes towards conducting reproducible research. Data, computational scripts, and model code developed to produce the study results of this work have been shared in the HydroShare repository.

REFERENCES

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., & Dozier, J. (2006). Mountain hydrology of the western United States: MOUNTAIN HYDROLOGY OF THE WESTERN US. Water Resources Research, 42(8). https://doi.org/10.1029/2005WR004387

Bhatti, A. M., Koike, T., & Shrestha, M. (2016). Climate change impact assessment on mountain snow hydrology by water and energy budget-based distributed hydrological model. Journal of Hydrology, 543, 523–541. https://doi.org/10.1016/j.jhydrol.2016.10.025

Chen, F., Liu, C., Dudhia, J., & Chen, M. (2014). A sensitivity study of high-resolution regional climate simulations to three land surface models over the western United States: SENSITIVITY STUDY OF LSMS IN WRF. Journal of Geophysical Research: Atmospheres, 119(12), 7271–7291. https://doi.org/10.1002/2014JD021827

Cuntz, M., Mai, J., Samaniego, L., Clark, M., Wulfmeyer, V., Branch, O., Attinger, S., & Thober, S. (2016). The impact of standard and hard-coded parameters on the hydrologic fluxes in the Noah-MP land surface model: HARD-CODED PARAMETERS IN NOAH-MP. Journal of Geophysical Research: Atmospheres, 121(18), 10,676-10,700. https://doi.org/10.1002/2016JD025097

DeWalle, D. R., & Rango, A. (2008). Principles of Snow Hydrology. Cambridge University Press. https://doi.org/10.1017/CBO9780511535673

Feiccabrino, J., Graff, W., Lundberg, A., Sandström, N., & Gustafsson, D. (2015). Meteorological Knowledge Useful for the Improvement of Snow Rain Separation in Surface Based Models. Hydrology, 2(4), 266–288. https://doi.org/10.3390/hydrology2040266

Gergel, D. R., Nijssen, B., Abatzoglou, J. T., Lettenmaier, D. P., & Stumbaugh, M. R. (2017). Effects of climate change on snowpack and fire potential in the western USA. Climatic Change, 141(2), 287–299. https://doi.org/10.1007/s10584-017-1899-y

Harder, P., & Pomeroy, J. W. (2014). Hydrological model uncertainty due to precipitation-phase partitioning methods: HYDROLOGIC MODEL UNCERTAINTY OF PRECIPITATION-PHASE METHODS. Hydrological Processes, 28(14), 4311–4327. https://doi.org/10.1002/hyp.10214

Harpold, A. A., Kaplan, M. L., Klos, P. Z., Link, T., McNamara, J. P., Rajagopal, S., Schumer, R., & Steele, C. M. (2017). Rain or snow: Hydrologic processes, observations, prediction, and research needs. Hydrology and Earth System Sciences, 21(1), 1–22. https://doi.org/10.5194/hess-21-1-2017

Helbig, N., van Herwijnen, A., Magnusson, J., & Jonas, T. (2015). Fractional snow-covered area parameterization over complex topography. Hydrology and Earth System Sciences, 19(3), 1339–1351. https://doi.org/10.5194/hess-19-1339-2015

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., … Cudennec, C. (2013). A decade of Predictions in Ungauged Basins (PUB)—A review. Hydrological Sciences Journal, 58(6), 1198–1255. https://doi.org/10.1080/02626667.2013.803183

Jennings, K. S., Winchell, T. S., Livneh, B., & Molotch, N. P. (2018). Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere. Nature Communications, 9(1), 1148. https://doi.org/10.1038/s41467-018-03629-7

Klos, P. Z., Link, T. E., & Abatzoglou, J. T. (2014). Extent of the rain-snow transition zone in the western U.S. under historic and projected climate: Climatic rain-snow transition zone. Geophysical Research Letters, 41(13), 4560–4568. https://doi.org/10.1002/2014GL060500

Lahmers, T. M., Gupta, H., Castro, C. L., Gochis, D. J., Yates, D., Dugger, A., Goodrich, D., & Hazenberg, P. (2019). Enhancing the Structure of the WRF-Hydro Hydrologic Model for Semiarid Environments. Journal of Hydrometeorology, 20(4), 691–714. https://doi.org/10.1175/JHM-D-18-0064.1

Li, D., Wrzesien, M. L., Durand, M., Adam, J., & Lettenmaier, D. P. (2017). How much runoff originates as snow in the western United States, and how will that change in the future?: Western U.S. Snowmelt-Derived Runoff. Geophysical Research Letters, 44(12), 6163–6172. https://doi.org/10.1002/2017GL073551

Liu, C., Ikeda, K., Rasmussen, R., Barlage, M., Newman, A. J., Prein, A. F., Chen, F., Chen, L., Clark, M., Dai, A., Dudhia, J., Eidhammer, T., Gochis, D., Gutmann, E., Kurkute, S., Li, Y., Thompson, G., & Yates, D. (2017). Continental-scale convection-permitting modeling of the current and future climate of North America. Climate Dynamics, 49(1–2), 71–95. https://doi.org/10.1007/s00382-016-3327-9

Liu, Y. Y., Maidment, D. R., Tarboton, D. G., Zheng, X., & Wang, S. (2018). A CyberGIS Integration and Computation Framework for High-Resolution Continental-Scale Flood Inundation Mapping. JAWRA Journal of the American Water Resources Association, 54(4), 770–784. https://doi.org/10.1111/1752-1688.12660

Livneh, B., & Badger, A. M. (2020). Drought less predictable under declining future snowpack. Nature Climate Change, 10(5), 452–458. https://doi.org/10.1038/s41558-020-0754-8

Loth, B., Graf, H.-F., & Oberhuber, J. M. (1993). Snow cover model for global climate simulations. Journal of Geophysical Research, 98(D6), 10451. https://doi.org/10.1029/93JD00324

Magand, C., Ducharne, A., Le Moine, N., & Gascoin, S. (2014). Introducing Hysteresis in Snow Depletion Curves to Improve the Water Budget of a Land Surface Model in an Alpine Catchment. Journal of Hydrometeorology, 15(2), 631–649. https://doi.org/10.1175/JHM-D-13-091.1

McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., & Rea, R. (2012). NHDPlus Version 2: User Guide.

Mendoza, P. A., Clark, M. P., Barlage, M., Rajagopalan, B., Samaniego, L., Abramowitz, G., & Gupta, H. (2015). Are we unnecessarily constraining the agility of complex process-based models? Water Resources Research, 51(1), 716–728. https://doi.org/10.1002/2014WR015820

Mizukami, N., Koren, V., Smith, M., Kingsmill, D., Zhang, Z., Cosgrove, B., & Cui, Z. (2013). The Impact of Precipitation Type Discrimination on Hydrologic Simulation: Rain–Snow Partitioning Derived from HMT-West Radar-Detected Brightband Height versus Surface Temperature Data. Journal of Hydrometeorology, 14(4), 1139–1158. https://doi.org/10.1175/JHM-D-12-035.1

Mote, P. W. (2003). Trends in snow water equivalent in the Pacific Northwest and their climatic causes: TRENDS IN SNOW WATER EQUIVALENT. Geophysical Research Letters, 30(12). https://doi.org/10.1029/2003GL017258

Mote, P. W., Hamlet, A. F., Clark, M. P., & Lettenmaier, D. P. (2005). DECLINING MOUNTAIN SNOWPACK IN WESTERN NORTH AMERICA*. Bulletin of the American Meteorological Society, 86(1), 39–50. https://doi.org/10.1175/BAMS-86-1-39

Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., Barlage, M., & Rasmussen, R. (2018). Projected increases and shifts in rain-on-snow flood risk over western North America. Nature Climate Change, 8(9), 808–812. https://doi.org/10.1038/s41558-018-0236-4

Regonda, S. K., Rajagopalan, B., Clark, M., & Pitlick, J. (2005). Seasonal Cycle Shifts in Hydroclimatology over the Western United States. Journal of Climate, 18(2), 372–384. https://doi.org/10.1175/JCLI-3272.1

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., Barr, A., Bartlett, P., Boone, A., Deng, H., Douville, H., Dutra, E., Elder, K., Ellis, C., Feng, X., Gelfan, A., Goodbody, A., Gusev, Y., Gustafsson, D., … Yamazaki, T. (2009). Evaluation of forest snow processes models (SnowMIP2). Journal of Geophysical Research, 114(D6), D06111. https://doi.org/10.1029/2008JD011063

Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2004). Changes in Snowmelt Runoff Timing in Western North America under a `Business as Usual' Climate Change Scenario. Climatic Change, 62(1–3), 217–232. https://doi.org/10.1023/B:CLIM.0000013702.22656.e8

Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2005). Changes toward Earlier Streamflow Timing across Western North America. Journal of Climate, 18(8), 1136–1155. https://doi.org/10.1175/JCLI3321.1

Swenson, S. C., & Lawrence, D. M. (2012). A new fractional snow-covered area parameterization for the Community Land Model and its effect on the surface energy balance: CLM SNOW COVER FRACTION. Journal of Geophysical Research: Atmospheres, 117(D21), n/a-n/a. https://doi.org/10.1029/2012JD018178

Viterbo, F., Mahoney, K., Read, L., Salas, F., Bates, B., Elliott, J., Cosgrove, B., Dugger, A., Gochis, D., & Cifelli, R. (2020). A Multiscale, Hydrometeorological Forecast Evaluation of National Water Model Forecasts of the May 2018 Ellicott City, Maryland, Flood. Journal of Hydrometeorology, 21(3), 475–499. https://doi.org/10.1175/JHM-D-19-0125.1

Wang, Y., Broxton, P., Fang, Y., Behrangi, A., Barlage, M., Zeng, X., & Niu, G. (2019). A Wet-Bulb Temperature-Based Rain-Snow Partitioning Scheme Improves Snowpack Prediction Over the Drier Western United States. Geophysical Research Letters, 46(23), 13825–13835. https://doi.org/10.1029/2019GL085722

Wrzesien, M. L., Pavelsky, T. M., Kapnick, S. B., Durand, M. T., & Painter, T. H. (2015). Evaluation of snow cover fraction for regional climate simulations in the Sierra Nevada: EVALUATION OF SNOW COVER FOR REGIONAL SIMULATIONS IN THE SIERRA NEVADA. International Journal of Climatology, 35(9), 2472–2484. https://doi.org/10.1002/joc.4136

Zheng, X., Maidment, D. R., Tarboton, D. G., Liu, Y. Y., & Passalacqua, P. (2018). GeoFlood: Large-Scale Flood Inundation Mapping Based on High-Resolution Terrain Analysis. Water Resources Research, 54(12). https://doi.org/10.1029/2018WR023457

Zheng, X., Tarboton, D. G., Maidment, D. R., Liu, Y. Y., & Passalacqua, P. (2018). River Channel Geometry and Rating Curve Estimation Using Height above the Nearest Drainage. JAWRA Journal of the American Water Resources Association, 54(4), 785–806. https://doi.org/10.1111/1752-1688.12661

CHAPTER 2

TERRAIN ANALYSIS ENHANCEMENTS TO THE HEIGHT ABOVE NEAREST DRAINAGE FLOOD INUNDATION MAPPING METHOD[1]

Abstract

Flood inundation remains challenging to map, model, and forecast because it requires detailed representations of hydrologic and hydraulic processes. Recently, Continental-Scale Flood Inundation Mapping (CFIM), an empirical approach with fewer data demands, has been suggested. This approach uses National Water Model forecast discharge with Height Above Nearest Drainage (HAND) calculated from a digital elevation model to approximate reach-averaged hydraulic properties, estimate a synthetic rating curve, and map near real-time flood inundation from stage. In 2017, rapid snowmelt resulted in a record flood on the Bear River in Utah, USA. In this study, we evaluated the CFIM method over the river section where this flooding occurred. We compared modeled flood inundation with the flood inundation observed in high-resolution Planet RapidEye satellite imagery. Differences were attributed to discrepancies between observed and forecast discharges but also notably due to shortcomings in the derivation of HAND from National Elevation Dataset as implemented in CFIM, and possibly due to suboptimal hydraulic roughness parameter. Examining these differences highlights limitations in the HAND terrain analysis methodology. We present a set of improvements developed to overcome some limitations and advance CFIM outcomes. These include conditioning the topography using high-resolution hydrography, dispersing nodes used to subdivide the river into reaches and catchments, and using a high-resolution digital elevation model. We also suggest an approach to obtain a reach specific Manning's n from observed inundation and validated improvements for the flood of March 2019 in the Ocheyedan River, Iowa. The methods developed have the potential to improve CFIM.

[1] Garousi-Nejad, I., Tarboton, D. G., Aboutalebi, M., & Torres-Rua, A. (2019). Terrain analysis enhancements to the Height Above Nearest Drainage flood inundation mapping method. Water Resources Research, 55, 7983– 8009. https://doi.org/10.1029/2019WR024837

2.1 Introduction

Floods are responsible for billions of dollars of damage and loss of life worldwide (Parker, 2017), and much hydrologic research has focused on improving the ability to predict and model flood inundation, prepare for and mitigate flood effects, and warn people at risk. Mapping and modeling flood inundation extent with high precision is challenging because it requires a comprehensive description of computationally demanding and data limited hydrologic and hydraulic processes. Satellite observations of the inundated area, as well as detailed digital elevation model (DEM) topography, offer the opportunity to examine and improve currently available flood inundation mapping methods.

In February 2017, a flood occurred in Box Elder County, Utah, USA. A combination of rainfall and warmer temperatures caused significant snowmelt in northern Utah, resulting in record flows for this time of year in the Bear River. The 2017 discharge was the third largest on record (1952–present) and the largest since 1987. Here we use this event as a case study to evaluate and develop improvements in empirical methods for Continental-Scale Flood Inundation Mapping (CFIM) based on the Height above Nearest Drainage (HAND). Following detailed analysis of this case study, we validated the improvements we developed for another flood that occurred in March 2019 in the Ocheyedan River near Spencer City in Iowa, USA.

In contrast to comprehensive hydraulic models, empirical approaches offer alternatives that have fewer data demands and perhaps offer a more practical alternative for generating flood inundation maps. Researchers such as Rodda (2005) started to incorporate DEM information and grid cell-based operations to calculate flood depth and

then the probability of insured losses from floods. Rennó et al. (2008) redesigned Rodda's (2005) concept and introduced a terrain descriptor called HAND, which uses DEMs to define the height of each grid cell on the land surface above the cell in the nearest stream to which the drainage from that land surface cell flows. Thereafter, researchers applied HAND as a descriptor to determine soil water potential (Nobre et al., 2011), groundwater potential (Rahmati et al., 2018), and flood potential (Nobre et al., 2016).

To calculate HAND, a hydrologically conditioned DEM and a representation of the DEM flow field are required. A hydrologically conditioned DEM is one for which internally draining areas have been removed (or true internally draining areas are marked and retained), and each grid cell can drain following a nonincreasing elevation path to the edge of the DEM (or true internally draining sink). Pit filling (Jenson & Domingue, 1988), breaching (Soille et al., 2003), and hybrid filling-breaching algorithms to hydrologically condition DEMs have been developed (Lindsay & Creed, 2005; Martz & Garbrecht, 1999; Soille, 2004). Also, recent work has advanced hydrography-driven coarsening to retain hydrographic fidelity in a high-resolution DEM when the DEM needs to be reduced for computational reasons (Moretti & Orlandini, 2018).

To represent the flow field, the earliest method is the D8 single-flow direction model initially proposed by O'Callaghan and Mark (1984). In this method, flow moves from each grid cell to one of its eight neighbors along the steepest downward slope direction. This method limits the precision with which flow direction is represented and introduces grid bias (Costa-Cabral & Burges, 1994; Fairfield & Leymarie, 1991; Tarboton, 1997).

Considerable work has been done to overcome D8 limitations (Freeman, 1991; Orlandini et al., 2003, 2014; Orlandini & Moretti, 2009; Quinn et al., 1991; Seibert & McGlynn, 2007; Tarboton, 1997). Among these alternative approaches is the D∞ (D-infinity) multiple-flow direction model developed by Tarboton (1997). The D∞ model shares the flow from a grid cell between two adjacent down slope grid cells based on flow direction angle proportioning.

Taking advantage of the D∞ flow model, Tesfa et al. (2011) presented an algorithm that uses DEMs to derive a set of hydrological proximity measures that include distances up from or down to target grid cells. These distances may be evaluated horizontally or vertically using the D∞ flow model with weighted averaging applied where flow is proportioned between multiple grid cells. The D∞ distance down function, calculated vertically with the target grid cells being stream grid cells, may be used to calculate HAND (Zheng, Maidment, et al., 2018). This D∞ distance down evaluation of HAND differs from earlier HAND approaches (Nobre et al., 2011, 2016) that use D8. The D∞ averaging across multiple-flow paths results in a smooth HAND surface and avoids step-like discontinuities that can result from using D8. Tesfa et al. (2011) described an efficient parallel implementation of these calculations that is available as part of the TauDEM software (Tarboton, 2017).

Liu et al. (2016) implemented the D∞ distance down function to generate a HAND raster for the continental United States using a high-performance computer at the University of Illinois CyberGIS facility. Given this HAND raster, Zheng, Tarboton, et al. (2018) proposed a method to compute the reach-averaged channel geometry properties and estimate a synthetic rating curve to relate flow to water level in a stream reach. The

Zheng, Tarboton, et al. (2018) method is an alternative to traditional approaches, where river geometry properties are computed using surveyed river cross sections at points along the stream reach that omit terrain detail between cross sections. This approach also avoids the intensive data requirements for river cross-section-based hydraulic modeling. At the same time, Liu et al. (2018) designed a workflow based on Liu et al. (2016) and Zheng, Tarboton, et al. (2018) for CFIM using discharge forecasts from the National Oceanic and Atmospheric Administration National Water Model (NWM). In this method, NWM discharges are converted to stage using a rating curve, and then the stage is used to map inundation based on HAND. Each reach in the National Hydrography Dataset (NHDPlus medium resolution; McKay et al., 2012) has a separate forecast discharge, and HAND is used to evaluate rating curves and map inundation for the NHDPlus catchment draining into each reach. In another recent study, Zheng, Maidment, et al. (2018) implemented the HAND approach with high-resolution topographic data derived from light detection and ranging using a geodesic minimization technique to map the streams (Passalacqua et al., 2010; Sangireddy et al., 2016). This approach maps streams by selecting a flow path to minimize a metric that is a combination of contributing area, curvature, and distance from NHDPlus (medium resolution) streams. The results showed that the inundation extent produced by their approach, called GeoFlood, is able to capture the Federal Emergency Management Agency flood plain coverage with 60–90% overlap accuracy. Given the potential for using the HAND method to map flood inundation over large areas based on available DEM data and discharge forecasts from the NWM, it is important to evaluate the performance of the approach in many different locations and settings.

Satellite-based remote sensing is a useful source for evaluating modeled flood inundation mapping because of its ability to provide a synoptic perspective across a large range of scales and resolutions. Remote sensing offers a practical solution to observing the location and extent of inundation for many flooding events (Policelli et al., 2017). Recent deployments of small CubeSat satellites by companies such as Planet (2017) advance hydrological remote sensing by providing an unprecedented combination of high-temporal- and high-spatial-resolution imagery at the global scale (Cooley et al., 2017). Remote sensing for river discharge is a topic receiving increasing attention, for example, Surface Water and Ocean Topography (SWOT) satellite mission, https://swot.jpl.nasa.gov/hydrology.htm (Biancamaria et al., 2010; Tourian et al., 2017). In particular, Sichangi et al. (2016) used Moderate Resolution Imaging Spectroradiometer (MODIS) and satellite altimetry to estimate river discharge and also suggested its use in rating curves and hydraulic parameter optimization. Overall, there is broad potential for applying remote sensing of surface properties, including inundation mapping and monitoring in hydrological research (Cooley et al., 2017; McCabe et al., 2017).

In this study, we evaluate the performance of CFIM in terms of its accuracy in representing flooded and non-flooded areas when comparing the results with inundation observed by high-resolution Planet RapidEye Satellites. We first use the HAND data and rating curves from Liu et al. (2018) to compute flood inundation based on a published methodology (Liu et al., 2018; Zheng, Tarboton, et al., 2018). Then, we examine the causes of discrepancies between mapped and observed flood inundation and develop and evaluate important improvements to the published CFIM method. These improvements include (1) dispersing nodes approximately uniformly along the reach to avoid the

sometimes small and irregular-sized NHDPlus catchments that degrade synthetic rating curve estimation; (2) using a high-resolution hydrography (NHD high-resolution, i.e., 1:24,000) to condition the DEM and breach DEM barriers, often due to roads; and (3) using a high-resolution (i.e., 1/9th arc-sec [3 m]) DEM that is available for this area. We apply these sequentially, and at each step quantify the improvement in flood inundation mapping fidelity. We also evaluate the stage that best matches HAND-based flood inundation with observed inundation and use this to quantify inherent uncertainty in this evaluation and infer reach specific information about Manning's n that may minimize errors introduced with this approach. This opens an opportunity for direct estimation of reach specific Manning's n from observed inundation, overcoming the current CFIM limitations of using a single Manning's n everywhere. This builds on ideas from Sichangi et al. (2016).

The following section describes the study site and the February 2017 flood that occurred in the Bear River, Utah, USA. The methodology section describes the CFIM procedure for generating a flood inundation map based on HAND and introduces improvements to address the identified CFIM challenges. Next, we present the workflow for creating a flood inundation map based on RapidEye imagery, along with metrics to compare HAND-based flood inundation maps with high-resolution satellite imagery. The results section compares HAND-based flood inundation from CFIM with RapidEye observations and then sequentially quantifies the improvement due to each innovation introduced. Thereafter, results for a validation case study (a March 2019 flood that occurred in the Ocheyedan River, Iowa) based on both the published CFIM method and developed improvements are presented and compared with the observed flood

inundation. We conclude with a discussion of the advantages and limitations of the new developments presented and ideas for future work.

2.2 Study Site and Data

2.2.1 Topography and Hydrography

Box Elder County is a mountainous area in northern Utah, USA. Flooding that occurred in the area in 2017 had a marked influence on people and properties. Figure 2.1a illustrates the topography of a region located within Box Elder County, Utah, that includes the Bear River. The contour lines represent the 10-m DEM available from the National Elevation Dataset (NED), from which Liu et al.'s (2018) HAND data were derived. The raster layer, also from NED, shows the available 3-m DEM covering the Bear River. The dataset includes DEMS of different resolution, with 10-m resolution data available over the entire continental United States, whereas 3-m data are available only for specific areas. The 3-m domain covers a part of the Bear River between two stream gages: the upstream gage operated by PacifiCorp at Collinston and the lower gage operated by the U.S. Geological Survey (USGS)–station 10126000 Bear River near Corrine. We focused on the reach of the Bear River between these gages. The Malad River (Figure 2.1b) enters the study reach close to its downstream (southern) end. Three main highways cross the Bear River within the domain. This is of note because the road-top elevations recorded in the NED DEM that result in artifacts in Liu et al.'s (2018) published HAND layer have been corrected by the high-resolution hydrography breaching procedure applied here.

According to the National Hydrography Dataset (NHDPlus), the contributing area at the PacifiCorp gage at the upstream end of the Bear River reach is 15,545 km^2. The contributing area at the Corrine USGS gage at the downstream end is 17,868 km^2. The contributing area of the Bear River at its junction with the Malad River is 15,728 km^2, and the contributing area of the Malad River is 2,110 km^2. Therefore, the area draining directly and from other tributaries into the Bear River reach is 213 km^2 or about 10 times less than the Malad River.

2.2.2 Historical and Observed Streamflow

Historical peak discharges (Figure 2.2a) indicate that the February 2017 discharge was the third largest on record (1952–present) and the largest since two prior floods in 1984 and 1986 at the USGS Corrine gage. Discharges observed at each gage (the PacifiCorp and USGS gages) for the flood of 2017 are shown in Figures 2.2b and 2.2c, and the date (15 February) when Planet RapidEye satellite imagery is available is indicated as well. The daily average discharges on 15 February 2017 were 224.33 and 251.77 m^3/s observed at the PacifiCorp (http://bearriverbasin.org/rivers/rivers/) and the USGS gages (https://waterdata.usgs.gov/nwis/uv?site_no=10126000), respectively. Note that while the USGS gage recorded 15-min discharges, only daily discharges were available at the PacifiCorp gage. However, this event was gradual enough that daily discharges are sufficient for this analysis. The instantaneous discharge at the USGS gage at noon on 15 February was 250.6 m^3/s, essentially identical to the daily average.

2.2.3 Planet RapidEye Satellite Imagery

A Planet RapidEye image from 15 February 2017 was selected as being closest in time to the peak (Figure 2.3a). The RapidEye satellite constellation consists of five

satellites collecting remotely sensed information within five spectral bands (blue: 440-510 nm; green: 520-590 nm; red: 630-685 nm; red edge: 690-730 nm; and near infrared: 760-850 nm) with the daily off-nadir and 5.5-day at-nadir revisit time. The specific image selected was a RapidEye Analytic Ortho Tile at a spatial resolution of 5 meters captured on 15 February 2017 at 18:42 UTC (i.e., 11:42 a.m. MST). This tile included 100% coverage of the study area and cloud cover of less than 1%. The RapidEye Analytic Ortho Tile product is geometrically corrected by the Planet (2017) team. These corrections remove distortions due to image perspective (tilt) and relief (terrain) prior to it being made available to research users.

We present a zoomed-in plot of the yellow box (Figure 2.3b) to show a portion of the flooded area (mostly around the river corridor) and the corresponding high-resolution 30-cm world imagery (Figure 2.3c) from ESRI (2018) as a reference image to indicate the general land cover of this area.

2.2.4 National Water Model Analysis and Assimilation Data

The Current CFIM approach implements the Zheng, Tarboton, et al. (2018) and Liu et al. (2018) HAND method using the NWM flows for each NHDPlus reach. The NWM is a hydrologic model that simulates the water cycle over the entire continental United States (http://water.noaa.gov/about/nwm). The model produces an analysis and assimilation discharge that represents a snapshot of current hydrologic conditions, in addition to short, medium, and long-range forecasts. Since we were working with a historical flood, we used the NWM assimilated flow at 18:00 UTC on 15 February 2017 (Figure 2.4) to create flood inundation maps representative of the current CFIM method.

The Bear River has 32 NHDPlus reaches between the two gages in the study site. Discharge is shown for a selection of reaches to illustrate the modeled increase along the reach. The upstream reach (close to the PacifiCorp gage) had a daily average discharge of 146.8 m^3/s on 15 February 2017 from the assimilated NWM flows, whereas the daily average observed value for the same date at the gage was approximately 224.3 m^3/s. Subtracting the upstream flow (146.8 m^3/s) from the flow just prior to the Malad River junction (183.4 m^3/s), we infer that model tributary inflows are about 36.7 m^3/s along the study reach upstream of the Malad River junction. The NWM Malad River inflow was 99.0 m^3/s, consistent with the increment in the Bear River flow across this junction. However, a discontinuity exists in the reported assimilated NWM discharges upstream and downstream of the lower stream gage. These discharges differ substantially from observed discharges. Notably, the NWM discharge at the upper end of the study reach (daily average of 146.8 m^3/s) is about 35% less than daily average observed by PacifiCorp (224.33 m^3/s). At the downstream end, the discharges are closer (daily average of 251.77 m^3/s observed and daily average of 285.8 m^3/s NWM). These discrepancies may affect CFIM results.

2.3 Methodology

This section describes (1) the workflow to generate a flood inundation map based on HAND, along with the proposed improvements; (2) the flow direction conditioning method; (3) the workflow used to map observed flood inundation based on RapidEye imagery using a supervised classification method; (4) evaluation metrics used to compare HAND-based flood inundation mapping with high-resolution RapidEye imagery; and (5)

the improvements to flood inundation mapping developed and evaluated as part of this effort.

2.3.1 HAND-based Flood Inundation Mapping and Proposed Improvements

The HAND raster is a drainage-normalized and flow path-coherent version of a DEM (Nobre et al., 2016). Generating a HAND-based flood inundation map primarily consists of three procedures:

1. Calculate HAND.
2. Estimate reach-averaged hydraulic properties and synthetic rating curve from HAND for each stream reach.
3. Given discharge and the rating curve, calculate stage and map inundation at locations where HAND is less than the calculated stage.

According to the CFIM described by Liu et al. (2018), calculating HAND starts with hydrologically conditioning the DEM by pit filling. Then D8 flow directions are computed. A raster representation of the stream network is derived using head points of streams mapped in the NHDPlus medium-resolution dataset as stream starting points. This is done using a weighted D8 flow accumulation calculation with weights taken as 1 at stream heads, and then applying a threshold to the result. This produces a stream raster, which is effectively the grid cells along downslope flow paths traced along flow directions from each starting point. This procedure maps streams at a drainage density close to that of the NHDPlus stream network, but it differs from NHDPlus in that streams are located along DEM elevation valleys, which resolves discrepancies that arise due to misalignment between the DEM and cartographically mapped NHDPlus streams. Finally, the TauDEM D∞ distance down function is used with the hydrologically conditioned (pit

filled) DEM and D∞ flow directions and the raster representation of the drainage network to generate the HAND raster. For details see Zheng, Tarboton, et al. (2018) and Liu et al. (2018).

The second procedure estimates hydraulic properties from the HAND map as described by Zheng, Tarboton, et al. (2018). The procedure uses the HAND raster map to extract/estimate the river geometry (i.e., surface area, wetted bed area, and volume of flood) and then compute the reach-averaged hydraulic properties (i.e., cross section area, wetted perimeter, top width, and hydraulic radius) for each reach. This procedure operates over the local catchments, defined as the areas draining directly to each stream reach. Zheng, Tarboton, et al. (2018) and Liu et al. (2018) used NHDPlus catchments. Here, in addition to NHDPlus catchments, we evaluated using catchments derived directly from the higher resolution DEMs (10 and 3 m) using nodes dispersed approximately uniformly along the study reach. NHDPlus catchments were derived from a 30-m DEM and may not align well with 10- and 3-m DEM-derived catchments in locations with complex topography. Hydraulic properties are calculated for a series of water depths or stages h, ranging from 0 to greater than the maximum water depth reasonably possible. For each catchment and water depth h, we first identify the cells where the HAND value H is the less than h. Then, the river geometry and the reach-averaged hydraulic properties are computed, and Manning's equation (equation (1)) is used to calculate the discharge associated with each depth value h to establish the synthetic rating curve for each reach.

$$Q = \frac{1}{n} A(h) R(h)^{2/3} S^{1/2} \quad (1)$$

where Q is discharge, n is the hydraulic roughness parameter in Manning's equation, A(h) is the cross-section area of a channel as a function of stage h, R(h) is the hydraulic radius as a function of stage h, and S is the slope of the channel.

The final (third) inundation mapping procedure uses discharge as input and the synthetic rating curve to determine stage h*. Then, inundated extent is mapped as those grid cells for which h* > H. Each inundated cell is classified as 1 (flooded), and the rest of the cells are classified as 0 (non-flooded). This is accomplished separately for the catchment associated with each reach.

Our application of this approach for the Bear River study reach identified a number of limitations where improvements are possible:

1. The CFIM used by Liu et al. (2018) utilizes NHDPlus medium-resolution catchments in the evaluation of HAND hydraulic properties for the associated NHDPlus stream. Problems arise where these catchments are small or variable in size, sometimes due to the presence of canals/ditches in the NHDPlus dataset. Other discrepancies are due to the NHDPlus streams not being well aligned with the DEM used in the HAND calculation. This is because NHDPlus used a 30-m DEM, while HAND used a 10-m DEM to capture additional detail. This can lead to discrepancies in hydraulic properties calculated for NHDPlus reaches, which propagate into the inundation mapping.
2. The pit filling method used in the CFIM can result in large flat areas behind barriers in the DEM, where the DEM represents the top surfaces of bridge crossings.

3. The results, especially for hydraulic properties, are only as good as the DEM used. The CFIM approach is based on the 1/3rd arc-sec (10-m) DEM. In some places a 1/9th arc-sec (3-m) DEM is available and may provide an opportunity for improvement.

Here we introduce enhancements to the HAND-based flood inundation mapping approach to address each of the noted limitations and improve the accuracy of the flood inundation mapping result while relying on the concepts introduced by Zheng, Tarboton, et al. (2018) and Liu et al. (2018). The improvements, which explicitly aim to reduce the occurrence of unrealistic HAND values and diminish the impact of DEM errors on flood inundation mapping, are as follows:

1. Remove the canal/ditch features from the NHDPlus medium-resolution flow lines prior to the preparation of HAND inputs because canal/ditches are generally not evident in the DEM.
2. Derive hydraulic properties and synthetic rating curves using the DEM-based drainage network and catchments based on evenly spaced nodes along a stream reach to avoid inconsistencies due to the size variability of the catchments. One challenge is that the slope can be zero for reaches that are across flat areas, and these become artifacts in the DEM. In these cases, we adjusted the local elevations of the junctions in the digital representation of the stream network to shift elevation changes between stream reaches and impose a nonzero slope on each stream reach. For example, when the reach downstream of a junction has a positive slope, but the reach upstream is flat; the elevation of the junction can be lowered to make these reaches have an

equal but smaller slope than the downstream reach did initially. We extended this idea both upstream and downstream to account for the occurrence of adjacent reaches with zero slope. The result was a set of reach slopes that are all positive but do not alter the overall elevation differences and hence slopes in the stream network.

3. Incorporate a hybrid filling-breaching algorithm to hydrologically condition the DEM consistent with a high-resolution hydrography dataset (such as NHD high-resolution). We developed a new flow direction conditioning approach for this purpose.
4. Use a higher-resolution DEM (i.e., 1/9 arc-sec or 3-m resolution) to enhance accuracy, along with the enhancements above, which are using evenly sized stream reach and high-resolution hydrography flow direction conditioning etching approach.

The next section details the flow direction conditioning approach.

2.3.2 Flow Direction Conditioning

The approach for flow direction conditioning is to first determine the set of flow directions, represented as a grid with the same dimensions as the DEM, that follows the downward flow direction of the given hydrography dataset, and then to adjust the DEM so that elevations, moving downslope along these flow directions, are nonincreasing. The flow directions were determined using the sequence of TauDEM and GIS grid manipulations described below, while a new TauDEM tool was developed to adjust (flow direction condition) the DEM. To determine flow directions along streams, the DEM was modified by lowering elevation values significantly for grid cells on streams, a process

sometimes referred to as burning, after which pits are filled and flow directions computed for this DEM. The burned-in streams manifest as canyons, and flow directions are constrained to remain within the canyon (stream) until they exit the domain. These on-stream flow directions thus follow the path of the rasterized hydrography dataset. Off stream flow directions are set aside, and only the on-stream flow directions are used in the new flow direction conditioning tool. The detailed steps used to determine along stream flow directions are as follows:

1. Designate the input DEM as Z.
2. Convert the hydrography stream lines to a raster that has the same dimensions (columns, rows, cell size, and edge coordinates) as the DEM denotes srfv (stream raster from vector) with values 1 on stream and 0 off stream. The stream vector dataset used here should not include streams that enter the domain (extent of the grid) from outside and must include streams where they leave the domain.
3. Burn srfv into Z using the cell-by-cell grid calculation Zb = Z – B * srfv, where B is a big number set as described below. The resulting grid is Zb, a DEM with deeply burned canyons along the rasterized streams. The ArcGIS raster calculator tool is used here. The big number B is somewhat arbitrary and should be big enough to burn the stream hydrography into the DEM to a sufficient depth that when pits are filled and flow directions determined, the only path available for flow is along the rasterized hydrography. One way to decide on B is to fill pits and then subtract the original DEM and use a number that is larger than the maximum difference but does not produce numerical

overflow. Our application with high-resolution NHD streams has different classes of flowlines, specifically StreamRiver, ArtificialPath, CanalDitch, Pipeline, Connector, Coastline, and Underground Conduit. We excluded all except StreamRiver, ArtificialPath, and Pipeline/Connector and used B=1,000 m for StreamRiver, 700 m for ArtificialPath, and 400 m for Pipeline/Connector, which gives preference to flow along the StreamRiver line that we took to be the main flow path in braided situations. Then, an ArtificialPath is prioritized over a pipeline/connector.

4. Fill pits in burned DEM (Zb) using the TauDEM Pitremove function. The result is Zbfel, which is now hydrologically conditioned but with deeply burned canyons along the streams. To avoid completely filling in the deeply burned canyons during pit filling, they should extend only to the edge of the DEM where flow exits the domain. This is why the stream vector dataset should not include streams that enter the domain and must include streams that leave the domain.
5. Calculate D8 flow directions using the TauDEM D8FlowDir function with input Zbfel. The output flow directions, which are designated as raster p, are constrained by the burning to be within the burned canyons along the streams. Because the stream vector data set is required to include streams leaving the domain and not streams entering the domain, the conditioned DEM from (4) and flow directions will define paths that follow downslope within the stream raster to the edge of the domain where streams leave.

6. Mask the D8 flow directions to have only the flow directions on streams, setting all other flow directions to no data. The raster calculation pm = p/srvf achieves this.

The result from this process is a grid with D8 flow direction values set along the streams for the grid cells intersecting the input hydrography dataset. This is used as input to the flow direction conditioning tool. The significantly altered DEMs (Zb and Zbfel) are not used further, and neither are the flow directions that were computed for parts of the domain outside streams.

The second step in flow direction conditioning is to adjust DEM elevations so they are strictly nonincreasing in the downstream direction along the flow directions generated from the first step above. A new TauDEM function, “flow direction conditioning,” was written to achieve this (Algorithm). This takes as input the original DEM, Z, and conditioned flow direction raster, pm, and produces as output a conditioned DEM, Zc.

Algorithm 1. Flow Direction Conditioning. Z denotes the input DEM and pm denotes the input flow direction grid. N is a grid of dependencies initialized to zero and used to count the number of grid cells that drain into any grid cell whose elevation has not been conditioned. Q is an initially empty queue

```
// Initialize Dependencies
for each grid cell x in pm:
        if pm(x) is not no data:
        for each neighbor y that drains to x based on pm(y):
                N(x)= N(x)+1
        end for
        if N(x)=0:
                add x to Que Q
        end if
end if
end for
//At this point Q holds grid cells that are channel heads and do not
have any upstream grid cells flowing into them. They are thus ready
for evaluation with N = 0.  N holds the number of grid cells draining
into each grid cell, or put another way, the number of adjacent grid
cells that have to be solved before the grid cell can be solved.

//Condition Elevation Values to be non-increasing along flow
directions
for each grid cell x on Q
for each neighbor y that drains to x based on pm(y):
        if Z (x) > Z (y):
                set Z (x) = Z (y)   // This lowers Z to the value of an
                upstream grid cell if necessary, ensuring that drainage
                can occur along the flow directions pm
        end if
        for neighbor k downstream from x based on pm (x):
                N(k)= N(k)-1
                if N(k) = 0  // This indicates that all neighbors have
                been evaluated and k is read for evaluation
                        Add k to Q
                end if
        end for
        end for
end for
// Here the Q is empty and all cells needing adjustment have been
adjusted.
Write Z as new conditioned DEM Zc
End
```

The key part of Algorithm is the two lines in bold above evaluating the following logic.

$$Z(x) = \min_{\text{y drains to x}} \big(Z(x), Z(y)\big) \quad (2)$$

This recursively evaluates Z to the least of the incoming neighbor Z values and the cell value itself. The result is a grid of elevation values Z that are conditioned to drain along the flow directions given.

2.3.3 RapidEye-Based Flood Inundation Mapping

Satellite-based flood inundation mapping is generally performed using water detection algorithms. Water in satellite images is detected using three methods: (1) single-band, (2) multiband, and (3) classification. The single-band approach involves choosing one characteristic band from a multispectral image, a band for which the spectral signature of water is unique and representative. Then, to discriminate water from other surfaces, a threshold, often derived from the histogram analysis of the image for the characteristic band, needs to be defined. Separating surface water from other land types based on a single threshold in a single unique band is frequently problematic (Verpoorter et al., 2012). Identification of surface water can be improved using multi-band methods where a combination of different bands is used through a so-called spectral index. Multiband methods also require definition of a threshold for the selected spectral index to determine whether a pixel value is categorized as water or not. The subjective selection of the threshold may lead to an overestimation or underestimation of surface water (Xu, 2006).

Classification methods are frequently applied for classifying surface water in images. The classification approaches can be categorized into two main groups:

supervised and unsupervised. The major task of the supervised methods is to segregate the spectral domain into classes (different land covers) according to their spectral similarities. Unsupervised methods assign a class to each pixel without any prior knowledge of the names (types) of those classes. Unsupervised techniques derive their result using the statistical properties of the data. In other words, unsupervised techniques group pixels if they have similar statistical properties. In cases where it is possible to obtain a labeled dataset (with class names such as water class) for training the classification algorithm, supervised classification is suggested as it significantly outperforms unsupervised classification (Laskov et al., 2005). Ireland et al. (2015) provides an example of the successful application of supervised classification of flooded areas from Landsat imagery.

In this study, we used a supervised classification method within the ArcGIS Pro software from ESRI (2018) on a RapidEye image for 15 February 2017 (which is 1 day after the first peak of the flood). Visually examining the true color image, we developed a training sample that included areas within open water and non-water areas. The RapidEye image was captured in wintertime (February 2017), when it is sometimes difficult to discriminate between water and shadow pixels visually. To mitigate this difficulty and increase the accuracy of the training sample, we used a normalized difference vegetation index (NDVI) map computed using equation (3), recognizing that negative NDVI values theoretically correspond to water.

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{RED}}{\mathrm{NIR} + \mathrm{RED}} = \frac{\mathrm{Band5}_{\mathrm{RapidEye}} - \mathrm{Band3}_{\mathrm{RapidEye}}}{\mathrm{Band5}_{\mathrm{RapidEye}} + \mathrm{Band3}_{\mathrm{RapidEye}}} \quad (3)$$

Within the ArcGIS Pro Classification Wizard, the Classification Method was set to supervised classification, the Classification Type to Pixel based, and the schema to default. Next, through the Training Samples Manager, we selected the training samples and chose Maximum Likelihood as the classifier. The spatial resolution of the classified image was the same as the RapidEye image (i.e., 5 m). Since the resolution of the HAND-based flood inundation maps in the present study was 10 or 3 m, we used nearest neighbor re-sampling in ArcGIS Pro to obtain both 10- and 3-mclassified maps, which were then used for calculating the evaluation metrics as described in the following section.

2.3.4 Evaluation Metrics

In general, evaluation metrics are used to validate the results of a model against observations to measure how well the model performs. In this study, we used Correctness (C; equation (4)) and Fit (F; equation (5)) to indicate the degree-of-overlap between model and observed flood inundation maps (Horritt & Bates, 2002; Merwade et al., 2018; Sangwan & Merwade, 2015).

$$C = \frac{\text{Model}_{\text{wet}}}{\text{Observed}_{\text{wet}}} \tag{4}$$

$$F = \frac{\text{Model}_{\text{wet}} \cap \text{Observed}_{\text{wet}}}{\text{Model}_{\text{wet}} \cup \text{Observed}_{\text{wet}}} \tag{5}$$

Both statistics should ideally be 1 (100%). C is an overall area metric and F is a location-specific metric. C, the correctness metric, quantifies the degree to which the total modeled and observed areas classified as inundated (wet) match. F is a stricter statistic that quantifies whether modeled and observed locations match (i.e., intersection), scaled by the total area mapped as inundated by either (i.e., union; Sangwan & Merwade, 2015).

Here we took the $Model_{wet}$ as $HAND_{wet}$, the grid cells inundated in a HAND-based flood inundation map (such as CFIM), and $Observed_{wet}$ as $RapidEye_{wet}$, the grid cells classified as water (inundated) from the RapidEye imagery.

2.3.5 Improvements Developed

We developed HAND-based flood inundation maps for each of the following scenarios to evaluate the HAND approach and suggested terrain-processing improvements through comparison to the classified RapidEye image. The first scenario is the current CFIM base case used as a starting point for evaluating improvements.

1. Model inundation based on publicly available CFIM information, where the hourly-assimilated NWM flows on 15 February 2017 at 18:00 UTC are used. The HAND raster, hydraulic properties, and rating curves were obtained from the NFIE Continental Flood Inundation Mapping data repository (https://web.corral.tacc.utexas.edu/nfiedata), and the codes that we used are the ones available on GitHub (https://github.com/cybergis/nfie-floodmap). This serves as an indicator for how well the current CFIM approach may be expected to perform.
2. Model inundation based on HAND, hydraulic properties, and rating curves from scenario (1) but with observed discharges in the main river (i.e., Bear River) and negligible discharges from side tributaries except for the Malad River (because the drainage area of other tributaries is much smaller than the Malad River). We assumed a flow of 224.33 m^3/s in the Bear River reach from the upper PacifiCorp Collinston gage to the Malad junction, with a flow of 251.77 m^3/s in the Bear River reach downstream of the Malad junction.

This is our best estimate of what actual discharges were on 15 February 2017, and it serves to separate the effect of discharge errors from DEM and HAND and rating curve errors.

3. We delineated only the main Bear River and Malad tributary channels from the DEM, initiating them at channel heads near where the NHDPlus medium-resolution (i.e., 1:100,000) channel network enters the domain. We distributed nodes evenly along the main river to delineate catchments from the 10-m DEM for use in the HAND process. This results in streams comparable to CFIM but more consistent catchments, and hence, more consistent channel properties than used in CFIM where NHDPlus medium resolution catchments are based on the 30-m DEM and are quite variable in size.
4. We etched the high-resolution hydrography (NHD high-resolution, i.e., 1:24,000) into the 10-m DEM using flow direction conditioning (section 3.2). This removes DEM barriers often due to roads. Catchments were delineated using evenly distributed nodes. This scenario allows us to evaluate the flow direction conditioning approach at the same DEM resolution as operational CFIM.
5. We repeated the procedure from scenario (3) but with the high resolution (1/9th arc-sec, 3 m) DEM that is available for this area. This scenario allows us to evaluate the potential benefit from higher resolution DEM data.
6. We etched the high-resolution hydrography (NHD high-resolution, i.e., 1:24,000) into the 3-m DEM using flow direction conditioning. This repeated the procedure from scenario (4) but with the high-resolution (1/9th arc-sec, 3-

m) DEM that is available for this area and allowed us to evaluate the potential benefit from higher resolution DEM data with flow direction conditioning.

2.4 Results

2.4.1 Conditioned Topography Through Flow Direction Conditioning Approach and Its Effect on HAND

To illustrate the effect of the flow direction conditioning (etching) on the DEM, we zoomed in to locations that illustrate the effect well (Figure 2.5). The flow direction conditioning approach etches the path of the high-resolution NHD flowline into the 10-m DEM (Figures 2.5a-2.5c). The DEM etching that is apparent here was negligible or not discernible over much of the 10-m DEM area, indicating that the etching only affects the DEM at places where barriers exist. Road barriers seemed more prevalent in the 3-m DEM, and etching provides a way to resolve (punch through) these barriers (Figures 2.5d-2.5f).

We created HAND rasters for each DEM (Figure 2.6). The results show that the etching method affects the HAND raster for both selected areas. Without etching, a considerable part of the streambed is flat due to pit filling, resulting in a HAND map with values close to 0 for most of the river corridor (Figures 2.6a and 2.6c). This is of concern because it impacts extraction of river hydraulic geometry and the synthetic rating curve from the HAND map such that for the same flowrate, the water depth estimated from a HAND-derived synthetic rating curve might be unrealistic.

2.4.2 Flood Inundation

The length of the study reach is almost 60 km, with 32 NHDPlus reaches and corresponding catchments along the main stem, although some of these are very small. Other studies (Godbout, 2018; Zheng, Maidment, et al., 2018) have suggested that reach lengths of 1.5 to 5 km may be optimal for use with the HAND approach. We placed 20 nodes approximately evenly along the main stem, including at the downstream end, upstream end, and at the Malad River junction. This number of nodes was chosen to obtain stream reaches about 3 km long, within the range suggested by Godbout (2018) and Zheng, Maidment, et al. (2018). This resulted in 18 catchments upstream of the Malad River junction and one catchment from the Malad River junction to the downstream gage being delineated, each draining to a reach about 3 km long (Figure 2.7).

Flood inundation maps for each scenario that we evaluated are shown in Figure 2.8, and comparisons with classified inundation are detailed in Figure 2.9. We compared observed flood inundation (a) with modeled flood inundation using NWM discharges (b), observed discharges (c), DEM-derived catchments from evenly distributed nodes (d), etched high-resolution hydrography (e), 3-m DEM (f) and etched hydrography with 3-m DEM (g). The results show that the modeled inundation extent (Figures 2.8b-2.8g) is able to capture the majority of the observed inundation extent from Planet RapidEye satellite (Figure 2.8a) in all scenarios. Evaluation metrics C and F (Table 2.1) quantify the performance of each scenario.

The results (Figure 2.9 and Table 2.1) show that the adoption of measured discharge values improves C as it reduces from 1.77 in scenario 1 (CFIM-NWM) to 1.65 in scenario 2 (CFIM-OBS). However, F degrades slightly (i.e., 0.47 to 0.45). The

improvement in C is due to better representation of tributary flows, which, at least for the NWM assimilation we had, are inconsistent with the observations. The reduction in F is due to overall greater inundation because of higher observed flows, and the DEM having fairly extensive flat areas where pits were filled.

Using evenly distributed nodes to create uniform-length reaches and associated catchments for use in synthetic rating curve estimation, going from scenario 2 (CFIM-OBS) to scenario 3 (10m-Uniform), does not lead to an improvement in overall inundation extent prediction as C increases due to the overestimation of stage caused by flat areas near the downstream end of the study area close to Bear River City (see arrow in Figure 2.9c). On the other hand, F increases slightly from CFIM-OBS to 10m-Uniform (Table 2.1) indicating that location-specific overestimation and underestimation are slightly improved. One reason for this is that catchments are better aligned with the DEM (Figure 2.10).

In the CFIM-OBS scenario (Figure 2.10a), the discrepancy between NHDPlus catchments and HAND derived from 10-m DEM is quite evident. An area within the catchment, according to the 10-m DEM, drains to another nearest stream, and this impacts the calculation of hydraulic properties and mapping of inundation (Figure 2.10c). On the other hand, the catchment derived based on the 10-m DEM is more consistent (Figure 2.10b) and results in improved mapping of inundation.

The results (Figure 2.9 and Table 2.1) show that the flow direction conditioning of the 10-m DEM (scenario 4) improves both C and F metrics (Table 2.1). The overestimation of stage caused by flat areas near the downstream end of the study area close to Bear River City (see arrow in Figure 2.9d) is reduced. Additionally, using a high-

resolution DEM (3-m DEM in scenario 5, i.e., 3m-Uniform) improves both metrics compared to 10m-ETCH (F: 0.54>0.52 and C: 1.60<1.67). However, overestimation still occur in some areas (such as areas close to Highway 30; Figure 2.9e) due to a barrier in the 3-m DEM caused by Highway 30. Overall flow direction conditioning of the 3-m DEM (scenario 6) improves the mapping of inundation in terms of both C and F metrics (Table 2.1 and Figure 2.9f). In particular, the F metric improves noticeably since the overestimation (red color) of the modeled flooded areas is reduced around the areas where road crossings exist.

To illustrate the effect of flow direction conditioning (etching) with the 3-m DEM, we zoomed in on an area where 3m-Uniform has mapped inundation that was not observed (Figure 2.11). In the 3m-Uniform scenario (Figure 2.11a), the HAND map shows a notable flat area, which then caused this entire area to be flooded (Figure 2.11c). In 3m-ETCH (scenario 6), where etching removed the DEM barrier, the overestimation of the modeled flood was reduced (Figure 2.11d).

2.4.3 Sensitivity Analysis: Optimal Stage and Roughness Coefficient

The CFIM method uses a uniform value of 0.05 for Manning's n in the calculation of synthetic rating curves. We also used this value in the six scenarios reported so far. However, this assumption may be responsible for uncertainties that remain in the modeled inundation extent. To investigate this, we used catchments delineated with the 3m-ETCH DEM and searched over a range of stage (h) values for the stage that provided the best fit (highest F metric) compared to the Planet RapidEye observed flood inundation extent. This provides a quantification of the best inundation mapping possible using the HAND approach with the DEM and catchments chosen, separate from

uncertainties in discharge and synthetic rating curve. The stage evaluated to produce best fit F (Figure 2.12a) was then used to back calculate the Manning's n that would produce this stage using equation (1) and the observed discharge (Figure 2.12b) for each reach. The corresponding fit metric, F, (Figure 2.12c) and correctness metric, C, (Figure 2.12d) are also shown for each catchment. Overall, the optimal stage resulted in F = 0.72 and C = 1.11. The average of the fitted Manning's n values was n = 0.02 and stage, as well as the Fit metrics for this n, applied as a uniform value to all catchments, and the a-priori n = 0.05 are shown (Figures 2.12 a, 2.12c, and 2.12d). Applying the average of the fitted Manning's n (n = 0.02) to each catchment with the 3m-ETCH scenario produced an inundation map where F and C were 0.63 and 1.12, respectively. Note that this notably improves both metrics relative to the best from Table 2.1 (i.e., F = 0.59 and C = 1.49). Thus, an improvement in F of about 7% (from 0.59 to 0.63) may be obtained simply by calibrating n, keeping the value the same everywhere. A further improvement (from 0.63 to 0.72) may be obtained by letting n vary spatially. Furthermore, recognizing that F=0.72 is the optimal fit of stage given the DEM and RapidEye observations, we can interpret F values from Table 2.1 relative to this value. Specifically, for this study the impact of flow direction conditioning at the 10-m scale was an improvement of F from 0.47 to 0.52, a 7% improvement. At the 3-m scale, flow direction conditioning improved F from 0.54 to 0.59, an 7% improvement. In moving from 10- to 3-m DEM, F with flow direction conditioning improves from 0.52 to 0.59, an 10% improvement.

2.4.4 Validation of Developed Improvements

To validate the improvements developed, we applied the approach to a reach of the Ocheyedan River (about 13 km) close to Spencer City in Iowa, USA, that experienced

a flood in March 2019. This validation case was selected using the following criteria: (1) the reach of interest should have at least one active observation gage to provide the observed streamflow, (2) it should have high-resolution DEM (preferably 3m), (3) the region of interest should have high-resolution and cloud free satellite imagery on the flood date, and (4) it should include roads crossing the stream rivers such that the effect of etching on HAND calculations makes a difference. This validation case was identified by searching waterwatch.usgs.gov to identify recent floods and at the same time searching Planet.com for the availability of Planet satellite imagery and nationalmap.gov to check whether the region a had high-resolution DEM (3m). After several attempts, we selected a reach of the Ocheyedan River (about 13 km) close to Spencer City in Iowa, USA, as our validation case study because it met all criteria mentioned above (Figure 2.13a).

The flood of March 2019 on the Ocheyedan River reach (Figure 2.13d) shows the daily average peak value of 172.45 m^3/s on 15 March observed at the USGS 06605000 Ocheyedan River near Spencer gage. The National Weather Service flood stage for this gage is 2.44 m (or 8.0 ft.), and the observed gage heights at the USGS 06605000 gage were above the flood stage during 14–23 March (for example 3.66 m on 15 March and 3.05 m on 21 March at noon).

A Planet Sentinel-2 image from 21 March 2019 captured at 17:21 UTC (i.e., 12:21 pm CT) was selected as being closest in time to the peak (Figure 2.13c). This image is within the flood period (i.e., 14–23 March) where the gage heights were above the flood stage. When compared to another Sentinel-2 image captured on 20 April 2019 (Figure 2.13b), the flooded region can be seen in the images. The Sentinel-2 satellite

(different from the RapidEye satellite used in the Bear River) is a wide-swath, high-resolution and multi-spectral satellite that contains 13 spectral bands with different spatial resolution (10, 20, and 60m). The specific image selected was S2A_MSIL1C_20190321T171011_N0207_R112_T15TUH_20190322T001148. This tile includes 100% coverage of the study area with 0% cloud coverage. As in the workflow used in the Bear River case study, we used the Visible and Near Infrared bands (i.e., band B2 [blue], B3 [Green], B4 [Red], and B8 [NIR] with spatial resolution of 10m) for the supervised classification.

We created a HAND raster for the 10-m DEM (as used in the published CFIM), 3-m DEM, and 3-m etched DEM and zoomed in to a location (the red circle in Figure 2.13) about 4-km downstream of the USGS 06605000 gage, where a road crosses the Ocheyedan River reach (Figure 2.14). Results illustrate the importance of using a high-resolution DEM on the HAND raster as well as the impact of the etching approach on removing barriers when using such a high-resolution DEM dataset. With the 10-m DEM (Figure 2.14a), a considerable portion is flat due to pit filling, resulting in a HAND map with values close to 0 for most of the river corridor. Without etching (Figure 2.14b), the existence of the road barrier seems more prevalent in the high-resolution DEM, which affects the HAND raster. Once again, this is of concern because it influences the extraction of river hydraulic geometry and the synthetic rating curve from the HAND map. The flow direction conditioning approach (Figure 2.14c) etches the path of the high-resolution NHDPlus flowline into the 3-m DEM and provides a way to resolve (punch through) the barrier.

We followed the procedure of dispersing nodes approximately uniformly along the Ocheyedan River reach to avoid the sometimes small and irregular-sized NHDPlus catchments (Figure 2.15a) that degrade synthetic rating curve estimation. We placed six nodes approximately evenly along the main stem (Ocheyedan River reach between the upstream junction of Stony Creek and the Ocheyedan River and the downstream junction of the Ocheyedan River and Little Sioux River). This resulted in five catchments each draining to a reach of about 3 km long (Figure 2.15b).

We used the daily average discharge value of 75.61 m^3/s (observed on 21 March) in the main river and illustrated the comparison of the modeled inundation based on HAND for both CFIM and the 3-m etched DEM with classified inundation from Planet Sentinel-2 satellite imagery (Figures 2.15c and 2.15d). In order to prevent inconsistency in the different domain due to misalignment of the NHDPlus and DEM-derived catchments, we chose a region that is available for both scenarios (i.e., the NHDPlus catchments) for calculating evaluation metrics. Results show that the modeled inundation extent is able to capture the majority of the observed inundation extent from Planet Sentinel-2 in both scenarios. However, discrepancies between modeled and observed flood inundation are more apparent in the CFIM method (Figure 2.15c).

Computed evaluation metrics (shown in Figures 2.15c and 15d) for each scenario shows that using evenly distributed nodes to create uniform-length reaches and associated catchments for use in synthetic rating curve estimation along with flow direction condition to resolve barriers in a DEM improve both C and F metrics compared to CFIM (C: 1.17<1.44 and F: 0.67>0.56). The discrepancy between NHDPlus catchments and HAND derived from a 10-m DEM resulted in overestimation of flood inundation in

CFIM (see arrows labeled as “A” in Figure 2.15c). These areas within the NHDPlus catchment drain to other nearest stream according to the 10-m DEM. This affects the calculation of hydraulic properties and mapping of inundation. In addition, due to the existence of DEM barriers, a considerable portion is flat due to pit filling, resulting in a HAND map with values close to 0 and overestimation of the estimated stage (see the arrow labeled as “B” in Figure 2.15c). Results show our improvements reduce the occurrence of unrealistic HAND values and diminish the impact of DEM errors on flood inundation mapping.

Evaluating the stage that best matches HAND-based flood inundation with observed inundation and letting Manning's n vary spatially resulted in F = 0.69 and C = 1.10. This improves both metrics relative to the best scenario (i.e., F = 0.67 and C =1.17). Thus, an improvement in F of about 3% (from 0.67 to 0.69) may be obtained simply by calibrating Manning's n. Applying the average of the fitted Manning's n values (n=0.04) as a uniform value to all five catchments produced an inundation map where F and C were 0.67 and 1.13, respectively. Recognizing that F=0.69 is the optimal fit of stage given the DEM and Sentinel-2 observations, we can interpret F relative to this value. In moving from CFIM to 3m-ETCH, F with flow direction conditioning improves from 0.56 to 0.67, a 16 % improvement.

2.5 Discussion

This work follows on from the work of Zheng, Maidment, et al., (2018); Zheng, Tarboton, et al., (2018) and Liu et al. (2018) that have developed and advanced the HAND approach for use on the U.S. NHDPlus network by the National Water Model. In

our work, the streams to which HAND was calculated, and the catchments over which hydraulic properties were evaluated, were derived purely from the DEM after conditioning. This is important to ensure consistency between streams and catchments in the terrain analysis processing for enriching the content of DEM data for use in hydrologic modeling. This should be the case no matter the resolution or accuracy of the DEM. There will always be DEM errors to some degree. Flow direction conditioning alters the DEM to be consistent with the given hydrography, taken to be a better source for flow direction than the DEM, which suffers from artificial barriers.

The approach evaluated here differs from Zheng, Tarboton, et al. (2018) in two ways. First, Zheng, Tarboton, et al. (2018) did not use conditioning on high-resolution streams to etch, or punch through, road barriers, and thus HAND evaluated by Zheng, Tarboton, et al. (2018) will still be limited by the occurrence of flat areas upstream of these barriers. Second, streams mapped using the geodesic minimization approach are not guaranteed to align with topographic minima the way that streams derived directly from the DEM flow directions do, and thus, it cannot be guaranteed that they are good targets with which to evaluate HAND. In the calculation of HAND, it is important that the "stream" used as a target for HAND be consistent with the DEM, or, more specifically, be located at the bottom of whatever valley or channel is represented in the DEM, so that HAND is a positive quantity measured down to this target stream.

The HAND process used in this study has uncertainty due to the simplified representation of flow hydraulics through the assumption of uniform flow and application of Manning's equation over irregular stream reaches. This is acknowledged as a shortcoming, but it is also advocated as a useful approximation for regional- and national-

scale modeling, as the information needed to apply more rigorous hydraulic approaches is rarely available over large areas. The use of reach-averaged hydraulic properties derived from HAND is seen as an advantage over cross-section-based approaches, such as are commonly used in Federal Emergency Management Agency HEC-RAS studies, because the reach-averaged hydraulic properties account for, and integrate, through the aggregation of volume and wetted bed area associated with each flow depth, the variability within a stream reach that can be lost between cross sections in a cross-section-based approach.

The inundation mapped from remote sensing, from Planet RapidEye satellites or Planet Sentinel-2 in this study, also has uncertainties. The best match that would be possible with the Planet RapidEye classified inundation used in the Bear River reach or Planet Sentinel-2 classified inundation used in the Ocheyedan River reach case studies and HAND mapped inundation has fitness scores of 0.72 and 0.69, respectively, less than the theoretical optimum of 1. This is partly due to errors and uncertainties in satellite mapped inundation, and such inundation being inconsistent with the topography. For example, adjacent grid cells of equal HAND value should either both be inundated or not, but this was not always the case with Planet RapidEye inundation.

Adjusting the HAND inundation threshold, h, to achieve the best fit between Planet RapidEye or Sentinel-2 inundation and HAND inundation provides an independent estimate of the stage in each stream reach. When combined with observed discharge (or modeled discharge where observations have been assimilated), this provides an independent estimate of the Manning's n channel hydraulic roughness parameter consistent with other assumptions in the HAND approach. Fitted roughness

values generated by doing this were on average less than the default value used in CFIM for both Bear River and Ocheyedan River case studies, although, again, this inference is limited to these study reaches. This may reflect a bias in the CFIM value, or may also reflect part of the channel being missing in the DEM due to the DEM representing water surface elevations. This observation opens some new questions for future research. First, it emphasizes the importance of having DEMs that represent the bathymetry (channel bed topography) as closely as possible. Second, there is a question as to whether, in locations where the DEM does not represent bed topography, the hydraulic geometry parameters of the missing part can be inferred from the inferred stage, discharge, and roughness parameter. Third, this suggests an opportunity to use observed inundation from past floods to infer stream reach hydraulic roughness in a way that is consistent with HAND hydraulic geometry and to use these values in hydraulic routing. We recognize that the specific Manning's n values fit here are limited because they have been estimated from one event. There would certainly be value in examining multiple events in estimating Manning's n. We should not want to overstate the importance of the specific Manning's n estimated, but rather note that this provides an approach or an opportunity for estimating a spatially variable Manning's n that warrants further investigation as a way to overcome the limitations and bias associated with a single roughness parameter that was evident in the results. This may provide a way to come up with distributed roughness parameters for use in distributed reach scale hydraulic routing, while noting that the presence of high-resolution satellite imagery might be a limiting factor. It is worth observing that Sichangi et al. (2016) have explored relationships between modified Manning's equation parameters and stage and discharge from remote sensing. Furthermore, the generalization

of rating curves, such as addressed in this paper, for stream reaches anywhere there is a good DEM, has potential for use with remote sensing approaches that are being pursued with SWOT (Biancamaria et al., 2010). This is thus a rich area for future research where it is an open question as to the degree to which improved hydraulic parameters would help towards better discharge forecasts, as these also involve routing. The flow direction conditioning method uses stream hydrography to condition the DEM and provides an improvement whenever the stream hydrography is of a resolution or quality to be a better indicator of flow than the DEM by itself. This method is particularly beneficial where there are artificial barriers that result in flat areas where topographic information is lost during pit filling, common at many transportation stream crossings. The flow direction conditioning method required only two inputs: a DEM and a high-resolution hydrography dataset. In some cases, the lack of a high-resolution hydrography dataset may hinder the applicability of the flow direction conditioning approach. However, in the United States, the high-resolution hydrography mapped at 1:24,000 scale, as used in this study, is available from the U.S. National Hydrography Dataset for the entire continental U.S., and the flow direction conditioning could be applied nationwide. The computational cost is significant. DEM processing steps are essentially done twice: first to determine flow directions in a burned in DEM and then to condition the DEM and repeat the process for the conditioned DEM. Computation time is thus expected to be about double. However, this is something that is done once to prepare the data for the HAND approach, and we feel the improvements merit the extra computation, which can be done quite quickly as described by Liu et al. (2018). There are no additional ongoing computational costs once the flood inundation is being modeled.

2.6 Conclusions

The first contribution of this paper is flow direction conditioning, a new DEM terrain analysis method that was developed using high-resolution hydrography data to alter, or condition a DEM so that elevation values do not increase moving downstream along hydrographic flow paths. This serves to remove artificial barriers in the DEM due to infrastructure such as road crossings, producing an important improvement in the calculation of HAND and mapping of inundated flooding. In this study, the fit metric that quantifies how well modeled inundation matched high-resolution satellite observations was improved by 7% for a 10-m DEM and 8% for a 3-m DEM in the Bear River case study, an important improvement given the relatively small but important fraction of the area that flow direction conditioning impacts. Further evaluation of this approach for different study areas is certainly warranted. High-resolution hydrography mapped at 1:24,000 scale, such as used in this study, is available from the U.S. National Hydrography Dataset for the entire continental United States, and this approach could be applied, similar to Liu et al. (2018), at a continental U.S. scale to compute HAND and associated channel hydraulic properties.

The importance of DEM scale (3-m vs. 10-m) was also quantified. Higher-resolution DEM data such as the 1/9 arc-sec (3-m) resolution from the U.S. NED for some areas can improve the precision with which flood inundation can be mapped using the HAND approach. In this work, the fit metric improved 10% and 16% with the higher resolution DEM in the Bear River and the Ocheyedan River case studies, respectively. This provides input to consider when evaluating the merit and additional expense of 3-m data collection.

The results presented here have shown that the misalignment between NHDPlus catchments and DEM-derived catchments can be a limitation in the application of the HAND approach to flood inundation mapping. Catchments and stream reaches derived from nodes evenly spread along streams to balance reach lengths helped resolve some of these problems. Using catchments derived from the DEM produced results that were improvements in comparison to those obtained using NHDPlus catchments derived from a coarser DEM.

Lastly, the fixed roughness parameter in CFIM can be a limitation, and this study introduced an approach to estimate reach specific Manning's n from observed flood inundation.

In an effort to make this study reproducible, the data and computational scripts used to produce the study results have been saved in HydroShare (Garousi-Nejad et al., 2019). The code for the flow direction conditioning tool is part of TauDEM and is available from the TauDEM GitHub repository (http://github.com/dtarb/taudem).

ACKNOWLEDGMENTS

The authors would like to acknowledge support from the Utah Water Research Laboratory to conduct this research. The preliminary results of scenario 1 were obtained while the author was attending the University Consortium for Geographical Information Science (USGIS) summer school in 2017, supported by the National Science Foundation (NSF). The authors thank Shaowen Wang for his support and Yan Liu for providing insightful comments and help on the results.

REFERENCES

Biancamaria, S., Andreadis, K. M., Durand, M., Clark, E. A., Rodriguez, E., Mognard, N. M., Alsdorf, D. E., Lettenmaier, D. P., & Oudin, Y. (2010). Preliminary Characterization of SWOT Hydrology Error Budget and Global Capabilities. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 3(1), 6–19. https://doi.org/10.1109/JSTARS.2009.2034614

Cooley, S., Smith, L., Stepan, L., & Mascaro, J. (2017). Tracking Dynamic Northern Surface Water Changes with High-Frequency Planet CubeSat Imagery. Remote Sensing, 9(12), 1306. https://doi.org/10.3390/rs9121306

Costa-Cabral, M. C., & Burges, S. J. (1994). Digital Elevation Model Networks (DEMON): A model of flow over hillslopes for computation of contributing and dispersal areas. Water Resources Research, 30(6), 1681–1692. https://doi.org/10.1029/93WR03512

ESRI. (2018). ArcPro Software Base Maps. https://services.arcgisonline.com/ArcGIS/rest/services/World_Imagery/MapServer

Fairfield, J., & Leymarie, P. (1991). Drainage networks from grid digital elevation models. Water Resources Research, 27(5), 709–717. https://doi.org/10.1029/90WR02658

Freeman, T. G. (1991). Calculating catchment area with divergent flow based on a regular grid. Computers & Geosciences, 17(3), 413–422. https://doi.org/10.1016/0098-3004(91)90048-I

Garousi-Nejad, I., Tarboton, D. G., Aboutalebi, M., & Torres-Rua, A. F. (2019). Data For Terrain Analysis Enhancements to the Height Above Nearest Drainage Flood Inundation Mapping Method [Data set]. https://doi.org/10.4211/hs.7235a0d6a18343078b2028085b7d8018

Godbout, L. D. (2018). Error assessment for height above the nearest drainage inundation mapping [Master of Science in Engineering thesis, The University of Texas at Austin]. https://repositories.lib.utexas.edu/handle/2152/68235

Horritt, M. S., & Bates, P. D. (2002). Evaluation of 1D and 2D numerical models for predicting river flood inundation. Journal of Hydrology, 268(1–4), 87–99. https://doi.org/10.1016/S0022-1694(02)00121-X

Ireland, G., Volpi, M., & Petropoulos, G. (2015). Examining the Capability of Supervised Machine Learning Classifiers in Extracting Flooded Areas from Landsat TM Imagery: A Case Study from a Mediterranean Flood. Remote Sensing, 7(3), 3372–3399. https://doi.org/10.3390/rs70303372

Jenson, S. K., & Domingue, J. O. (1988). Extracting topographic structure from digital elevation data for geographic information-system analysis. Photogrammetric Engineering and Remote Sensing, 54(11), 1593–1600.

Laskov, P., Düssel, P., Schäfer, C., & Rieck, K. (2005). Learning Intrusion Detection: Supervised or Unsupervised? In F. Roli & S. Vitulano (Eds.), Image Analysis and Processing – ICIAP 2005 (Vol. 3617, pp. 50–57). Springer Berlin Heidelberg. https://doi.org/10.1007/11553595_6

Lindsay, J. B., & Creed, I. F. (2005). Removal of artifact depressions from digital elevation models: Towards a minimum impact approach. Hydrological Processes, 19(16), 3113–3126. https://doi.org/10.1002/hyp.5835

Liu, Y. Y., Maidment, D. R., Tarboton, D. G., Zheng, X., & Wang, S. (2018). A CyberGIS Integration and Computation Framework for High-Resolution Continental-Scale Flood Inundation Mapping. JAWRA Journal of the American Water Resources Association, 54(4), 770–784. https://doi.org/10.1111/1752-1688.12660

Liu, Y. Y., Maidment, D. R., Tarboton, D. G., Zheng, X., Yildirim, A., Sazib, N. S., & Wang, S. (2016). CyberGIS approach to generating highresolution Height Above Nearest Drainage (HAND) raster for national flood mapping. CyberGIS 16, The Third International Conference on CyberGIS and Geospatial Data Science, Urbana, Illinois.

Martz, L. W., & Garbrecht, J. (1999). An outlet breaching algorithm for the treatment of closed depressions in a raster DEM. Computers & Geosciences, 25(7), 835–844. https://doi.org/10.1016/S0098-3004(99)00018-7

McCabe, M. F., Aragon, B., Houborg, R., & Mascaro, J. (2017). CubeSats in Hydrology: Ultrahigh-Resolution Insights Into Vegetation Dynamics and Terrestrial Evaporation: CUBESATS IN HYDROLOGY. Water Resources Research, 53(12), 10017–10024. https://doi.org/10.1002/2017WR022240

McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., & Rea, R. (2012). NHDPlus Version 2: User Guide.

Merwade, V., Rajib, M. A., & Liu, Z. (2018). Bridging Science and Policy Implication for Managing Climate Extremes (In H. Jung&B. Wang (Eds.), Vol. 10). World Scientific. https://doi.org/10.1142/10867

Moretti, G., & Orlandini, S. (2018). Hydrography-Driven Coarsening of Grid Digital Elevation Models. Water Resources Research, 54(5), 3654–3672. https://doi.org/10.1029/2017WR021206

Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., Waterloo, M., & Saleska, S. (2011). Height Above the Nearest Drainage – a

hydrologically relevant new terrain model. Journal of Hydrology, 404(1–2), 13–29. https://doi.org/10.1016/j.jhydrol.2011.03.051

Nobre, A. D., Cuartas, L. A., Momo, M. R., Severo, D. L., Pinheiro, A., & Nobre, C. A. (2016). HAND contour: A new proxy predictor of inundation extent: Mapping Flood Hazard Potential Using Topography. Hydrological Processes, 30(2), 320–333. https://doi.org/10.1002/hyp.10581

O'Callaghan, J. F., & Mark, D. M. (1984). The extraction of drainage networks from digital elevation data. Computer Vision, Graphics, and Image Processing, 28(3), 323–344. https://doi.org/10.1016/S0734-189X(84)80011-0

Orlandini, S., & Moretti, G. (2009). Determination of surface flow paths from gridded elevation data: SURFACE FLOW PATHS FROM GRIDDED ELEVATION DATA. Water Resources Research, 45(3). https://doi.org/10.1029/2008WR007099

Orlandini, S., Moretti, G., Franchini, M., Aldighieri, B., & Testa, B. (2003). Path-based methods for the determination of nondispersive drainage directions in grid-based digital elevation models: TECHNICAL NOTE. Water Resources Research, 39(6). https://doi.org/10.1029/2002WR001639

Orlandini, S., Moretti, G., & Gavioli, A. (2014). Analytical basis for determining slope lines in grid digital elevation models: SLOPE LINES IN GRID DIGITAL ELEVATION MODELS. Water Resources Research, 50(1), 526–539. https://doi.org/10.1002/2013WR014606

Parker, D. J. (2017). Flood Warning Systems and Their Performance. In D. J. Parker, Oxford Research Encyclopedia of Natural Hazard Science. Oxford University Press. https://doi.org/10.1093/acrefore/9780199389407.013.84

Passalacqua, P., Do Trung, T., Foufoula-Georgiou, E., Sapiro, G., & Dietrich, W. E. (2010). A geometric framework for channel network extraction from lidar: Nonlinear diffusion and geodesic paths. Journal of Geophysical Research, 115(F1), F01002. https://doi.org/10.1029/2009JF001254

Planet. (2017). Planet application program interface: In space for life on Earth. https://api.planet.com

Policelli, F., Slayback, D., Brakenridge, B., Nigro, J., Hubbard, A., Zaitchik, B., Carroll, M., & Jung, H. (2017). The NASA Global Flood Mapping System. In V. Lakshmi (Ed.), Remote Sensing of Hydrological Extremes (pp. 47–63). Springer International Publishing. https://doi.org/10.1007/978-3-319-43744-6_3

Quinn, P., Beven, K., Chevallier, P., & Planchon, O. (1991). The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. Hydrological Processes, 5(1), 59–79. https://doi.org/10.1002/hyp.3360050106

Rahmati, O., Kornejady, A., Samadi, M., Nobre, A. D., & Melesse, A. M. (2018). Development of an automated GIS tool for reproducing the HAND terrain model. Environmental Modelling & Software, 102, 1–12. https://doi.org/10.1016/j.envsoft.2018.01.004

Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J., & Waterloo, M. J. (2008). HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. Remote Sensing of Environment, 112(9), 3469–3481. https://doi.org/10.1016/j.rse.2008.03.018

Rodda, H. J. E. (2005). The Development and Application of a Flood Risk Model for the Czech Republic. Natural Hazards, 36(1–2), 207–220. https://doi.org/10.1007/s11069-004-4549-4

Sangireddy, H., Stark, C. P., Kladzyk, A., & Passalacqua, P. (2016). GeoNet: An open source software for the automatic and objective extraction of channel heads, channel network, and channel morphology from high resolution topography data. Environmental Modelling & Software, 83, 58–73. https://doi.org/10.1016/j.envsoft.2016.04.026

Sangwan, N., & Merwade, V. (2015). A Faster and Economical Approach to Floodplain Mapping Using Soil Information. JAWRA Journal of the American Water Resources Association, 51(5), 1286–1304. https://doi.org/10.1111/1752-1688.12306

Seibert, J., & McGlynn, B. L. (2007). A new triangular multiple flow direction algorithm for computing upslope areas from gridded digital elevation models: A NEW TRIANGULAR MULTIPLE-FLOW DIRECTION. Water Resources Research, 43(4). https://doi.org/10.1029/2006WR005128

Sichangi, A. W., Wang, L., Yang, K., Chen, D., Wang, Z., Li, X., Zhou, J., Liu, W., & Kuria, D. (2016). Estimating continental river basin discharges using multiple remote sensing data sets. Remote Sensing of Environment, 179, 36–53. https://doi.org/10.1016/j.rse.2016.03.019

Soille, P. (2004). Optimal removal of spurious pits in grid digital elevation models: OPTIMAL REMOVAL OF SPURIOUS PITS IN GRID DEMS. Water Resources Research, 40(12). https://doi.org/10.1029/2004WR003060

Soille, P., Vogt, J., & Colombo, R. (2003). Carving and adaptive drainage enforcement of grid digital elevation models: CARVING AND ADAPTIVE DRAINAGE ENFORCEMENT OF GRID DEMS. Water Resources Research, 39(12). https://doi.org/10.1029/2002WR001879

Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water Resources Research, 33(2), 309–319. https://doi.org/10.1029/96WR03137

Tarboton, D. G. (2017). Terrain analysis using digital elevation models (TauDEM). Utah Water Research Laboratory. http://hydrology.usu.edu/taudem

Tesfa, T. K., Tarboton, D. G., Watson, D. W., Schreuders, K. A. T., Baker, M. E., & Wallace, R. M. (2011). Extraction of hydrological proximity measures from DEMs using parallel processing. Environmental Modelling & Software, 26(12), 1696–1709. https://doi.org/10.1016/j.envsoft.2011.07.018

Tourian, M., Elmi, O., Mohammadnejad, A., & Sneeuw, N. (2017). Estimating River Depth from SWOT-Type Observables Obtained by Satellite Altimetry and Imagery. Water, 9(10), 753. https://doi.org/10.3390/w9100753

Verpoorter, C., Kutser, T., & Tranvik, L. (2012). Automated mapping of water bodies using Landsat multispectral data: Automated mapping of water bodies. Limnology and Oceanography: Methods, 10(12), 1037–1050. https://doi.org/10.4319/lom.2012.10.1037

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing, 27(14), 3025–3033. https://doi.org/10.1080/01431160600589179

Zheng, X., Maidment, D. R., Tarboton, D. G., Liu, Y. Y., & Passalacqua, P. (2018). GeoFlood: Large-Scale Flood Inundation Mapping Based on High-Resolution Terrain Analysis. Water Resources Research, 54(12). https://doi.org/10.1029/2018WR023457

Zheng, X., Tarboton, D. G., Maidment, D. R., Liu, Y. Y., & Passalacqua, P. (2018). River Channel Geometry and Rating Curve Estimation Using Height above the Nearest Drainage. JAWRA Journal of the American Water Resources Association, 54(4), 785–806. https://doi.org/10.1111/1752-1688.12661

Appendices

Table 2.1 C = Correctness Metric and F = Fit Metric Values for Each Scenario

Fig 2.9	a.	b.	c.	d.	e.	f.
Scenario	1. CFIM-NWM	2. CFIM-OBS	3. 10m-Uniform	4. 10m-ETCH	5. 3m-Uniform	6. 3m-ETCH
C	1.77	1.65	1.71	1.67	1.60	1.49
F	0.47	0.45	0.47	0.52	0.54	0.59

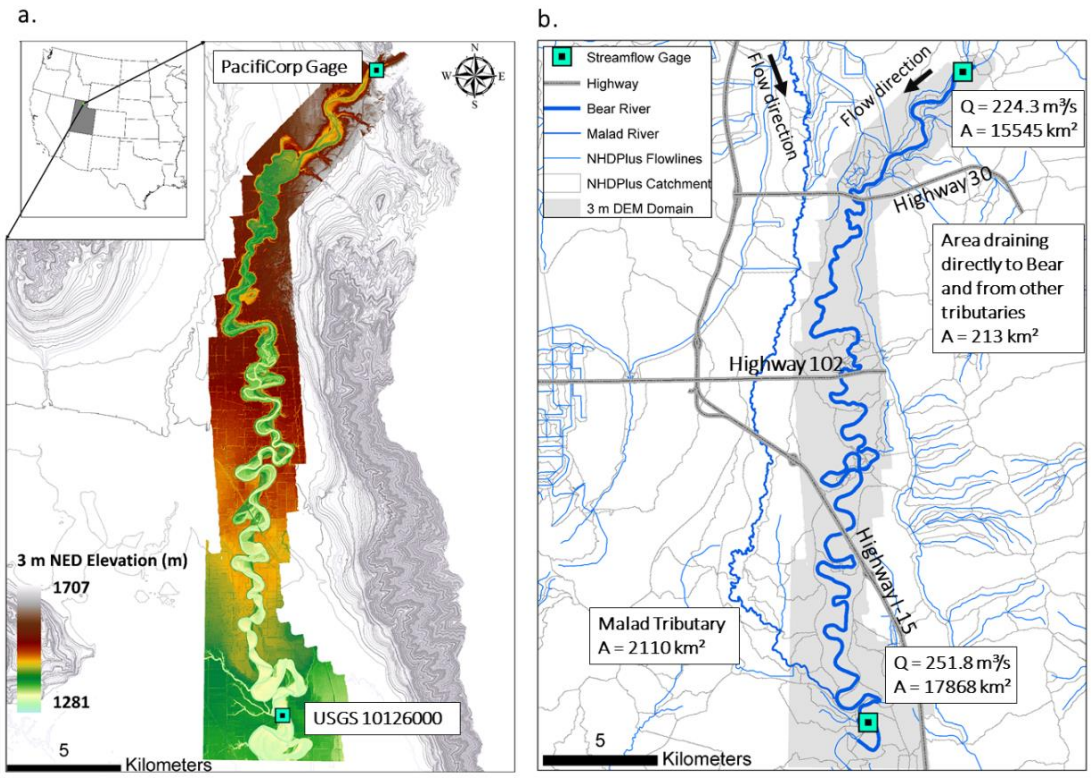

Figure 2.1. (a) Location and topography of the study area with 10-m National Elevation Dataset (NED) as grey contour lines (10-m interval) and the available 3-m NED with raster layer (color region); (b) Hydrography of the study site from the NHDPlus dataset and the locations of two stream gages (i.e., the upper northern PacifiCorp gage at Collinston and the lower southern U.S. Geological Survey [USGS] gage at Corrine). Flow is from north to south.

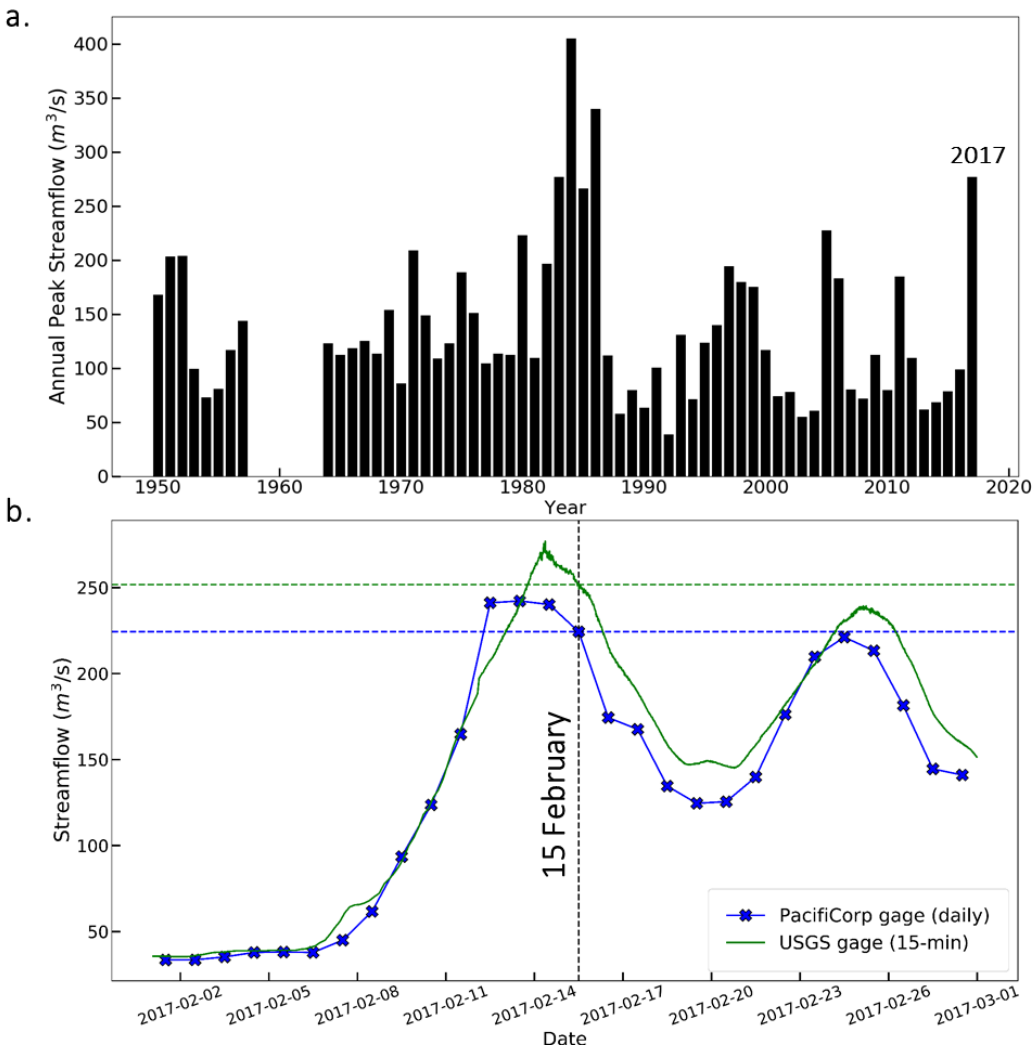

Figure 2.2. (a) Historical annual peak discharges observed at the U.S. Geological Survey (USGS) 10126000 Bear River near the Corrine gage from 1952 to present and (b) discharges observed during February 2017 at the PacifiCorp gage near Collinston (blue line = daily average) and the USGS 10126000 Bear River gage near Corrine (green line = 15-min). The vertical dotted line indicates the date for which Planet RapidEye imagery is available closest to the peak flow, and horizontal dotted lines are the values used for inundation mapping.

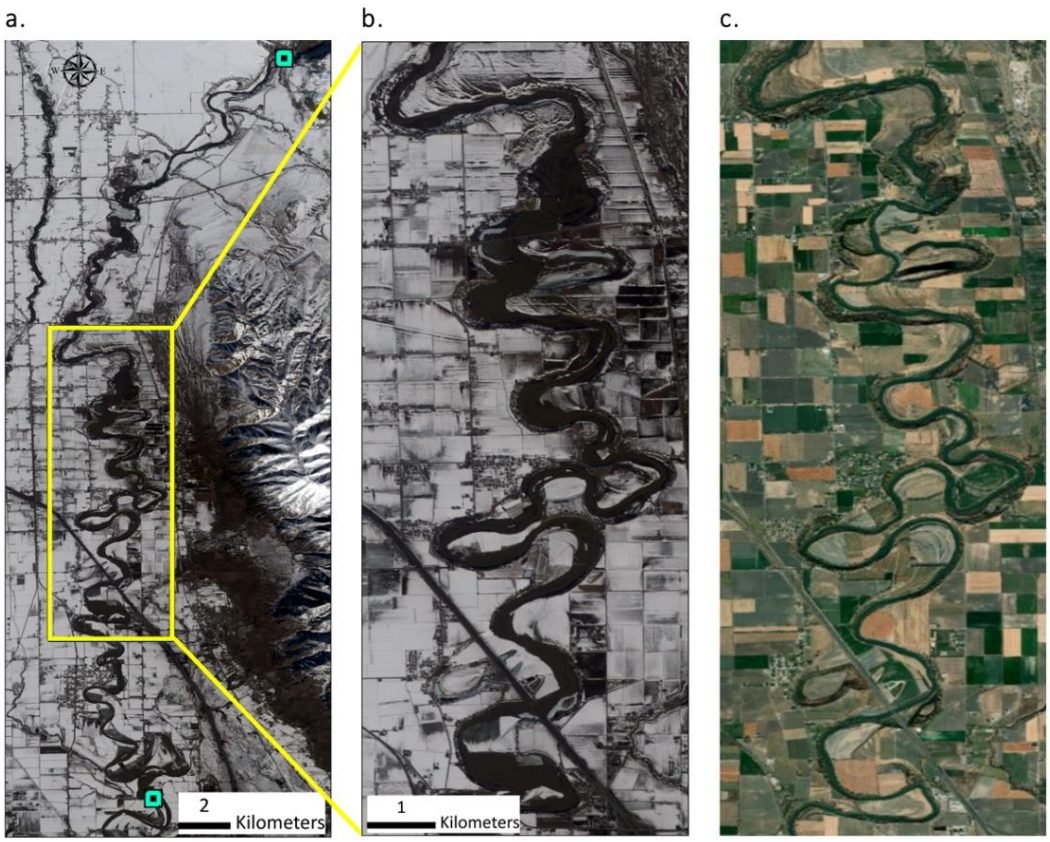

Figure 2.3. (a) The true color RapidEye Analytic Ortho Tile image captured on 15 February 2017 from Planet. (b) Zoomed in RapidEye image. (c) 30-cm resolution world imagery from ESRI (2018) for the same area as shown in (b).

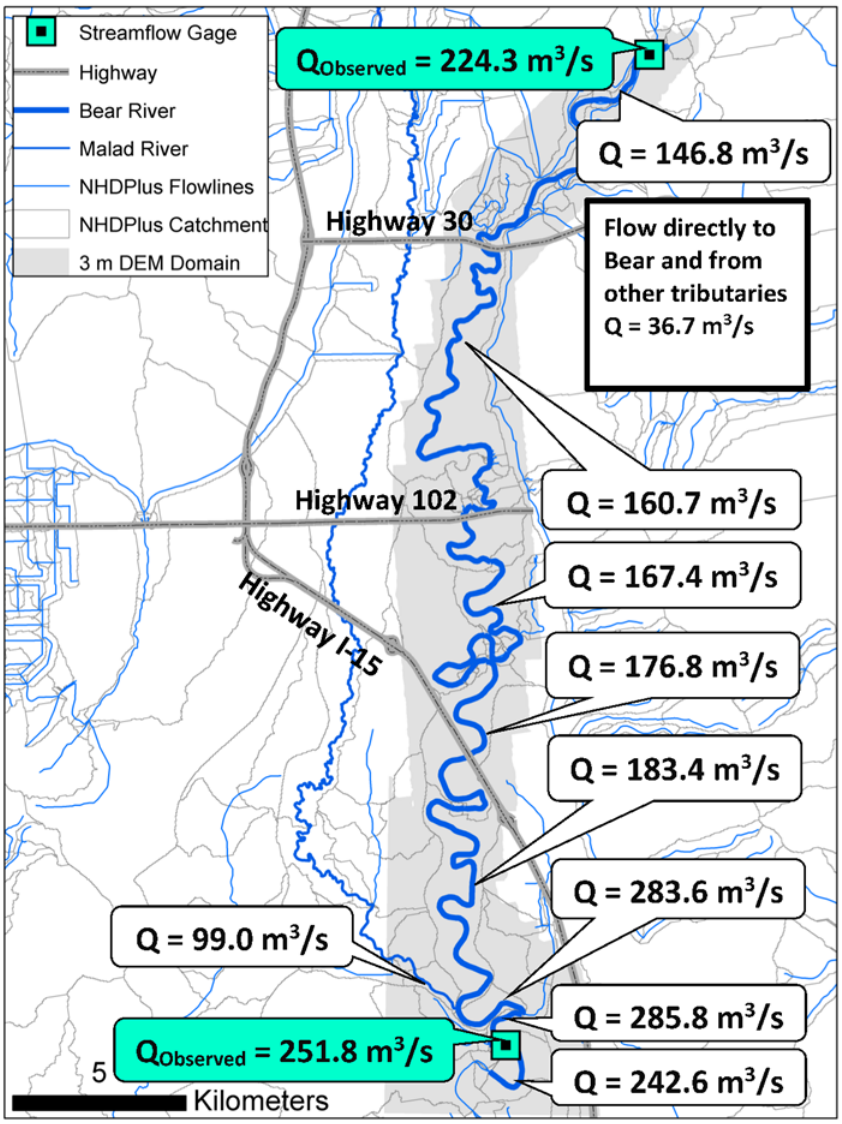

Figure 2.4. Assimilated National Water Model daily average flows (Q) on 15 February 2017 for selected reaches in the Bear River study site and Malad River tributary inflow. Observed flows shown for reference.

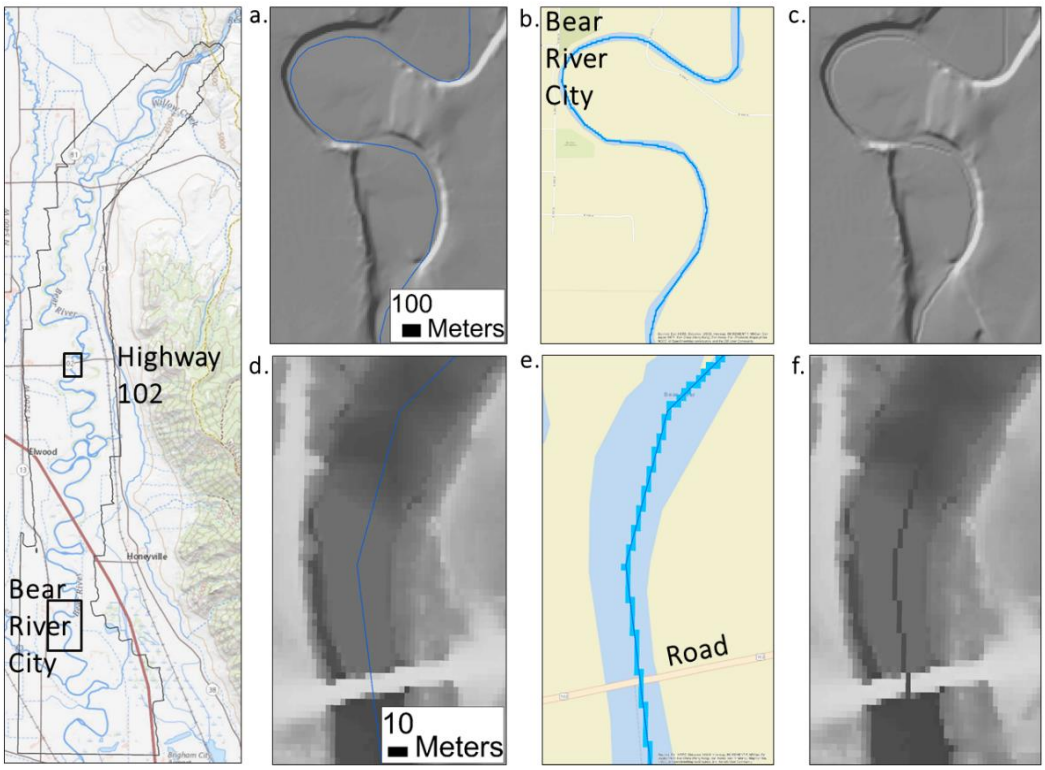

Figure 2.5. Examples of the effect of flow direction conditioning (etching). (a) 10-m digital elevation model (DEM) close to Bear River City, (b) 10-m stream raster of the high-resolution National Hydrography Dataset (NHD), (c) 10-m etched DEM close to Bear River City, (d) 3-m DEM around Highway 102, (e) 3-m stream raster of the high-resolution NHD, and (f) 3-m etched DEM around Highway 102.

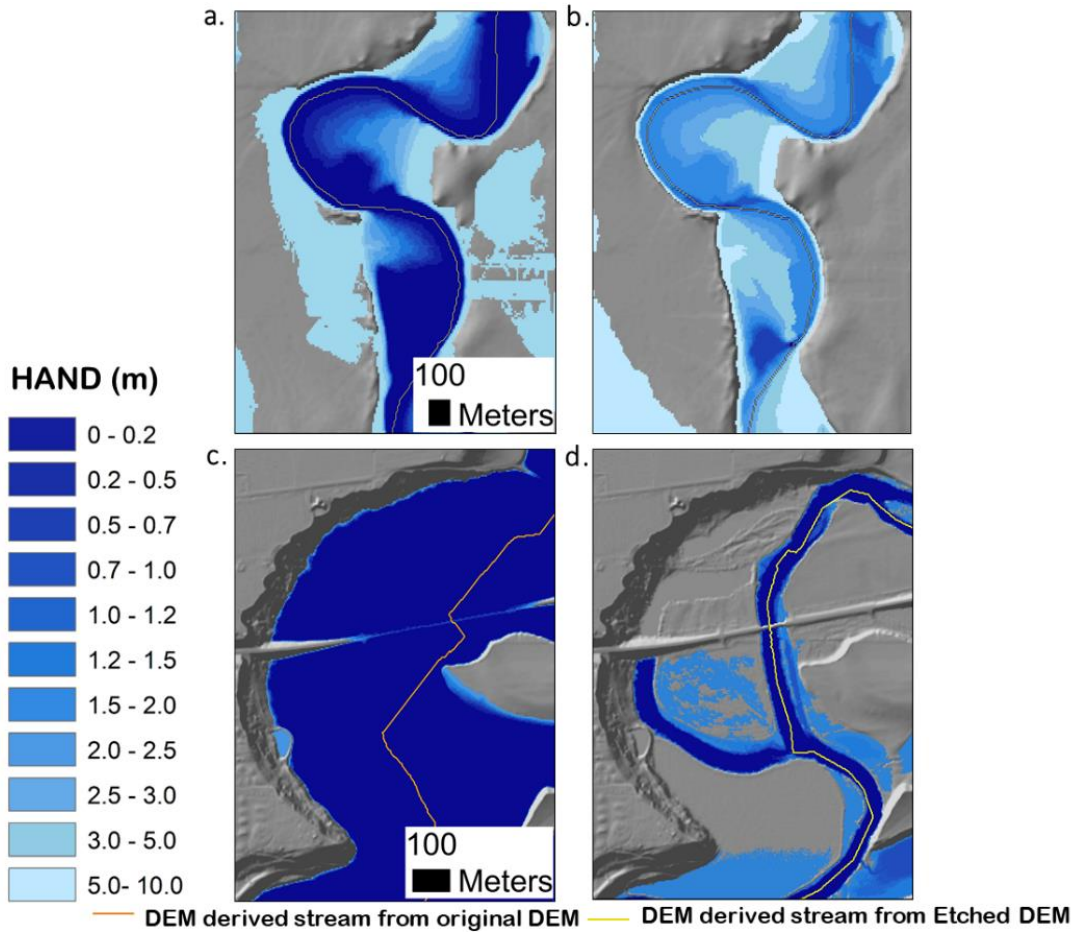

Figure 2.6. The HAND map for the regions shown in Figure 2.5 computed based on (a) 10-m pit-removed digital elevation model (DEM), (b) 10-m pit-removed etched DEM, (c) 3-m pit-removed DEM, and (d) 3-m pit-removed etched DEM for areas close to Bear River City (a-b) and Highway 102 (c and d)

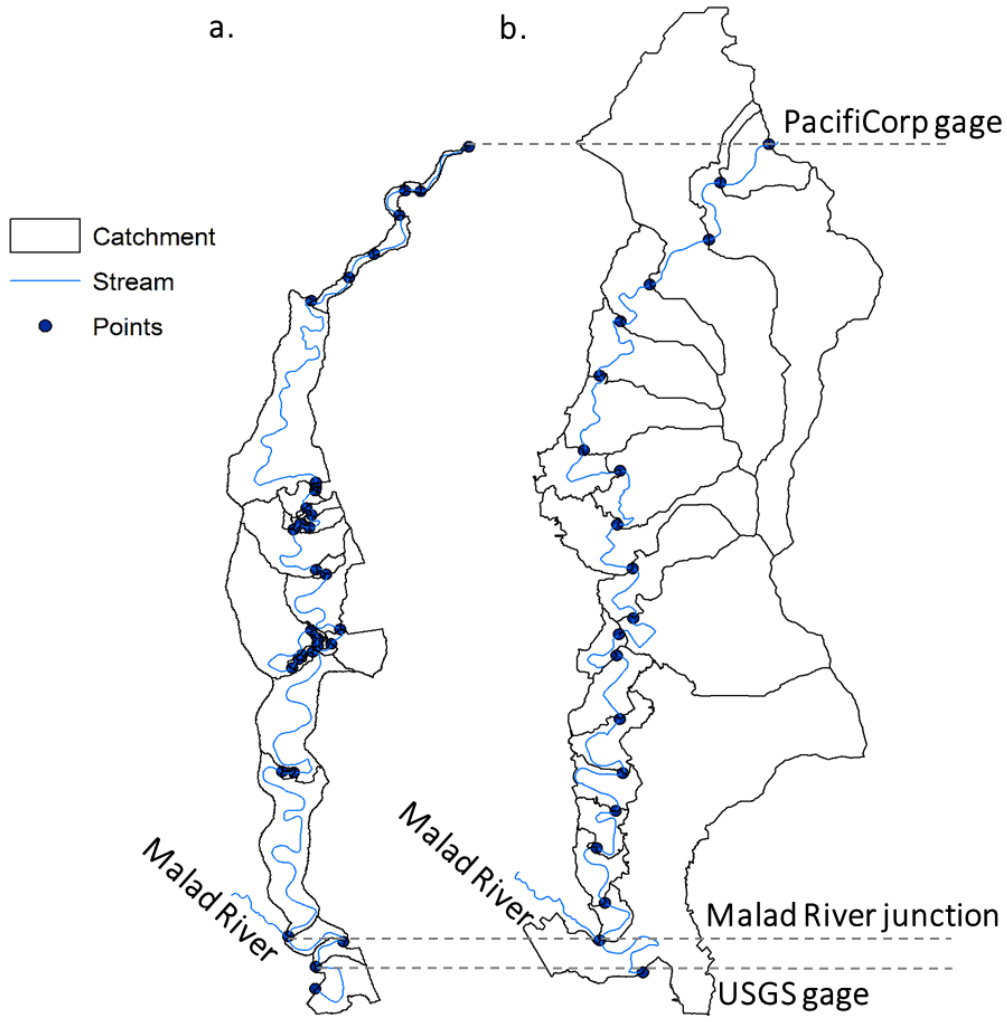

Figure 2.7. Location of nodes and associated catchments in (a) NHDPlus-derived Continental-scale Flood Inundation Mapping (CFIM) approach and (b) evenly distributed nodes to get uniform length reaches.

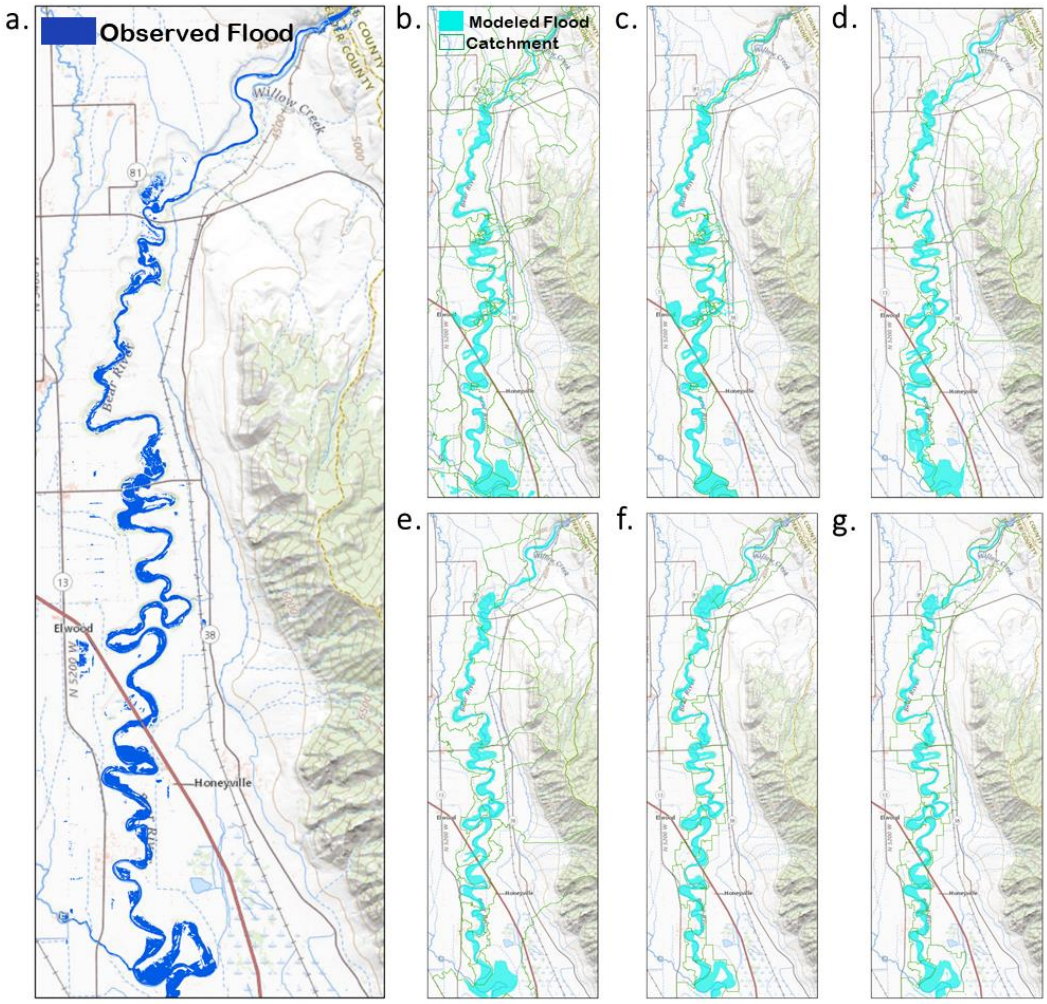

Figure 2.8. Modeled and observed flood inundation maps. (a) RapidEye-OBS: inundation classified from Planet RapidEye satellite imagery; (b) Scenario 1, Continental-Scale Flood Inundation Mapping-National Water Model (CFIM-NWM): CFIM methodology with NWM assimilated flows; (c) Scenario 2, CFIMOBS: CFIM methodology with observed flows; (d) Scenario 3, 10m-Uniform: evenly distributed nodes and TauDEM used to re-delineate channels and catchments from 10-m DEM; (e) Scenario 4, 10m-ETCH: high-resolution National Hydrography Dataset (NHD) flow paths etched into 10-m DEM; (f) Scenario 5, 3m-Uniform: evenly distributed nodes and TauDEM used to re-delineate channels and catchments from 3-m DEM; and (g) Scenario 6, 3m-ETCH: high-resolution NHD flow paths etched into 3-m DEM.

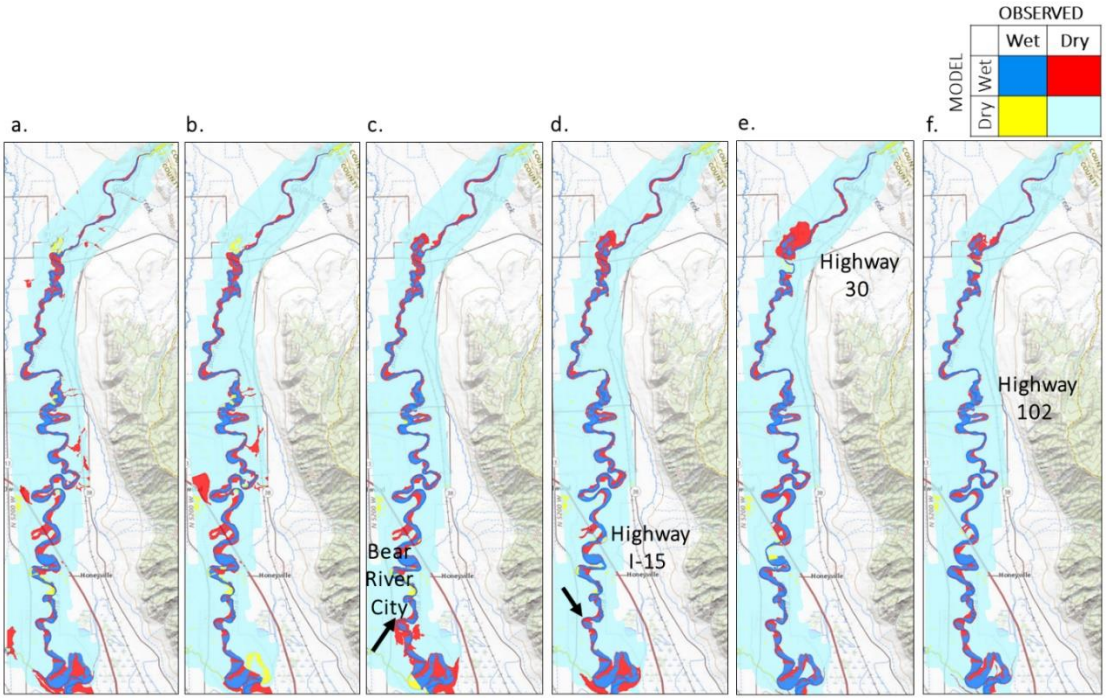

Figure 2.9. Comparison of each scenario with inundation classified from Planet RapidEye imagery. (a) Continental-Scale Flood Inundation Mapping-National Water Model (CFIM-NWM), (b) CFIM-OBS, (c) 10m-Uniform, (d) 10m-ETCH, (e) 3m-Uniform, and (f) 3m-ETCH.

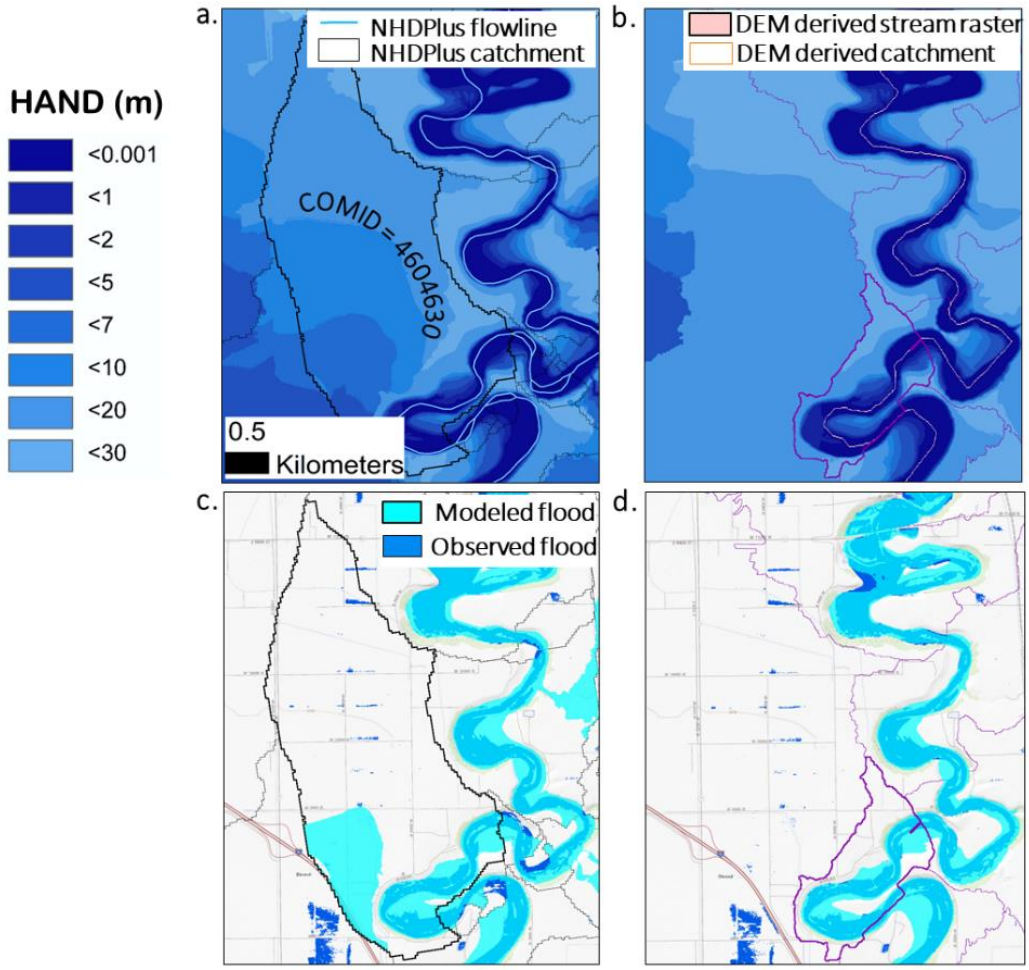

Figure 2.10. The effect of using evenly distributed nodes and digital elevation model (DEM)-derived catchments and streams on flood inundation for a zoomed in region in the study domain. (a) A selected NHDPlus catchment and the HAND map based on 10-m DEM as used in the Continental-Scale Flood Inundation Mapping-National Water Model (CFIM-NWM) and CFIM-OBS scenarios. (b) A selected DEM-derived catchment and the HAND map based on 10-m DEM as used in the 10m-Uniform scenario. (c) Modeled flood in CFIM-OBS. (d) Modeled flood in 10m-Uniform.

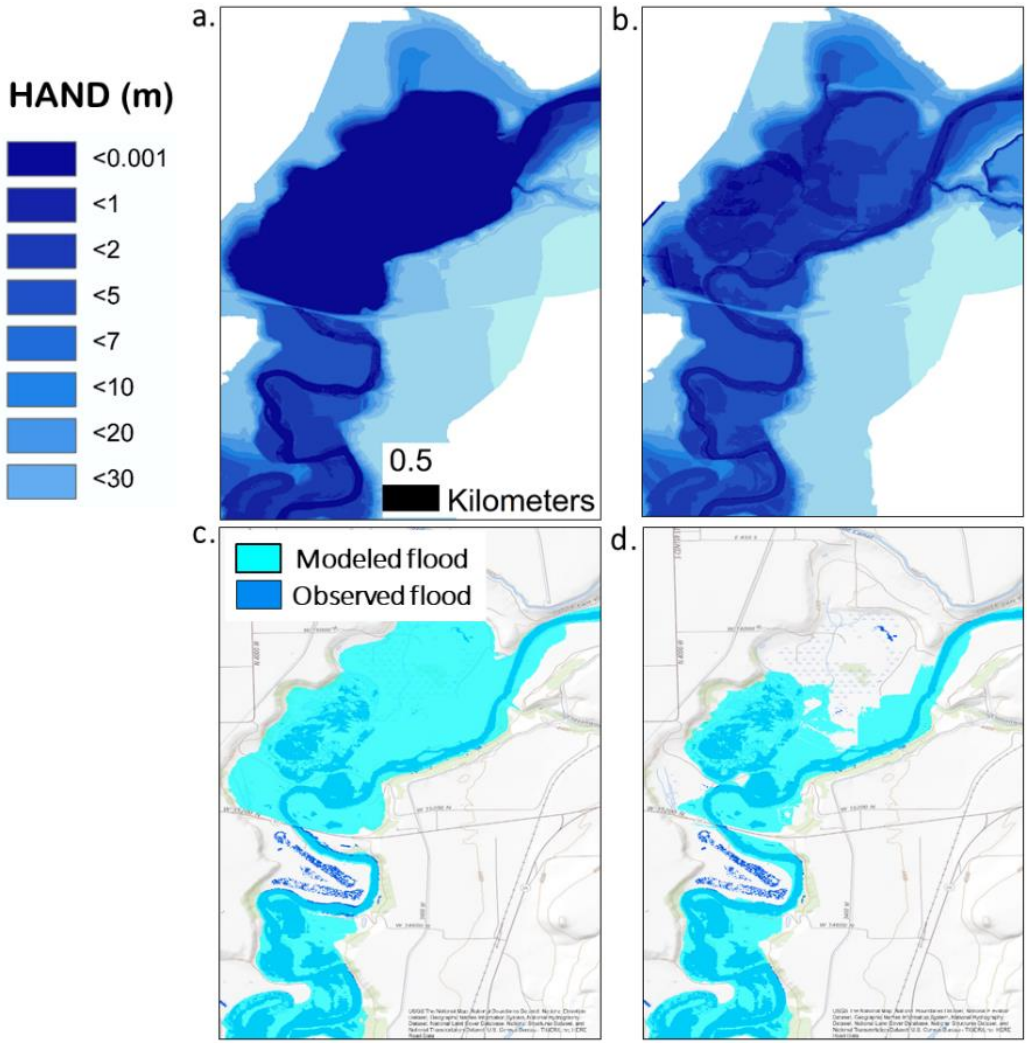

Figure 2.11. The effect of flow direction conditioning on flood inundation for an area around Highway 30. (a) HAND based on 3-m digital elevation model as used in the 3m-Uniform scenario, (b) HAND based on 3-m etched digital elevation model as used in the 3m-ETCH scenario, (c) modeled flood in 3m-Uniform, and (d) modeled flood in 3m-ETCH.

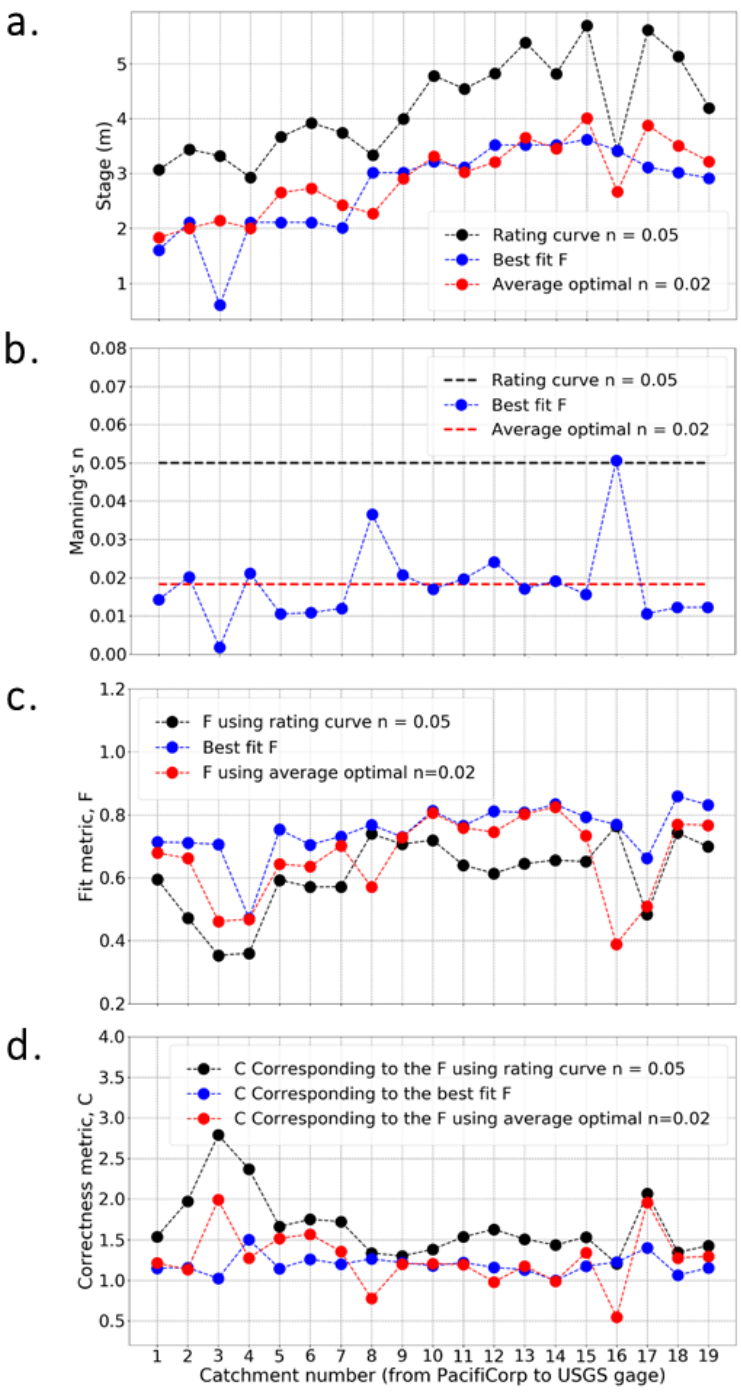

Figure 2.12. Optimizing stage and estimating Manning's n. (a) Stage values for each stream reach using (I) synthetic rating curve with n = 0.05; (II) Best fit F from optimization of h to where F is maximized; (III) synthetic rating curve with n = 0.02, the n value obtained by averaging the n values obtained from the optimal stage. (b) Manning's n obtained from stage corresponding to best fit F. (c) Fit metric F for each of the cases in (a). (d) Correctness metric C for each of the cases in (a).

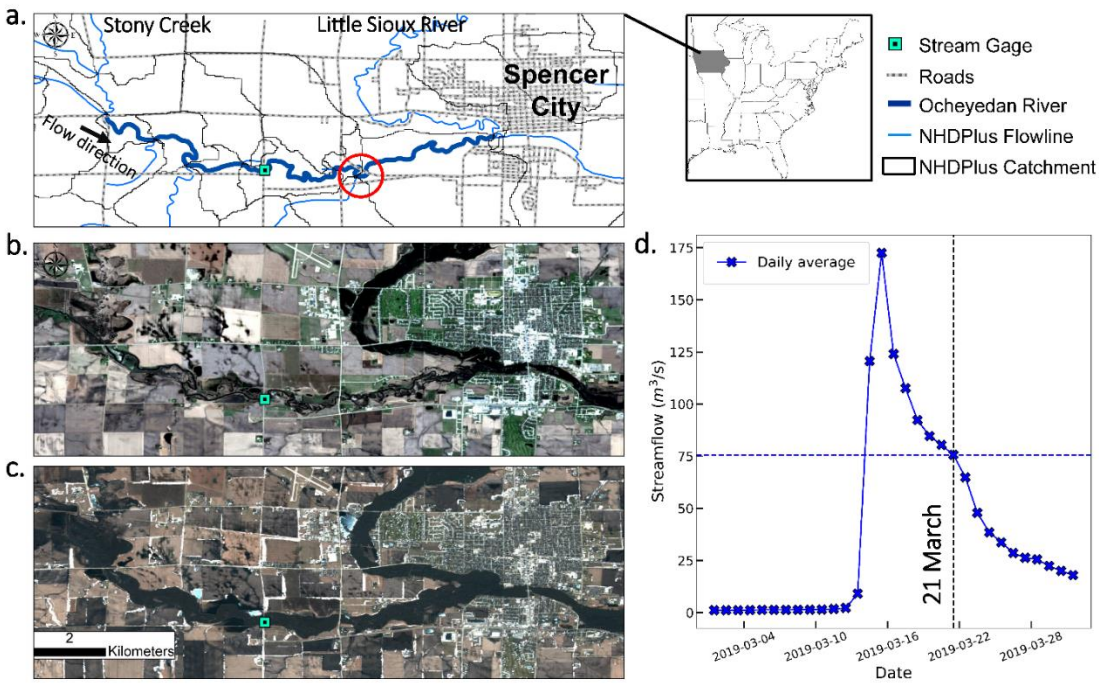

Figure 2.13. Validation case study and the observed flood of March 2019 in the Ocheyedan River in Iowa. (a) Hydrography of the study site from the NHDPlus dataset and the location of the USGS 06605000 gage close to Spencer City (Flow is from west to east), the true color Sentinel-2 Tile image captured on (b) 20 April 2019 and (c) 21 March 2019 from Planet, and (d) the daily average discharges observed during March 2019 at the USGS 06605000 gage. The vertical dotted line indicates the date for which Planet Sentinel-2 imagery is available closest to the peak flow, and the horizontal dotted line is the value used for inundation mapping (i.e., 75.61 m^3/s).

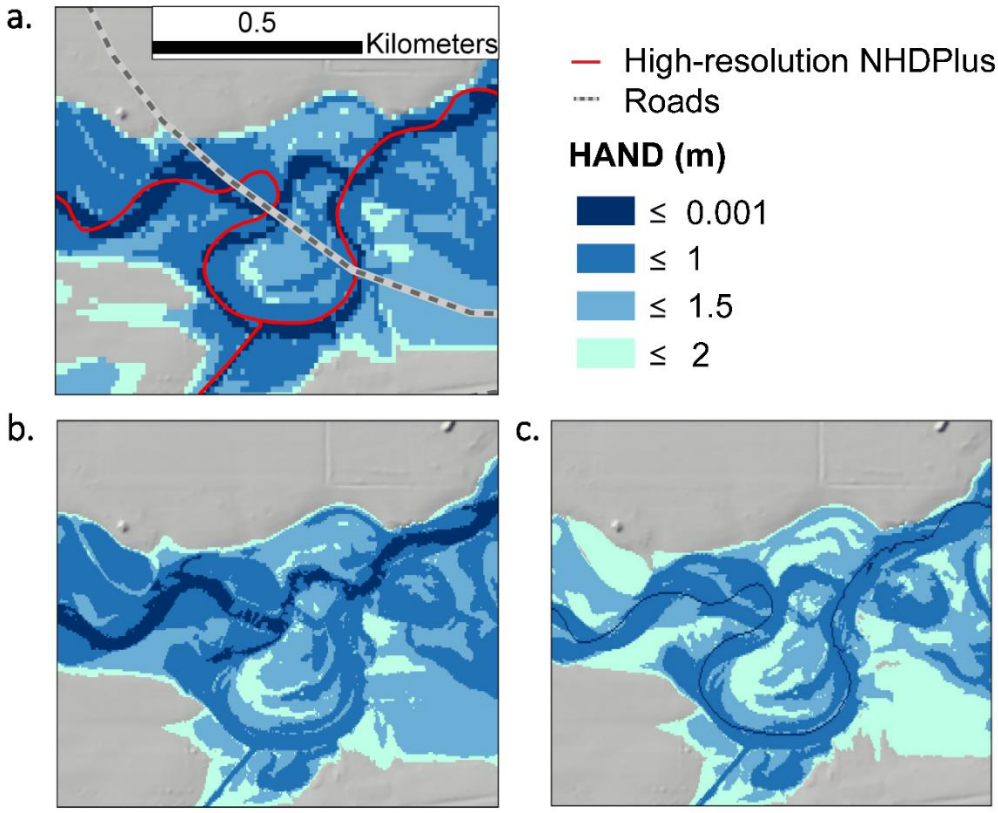

Figure 2.14. Example of the effect of flow direction conditioning (etching) at a location about 4 km downstream of the USGS 06605000 gage. The HAND map based on (a) a 10-m pit-removed digital elevation model (DEM) as used in the published Continental-Scale Flood Inundation Mapping, (b) a 3-m pit-removed DEM, and (c) a 3-m pit-removed etched DEM.

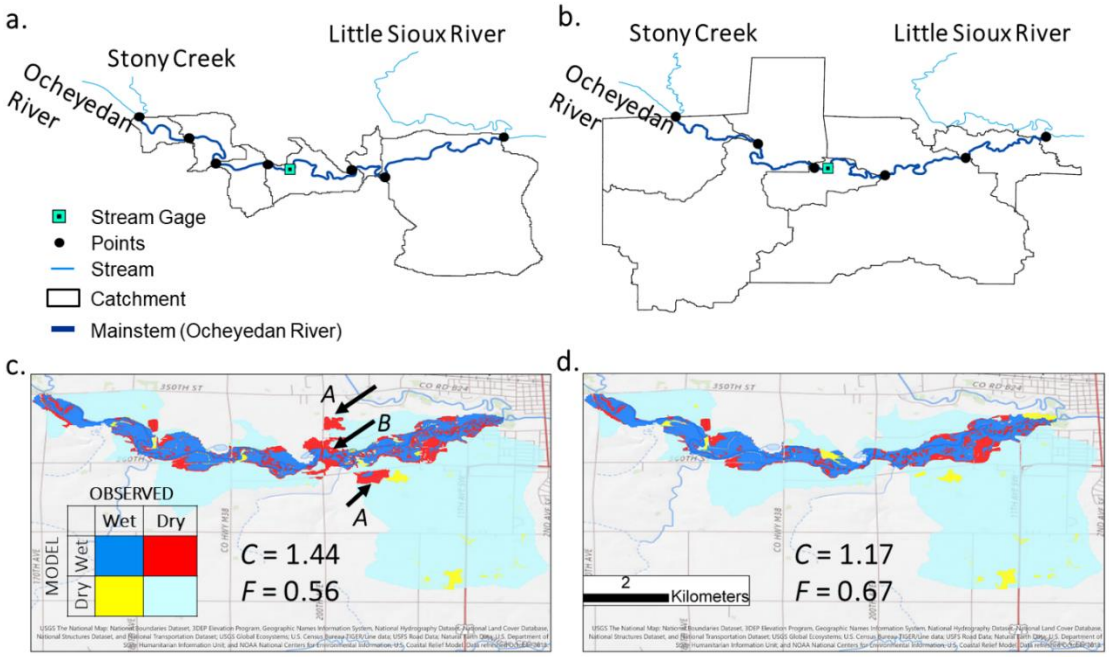

Figure 2.15. Location of nodes and associated catchments in (a) NDHPlus derived Continental-Scale Flood Inundation Mapping (CFIM) approach and (b) evenly distributed nodes to get uniform length reaches. Comparison of the result of the modeled (HAND-based) flood inundation with inundation classified from Planet Sentinel-2 imagery. (c) CFIM approach and (d) 3-m etch digital elevation model with all improvements involved.

CHAPTER 3

A COMPARISON OF NATIONAL WATER MODEL RETROSPECTIVE ANALYSIS SNOW OUTPUTS AT SNOTEL SITES ACROSS THE WESTERN U.S.[1]

Abstract

This study compares the U.S. National Water Model (NWM) reanalysis snow outputs to observed snow water equivalent (SWE) and snow-covered area fraction (SCAF) at SNOTEL sites across the Western U.S. SWE was obtained from SNOTEL sites, while SCAF was obtained from MODIS observations at a nominal 500 m grid scale. Retrospective NWM results were at a 1000 m grid scale. We compared results for SNOTEL sites to gridded NWM and MODIS outputs for the grid cells encompassing each SNOTEL site. Differences between modeled and observed SWE were attributed to both model errors, as well as errors in inputs, notably precipitation and temperature. The NWM generally under-predicted SWE, partly due to precipitation input differences. There was also a slight general bias for model input temperature to be cooler than observed, counter to the direction expected to lead to under-modeling of SWE. There was also under-modeling of SWE for a subset of sites where precipitation inputs were good. Furthermore, the NWM generally tends to melt snow early. There was considerable variability between modeled and observed SCAF as well as the binary comparison of snow cover presence that hampered useful interpretation of SCAF comparisons. This is in part due to the shortcomings associated with both model SCAF parameterization and MODIS observations, particularly in vegetated regions. However, when SCAF was aggregated across all sites and years, modeled SCAF tended to be more than observed using MODIS. These differences are regional with generally better SWE and SCAF results in the Central Basin and Range and differences tending to become larger the further away regions are from this region. These findings identify areas where predictions from the NWM involving snow may be better or worse, and suggest opportunities for research directed towards model improvements.

[1] Irene Garousi-Nejad and David G. Tarboton. (2021). A Comparison of National Water Model Retrospective Analysis Snow Outputs at SNOTEL Sites Across the Western U.S. Authorea. Dec 13, 2021. https://doi.org/10.22541/au.161656955.51617798/v2

3.1 Introduction

Accurate water supply forecasts will become increasingly crucial as western populations grow and demand more water, and as operational agencies have to manage water under global environmental change (Bhatti et al., 2016; Gergel et al., 2017; Li et al., 2017; Livneh & Badger, 2020; Mote, 2003; Mote et al., 2005; Regonda et al., 2005; Stewart et al., 2004, 2005). Many scientific challenges in understanding and preparing for global environmental change rest upon our ability to predict streamflow and snowmelt quantity, timing, and spatial patterns that are important for decision making in water-sensitive sectors. In the United States, the National Weather Service (NWS) of the National Oceanic and Atmospheric Administration (NOAA) is responsible for short- and long-term streamflow predictions across the U.S. Prior to 2016, NWS operational forecasts were limited to forecasts from NWS River Forecast Centers (RFC) at about 4000 forecast points. These were produced predominantly using the Sacramento soil moisture accounting model (SAC-SMA) to simulate runoff production and SNOW-17 model to simulate snowpack and snowmelt, within the Advanced Hydrologic Prediction System (AHPS, https://water.weather.gov/ahps/rfc/rfc.php) modeling infrastructure (McEnery et al., 2005).

While Franz et al. (2008) showed that SNOW-17 performed well over the Reynolds Creek Experimental Watershed located in southwestern Idaho, other studies found limitations such as being unable to capture snowmelt timing precisely due to its simple conceptual framework, its inability to represent spatial variability of land properties, and its dependence on extensive calibration for each basin using historical data (Lundquist & Flint, 2006; Shamir et al., 2006; Zalenski et al., 2017). Furthermore, a

National Research Council committee identified a gap between what is now considered state-of-the-art modeling capabilities and those used in AHPS (National Research Council, 2006). It concluded that the NWS needs to incorporate more advanced hydrologic science into their hydrologic models.

The increasing availability of distributed geographic data and computer power has made it possible to develop national/continental scale, physically-based, and distributed models. In 2016, NOAA's Office of Water Prediction implemented the National Water Model (NWM) as a physically-based distributed model based on the Weather Research and Forecasting Model Hydrological modeling system (WRF-Hydro) framework (Gochis, Barlage, Cabell, Casali, et al., 2020) to provide nationally consistent operational hydrologic forecasting capability. The main goals of the NWM were to provide forecast streamflow, produce spatially continuous countrywide estimates of hydrologic states (soil moisture, snowpack, etc.), and to implement a modeling architecture that permits rapid infusion of new data and science.

The NWM provides hourly flow forecasts at about 2.7 million locations in the U.S. In addition to the increased number of forecast locations, another advantage of the NWM is that it utilizes a specific configuration of the physically-based Noah-MultiParameterization (Noah-MP) land surface model to represent the land-atmosphere interactions including snow processes. There have been several studies evaluating results from the NWM. For instance, Viterbo et al. (2020) evaluated the prediction of flooding in NWM streamflow forecasts. They found that errors were due to both meteorological input errors as well as hydrologic process representation. In another study, Lahmers et al. (2019) improved the performance of WRF-Hydro configured as NWM version 1.1 by

implementing a conceptual channel infiltration function into the model architecture. They concluded that accounting for channel infiltration loss in the semi-arid Western U.S. improves the streamflow behaviour simulated when the model is forced with high-resolution precipitation input. However, we are not aware of a systematic and thorough evaluation of the NWM snow outputs.

The NWM (Gochis, Barlage, Cabell, Dugger, et al., 2020) has been running in NWS operations since 2016 to support operational flood forecasts. The latest operational version, version 2.0, was implemented in June of 2019. Prior to this operational deployment, the NWM version 2.0 retrospective analysis data were generated (by the NWM team) for investigations into the performance of the NWM. These are publicly available in Google Cloud Storage (National Weather Service, 2019).

These retrospective analysis results contain output from a 26-year simulation (January 1993 through December 2018), hereafter is referred to as NWM-R2. The meteorological forcing data used for the version-2 retrospective analysis configuration was drawn from the North American Land Data Assimilation System II (NLDAS2) datasets, a gridded product with spatial resolution of 1/8th-degree and hourly temporal resolution. The non-precipitation forcing fields in NLDAS2 are from the analysis fields of the National Centers for Environmental Prediction (NCEP)/North American Regional Reanalysis (NARR), i.e., a retrospective dataset, while the precipitation is from the gage-based NCEP/Climate Prediction Center (CPC). As a pre-processing step, the NWM team downscaled the NLDAS2 data and applied a mountain mapper (Hou et al., 2014) adjustment to the precipitation data to adjust the values for climatological variation due to topography and wind directions (RafieeiNasab et al., 2020). The result forcing dataset is a

1 km spatial resolution data layer for each hour which contains incoming short- and longwave radiation, specific humidity, air temperature, surface pressure, near surface wind, and precipitation rate. In terms of snow, outputs include gridded snow water equivalent (SWE), the amount of water stored in a snowpack, and the snow-covered area fraction (SCAF).

Across the Western U.S., snow is observed at 808 snow telemetry (SNOTEL) sites that provide data intended to quantify snow and inform water supply forecasts. Illustrative comparisons of NWM-R2 SWE to SNOTEL SWE (Figure 3.1) indicate that SWE is well modeled at some locations (Figure 3.1a) while significantly different from observations at other locations (Figure 3.1b). Accurate modeling of SWE is a necessary condition for accurate physically-based modeling of runoff. This motivated the need, addressed in this study, to systematically evaluate the performance of NWM-R2 simulations of SWE and SCAF against available SNOTEL measurements and the moderate resolution imaging spectroradiometer (MODIS) satellite imagery to answer the following questions:

- How well does the NWM model simulate snowpack (in terms of SWE, SCAF, and snowmelt timing) compare to observations over the entire Western U.S.?
- What are the potential causes responsible for discrepancies in NWM-R2 SWE, SCAF, and snowmelt timing?
- Are these discrepancies associated with the model input errors, the snow parameterization in the model, or measurement errors?

Answers to these questions are needed to further improve the NWM snow components, and ultimately runoff and water supply forecasts in snowmelt-dominated

regions. While U.S. based, the NWM is built using the WRF-Hydro modeling framework that has been applied worldwide, and the lessons learned from this comparison across the U.S. have application to the representation of snow processes in national and continental scale models throughout the world.

The following section—Model, Data, and Experimental Design—first presents a summary of the NWM-R2 snow parameterization. Then, it describes the datasets used in this study, comprised of the NWM-R2 reanalysis products, SNOTEL snow observations, and MODIS imagery giving the snow-covered area fraction. Next, it presents the metrics that were used for evaluating the model results versus observations. The results section compares the NWM-R2 SWE, precipitation, air temperature, SCAF, and presence or absence of snow with observations from SNOTEL and MODIS. It also compares modeled and observed snowmelt timing. We conclude with a discussion of the uncertainties and limitations in our analysis and present ideas for future work.

3.2 Model, Data, and Experimental Design

The study region comprises the SNOTEL sites across the Western U.S. (Figure 3.2a). The model is the NWM version 2.0 reanalysis (NWM-R2), that includes Noah-MP land surface components for snow. Data include NWM-R2 inputs and outputs, in-situ measurements, and remotely sensed data from MODIS for water years 2008-2018. NWM-R2 inputs that we used in our analysis were hourly NLDAS2-based precipitation, hourly NLDAS2-based air temperature, and elevation—derived from the 30 m Digital Elevation Model (Zhang et al., 2021)—with 1 km spatial resolution. We used NWM-R2 outputs of 3-hourly SWE and SCAF with 1 km spatial resolution from the land surface

module. We retrieved these inputs and outputs for NWM grid cells containing SNOTEL sites based on the nearest neighbor approach. In-situ measurements comprised daily precipitation, daily air temperature, elevation, and daily SWE from SNOTEL. Remotely sensed MODIS daily snow-covered areas with nominal 500 m spatial resolution were from the MODIS sensor. The model, in-situ, and remotely sensed datasets thus have different spatial resolutions (Figure 3.2b). The difference in scale is a potential source of uncertainty in our comparative analysis, and needs to be recognized in interpretation. There are small differences in elevation between SNOTEL (point elevations) and NWM-R2 (1 km grid elevations), that may impact temperature comparisons due to lapse rate effects, but there does not appear to be any significant bias (Figure 3.2c).

3.2.1 NWM-R2 Snow Parameterization (Noah-MP) and Snow Reanalysis Products

The NWM-R2 uses a particular configuration of Noah-MP (Table 3.1) as the land surface model to simulate snow processes as a 1-dimensional vertical column over 1 km spatial resolution grid cells with no representation of any lateral snow processes within a grid cell. Details of the NWM-R2 are given in WRF-Hydro version 5.1.1 documentation (Gochis, Barlage, Cabell, Casali, et al., 2020) and the code (Gochis, Barlage, Cabell, Dugger, et al., 2020). WRF-Hydro version 5.1.1 is the WRF-Hydro version used in NWM-R2. However, (Gochis, Barlage, Cabell, Casali, et al., 2020) does not describe details of the snow parameterization. Instead reference is made to the Noah-MP technical description (Yang et al., 2011) and associated paper (Niu et al., 2011). Here we have summarized key features of the snow parameterization that pertain to the interpretation of our results. The focus in this paper is on NWM-R2 results, practically amounts to a large-scale test of Noah-MP as configured for use in the NWM.

3.2.1.1 Snowfall

The separation of precipitation into rainfall or snowfall is based on Jordan's (1991) algorithm that uses near surface air temperature thresholds [Equations (1-2)].

$$f_{p,ice} = \begin{cases} 1.0 & T_{frz} + 0.0 \le T_{sfc} \le T_{frz} + 0.5 \\ 1.0 - (-54.632 + 0.2 \times T_{sfc}) & T_{frz} + 0.5 \le T_{sfc} \le T_{frz} + 2.0 \\ 0.6 & T_{frz} + 2.0 \le T_{sfc} \le T_{frz} + 2.5 \\ 0.0 & T_{frz} + 2.5 \le T_{sfc} > T_{frz} + 2.5 \end{cases} \quad (1)$$

$$\begin{aligned} rain &= P \times (1 - f_{p,ice}) \\ snow &= P \times f_{p,ice} \end{aligned} \quad (2)$$

where $f_{p,ice}$ is the snow fraction in precipitation, T_{sfc} [K] is the surface air temperature, T_{frz} [273.16 K] is freezing/melting point, and P [mm s^{-1}] is the input precipitation. Freshly fallen snow density (ρ_{fs} [kg/m^3]) is calculated using Equation (3), based on Hedstrom and Pomeroy (1998).

$$\rho_{fs} = \min\left(120, 67.92 + 51.25e^{\left(\frac{T_{sfc} - T_{frz}}{2.59}\right)}\right) \quad (3)$$

3.2.1.2 Vegetation and Snow Interception

In Noah-MP, a single-layer vegetation canopy model characterizes the fraction covered by vegetation (FVEG) in each model grid cell. Since the Noah-MP dynamic vegetation option is set off in NWM-R2, the model uses the maximum vegetation fraction from the Leaf Area Index (LAI) table as FVEG. If a model grid has a FVEG>0 and a snow depth greater than 0.025 m (from initial conditions or the last time step), the model computes the fraction of canopy buried by snow based on the snow depth and the canopy height. Then, the model uses this fraction to adjust the LAI and Stem Area Index (SAI), which are used in the snow interception model. The snow interception model allows for both liquid water and ice to be present on the vegetation canopy; and includes loading/unloading of snowfall, melting of intercepted snow and refreezing of the

meltwater, frost/sublimation of canopy-intercepted snow, and dew/evaporation. The model solves the canopy liquid water balance [Equation (4)] and ice balance [Equation (5)] based on Niu and Yang (2004).

$$\frac{\partial M_{liq}}{\partial t} = R_{intr} + (R_{dew} - R_{eva}) + (R_{melt} - R_{frz}) \quad (4)$$

$$\frac{\partial M_{ice}}{\partial t} = (R_{load} - R_{unload}) + (R_{frost} - R_{sub}) + (R_{frz} - R_{melt}) \quad (5)$$

where M_{liq} [kg m^{-2}] is the storage of liquid water in the canopy, and R_{intr} [kg m^{-2} s^{-1}], R_{dew} [kg m^{-2} s^{-1}], and R_{eva} [kg m^{-2} s^{-1}] are interception rate for rain, dew rate, and evaporation rate, respectively. R_{melt} [kg m^{-2} s^{-1}] and R_{frz} [kg m^{-2} s^{-1}] are melting and refreezing rates. M_{ice} [kg m^{-2}] is the storage of ice in the canopy and R_{load} [kg m^{-2} s^{-1}] and R_{unload} [kg m^{-2} s^{-1}] are snow loading and unloading rates, respectively. R_{frost} [kg m^{-2} s^{-1}] and R_{sub} [kg m^{-2} s^{-1}] are frost and sublimation rates. Heat transported by snow and rain to the vegetation canopy layer, the vegetated ground, and non-vegetated ground is also computed; and is used later in the energy balance computation.

3.2.1.3 Snow-Covered Area and Snow Albedo

Noah-MP calculates SCAF based on snowpack density (ρ_{sno} [kg m^{-3}]), snow depth (h_{sno} [m]) from initial conditions or the previous time step, snow surface roughness length ($z_{0,g}$ [m]), density of fresh snow (ρ_{new} [kg m^{-3}]), and a dimensionless area-depth factor (m) that determines the curve relating SCAF and snow depth [Equation (6)] as developed by Niu and Yang (2007).

$$SCAF = \tanh\left(\frac{h_{sno}}{2.5 z_{0,g}\left(\frac{\rho_{sno}}{\rho_{new}}\right)^{m}}\right), \qquad \rho_{sno} = \frac{SWE}{h_{sno}} \quad (6)$$

In NWM-R2 calculations of snow-covered area, ρ_{new} and $z_{0,g}$ are constants set equal to 100 kg m^{-3} and 0.002 m, respectively. However, the factor m is among the parameters that are adjusted during calibration to minimize differences between modeled and observed streamflow over calibration watersheds (Lahmers et al., 2019; RafieeiNasab et al., 2020). The functional relationship between SCAF and depth quantifies small-scale variability of snow within a computational grid element which plays an important role in the process governing snow accumulation and ablation. SCAF is used to weight the ground emissivity and ground surface resistance. It also affects the computed snow surface albedo that is modeled using the Biosphere-Atmosphere Transfer Scheme (BATS). BATS (Yang & Dickinson, 1996) models direct and diffusive radiation in visible and near-infrared bands separately accounting for fresh snow albedo, snow age, grain size growth, impurity, and solar zenith angle.

3.2.1.4 Surface Energy Balance, Radiation, and Momentum Fluxes

Shortwave radiation is modeled over the entire grid cell using a modified two-stream approximation (Niu & Yang, 2004) treating the vegetation as evenly distributed with gaps. The result is canopy-absorbed and ground-absorbed solar radiation over the grid cell. Longwave radiation, latent heat, sensible heat, and ground heat fluxes are modeled, using a tile approach that treats vegetated and bare fractions of the cell separately (Niu et al., 2011). Noah-MP treats turbulence fluxes between the snowpack, vegetation canopy, and air using Monin-Obukhov similarity theory to model atmospheric stability conditions. Stability corrections of under canopy turbulent transfer account for the strong stable condition of a warmer canopy overlying the snow surface during the melt season (Fei Chen et al., 2014). Precipitation advected heat is also computed

separately for the canopy vegetation, vegetated ground surface, and non-vegetated ground surface. The vegetation canopy temperature (T_v), the vegetated ground surface temperature ($T_{g,v}$), and the non-vegetated ground surface temperature ($T_{g,b}$) are estimated using the Newton-Raphson method with 20 iterations. If the snow depth is greater than a specified snow depth ($\geq$ 0.05 m) and the ground surface temperature ($T_{g,v}$/$T_{g,b}$) is greater than the freezing point (273.16 K), the ground temperature is updated to $(\mathbf{1}-\boldsymbol{SCAF}) \times \boldsymbol{T_g} + \boldsymbol{SCAF} \times \boldsymbol{T_{frz}}$, and all turbulent fluxes are reevaluated. Finally, these radiative and turbulent fluxes are then aggregated based on the vegetated fraction (FVEG) parameter.

3.2.1.5 Snowpack Vertical Discretization and Snow Thermal Properties

The Noah-MP snow module uses up to three snow layers, depending on depth (from initial conditions or the last time step). The state variables for each layer are the mass of liquid water, mass of ice, layer thickness, and layer temperature. Snow can also exist in the model without being represented by explicit snow layers. This occurs when the total snowpack thickness is less than a specified minimum snow depth (< 0.025 m). In this case, the only state variable is the mass of snow.

Snow thermal properties including partial volume of ice, partial volume of liquid water, effective porosity, bulk density [based on Lynch-Stieglitz (1994)], volumetric specific heat, and thermal conductivity are computed for each snow layer [Equations (7-12)]. Energy for phase change (melting/refreezing) is also computed for each layer.

$$\theta_{ice,i} = \frac{Mass_{ice,i}}{\Delta Z_i \times \rho_{ice}} \tag{7}$$

$$\theta_{e,i} = 1 - \theta_{\mathrm{ice,i}} \tag{8}$$

$$\theta_{liquid,i} = \min\left(\theta_{e,i}, \frac{Mass_{liquid,i}}{\Delta Z_i \times \rho_{water}}\right) \quad (9)$$

$$\rho_{snow,i} = \frac{Mass_{ice,i} + Mass_{liquid,i}}{\Delta Z_i} \quad (10)$$

$$C_{v,i} = C_{ice} \times \theta_{ice,i} + C_{liquid} \times \theta_{liquid,i} \quad (11)$$

$$k_i = 3.2217 \times 10^{-6} \times {\rho_{snow,i}}^2 \quad (12)$$

where $\theta_{ice,i}$ [m^{-3}/m^{-3}] is partial volume ice of snow layer i, $Mass_{ice,i}$ [kg m^{-2}] is snow ice mass of snow layer i, ΔZ_i [m] is the snow layer thickness of snow layer i, ρ_{ice} [917 kg m^{-3}] is ice density, $\theta_{e,i}$ [m^{-3}/m^{-3}] is the effective porosity of snow layer i, $\theta_{liquid,i}$ [m^{-3}/m^{-3}] is partial volume of liquid water of snow layer i, $Mass_{liquid,i}$ [kg m^{-2}] is liquid water mass of snow layer i, ρ_{water} [1000 kg m^{-2}] is liquid water density, $\rho_{snow,i}$ [kg/m^{-3}] is bulk density of snow layer i, $C_{v,i}$ [J m^{-3} K^{-1}] is volumetric specific heat of snow layer i, C_{ice} [2.094×10^6 J m^{-3} K^{-1}] is specific heat capacity of ice, C_{liquid} [4.188×10^6 J m^{-3} K^{-1}] is specific heat capacity of liquid water, and k_i [W m^{-1} K^{-1}] is thermal conductivity of snow layer i.

Heat flux between layers is calculated based on temperature gradient and thermal conductivity, and then this is used to update layer temperatures using a semi-implicit numerical scheme. When heat flux calculations result in temperatures of snow layers greater than freezing, the excess energy is used to adjust (melt or freeze) liquid water present. The change in the density of the snow with time due to destructive metamorphism, the weight of the overlying layers of snow, and melting (which dictates layer thickness) is modeled, following Anderson (1976) as a function of snow temperature (Niu et al., 2011).

3.2.1.6 Snow Water Equivalent and Snow Depth

The change in SWE is balanced by the input snowfall (Q_{snow}) reaching the surface in forms of drip and throughfall; and output snowmelt (M), snow sublimation, and snow frost [both expressed as E in Equation (13)].

$$\frac{dSWE}{dt} = Q_{snow} - M - E \quad (13)$$

When new snowfall occurs in a time step, the snow depth and snow ice are increased based on the snow depth increasing rate and the input snowfall rate (both outputs of the snow interception module), respectively. After the depth, phase change and compaction calculations, the number of snow layers is adjusted by either combining the adjacent layers or subdividing them following Jordan (1991). If rainfall (in terms of drip and throughfall) occurs, it is added to the liquid water of the snow layer. The liquid water movement within a snow layer is added to the underlying snow layer when the liquid water content within a snow layer exceeds the layer's liquid water-holding capacity for snowpack (0.03 m^3/m^3). Finally, the liquid water of the snow layer updates after the water flows out of the layer.

3.2.1.7 Post-processing NWM-R2 Snow Reanalysis Products

This study used the NWM-R2's land surface model outputs, which are geospatial gridded results with a spatial resolution of 1 km and temporal resolution of 3-hours. We obtained the NWM-R2 SWE (model code name: SNEQV) and SCAF (model code name: FSNO) for grid cells containing SNOTEL sites based on the nearest neighbour approach [code available at Garousi-Nejad and Tarboton (2021d)] from the NOAA Google Cloud archive using a Jupyter Notebook [code available at Tarboton and Garousi-Nejad (2021)]. Then, we averaged 3-hourly results to daily values [code available at Garousi-Nejad and

Tarboton (2021f)] to have a similar temporal resolution when comparing the NWM-R2 results with SNOTEL and MODIS observations because both these datasets produce daily data. We also obtained the hourly precipitation, hourly air temperature, and elevation input data used for NWM-R2 simulations for the selected grid cells. The WRF-Hydro team at NCAR provided precipitation and air temperature values for us as those data were not available on the Google Cloud archive. Then, we computed daily precipitation and the daily average temperature [code available at Garousi-Nejad and Tarboton (2021f)].

3.2.2 SNOTEL

SNOTEL stations, managed by the Natural Resources Conservation Service (NRCS), generally consist of a snow pillow, an air temperature sensor, and a storage precipitation gage. Our study used the daily precipitation, air temperature, and SWE values measured at SNOTEL sites as a reference dataset to evaluate the NWM-R2 precipitation, air temperature, and SWE. We realize that SNOTEL data must be used with some caution because the sites are mostly located in small clearings within forests protected by forest canopies, leading to differences in exposure to wind and radiation (McCreight et al., 2014). Furthermore, SNOTEL data do not undergo a high correction level (Swenson & Lawrence, 2012). In some instances, we found unrealistically high temperature values that needed to be filtered out. Nevertheless, SNOTEL data remain the only widespread in situ SWE observations available for model validation in the Western U.S. (Barlage et al., 2010; Clow et al., 2012; Livneh et al., 2010; Pan et al., 2003; Toure et al., 2016). We automated retrieval of the SNOTEL data by calling its Consortium of

Universities for the Advancement of Hydrologic Science, Inc (CUAHSI) web service from a Jupyter Notebook script (Garousi-Nejad & Tarboton, 2021c).

3.2.3 MODIS

The National Aeronautics and Space Administration (NASA)'s MODIS instrument launched aboard the Terra satellite in late 1999 is designed to observe and monitor Earth changes, such as snow cover. MODIS has spectral bands in the visible and near-infrared regions, nominal 500 m spatial resolution, and near-daily global coverage. The daily snow-cover gridded tile product, MOD10A1, has been used and improved over time in multiple snow studies (Aalstad et al., 2020; Bennett et al., 2019; Magand et al., 2014; Masson et al., 2018; Salomonson & Appel, 2006; Swenson & Lawrence, 2012). We used products from the current version of the MODIS snow-cover algorithm which is the collection 6 suite of MODIS (hereafter referred to as MODIS-C6, or just MODIS). We chose to use MODIS-C6 (Hall & Riggs, 2016) as a reference to evaluate NWM-R2 SCAF because the improvements/revisions to MODIS-C6 (i.e., accounting for the surface temperature and surface height) led to a notable increase in accuracy of snow cover detection on mountain ranges and low illumination conditions in the Northern Hemisphere during spring and summer (Riggs et al., 2017).

The MODIS-C6 snow algorithm is designed to detect snow cover based on the normalized ratio of the differences in reflectance in band 4 (centred at 0.56 μm, visible green) and band 6 (centred at 1.64 μm) of the MODIS instrument with revisions applied to alleviate snow detection commission errors (reported for previous versions) for which snow detection is uncertain. The MODIS-C6 products include this ratio, the Normalized Difference Snow Index (NDSI, product name: NDSI_Snow_Cover) rather than snow

cover. This approach allows users to have the option to estimate snow cover using the global empirical model [Equation (14)] or develop region-specific models (Riggs et al., 2016). In this study, we developed a script (Garousi-Nejad & Tarboton, 2021b) run in Google Earth Engine to retrieve NDSI_Snow_Cover for each NWM grid cell containing a SNOTEL site. Since MODIS output is available on a 500 m grid and NWM grid cells are 1 km in size, the script averaged NDSI_Snow_Cover over the four MODIS grid cells that have their centroid within the NWM grid cell (Figure 3.2). Valid NDSI_Snow_Cover values range between 0-100 with values above 100 indicating missing data, no decision, night, inland water, ocean, cloud, and detector saturated issues, which we masked out in Google Earth Engine. The returned MODIS images thus have spatial gaps due to this masking. We filled gaps in each image with NDSI_Snow_Cover from the most previous valid value (forward filling). Then, we applied the globally-determined linear model of Riggs et al. (Riggs et al., 2016) to compute MODIS SCAF from NDSI_Snow_Cover values [Equation (14)].

$$SCAF = \min[\max(-0.01 + 1.45 \times NDSI, 0), 1] \quad \text{where } NDSI \in [0,1] \tag{14}$$

In Equation (14), the MODIS SCFA is always estimated as 1 for NDSI values equal or greater than 0.7, and it changes linearly for NDSI values between 0 to 0.7.

The resulting dataset includes 2,504,102 site-days in the period of overlap between NWM-R2 and SNOTEL data [data and code used to aggregate it are available at Garousi-Nejad and Tarboton, (2021e)]. We organized the SNOTEL sites into subgroups using Omernik Ecoregions level III (Omernik & Griffith, 2014) available from the Commission for Environmental Corporation (http://www.cec.org/north-american-environmental-atlas/terrestrial-ecoregions-level-iii/) to identify regional differences in

model results versus observations. The ecoregions are areas with general similarities in location, climate, vegetation, hydrology, terrain, wildlife, and land use; and have been used in multiple prior studies (Sun et al., 2019; Trujillo & Molotch, 2014).

3.2.4 Metrics

We used several metrics to compare NWM-R2 snow water equivalent (SWE), snow covered area fraction (SCAF), precipitation (P), and snowmelt timing against SNOTEL SWE and MODIS-C6 SCAF.

Seasonal:

- First day of the month comparisons were used for NWM-R2 SWE/SCAF (modeled) versus SNOTEL SWE and MODIS SCAF (observed) for months Nov-Jun.
- Monthly precipitation and average air temperature were also compared for these months.

These monthly comparisons let us evaluate the seasonal variability of snow in both modeled and observed datasets for data in the period of overlap between NWM-R2 and SNOTEL data.

Snow Water Equivalent and Snow-Covered Area at peak SWE:

- Modeled and observed SWE and SCAF were compared on the date of observed peak SWE (same day comparison).
- Modeled and observed peak SWE do not necessarily occur on the same date. We compared both SWE and SCAF on the separate dates where peak SWE was modeled and observed (different day comparison).

- Model input and SNOTEL observed total precipitation accumulated from the start of the water year, Oct 1, to the date of peak SWE were also compared.

Total precipitation was computed to assess the degree to which differences may be attributable to precipitation differences. This was done for both same day (observed peak SWE) and different day (observed and modeled peak day) comparisons. The different peak day comparison addresses the possibility that peak modeled and observed SWE may be close, but appear further apart in same day comparisons due to a timing mismatch.

Direct (binary) comparison of snow presence or absence:

- Full snow cover. Daily modeled SCAF taken as full snow if SCAF is $\geq$ 0.95. Daily MODIS inferred (observed) SCAF taken as full snow if NDSI is $\geq$ 0.7.
- Some snow cover. Daily SCAF taken as indicating some snow if modeled SCAF, or MODIS NDSI > 0.3.

First, we classified the snow presence or absence grid cells based on these thresholds. We then counted the number of classified grid cells for both observed and modeled datasets for each date. This was done only for grid cells locations where SNOTEL sites exist, because our scripts extracting NWM output were only run at these locations and running for all grid cells across the Western U.S. was computationally prohibitive.

- Presence Absence comparison metrics were used to indicated the degree-of-overlap between modeled and observed datasets (Horritt & Bates, 2002; Sangwan & Merwade, 2015).

The correctness metric [Equation (15)] compares the total number of modeled and observed grid cells having some or full snow cover, while the fit metric [Equation (16)] quantifies whether modeled and observed locations match, scaled by the total area mapped with snow (either full or some).

$$C_t = \frac{\text{Modeled}_{\text{snow}}}{\text{Observed}_{\text{snow}}} \tag{15}$$

$$F_t = \frac{\text{Modeled}_{\text{snow}} \cap \text{Observed}_{\text{snow}}}{\text{Modeled}_{\text{snow}} \cup \text{Observed}_{\text{snow}}} \tag{16}$$

where C_t and F_t are correctness and fit metrics computed for date t, respectively, and $\text{Modeled}_{\text{snow}}$ and $\text{Observed}_{\text{snow}}$ are grid cells classified as snowy cells on that date. Correctness (C_t) and Fit (F_t) should both ideally be 1 (100%).

To account for the fact that MODIS may be interpreting vegetated grid cells as snow free and thus underestimating the snow cover (Steele et al., 2017; X. Wang et al., 2017), while NWM-R2 may have snow beneath the vegetation canopy, and that SNOTEL sites are often in openings much smaller than the cell size (1 km) in generally forested areas, we requested, and obtained from the NRCS (the agency that operates SNOTEL) a list of sites in generally open areas. We report separate metrics for these sites reported to be open. The NRCS indicated that SNOTEL sites may be open due to canopy disturbance caused by pine bark beetle damage and fire, which may have occurred during the study period, resulting in some uncertainty as to sites being open early on.

Melt timing

- Half melt from peak SWE date (Clow, 2010).

The date, when half the snowpack has melted serves as a measure of melt timing somewhat robust to small fluctuations or a long period where SWE is flat near the peak. We categorized the differences between observed and modeled half melt dates as close (within 5 days), model early (the model is 6 to 19 days ahead of observed), model late (the model is 6 to 19 days after observed), and far apart (the modeled and observed differ by 20 days or more).

Commonly used statistics:

- Coefficient of determination [r^2, Equation (17)] that ranges from -1 to 1 with 1 indicating a perfect positive linear relationship but insensitive to proportional differences between modeled and observed data;
- Spearman's rank correlation [Spearmanr, Equation (18)], a non-parametric measure of correlation used to measure the strength of association between modeled and observed values where value 1 means a perfect positive correlation;
- Root mean square error [RMSE, Equation (19)], a measure of how concentrated the data are around the line of best fit;
- Nash Sutcliffe efficiency [NSE, Equation (20)], a normalized statistic that determines the relative magnitude of the residual variance compared to observed values ranging from -∞ to 1 with 1 indicating observed and modeled data fits the 1:1 line; and

- Bias [Bias, Equation (21)], the average of the difference between modeled and observed.

$$r^2 = \left[\frac{\sum_{t=1}^{N} (M_t - \bar{O}_t)(M_t - \bar{M}_t)}{\sqrt{\sum_{t=1}^{N} (O_t - \bar{O}_t)^2 \sum_{t=1}^{N} (M_t - \bar{M}_t)^2}} \right]^2 \quad (17)$$

$$\text{Spearmanr} = 1 - \frac{6 \sum_{t=1}^{N} d_t^2}{N(N^2 - 1)} \quad (18)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{N} (O_t - M_t)^2}{N}} \quad (19)$$

$$\text{NSE} = 1 - \frac{\sum_{t=1}^{N} (O_t - M_t)^2}{\sum_{t=1}^{N} (O_t - M_t)^2} \quad (20)$$

$$\text{Bias} = \frac{\sum_{t=1}^{N} (M_t - O_t)}{N} \quad (21)$$

where M_t is model simulation, O_t is observation, N is the total number of simulations or observations, d_t is difference between observed and modeled rank, and the overbar indicates average.

3.3 Results

3.3.1 Seasonal (Monthly) Comparison

We compared the NWM-R2 SWE results with observations from SNOTEL and found a persistent bias in modeled SWE across most months (Figure 3.3). Results show that throughout the accumulation phase (Nov-Feb), the rank correlation between

observed and modeled SWE increases (Spearmanr from 0.7 to 0.8). However, this does not necessarily indicate an acceptable model performance. The discrepancies between the observed and modeled SWE increase as snow accumulates (RMSE 21 to 135 mm). In the ablation phase (Mar-Jun), the rank correlation decreases, and discrepancies are highest in May (Bias -149 mm, RMSE 292 mm). The increasing scatter in later months (Figure 3.3) shows that the NWM generally performs well during the accumulation phase but simulates SWE less well during the ablation phase. Most points fall below the 1:1 line (red line). The points clustered into vertical and horizontal lines on the bottom and left axes of scatter plots in May and Jun indicate early and late modeling of complete melt out, respectively.

The comparison between the NWM-R2 SCAF and estimates from MODIS-C6 revealed that the modeled SCAF is highly uncorrelated with what is detected by satellite imagery (Figure 3.4). Throughout the last three months of the accumulation phase (Dec-Feb), the NWM results show that more than 70% of points (each representing one NWM grid cell that includes a SNOTEL site and a water year) have SCAF 0.9-1, while less than 10% have SCAF 0-0.1 (histograms in Figure 3.4). In contrast to the binary behaviour of the NWM-R2 SCAF, MODIS SCAF exhibits gradual increases and decreases. At most, 30% of the observed data have SCAF values ranging from 0.9-1 during the accumulation phase. In December, 14% of the observed data have SCAF greater than 0.9, while about 70% of modeled points have SCAF greater than 0.9. During the ablation phase (Mar-Jun), both modeled and observed datasets have relatively a similar data percentage with SCAF less than 0.1. However, the portion of the points where modeled SCAF is above

0.9 is still much more significant (3-7 times depending on the month) than those in the observed dataset (histograms in Figure 3.4).

The SCAF comparisons above are only at SNOTEL sites. We did not undertake the computation needed to compare NWM-R2 and MODIS-C6 for all grid cells and dates. However, as an illustration for locations beyond SNOTEL sites NWM-R2 and MODIS-C6 SCAF maps on Dec 1, 2011 (Figure 3.5) show that while patterns are generally the same, MODIS SCAF seems less than modeled. Note that the MODIS-C6 SCAF map (Figure 3.5a) has gaps and cloud areas (grey) that we did not fill in from the most recent previous image with data (as described in Section 3) for this visualization. NWM-R2 SCAF covers the entire region selected based on the MODIS tiles. The visual comparison of a zoomed-in map for the region where observed SCAF were available for more than 90% of the area reveals both similarities and differences between NWM-R2 and MODIS-C6 datasets (Figure 3.5c and 3.5d). The NWM-R2 SCAF map for the zoomed-in area shows more white regions (i.e., SCAF values greater than 0.9), suggesting that NWM tends to overestimate SCAF compared to observations from MODIS.

Scatterplots of monthly precipitation (Figure 3.6) indicate model input precipitation generally less than measured at SNOTEL sites, possibly contributing to under-modeling of SWE (Figure 3.3). Spearmanr and NSE values show an acceptable correlation between modeled and observed monthly precipitation (on average, 0.8 for both statistics). However, the precipitation bias is larger during the accumulation phase than the ablation phase, suggesting that increased SWE scatter, in the ablation phase, is

less associated with precipitation input errors than other factors during the ablation phase snowmelt.

Elevation, through orographic effects, is often suspected as a contributor to precipitation bias. However, the comparison of model input elevation (1 km grid cell) with SNOTEL point elevation (Figure 3.2) indicated no bias and small scatter (r^2=0.98 in Figure 3.2c). There are, nevertheless, discrepancies between the NWM-R2 monthly averaged air temperature inputs and the monthly averages of the daily mean air temperature measured at SNOTEL sites (Figure 3.7), reported as the 24-hour average of a minimum four samples per hour (U.S. Department of Agriculture, 2011). NWM-R2 air temperatures are generally slightly below observations. This is counter to the direction needed to explain discrepancies in SWE as colder model input air temperatures should result in (1) greater fractions of precipitation as snowfall and (2) slower rather than quicker snowmelt, both processes that increase rather than decrease SWE.

The seasonal pattern of SWE and SCAF averaged across all SNOTEL site years for each specific day (Figure 3.8) further indicates the general under modeling of SWE and over modeling of SCAF relative to SNOTEL and MODIS observations, respectively.

Discrepancies between the seasonal pattern of SWE and SCAF are regional and somewhat different for SWE than SCAF (Figure 3.9 and Figure 3.10, respectively). The NWM SWE was better in the Klamath Mountains, Blue Mountains, and Central Basin and Range (region 9, 2, and 5, respectively, in Figure 3.9) with SWE bias differences tending to become larger further to the north and east across the study region. However, the NWM SCAF are closer to the observations in the Northern Basin and Range, Sierra Nevada, and Central Basin and Range regions (regions 12, 13, and 5, respectively, in

Figure 3.10), with SCAF differences tending to become larger the further away regions are from the Central Basin and Range region.

3.3.2 Observed Peak SWE (Same Day and Different Day) Comparison

The scatterplot of modeled versus observed SWE on the date of peak observed SWE (Figure 3.11a) indicates a general downward bias in modeled SWE. NWM SCAF clusters around 1 on this date (histograms in Figure 3.11b) while MODIS SCAF is more fractional, and similar to monthly SCAF the point comparisons are scattered and poor. Precipitation accumulated from Oct 1 to the date of observed peak SWE indicates model input precipitation generally less than SNOTEL observed (Figure 3.11c: Bias -111 mm, RMSE 212 mm). This suggests that under estimation of model precipitation inputs may be a contributor to under modeling of peak SWE. This comparison may also be influenced by the fact that observed SWE is at its peak, but modeled SWE is not.

We also compared observed and modeled peak SWE, noting that these do not necessarily occur on the same date (Figure 3.12). Results are similar to the observed peak SWE date comparison. Here the accumulated observed and modeled precipitation (Figure 3.12c) are over the accumulation period, to their respective peak SWE dates, a possible reason for increased scatter and poorer error metrics in this figure.

Under modeling of SWE is also evident when comparing the observed and modeled peak SWE for a subset of SNOTEL sites where the model precipitation is relatively close to the observed (Figure 13b: Bias -96 mm, RMSE 168 mm). However, the errors are less than for the entire dataset SWE comparison. We chose this subset of sites based on the NSE measure between daily model input and observed precipitation being greater than or equal to 0.9 computed over the full study period. This subset shows

a reduced bias (compared to the entire dataset) between the observed and modeled precipitation accumulated from Oct 1 to peak observed SWE date (Figure 3.13a).

3.3.3 Direct (Binary) Comparison of Snow Presence or Absence

The cell by cell binary comparison of snowy grid cells at SNOTEL sites shows that this comparison does not work well for the all-snow-present condition, i.e., when the observed and modeled SCAF thresholds were 0.7 and 0.95, respectively (Figure 3.14a). We observed that the average C for the entire period of study was 9.4 and average F, 0.11. These are poor degree of overlap statistics, and are due to the fact that MODIS never reports more than about 30% of the area as having full snow.

However, the cell by cell binary evaluation for some snow present resulted in better degree of overlap statistics (Figure 3.14b, $\bar{C}$=1.47 and $\bar{F}$=0.50). Discrepancies between the modeled and observed snowy grid cells as implied by average C (=1.20) and F (=0.64) were even less when we only focused on the 62 SNOTEL sites (about 8% of all sites) reported as open (Figure 3.14c). Table 3.2 summarizes fit metrics for the snow cover binary comparison.

3.3.4 Melt Timing Comparison

For 68% of the site years analyzed, the modeled half melt date was earlier than observed. When further classified based on whether modeled half melt dates were close, ahead, behind or far apart from observed melt dates (Figure 3.15a) we observe that the NWM half melt date was greater than 20 days from observed half melt date, for 34% of the site years, and off by 6 days or more for 75% of site years. For those site years where the difference was between 5 and 20 days, a greater percentage had the model melting ahead, than behind the observed. The site years that have modeled half melt date ahead of

observed tend to have lower modeled half melt date SWE (which is by definition half the peak SWE) than observed (Figure 3.15b).

3.4 Discussion

The seasonal pattern of SWE and SCAF averaged across all SNOTEL site-years shows that NWM generally under-estimates SWE and over-estimates SCAF relative to SNOTEL and MODIS observations, respectively. These discrepancies vary regionally with relatively better SWE results in the Arizona/New Mexico Mountains, Blue Mountains, and Central Basin and Range ecoregions; and better SCAF results in the Central Basin and Range and Sierra Nevada ecoregions tending to become larger the further away regions are from the Central Basin and Range. There are several sources of uncertainties in our comparisons that need to be pointed out. The spatial scale differences in different datasets is a source of uncertainty in this analysis. A point-scale measurement of SWE cannot with confidence validate the NWM-R2 grid cell value with nearest center, particularly in forest regions (McCreight et al., 2014). We realize that using other approaches, such as bilinear or cubic interpolation of NWM grid values would give different values at each SNOTEL site, a question we did not explore. In the cell by cell comparison between NWM-R2 and MODIS-C6 datasets, the mean value of MODIS grid cells would be different if using a different number of cells, e.g. nine grid cells instead of four.

Precipitation discrepancies suggest that SWE differences are partly due to discrepancies between observed precipitation (SNOTEL) and model input precipitation [adjusted NLDAS-2 (RafieeiNasab et al., 2020)]. There are multiple possible sources of uncertainty that may lead to this difference. First, SNOTEL latitude and longitude

locations may not be precise in the geographic information from SNOTEL, as, for site security, exact site locations may not be reported. This may result in selecting a non-representative 1 km NWM grid cell. Second, there may be systematic bias for gage precipitation, particularly with snowfall measurements being subject to “under-catch” (Mote, 2003; Sun et al., 2019). However, we note that model input precipitation was typically less than measured at SNOTEL sites, indicating that if under-catch is an issue, it may be larger in the data used to produce model inputs. In NWM version 2.0, a mountain mapper adjustment has been applied to obtain input precipitation from NLDAS-2 (RafieeiNasab et al., 2020); nevertheless, there are still differences and biases compared to SNOTEL measurements that may be impacting model results. Third, SNOTEL data do not undergo a high correction level (Swenson & Lawrence, 2012). It was not uncommon to see accumulated precipitation less than SWE at SNOTEL sites (notably for stations at higher elevations), which could be due to either precipitation under-catch, or inflated SWE (Meyer et al., 2012). This makes using this information for model comparison challenging, as the model cannot accumulate more snow than its precipitation input. This is an unresolvable difference and should be recognized as a source of uncertainty associated with the in-situ measurements used in this study.

Our results show a cold (downward) bias for the model input air temperature (based on NLDAS-2) compared to SNOTEL sites' observations. This is different from Naple et al. (2020), who reported a warm (upward) bias for the NWM retrospective runs compared to the New York State Mesonet observations. The cold bias in the model temperature input is counter to the direction expected to lead to the under-modeling of SWE, a point which needs more investigation.

The discrepancies in model inputs (precipitation and air temperature in this study) are not the only potential sources for SWE differences. Even at sites with statistically highly correlated precipitation input (NSE>0.9), the results indicate that some SWE bias, potentially due to other factors, still remains. This opens up the question as to whether there are other deficiencies that lead to SWE under-modeling, both due to observation and model errors. Errors in SWE measurements may occur, due to factors such as wind causing snowdrifts on the snow pillow (Meyer et al., 2012), or the small clearing SNOTEL site location not being representative of larger scale snowpack (McCreight et al., 2014). In the NWM land surface model (Noah-MP), the partitioning of precipitation into rainfall and snowfall, which is one of the most sensitive parameterizations in simulating cold-region hydrological processes (Loth et al., 1993), is based on Jordan's (1991) algorithm, which ignores some physical processes controlling precipitation phase by not incorporating humidity. This may lead to biases in SWE, snow depth, and snow cover fraction (Feng Chen et al., 2014; Harder & Pomeroy, 2014; Y. Wang et al., 2019). Y. Wang et al. (2019) suggest that using a snow-rain partitioning scheme based on the wet-bulb temperature within Noah-MP produces more snowfall and snow mass on the ground that agrees better with ground-based snow observations, particularly over mountainous regions in the Western U.S. Recently, Naple et al. (2020) shows that using the precipitation phase partition from the high-resolution rapid refresh (HRRR), in lieu of the operational method (Jordan, 1991), leads to improved snow results for the NWM version 2.0 configuration.

Our results show that, on average, the NWM tends to melt snow early (6-19 days) compared to SNOTEL observation. For 75% of the site years, the modeled date of half

melt from peak SWE was off by 6 days or more from the observed half melt dates, sometimes being as far apart as 2 months (for example, Magic Mountain SNOTEL site, ID: 610 in Idaho, at water year 2010). This suggests that the modeling of melt timing is somewhat problematic and there is a need to further investigate overall energy balance and snow surface temperature, possibly drawing on ideas from the Utah Energy Balance model (Mahat & Tarboton, 2014; You et al., 2014).

Overall, NWM-R2 SCAF was difficult to compare to MODIS-C6 SCAF using single SNOTEL sites and days. Some of this difficulty—manifested in the scatter in Figures 3.4, 3.11, and 3.12—may reflect the fact that the MODIS and NWM SCAF quantities are not really the same thing. MODIS may be interpreting vegetation as snow free (Steele et al., 2017; X. Wang et al., 2017), while NWM has snow beneath vegetation. In NWM-R2 results, the persistent low and high SCAF (<0.1 and >0.9, respectively) reflects that NWM treats SCAF as a binary metric in mountainous regions. NWM-R2 SCAF values stay near 1 with less variability between Dec-Apr for more than 70% of cases. This suggests that once the NWM grid cell (1 km spatial resolution) is more than 90% snow-covered, it is implausible for it to diverge from 1 for the rest of the accumulation phase and early ablation phase. One possible reason for this behaviour is the lack of representation of some factors affecting SCAF such as vegetation type and seasonal change, and topography. These limitations affect the accurate simulation of SCAF and SWE (Helbig et al., 2015; Magand et al., 2014; Swenson & Lawrence, 2012; Wrzesien et al., 2015). Another possible reason for some of the differences is the lack of any representation of snow drifting processes (i.e., wind-driven redistribution of snow) in the snow model. Snow drifting increases the variability of snow depth within a grid cell,

which then, when melting starts leads to intervening (non-binary 0 or 1) snow covered area fractions. This may be a factor contributing to differences in regions with modeled SCAF less than 10% while the observed SCAF are more than 50% (points along the horizontal axis of SCAF on March 1, April 1, and May 1 in Figure 3.4).

We recognize that the SCAF mapped from MODIS in this study also has uncertainties and limitations. First, the temporal forward filling approach that we used to fill gaps associated with clouds may miss some of the daily variability of snow cover, particularly in mountainous regions. Second, the parameters of Equation (14), which estimates SCAF from MODIS-C6 NDSI_Snow_Cover product, were those from Salomonson and Appel (2006) and were constant for our entire study region. Adjusting these parameters to improve the snow cover products from MODIS regionally has been suggested (Riggs et al., 2017). Third, MODIS NDSI_Snow_Cover grids (nominally 500 m) were averaged for 1 km NWM grid cells, using an unweighted approach in the Google Earth Engine platform. This approach selects MODIS grids whose centers fall within the target area (i.e., NWM grid cells). These scale differences may be a further source of uncertainty, compounded by the nonlinearity in Equation (14) [plateau at NDSI > 0.7] having an impact on SCAF from averaged NDSI.

Results for the direct (binary) comparison of full snow cover were poor as MODIS never reports more than about 30% of the area as having full snow, while the degree-of-overlap between the modeled and observed results, in terms of average C and F, improved considerably when comparing cells having some snow present. We interpret this as a shortcoming of MODIS for this sort of comparison, perhaps due to the presence of vegetation. MODIS SCAF estimates may not account for snow beneath the canopy due

to incapability of the sensor to see forest gaps (the snow-covered ground) through the vegetation canopy (Steele et al., 2017; X. Wang et al., 2017), while the NWM-R2 land surface model (Noah-MP) may estimate snow under the vegetation canopy in these locations. Our results show that discrepancies between modeled and observed snowy grid cells reduce when we focus only on the SNOTEL sites reported as open. For full snow present average C improves from 9.41 to 6.18 while average F improves from 0.11 to 0.16. These are still poor, but less poor. For some snow present average C improves from 1.47 to 1.2 and average F improves from 0.5 to 0.64, making them reasonably respectable, in comparison to the ideal values of 1. This suggests that forest vegetation is a dominant contributor to the disagreement between model and MODIS observed snow cover.

3.5 Conclusions

A cell by cell comparison for sites and dates in the period of overlap between SNOTEL SWE with modeled SWE from NWM-R2 simulations, in general, shows that there is a tendency for the NWM-R2 configuration to under-estimate SWE early in the season and become progressively more biased late in the season compared to in-situ observations of SWE. When aggregated across all sites and years, seasonal variations show an overall downward bias of about 55 mm with NSE 0.75 which varies regionally over Omernik ecoregions. SWE discrepancies are attributed to errors in inputs, notably precipitation and air temperature. The downward bias in precipitation input contributes to the downward biases in SWE and the SWE bias is persistent even when the model precipitation input is relatively close to the observed precipitation at SNOTEL sites with

daily precipitation NSE higher than 0.9. However, the cold bias in the model temperature input is counter to the direction expected to lead to under-modeling of SWE. This needs further exploration. There is a significant variability between the MODIS SCAF and NWM SCAF in the cell by cell comparison for sites and dates in the period of overlap between model results and observations which hindered useful interpretation of these comparisons. The challenge in simulating SCAF is in part due to the model SCAF essentially being binary as it lacks representation of vegetation and topography while observations are much more fractional. They may not reflect the same physical quantity. The binary comparison of full snow presence reveals that the degree-of-overlap between the modeled and observed results still remains poor, which is possibly due to uncertainties associated with MODIS observations in vegetated areas. Results of the binary comparison of some snow presence improves when we focus only on the SNOTEL sites reported as open (average C=1.2 and average F=0.64). Also, when aggregated across all sites and years, seasonal variations show an overall upward bias of 0.12 with NSE 0.76 which vary regionally for ecoregions. Our investigation opens some new questions for future research. First, it emphasizes the importance of having a more accurate (bias corrected) precipitation and air temperature input for the NWM. Second, there is a question as to whether, in circumstances where there is disagreement between the NWM SCAF (estimated by the Noah-MP module) and MODIS observations in the binary comparison, the SCAF parameterization should be improved or can be inferred from satellites while considering the uncertainties associated with these products. Using satellite-based snow-covered maps may potentially provide an approach or an opportunity for estimating SCAF as a way to overcome limitations associated with

parameterization of SCAF in the snow model. However, there would need to be resolution of differences in definition of the physical quantity being compared. Overall, our evaluation effort identifies some challenges in the current snow parameterization within the specific settings of the Noah-MP as implemented in the NWM-R2 configuration and suggests where potential development effort should be directed in the future. It would also be helpful, for future work, to have a more comprehensive observation data set, beyond the SNOTEL sites, such as possibly Critical Zone Observatory (CZO) or experimental forest sites, that include snowfall/rainfall measurements, canopy snow interception, turbulence and radiation fluxes above and below the canopy. Another opportunity is to run the model at higher resolution which would involve downscaling the forcing inputs to higher resolution. Higher-resolution remotely sensed snow-covered area (e.g., from LANDSAT satellite) could then be used for model evaluation.

DATA AVAILABILITY

All data sources used in this research are publicly available.

- The NWM-R2 are available at the NOAA Google Cloud archive at https://console.cloud.google.com/storage/browser/national-water-model-v2?pli=1. The precipitation and air temperature inputs prepared by the WRF-Hydro NCAR team, we have made available on HydroShare for reproducibility (Garousi-Nejad & Tarboton, 2021f). The NWM elevation dataset is available at https://www.nco.ncep.noaa.gov/pmb/codes/nwprod/nwm.v2.0.4/parm/domain/
- The NRCS SNOTEL data are available at https://www.wcc.nrcs.usda.gov/snow/
- The NASA MODIS data are available at https://nsidc.org/data/MOD10A1/versions/6
- The Omernik ecoregions are available at http://www.cec.org/north-american-environmental-atlas/terrestrial-ecoregions-level-iii/

All codes developed for this research are shared and publicly available as a collection on HydroShare (Garousi-Nejad & Tarboton, 2021a) comprised of:

- Input data and code to get the indices of the NWM grid cells containing SNOTEL sites (Garousi-Nejad & Tarboton, 2021d)
- Input data, code to retrieve the NWM-R2 inputs and outputs at SNOTEL sites (Tarboton & Garousi-Nejad, 2021)
- Input data, code and output from post-processing the retrieved NWM-R2 inputs and outputs at SNOTEL sites (Garousi-Nejad & Tarboton, 2021f)
- Input data and code to retrieve precipitation, air temperature, and SWE measurements at SNOTEL sites (Garousi-Nejad & Tarboton, 2021c)

- Input data and Google Earth Engine code to retrieve averaged MODIS-C6 NDSI snow cover at SNOTEL sites (Garousi-Nejad & Tarboton, 2021b)
- Input data, code and output from combining the NWM inputs and outputs with observations form SNOTEL and MODIS at SNOTEL sites (Garousi-Nejad & Tarboton, 2021e)
- Input data, code and output used to produce Figures 3.1-3.4 and Figures 3.6-3.15 (Garousi-Nejad & Tarboton, 2021g)

ACKNOWLEDGMENTS

This work was supported by the Utah Water Research Laboratory and National Science Foundation under collaborative grants OAC-1664061 and OAC-1664119. This work used compute allocation TG-EAR190007 from the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562 (Towns et al., 2014). The support and resources from the Center for High Performance Computing (CHPC) at the University of Utah are gratefully acknowledged. Thanks to Arezoo RafieeiNasab of NCAR who supplied some of the National Water Model inputs for the study region and to David Gochis and the NCAR WRH-Hydro research team for encouragement and discussion. We also thank the Natural Resources Conservation Service (NRCS) staff, Danny Tappa; Jeff Anderson; Mage Hultstrand; Scott Pattee; Brian Domonkos; Dave Eiriksson; and Jordan Clayton, for their help providing us information on which SNOTEL sites are in open areas.

REFERENCES

Aalstad, K., Westermann, S., & Bertino, L. (2020). Evaluating satellite retrieved fractional snow-covered area at a high-Arctic site using terrestrial photography. Remote Sensing of Environment, 239, 111618. https://doi.org/10.1016/j.rse.2019.111618

Anderson, E. A. (1976). A point energy and mass balance model of a snow cover (NOAA Technical Report No. NWS 29). Silver Spring, MD: National Weather Service. Retrieved from https://repository.library.noaa.gov/view/noaa/6392

Barlage, M., Chen, F., Tewari, M., Ikeda, K., Gochis, D., Dudhia, J., et al. (2010). Noah land surface model modifications to improve snowpack prediction in the Colorado Rocky Mountains. Journal of Geophysical Research, 115(D22), D22101. https://doi.org/10.1029/2009JD013470

Bennett, K. E., Cherry, J. E., Balk, B., & Lindsey, S. (2019). Using MODIS estimates of fractional snow cover area to improve streamflow forecasts in interior Alaska. Hydrology and Earth System Sciences, 23(5), 2439–2459. https://doi.org/10.5194/hess-23-2439-2019

Bhatti, A. M., Koike, T., & Shrestha, M. (2016). Climate change impact assessment on mountain snow hydrology by water and energy budget-based distributed hydrological model. Journal of Hydrology, 543, 523–541. https://doi.org/10.1016/j.jhydrol.2016.10.025

Chen, Fei, Barlage, M., Tewari, M., Rasmussen, R., Jin, J., Lettenmaier, D., et al. (2014). Modeling seasonal snowpack evolution in the complex terrain and forested Colorado Headwaters region: A model intercomparison study. Journal of Geophysical Research: Atmospheres, 119(24), 13,795-13,819. https://doi.org/10.1002/2014JD022167

Chen, Feng, Liu, C., Dudhia, J., & Chen, M. (2014). A sensitivity study of high-resolution regional climate simulations to three land surface models over the western United States: SENSITIVITY STUDY OF LSMS IN WRF. Journal of Geophysical Research: Atmospheres, 119(12), 7271–7291. https://doi.org/10.1002/2014JD021827

Clow, D. W. (2010). Changes in the timing of snowmelt and streamflow in Colorado: A response to recent warming. Journal of Climate, 23(9), 2293–2306. https://doi.org/10.1175/2009JCLI2951.1

Clow, D. W., Nanus, L., Verdin, K. L., & Schmidt, J. (2012). Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA: EVALUATION OF SNODAS. Hydrological Processes, 26(17), 2583–2591. https://doi.org/10.1002/hyp.9385

Franz, K. J., Hogue, T. S., & Sorooshian, S. (2008). Operational snow modeling: Addressing the challenges of an energy balance model for National Weather Service forecasts. Journal of Hydrology, 360(1–4), 48–66. https://doi.org/10.1016/j.jhydrol.2008.07.013

Garousi-Nejad, I., & Tarboton, D. (2021a). Data for A Comparison of National Water Model Retrospective Analysis Snow Outputs at SNOTEL Sites Across the Western U.S. HydroShare. Retrieved from https://www.hydroshare.org/resource/7a51f56c2cf24ae78012ac6a6d4815a6/

Garousi-Nejad, I., & Tarboton, D. (2021b). JavaScript code for retrieval of MODIS Collection 6 NDSI snow cover at SNOTEL sites and a Jupyter Notebook to merge/reprocess data. HydroShare. Retrieved from https://www.hydroshare.org/resource/d287f010b2dd48edb0573415a56d47f8/

Garousi-Nejad, I., & Tarboton, D. (2021c). Notebook for retrieval of precipitation, air temperature, and snow water equivalent measurements at SNOTEL sites. HydroShare. Retrieved from https://www.hydroshare.org/resource/d1fe0668734e4892b066f198c4015b06/

Garousi-Nejad, I., & Tarboton, D. (2021d). Notebook to get the indices of National Water Model V2.0 grid cells containing SNOTL sites. HydroShare. Retrieved from https://www.hydroshare.org/resource/7839e3f3b4f54940bd3591b24803cacf/

Garousi-Nejad, I., & Tarboton, D. (2021e). Notebooks for combining the National Water Model results/inputs with observations from SNOTEL and MODIS at SNOTEL sites. HydroShare. Retrieved from https://www.hydroshare.org/resource/493e0ad05c2d45199427cc41a6c76de0/

Garousi-Nejad, I., & Tarboton, D. (2021f). Notebooks for post-processing the retrieved National Water Model V2.0 Retrospective run results and inputs at SNOTEL sites. HydroShare. Retrieved from https://www.hydroshare.org/resource/1b66a752b0cc467eb0f46bda5fdc4b34/

Garousi-Nejad, I., & Tarboton, D. (2021g). Notebooks for visualizations reported at A Comparison of National Water Model Retrospective Analysis Snow Outputs at SNOTEL Sites Across the Western U.S. HydroShare. Retrieved from https://www.hydroshare.org/resource/8507aa41130e45bfb0752026cf2253ab/

Gergel, D. R., Nijssen, B., Abatzoglou, J. T., Lettenmaier, D. P., & Stumbaugh, M. R. (2017). Effects of climate change on snowpack and fire potential in the western USA. Climatic Change, 141(2), 287–299. https://doi.org/10.1007/s10584-017-1899-y

Gochis, D., Barlage, M., Cabell, R., Casali, M., Dugger, A., FitzGerald, K., et al. (2020). The WRF-Hydro® modeling system technical description, (Version 5.1.1). NCAR Technical Note. Retrieved from

https://ral.ucar.edu/sites/default/files/public/WRFHydroV511TechnicalDescription.pdf

Gochis, D., Barlage, M., Cabell, R., Dugger, A., Fanfarillo, A., FitzGerald, K., et al. (2020). WRF-Hydro® v5.1.1 (Version v5.1.1). Zenodo. https://doi.org/10.5281/ZENODO.3625238

Hall, D. K., & Riggs, G. A. (2016). MODIS/Terra Snow Cover Daily L3 Global 500m SIN Grid [Data set]. NASA National Snow and Ice Data Center DAAC. https://doi.org/10.5067/MODIS/MOD10A1.006

Harder, P., & Pomeroy, J. W. (2014). Hydrological model uncertainty due to precipitation-phase partitioning methods: HYDROLOGIC MODEL UNCERTAINTY OF PRECIPITATION-PHASE METHODS. Hydrological Processes, 28(14), 4311–4327. https://doi.org/10.1002/hyp.10214

Hedstrom, N. R., & Pomeroy, J. W. (1998). Measurements and modelling of snow interception in the boreal forest. Hydrological Processes, 12, 1611–1625. https://doi.org/10.1002/(SICI)1099-1085(199808/09)12:10/11%3C1611::AID-HYP684%3E3.0.CO;2-4

Helbig, N., van Herwijnen, A., Magnusson, J., & Jonas, T. (2015). Fractional snow-covered area parameterization over complex topography. Hydrology and Earth System Sciences, 19(3), 1339–1351. https://doi.org/10.5194/hess-19-1339-2015

Horritt, M. S., & Bates, P. D. (2002). Evaluation of 1D and 2D numerical models for predicting river flood inundation. Journal of Hydrology, 268(1–4), 87–99. https://doi.org/10.1016/S0022-1694(02)00121-X

Hou, D., charles, M., Luo, Y., Toth, Z., Zhu, Y., Krzysztofowicz, R., et al. (2014). Climatology-Calibrated Precipitation Analysis at Fine Scales: Statistical Adjustment of Stage IV toward CPC Gauge-Based Analysis. Journal of Hydrometeorology, 15(6), 2542–2557. https://doi.org/10.1175/JHM-D-11-0140.1

Jordan, R. E. (1991). A One-dimensional temperature model for a snow cover : technical documentation for SNTHERM.89. Cold Regions Research and Engineering Laboratory (U.S.). Retrieved from http://hdl.handle.net/11681/11677

Lahmers, T. M., Gupta, H., Castro, C. L., Gochis, D. J., Yates, D., Dugger, A., et al. (2019). Enhancing the Structure of the WRF-Hydro Hydrologic Model for Semiarid Environments. Journal of Hydrometeorology, 20(4), 691–714. https://doi.org/10.1175/JHM-D-18-0064.1

Li, D., Wrzesien, M. L., Durand, M., Adam, J., & Lettenmaier, D. P. (2017). How much runoff originates as snow in the western United States, and how will that change in the future?: Western U.S. Snowmelt-Derived Runoff. Geophysical Research Letters, 44(12), 6163–6172. https://doi.org/10.1002/2017GL073551

Livneh, B., & Badger, A. M. (2020). Drought less predictable under declining future snowpack. Nature Climate Change, 10(5), 452–458. https://doi.org/10.1038/s41558-020-0754-8

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

Loth, B., Graf, H.-F., & Oberhuber, J. M. (1993). Snow cover model for global climate simulations. Journal of Geophysical Research, 98(D6), 10451. https://doi.org/10.1029/93JD00324

Lundquist, J. D., & Flint, A. L. (2006). Onset of Snowmelt and Streamflow in 2004 in the Western United States: How Shading May Affect Spring Streamflow Timing in a Warmer World. Journal of Hydrometeorology, 7(6), 1199–1217. https://doi.org/10.1175/JHM539.1

Lynch-Stieglitz, M. (1994). The Development and Validation of a Simple Snow Model for the GISS GCM. Journal of Climate, 7(12), 1842–1855. Retrieved from http://www.jstor.org/stable/26198671

Magand, C., Ducharne, A., Le Moine, N., & Gascoin, S. (2014). Introducing Hysteresis in Snow Depletion Curves to Improve the Water Budget of a Land Surface Model in an Alpine Catchment. Journal of Hydrometeorology, 15(2), 631–649. https://doi.org/10.1175/JHM-D-13-091.1

Mahat, V., & Tarboton, D. G. (2014). Representation of canopy snow interception, unloading and melt in a parsimonious snowmelt model: CANOPY SNOW INTERCEPTION, UNLOADING AND MELT. Hydrological Processes, 28(26), 6320–6336. https://doi.org/10.1002/hyp.10116

Masson, T., Dumont, M., Mura, M., Sirguey, P., Gascoin, S., Dedieu, J.-P., & Chanussot, J. (2018). An Assessment of Existing Methodologies to Retrieve Snow Cover Fraction from MODIS Data. Remote Sensing, 10(4), 619. https://doi.org/10.3390/rs10040619

McCreight, J. L., Small, E. E., & Larson, K. M. (2014). Snow depth, density, and SWE estimates derived from GPS reflection data: Validation in the western U. S. Water Resources Research, 50(8), 6892–6909. https://doi.org/10.1002/2014WR015561

McEnery, J., Ingram, J., Duan, Q., Adams, T., & Anderson, L. (2005). NOAA'S ADVANCED HYDROLOGIC PREDICTION SERVICE: Building Pathways for Better Science in Water Forecasting. Bulletin of the American Meteorological Society, 86(3), 375–386. https://doi.org/10.1175/BAMS-86-3-375

Meyer, J. D. D., Jin, J., & Wang, S.-Y. (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. https://doi.org/10.1175/JHM-D-12-066.1

Mote, P. W. (2003). Trends in snow water equivalent in the Pacific Northwest and their climatic causes: TRENDS IN SNOW WATER EQUIVALENT. Geophysical Research Letters, 30(12). https://doi.org/10.1029/2003GL017258

Mote, P. W., Hamlet, A. F., Clark, M. P., & Lettenmaier, D. P. (2005). DECLINING MOUNTAIN SNOWPACK IN WESTERN NORTH AMERICA*. Bulletin of the American Meteorological Society, 86(1), 39–50. https://doi.org/10.1175/BAMS-86-1-39

Naple, P., Letcher, T., Minder, J. R., Gochis, D., Dugger, A., & RafieeiNasab, A. (2020). Improving parameterizations of snow in the National Water Model with observations from the New York State Mesonet to better simulate snow and streamflow in the northeastern United States. Presented at the AGU Fall Meeting, Virtual. Retrieved from https://ui.adsabs.harvard.edu/abs/2020AGUFMC063.0006N/abstract

National Weather Service. (2019). National-Water-Model-V2. Google Cloud Bucket. Retrieved from https://console.cloud.google.com/storage/browser/national-water-model-v2

Niu, G.-Y., & Yang, Z.-L. (2004). Effects of vegetation canopy processes on snow surface energy and mass balances: CANOPY EFFECTS ON SNOW PROCESSES. Journal of Geophysical Research: Atmospheres, 109(D23). https://doi.org/10.1029/2004JD004884

Niu, G.-Y., & Yang, Z.-L. (2007). An observation-based formulation of snow cover fraction and its evaluation over large North American river basins. Journal of Geophysical Research, 112(D21), D21101. https://doi.org/10.1029/2007JD008674

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. Journal of Geophysical Research, 116(D12), D12109. https://doi.org/10.1029/2010JD015139

NRC. (2006). Towards Advanced Hydrologic Prediction Service (AHPS). National Research Council Committee to Assess the National Weather Service Advanced Hydrologic Prediction Service Initiative. Washington, DC: National Academies Press.

Omernik, J. M., & Griffith, G. E. (2014). Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework. Environmental Management, 54(6), 1249–1266. https://doi.org/10.1007/s00267-014-0364-1

Pan, M., Sheffield, J., Wood, E. F., Mitchell, K. E., Houser, P. R., Schaake, J. C., et al. (2003). Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent. Journal of Geophysical Research: Atmospheres, 108(D22), 2003JD003994. https://doi.org/10.1029/2003JD003994

RafieeiNasab, A., Karsten, L., Dugger, A., FitzGerald, K., Cabell, R., Gochis, D., et al. (2020). Overview of National Water Model calibration general strategy & optimization, NCAR Community WRF-Hydro Modeling System training workshop. Presented at the NCAR Community WRF-Hydro Modeling System training workshop. Retrieved from https://ral.ucar.edu/projects/wrf_hydro/training-materials

Regonda, S. K., Rajagopalan, B., Clark, M., & Pitlick, J. (2005). Seasonal Cycle Shifts in Hydroclimatology over the Western United States. Journal of Climate, 18(2), 372–384. https://doi.org/10.1175/JCLI-3272.1

Riggs, G. A., Hall, D. K., & Román, M. O. (2016). MODIS Snow Products Collection 6 User Guide. Retrieved from http://modis-snow-ice.gsfc.nasa.gov/uploads/C6_MODIS_Snow_User_Guide.pdf

Riggs, G. A., Hall, D. K., & Román, M. O. (2017). Overview of NASA's MODIS and Visible Infrared Imaging Radiometer Suite (VIIRS) snow-cover Earth System Data Records. Earth System Science Data, 9(2), 765–777. https://doi.org/10.5194/essd-9-765-2017

Salomonson, V. V., & Appel, I. (2006). Development of the Aqua MODIS NDSI fractional snow cover algorithm and validation results. IEEE Transactions on Geoscience and Remote Sensing, 44(7), 1747–1756. https://doi.org/10.1109/TGRS.2006.876029

Sangwan, N., & Merwade, V. (2015). A Faster and Economical Approach to Floodplain Mapping Using Soil Information. JAWRA Journal of the American Water Resources Association, 51(5), 1286–1304. https://doi.org/10.1111/1752-1688.12306

Shamir, E., Carpenter, T. M., Fickenscher, P., & Georgakakos, K. P. (2006). Evaluation of the National Weather Service Operational Hydrologic Model and Forecasts for the American River Basin. Journal of Hydrologic Engineering, 11(5), 392–407. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:5(392)

Steele, C., Dialesandro, J., James, D., Elias, E., Rango, A., & Bleiweiss, M. (2017). Evaluating MODIS snow products for modelling snowmelt runoff: Case study of the Rio Grande headwaters. International Journal of Applied Earth Observation and Geoinformation, 63, 234–243. https://doi.org/10.1016/j.jag.2017.08.007

Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2004). Changes in Snowmelt Runoff Timing in Western North America under a `Business as Usual' Climate Change Scenario. Climatic Change, 62(1–3), 217–232. https://doi.org/10.1023/B:CLIM.0000013702.22656.e8

Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2005). Changes toward Earlier Streamflow Timing across Western North America. Journal of Climate, 18(8), 1136–1155. https://doi.org/10.1175/JCLI3321.1

Sun, N., Yan, H., Wigmosta, M. S., Leung, L. R., Skaggs, R., & Hou, Z. (2019). Regional Snow Parameters Estimation for Large-Domain Hydrological Applications in the Western United States. Journal of Geophysical Research: Atmospheres, 124(10), 5296–5313. https://doi.org/10.1029/2018JD030140

Swenson, S. C., & Lawrence, D. M. (2012). A new fractional snow-covered area parameterization for the Community Land Model and its effect on the surface energy balance: CLM SNOW COVER FRACTION. Journal of Geophysical Research: Atmospheres, 117(D21), n/a-n/a. https://doi.org/10.1029/2012JD018178

Tarboton, D., & Garousi-Nejad, I. (2021). Notebook for retrieval of National Water Model V2.0 Retrospective run results at SNOTEL sites. HydroShare. Retrieved from https://www.hydroshare.org/resource/3d4976bf6eb84dfbbe11446ab0e31a0a/

Toure, A. M., Rodell, M., Yang, Z.-L., Beaudoing, H., Kim, E., Zhang, Y., & Kwon, Y. (2016). Evaluation of the Snow Simulations from the Community Land Model, Version 4 (CLM4). Journal of Hydrometeorology, 17(1), 153–170. https://doi.org/10.1175/JHM-D-14-0165.1

Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., et al. (2014). XSEDE: Accelerating Scientific Discovery. Computing in Science & Engineering, 16(5), 62–74. https://doi.org/10.1109/MCSE.2014.80

Trujillo, E., & Molotch, N. P. (2014). Snowpack regimes of the Western United States. Water Resources Research, 50(7), 5611–5623. https://doi.org/10.1002/2013WR014753

U.S. Department of Agriculture. (2011). Snow Survey and Water Supply Forecasting. National Engineering Handbook Part 622. Water and Climate Center, Natural Resources Conservation Service. Retrieved from directives.sc.egov.usda.gov/viewerFS.aspx?hid=32040

Viterbo, F., Mahoney, K., Read, L., Salas, F., Bates, B., Elliott, J., et al. (2020). A Multiscale, Hydrometeorological Forecast Evaluation of National Water Model Forecasts of the May 2018 Ellicott City, Maryland, Flood. Journal of Hydrometeorology, 21(3), 475–499. https://doi.org/10.1175/JHM-D-19-0125.1

Wang, X., Zhu, Y., Chen, Y., Zheng, H., Liu, H., Huang, H., et al. (2017). Influences of forest on MODIS snow cover mapping and snow variations in the Amur River basin in Northeast Asia during 2000-2014. Hydrological Processes, 31(18), 3225–3241. https://doi.org/10.1002/hyp.11249

Wang, Y., Broxton, P., Fang, Y., Behrangi, A., Barlage, M., Zeng, X., & Niu, G. (2019). A Wet-Bulb Temperature-Based Rain-Snow Partitioning Scheme Improves Snowpack Prediction Over the Drier Western United States. Geophysical Research Letters, 46(23), 13825–13835. https://doi.org/10.1029/2019GL085722

Wrzesien, M. L., Pavelsky, T. M., Kapnick, S. B., Durand, M. T., & Painter, T. H. (2015). Evaluation of snow cover fraction for regional climate simulations in the Sierra Nevada: EVALUATION OF SNOW COVER FOR REGIONAL SIMULATIONS IN THE SIERRA NEVADA. International Journal of Climatology, 35(9), 2472–2484. https://doi.org/10.1002/joc.4136

Yang, Z.-L., & Dickinson, R. E. (1996). Description of the Biosphere-Atmosphere Transfer Scheme (BATS) for the Soil Moisture Workshop and evaluation of its performance. Global and Planetary Change, 13(1–4), 117–134. https://doi.org/10.1016/0921-8181(95)00041-0

Yang, Z.-L., Niu, G.-Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. Journal of Geophysical Research, 116(D12), D12110. https://doi.org/10.1029/2010JD015140

You, J., Tarboton, D. G., & Luce, C. H. (2014). Modeling the snow surface temperature with a one-layer energy balance snowmelt model. Hydrology and Earth System Sciences, 18(12), 5061–5076. https://doi.org/10.5194/hess-18-5061-2014

Zalenski, G., Krajewski, W. F., Quintero, F., Restrepo, P., & Buan, S. (2017). Analysis of National Weather Service Stage Forecast Errors. Weather and Forecasting, 32(4), 1441–1465. https://doi.org/10.1175/WAF-D-16-0219.1

Zhang, J., Condon, L. E., Tran, H., & Maxwell, R. M. (2021). A national topographic dataset for hydrological modeling over the contiguous United States. Earth System Science Data, 13(7), 3263–3279. https://doi.org/10.5194/essd-13-3263-2021

Appendices

Table 3.1 The Noah-MP land surface model options as defined in the National Water Model version 2.0 retrospective analysis configuration †

Code Name	Long Name	Physics Option Used
DYNAMIC_VEG_OPTION	Dynamic vegetation	4: Using monthly LAI is prescribed for various vegetation types
CANOPY_STOMATAL_RESISTANCE_OPTION	Canopy stomatal resistance	1: Ball-Berry
BTR_OPTION	Soil moisture factor for stomatal resistance	1: Noah type using soil moisture
RUNOFF_OPTION	Runoff and groundwater	3: Noah type surface and subsurface runoff (free drainage)
SURFACE_DRAG_OPTION	Surface layer drag coefficients	1: Monin-Obukhov
FROZEN_SOIL_OPTION	Frozen soil permeability	1: Using the total soil moisture to compute hydraulic properties
SUPERCOOLED_WATER_OPTION	Supercooled liquid water (or ice fraction)	1: No iteration (Form of the freezing-point depression equation)
RADIATIVE_TRANSFER_OPTION	Radiation transfer	3: Two-stream applied to vegetated fraction
SNOW_ALBEDO_OPTION	Ground snow surface albedo	2: BATS
PCP_PARTITION_OPTION	Partitioning precipitation into rainfall & snowfall	1: Jordan (1991)

†Based on Gochis, Barlage, Cabell, Casali, et al. (2020) and Gochis, Barlage, Cabell, Dugger, et. (2020).

Table 3.1 (continued).

Code Name	Long Name	Physics Option Used
TBOT_OPTION	Lower boundary condition of soil temperature	2: TBOT at ZBOT (8m) read from a file
TEMP_TIME_SCHEME_OPTION	Snow/soil temperature time scheme (only layer 1)	3: Semi-implicit; flux top boundary condition, but FSNO for TS calculation
GLACIER_OPTION	Glacier treatment	2: Ice treatment more like original Noah
SURFACE_RESISTANCE_OPTION	Surface resistant to evaporation and sublimation	4: For non-snow; rsurf = rsurf_snow for snow (set in MPTABLE)

[†]Based on Gochis, Barlage, Cabell, Casali, et al. (2020) and Gochis, Barlage, Cabell, Dugger, et. (2020).

Table 3.2 Summary of average correctness ($\bar{C}$) and average fit ($\bar{F}$) metrics evaluated for the binary comparison of snow presence or absence when considering (a) all SNOTEL sites and (b) sites reported as open approaches.

Average metrics	(a) All 734 SNOTEL sites		(b) The 62 SNOTEL sites reported as open	
	Snow Presence Condition		Snow Presence Condition	
	Full‡	Some§	Full‡	Some§
$\bar{C}$	9.41	1.47	6.18	1.20
$\bar{F}$	0.11	0.50	0.16	0.64

‡Daily modeled snow-covered area fraction (SCAF) taken as full snow if SCAF is $\geq$ 0.95. Daily MODIS SCAF taken as full snow if NDSI is $\geq$ 0.7.

§Daily modeled SCAF taken as some snow if SCAF is $\geq$ 0.3. Daily MODIS SCAF taken as some snow if NDSI is $\geq$ 0.3.

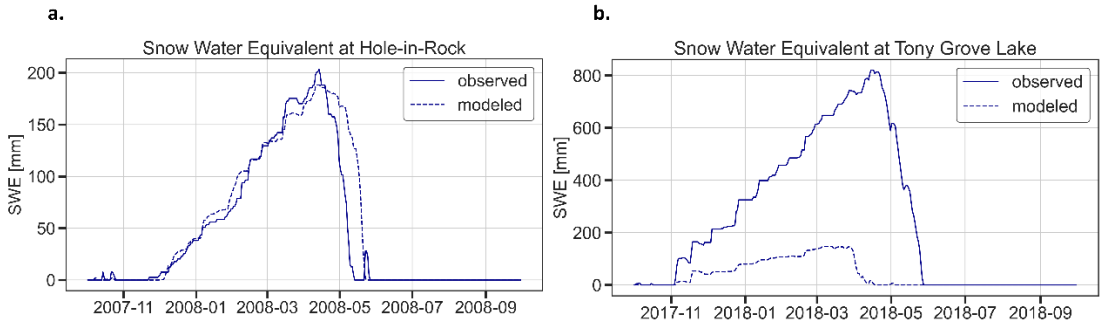

Figure 3.1 Snow water equivalent from the NWM version 2.0 reanalysis (NWM-R2) dataset compared to in-situ observations at two SNOTEL sites in Utah. (a) Hole-in-Rock site (ID: 528) located at 2794 m elevation for the water year 2008. (b) Tony Grove Lake site (ID: 823) located at 2582 m elevation for the water year 2018.

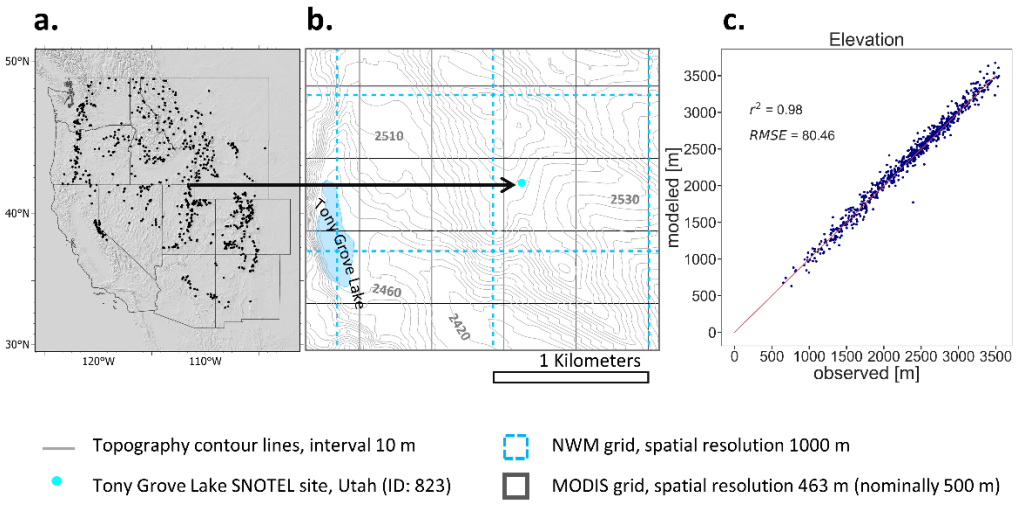

Figure 3.2 (a) SNOTEL sites (734 black dots) across the Western United States. (b) Illustrative relationship of Tony Grove Lake, Utah SNOTEL site (ID: 823), within NWM grid cells with a spatial resolution of 1 km and MODIS grid cells with a spatial resolution of 463 m (nominally 500 m). (c) NWM grid cell elevation vs. elevation reported for SNOTEL sites (observed). Note that there are four MODIS grid cells that have their centroid within each single NWM grid cell.

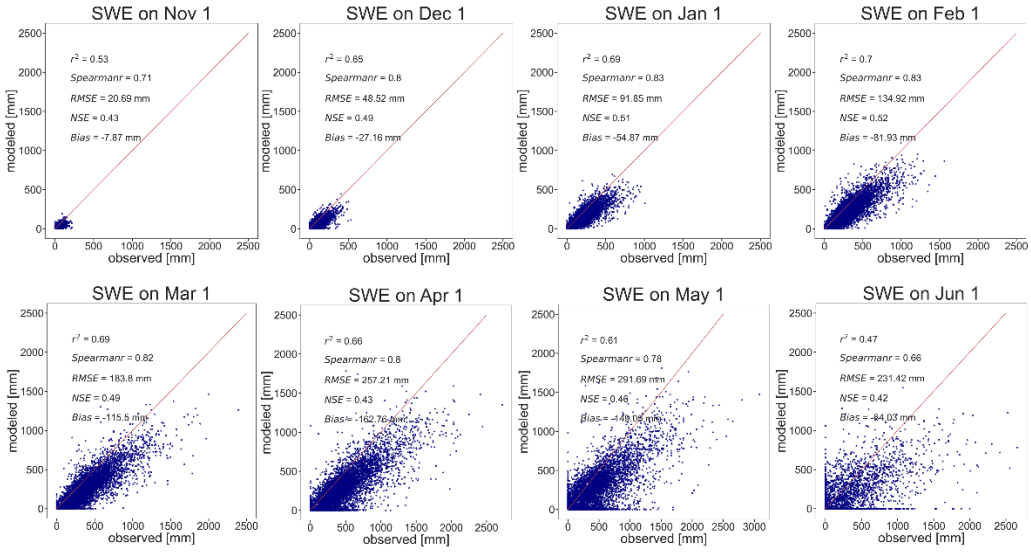

Figure 3.3 First day of month modeled (NWM-R2) vs. observed (SNOTEL) SWE. Each point is a site and date in the period of overlap between NWM-R2 and SNOTEL data.

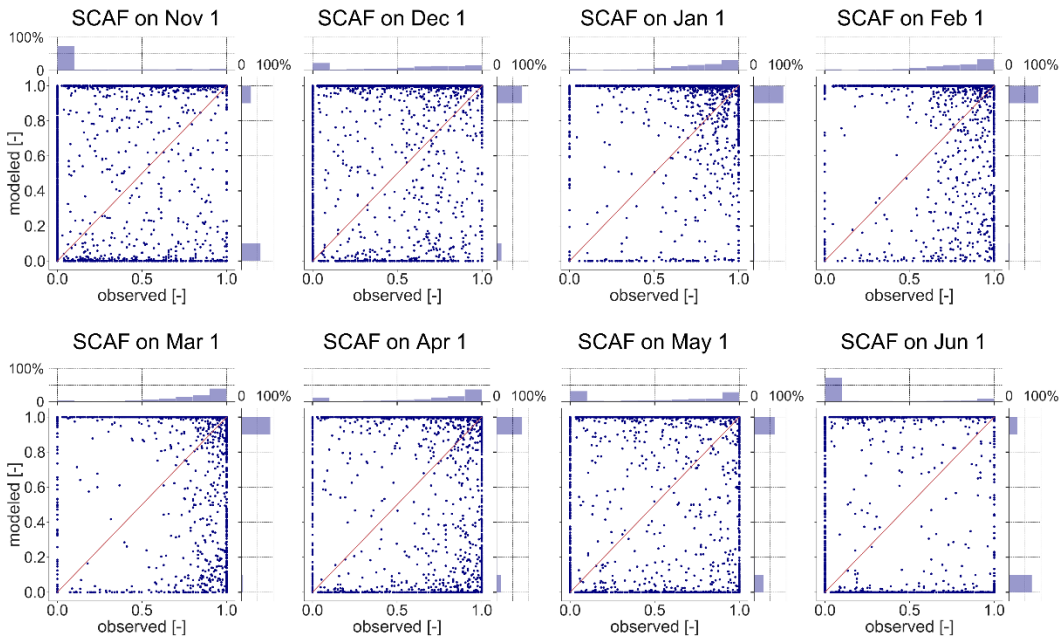

Figure 3.4 First day of month modeled (NWM-R2) vs. observed (MODIS-C6) SCAF for NWM grid cells and MODIS grid cells containing SNOTEL sites. Each point is a site and a date within the period of overlap between NWM and MODIS data. Axis histograms depict the SCAF distributions.

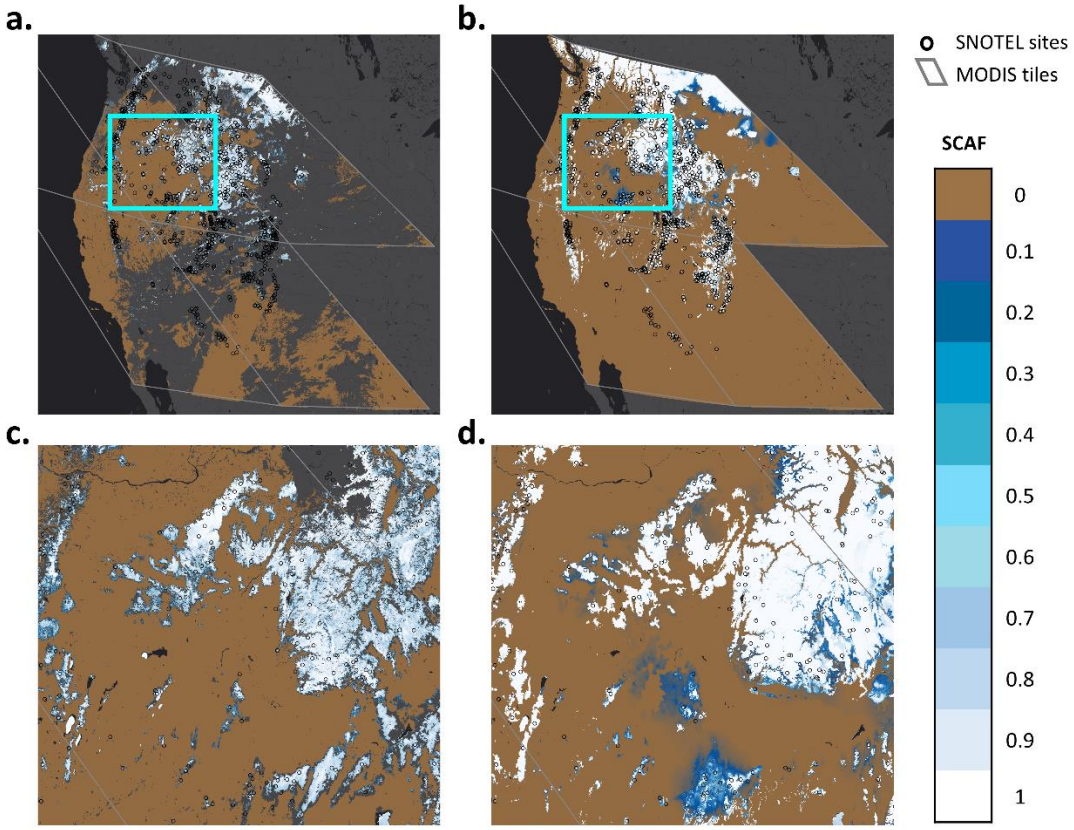

Figure 3.5 Comparison of NWM-R2 and MODIS-C6 SCAF maps over the study region on Dec 1, 2011. (a) MODIS-C6 SCAF estimated from NDSI_Snow_Cover values of five tiles (in grey). (b) NWM-R2 SCAF outputs at 00:00 UTC masked for the MODIS-C6 tiles. (c) The zoomed-in map of MODIC-C6 SCAF for the blue box in (a). (d) The zoomed-in map of NWM-R2 SCAF for the blue box in (b).

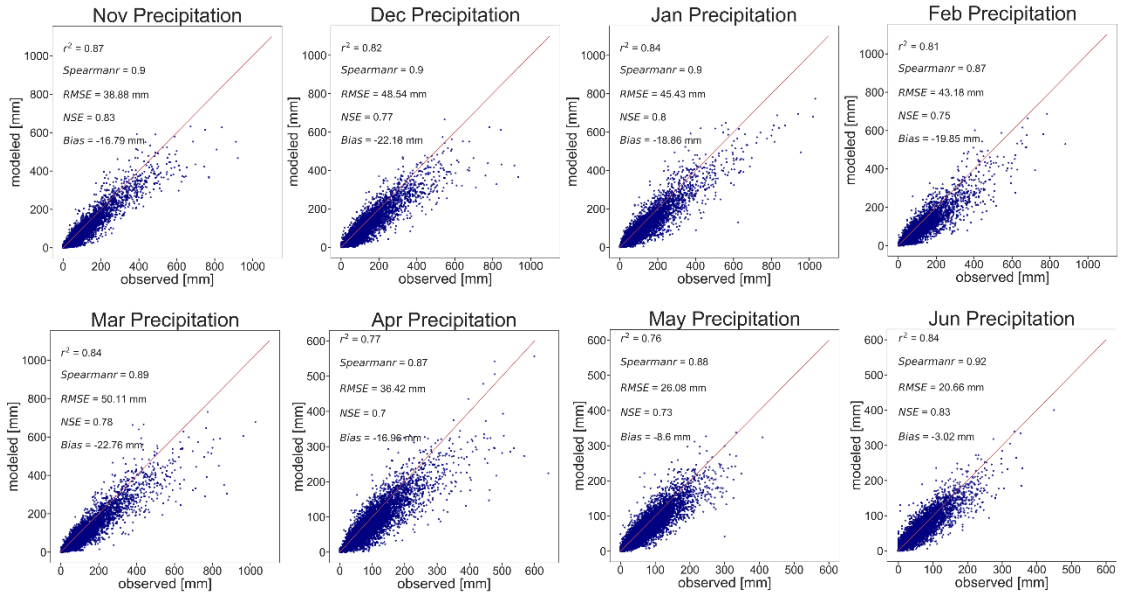

Figure 3.6 Comparison between NWM-R2 monthly precipitation input (labeled as modeled) and SNOTEL monthly precipitation (labeled as observed). Each point is a site and month in the period of overlap between NWM-R2 and SNOTEL data.

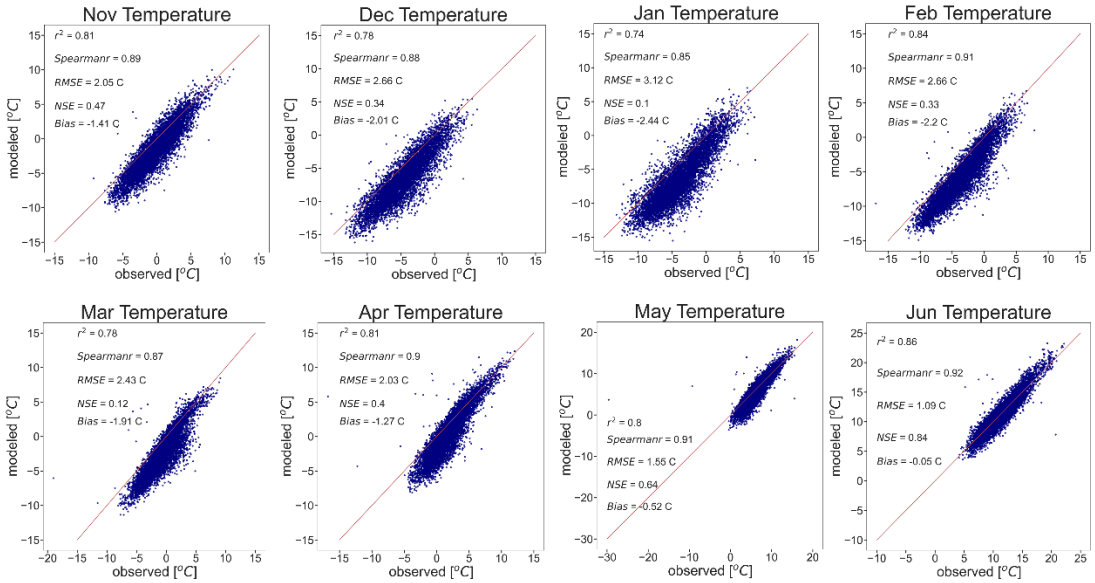

Figure 3.7 Comparison between NWM-R2 monthly average of hourly air temperature input (labeled as modeled) and SNOTEL monthly average of mean daily air temperature (labeled as observed). Each point is a site and month in the period of overlap between NWM-R2 and SNOTEL data.

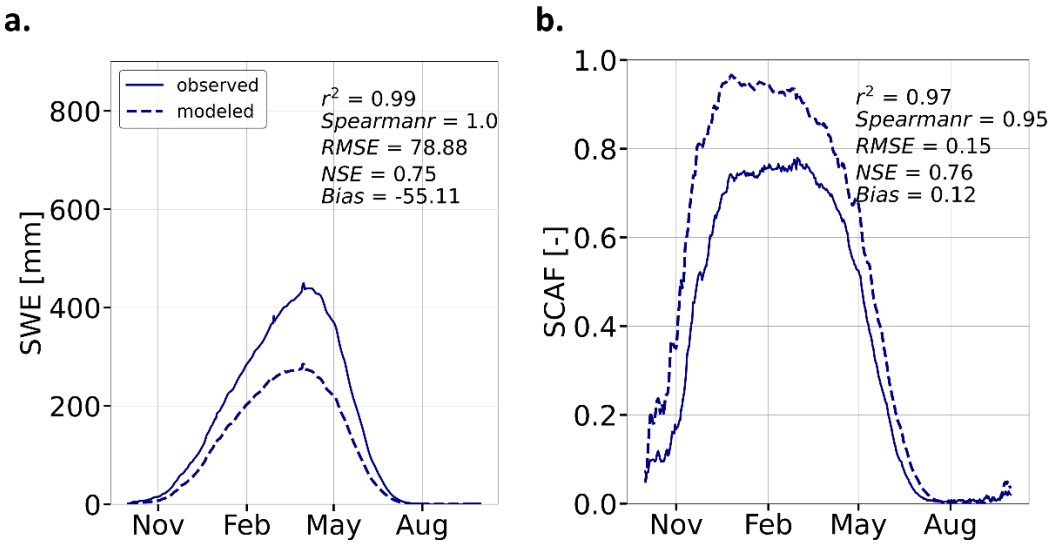

Figure 3.8 Modeled and observed (a) SWE and (b) SCAF averaged across all SNOTEL sites and years for each specific day of the (water) year.

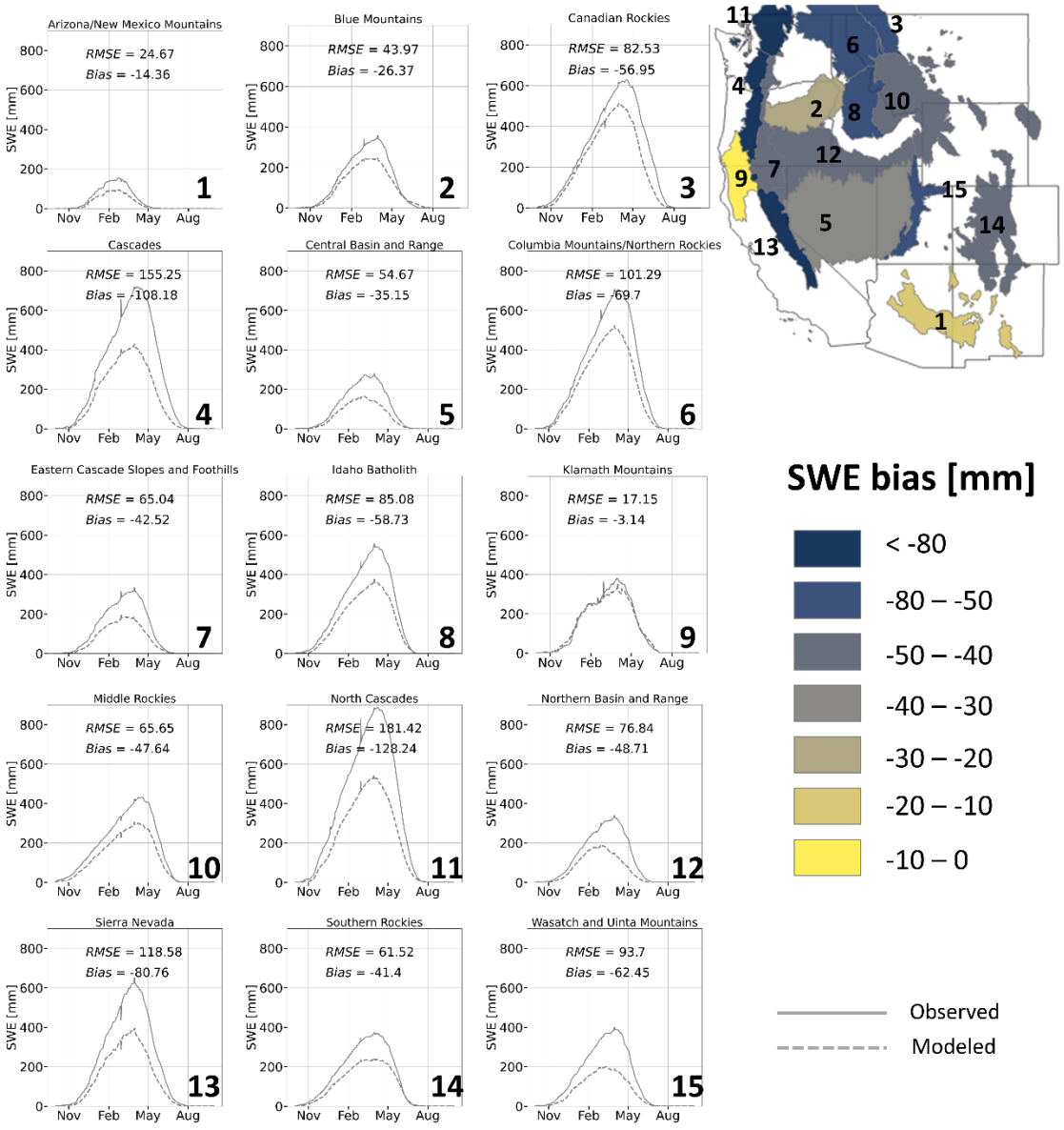

Figure 3.9 Modeled and observed SWE averaged across all SNOTEL sites and years for each specific day of the (water) year grouped by ecoregion. The map shows 15 Omernik ecoregions where colours represent the bias.

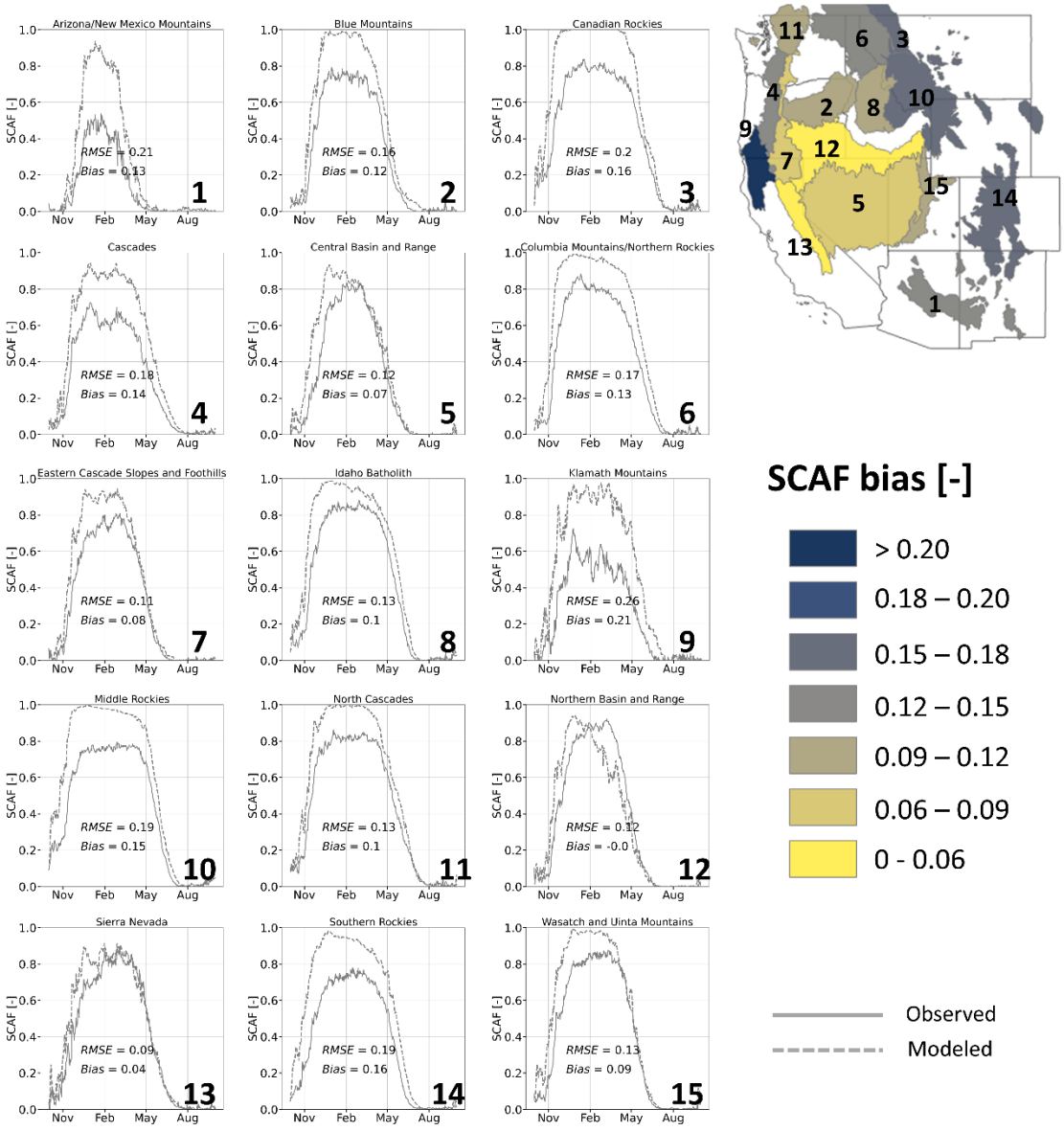

Figure 3.10 Modeled and observed SCAF averaged across all SNOTEL sites and years for each specific day of the (water) year grouped by ecoregion. The map shows 15 Omernik ecoregions where colours represent the bias.

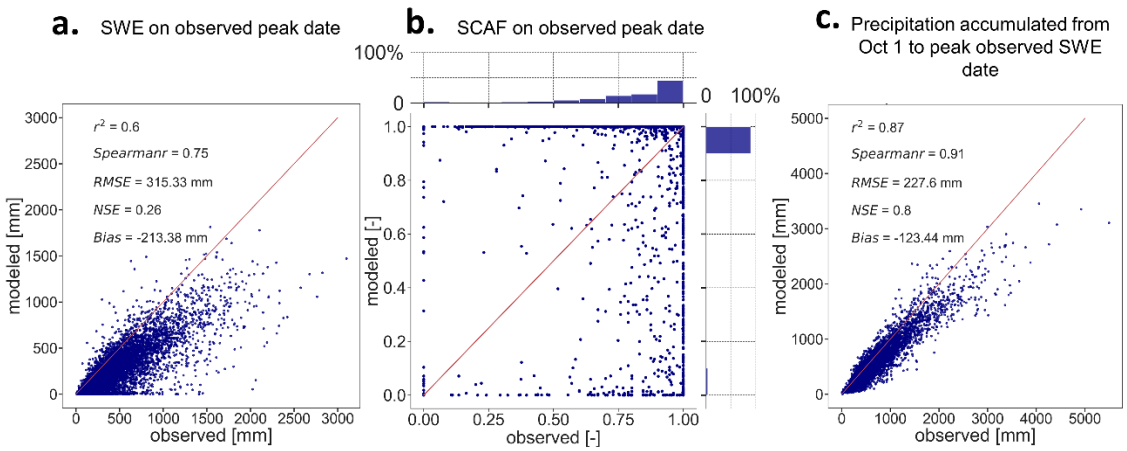

Figure 3.11 Comparisons on date of observed peak SWE. (a) NWM-R2 vs. SNOTEL SWE, (b) NWM-R2 vs. MODIS-C6 SCAF, and (c) NWM-R2 vs. SNOTEL precipitation accumulated from Oct 1 to observed peak SWE date. Each point is a site and a water year (that starts Oct 1) in the period of overlap between NWM-R2 and SNOTEL data.

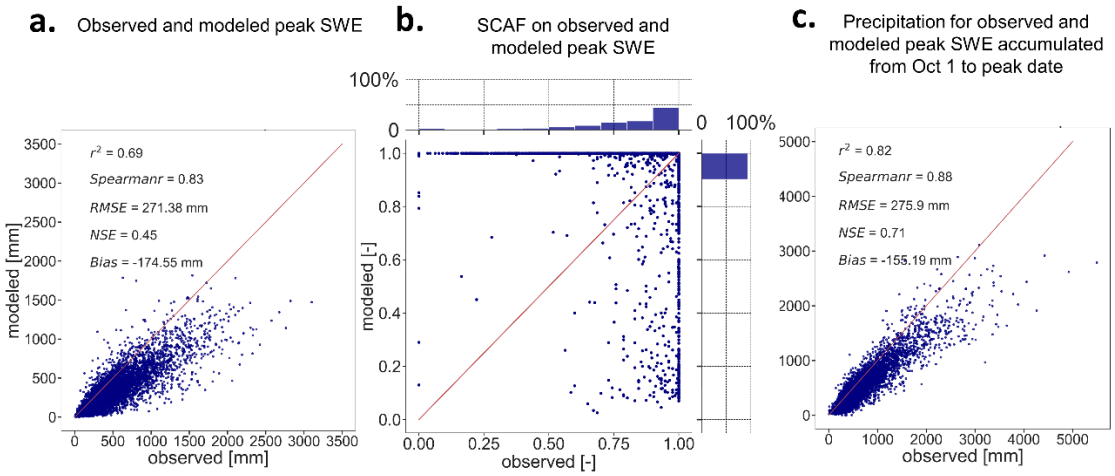

Figure 3.12 Different date comparison on dates of observed and modeled peak SWE (a) NWM-R2 vs. SNOTEL peak SWE, (b) NWM-R2 vs. MODIS-C6 SCAF, and (c) NWM-R2 vs. SNOTEL precipitation accumulated from Oct 1 to observed and modeled peak SWE dates. Each point is a site and a water year (that starts Oct 1) in the period of overlap between NWM-R2 and SNOTEL data.

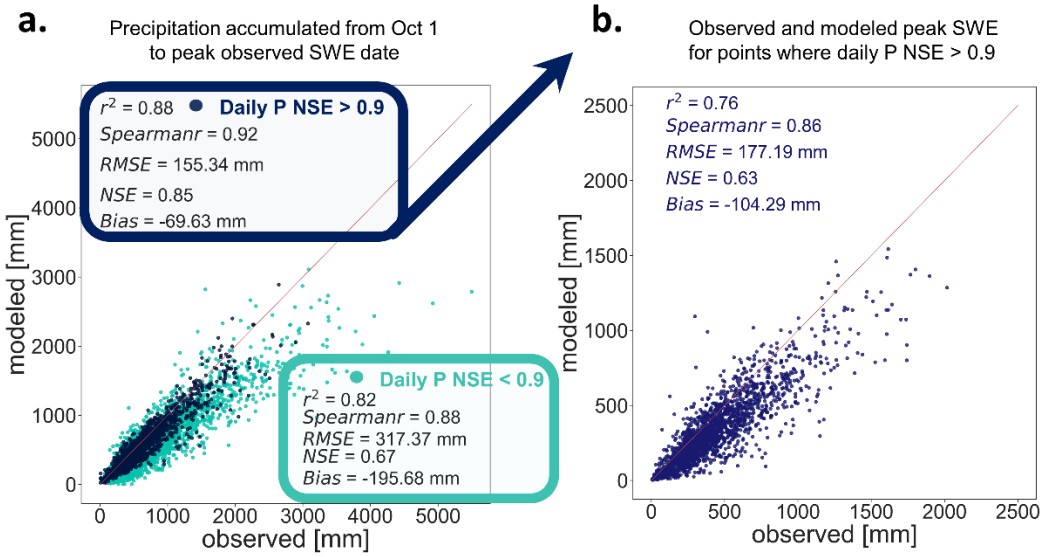

Figure 3.13 (a) NWM-R2 vs. SNOTEL precipitation accumulated from Oct 1 to observed and modeled peak SWE dates. This figure is similar to Figure 3.10 (a) but with colours separating points into two groups. The first group (dark blue) contains points where Nash Sutcliffe Efficiency (NSE) values for daily modeled vs. observed precipitation are equal to or greater than 0.9. The second group (light blue) includes points where NSE values for daily modeled vs. observed precipitation are less than 0.9. Statistics are reported separately for the NSE >= 0.9 and NSE < 0.9 subsets. (b) NWM-R2 peak SWE vs. SNOTEL peak SWE for points from (a) that have daily precipitation NSE equal to or greater than 0.9 (dark blue class).

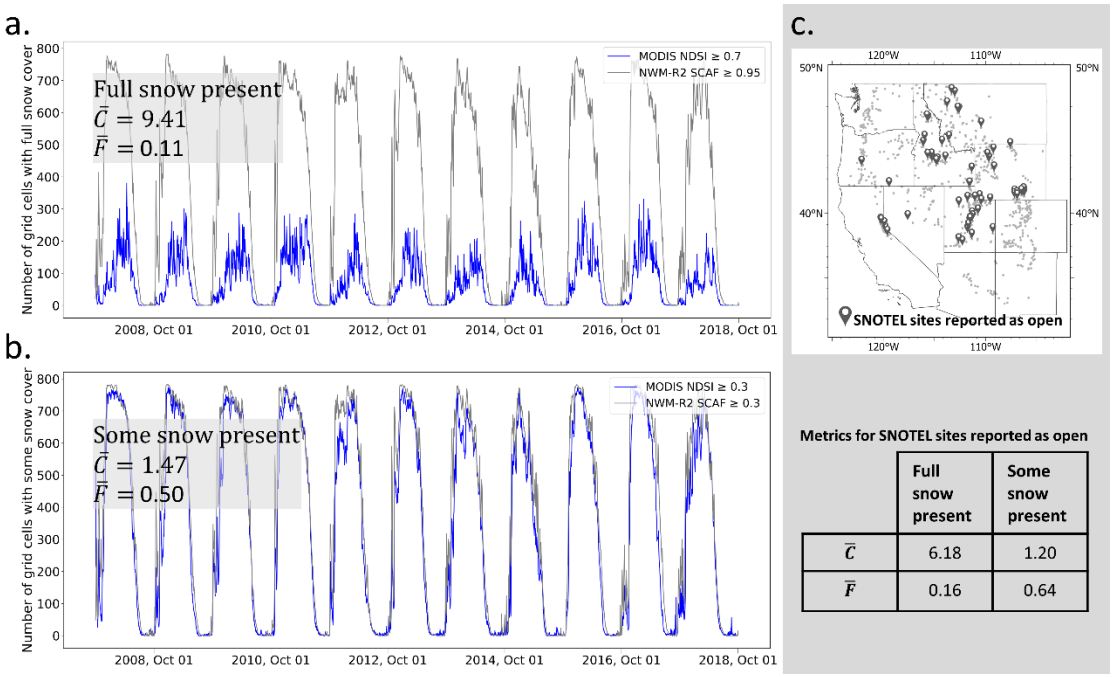

Figure 3.14 Direct (binary) comparison of snow presence considering all 734 SNOTEL sites with (a) full snow cover and (b) some snow cover. The modeled and observed thresholds for full snow cover were NWM-R2 SCAF≥ 0.95 and MODIS NDSI≥0.7, respectively. Lower thresholds were used for some snow cover (i.e., NWM-R2 SCAF>0.3 and MODIS NDSI>0.3). (c) Locations of the 62 SNOTEL sites reported as open. Average fit metrics (i.e., $\bar{C}$ and $\bar{F}$), presented here, quantitatively evaluate the degree-of-overlap between the modeled and observed snow presence.

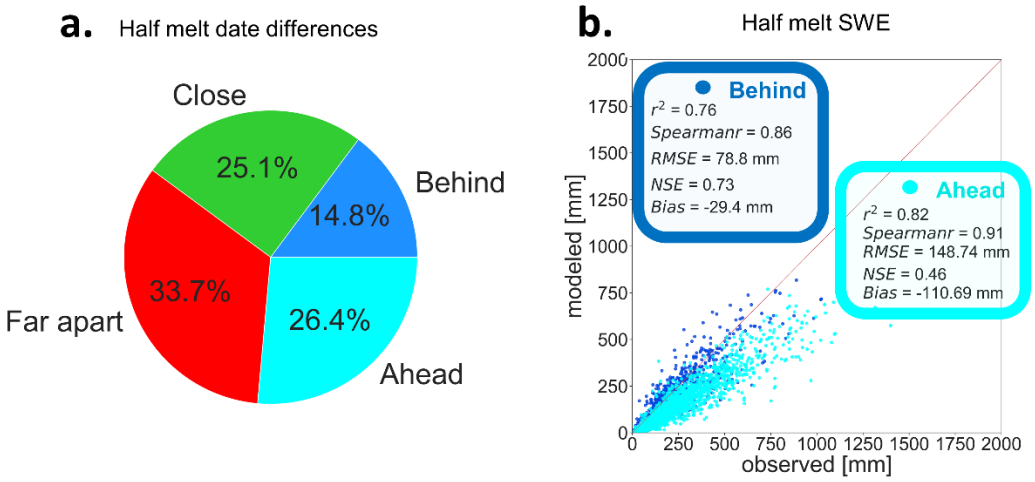

Figure 3.15 Analysis of melt timing. (a) Classification of differences between observed and modeled dates of half melt from peak SWE. Close: modeled and observed within 5 days of each other; Behind: modeled 6 to 19 days after observed; Ahead: modeled 6 to 19 days before observed; Far apart: Modeled and observed more than 20 days apart. (b) NWM-R2 SWE vs. SNOTEL SWE date of half melt from peak.

CHAPTER 4

EVALUATING INPUT DATA AND RAIN SNOW SEPERATION IMPROVEMENTS TO THE NATIONAL WATER MODEL SIMULATION OF SNOW WATER EQUIVALENT[1]

Abstract

We compared snowfall, and snow water equivalent (SWE) accumulation and ablation simulations from the WRF-Hydro model with the National Water Model (NWM) configuration against observations at a set of representative point locations from Snow Telemetry (SNOTEL) sites across the western U.S. We focused on the model's partitioning of precipitation between rain and snow and selected sites that span the variability of the percentage of rain on snow precipitation events. Our results show that the NWM generally under-estimates SWE and tends to melt snow earlier than observations in part due to errors in the precipitation and air temperature inputs. We reduced some of the observed and modeled discrepancies by using SNOTEL snow-adjusted precipitation and removing air temperature biases, based on observations. These input changes produced an average 59% improvement in the peak SWE. Modeled peak SWE was further improved using humidity-dependent rain-snow-separation. Both dew point and wet-bulb parameterizations were evaluated, with the dew-point parameterization giving better overall improvement, reducing the bias in SWE by 18% compared to the NWM air temperature-based scheme. This modification also improved melt timing with the number of site years having difference between modeled and observed date of half melt from peak SWE six or more days reduced by 6%. These SWE magnitude and timing improvements varied when analyzed for each rain-on-snow percentage class, with generally better results at sites where most precipitation events fall either as snow or as rain, and less improvement when there is a mix of snow and rain-on-snow events.

4.1 Introduction

Snow models are a central component of hydrologic forecasting systems when snow and snowmelt are the dominant influence on the regional streamflow. Decades of

[1] Co-authored by Irene Garousi-Nejad and David G. Tarboton

model development, combined with advances in technology and software engineering, have gradually enabled snowmelt runoff models to evolve into large-scale, high-resolution, and physically-based distributed models such as the National Oceanic and Atmospheric Administration (NOAA) National Water Model (NWM) in the U.S. (https://water.noaa.gov/about/nwm). This evolution was driven in part by the need to shorten the time interval for streamflow forecasts; to accommodate the shift from simple temperature-index based to energy balance methods; and to enable predicting the effects of anthropogenic and environmental changes such as those caused by land-use change or climate change on large heterogeneous basins (DeWalle & Rango, 2008). The NWM is now part of NOAA's water resources information system that provides timely hydrologic forecasts and data to support and inform emergency services and water resources decisions (https://water.noaa.gov).

To provide accurate predictions of seasonal water supplies over the continental U.S. under future changing conditions, the NWM, operated by the National Water Center, uses an energy balance model (Noah-MP) to solve the surface energy and water balances based on first principles of conservation of energy and mass to calculate snowmelt (Gochis, Barlage, Cabell, Dugger, et al., 2020; Niu et al., 2011). In Chapter 3 of this dissertation, we compared the Noah-MP models as implemented in the NWM version 2.0 retrospective simulations with snow observations at Snow Telemetry (SNOTEL) sites over the western U.S. and showed that the NWM generally underestimated snow water equivalent (SWE) early in the season and became progressively more biased later in the season compared to observations at SNOTEL sites, in part due to errors in inputs, notably precipitation and air temperature. However, the

discrepancies in model inputs were not the only sources of SWE differences. The SWE bias was persistent when the model precipitation input was relatively (statistically) close to the observed precipitation, suggesting that there were challenges in the current snow parameterization within the specific configuration of Noah-MP as implemented in the NWM version 2.0 retrospective configuration. We identified the current air temperature-dependent rain-snow-separation (RSS) parameterization within Noah-MP as a potential source of model error in SWE modeling, because this has been reported by other studies as a limitation of Noah-MP as used in the NWM (Chen et al., 2014; Liu et al., 2017; Wang et al., 2019). More generally, the accurate representation of RSS in hydrological models is important as the proportion of rainfall versus snowfall across mountainous regions changes, altering snowpack dynamics, streamflow timing and amount, and frequency of rain-on-snow events (Bales et al., 2006; Barnett et al., 2005; Gillies et al., 2012; Harpold et al., 2017; Knowles et al., 2006). Thus, research that evaluates the NWM performance and enhances model output accuracy through more realistic inputs and physics representations is essential. This motivated our focus on the NWM's partitioning of precipitation between rain and snow at sites selected to span the variability of precipitation events that were rain on snow present in the western US.

We addressed the following questions in this study:

- Question 1. To what degree are discrepancies in NWM SWE and RSS predictions due to input errors and how much could they potentially be improved if inputs were better?
- Question 2. How well does the NWM RSS (rainfall and snowfall separation) parameterization work in comparison to SNOTEL observations?

- Question 3. Do any other RSS parameterization methods yield more accurate snowfall compared to SNOTEL observations?
- Question 4. Does incorporating a statistically better RSS scheme into the NWM translate into appreciable improvements in modeling of SWE?
- Question 5. How do improvements in modeled SWE vary over sites grouped according to the percentage of precipitation events that are rain-on-snow?

In what follows, we first review prior literature used in this work (Section 2). We then describe the data and model we used (Section 3) followed by the method and numerical experiment design developed to answer our research questions (Section 4). We then compare gridded model results from each scenario simulated with point-scale measurements across the western U.S. (Section 5). Following that, we discuss limitations and uncertainties associated with the data and model providing perspective on the results presented and identifying areas for input data improvement and model enhancements (Section 6). Finally, we summarize our conclusions (Section 7) and provide links to data we used and codes we developed.

4.2 Background

Seasonal mountain snowpack has key implications for mid-to high-latitude regions such as the western U.S., storing water in the winter when snow falls and then releasing it as runoff in spring and summer when the snow melts and contributes (up to about 70%) to the total runoff in these regions (Li et al., 2017). The recently published Intergovernmental Panel on Climate Change (IPCC) report indicates a 0.29 million km^2 per decade decline in April snow cover extent—commonly used as an indicator of water

supply forecast for the following spring and summer season—in the Northern Hemisphere (Gulev et al., 2021). It is projected that seasonal snowpack decline will decrease water supplies for about 2 billion people this century (Mankin et al., 2015). In the western U.S., an average 30% decrease in areal extent of winter wet-day temperatures conducive to snowfall is projected (Klos et al., 2014). Given snowpack decline due to climate warming and its impact on water resources, accurate prediction of spring snowmelt will become increasingly important as the growing population demands more water and as operational agencies have to manage water under hydroclimate conditions outside of the historical record (Bhatti et al., 2016; Gergel et al., 2017; Mote, 2003; Mote et al., 2005).

Continued changes in the precipitation phase (rainfall, snowfall, or a mixture of both) are expected to alter snowpack dynamics, streamflow timing and amount, and frequency of rain-on-snow events; and thus present a new set of challenges for hydrologic modeling (Harpold et al., 2017; Musselman et al., 2018). RSS is one of the most sensitive parameterizations in simulating cold-region hydrological processes (Loth et al., 1993) and has a notable influence on the success of snowmelt models (Rutter et al., 2009). Despite advances in snowmelt modeling, most models rely on empirical algorithms based on air temperature to separate precipitation into rain and snow. For example, see the model comparison by Wen et al. (2013). These methods are empirical and ignore some of the physical processes involved in atmospheric formation of rain or snow where humidity and latent heat exchanges between a hydrometeor and the surrounding air play a role (Feiccabrino et al., 2015; Jennings et al., 2018). Such physical

process representations warrant consideration if models are to improve their predictability by reducing their dependence on empirical parameterizations.

Inaccurate RSS may result in errors in SWE, snow depth, and snow cover duration at both point and basin scale (Harder & Pomeroy, 2014; Wang et al., 2019) because snow can be produced in air temperatures slightly above freezing if the wet-bulb temperature (the temperature to which air is cooled by evaporating water into the air at constant pressure until it is saturated) is below about -2 ^{o}C (Stull, 2011). Ultimately, these errors propagate into the hydrological response (runoff and streamflow) of the watershed and land-atmosphere energy exchanges (Jennings et al., 2018; Mizukami et al., 2013). Some studies suggest that using dew point temperature, wet-bulb temperature, or psychrometric energy balance based RSS schemes, which consider the impact of atmospheric humidity in the energy budget of falling hydrometeors, improves the modeling of precipitation phase and the accuracy of partitioning between rain and snow (Behrangi et al., 2018; Harder & Pomeroy, 2013; Marks et al., 2013).

While there has been significant prior work on RSS, our goal was to evaluate the NWM snow model performance across a set of SNOTEL sites that are representative of various precipitation regimes (dominantly rainfall or snowfall, or rain-on-snow) across the western U.S., and to identify where model biases can be removed by using a more physically accurate RSS method. The RSS methods that we used here include the air temperature-based method from Jordan (1991) currently used in the NWM, the air temperature-based method developed by the U.S. Army Corps of Engineers (1956) as used in the Utah Energy Balance (UEB) model (Tarboton & Luce, 1996), the dew point temperature-based method used in the SNOBAL model (Marks et al., 1999), and the wet-

bulb temperature-based approach evaluated for the Variable Infiltration Capacity (Behrangi et al., 2018) and Noah-MP (Wang et al., 2019) models.

4.3 Data and Model

We used SNOTEL data, NWM input data, and an offline version of the WRF-Hydro model that serves as the basis for the NWM to evaluate different RSS parameterizations and their corresponding impact on the modeled SWE as detailed in the three subsections that follow.

4.3.1 SNOTEL Data

For more than 60 years, the automated SNOTEL network, currently consisting of 808 sites across the western U.S., has measured SWE using a pressure sensing snow pillow, precipitation (P) using a storage-type gage or tipping bucket, and air temperature (Ta) using a shielded thermistor sensor to monitor winter snow and inform spring and summer water supply forecasts. Our study used the daily snow-adjusted precipitation (start of the day) that accounts for uncertainty associated with snowfall measurements being subject to under-catch (Mote, 2003; Sun et al., 2019). We also used daily average air temperature and daily SWE (start of the day) at SNOTEL sites as a reference dataset to evaluate: (1) the snowfall fraction estimated from four different RSS parameterization methods, and (2) the accuracy of the NWM inputs (precipitation and air temperature) and outputs (SWE).

We recognize there are uncertainties associated with SNOTEL measurements that need to be considered in our analysis. However, SNOTEL provides the most comprehensive dataset we could obtain to explore our research questions because of its

long, historically continuous records of P, Ta, and SWE across the western U.S. For our analysis, we focused on SNOTEL sites where complete daily data were available for water years 2008-2020. This led to a set of 683 SNOTEL sites. Even though it would have been technically possible to set up simulations and run WRF-Hydro for all 683 sites, it would have been computationally prohibitive, and we decided to focus on a representative set of them for this research. To select a representative subset of SNOTEL sites, we used a random sampling within rain-on-snow classes that led to a group of 33 sites that spanned site rain-on-snow variability, described later, and for which we set up simulations and ran WRF-Hydro.

4.3.2 National Water Model Input Data

The NWM surface physiographic and atmospheric meteorological inputs (1 km spatial resolution and hourly temporal resolution) were made available to us by the NCAR team (D. Gochis and A. RafieeiNasab, personal communication, March 16, 2021) as a read only directory in the NCAR Cheyenne high-performance computer. The surface physiographic inputs included the model domain; initial conditions such as soil moisture, soil temperature, and snow states; geospatial inputs (such as topography, soil properties, land cover type, etc.) and parameter files (such as calibrated snowmelt factor used in calculation of the snow-covered area fraction). The meteorological inputs included the Analysis of Record for Calibration reanalysis dataset developed by NOAA National Weather Service (Kitzmiller et al., 2018; National Weather Service, Office of Water Prediction, 2021), hereafter referred to as AORC. AORC forcing data included incoming short- and longwave radiation, specific humidity, wind, air pressure, air temperature, and precipitation rate.

For each of the selected 33 SNOTEL sites we retrieved all required inputs for a four grid cell 2 km by 2 km area containing the SNOTEL site (Garousi-Nejad & Tarboton, 2021). Then, we transferred data from Cheyenne to Expanse, an eXtreme Science and Engineering Discovery Environment (XSEDE) supercomputer (Towns et al., 2014) where we ran WRF-Hydro. The first water year (2008) was used for model spin up and, while the SNOTEL data extended to 2020, NWM forcing data was not available for 2020 at the time this work was done. Therefore, we used the period 2009-2019 for model comparisons.

4.3.3 WRF-Hydro National Water Model Configuration Code

The NWM is a physically-based, distributed model based on the WRF-Hydro modeling framework (Gochis, Barlage, Cabell, Dugger, et al., 2020) that provides operational hydrological forecasts at 1 km spatial and hourly temporal resolution for snow across the entire continental U.S. The NWM has evolved beginning from version 1.0 (August 2016) to the current version 2.1 (October 2021) with improved soil/snow physics, calibration, and data assimilation. The core of the NWM system is WRF-Hydro, developed by the National Center for Atmospheric Research (NCAR), which consists of different modules with different geospatial representation (e.g., grids in the land surface and terrain routing modules connected to stream reaches in the channel routing module) and resolution (e.g., 1 km in the land surface module versus 250 m in the terrain routing module) to simulate land and atmosphere energy/water fluxes and storages. Details about the NWM and WRF-Hydro are available in Gochis, Barlage, Cabell, Casali, et al. (2020). We obtained the Fortran source code from the WRF-Hydro GitHub webpage (https://github.com/NCAR/wrf_hydro_nwm_public/releases/tag/v5.1.1, version 5.1.1

corresponding to the NWM version 2.0 available at the time this work started (Gochis, Barlage, Cabell, Dugger, et al., 2020). Releases beyond this to date include WRF-Hydro version 5.1.2 and version 5.2.0, both available in GitHub (https://github.com/NCAR/wrf_hydro_nwm_public/releases), but to our understanding the rain and snow separation parameterization that we evaluated has not been changed in these releases.

In this study, we focused on the land surface module of the NWM, which is a particular configuration of the Noah-MP model (Niu et al., 2011), where all snow processes are simulated within a 1-dimensional vertical column over 1 km spatial resolution grid cells. The Noah-MP module uses up to three snow layers to solve the energy balance (Equation 1) and water balance (Equation 2) between the snowpack, atmosphere, and the ground surface. The snow state variables for each snow layer are the mass of liquid water, the mass of ice, layer thickness, and layer temperature.

$$\frac{dU}{dt} = Q_{sw} + Q_{lw} + Q_{lt} + Q_{sn} + Q_g + Q_p + Q_m \quad (1)$$

$$\frac{dSWE}{dt} = P_{snow} - M - E \quad (2)$$

where U is the snowpack internal sensible and latent heat storage, t is time, Q_{sw} is net shortwave radiation flux, Q_{lw} is net longwave radiation flux, Q_{lt} is convective latent heat of vaporization/sublimation flux, Q_{sn} is convective sensible heat flux, Q_g is conductive ground heat flux, Q_m is heat of fusion energy flux due to meltwater leaving the snowpack (which is solved for as a residual in Equation 1), P_{snow} is the snowfall (in terms of water depth) that reaches the ground after adjusting for canopy interception, M is the meltwater, and E is snow sublimation/frost (Shuttleworth, 2012).

4.4 Methods and Numerical Experiment Design

4.4.1 Input Data Evaluation

The first step in our work was to compare the NWM inputs (elevation, P, and T_a for water years 2009-2019) with observations at representative SNOTEL sites. Results showed biases in model inputs that needed to be considered in the analysis. There were discrepancies of up to approximately 250 m between model elevation and the elevation of SNOTEL sites (Figure 4.1a). This may be a contributor to differences observed in the daily mean air temperature comparison due to the lapse rate (Figure 4.1b). For some years, we found artifacts in the air temperature inputs at three SNOTEL sites (Figure 4.2). After excluding these periods, we observed a negative bias (-0.53 ^{o}C) in AORC air temperatures compared to SNOTEL measurements (Figure 4.1b), meaning that T_a input to the NWM is generally colder than observations. There were no artifacts in AORC precipitation for the period of our study; however, we observed a downward bias of about -55 mm (Figure 4.1c) when comparing the annual precipitation (accumulated from October 1 through September 30 for each water year at each representative SNOTEL site). These observations were the basis for designing our initial numerical experiments (scenarios), where we attempted to reduce biases in model inputs (details are provided in Scenario 2 and Scenario 3 in Section 4.5).

4.4.2. Snow Rain Ratio

Evaluating simulated snowfall amounts from different RSS schemes is challenging due to the lack of reliable ground truth observations of the precipitation phase (Harpold et al., 2017). The Natural Resources Conservation Service (NRCS) reports a snow rain ratio (SNRR) for SNOTEL sites that estimates the fraction of

precipitation that falls as snowfall calculated as the ratio of daily SWE increases to daily P for the same period. In theory, the SNRR should range from 0 to 1, with 1 indicating all precipitation falls as snowfall. We obtained daily SNRR values from NRCS Report Generator version 2 for 683 SNOTEL sites for water years 2008-2020 using a Jupyter Notebook script we developed (Garousi-Nejad & Tarboton, 2021). We realized that this ratio was sometimes above 1 (100%) because it was calculated based on the daily P measurements which may be less than accumulated daily SWE. This may occur due to either precipitation measurement under-catch or processes that result in additional SWE being measured, such as snow drifting. The NRCS provides a snow-adjusted daily P estimate to account for this. We obtained this adjusted P and recalculated SNRR to get values within the range 0-1 (Algorithm 4.1). We used the computed SNRR values as a validation dataset to compare different rain/snow separation parameterizations. We acknowledge that there are uncertainties associated with this SNRR approach that may impact our analysis. However, this indicator was the best option available to us for evaluating RSS methods given the western-U.S.-wide dataset that we use in this study.

4.4.3 Representative SNOTEL Site Selection

We used the computed SNRR values to identify precipitation events that were rain-on-snow and classified sites based the percentage of rain-on-snow events they received to obtain a set to work with that spanned and is thus representative of the variability of rain-on-snow event percentages present across the western U.S. We designated precipitation events with SNRR >= 0.95 as snowfall and events with SNRR < 0.95 as rain-on-snow. We, thus, took rainfall or mixed rainfall and snowfall events for which SNRR < 0.95 as having a quantity of rain sufficient to be called rain-on-snow. We

calculated the percentage of precipitation events that were rain-on-snow (ROS%) for each SNOTEL site over water years 2008-2020 using a script we developed (Garousi-Nejad & Tarboton, 2021). For the 683 SNOTEL sites, ROS% values ranged between 30-100% (Figure 4.3a). We classified sites according to ROS% into seven groups each spanning a 10% class range. The largest number of sites fell in the 50-60% class, and the least frequent group (three sites) had ROS% between 90-100%.

To select the representative set of SNOTEL sites to work with, we randomly selected five sites from each class with ROS% between 30-90% and selected all members within the 90-100% class because it contained only three SNOTEL sites using a script we developed (Garousi-Nejad & Tarboton, 2021). This yielded a subset of 33 SNOTEL sites with different ROS% values spread across the western U.S. (Figure 4.3b). We obtained observed P, T_a, and SWE for these selected SNOTEL sites from NRCS Report Generator version 2 using Jupyter Notebook data retrieval scripts we developed (Garousi-Nejad & Tarboton, 2021).

4.4.4 Evaluation of Rain-Snow Separation (RSS) Parameterizations

We evaluated four different RSS schemes, including two air temperature-dependent and two humidity-dependent approaches, commonly used in hydrological models. The air temperature-based RSS schemes were from the U.S. Army Corps of Engineers, (U.S. Army Corps of Engineers, 1956; hereafter USCAE (1956)) as used in the UEB snow model (Tarboton & Luce, 1996), and Jordan (1991) as used in the current version of the NWM Noah-MP. The USACE (1956) T_a based method separates precipitation into rain and snow based on two temperature thresholds. All precipitation is rainfall if the air temperature is greater than or equal to 3 ^oC, snowfall if the air

temperature is less than or equal to -1 ^{o}C, and varies linearly for air temperature between -1 and 3 (Algorithm 4.2). The Jordan (1991) T_a based method uses multiple thresholds (0.5, 2, and 2.5 ^{o}C) to separate precipitation into rain and snow (Algorithm 4.3). Both these methods only consider air temperature (Figure 4.4a, 4.4b).

The humidity-based RSS approaches were from the dew point temperature method (Marks et al., 1999) as used in the SNOBAL model and the wet-bulb temperature based method evaluated for Noah-MP (Wang et al., 2019). Dew point temperature (T_d), a measure of the vapor pressure of the air (Equation 3), is defined as the temperature to which air must cool at constant pressure for it to saturate, without any moisture addition/removal (Marks et al., 2013; Shuttleworth, 2012):

$$T_d = \frac{\ln(e) + 0.49299}{0.0707 - 0.00421 \ln(e)} \quad (3)$$

where e is the vapor pressure of the air in kPa and T_d is the dew point temperature in ^{o}C.

Marks et al. (1999) described a dew point based approach that uses discrete steps to partition precipitation into rain and snow (Figure 4.4c, Algorithm 4.4). The discrete stepped nature of the approach seemed limiting as there do not appear to be physical reasons for such step changes. We thus developed a continuous version of Marks et al.'s (1999) method to provide a smoother function of T_d (Figure 4.4d).

Wet-bulb temperature (T_w) is defined as the temperature to which air is cooled by evaporating water into the air at constant pressure until it is saturated ($T_a \approx T_d \approx T_w$). According to thermodynamic laws, the air is thermally isolated in saturated environments. In other words, as the air cools to get to the saturation point, the heat (internal energy) removed from the air due to the cooling process must equal the latent heat required to evaporate water (from the hydrometeor surface in a precipitation event)

to raise the specific humidity of the air to saturation (Shuttleworth, 2012). This can be mathematically represented as Equation (4) which can be reformulated as the wet-bulb equation (Equation 5):

$$\rho_a V(T_a - T_w)c_p = \rho_a [q_{sat}(T_w) - q] V \lambda \quad (4)$$

$$es_w(T_w) - e = \frac{c_p P_{air}}{0.622 \lambda}(T_a - T_w) \quad (5)$$

where ρ_a is air density (kg/m^3), V is volume of air (m^3), T_a is (dry-bulb) air temperature (K), T_w is wet-bulb temperature (K), c_p is specific heat at constant pressure for air (1.04 kJ/kg K), $q_{sat}(T_w)$ is saturated specific humidity of air at T_w (kg/kg), q is specific humidity of air (kg/kg), λ is latent heat of vaporization (2.5 MJ/kg), $es_w(T_w)$ is the saturated vapor pressure of air at T_w (kPa), and P_{air} is air pressure (kPa). Equation (5) does not have an analytical inverse solution to calculate the wet-bulb temperature from air temperature and humidity (Stull, 2011), so was solved numerically using a Newton-Raphson scheme. We then used the sigmoid function of Wang et al. (2019) to calculate RSS (Algorithm 5).

4.4.5 RSS Modeling Experimental Design

We developed a set of modeling scenarios to answer the research questions given earlier. For each of the 33 representative SNOTEL sites selected, we used the WRF-Hydro version 5.1.1 NWM configuration in the following scenarios:

(1) Base scenario with AORC inputs. The hourly AORC forcing data was used to simulate snow processes from January 2008 to September 2019 (with the first nine months being set aside as model spin up) over 33 grid cells containing the representative SNOTEL sites. We call this scenario the base scenario as we kept all inputs and model settings the same as those used in the operational NWM version 2.0.

The outputs that we evaluated are hourly snowfall (from the Jordan (1991) RSS scheme) and SWE values.

(2) Replacing AORC precipitation with observations from SNOTEL (Observed precipitation scenario). Scenario 2 was the same as the base scenario except for the input precipitation. In our preparation step (Section 4.3.3), we showed a downward bias for AORC precipitation compared to observations at SNOTEL sites. To isolate the effects of AORC precipitation biases on modeled snowfall and SWE, we used the SNOTEL observed precipitation as supplemental precipitation to run the model. This means that the model used all other AORC inputs, but the precipitation data were read from the additional forcing inputs. To generate supplemental precipitation input files, we followed the steps described in Gochis et al. (2020). We resampled observed daily precipitation into hourly precipitation by dividing the total daily precipitation from SNOTEL sites equally into 24 hours using scripts we developed (Garousi-Nejad & Tarboton, 2021).

(3) Replacing AORC air temperature with bias corrected air temperature based on SNOTEL on top of the precipitation adjustments of Scenario 2 (Bias-corrected temperature scenario). Since we observed a negative bias in AORC air temperature compared to SNOTEL observations, we designed Scenario 3 to diminish the impact of errors in air temperature on the modeled snowfall and SWE. For each SNOTEL site we computed the average difference in daily temperature for the common data period (12 years) and used this difference to adjust the AORC hourly temperature inputs. This one difference value thus served as a bias correction offset for each representative SNOTEL site. The model physics settings were the same as in

Scenarios 1 and 2, and precipitation was from SNOTEL observations (as prepared in Scenario 2).

(4) Inputs prepared for Scenario 3 but with USACE (1956) air temperature RSS modifications to the code. In this scenario, we used inputs prepared for Scenario 3 to run the WRF-Hydro model modified to use the USACE (1956) air temperature based RSS scheme (Algorithm 4.2). This was achieved by editing the rain snow separation code in the module_noahmplsm.F source code file and recompiling the model.

(5) Inputs prepared for Scenario 3 but with continuous dew point based RSS based on Marks et al. (1999). In this scenario, we used inputs prepared for Scenario 3 to run the WRF-Hydro model modified to implement the continuous version of the Marks et al. (1999) dew point based RSS method (Algorithm 4.4). This was also achieved by editing the rain snow separation code in the module_noahmplsm.F source code file and recompiling the model.

(6) Inputs prepared for Scenario 3 but with Wang et al. (2019) wet-bulb based RSS. In this scenario, we used inputs prepared for Scenario 3 and implemented the Wang et al. (2019) wet-bulb based RSS parametrization (Algorithm 4.5) in the NWM code as for scenarios 4 and 5.

We ran these scenarios sequentially, and at each step evaluated the improvement in snowfall and SWE fidelity. Scenarios 1 to 3 represent improvements in input data, while scenarios 3 to 6 evaluate alternative RSS parameterizations using consistent best estimates of the input data.

4.4.6 Comparing Snow Accumulation and Melt

To assess the performance of the model, we first compared the computed snowfall amount from each RSS method and quantified the performance of each approach against observed RSS that was inferred from SNRR at SNOTEL sites through a set of statistical metrics, including Coefficient of Determination (r^2), Spearman's Rank Correlation (Spearmanr), Root Mean Square Error (RMSE), Nash Sutcliffe Efficiency (NSE), and Bias (Table 4.1). In addition to these statistical metrics, we used (1) SWE on observed peak date, (2) observed and modeled peak SWE, and (3) date of half melt from peak SWE metrics to compare the simulated SWE to observed SWE at SNOTEL sites (Chapter 3). First, we used the date on which peak SWE was observed to compare modeled SWE against observations. We refer to this comparison metric as a same-day comparison. Note that if there is a discrepancy in timing, model and observed peak SWE may be similar, while the model SWE on the observed peak date is different. To account for this the second metric compared observed and modeled peak SWE regardless of the dates when they occur. This is referred to as a different-day comparison in this study. This comparison may have limitations due to cumulative precipitation inputs being different up to the different dates. We did not report comparison of the Peak SWE timing because of variability associated with peak SWE time related to long periods where the SWE time series was flat near the peak. Instead, we chose the date of half melt from peak SWE as a metric to quantify the model's performance in terms of simulating the melt timing (Clow, 2010). This is the date (either modeled or observed) when half of the peak SWE has melted. To quantitively assess the difference between the modeled and observed half melt dates, we categorized the date differences into four groups—close,

model early, model late, and far apart (Chapter 3). Close indicates that modeled and observed half melt dates are within 5 days of each other. Model early refers to the situation where modeled half melt dates are 6 to 19 days before observed, while model late means that modeled half melt dates are 6 to 19 days after observed. Lastly, far apart means that modeled an observed half melt dates are more than 20 days apart.

4.5 Results

4.5.1 Changes in Snowfall

We compared the estimated annual snowfall magnitude from five different RSS methods with the observations inferred from SNRR from SNOTEL and found a persistent upward bias in snowfall from all methods (Figure 4.5). This is an average bias across all 33 sites and all years. USACE (1956) T_a based showed the smallest bias (about 6 mm) and Marks et al. (1999) T_d based (continuous version) had the most significant bias (about 45 mm). Results for Jordan (1991) T_a based (the current RSS scheme in the NWM Noah-MP) were slightly better than the dew point temperature-based (both discrete and continuous) methods (Figure 4.5b, 4.5c, and 4.5d). Among the two humidity-based methods, Wang et al. (2019) T_w based showed a smaller bias (more than 10 mm smaller), but its bias was still six times larger than USACE (1956) T_a based (Figure 4.5d and 4.5a).

The seasonal variations (11-year daily averages across selected SNOTEL sites) of accumulated snowfall from all methods indicated that more than 70% of the annual precipitation during February through May, independent of the RSS method, fell as snowfall averaged across the SNOTEL sites and water years (Figure 4.5f). Observations and USACE (1956) T_a based average accumulation matched well over the entire year.

The other RSS methods tracked above observations and were all close together during the accumulation phase (October through May). Following May, Marks et al. (1999) T_d based (continuous version) diverged and produced more snowfall than other RSS methods and observations (50% more than observed in May). Also, Marks et al. (1999) T_d based was the only RSS method that showed 19% and 17% of precipitation falling as snowfall during July and September, respectively. This sets the Marks et al. (1999) T_d based method apart from other methods as the only one that estimated snowfall during warmer months (Figure 4.5f). Average air, wet-bulb, and dew point temperatures for each day across all site years indicated the general differences between these quantities that were inputs to the RSS methods (Figure 4.5g).

4.5.2 Snow Water Equivalent on Observed Peak Date (Same-day Comparison)

The comparison between modeled and observed SWE on the date of observed peak SWE revealed a general downward bias in modeled SWE (Figure 4.6), suggesting that the NWM generally underestimated SWE on the date of observed peak SWE, independent of the model input errors (shown before in Figure 4.1) and model physics (specifically in terms of the different RSS methods as shown before in Figure 4.5). However, biases in modeled SWE were reduced when using observed precipitation instead of AORC precipitation, from -228 mm in the base scenario to -92 mm in the observed precipitation scenario (Figure 4.6b). This emphasizes the importance of using high-quality input forcing in the NWM. Even though we further reduced model input errors/biases by correcting the AORC air temperature biases, this did not improve SWE estimates (Figure 4.6c). Contrarily, it increased the downward bias in SWE. This should

not be considered as a negative point as it is essential to have correct/accurate inputs, even though that may not necessarily translate into improvements in model outputs.

Even though our comparison of annual snowfall magnitude from different RSS methods (Figure 4.5) showed that USACE (1956) T_a based had the best agreement with observations, this agreement did not translate to the best same-day SWE comparison. Among the four RSS comparisons, when the best input estimates were used (Scenarios 3 to 6), USACE (1956) T_a based showed the largest negative bias (about -168 mm) and Marks et al. (1999) T_d based showed the least bias (about -111 mm) and best NSE and RMSE (Figure 4.6c, 4.6d, 4.6e, and 4.6f). Similar to the snowfall comparison, the modeled SWE from the current NWM RSS scheme (Jordan (1991) T_a based) and Wang et al. (2019) T_w based had almost statistically identical behavior when compared to SWE observations (Figure 4.6c versus 4.6f).

4.5.3 Observed and Modeled Peak Snow Water Equivalent (Different-day Comparison)

Under-modeling of SWE was also evident in our comparison of observed and modeled peak SWE noting that the observed and modeled peak SWE do not necessarily occur on the exact same date (Figure 4.7). Among the four RSS schemes modeled (Scenarios 3 to 6) the dew point temperature-based scheme (Scenario 5) provided less biased modeled SWE similar to the same-day comparison. In general, these different day peak SWE comparisons had smaller error metrics than the comparisons presented above for the day of observed peak SWE.

4.5.4 Seasonal Snow Water Equivalent

The seasonal pattern of SWE averaged across the representative SNOTEL sites indicated the general under-modeling of SWE relative to observations at SNOTEL sites

in all scenarios, with USACE (1956) T_a based scheme (Scenario 3) being further apart from and Marks et al. (1999) T_d based scheme (Scenario 5) being the closest to the observations (Figure 4.8a). For the purpose of evaluating RSS options, we did not include results from scenarios that had inferior inputs (Scenarios 1 and 2) in this comparison. Furthermore, our results showed that discrepancies between seasonal patterns of SWE vary when analyzed for each ROS percentage class (Figure 4.8b-g). For SNOTEL sites with the smallest ROS% (30-40%, meaning that most precipitation events fall on average as snow), all RSS methods simulated almost identical SWE (Figure 4.8b). However, as ROS% increased, the impact of different RSS methods in modeling SWE became more evident in such a way that the T_d based RSS SWE simulations almost always stayed above the SWE from other RSS methods, meaning that it produced more SWE compared to other RSS methods. For the sites with ROS% between 80-100 (where rain-on-snow events are dominant), the T_d based RSS scheme simulated SWE was almost identical to observations during the accumulation period, October-March, while the other RSS methods underestimated SWE (Figure 4.8g). During the melt period all methods tended to melt the snow a bit slowly compared to observations, a difference likely due to model considerations other than RSS.

4.5.5 Melt Timing Comparison (Half Melt from Peak Snow Water Equivalent Date)

Our comparison of the modeled half melt date (from scenarios that had valid inputs) with observations showed that the modeled half melt date was generally earlier than observations for more than 60% of the site-years (Table 4.2). When further classified depending on whether the differences between observed and modeled half melt dates from peak SWE were close, ahead, behind or far apart from observed melt dates,

we observed that the NWM half melt date was off by 6 days or more for about 75% of site years (Figure 4.9a). This became even more noticeable when using the USACE (1956) T_a based RSS method (Figure 4.9b showing that about 79% of site-years deviated by 6 days or more from observations). Our results show that using humidity-based RSS methods improved the early melt issue in the NWM to some extent (Figure 4.9c and 4.9d), with the T_d based RSS method showing the most considerable degree of improvement compared to other RSS methods.

The NWM early melt issue inferred from the half melt date comparison between modeled results (Scenario 4 with Marks et al. (1999) T_d based method) and observations at selected SNOTEL sites during 11 years (the water year 2009-2019) was persistent across all sites but varied differently across ROS% classes (Figure 4.10). In this figure, the ROS% classes in the middle of the range, which represent sites with rain and snow mixes, as opposed to dominantly snow or dominantly rain, tended to have smaller percentages with close melt timing. For the sites where ROS% events were significantly high (>80%) or low (<40%), the modeled half melt date was close (off 6 days or less) more frequently (Figure 4.10a and 4.10f).

4.6 Discussion

In this study, our goal was to evaluate input data and three alternative RSS parameterizations to the NWM version 2.0 to find whether these improve SWE simulations. This section discusses findings for each of the research questions given in the introduction.

To what degree are discrepancies in NWM SWE and RSS predictions due to input errors and how much could they potentially be improved if inputs were better?

In this experiment, the most noticeable improvements in modeling SWE compared to the base scenario were achieved when we used observed precipitation from SNOTEL sites instead of the NWM AORC precipitation data (about 60% and 77% improvements in bias for same-day and different-day comparisons of peak SWE, respectively). Using better meteorological inputs to improve NWM performance has been reported by other studies (Lahmers et al., 2019; Viterbo et al., 2020). While stating that better inputs lead to better model performance is not new, this emphasizes the sensitivity to hydrometeorological input error, specifically precipitation and near-surface air temperature, in hydrological modeling predictions (Förster et al., 2014; Raleigh et al., 2015; Zehe et al., 2005).

Our model evaluation that quantifies how much the NWM performance in modeling SWE could improve by using more accurate meteorological inputs is important in considering where to invest time and effort in enhancing the NWM overall. We understand that model input improvements do not per se improve hydrologic process understanding; however, the ability to produce accurate hydrological forecasts is essential, and beyond forecast quality, the NWM does provide several outputs of hydrologic quantities, either not observed, or only observed in specialized field studies, and certainly not comprehensively across a continent. Examination of these outputs and their patterns across a continent does enhance process understanding. In addition, developing more accurate gridded precipitation products may reduce the need to make existing physical parameterizations more complex and add more uncertainties to the

model due to new parameters (e.g., best fit coefficients in the Wang et al. (2019) T_w based approach).

How well does the NWM RSS (rainfall and snowfall separation) parameterization work in comparison to SNOTEL observations?

Our results showed that the NWM RSS (Jordan (1991) T_a based) performed statistically poorly (bias 41 mm, RMSE 74 mm) in separating precipitation into rain and snow compared to observed snowfall inferred from SNRR at 33 representative SNOTEL sites across the western U.S. Several challenges exist in this comparison, and each can be considered as a contributor to discrepancies observed. First, the spatial scale differences between SNOTEL and NWM datasets are a source of uncertainty in this analysis. As with all numerical models, the representation of sub-grid variability of snow processes may not be well parameterized when working with models such as the NWM that simulate snow processes across 1 km spatial resolution. Second, even though we used snow-adjusted precipitation from SNOTEL sites, there may still be systematic bias for SNOTEL precipitation due to under-catch (Mote, 2003; Sun et al., 2019). Third, even though we used observed precipitation from SNOTEL sites (instead of AORC precipitation that had downward bias) along with bias-corrected AORC air temperatures (corrected based on SNOTEL observations), there may still be uncertainties associated with other NWM AORC inputs, including specific humidity, in RSS calculations. Fourth, the method for inferring SNRR from SNOTEL measurements of precipitation and SWE has limitations. For example, rain that falls on a cold snowpack, freezes and adds to SWE mass will increase SWE and be interpreted to be snowfall. Other processes such as wind drifting or scouring of SWE at the SNOTEL site also introduce uncertainty. Lastly, while

when SWE increases were more that P measurements they were used to infer and adjust for P under-catch, this does not adjust for under-catch of rainfall that may be present, even though it is commonly not thought to be as problematic as under-catch of snowfall (e.g., Meyer et al., 2012).

Do any other RSS parameterization methods yield more accurate snowfall compared to SNOTEL observations?

When considering other RSS alternatives from the literature, we observed that the dual-threshold air temperature-based method (USACE (1956) T_a based) yields noticeably better agreement between modeled and observed snowfall (bias 6 mm, RMSE 54 mm) compared to the other two humidity-based approaches (T_d based and T_w based). This may be interpreted as good, because it would be easier to apply a dual-threshold method with a linear decrease in between that takes only air temperature as the input to separate precipitation into rain and snow than T_d based or T_w based methods that determine the snowfall fraction using humidity information which potentially could add more errors if input data are not accurate. This finding is in line with the work of Feiccabrino et al. (2013) that reported on the superiority of the air temperature-based method over the dew point temperature approach based on data from 19 Swedish meteorological stations.

However, we should consider that this finding may be based on some assumptions that hinder us from concluding that USACE (1956) T_a based is the best among other methods tested in this study. Firstly, there are uncertainties associated with the NWM AORC data (even with our bias removal from precipitation and air temperature) we used as inputs to RSS methods and the reference data (SNRR) that we used to evaluate the performance of each RSS scheme. Secondly, even though air temperature-based RSS

schemes are easy to use, they are empirically-based methods that have been developed based on historical data. Physically based methods are theoretically preferable for the simulation of processes under conditions that may differ from the historical conditions where empirical methods have been calibrated or optimized. We note that other studies report on the superiority of humidity-based approaches over air temperature-based ones in modeling both snowfall and SWE over mountainous regions (Ding et al., 2014; Marks et al., 2013; Wang et al., 2019). Further, as noted above, there are limitations associated with the SNOTEL inferred SNRR that may merit giving higher consideration to overall SWE simulation comparisons than snowfall ratio comparisons in assessing a RSS model. This is discussed below.

In this study, our results showed that snowfall estimates from Wang et al. (2019) T_w based scheme better agreed with observations inferred from SNRR at SNOTEL sites (Figure 4.5e: bias 34 mm, RMSE 63 mm) than those from Marks et al. (1999) T_d based scheme (Figure 4.5d: continuous version with bias 45 mm and RMSE 76 mm). This difference could be because T_w is more physically related to the precipitation phase as it considers the sensible and latent heat fluxes that determine the internal energy and temperature of a hydrometeor, and thus it is closer to the surface temperature of a falling hydrometeor than the air temperature (Wang et al., 2019). However, T_d only describes the cooling necessary for an unsaturated parcel of air to reach saturation over constant pressure, and it does not consider sensible and latent heat fluxes to the hydrometeor (Harder & Pomeroy, 2013). There may also be uncertainty related to best fit coefficients in the Wang et al. (2019) snowfall fraction equation that has been optimized to fit the

observation-based relationship between snowfall probability and the T_w from Behrangi et al. (2018).

Does incorporating a statistically better RSS scheme into NWM translate into appreciable improvements in modeling of SWE?

Not only did incorporating a statistically better RSS scheme (Scenario 4 with USACE (1956) T_a based scheme) not translate into appreciable improvements in SWE estimates, but it turned out that this scheme was the least acceptable among the RSS alternatives evaluated when compared to SNOTEL SWE observations (evident in both same day and different day comparison of peak SWE).

When using observed precipitation and unbiased air temperature, our analysis showed that the humidity-dependent RSS schemes (dew point and wet-bulb temperature based) overcame the under-modeling of SWE to some extent. This is in line with previous work reporting on the impact of incorporating humidity into RSS processes on snowfall and snow mass compared to ground-based snow products (Behrangi et al., 2018; Jennings et al., 2018; Marks et al., 2013; Wang et al., 2019). In our study, while the Wang et al. (2019) T_w based RSS method showed better snowfall results than those from the Marks et al. (1999) T_d based RSS scheme, we found greater improvements in modeled SWE from the T_d based than T_w based RSS scheme (Figures 4.6 and 4.7). We give this finding that the T_d based RSS scheme performs better for direct comparisons against SNOTEL SWE observations greater credence than the USACE T_a based method performing best against inferred snowfall, due to the limitations associated with the SNOTEL SNRR separation method, and due to predictions of SWE being an ultimate target of this modeling. There was, however, remaining under-modeling of SWE which

could be due to shortcomings associated with other meteorological inputs such as incoming solar and long-wave radiation which we did not study in this work and snow processes parameterizations in the NWM Noah-MP, such as the snow cover fraction calculations which have been reported to be problematic in modeling of SWE (Helbig et al., 2015; Magand et al., 2014; Wrzesien et al., 2015). These are open areas for future research to advance snow modeling in the NWM.

Collectively, our results showed that, on average, the NWM tended to melt snow early compared to observations at SNOTEL sites independent of the RSS scheme being used. However, the humidity-dependent approaches showed slightly better results. This observation that the modeling of melt timing was not significantly sensitive to the RSS scheme suggests that there is a need to investigate the overall energy balance and snow surface temperature calculations in the model.

How do improvements in modeled SWE vary over sites grouped according to the percentage of precipitation events that are rain on snow?

We observed that the degree of improvement in modeled SWE (in terms of both magnitude and melt timing) varied across ROS% classes. SWE was not well modeled for the ROS% classes in the middle rain dominated part of the range (60-80%), while at the lower end (predominantly snow) or higher end (predominantly rain) the model performed better. For these ROS% classes where the model performs better, Marks et al. (1999) T_d based separation gave the best improvements. A caveat of this analysis is that we characterized the representative SNOTEL sites based on the ROS% events metric that we computed based on the inferred precipitation phase from SNRR. We understand that this

approach has limitations; however, without direct rainfall and snowfall measurements, which are rare across larger areas, it was the approach that was available to us.

4.7 Conclusions

Two key points emerge from this work. First, our comparison of the National Water Model (NWM) Noah-MP snow water equivalent (SWE) and SNOTEL snow water equivalent for representative sites and dates in the 2009-2019 water years reiterated that the accuracy of model inputs plays a key role in the accuracy of model outputs. Results showed that using observed precipitation and bias-corrected air temperature significantly improved the general downward bias in the NWM SWE magnitude and slightly improved early half melt timing of NWM compared to observations at representative SNOTEL sites across the western U.S. Second, our evaluation of three alternative RSS parameterizations in the NWM across a set of representative SNOTEL sites that spanned site rain-on-snow variability indicated that the negative bias in NWM SWE can be reduced, on average, by using RSS methods that incorporate specific humidity information in precipitation separation into rain and snow with consistent best estimates of the input data. Among the two humidity-based RSS schemes, the dew point temperature-based method was slightly better (smaller RMSE and Bias and larger NSE) than the wet-bulb temperature-based method at simulating peak SWE. Using the dew point temperature-based RSS also improved the modeling of melt timing slightly (early melt inferred from the half melt date comparison). Both SWE magnitude and timing varied across ROS% classes, with better results for the ROS% classes at the lower end (predominantly sow) or higher end (predominantly rain). These findings support the

benefit of including physically based process representations in a model such as the NWM. Future work is required to assess the impact of improved SWE on streamflow.

OPEN RESEARCH – DATA AND MODEL AVAILABILITY

Codes developed for this research and the data we specifically used are publicly available in the HydroShare repository (Garousi-Nejad & Tarboton, 2021).

The data and model sources that we drew from include:

- SNOTEL data accessed through the NRCS Report Generator v2: https://wcc.sc.egov.usda.gov/reportGenerator/
- WRF-Hydro version 5.1.1 source code was accessed in GitHub: https://github.com/NCAR/wrf_hydro_nwm_public/releases/tag/v5.1.1
- NWM physiographic and atmospheric meteorological inputs were made available to us by the NCAR team in the NCAR Cheyenne high-performance computer. The specific data we used from this source are in the HydroShare resource given above.

ACKNOWLEDGMENTS

This work was completed on the land of Eastern Shoshone Tribe, and was supported by the Utah Water Research Laboratory and National Science Foundation under collaborative grants OAC-1664061 and OAC-1664119. This work used compute allocation TG-EAR190007 from the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562 (Towns et al., 2014). We thank David Gochis at NCAR and Ed Clark at the National Water Center for facilitating access to the NWM inputs. Thanks to Mahidhar Tatineni at the San Diego Supercomputer Center who helped to optimize our computational work load on XSEDE. Thanks to Mahyar Aboutalebi for his help with computational simulation runs on XSEDE, and to Jeffery S. Horsburgh for his comments and suggestions.

REFERENCES

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., & Dozier, J. (2006). Mountain hydrology of the western United States: MOUNTAIN HYDROLOGY OF THE WESTERN US. Water Resources Research, 42(8). https://doi.org/10.1029/2005WR004387

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. Nature, 438(7066), 303–309. https://doi.org/10.1038/nature04141

Behrangi, A., Yin, X., Rajagopal, S., Stampoulis, D., & Ye, H. (2018). On distinguishing snowfall from rainfall using near-surface atmospheric information: Comparative analysis, uncertainties and hydrologic importance. Quarterly Journal of the Royal Meteorological Society, 144(S1), 89–102. https://doi.org/10.1002/qj.3240

Bhatti, A. M., Koike, T., & Shrestha, M. (2016). Climate change impact assessment on mountain snow hydrology by water and energy budget-based distributed hydrological model. Journal of Hydrology, 543, 523–541. https://doi.org/10.1016/j.jhydrol.2016.10.025

Chen, F., Liu, C., Dudhia, J., & Chen, M. (2014). A sensitivity study of high-resolution regional climate simulations to three land surface models over the western United States: SENSITIVITY STUDY OF LSMS IN WRF. Journal of Geophysical Research: Atmospheres, 119(12), 7271–7291. https://doi.org/10.1002/2014JD021827

Clow, D. W. (2010). Changes in the timing of snowmelt and streamflow in Colorado: A response to recent warming. Journal of Climate, 23(9), 2293–2306. USGS Publications Warehouse. https://doi.org/10.1175/2009JCLI2951.1

DeWalle, D. R., & Rango, A. (2008). Principles of Snow Hydrology. Cambridge University Press. https://doi.org/10.1017/CBO9780511535673

Ding, B., Yang, K., Qin, J., Wang, L., Chen, Y., & He, X. (2014). The dependence of precipitation types on surface elevation and meteorological conditions and its parameterization. Journal of Hydrology, 513, 154–163. https://doi.org/10.1016/j.jhydrol.2014.03.038

Feiccabrino, J., Graff, W., Lundberg, A., Sandström, N., & Gustafsson, D. (2015). Meteorological Knowledge Useful for the Improvement of Snow Rain Separation in Surface Based Models. Hydrology, 2(4), 266–288. https://doi.org/10.3390/hydrology2040266

Feiccabrino, J., Gustafsson, D., & Lundberg, A. (2013). Surface-based precipitation phase determination methods in hydrological models. Hydrology Research, 44(1), 44–57. https://doi.org/10.2166/nh.2012.158

Förster, K., Meon, G., Marke, T., & Strasser, U. (2014). Effect of meteorological forcing and snow model complexity on hydrological simulations in the Sieber catchment (Harz Mountains, Germany). Hydrology and Earth System Sciences, 18(11), 4703–4720. https://doi.org/10.5194/hess-18-4703-2014

Garousi-Nejad, I., & Tarboton, D. G. (2021). Data for Evaluating Input Data and Rain Snow Separation Improvements to the National Water Model Simulation of Snow Water Equivalent. HydroShare. http://www.hydroshare.org/resource/bdbecdef23b14848b5da46c4f465ec21

Gergel, D. R., Nijssen, B., Abatzoglou, J. T., Lettenmaier, D. P., & Stumbaugh, M. R. (2017). Effects of climate change on snowpack and fire potential in the western USA. Climatic Change, 141(2), 287–299. https://doi.org/10.1007/s10584-017-1899-y

Gillies, R. R., Wang, S.-Y., & Booth, M. R. (2012). Observational and Synoptic Analyses of the Winter Precipitation Regime Change over Utah. Journal of Climate, 25(13), 4679–4698. https://doi.org/10.1175/JCLI-D-11-00084.1

Gochis, D., Barlage, M., Cabell, R., Casali, M., Dugger, A., FitzGerald, K., McAllister, M., McCreight, J., RafieeiNasab, A., Read, L., Sampson, K., Yates, D., & Zhang, Y. (2020). The WRF-Hydro® modeling system technical description, (Version 5.1.1). NCAR Technical Note. https://ral.ucar.edu/sites/default/files/public/WRFHydroV511TechnicalDescription.pdf

Gochis, D., Barlage, M., Cabell, R., Dugger, A., Fanfarillo, A., FitzGerald, K., McAllister, M., McCreight, J., RafieeiNasab, A., Read, L., Frazier, N., Johnson, D., Mattern, J. D., Karsten, L., Mills, T. J., & Fersch, B. (2020). WRF-Hydro® v5.1.1 (v5.1.1) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.3625238

Gulev, S. K., Thorne, P. W., Ahn, J., Dentener, F. J., Domingues, C. M., Gerland, S., Gong, D., Kaufman, D. S., Nnamchi, H. C., Quaas, J., Rivera, J. A., Sathyendranath, S., Smith, S. L., Trewin, B., von Schuckmann, K., & Vose, R. S. (2021). Changing State of the Climate System (In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change). Cambridge University Press. In Press.

Harder, P., & Pomeroy, J. (2013). Estimating precipitation phase using a psychrometric energy balance method: PRECIPITATION PHASE USING A PSYCHROMETRIC ENERGY BALANCE. Hydrological Processes, 27(13), 1901–1914. https://doi.org/10.1002/hyp.9799

Harder, P., & Pomeroy, J. W. (2014). Hydrological model uncertainty due to precipitation-phase partitioning methods: HYDROLOGIC MODEL

UNCERTAINTY OF PRECIPITATION-PHASE METHODS. Hydrological Processes, 28(14), 4311–4327. https://doi.org/10.1002/hyp.10214

Harpold, A. A., Kaplan, M. L., Klos, P. Z., Link, T., McNamara, J. P., Rajagopal, S., Schumer, R., & Steele, C. M. (2017). Rain or snow: Hydrologic processes, observations, prediction, and research needs. Hydrology and Earth System Sciences, 21(1), 1–22. https://doi.org/10.5194/hess-21-1-2017

Helbig, N., van Herwijnen, A., Magnusson, J., & Jonas, T. (2015). Fractional snow-covered area parameterization over complex topography. Hydrology and Earth System Sciences, 19(3), 1339–1351. https://doi.org/10.5194/hess-19-1339-2015

Jennings, K. S., Winchell, T. S., Livneh, B., & Molotch, N. P. (2018). Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere. Nature Communications, 9(1), 1148. https://doi.org/10.1038/s41467-018-03629-7

Jordan, R. E. (1991). A One-dimensional temperature model for a snow cover: Technical documentation for SNTHERM.89. Cold Regions Research and Engineering Laboratory (U.S.). http://hdl.handle.net/11681/11677

Kitzmiller, D. H., Wu, H., Zhang, Z., Patrick, N., & Tan, X. (2018). The Analysis of Record for Calibration: A High-Resolution Precipitation and Surface Weather Dataset for the United States. American Geophysical Union, Fall Meeting, Washington, D.C. https://ui.adsabs.harvard.edu/abs/2018AGUFM.H41H..06K/abstract

Klos, P. Z., Link, T. E., & Abatzoglou, J. T. (2014). Extent of the rain-snow transition zone in the western U.S. under historic and projected climate: Climatic rain-snow transition zone. Geophysical Research Letters, 41(13), 4560–4568. https://doi.org/10.1002/2014GL060500

Knowles, N., Dettinger, M. D., & Cayan, D. R. (2006). Trends in Snowfall versus Rainfall in the Western United States. Journal of Climate, 19(18), 4545–4559. https://doi.org/10.1175/JCLI3850.1

Lahmers, T. M., Gupta, H., Castro, C. L., Gochis, D. J., Yates, D., Dugger, A., Goodrich, D., & Hazenberg, P. (2019). Enhancing the Structure of the WRF-Hydro Hydrologic Model for Semiarid Environments. Journal of Hydrometeorology, 20(4), 691–714. https://doi.org/10.1175/JHM-D-18-0064.1

Li, D., Wrzesien, M. L., Durand, M., Adam, J., & Lettenmaier, D. P. (2017). How much runoff originates as snow in the western United States, and how will that change in the future?: Western U.S. Snowmelt-Derived Runoff. Geophysical Research Letters, 44(12), 6163–6172. https://doi.org/10.1002/2017GL073551

Liu, C., Ikeda, K., Rasmussen, R., Barlage, M., Newman, A. J., Prein, A. F., Chen, F., Chen, L., Clark, M., Dai, A., Dudhia, J., Eidhammer, T., Gochis, D., Gutmann, E., Kurkute, S., Li, Y., Thompson, G., & Yates, D. (2017). Continental-scale

convection-permitting modeling of the current and future climate of North America. Climate Dynamics, 49(1–2), 71–95. https://doi.org/10.1007/s00382-016-3327-9

Loth, B., Graf, H.-F., & Oberhuber, J. M. (1993). Snow cover model for global climate simulations. Journal of Geophysical Research, 98(D6), 10451. https://doi.org/10.1029/93JD00324

Magand, C., Ducharne, A., Le Moine, N., & Gascoin, S. (2014). Introducing Hysteresis in Snow Depletion Curves to Improve the Water Budget of a Land Surface Model in an Alpine Catchment. Journal of Hydrometeorology, 15(2), 631–649. https://doi.org/10.1175/JHM-D-13-091.1

Mankin, J. S., Viviroli, D., Singh, D., Hoekstra, A. Y., & Diffenbaugh, N. S. (2015). The potential for snow to supply human water demand in the present and future. Environmental Research Letters, 10(11), 114016. https://doi.org/10.1088/1748-9326/10/11/114016

Marks, D., Domingo, J., Susong, D., Link, T., & Garen, D. (1999). A spatially distributed energy balance snowmelt model for application in mountain basins. Hydrological Processes, 13(12–13), 1935–1959. https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<1935::AID-HYP868>3.0.CO;2-C

Marks, D., Winstral, A., Reba, M., Pomeroy, J., & Kumar, M. (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. https://doi.org/10.1016/j.advwatres.2012.11.012

Meyer, J. D. D., Jin, J., & Wang, S.-Y. (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. https://doi.org/10.1175/JHM-D-12-066.1

Mizukami, N., Koren, V., Smith, M., Kingsmill, D., Zhang, Z., Cosgrove, B., & Cui, Z. (2013). The Impact of Precipitation Type Discrimination on Hydrologic Simulation: Rain–Snow Partitioning Derived from HMT-West Radar-Detected Brightband Height versus Surface Temperature Data. Journal of Hydrometeorology, 14(4), 1139–1158. https://doi.org/10.1175/JHM-D-12-035.1

Mote, P. W. (2003). Trends in snow water equivalent in the Pacific Northwest and their climatic causes: TRENDS IN SNOW WATER EQUIVALENT. Geophysical Research Letters, 30(12). https://doi.org/10.1029/2003GL017258

Mote, P. W., Hamlet, A. F., Clark, M. P., & Lettenmaier, D. P. (2005). DECLINING MOUNTAIN SNOWPACK IN WESTERN NORTH AMERICA*. Bulletin of the American Meteorological Society, 86(1), 39–50. https://doi.org/10.1175/BAMS-86-1-39

Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., Barlage, M., & Rasmussen, R. (2018). Projected increases and shifts in rain-on-snow flood risk over western North America. Nature Climate Change, 8(9), 808–812. https://doi.org/10.1038/s41558-018-0236-4

National Weather Service, Office of Water Prediction. (2021). Analysis of Record for Calibration: Version 1.1 Sources, Methods, and Verification. NOAA. https://hydrology.nws.noaa.gov/aorc-historic/Documents/AORC-Version1.1-SourcesMethodsandVerifications.pdf

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., & Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. Journal of Geophysical Research, 116(D12), D12109. https://doi.org/10.1029/2010JD015139

Raleigh, M. S., Lundquist, J. D., & Clark, M. P. (2015). Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. Hydrology and Earth System Sciences, 19(7), 3153–3179. https://doi.org/10.5194/hess-19-3153-2015

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., Barr, A., Bartlett, P., Boone, A., Deng, H., Douville, H., Dutra, E., Elder, K., Ellis, C., Feng, X., Gelfan, A., Goodbody, A., Gusev, Y., Gustafsson, D., … Yamazaki, T. (2009). Evaluation of forest snow processes models (SnowMIP2). Journal of Geophysical Research, 114(D6), D06111. https://doi.org/10.1029/2008JD011063

Shuttleworth, W. J. (2012). Terrestrial Hydrometeorology: Shuttleworth/Terrestrial Hydrometeorology. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119951933

Stull, R. (2011). Wet-Bulb Temperature from Relative Humidity and Air Temperature. Journal of Applied Meteorology and Climatology, 50(11), 2267–2269. https://doi.org/10.1175/JAMC-D-11-0143.1

Sun, N., Yan, H., Wigmosta, M. S., Leung, L. R., Skaggs, R., & Hou, Z. (2019). Regional Snow Parameters Estimation for Large-Domain Hydrological Applications in the Western United States. Journal of Geophysical Research: Atmospheres, 124(10), 5296–5313. https://doi.org/10.1029/2018JD030140

Tarboton, D. G., & Luce, C. H. (1996). Utah Energy Balance Snow Accumulation and Melt Model (UEB). Utah Water Research Laboratory and USDA Forest Service Intermountain Research Station. https://hydrology.usu.edu/dtarb/snow/snowreptext.pdf

Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D., Peterson, G. D., Roskies, R., Scott, J. R., & Wilkins-Diehr, N. (2014). XSEDE: Accelerating Scientific Discovery. Computing in Science & Engineering, 16(5), 62–74. https://doi.org/10.1109/MCSE.2014.80

U.S. Army Corps of Engineers. (1956). Snow Hydrology, Summary report of the Snow Investigations. U.S. Army Corps of Engineers. https://usace.contentdm.oclc.org/digital/collection/p266001coll1/id/4172/

Viterbo, F., Mahoney, K., Read, L., Salas, F., Bates, B., Elliott, J., Cosgrove, B., Dugger, A., Gochis, D., & Cifelli, R. (2020). A Multiscale, Hydrometeorological Forecast Evaluation of National Water Model Forecasts of the May 2018 Ellicott City, Maryland, Flood. Journal of Hydrometeorology, 21(3), 475–499. https://doi.org/10.1175/JHM-D-19-0125.1

Wang, Y., Broxton, P., Fang, Y., Behrangi, A., Barlage, M., Zeng, X., & Niu, G. (2019). A Wet-Bulb Temperature-Based Rain-Snow Partitioning Scheme Improves Snowpack Prediction Over the Drier Western United States. Geophysical Research Letters, 46(23), 13825–13835. https://doi.org/10.1029/2019GL085722

Wen, L., Nagabhatla, N., Lü, S., & Wang, S.-Y. (2013). Impact of rain snow threshold temperature on snow depth simulation in land surface and regional atmospheric models. Advances in Atmospheric Sciences, 30(5), 1449–1460. https://doi.org/10.1007/s00376-012-2192-7

Wrzesien, M. L., Pavelsky, T. M., Kapnick, S. B., Durand, M. T., & Painter, T. H. (2015). Evaluation of snow cover fraction for regional climate simulations in the Sierra Nevada: EVALUATION OF SNOW COVER FOR REGIONAL SIMULATIONS IN THE SIERRA NEVADA. International Journal of Climatology, 35(9), 2472–2484. https://doi.org/10.1002/joc.4136

Zehe, E., Becker, R., Bárdossy, A., & Plate, E. (2005). Uncertainty of simulated catchment runoff response in the presence of threshold processes: Role of initial soil moisture and precipitation. Journal of Hydrology, 315(1–4), 183–202. https://doi.org/10.1016/j.jhydrol.2005.03.038

Appendices

Algorithm 4.1. Snow rain ratio (SNRR) Calculation. P is the total precipitation and SWE is the snow water equivalent at the start of day. The index t and t+1 indicate the start and the end of the period (day).

If $P_t > 0$:

    // If there is an increase in SWE during the period,
    // compute SNRR
    If $SWE_{t+1} - SWE_t > 0$:
        $SNRR_t = (SWE_{t+1} - SWE_t) / P_t$
    else:
    // If there is a decrease in SWE during the period,
    // SNRR should be 0 due to the rain melting the snow
        $SNRR_t = 0$
else:
    // SNRR cannot be computed because there
    // is no precipitation to separate into rain and snow
    $SNRR_t$ = nan

Algorithm 4.2. Rain snow separation (RSS) scheme based on USACE (1956). T_a is air temperature in degree C and f_s is the fraction of snowfall.

If T_a >= 3:
    f_s = 0
else if T_a <= -1:
    f_s = 1
else:
    f_s = 1 - (T_a - (-1)) / (3 - (-1))

Algorithm 4.3. Rain snow separation (RSS) scheme based on Jordan (1991). T_a is air temperature in degree K, T_f is the freezing point in degree K, and f_s is the fraction of snowfall.

```
// Physical constants and parameters required
Tf = 273.16

If Ta >= Tf + 2.5:
    fs = 0
else:
    fs = 1
    if Ta <= Tf + 0.5:
       fs = 1
    else if Ta <= Tf + 2:
       fs = 1 - (-54.632 + 0.2 Ta)
    else:
       fs = 0.6
```

Algorithm 4.4. Rain snow separation (RSS) scheme based on Marks et al. (1991). e is the vapor pressure of the air in kPa, P_{air} is the air pressure in kPa, q is specific humidity kg/kg, T_d is dew point temperature in degree C, and f_s is the fraction of snowfall.

```
// Compute the vapor pressure of the air from
// Shuttleworth (2012) Equation 2.8
e = (P_air q) / (0.622 + 0.378 q)

// Compute T_d from Shuttleworth (2012) Equation 2.21
T_d = (ln(e) + 0.49299) / (0.0707 - 0.00421 ln(e))

// Discrete version: compute snowfall fraction based on
// T_d from Marks et al. (1999) Table 1.
If T_d < -0.5:
    f_s = 1
else if -0.5 <= T_d < 0:
    f_s = 0.75
else if 0 <= T_d < 0.5:
    f_s = 0.25
else:
    f_s = 0

// Continuous version: compute snowfall fraction using a
// continuous version of Marks et al. (1999) Table 1
If T_d < -0.5:
    f_s = 1
else if -0.5 <= T_d < 0.5:
    f_s = 0.5 - T_d
else:
    f_s = 0
```

Algorithm 4.5. Rain snow separation (RSS) scheme based on Wang et al. (2019). T_f is freezing point in degree K, cp is heat capacity of vaporization in j/kg, L_v is latent heat of vaporization in j/kg, NITER is number of iterations to iteratively solve the T_w equation, T_a is air temperature in degree K, P_{air} is air pressure in Pa, q is specific humidity in kg/kg, gamma is the psychrometric constant in Pa, e is the vapor pressure of the air in Pa, es_a is the saturated vapor pressure at T_a in Pa, RH is relative humidity, T_w is wet-bulb temperature in degree C, es_w is the saturated vapor pressure at T_w in Pa, and fs is the fraction of snowfall. Note that constant values are the same as used in the NWM Noah-MP code.

```
// Physical constants and parameters required
Tf = 273.16
cp = 1004.64
Lv = 2.5104E06
NITER = 20

Tc = Ta - Tf   // Kelvin to Celsius
gamma = (cp Pair) / (0.622 Lv)
e = (Pair q) / (0.622 + 0.378 q)
esa = 610.8 exp ((17.27 Tc) / (237.3 + Tc))
RH = e/es

if RH > 100:
      Tw = Tc
      esw = 610.8 exp ((17.27 Tw) / (237.3 + Tw))
else:
      Tw = Tc - 5                    // First guess for Tw to start the iterative method
      for i in range (1, NITER):     // Use Newton-Raphson method:
         esw = 610.8 exp ((17.27 Tw) / (237.3 + Tw))
         F = Tw - Tc + (1 / gamma) (esw – e)
         Fprim = 1 + (1 / gamma) (esw) [17.27 / (237.3 + Tw) - (17.27 Tw) / (237.3 + Tw)
            **2]
         Tw = Tw - F / Fprim          // Update Tw

         // Check the stopping criteria
         if ABS (F / Fprim) <= 0.01:
            break
      Tw = max ( -50, Tw)

// Compute fs using Wang et al. (2019) approach
A = 6.99*10**(-5)
B = 2
C = 3.97
fs = 1 / (1 +A exp (B (Tw + C)))
```

Table 4.1. Common statistical metrics used in this study to compare model inputs and outputs versus observations[†].

Name	Equation	Range	Description
Coefficient of determination (r^2)	$r^2 = \left(\frac{\sum_{t=1}^{N}(O_t - \overline{O_t})(M_t - \overline{M_t})}{\sqrt{\sum_{t=1}^{N}(O_t - \overline{O_t})^2 \sum_{t=1}^{N}(M_t - \overline{M_t})^2}}\right)^2$	-1 to 1 with 1 indicating a perfect positive linear relationship	Measures the linear relationship. Insensitive to proportional differences between modeled and observed data.
Spearman's rank correlation (Spearmanr)	$Spearmanr = 1 - \frac{6\sum_{t=1}^{N} d_t^2}{N(N^2-1)}$	-1 to 1 with 1 indicating a perfect positive correlation	Measures the strength of association between modeled and observed values.
Root mean squared error (RMSE)	$RMSE = \sqrt{\frac{\sum_{t=1}^{N}(O_t - M_t)^2}{N}}$	Depends on the variable with the best value of 0.	Measures how concentrated the data are around the line of best fit.
Nash Sutcliffe efficiency (NSE)	$\mathrm{NSE} = 1 - \frac{\sum_{t=1}^{N}(O_t - M_t)^2}{\sum_{t=1}^{N}(O_t - \overline{O_t})^2}$	-infinity to 1 with 1 indicating observed and modeled data fits the 1:1 line	Determines the relative magnitude of the residual variance compared to observed values.
Bias	$\mathrm{Bias} = \frac{\sum_{t=1}^{N}(M_t - O_t)}{N}$	Depends on the variable with the best value of 0.	Quantifies the average of the differences between modeled and observed values.

[†]M_t is model simulation, O_t is observation, t is time, N is the total number of simulations or observations, d_t is difference between observed and modeled rank, and the overbar indicates average.

Table 4.2. Observed and modeled half melt dates comparison. Model half melt date is considered as early if it occurs one or more days before observations.

Scenarios that had observed precipitation and bias-corrected air temperature)	RSS scheme	Percentage of days with modeled half melt date earlier than observation across all sites and years
Scenario 3	Jordan (1991) T_a† based	67
Scenario 4	USACE (1956) T_a† based	72
Scenario 5	Marks et al. (1999) T_d + based	62
Scenario 6	Wang et al. (2019) T_w* based	65

†Air temperature

+Dew point temperature

*Wet-bulb temperature

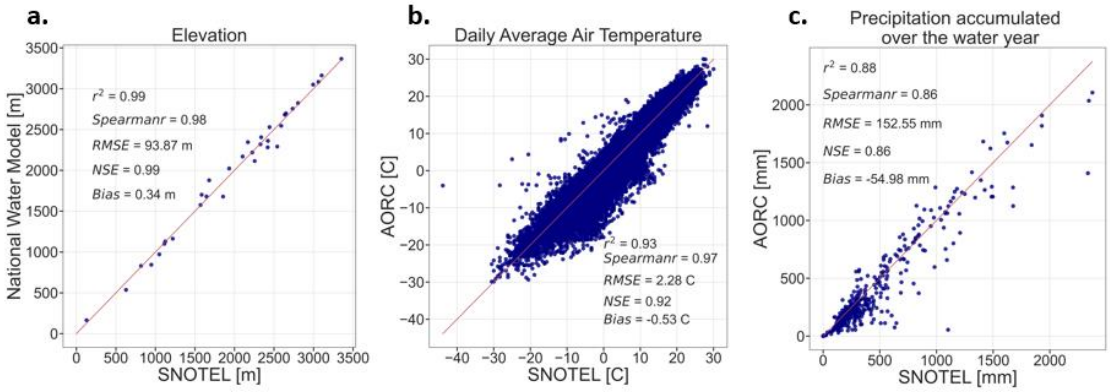

Figure 4.1 (a) NWM elevation inputs compared to SNOTEL site elevations (each point is a SNOTEL site), (b) AORC mean daily temperature compared to mean measurements at SNOTEL sites (each point is a day for a SNOTEL site during the 2009-2019 water years) excluding incorrect AORC air temperatures (see Figure 4.2), and (c) AORC annual precipitation compared to observations at SNOTEL sites (each point represents total precipitation during a water year at a SNOTEL site). Statistical metrics on graphs are coefficient of determination (r^2), Spearman's rank correlation (Spearmanr), root mean square error (RMSE), Nash Sutcliffe efficiency (NSE), and bias (Bias) for which equations are provided in Table 4.1.

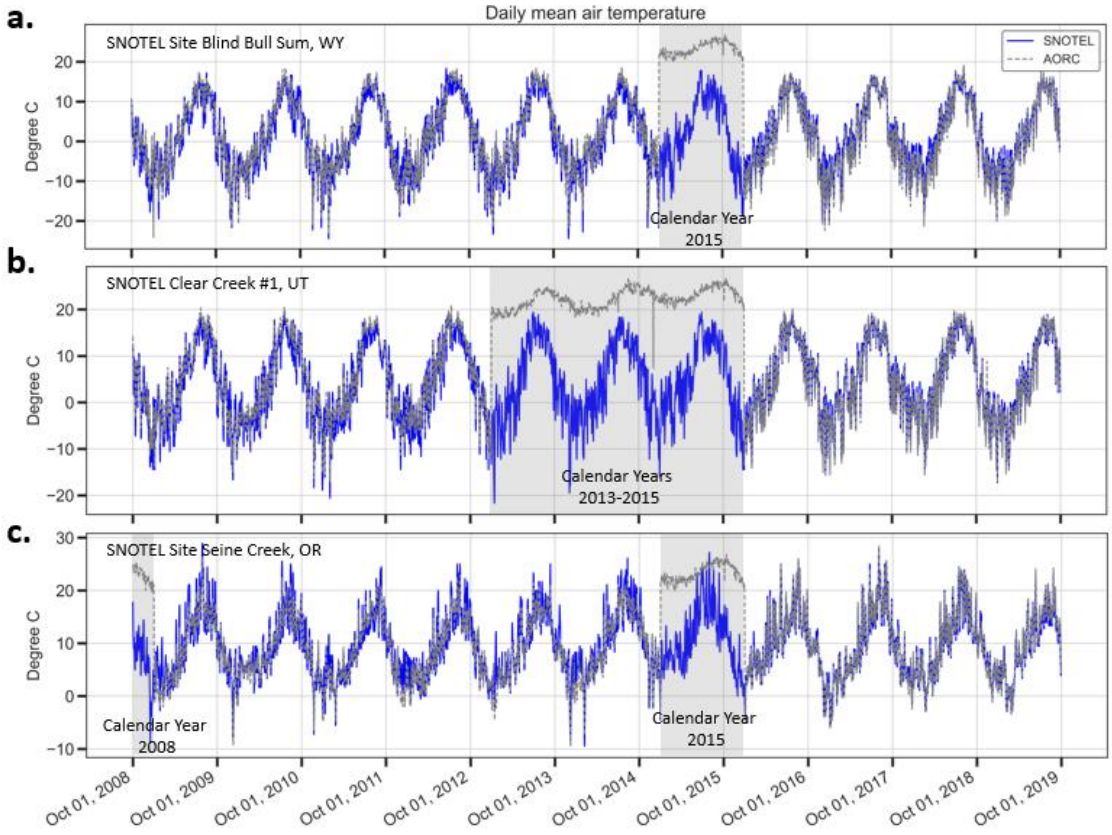

Figure 4.2 AORC and SNOTEL daily mean air temperature during 2009-2019 water years at (a) Blind Bull Sum SNOTEL site in Wyoming, (b) Clear Creek #1 SNOTEL site in Utah, and (c) Seine Creek SNOTEL site in Oregon with gray regions showing periods that AORC air temperature appear to be obviously incorrect. We considered these as artifacts and excluded these periods from our analysis.

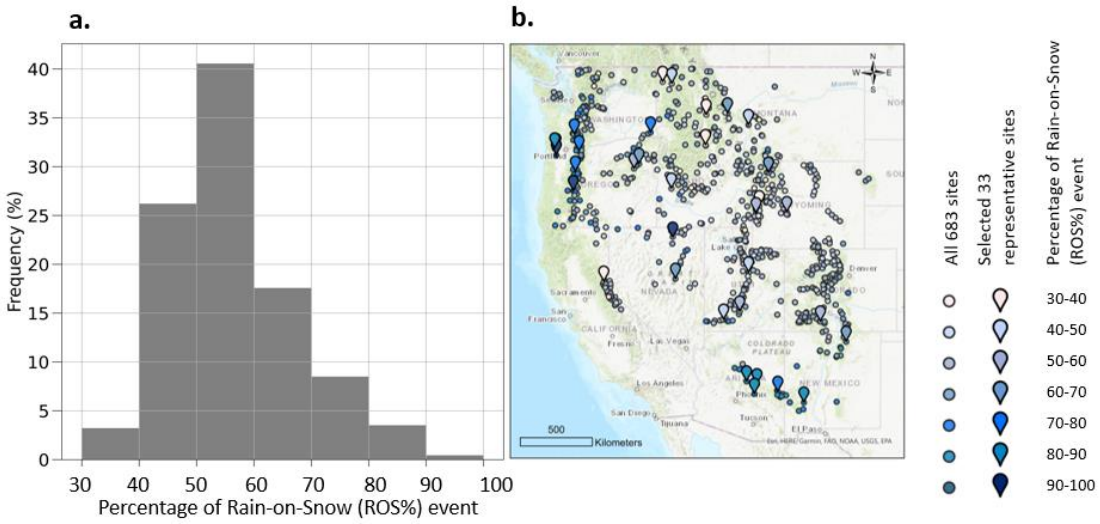

Figure 4.3 (a) Histogram of the percentage of historical Rain-on-Snow (ROS%) events inferred from the computed SNRR over SNOTEL sites (total of 683 sites) with data for 2008-2020 water years across the western U.S. (b) Location of representative SNOTEL sites selected based on the ROS%.

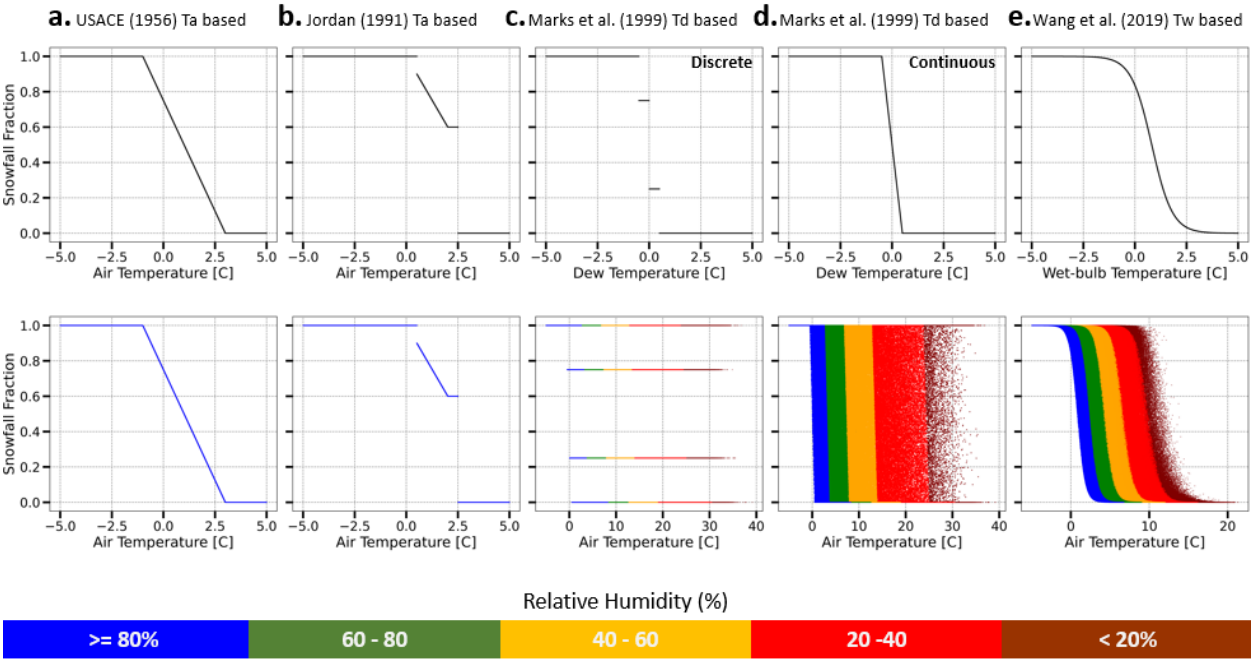

Figure 4.4 Snowfall fraction computed for the 33 SNOTEL sites using the observed precipitation and the NWM inputs (including air pressure, specific humidity, and bias-corrected air temperature) based on (a) USACE (1956), (b) Jordan (1991), (c) Marks et al. (1999): discrete version, (d) Marks et al. (1999): continuous version and (e) Wang et al. (2019) RSS methods. The plots on the top row show the relationship between snowfall fraction as a function of air temperature (T_a), dew point (T_d), or wet-bulb (T_w) temperature depending on the method. The plots on the bottom row illustrate the relationship between snowfall fraction and air temperature for all methods. The colors represent data with different relative humidity values.

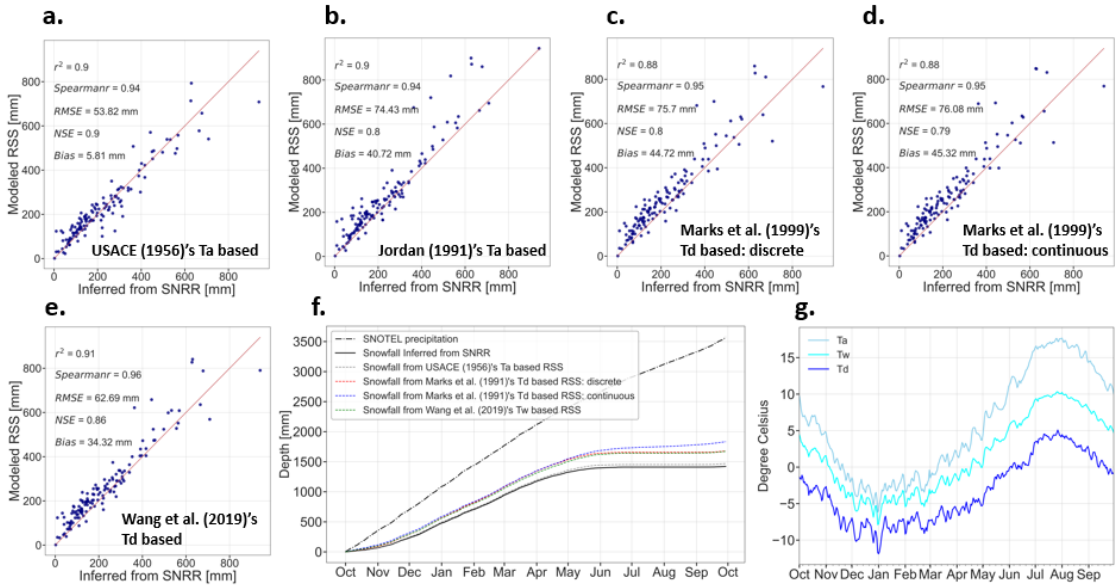

Figure 4.5 Analysis of annual snowfall estimated from different RSS schemes versus observations inferred from SNRR at SNOTEL sites for a period of 11 years (water years 2009-2019). (a) USACE (1956) air temperature-based RSS method versus SNRR, (b) Jordan (1991) air temperature-based RSS method (the current approach in the NWM version 2.0) versus SNRR, (c) Marks et al. (1999) dew point based (discrete version) RSS method versus SNRR, (d) Marks et al. (1999) dew point based (continuous version) RSS method versus SNRR, and (e) Wang et al. (2019) wet-bulb based RSS method versus SNRR. Each point in panels (a)-(e) represents a water year and a SNOTEL site. (f) The seasonal pattern of the long-term annual observed precipitation, observed snowfall inferred from SNRR, and modeled snowfall from all RSS schemes averaged across all sites and years. (g) Seasonal pattern of the long-term daily bias-corrected AORC air temperature (T_a) and computed wet-bulb (T_w) and dew point (T_d) temperatures using AORC data averaged across all sites and years.

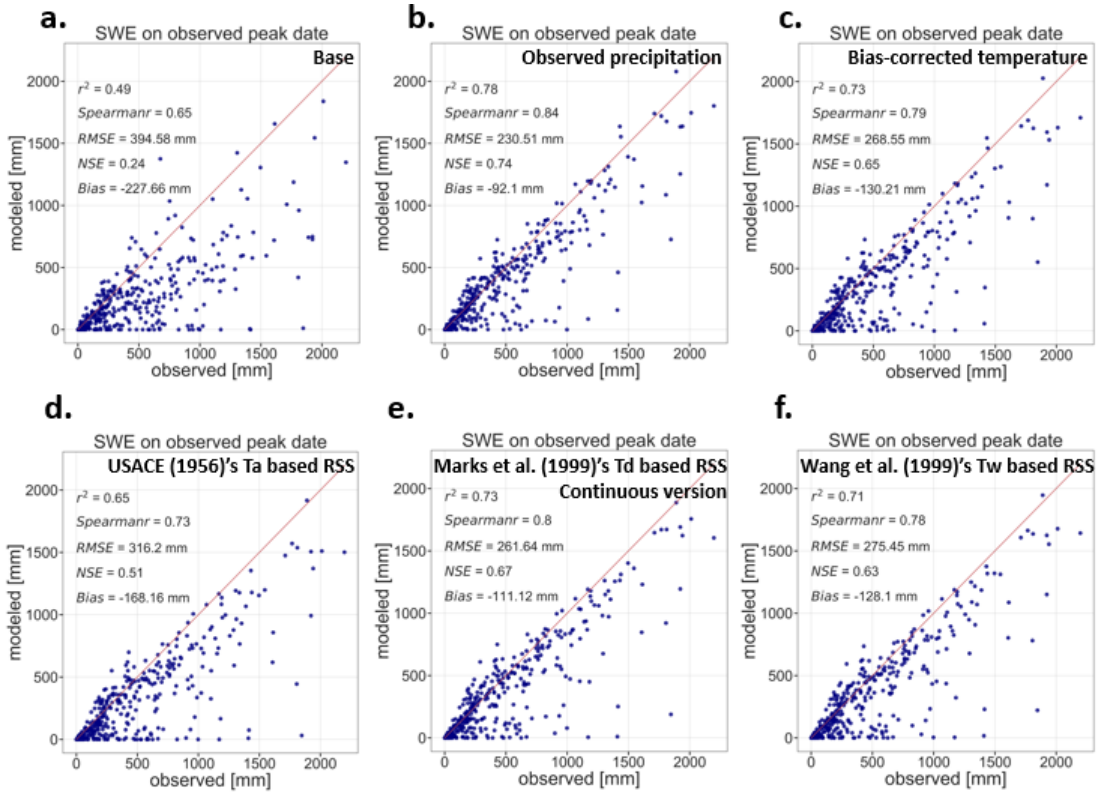

Figure 4.6 SWE Comparison on date of observed peak SWE. (a) NWM base scenario (Scenario 1) versus SNOTEL SWE, (b) NWM observed precipitation scenario (Scenario 2) versus SNOTEL SWE, (c) NWM bias-corrected temperature scenario (Scenario 3) versus SNOTEL SWE, (d) NWM using USACE (1956) air temperature (T_a) based RSS method (Scenario 4) versus SNOTEL SWE, (e) NWM using Marks et al. (1999) dew point (T_d) based (continuous version) RSS method (Scenario 5) versus SNOTEL SWE, (f) NWM using Wang et al. (2019) wet-bulb (T_w) based RSS method (Scenario 6) versus SNOTEL SWE. Each point on the graph represents a SNOTEL site and a water year.

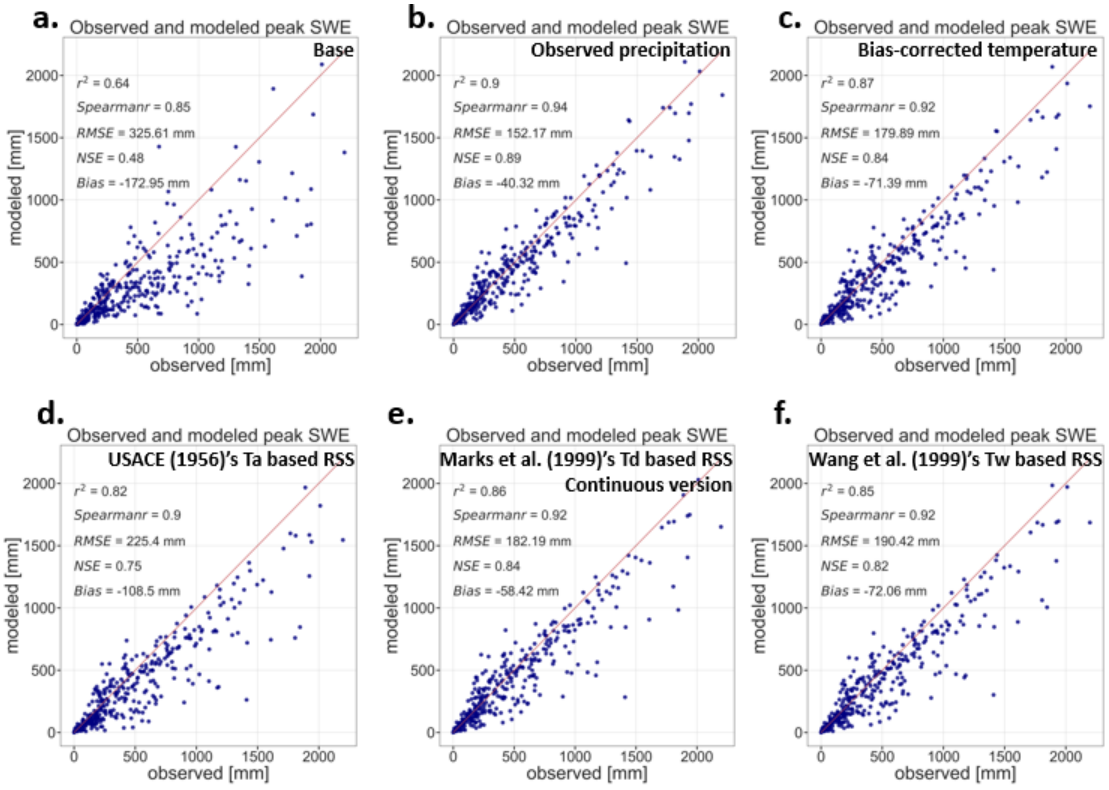

Figure 4.7 Observed and modeled peak SWE comparison (on the generally different dates they occur). (a) NWM base scenario (Scenario 1) versus SNOTEL SWE, (b) NWM observed precipitation scenario (Scenario 2) versus SNOTEL SWE, (c) NWM bias-corrected temperature scenario (Scenario 3) versus SNOTEL SWE, (d) NWM using USACE (1956) air temperature (T_a) based RSS method (Scenario 4) versus SNOTEL SWE, (e) NWM using Marks et al. (1999) dew point (T_d) based (continuous version) RSS method (Scenario 5) versus SNOTEL SWE, and (f) NWM using Wang et al. (2019) wet-bulb (Tw) based RSS method (Scenario 6) versus SNOTEL SWE. Each point on the graphs represents a SNOTEL site and a water year.

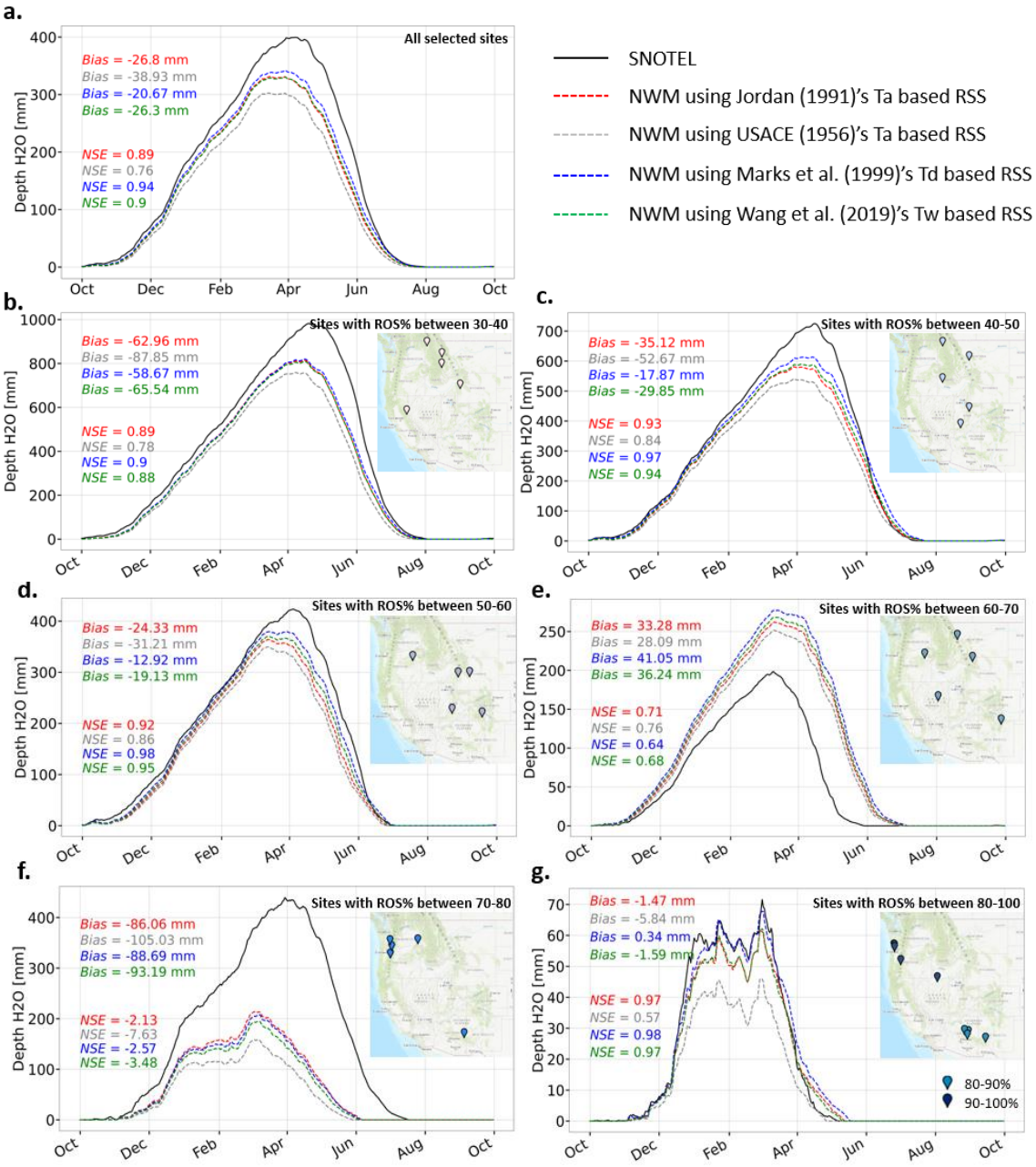

Figure 4.8 Observed and modeled SWE at the beginning of each date averaged across all years and (a) all selected SNOTEL sites, (b) sites with ROS% between 30-40%, (c) sites with ROS% within 40-50%, (d) sites with ROS% within 50-60%, (e) sites with ROS% within 60-70%, (f) sites with ROS% within 70-80%, and (g) sites with ROS% within 80-100%.

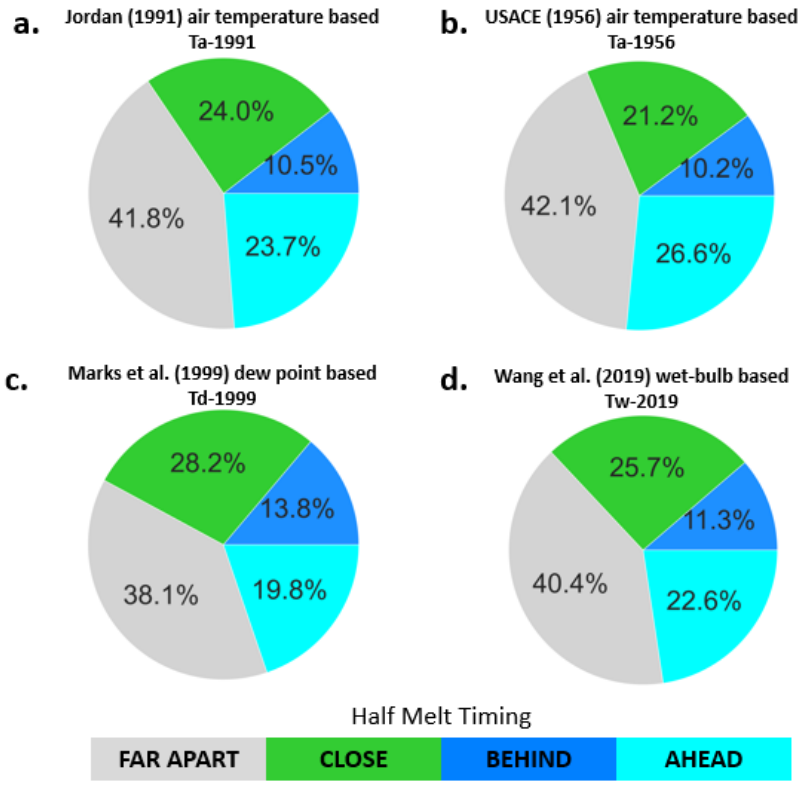

Figure 4.9 Analysis of melt timing based on classification of differences between observed and modeled dates of half melt from peak SWE. (a) NWM bias-corrected temperature scenario versus SNOTEL half melt dates, (b) NWM using USACE (1956) T_a based RSS method versus SNOTEL half melt dates, (c) NWM using Marks et al. (1999) T_d based RSS method versus SNOTEL half melt dates, and (d) NWM using Wang et al. (2019) T_w based RSS method versus SNOTEL half melt dates. In this figure, FAR APART: modeled and observed half melt dates are more than 20 days apart; CLOSE: modeled and observed half melt dates are within 5 days of each other; BEHIND: modeled half melt dates are 6 to 19 days after observed; and AHEAD: modeled half melt dates are 6 to 19 days before observed.

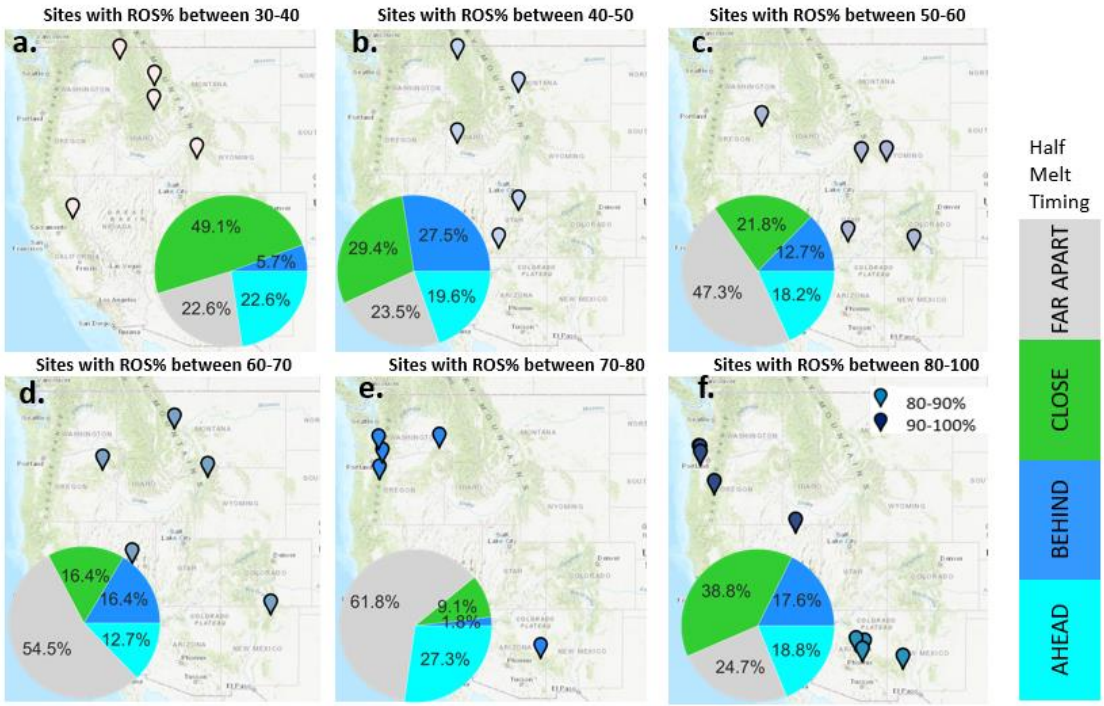

Figure 4.10 Analysis of melt timing from NWM using T_d based RSS scheme (the approach with the least bias and best NSE and RMSE in SWE comparisons) across different ROS% classes. (a) ROS% between 30 to 40%, (b) ROS% between 40 to 50%, (c) ROS% between 50 to 60%, (d) ROS% between 60 to 70%, (e) ROS% between 70 to 80%, and (f) ROS% between 80 to 100%.

CHAPTER 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The research presented in this dissertation evaluated opportunities to address the growing need within the hydrology community to advance state-of-the-art hydrological forecasting capabilities and improve process-based understanding within continental-scale hydrological modeling of flood and water supply forecasts under changing conditions. The enhancements to the Continental-scale Flood Inundation Mapping (CFIM) method and improvements evaluated for the National Oceanic and Atmospheric Administration's (NOAA) National Water Model (NWM) snow modeling demonstrated through the research presented in this dissertation are the main contributions of this work. The significance of the work presented in this dissertation is that it identifies some of limitations associated with CFIM and the NWM snow parameterization, and then it provides advances to the Height above Nearest Drainage (HAND) method in CFIM and parameterization of precipitation separation into rain and snow in NWM to improve modeled flood inundation extent from CFIM and snow outputs from NWM. All of these are necessary for improving flood and water supply forecasts that supports water prediction in the U.S.

Summary:

The work in Chapter 2 evaluated the CFIM method over two case studies, a section of the Bear River in Utah and Ocheyedan River in Iowa, by comparing modeled flood inundation with the flood inundation observed in high-resolution Planet RapidEye satellite imagery. Results showed discrepancies between modeled and observed inundation maps. Examination of these discrepancies identified several limitations in the

HAND terrain analysis methodology used in CFIM. First, streams to which HAND is calculated and the catchments to which hydraulic properties are evaluated are not derived from the Digital Elevation Model (DEM) data in CFIM and thus they do not match. This misalignment resulted in spurious flooded areas. Second, CFIM is limited by the occurrence of flat areas upstream of artificial barriers in DEM such as roads, resulting in over-estimated flooded areas when compared to observations. Third, the fixed hydraulic roughness parameter in HAND calculation was identified as a limitation. A set of improvements were developed to overcome these limitations and advance CFIM outcomes.

First, using the streams and catchments that were purely from the DEM improved the calculation of HAND and mapping of flood inundation. Following on from this, to resolve the problems associated with DEM barrier artefacts we developed a new flow direction conditioning DEM terrain analysis method. We also showed that using higher resolution DEM data (1/9 arc-sec (3 m) resolution) further improved the precision of mapping the flooded areas. Lastly, the approach to use observed inundation from past floods combined with observed discharge to estimate reach specific Manning's n, which is consistent with other assumptions in the HAND approach, provided an opportunity to overcome the limitations and bias associated with a single roughness parameter that was evident in the CFIM results.

In Chapter 3, the comparison of the NWM version 2.0 reanalysis snow outputs to observed Snow Water Equivalent (SWE) and Snow-Covered Area Fraction (SCAF) at SNOTEL sites across the Western U.S., in general, demonstrate that the NWM under-estimates SWE early in the season and becomes progressively more biased later in the

season. Results showed that SWE discrepancies were attributed to both model errors and errors in precipitation and air temperature inputs. The downward bias in precipitation input contributed to the downward biases in SWE. Even when the model precipitation was relatively close to the observed precipitation at SNOTEL sites, the under-modeling of SWE was persistent, suggesting deficiencies in model physics associated with snowpack development during accumulation and ablation phases. For example, the model may under-estimate snowfall during the accumulation phase or incorrectly estimate snow temperatures causing early melt during ablation phase. There was also a slight general bias for model input air temperature to be cooler than observed. In the SCAF comparison, we observed a significant variability between the MODIS SCAF and NWM SCAF which hampered useful interpretation of these comparisons. However, when SCAF was aggregated across all sites and years, modeled SCAF tended to be more than observed using MODIS.

Findings from Chapter 3 identified areas where snow predictions from the NWM may be better or worse and suggested opportunities for research directed towards model improvements, which motivated the research in Chapter 4. In this work, an offline implementation of the WRF-Hydro model with the National Water Model configuration was used to evaluate: (1) input data and (2) rain snow separation improvements to NWM SWE estimates across a set of representative point locations from SNOTEL sites over 11 water years. Even though meteorological inputs of NWM version 2.0 (NLDAS2 as described in Chapter 2) and NWM version 2.1 (AORC in Chapter 4) differed with AORC being less biased (Naple, 2021), results showed that the NWM still generally under-estimates SWE and tends to melt snow earlier than observations. Using observed

precipitation and bias-corrected air temperature inputs reduced a significant portion of the observed and modeled SWE discrepancies. NWM SWE outputs were further improved by using rain-snow separation parameterizations that incorporate humidity information. Both dew point and wet-bulb parameterizations were evaluated, with the dew-point parameterization giving the better overall improvement. This modification also improved the melt timing because the dew point temperature-based scheme had more snowfall and sometimes extended to warmer months such as May, compared to the current air temperature-based separation scheme in the NWM.

Conclusions:

The work described in this dissertation contributes to hydrological modeling research in several aspects. First, terrain analysis enhancements to the HAND flood inundation mapping method presented in Chapter 2 improved CFIM flood inundation mapping outputs. Improvements due to the streams and catchments that are purely derived from DEMs and are consistent with HAND suggests that it is important to ensure consistency between streams and catchments in the terrain analysis processing for enriching the content of DEM data for use in hydrologic modeling. In addition, using the flow direction conditioning approach produced an important improvement in the calculation of HAND and mapping of flood inundation, and thus it is recommended to be used where there are artificial barriers that result in flat areas. The flow direction conditioning method required only two inputs: a DEM and a high-resolution hydrography dataset. In some cases, the lack of a high-resolution hydrography dataset may hinder the applicability of the flow direction conditioning approach. However, high-resolution hydrography is available from the National Hydrography Dataset for the entire

continental U.S., and the flow direction conditioning could be applied countrywide. Also, improvements due to using a higher resolution DEM dataset provides input to consider when evaluating the merit and additional expense of 3 m data collection in hydrological modeling. This has also been reported by Munoth and Goyal (Munoth & Goyal, 2019).

Second, evaluation of the NWM retrospective inputs and snow outputs identified some of the current challenges and limitations in the model and suggested opportunities for research directed towards the NWM model snow component improvements. Examination of these inputs/outputs and their patterns across a continent enhances process understanding to improve the NWM snow components, and ultimately runoff and water supply forecasts in snowmelt-dominated regions. One key point is that model input accuracy plays a key role in snow model outputs. This has also been reported by others (Förster et al., 2014; Raleigh et al., 2015; Zehe et al., 2005). While stating that better inputs lead to better model performance is not new, this emphasizes that more attention needs to be given to model inputs in improving the NWM overall. Another key point is that the NWM SWE magnitude and snowmelt timing can be improved by using rain/snow separation parameterization schemes, which consider the impact of atmospheric humidity in the energy budget of falling hydrometeors.

Recommendations:

Results from this research open some new questions for future research. First, the improvements observed from using the approach to estimate reach specific Manning's n from observed flood inundation emphasizes the importance of having DEMs that represent the channel bed topography. Moreover, there is a question as to whether, in

locations where the channel bed topography is missing, the hydraulic geometry parameters can be estimated from the inferred stage, discharge, and roughness parameter.

Results from comparison of the NWM SCAF and MODIS observations open the question as to whether, in circumstances where there is disagreement between modeled and observed SCAF in a binary comparison, the SCAF parameterization in Noah-MP that lacks representation of vegetation and topography should be improved or can be inferred from satellite data while considering the uncertainties associated with these products. Using satellite-based snow-cover maps may potentially provide an approach or an opportunity for estimating SCAF as a way to overcome limitations associated with parameterization of SCAF in the snow model.

Results from using the dew-point parameterization to separate precipitation into rain and snow highlight several points for future research. Despite improvements to snowfall and SWE, several uncertainties nevertheless remain. First, it is suggested that model performance be evaluated against other observation datasets, for instance the observations of snowfall and rainfall measured at the UC Berkeley Central Sierra Snow Laboratory site. Also, it would be helpful for future work regarding the NWM evaluation and improvements to have a more comprehensive observation data set, beyond what is provided by SNOTEL sites. For example, Critical Zone Observatory (CZO) or experimental forest sites could include snowfall/rainfall measurements, canopy snow interception, turbulence, and radiation fluxes above and below the canopy. Second, the discrepancies in model half melt dates compared to observations suggest that the modeling of melt timing is somewhat problematic. There is a need to further investigate overall energy balance and snow surface temperature, possibly drawing on ideas from the

Utah Energy Balance model (Mahat & Tarboton, 2014; You et al., 2014). Third, despite the general under-modeling of NWM precipitation inputs which contributed to the under-modeling of SWE, the bias in NWM air temperature was counter to the direction expected to lead to under-modeling of SWE and needs further exploration. We hypothesize that the remaining discrepancies between the modeled and observed SWE are likely related to errors in calculating energy balance components that affect snowpack average temperature and available energy to melt along with other snow process parameterizations such as snow cover fraction parameterization. This has been reported by Wrzesien et al. (2015) when evaluating the Noah-MP snow processes parameterizations with satellite observations, but not necessarily based on the specific configuration used in the NWM. Thus, these could be considered as future directions to improve the NWM performance in mountainous regions. Lastly, while U.S. based, the NWM was built using the WRF-Hydro modeling framework that has been applied worldwide, and the lessons learned from this work have application to the representation of snow processes in continental scale models throughout the world.

Förster, K., Meon, G., Marke, T., & Strasser, U. (2014). Effect of meteorological forcing and snow model complexity on hydrological simulations in the Sieber catchment (Harz Mountains, Germany). Hydrology and Earth System Sciences, 18(11), 4703–4720. https://doi.org/10.5194/hess-18-4703-2014

Munoth, P., & Goyal, R. (2019). Effects of DEM Source, Spatial Resolution and Drainage Area Threshold Values on Hydrological Modeling. Water Resources Management, 33(9), 3303–3319. https://doi.org/10.1007/s11269-019-02303-x

Naple, P. W. (2021). Evaluating the Performance of National Water Model Snow Simulations in the Northeastern United States Using Advanced Mesonet Observations [State University of New York at Albany]. https://www.proquest.com/openview/8c5f9dbfba7aa68d7998f5f727e3a658/1?pq-origsite=gscholar&cbl=18750&diss=y

Raleigh, M. S., Lundquist, J. D., & Clark, M. P. (2015). Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. Hydrology and Earth System Sciences, 19(7), 3153–3179. https://doi.org/10.5194/hess-19-3153-2015

Wrzesien, M. L., Pavelsky, T. M., Kapnick, S. B., Durand, M. T., & Painter, T. H. (2015). Evaluation of snow cover fraction for regional climate simulations in the Sierra Nevada: EVALUATION OF SNOW COVER FOR REGIONAL SIMULATIONS IN THE SIERRA NEVADA. International Journal of Climatology, 35(9), 2472–2484. https://doi.org/10.1002/joc.4136

Zehe, E., Becker, R., Bárdossy, A., & Plate, E. (2005). Uncertainty of simulated catchment runoff response in the presence of threshold processes: Role of initial soil moisture and precipitation. Journal of Hydrology, 315(1–4), 183–202. https://doi.org/10.1016/j.jhydrol.2005.03.038

APPENDICES

Utah Water Research Laboratory
UtahStateUniversity

January 4, 2022

Mahyar Aboutalebi, Ph.D.
Research Scientist I – Remote Sensing
E. & J. Gallo Winery
Email: mahyar.aboutalebi@gmail.com

Dear Mahyar:

I am in the process of preparing my dissertation in the Civil and Environmental Engineering Department at Utah State University. I hope to complete my degree in December 2021.

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

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

Thank you,

Irene Garousi-Nejad

I hereby give permission to Irene Garousi-Nejad to use and reprint all the material that I have contributed to Chapter 2 of this dissertation.

_____

Mahyar Aboutalebi

CURRICULUM VITAE

Irene Garousi-Nejad
Graduate Research Assistant
Department of Civil and Environmental Engineering
Utah Water Research Laboratory
Utah State University
1600 Canyon Rd, Logan, UT 84321
Email irene.garousi@usu.edu

Research Interests:

- Surface hydrology
- Snow hydrology
- Geospatial and terrain analysis
- Flood modelling
- Computational physical hydrology modelling
- Remote sensing applications to hydrology
- Land surface and atmosphere interactions
- Water resources planning and management

Academic Education:

2016 – (Dec 2021) Utah State University, Utah. U.S.
PhD Civil and Environmental Engineering – Hydrology and Water Resources
Dissertation: Improving Flood Inundation and Streamflow Forecasts in Snowmelt Dominated Regions
Committee: David G. Tarboton (advisor), Jeffery S. Horsburgh, Alfonso Torres-Rua, Simon Wang, Robert Gillies

2015 University of Tehran, Tehran, Iran
MS Agricultural Engineering, Water Resources Engineering
Thesis: Developing of multi-objective firefly algorithm in optimal operation of the reservoirs systems
Advisor: Omid Bozorg-Haddad; Co-advisor: Abdolhosein Hourfar

2011 University of Tehran, Tehran, Iran
BS Agricultural Engineering, Water Resources Engineering

Honours and Awards:

2021 Awarded in the first-round of the 16th Annual J. Paul Riley AWRA-Utah Section Student Conference and Paper Competition, Virtual. [$250]

2021 Selected as the Doctoral Student Researcher of the Year for the College of Engineering at Utah State University [$500]

2020 The Water Resources Research paper (Garousi-Nejad et al., 2019, https://doi.org/10.1029/2019WR024837) received a certificate for being among the top 10% most downloaded papers in recent publication history.

2020 Awarded Eva Nieminski Honorary Graduate Scholarship provided by American Water Work Association Intermountain Section, Midway, Utah. [$1,000]

2018 Awarded in the first-round of the 13th Annual J. Paul Riley AWRA-Utah Section Student Conference and Paper Competition, Salt Lake City, Utah. [$250]
2017 Selected and awarded travel funding by CUAHSI to attend the National Water Center Innovators Program: Summer Institute June-July 2017 held at the University of Alabama, AL. [$4,000]
2017 Outstanding student poster in Spring Runoff Conference 2017, Utah State University. [$100]
2016 Doctoral Dissertation Fellowship, Utah Water Research Laboratory, Utah State University, Logan, UT.
2015 Ranked the top student in the Department of Irrigation and Reclamation Engineering, University of Tehran, Iran.
2012 Awarded scholarship of Master Program, the University of Tehran, Iran.

Research Visits and Professional Training:

2020 Participated in the Process-Based Hydrological Modelling, a graduate-level course offered through the University of Saskatchewan Center for Hydrology, Canada, taught by Prof. Martyn Clark.
2019 Participated in CUAHSI's Master Class: Advanced Techniques in Watershed Science for one week at Biosphere 2, Arizona.
2018 Selected and awarded travel grant, generously provided by NASA through ATA Aerospace to attend CUAHSI's Snow Measurement Field School for one week at the Fraser Experimental Forest in Fraser, Colorado.
2017 Selected to participate in the National Water Center Innovators Program: Summer Institute for seven weeks supported by CUAHSI held at the University of Alabama (Resulted in Chang et al. (2017) and Garousi-Nejad et al. (2017) technical reports).
2017 Selected to participate in University Consortium for Geographic Information Science (UCGIS) for two weeks supported by National Science Foundation held at the University of Illinois at Urbana-Champaign. (Resulted in Garousi-Nejad et al. 2019, Water Resources Research).
2017 Selected to participate in WRF-Hydro workshop held at the National Center for Atmospheric Research (NCAR), Boulder, CO.

Publications:

Google Scholar = 167 citations, h-index = 6

Garousi-Nejad, I., and Tarboton, D. G. (2021). "A Comparison of National Water Model Retrospective Analysis Snow Outputs at SNOTEL Sites Across the Western U.S.", Authorea. https://doi.org/10.22541/au.161656955.51617798/v2

Lane, B., I. Garousi-Nejad, M. Gallagher, D. Tarboton and E. Habib. (2021). " An open web-based module developed to advance data-driven hydrologic process learning", Hydrological Processes, 35(7). https://doi.org/10.1002/hyp.14273

Garousi-Nejad, I., Tarboton, D. G., Aboutalebi, M., and Torres-Rua, A. F. (2019). "Terrain analysis enhancements to the height above nearest drainage flood inundation mapping method", Water Resources Research. https://doi.org/10.1029/2019WR024837

Garousi-Nejad, I., Bozorg-Haddad, O., & Loáiciga, H. A. (2018). Closure to "Application of the firefly algorithm to optimal operation of reservoirs with the purpose of irrigation supply and hydropower Production", Journal of

Irrigation and Drainage Engineering, 144(1), https://doi.org/10.1061/(ASCE)IR.1943-4774.0001257

Garousi-Nejad, I., He, S., and Tang, Q. (2017). "Chapter 2: Comparison of coarse and high-resolution hydrologic modelling in mountainous area" in CUAHSI National Water Center Innovators Program Summer Institute Report. https://doi.org/10.4211/technical.20171009

Chang, J., Garousi-Nejad, I., Grimley, L., He, S., Khanam, M., Madsen, T., Tang, Q., Tiernan, E., Tijerina, D., and Turnipseed, C. (2017). "Chapter 1: ADHydro Introduction and Workflow" in CUAHSI National Water Center Innovators Program Summer Institute Report. https://doi.org/10.4211/technical.20171009

Bozorg Haddad, O., Garousi-Nejad, I., and Loaiciga, H. A. (2017). "Extended multi-objective firefly algorithm for hydropower energy generation", Journal of Hydroinformatics, 19(5), 734–751. https://doi.org/10.2166/hydro.2017.114

Garousi-Nejad, I., Bozorg Haddad, O., and Loaiciga, H. A. (2017). Closure to "Modified Firefly Algorithm for Solving Multireservoir Operation in Continuous and Discrete Domains", Journal of Water Resources Planning and Management, 143(10). https://doi.org/10.1061/(ASCE)WR.1943-5452.0000837

Garousi-Nejad, I., Bozorg Haddad, O., Loaiciga, H. A., and Marino, M. A. (2016). "Application of the firefly algorithm to the optimal operation of reservoirs with the purposes of irrigation supply and hydropower production", Journal of Irrigation and Drainage Engineering, 142(10). https://doi.org/10.1061/(ASCE)IR.1943-4774.0001064

Garousi-Nejad, I., Bozorg-Haddad, O., and Loaiciga, H. A. (2016). "Modified firefly algorithm for solving multi-reservoir operation in continuous and discrete domains", Journal of Water Resources Planning and Management, 142(9). https://doi.org/10.1061/(ASCE)WR.1943-5452.0000644

Garousi-Nejad, I., Bozorg-Haddad, O., and Aboutalebi, M. (2015). Discussion of "Investigating parameters of two-point hedging policy for operating a storage reservoir", ISH Journal of Hydraulic Engineering, 21(3), 312-314. https://doi.org/10.1080/09715010.2015.1021282

Garousi-Nejad, I. and Bozorg-Haddad, O. (2015). "The implementation of developed firefly algorithm in multireservoir optimization in continuous domain", International Journal of Civil and Structural Engineering (IJCSE), 2(1), 104-108.

Aboutalebi, M. and Garousi-Nejad, I. (2015). Discussion of "Application of the Water Cycle Algorithm to the Optimal Operation of Reservoir Systems", Journal of Irrigation and Drainage Engineering, 141(12). https://doi.org/10.1061/(ASCE)IR.1943-4774.0000947

Bozorg-Haddad, O., Aboutalebi, M., and Garousi-Nejad, I. (2014). Discussion of "Hydroclimatic stream flow prediction using least square-support vector regression", ISH Journal of Hydraulic Engineering, 20(3), 312-314. https://doi.org/10.1080/09715010.2014.881082

Invited Presentation:

Garousi-Nejad, I. (2019). "Examples of CUAHSI Services", Presented at American Geophysical Union (AGU) Fall Meeting, CUAHSI Town Hall, San Francisco, CA, 9-13 December.

Presentations:

Garousi-Nejad, I. and Tarboton, D. G. (2021). "National Water Model Reanalysis Snow Outputs versus Ground Observations", American Geophysical Union (AGU) Fall Meeting, Virtual, 13-17 December.

Garousi-Nejad, I. (2021). "An Assessment of National Water Model Retrospective Analysis Snow Outputs Across the Western U.S.", 16th Annual J. Paul Riley AWRA-Utah Section Student Conference and Paper Competition, Virtual.

Garousi-Nejad, I. and Tarboton, D. G. (2020). "An assessment of the Strengths and Limitations of the National Water Model Snow Representation against In-Situ Measurements, Remote Sensing Products, and Assimilated Data", American Geophysical Union (AGU) Fall Meeting, Virtual, 1-17 December.

Garousi-Nejad, I. and Tarboton, D. G. (2019). "Assessment and calibration of the WRF-Hydro National Water Model configuration for a snowmelt dominated watershed in mountain Karst region", American Geophysical Union (AGU) Fall Meeting, San Francisco, CA, 9-13 December.

Garousi-Nejad, I. and Tarboton, D. G. (2019). "Application and evaluation of WRF-Hydro National Water Model configuration to a headwater snowmelt dominated watershed in mountain Karst region". Presented at CUAHSI 2019 Conference on Hydroinformatics, Brigham Young University, Provo, Utah, 29-31 July.

Garousi-Nejad, I. Tarboton, D. G., Aboutalebi, M., and Torres-Rua, A. F. (2018). "Assessment and Enhancement of National Water Model Height above Nearest Drainage Flood Inundation Mapping Using Planet CubeSat for the 2017 Bear River Flood Event", Development, and Evaluation of the National Water Model and Facilitation of Community Involvement II, American Geophysical Union (AGU) Fall Meeting, Washington, D.C., 10-14 December.

Garousi-Nejad, I. (2018). "Evaluating the Height Above Nearest Drainage Flood Inundation Mapping Approach with High-Resolution Planet CubeSat Imagery", 13th Annual J. Paul Riley AWRA-Utah Section Student Conference and Paper Competition, Salt Lake City, Utah

Garousi-Nejad, I. He, S., Tang, Q., Ogden, F. L., Steinke, R. C., Frazier, N., Tarboton, D. G., Ohara, N., and Lin, H. (2017). "A Study on the Effects of Spatial Scale on Snow Process in Hyper-Resolution Hydrological Modelling over Mountainous Areas", Hyper-Resolution Hydrologic Modelling: Progress and Challenges, American Geophysical Union (AGU) Fall Meeting, New Orleans, LA, 11-15 December.

Garousi-Nejad, I. (2017). "Hyper work, hyper fun, hyper performance computing, hyper resolution hydrologic modelling in the National Center for Supercomputing Application (NCSA) and National Water Center (NWC) in hyper heat of summer 2017", Presented at Utah State University, CEE 6800 Seminar, October 4, 2017, Logan, Utah.

Garousi-Nejad, I., He, S., and Tang, Q. (2017). "Comparison of coarse and high-resolution hydrologic modelling in mountainous area", Presented at CUAHSI 2017 Conference on Hydroinformatics, University of Alabama at Tuscaloosa, 25-27 July.

Garousi-Nejad, I., Li, X., Petrasova, A., and Xu, H. "High-resolution Flood Inundation Mapping Using National Water Model Forecasts for Emergency Management", Presented at UCGIS 2017 Symposium, Hyatt Arlington (Virginia), Washington, D.C., 23-25 May.

Garousi-Nejad, I. and Horsburgh, J. S. (2017). "National Water Model Forecasts and USGS Streamflow Observations Acquisition Using IPython", Presented at Utah State University Spring Runoff Conference, Logan, UT, 28-29 March.

Garousi-Nejad, I. and Bozorg-Haddad, O. (2014). "The implementation of developed firefly algorithm in multireservoir optimization in continuous domain", Second International Conference on Advances in Civil, Structural and Environmental Engineering-ACSEE, Zurich, Switzerland, 25-26, October.

Community Service (professional development and outreach):

Journal reviewer (21 manuscripts based on Publons.com). Water Resources Research, Journal of Hydroinformatics, Hydrological Sciences Journal, Water, Water Science and Technology: Water Supply, Remote Sensing, Urban Climate.

2020 Co-moderator for TH003 - Justice, Equity, Diversity, and Inclusion (JEDI) in the AGU Hydrology Community and Beyond, AGU-H3S Town Hall at the AGU Fall Meeting, Virtual, Dec 1. https://agu.confex.com/agu/fm20/meetingapp.cgi/Session/111432

2020 Primary moderator for TH033 - Navigating a Non-Academic Research Career: Gain Tips and Insights on How to Stand out From the Crowd, AGU-H3S Town Hall at the AGU Fall Meeting, Virtual, Dec 4. https://agu.confex.com/agu/fm20/meetingapp.cgi/Session/102409

2020 Primary moderator for "You do belong: let's talk about impostor syndrome", AGU-H3S CUAHSI co-sponsored webinar, Sep 25. https://www.youtube.com/watch?v=BGRXubyWz-k&ab_channel=CUAHSI

2020 "AGU Hydrology Section Call to Action for a Just, Equitable, Diverse, and Inclusive Scientific Society" white paper, prepared by AGU-H3S JEDI team members, https://drive.google.com/file/d/17b6LezeFA1UJTCsGlWgJNFtFJ9-m5cVU/view

2020 "ECR Hydrology Research Tidbit: The Value of the 40-Year Airborne Gamma Snow Product", a blog post prepared for the AGU-H3S Research Highlights, https://agu-h3s.org/2020/06/15/ecr-hydrology-research-tidbit-the-value-of-the-40-year-airborne-gamma-snow-product/

2020 "Justice, Equity, Diversity, and Inclusion resources to inspire change in the hydrology community", a blog post prepared for the AGU-H3S community resources, https://agu-h3s.org/2020/10/22/justice-equity-diversity-and-inclusion-resources-to-inspire-change-in-the-hydrology-community/

2020 "Where should I submit? Choosing a hydrology journal", a blog post prepared for the AGU-H3S community resources, https://agu-h3s.org/2020/04/24/hydrology-journals/

Teaching Experience:

2019 Design open educational resources for Physical Hydrology course on HydroLearn

2018 Terrain analysis and TOPMODEL in Physical Hydrology course, Utah State University.

2017 Terrain Analysis in GIS in Water Resources course, Utah State University.

2014 TA for Water Resources Planning and Management course, the University of Tehran.

Professional Organization Member:

2020 – 2021 Member, AGU Hydrology Section Student Subcommittee

2020	Member, American Water Work Association (AWWA)
2016 – Present	American Geophysical Union (AGU), student member

Skills:

Computer: UNIX, Python, ArcGIS, QGIS, Google Earth Engine, Git, HPC, SQL, WRF-Hydro

Languages: English: Fluent; Farsi: Fluent; Kurdish: Fluent

References:

Dr. D. G., Tarboton, Full Professor of Civil and Environmental Engineering Department, Utah State University, Email: david.tarboton@usu.edu

Dr. J. S. Horsburgh, Associate Professor of Civil and Environmental Engineering Department, Utah State University, Email: jeff.horsburgh@usu.edu

Dr. B. Lane, Assistant Professor of Civil and Environmental Engineering Department, Utah State University, Email: belize.lane@usu.edu

Dr. J. Bales, Executive Director, Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI), Email: jdbales@cuahsi.org