

Independent Evaluation of Frozen Precipitation from WRF and PRISM in the Olympic Mountains

WILLIAM RYAN CURRIER, THEODORE THORSON,^a AND JESSICA D. LUNDQUIST

Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington

(Manuscript received 10 February 2017, in final form 19 July 2017)

ABSTRACT

Estimates of precipitation from the Weather Research and Forecasting (WRF) Model and the Parameter-Elevation Regressions on Independent Slopes Model (PRISM) are widely used in complex terrain to obtain spatially distributed precipitation data. The authors evaluated both WRF (4/3 km) and PRISM's (800-m annual climatology) ability to estimate frozen precipitation using the hydrologic model Structure for Unifying Multiple Modeling Alternatives (SUMMA) and a unique set of spatiotemporal snow depth and snow water equivalent (SWE) observations collected for the Olympic Mountain Experiment (OLYMPEX) ground validation campaign during water year 2016. When SUMMA was forced with WRF precipitation and used a calibrated, wet-bulb-temperature-based method for partitioning rain versus snow, its estimation of near-peak SWE was biased low by 21% on average. However, when SUMMA was allowed to partition WRF total precipitation into rain and snow based on output from WRF's microphysical scheme (WRF_{MPP}), simulations of snow depth and SWE were near equal to or better than simulations that used PRISM-derived precipitation with the calibrated partitioning method. Over all sites, WRF_{MPP} and simulations that used PRISM-derived precipitation had relatively unbiased estimates of near-peak SWE, but both simulated absolute errors in near-peak SWE of 30%–60% at a few locations. Since, on average, WRF_{MPP} had similar errors to PRISM, WRF_{MPP} suggested a promising path forward in hydrology, as it was independent of gauge data and did not require SWE observations for calibration. Furthermore, in similar maritime environments, hydrologic modelers should pay close attention to decisions regarding rain-versus-snow partitioning, wind speed, and incoming longwave radiation.

1. Introduction

Quantifying the amount of precipitation that falls as snow in complex terrain, where we have limited observations, remains a challenge. Methods that produce estimates of spatially distributed precipitation range from physically based numerical weather models, such as the Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008), to statistical models that spatially interpolate surface precipitation observations. A widely used statistical model is the Parameter-Elevation Regressions on Independent Slopes Model (PRISM), which is based on statistical regressions that account for topography and coastal proximity (Daly et al. 2008). The

PRISM climatology has been used to spatially interpolate gauge observations of precipitation to a grid in many spatiotemporal datasets, including Hamlet and Lettenmaier (2005), Maurer et al. (2002), NLDAS-2, Hamlet et al. (2010), Livneh et al. (2013), and NCEP Stage IV.

Previous studies have found that in complex terrain there is significant uncertainty in spatially distributed precipitation estimates (Gutmann et al. 2012; Livneh et al. 2014; Henn et al. 2016) due to a sparse network of gauges (Lundquist et al. 2003) and observational uncertainty at the gauge itself (Goodison et al. 1998; Rasmussen et al. 2012). WRF or PRISM are frequently used to force hydrologic models, which guide decisions regarding avalanche control, reservoir storage, and flood forecasting. Therefore, uncertainties in the estimation of precipitation directly translate into uncertainties in forecasts for agriculture, transportation, hydroelectric power, and recreation.

The Olympic Mountain Experiment (OLYMPEX) was a ground validation campaign for the NASA Global

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-17-0026.s1>.

^a Current affiliation: Stanford University, Palo Alto, California.

Corresponding author: William Ryan Currier, currierw@uw.edu

Precipitation Measurement (GPM) mission on the Olympic Peninsula in Washington, United States (Houze et al. 2017), and offered a unique opportunity to compare the performance of WRF and PRISM-derived precipitation in a maritime mountain environment. While the Olympic Mountains have been the focus location of previous studies, which evaluated dynamical (Anders et al. 2007; Minder et al. 2008) and statistical (Daly et al. 2008) precipitation models, its historical lack of mountain observations allowed us to estimate how both approaches work at higher mountain elevations, where data were not previously available for model training and development.

For the OLYMPEX campaign, we collected a unique set of independent snow depth and SWE observations (using cameras, poles, snow course observations, and lidar, described in section 3). We used these observations to evaluate the ability of PRISM and a high-resolution (4/3 km) atmospheric model simulation (WRF; Mass et al. 2003) to determine frozen precipitation throughout water year (WY) 2016 and during individual storm events (focused on the OLYMPEX intensive observational period from November to December 2015).

This paper is organized as follows. Section 2 provides background information on previous evaluations of WRF and PRISM. Section 3 describes the location of this study and the data used. Section 4 explains our methodology. Section 5 presents the results, and section 6 discusses sensitivities within this study. Section 7 offers conclusions.

2. Background

WRF and PRISM are both commonly used to obtain spatially distributed estimates of precipitation. WRF is an atmospheric model, which simulates atmospheric dynamics and contains a cloud microphysical scheme. The microphysical scheme parameterizes processes that control the formation, growth, and fallout of precipitation from clouds. WRF does not require surface gauge observations and represents varying synoptic conditions but is sensitive to various model decisions, such as model resolution (Colle et al. 1999; Colle and Mass 2000), boundary conditions (Yang et al. 2012), and the chosen microphysical scheme (Jankov et al. 2009; Liu et al. 2011; Minder and Kingsmill 2013). In contrast, PRISM is a gridded climatology map, which estimates the spatial variability in precipitation due to precipitation observations, topography, and coastal proximity. To obtain a spatiotemporal dataset, total precipitation (solid and liquid) observations at the surface from nearby gauges are used with the climatology to spatially interpolate precipitation between observations. Therefore, PRISM-derived precipitation is highly dependent

on the presence and quality of nearby precipitation observations.

A previous evaluation of WRF and PRISM in Colorado showed that significant differences appeared at locations furthest away from precipitation observations and that PRISM was biased about 150% in estimates of winter precipitation compared to an independent SNOTEL observation (Gutmann et al. 2012). In the Sierra Nevada, PRISM-derived spatiotemporal precipitation datasets were shown to perform well compared with independent observations on a total water year time scale (median $\pm 10\%$ errors); however, significant errors occurred during unusual synoptic conditions (Lundquist et al. 2015).

In the Olympic Mountains, where little model training data exists, the PRISM climatology has been shown to perform better than other statistical precipitation models (Thornton et al. 1997; Hijmans et al. 2005) because it was able to simulate the nonmonotonic relationship between precipitation and elevation (Daly et al. 2008). Meanwhile, mesoscale atmospheric models were able to capture small-scale (~ 10 km) orographic precipitation enhancement in the Olympic Mountains on annual and seasonal time scales, but individual events contained significant errors (Anders et al. 2007; Minder et al. 2008). Furthermore, in a similar climate, mesoscale atmospheric models have helped resolve issues with rain-versus-snow partitioning by using the microphysical scheme output to calculate the fraction of rain and snow in an individual event instead of relying solely on surface temperature (Wayand et al. 2016a).

Herein, we further evaluated both WRF and PRISM's ability to estimate frozen precipitation using a unique spatiotemporal snow depth and SWE dataset collected during the OLYMPEX campaign. However, because our snow depth and SWE observations are not a direct measurement of frozen precipitation, we relied on a hydrologic model. Precipitation is the greatest source of uncertainty in a snow model (Raleigh et al. 2015), and we therefore used the hydrologic model (evaluated at four nearby SNOTEL sites) to simulate snow depth and SWE from estimates of precipitation by WRF and PRISM. We then compared these simulations to independent observations of snow depth and SWE across the mountain range.

3. Location and data

a. Location and climate

The Olympic Mountains are located in the northwestern corner of Washington State, United States (Fig. 1). The mountain range causes significant gradients

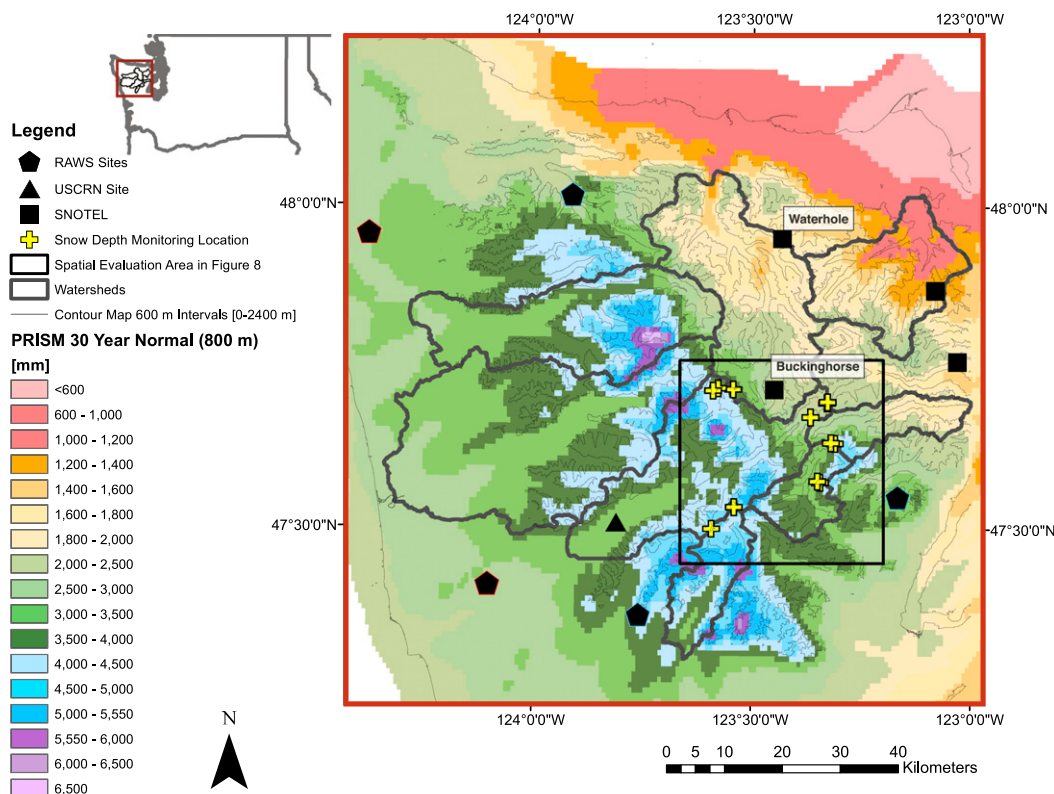


FIG. 1. Independent snow-depth-monitoring locations used within this study relative to PRISM 30-yr annual precipitation averages on the Olympic Peninsula in northwest Washington State, United States. Acronyms and spatial distributions of snow monitoring site locations are shown in Fig. 8. Red outlined RAWS sites are for sites at an elevation < 200 m. Blue outlined RAWS sites are between 670 and 730 m, where snowfall is possible.

in precipitation as moisture-laden southwest flow is orographically uplifted. Surface-based radiosonde data from 1973 to 2007 at the nearby Quillayute sounding site showed that the median 0°C isotherm during December–March precipitation events was at 1200 m. This rain–snow transition zone is further demonstrated at the SNOTEL sites (elevation range of 1270–1527 m), where we found that 40% of the hours with observed precipitation between 1 November 2015 and 1 April 2016 fell between -1° and 2°C . This makes rain-versus-snow partitioning of total precipitation critical for accurately simulating snowfall in this environment.

b. Snow depth monitoring sites

Within Olympic National Park, we monitored snow depth, temperature, and relative humidity (RH) at 12 sites. At each location, 2–3 Wingscape time-lapse cameras were deployed in a nearby tree and took pictures of 3–4 marked snow depth poles every hour during daylight hours [0900–1600 Pacific standard time (PST)]. Poles were located in flat, grassy, forest clearings (~ 10 – 25 m diameter) and ranged in height

from 4 to 6 m. Each pole had black tape every 5 cm and brightly colored tape every 50 cm. Currier (2016) describes the processing of the camera images to snow depth values. When and where the uncertainty in the measurements was greater than ± 5 cm, the measurements were not used for evaluation, but were instead shown with uncertainty bounds to provide guidance in the evolution of the snowpack. Uncertainty was based on laboratory experiments, where poles were bent at various angles, and the camera viewed the pole at different angles (Currier 2016). Pole measurements were averaged together if two or more poles provided measurements with uncertainty of less than 5 cm.

Adjacent to the cameras, in conifer trees, we placed HOBO U23 Prov2 temperature/RH sensors within plastic radiation shields following the methods of Lundquist and Huggett (2008). The temperature sensors were reported by the manufacturer to have an uncertainty of $\pm 0.21^{\circ}\text{C}$ at 0°C , with uncertainty increasing to about $\pm 0.75^{\circ}\text{C}$ at -40°C . RH accuracy was within $\pm 2.5\%$ between 10% and 90% RH. The

uncertainty below 10% and above 90% increases to a maximum of $\pm 3.5\%$, including hysteresis.

c. NRCS SNOTEL data

Observations of hourly temperature, daily incremental precipitation, daily snow depth, and daily SWE data were received from the four National Resource Conservation Service (NRCS) SNOTEL locations. Two of the SNOTEL sites, Buckinghorse and Waterhole, also provided hourly observations of RH, and Waterhole provided hourly averaged wind speed. For this study, daily incremental precipitation was uniformly distributed to hourly values.

NRCS sites use precipitation accumulation reservoir gauges with antifreeze. Figure 1 shows that the majority of the SNOTEL locations are on the leeward side of the mountains. The Buckinghorse site is located in the center of the Olympic Peninsula's mountain range but was installed in 2007. Buckinghorse was therefore not used in the development of the PRISM climatology.

Current SNOTEL temperature data are biased warm at cold temperatures because of an erroneous conversion from voltages to degrees Celsius (Julander et al. 2007; Oyler et al. 2015; D. S. Harms et al. 2016, meeting presentation). We corrected for the bias in the observations by using a linear equation [(1)] based on a least squares regression between the SNOTEL temperature observations and collocated HOBO temperature sensors during WY 2014, as follows:

$$T_{\text{corr}} = 1.03 \times T_{\text{SNTL}} - 0.90, \quad (1)$$

where T_{SNTL} is the observed raw SNOTEL temperature ($^{\circ}\text{C}$) and T_{corr} is the corrected SNOTEL temperature ($^{\circ}\text{C}$).

d. RAWS data

Additional daily precipitation observations were received from several Remote Automated Weather Stations (RAWS). Again, the daily precipitation data were uniformly distributed to hourly precipitation. RAWS use unheated tipping-bucket gauges, which may be subject to freezing and thus were not used in the creation of the PRISM climatology outside of May–September (Daly et al. 2008). Three of the five RAWS stations were at elevations of around 700 m, where snowfall is possible. We therefore explored the sensitivity of our results to using these higher-elevation gauges in our precipitation estimates predicted with PRISM weights [sections 4d(1) and 6c].

e. U.S. Climate Reference Network (USCRN)

Hourly precipitation data were downloaded from the Quinault Climate Reference Network location, which uses heated weighing gauges. Hourly data were

aggregated to daily values and then uniformly distributed to be consistent with RAWS and SNOTEL data.

f. PRISM

The PRISM climatology group provides a map of total annual precipitation estimates over a 30-yr period throughout the contiguous United States. PRISM relies on a regression between elevation and observations of precipitation, and individual grid cells are further modified based on the coastal proximity and topographic facets (Daly et al. 2008). These modifications allow PRISM to estimate rain shadows and the orographic enhancement of precipitation. In this study, we used the 800-m, 30-yr (1981–2010) annual climate normal. We combined the nearest PRISM climatology grid cell with observations of precipitation from the RAWS, USCRN, and SNOTEL sites to estimate precipitation at a snow monitoring site. See section 4d(1) for more details. The maximum difference between the PRISM elevation map and the elevation of our snow monitoring sites was 317 m, with a mean difference and mean absolute difference of 65 and 96 m, respectively.

g. WRF

The WRF data were provided by the Northwest Modeling Consortium (Mass et al. 2003), which runs and archives WRF version 3.6.1 output in four nested domains (36, 12, 4, and 4/3 km). The 4/3-km nested domain encompasses Washington State and uses the Thompson et al. (2004, 2008) microphysical scheme without convective parameterizations for numerical weather prediction. Shortwave and longwave radiation simulations used the Rapid Radiative Transfer Model (Mlawer et al. 1997). WRF was run with 84-h forecasts that were initialized every 12 h. As in Minder et al. (2010) and Wayand et al. (2016a), the 12–24 h forecasts were extracted from the 84-h forecasts and concatenated to provide a temporally continuous dataset at an hourly time step. The maximum difference between the WRF terrain height and the elevation of each snow monitoring site was 361 m, with a mean difference and mean absolute difference of 70 and 105 m, respectively.

h. Airborne snow observatory snow depth and snow density observations

Two spatially complete snow depth datasets from airborne scanning lidar were provided by the Airborne Snow Observatory (ASO) team (Painter et al. 2016). Snow-on flights were flown on 8–9 February 2016 and 29–30 March 2016, and the data were processed to a 3-m gridded resolution of snow depth. The accuracy of ASO in a nonforested, flat, $15\text{ m} \times 15\text{ m}$ area has been shown

to have a mean absolute error of less than 8 cm, with an overall bias of less than 1 cm (Painter et al. 2016). In appendix A, we compare the ASO data with our snow depth pole measurements. In this paper, we focus on snowfall accumulation and not on forest–snow interactions. Therefore, we used the classification from the Compact Airborne Spectrographic Imager (CASI) 1500 imaging spectrometer, which was aboard the ASO, to remove forested pixels from the analysis. Using the March snow depth map, this step removed 17%–78% of the snow depth pixels within a 60-m bounding box (mean 54%).

Snow density observations were collected with a federal sampler on 8 February 2016 and 7 April 2016 at seven locations near our snow depth locations to convert ASO snow depth observations to SWE. To account for the snow becoming denser between 29–30 March 2016 and 7 April 2016, we took the difference, at the nearest SNOTEL site, between density observations on 30 March 2016 and 7 April 2016. We then subtracted the change in density over the eight days (20 kg m^{-3}) from the 7 April 2016 observations. Therefore, observations of density to convert ASO snow depth to SWE on 29–30 March 2016 ranged from 450 to 490 kg m^{-3} , with a mean of 480 kg m^{-3} .

4. Methodology

a. Overview

As outlined in Fig. 2, we first calibrated the snow model at the available SNOTEL sites using observed precipitation that was uniformly distributed throughout the day. We adjusted the precipitation partitioning parameters so that the model was unbiased for SWE from the start of the season until peak SWE (Fig. 2, phase 1A). We chose two sets of parameter values taken from the literature for our new snow density and compaction routines to provide a range of uncertainty in modeled snow depth (Fig. 2, phase 1B). We then used the same model structure and parameters at all of our independent snow monitoring stations, using precipitation estimated with PRISM and predicted by WRF. Detailed descriptions of how PRISM was used to distribute precipitation and how WRF data were used to run the hydrologic model are described in section 4d and appendix B. All simulations were then evaluated against our observations of snow depth and against the median ASO snow depth/SWE values within a 60-m square bounding box surrounding each site (Fig. 2, phase 2). The following sections provide more detail on the model choice, forcing data, calibration, and evaluation.

b. Model description

In this study we used the Structure for Unifying Multiple Modeling Alternatives (SUMMA) (Clark et al. 2015a,b,c). SUMMA is a modular, physically based, energy balance model with a numerical solver at its core. This allowed additional parameterizations to be added to the model and allowed multiple existing modeling approaches to be vetted against one another. SUMMA was run at an hourly time step.

c. Model calibration and evaluation at SNOTEL sites

1) RAIN-VERSUS-SNOW PARTITIONING AND SWE

Rain–snow partitioning is the most critical model parameterization in a warm maritime snow environment (Wayand et al. 2016a,b). Therefore, the model was calibrated to partition rain versus snow in WY 2016 by searching for an optimal parameter value T_{crit} . Parameter T_{crit} is the midpoint value of the wet bulb temperature T_w range. The temperature range T_{range} was held constant at SUMMA's default value, 2°C . The T_{range} describes how precipitation transitions linearly between rain and snow, as described in USACE (1956). To calibrate the model, we varied T_{crit} in increments of 0.05°C , between -2° and 2°C , to find when our objective function [(2)] had a near-zero bias from the start of the snow season until the date of peak SWE (Fig. 3). Our objective function is written as follows:

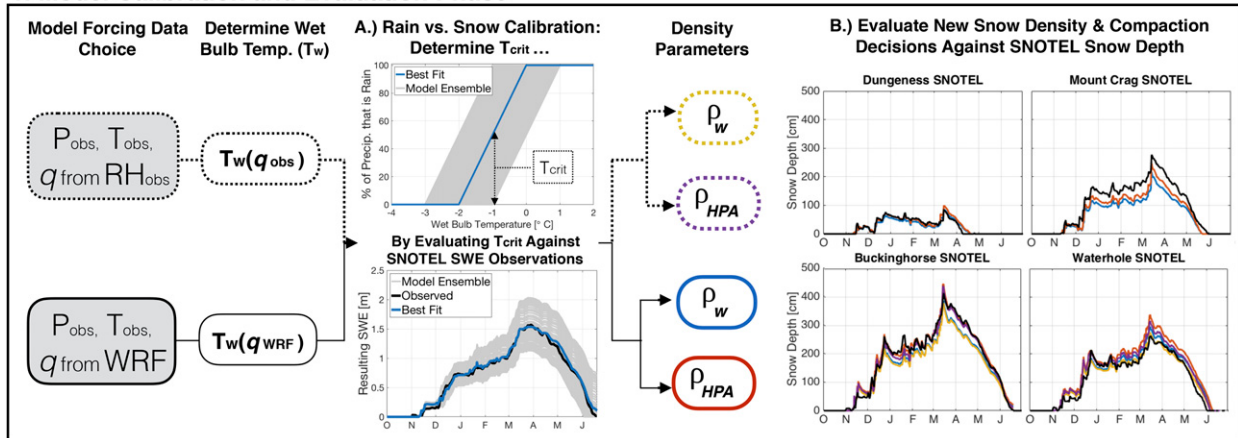
$$\overline{\text{Bias}}_{\rightarrow \text{SWE}_{\text{pk}}} = \text{abs} \left[\frac{\sum_{n=1}^{\text{SNTL}} \sum_{i=1}^k (\text{SWE}_{m_{i,n}} - \text{SWE}_{o_{i,n}}) / k}{\text{SNTL}} \right], \quad (2)$$

where n is an individual SNOTEL site and k signifies the number of days between the first day of snow accumulation and the day of peak SWE. Parameter SWE_m is modeled SWE averaged from hourly time steps to daily values, and SWE_o is observed daily SWE. Parameter SNTL is the number of SNOTEL sites used to calibrate the model. In this study, SNTL differs between two and four, depending on the source of specific humidity q used in the model.

Accurate measurements of q are essential, as the temperature during snowfall events is typically near the rain-versus-snow threshold, and SUMMA determines rain versus snow with T_w because T_w has been shown to improve the phase of the precipitation at subdaily time steps (Marks et al. 2013). Therefore, we tested the sensitivity of our model calibration to different q inputs, specifically comparing those measured in situ (Table 1, set 1)

Hydrologic Model Choice: SUMMA [Clark *et al.*, 2015a,b,c]

1. Model Calibration and Evaluation Phase



2. Frozen Precipitation Evaluation Phase

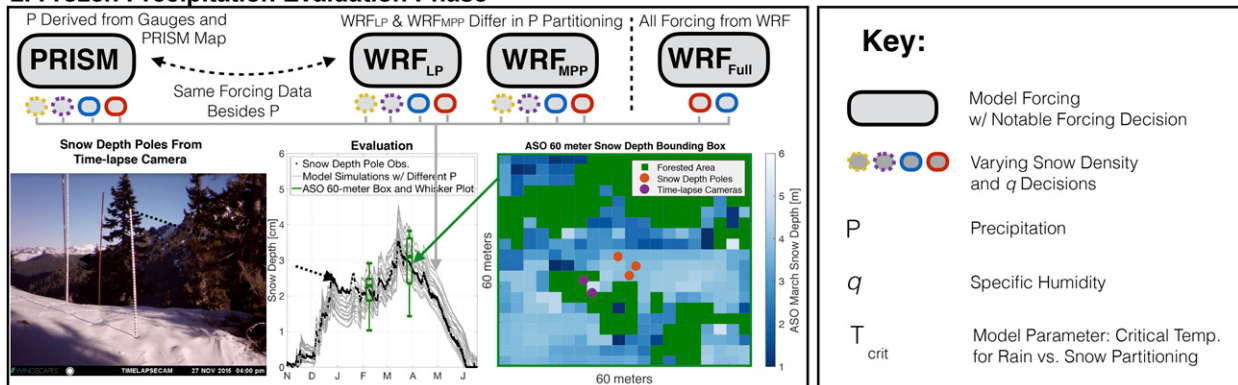


FIG. 2. Conceptual figure of the overarching methodology with two phases: 1) the calibration of the T_{crit} model parameter for rain vs snow calibration (using SWE observations) and modeled snow depth evaluation against the four available SNOTEL sites (black line) and 2) evaluating simulations of snow depth and SWE from different precipitation estimates against OLYMPLEX snow depth and SWE observations (snow depth poles and the median of an ASO 60-m bounding box—converted to SWE using snow course observations).

with those predicted by WRF (Table 1, set 2; Fig. 2, phase 1A). Despite differences in q causing differences of simulated SWE at individual SNOTEL sites (Table 1), both sets of forcing data converged on the same T_{crit} value (-1°C) for a minimum $\text{Bias} \rightarrow \text{SWE}_{pk}$.

All other model forcing data related to the energy balance were either closely evaluated at the nearby Snoqualmie Pass energy balance tower (Wayand *et al.* 2015) or are further discussed in section 6d and in appendix B. Furthermore, adjusted model parameters, beyond T_{crit} , were fit to observations from Snoqualmie Pass, Washington, which has a similar climate, or were taken from the literature (Tables C1 and C2).

2) SNOW DEPTH

We used two sets of literature values for the new snow density and compaction parameters (Table C2, Figs. 2, 3)

to generate an ensemble that accounts for the model uncertainty in simulating snow depth. Within these two sets of snow density parameters, the choice in q was also varied to generate a four-member ensemble at SNOTEL sites with observed RH. The first snow density parameter set ρ_w used literature values from Wayand *et al.* (2016b), who evaluated various parameters at the nearby Snoqualmie Pass. The other set of snow density parameters ρ_{HPA} came from two separate studies. New snow density parameters came from the Hedstrom and Pomeroy (1998) study because this represented the opposite end of the new snow density spectrum. The ρ_{HPA} parameter values were optimized from new snow density observations in a cold continental climate, which typically provides much lower new snow density values (Judson and Doesken 2000; LaChapelle 1958). Last, in ρ_{HPA} , we used default compaction parameters from the commonly used Anderson (1976) parameterization.

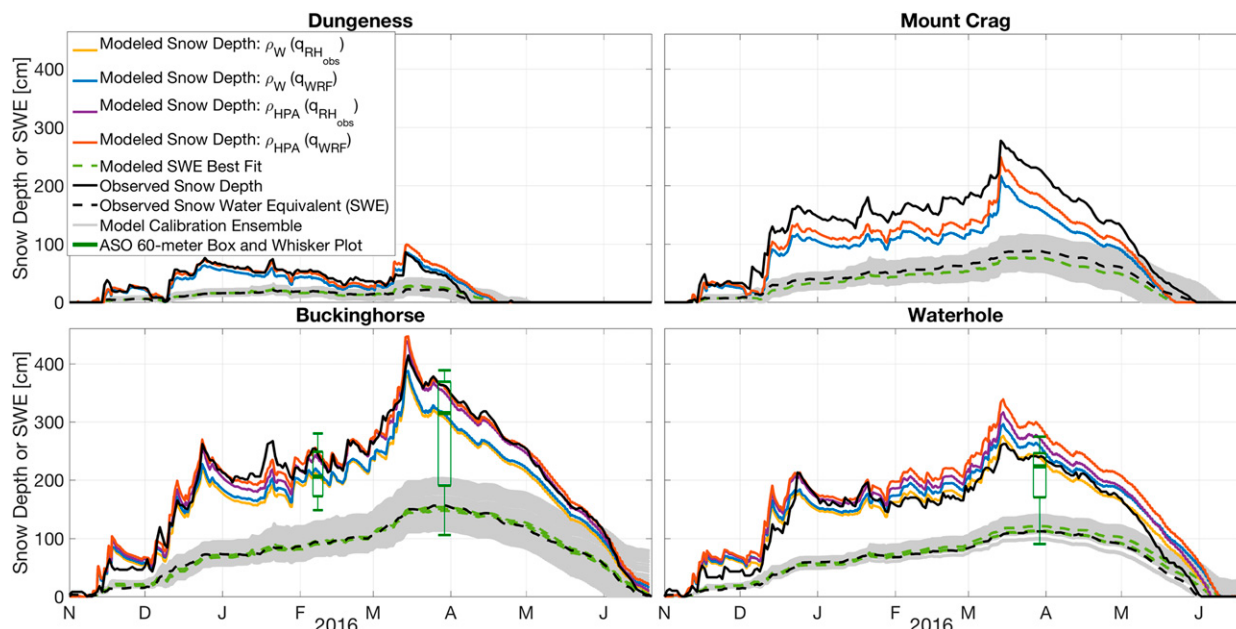


FIG. 3. Best-fit modeled SWE from the calibration ($T_{\text{crit}}: -1^{\circ}\text{C}$, $T_{\text{range}}: 2^{\circ}\text{C}$) and the ensemble of snow depth simulations. ASO lidar observations were available over two of the four SNOTEL sites. The minimum and maximum values of the box-and-whisker plot show the bottom 10% and top 90% of the snow depth values within the 60-m bounding box, respectively.

Simulations of snow depth for various snow density and specific humidity decisions overall performed well, with parameters from ρ_w consistently simulating lower snow depth values than the ρ_{HPA} parameters (Fig. 3). Neither parameter set provided a perfect model simulation, nor was one consistently underpredicting or overpredicting observed snow depth (Table 2). Furthermore, there was a significant distribution of snow depth within a spatial field of 60 m, surrounding a NRCS SNOTEL site (Fig. 3). We used the mean of the ensemble going forward to reduce the dimensionality of the model uncertainty and because the mean was within the distribution of ASO snow depth values.

When evaluating our model at the SNOTEL sites (Table 2), we chose 29–30 March 2016 as an evaluation date instead of peak SWE because 29–30 March was the closest date to peak SWE (~ 13 days after) that we had observations at all of our sites. We found the percent errors in modeled versus observed SWE at the SNOTEL sites to be relatively consistent, temporally, and normally distributed around 0%, with a standard deviation of 10% and a 95% confidence interval between -19% and 19% . Therefore, in the annual precipitation evaluation phase (Fig. 2, phase 2), we determined an over-accumulation or underaccumulation of total annual frozen precipitation to occur when the difference between observed SWE and the modeled SWE was greater than 19% on 29–30 March 2016.

d. PRISM-derived precipitation and WRF Model configurations

1) HYDROLOGIC MODEL SIMULATIONS WITH PRISM-DERIVED PRECIPITATION

In the frozen precipitation evaluation phase, model runs denoted PRISM used the PRISM annual climatology with n sources of precipitation P_i to estimate precipitation at a snow monitoring station. Except for the sensitivity discussion in section 6c, n is equal to 9 and contains all SNOTEL, RAWs, and USCRN gauges shown in Fig. 1. At each snow monitoring station, the ratio m_i between the nearest 30-yr normal PRISM annual precipitation value at the snow

TABLE 1. Bias until peak SWE from WY 2016 at four NRCS SNOTEL sites using model forcing decisions as described in Table B1 and using model parameters from Table C1.

	Bias until peak SWE (cm)	
	Model forcing: Set 1	Model forcing: Set 2
Dungeness	—	0
Mount Crag	—	−6
Buckinghorse	−2	2
Waterhole	1	5
All (mean)	0	0

TABLE 2. Difference between modeled SWE and observed SWE on 29–30 Mar 2016 at four NRCS SNOTEL sites. Model differed in various snow density (Table C2) and specific humidity q decisions.

Specific humidity source	SWE differences on 29–30 Mar 2016 [cm (%)]				
	q from obs. RH	q from WRF	q from obs. RH	q from WRF	Ensemble mean
New snow density and compaction parameters	Wayand et al. (2016b)	Wayand et al. (2016b)	Hedstrom and Pomeroy (1998) and Anderson (1976)	Hedstrom and Pomeroy (1998) and Anderson (1976)	
Dungeness	—	3 (14)	—	4 (17)	3 (15)
Mount Crag	—	−13 (−14)	—	−12 (−13)	−12 (−14)
Buckinghorse	−10 (−6)	−6 (−4)	−7 (−4)	−3 (−2)	−6 (−4)
Waterhole	0 (0)	9 (8)	2 (1)	9 (8)	5 (4)
All (mean)	−5 (−3)	−2 (1)	−3 (−1)	0 (3)	−3 (0)

monitoring station (PRISM_{SMS}) and all precipitation sites (PRISM_{Gauge}) was determined:

$$m_i = \frac{\text{PRISM}_{\text{SMS}}}{\text{PRISM}_{\text{Gauge}}}. \quad (3)$$

The relative inverse distance was also calculated to determine a set of weights for an individual snow monitoring site w_i :

$$w_i = \frac{1/d_i}{\sum_{i=1}^n 1/d_i}, \quad (4)$$

where d_i is the distance between a snow monitoring site and a precipitation source. In section 6c we discuss using inverse distance squared weighting (IDSW) $1/d_i^2$ versus inverse distance weighting (IDW). The total precipitation at a snow monitoring station P_{sms} , which was used to run the hydrologic model, was then calculated:

$$P_{\text{sms}} = \sum_{i=1}^n P_i m_i w_i. \quad (5)$$

This allowed a precipitation site that was closer to the snow monitoring station to influence the precipitation at the snow monitoring station more than a source of precipitation farther away. We note that we tested the same methodology with the monthly climatology datasets, where m_i changes based on the month, and found nearly equivalent simulations of snow depth and SWE.

2) HYDROLOGIC MODEL CONFIGURATIONS WITH WRF DATA

Model runs using WRF precipitation partitioned precipitation in two different ways. In WRF_{LP}, the nearest precipitation grid cell had precipitation aggregated to daily values and then uniformly distributed to hourly values to

be consistent with the methods used during model calibration and with the PRISM simulations. WRF_{LP} then used the same linear precipitation–partitioning (LP) scheme as was used in the calibration and in simulations that used PRISM-derived precipitation.

WRF_{MPP} used output from WRF’s microphysical scheme to partition the hourly precipitation. SUMMA was modified to allow the fraction of rain or snow in any given event to be prescribed using output from WRF’s microphysical scheme. Herein, precipitation partitioning was calculated as

$$\text{rain fraction} = 1 - \frac{\text{Snow} + \text{Hail} + \text{Graupel}}{\text{Total Precipitation}}, \quad (6)$$

where snow, hail, graupel, and total precipitation were extracted from WRF’s microphysical scheme (Thompson et al. 2008). Both WRF_{LP} and WRF_{MPP} have the same model forcing data and model decisions as in the calibration and PRISM model runs (except for those related to precipitation and partitioning). WRF_{Full} used the same method for precipitation and precipitation partitioning as WRF_{MPP}. However, WRF_{Full} used WRF output for all of the hydrologic model forcing data. Model forcing data are detailed in appendix B.

5. Results

a. Frozen precipitation evaluation

1) ANNUAL DIFFERENCES

Modeled and observed snow depth time series for WY 2016 represented storm timing well (Fig. 4). Furthermore, simulations of snow depth and SWE diverged based on their source of precipitation (PRISM and WRF_{LP}) and based on how the precipitation was partitioned (WRF_{LP} and WRF_{MPP}; Figs. 4, 5). PRISM, WRF_{LP}, and WRF_{MPP} simulations of snow depth and

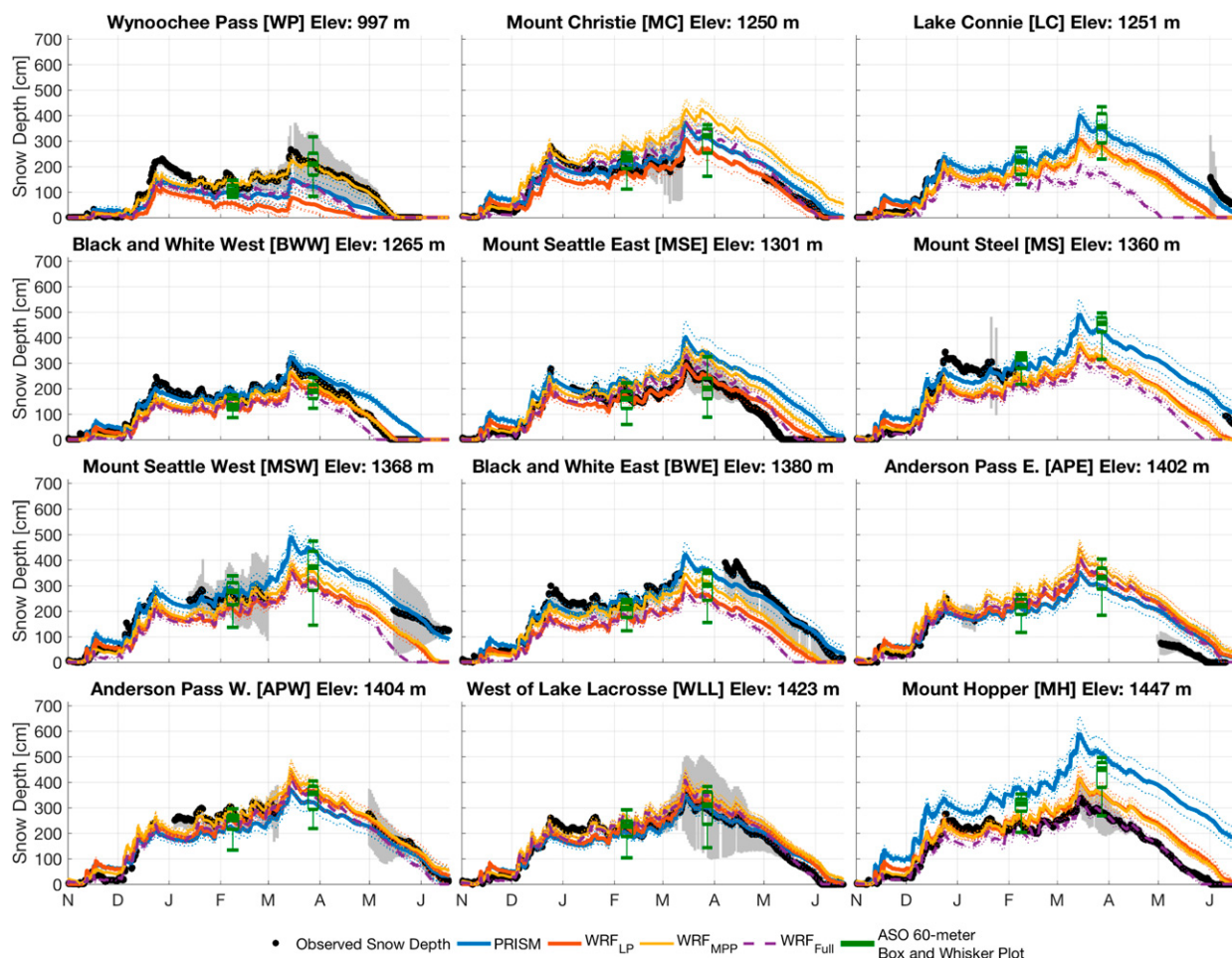


FIG. 4. Snow depth time series at all 12 independent snow monitoring sites. The thick line represents the ensemble mean, while the thin dotted lines represent individual ensemble members. Black dotted lines with gray shading show snow depth measurements from the time-lapse cameras along with their uncertainty.

SWE had similar mean absolute percent differences, while WRF_{Full} had the highest mean absolute percent difference and was biased low in both snow depth and SWE (Table 3). We attributed the overall low bias in WRF_{Full} to high wind speeds and ablation errors rather than precipitation errors and leave this to further discussion in section 6d.

WRF precipitation, partitioned by the microphysical scheme output (WRF_{MPP}), had a smaller mean difference than WRF precipitation partitioned using the calibrated linear threshold (WRF_{LP}). WRF_{LP} was generally biased low, with a mean difference in SWE across all sites of -33 cm (-21%), which was outside the snow model's 95% confidence interval. Furthermore, WRF_{LP} had a higher mean absolute difference, which showed that errors were generally larger in magnitude for WRF_{LP} than WRF_{MPP} . Similar percent differences were also found for simulations of snow depth.

The WRF_{MPP} hydrologic model simulated SWE and snow depth with similar skill to PRISM simulations, as they had similar mean differences and mean absolute differences. The best performance was related to how the errors were reported and to the metric used. For instance, PRISM and WRF_{MPP} simulations were both generally unbiased, with similar mean percent errors. WRF_{MPP} generally had larger absolute errors than PRISM simulations (higher mean absolute difference), but these were generally located at sites with more observed snow (lower mean absolute percent difference). PRISM's ensemble mean fell outside the range of the snow model's uncertainty ($\pm 19\%$) at three sites, compared to WRF_{MPP} , which fell outside of this range at five sites. This implies we are more confident that PRISM was able to estimate frozen precipitation at more locations on an annual basis. However, it is difficult to definitively say which precipitation estimate performed best because both WRF_{MPP}

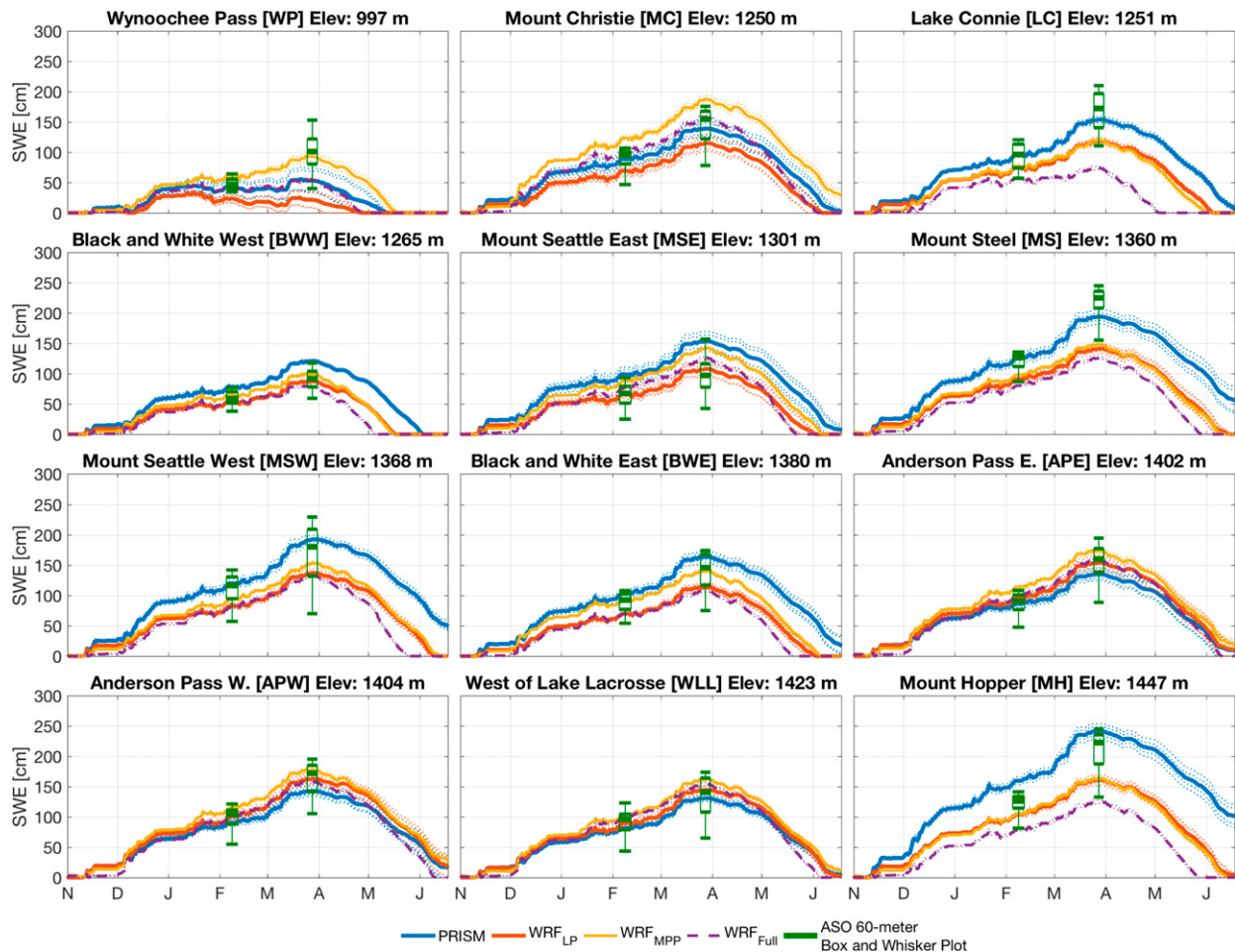


FIG. 5. SWE simulations at all 12 independent snow monitoring sites. The thick line represents the ensemble mean, while the thin dotted lines represent individual ensemble members. Observed snow depth from ASO was converted to SWE using nearby snow density observations.

and PRISM simulations were generally unbiased when compared to the observations of SWE or snow depth.

2) OLYMPEX INTENSIVE OBSERVATIONAL PERIOD

We had high-quality observations of snow depth from the time-lapse camera network during the intensive observational period of the OLYMPEX campaign (from 1 November to 23 December 2015; Fig. 4). During this time period we averaged our snow depth observations to daily averages and our model estimates of snow depth between 0900 and 1600 PST to be consistent with the observations. We then compared cumulative sums of the positive differences between daytime averages of snow depth from the model and the observations.

We computed cumulative sums of snow accumulation for the entire OLYMPEX intensive observational period and for the December snowstorms (4–23 December 2015).

Differences in accumulated snowfall between modeled and observed are shown in Tables 4 and 5.

From 1 November to 23 December 2015, we found the difference between the modeled cumulative sum (all ensemble members) at the SNOTEL sites and the cumulative sum of the observations to have a mean percent difference of -4% and a standard deviation among all ensemble members to be 15% . Using a normal distribution with a mean of -4% and the standard deviation of 15% , we found that the 95% confidence interval was between -34% and 24% . Using the same analysis during the December storms, the mean of the percent errors in total accumulation was -3% with a 95% confidence interval between -29% and 30% .

(i) 1 November–23 December 2015

PRISM, WRF_{MPP} , and WRF_{Full} simulations all had similar mean absolute differences (16% – 18%) and

TABLE 3. Ensemble mean of modeled SWE from various sources of precipitation compared to observations on 29–30 Mar 2016. The average modeled SWE value between 29 and 30 Mar 2016 from the ensemble mean was compared to the median value within a 60-m bounding box from the ASO snow depth data that had the forest values removed (converted to SWE with nearby density observations). Simulations at sites that are outside of the model's 95% confidence interval are in bold.

Site name	March ensemble mean difference [modeled – observed; cm (%)]			
	PRISM	WRF _{LP}	WRF _{MPP}	WRF _{Full}
Wynoochee Pass	–49 (–48)	–80 (–79)	–9 (–9)	–48 (–48)
Mount Christie	–16 (–10)	–40 (–26)	32 (21)	2 (1)
Lake Connie	–18 (–11)	–54 (–31)	–53 (–31)	–99 (–58)
Black and White West	29 (31)	–5 (–5)	7 (8)	–14 (–15)
Mount Seattle East	56 (58)	11 (11)	45 (46)	28 (29)
Mount Steel	–31 (–14)	–84 (–37)	–78 (–34)	–100 (–44)
Mount Seattle West	12 (7)	–43 (–24)	–27 (–15)	–48 (–26)
Black and White East	18 (12)	–32 (–22)	–6 (–4)	–39 (–26)
Anderson Pass East	–25 (–16)	–7 (–4)	13 (8)	–3 (–2)
Anderson Pass West	–30 (–17)	–9 (–5)	5 (3)	–15 (–9)
West of Lake Lacrosse	–10 (–7)	4 (3)	21 (15)	13 (9)
Mount Hopper	19 (8)	–62 (–28)	–59 (–27)	–98 (–44)
Total Sites				12.0
Mean difference	–4 (–1)	–33 (–21)	–9 (–2)	–35 (–19)
Mean absolute difference	26 (20)	36 (23)	30 (18)	42 (26)
Std dev of errors	30 (27)	32 (24)	38 (23)	45 (26)
Number of occurrences outside $\pm 19\%$	3	7	5	7

mean differences, which were less than the model uncertainty (Table 4; Fig. 6). WRF_{LP} was within the model uncertainty, but it had the most dramatic low bias (–20%). In general, PRISM simulations overaccumulated the November storm but underaccumulated the December storms (Figs. 6, 7), resulting in an overall mean percent difference of –10%. While

PRISM simulations had a smaller mean difference, and mean absolute difference, they exhibited a larger standard deviation than the WRF simulations that used the microphysical partitioning, indicating that PRISM simulations had compensating errors. WRF simulations with the microphysical partitioning (WRF_{MPP} and WRF_{Full}) differed slightly in cumulative snow depth.

TABLE 4. Total difference in accumulated snow depth between the model ensemble mean and observed, accumulated snow depth during the intensive OLYMPEX observational period (from 1 Nov to 23 Dec 2015). Simulations at sites that are outside of the model's 95% confidence interval are in bold.

Site name	Ensemble mean difference [modeled – observed; cm (%)]			
	PRISM	WRF _{LP}	WRF _{MPP}	WRF _{Full}
Wynoochee Pass	–125 (–41)	–170 (–55)	–108 (–35)	–117 (–38)
Mount Christie	–28 (–8)	–68 (–20)	5 (1)	–6 (–2)
Lake Connie	–16 (–5)	–57 (–18)	–99 (–32)	–96 (–31)
Black and White West	–64 (–19)	–109 (–33)	–123 (–37)	–105 (–32)
Mount Seattle East	8 (2)	–56 (–17)	–36 (–11)	–18 (–5)
Mount Steel	–59 (–14)	–121 (–29)	–118 (–28)	–120 (–28)
Mount Seattle West	–49 (–11)	–119 (–27)	–121 (–28)	–104 (–24)
Black and White East	20 (8)	–62 (–25)	–28 (–11)	–27 (–11)
Anderson Pass East	–45 (–15)	7 (2)	17 (5)	23 (7)
Anderson Pass West	–109 (–29)	–59 (–16)	–54 (–14)	–48 (–13)
West of Lake Lacrosse	–70 (–21)	–19 (–6)	–22 (–6)	–7 (–2)
Mount Hopper	122 (37)	4 (1)	–15 (–6)	–11 (–3)
Total Sites				12.0
Mean difference	–35 (–10)	–69 (–20)	–58 (–17)	–53 (–15)
Mean absolute difference	60 (17)	71 (21)	62 (18)	57 (16)
Std dev of the errors	65 (20)	53 (16)	52 (15)	52 (15)

TABLE 5. Total difference in accumulated snow depth between the model ensemble mean and observed, accumulated snow depth during the December snow storms (accumulation period of 4–23 Dec 2015). Simulations at sites that are outside of the model's 95% confidence interval are in bold.

Site name	Ensemble mean difference [modeled – observed; cm (%)]			
	PRISM	WRF _{LP}	WRF _{MPP}	WRF _{Full}
Wynoochee Pass	–118 (–47)	–144 (–57)	–89 (–36)	–99 (–39)
Mount Christie	–48 (–21)	–68 (–29)	1 (1)	–7 (–3)
Lake Connie	–48 (–21)	–79 (–34)	–69 (–30)	–71 (–30)
Black and White West	–66 (–28)	–89 (–37)	–65 (–27)	–51 (–21)
Mount Seattle East	–34 (–14)	–69 (–28)	–20 (–8)	–16 (–6)
Mount Steel	–99 (–30)	–132 (–41)	–102 (–31)	–113 (–35)
Mount Seattle West	–68 (–22)	–111 (–36)	–75 (–24)	–73 (–24)
Black and White East	–4 (–3)	–44 (–30)	–4 (–2)	–8 (–6)
Anderson Pass East	–75 (–31)	–37 (–15)	5 (2)	–1 (0)
Anderson Pass West	–148 (–47)	–111 (–35)	–72 (–23)	–78 (–25)
West of Lake Lacrosse	–72 (–30)	–34 (–14)	–1 (0)	–2 (–1)
Mount Hopper	24 (10)	–41 (–16)	–25 (–10)	–31 (–12)
Total sites				12.0
Mean difference	–63 (–24)	–80 (–31)	–43 (–16)	–46 (–17)
Mean absolute difference	67 (25)	80 (31)	44 (16)	46 (17)
Std dev of the errors	47 (16)	38 (12)	39 (14)	40 (14)

Differences were attributed to new snow densities, as WRF's air temperature was on average 0.8°C colder than observed temperatures. Additionally, WRF_{Full} simulations generally had higher wind speeds and more incoming longwave radiation, leading to melt during accumulation events.

(ii) 4–23 December 2015

Considering just December, when the Olympic Mountains received most of its snow in WY 2016 (~2–3 m of accumulated snow depth), WRF_{MPP} and WRF_{Full} had the smallest mean difference (–15% and –16%) and mean absolute difference, but generally underaccumulated (Table 5; Fig. 7). Meanwhile, WRF_{LP} was also biased low, with the mean percent difference falling outside the 95% confidence interval. PRISM simulations on average accumulated less snow than WRF simulations that used the microphysical partitioning method but more snow than WRF_{LP}.

The temperature-based threshold led to errors in PRISM and WRF_{LP} simulations over shorter time periods. For instance, observations showed that during the 4–6 December 2015 period, snowfall increased, while PRISM and WRF_{LP} simulations generally stopped accumulating after 5 December. During the 5–6 December period, T_w increased from –1°C to around 0°C. Snowfall occurred and was captured by the microphysical partitioning method, but our calibrated linear partitioning method had no snowfall occur if T_w was at or above 0°C. Further partitioning errors are discussed in section 6b.

We also note that all methods were biased low by about 30% during the 19–23 December period, where observations showed accumulations of 1.0–1.5 m of snow depth. Parameter T_w during this time period was consistently below –2°C (100% snow), and therefore errors were not due to partitioning. We also note that modeled snow accumulation at the four SNOTEL sites, which used observed precipitation, had a mean difference of 1%, and therefore errors were likely not caused by gauge undercatch or errors within the hydrologic model but were likely due to precipitation estimates. During this time period, the primary storm direction was from the northwest rather than from the southwest. We suggest that future work more thoroughly investigate how the primary storm direction influences WRF and PRISM-derived estimates of snowfall within the Olympic Mountains.

b. Spatial distribution of errors

No set of model simulations had errors that resulted in a definitive spatial pattern (Fig. 8). However, we speculate that the annual average precipitation values in PRISM are too low for the prediction of snowfall in the southern Elwha Watershed [Mount Christie (MC), Buckinghorse (BK)] and in eastern Quinalt [Anderson Pass East (APE), Anderson Pass West (APW), West of Lake Lacrosse (WLL)], as the PRISM simulations of snow depth and SWE are biased low in this region. The percent differences between modeled SWE with PRISM-estimated precipitation and ASO-derived SWE at BK (–15%), MC (–10%), WLL (–10%) and both

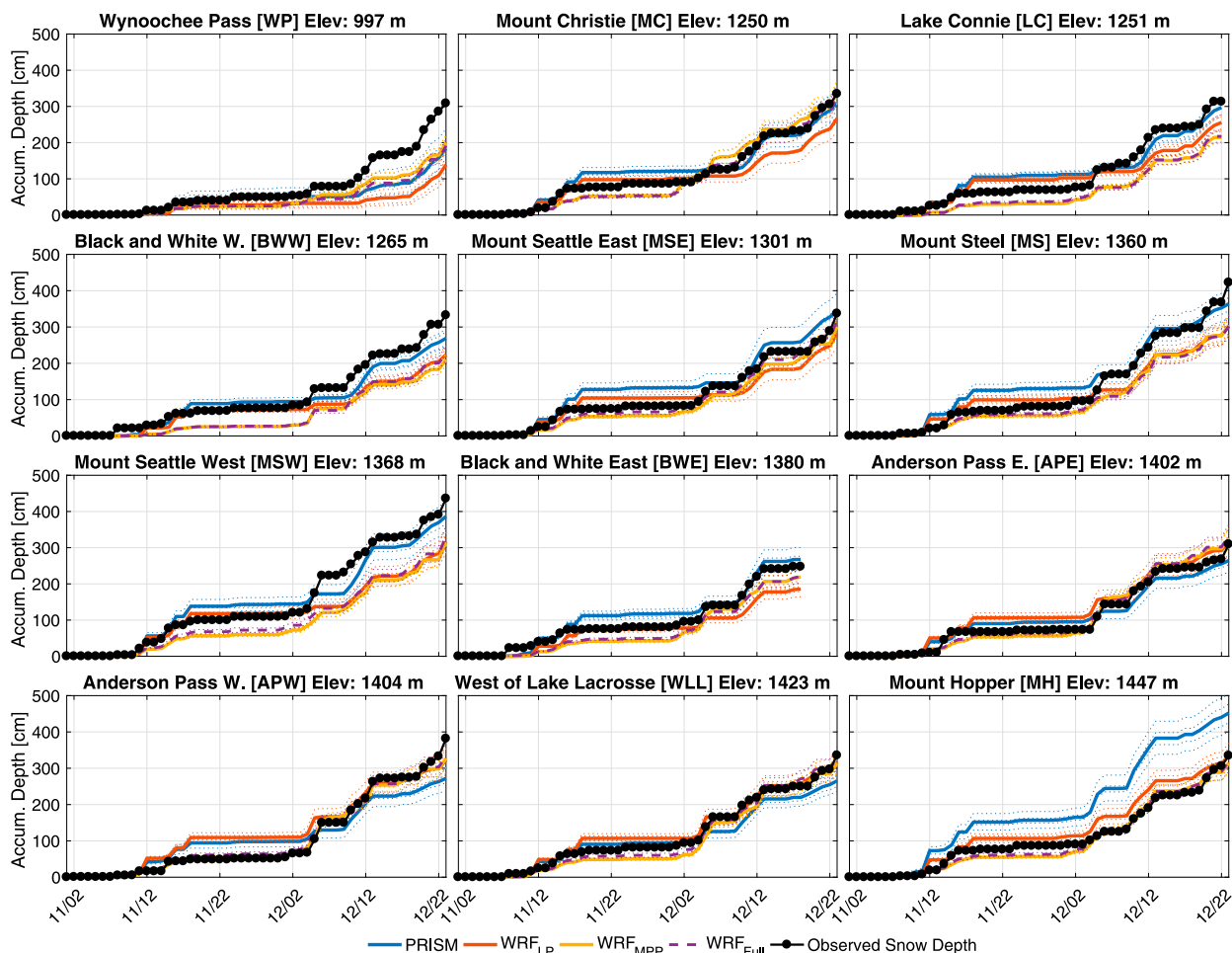


FIG. 6. Accumulated snow depth between 1 Nov and 23 Dec 2015. Individual ensemble members are shown with thin dotted lines. At Black and White East and Lake Connie, the evaluation ends on 18 and 22 Dec 2015, respectively. After these dates the poles became significantly bent toward the camera or buried.

APE (-16%), and APW (-17%) suggest that the PRISM climatology values are too low. We also note that PRISM-estimated total precipitation at the Buckinghorse SNOTEL site (not used in the development of PRISM) is also biased low 25% on an annual basis in WY 2016 when using all precipitation gauges besides Buckinghorse.

APE, APW, and WLL were located on the windward side of the Olympic Mountains near the Eel Glacier and Anderson Glacier, on the banks of a deep U-shaped valley, referred to as the Enchanted Valley. Because of the sparse gauge distribution in the Olympics, and the absence of the Buckinghorse SNOTEL site during the development of PRISM, the DEM used to derive PRISM's topographic facets (defining windward versus leeward) smoothed the relatively narrow Enchanted Valley (~ 4 km; see Daly et al. 2002, 2008). We speculate that this shifted the

region of maximum precipitation west of the true crest. However, we cannot conclusively say that the PRISM values are too low because simulations are not outside our models' 95% confidence interval. In contrast, we note that the WRF_{MPP} simulations of SWE at APE, APW, and WLL were biased high by about 8% on average, suggesting the 1.33-km WRF grid spacing was sufficient to resolve the relatively narrow valley.

WRF_{MPP}, WRF_{Full}, and WRF_{LP} showed differences in the directionality of their errors at Mount Seattle West (MSW) and Mount Seattle East (MSE). Both of these sites (~ 950 m apart) had the same nearest WRF grid cell. However, both sites accumulated different amounts of snow (Figs. 6, 7), which suggests that true snowfall exhibits variability within a WRF grid cell, and that the overaccumulation in WRF_{MPP} simulations at MSE (less observed

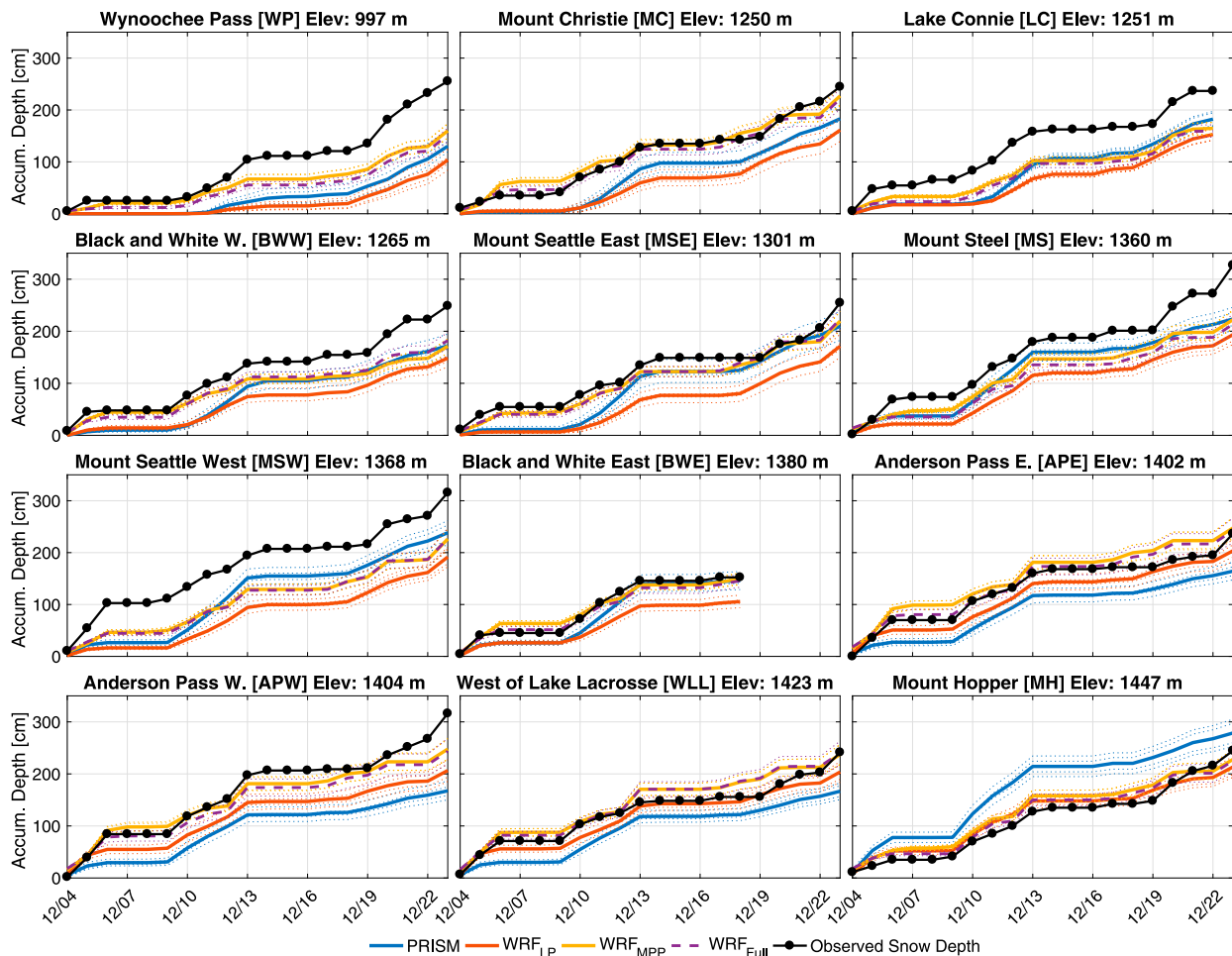


FIG. 7. Accumulated snow depth between 4 and 23 Dec 2015. Individual ensemble members are shown in thin dotted lines. Black and White East and Lake Connie have the evaluations end on 18 and 22 Dec 2015, respectively. After these dates the poles became significantly bent toward the camera or buried.

snowfall) is balanced out by the underaccumulation at MSW (more observed snowfall). WRF_{MPP} accurately simulated average snowfall within the area of this WRF grid cell, which is reflected by comparing the simulated SWE to the median value from a 1.33-km spatial domain at both MSE and MSW (Fig. 9). This highlights the complexity of evaluating gridded precipitation sources at larger spatial resolutions to smaller domains (point or 60-m area). See discussion for more information.

6. Discussion

a. Spatial representativeness of lidar observations to PRISM and WRF grid cells

There is a significant range in ASO-derived SWE within a spatial scale of 60, 800 (PRISM grid spacing), and 1333 m (WRF grid spacing; Fig. 9). In

general, the SWE distributions were similar across spatial scales. However, the median values were not always consistent among different spatial areas. We found the median SWE value from a 60-m bounding box was generally higher by about 20–60 cm when compared to the median value from an 800-m bounding box. Similarly, we found that the median elevation within a 60-m bounding box was around 40–80 m higher than the elevation within an 800-m bounding box. When we used the observed lapse rate and the calculated sensitivity of rain versus snow to temperature (section 6b), we could not fully explain the differences in median snow depth values across spatial scales based on changes in median elevation alone. We therefore hypothesize that complex interactions between wind, terrain, and vegetation are responsible for these differences in median values at different spatial scales rather than solely

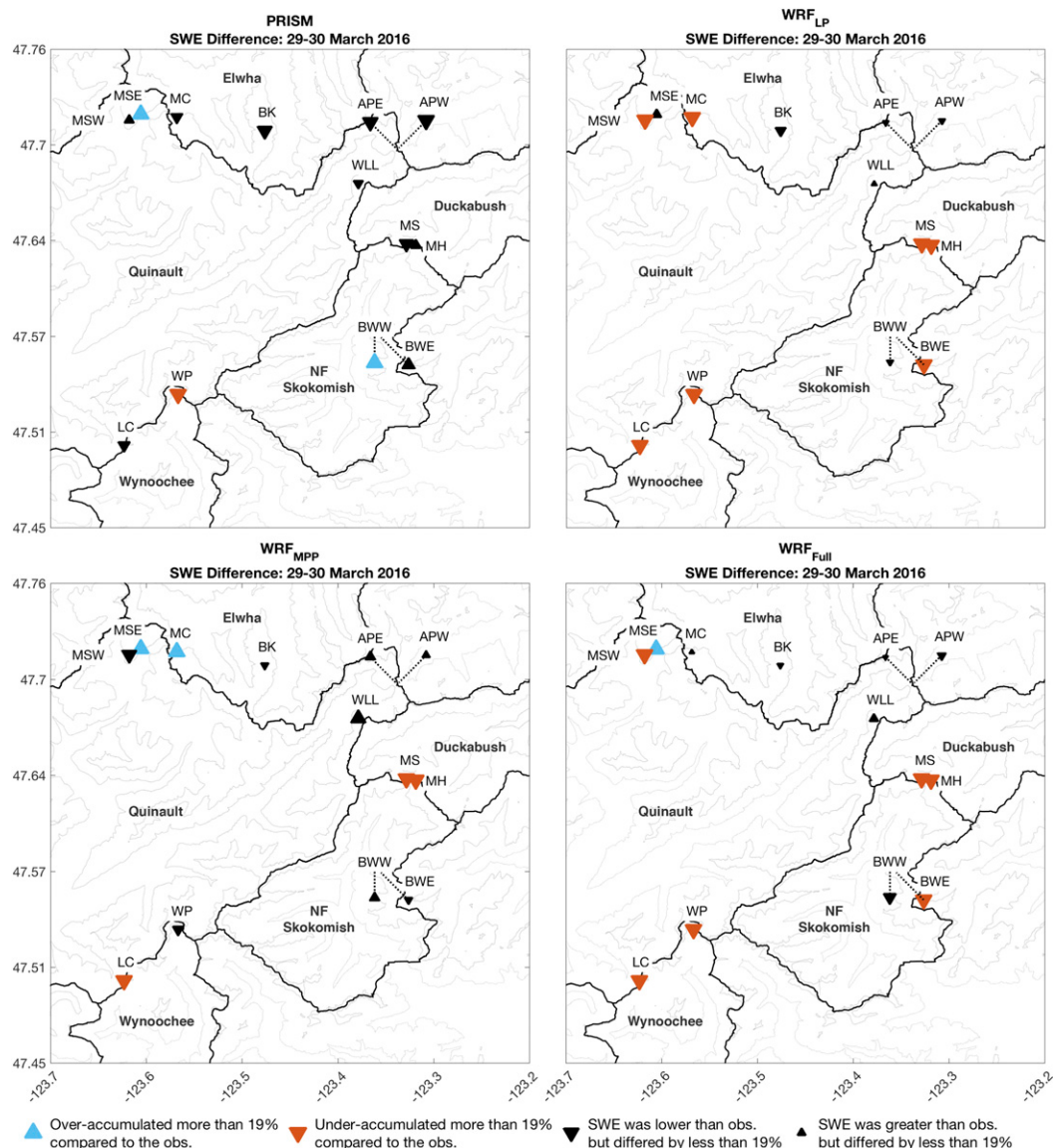


FIG. 8. Percent errors of modeled SWE compared to the median value from the ASO 60-m bounding boxes at independent snow monitoring sites. Black triangles are scaled based on how much they differ from the observations. Thick black lines show watershed delineations based on USGS streamflow gauges. Thin lines represent 600-m contour intervals.

the sensitivity of rain-versus-snow partitioning to elevation and temperature.

Lake Connie (LC), which sits near the top of a ridgeline, had the most dramatic difference between the median ASO-derived SWE values (105 cm) and median elevation (89 m) when looking at different spatial scales. The distributions at LC show that both WRF_{MPP} and WRF_{LP} were more reflective of the median SWE value from the larger domains (Fig. 9), and therefore, both WRF_{LP} and WRF_{MPP} may not be underaccumulating at LC (Figs. 4, 5, 8). In contrast, simulations of SWE that used PRISM-derived

precipitation agreed with the observed SWE from the median 60-m bounding box but overaccumulated when compared to the median values from the larger bounding boxes.

Despite the complexity of evaluating gridded precipitation products to both point and larger spatial areas, we highlight that at most locations there was little change in the evaluation of PRISM, WRF_{MPP}, WRF_{Full}, and WRF_{LP} at individual sites. Furthermore, there was little change in the mean differences and mean absolute differences when moving from the 60-m bounding box to a larger spatial area (Table 6).

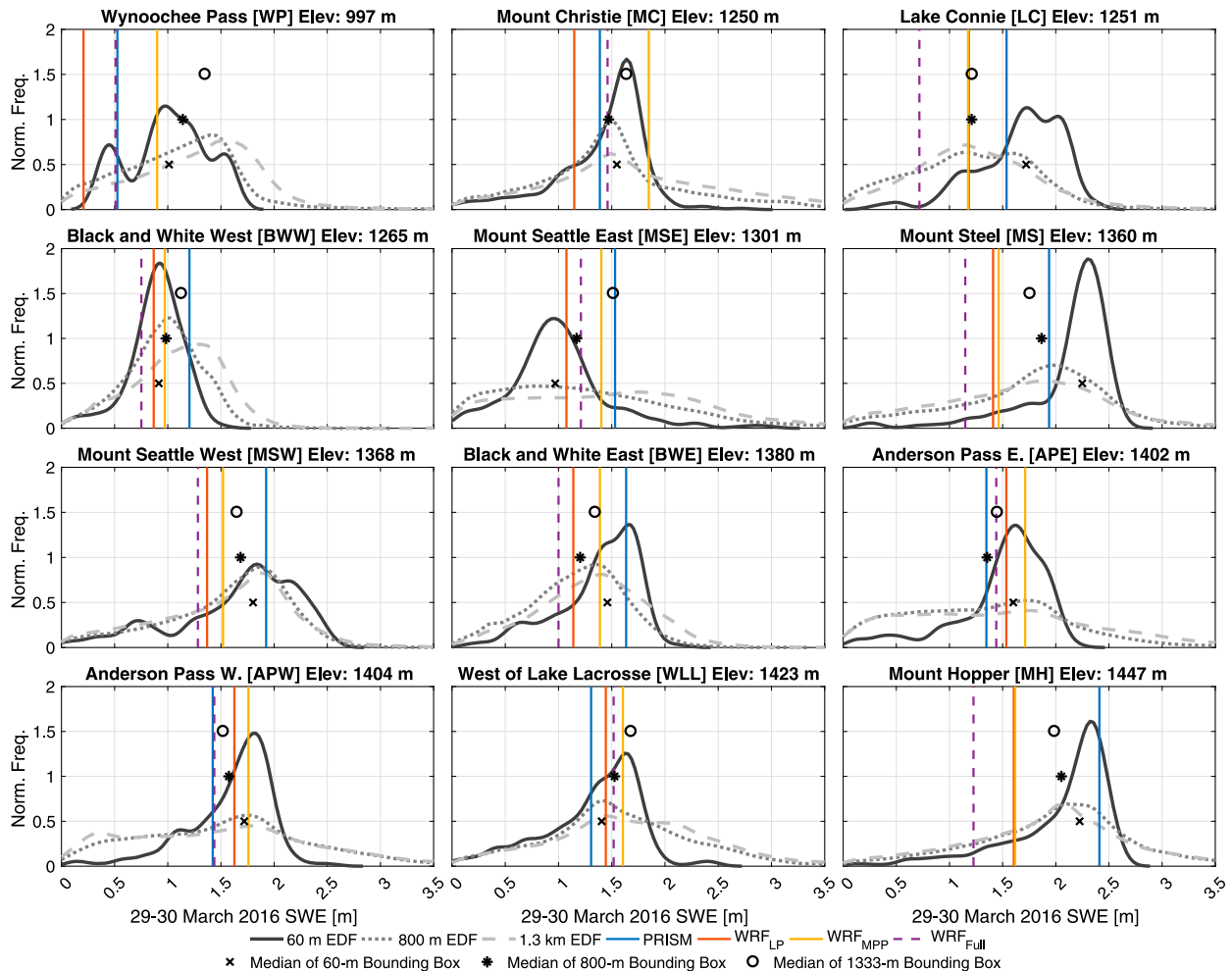


FIG. 9. Normalized empirical distribution functions (EDF) showing the distribution of SWE within different spatial areas surrounding the location of the snow monitoring station. Median SWE values from different spatial areas are independent of the y axis and are shown at various vertical positions to avoid overlap. We note that at Mount Hopper and Lake Connie, WRF_{MPP} (gold) and WRF_{LP} (red) overlap.

b. Rain-versus-snow sensitivity

There were significant differences in model performance based on the method used for rain-versus-snow partitioning. WRF_{MPP} indicated a promising path forward in hydrology, as results were as good as or better than the PRISM simulations, and WRF_{MPP} was able to simulate changes in rain versus snow at an hourly time step and was not dependent on using the calibrated linear temperature threshold for a nonlinear process. For example, using the output from the microphysical scheme, we back calculated the T_{crit} model parameter by finding the T_w during events that had a rain fraction between 20% and 80%. Parameter T_w was calculated using the observed temperature and RH along with WRF's pressure (Iribarne and Godson 1981). The

back-calculated T_{crit} at the snow monitoring stations was normally distributed around 0.1°C with a 95% confidence interval that ranged from -2.6° to 2.9°C . Parameter T_{crit} was also dynamic, in that it changed from hour to hour within a storm. Similar results were found when we constrained the rain fraction to 30%–70% and when we used WRF's temperature and specific humidity to calculate T_w .

Using a traditional calibration approach for T_{crit} (section 4c), we used a constant T_{crit} value of -1.0°C at all locations for both PRISM and WRF_{LP} simulations. This method was dependent on uniformly distributed daily total precipitation observations because the calibration locations only contained daily observations of precipitation. When the

TABLE 6. Mean differences and mean absolute differences of SWE between different model simulations when using median values of observed SWE (derived from ASO snow depth and density observations) from different spatial areas.

Model evaluation area	PRISM [cm (%)]	WRF _{LP} [cm (%)]	WRF _{MPP} [cm (%)]	WRF _{Full} [cm (%)]
Mean difference				
60 m	−4 (−1)	−33 (−21)	−9 (−2)	−35 (−19)
800 m	8 (6)	−22 (−15)	2 (3)	−23 (−16)
1333 m	−1 (−1)	−30 (−20)	−6 (−4)	−32 (−21)
Mean absolute difference				
60 m	26 (20)	36 (23)	30 (18)	42 (26)
800 m	23 (20)	26 (18)	26 (15)	30 (21)
1333 m	27 (18)	34 (22)	19 (13)	34 (23)

calibrated linear temperature threshold was run with hourly WRF precipitation and compared to the observations of SWE at the snow monitoring site, we found a systematic and more significant low bias in SWE (−33%), highlighting that the temperature-based partitioning method is dependent on the temporal resolution of the precipitation observations used to calibrate the model (Harder and Pomeroy 2013; Wayand et al. 2016a).

Furthermore, we found that the temperature-based partitioning method, when combined with the PRISM-derived estimates of precipitation, could provide an unbiased estimate of SWE on an annual basis because the PRISM simulations used the model calibration data for partitioning. On average, 79% of the weight in the distribution of PRISM-derived precipitation came from the four SNOTEL sites. However, the temperature-based partitioning method exhibited errors in identifying the phase of the precipitation during individual events, as the calibration method had compensating errors that provided an unbiased estimate of SWE [section 5a(2)(i)]. Therefore, PRISM simulations had an advantage over the WRF_{LP} simulations because WRF_{LP} simulations used independent estimates of precipitation. Therefore, any differences in the timing of precipitation could lead to errors in the partitioning.

Furthermore, we note that T_{crit} is a highly sensitive model parameter in simulating SWE in this region. We found that a 1°C (from −1.4° to 0.4°C) change in the T_{crit} model parameter resulted in a difference at the Buckinghorse SNOTEL site in over a half a meter (52 cm) of SWE. A 0.2°C change in T_{crit} changed peak SWE by 2%–8%. Despite the need for calibration data to determine this model parameter, the sensitivity is also important to consider when using a distributed hydrologic model. For instance, using temperature sensors that were deployed at elevations ranging from 180 to 1458 m, we determined the mean lapse rate to be $-4.5^{\circ}\text{C km}^{-1}$ during precipitation events. This is consistent with previous work (Minder et al. 2010) that

showed lapse rates within the Pacific Northwest are considerably less than the $-6.5^{\circ}\text{C km}^{-1}$ often assumed in hydrologic modeling studies (e.g., Livneh et al. 2013). This $-4.5^{\circ}\text{C km}^{-1}$ lapse rate, in conjunction with the rain-versus-snow sensitivity, shows that an elevation change of 100 m could result in a 4%–16% change in modeled SWE for that elevation band.

WRF does not require a known lapse rate, and the microphysical partitioning of rain versus snow does not require regional calibration. Therefore, we recommend that future snow model development in maritime environments focus on assessing the best microphysical scheme rather than temperature-based partitioning methods. Here we only evaluated the Thompson et al. (2008) microphysical scheme, but it is promising that in a comparison of microphysical schemes, the Thompson scheme was found to perform best in predicting snow across the cold continental climate of Colorado, United States (Liu et al. 2011). Furthermore, this microphysical scheme and model setup helped improve partitioning rain and snow at Snoqualmie Pass, Washington, especially during cold air intrusions (Wayand et al. 2016a). We recommend that the snow depth data presented herein be used along with other OLYMPLEX observations to evaluate how WRF performs with different physics, and to improve microphysical schemes, as previous work has shown that different microphysical schemes show substantial variability in the elevation at which snow transitions to rain (Frick et al. 2013; Minder and Kingsmill 2013).

c. Sensitivity to the weighting of gauges with PRISM

We tested our decision to use IDW over IDSW to distribute precipitation with the PRISM annual climatology. In this manuscript, we showed results from using IDW, as it resulted in simulations with the lowest mean percent difference in snow depth (5%) and SWE (−1%). We found that using IDSW resulted in an increase in the simulated snow depth by 4% and SWE by 6% on average. However, southern sites (Wynoochee Pass, Lake Connie, Black and White East, Black and White West)

were the least affected by the weighting method, differing by about 0%–2% in SWE. When IDSW was used, Buckinghamhorse shifted from having an average weight of 26%–46%, while the sites that were largely unaffected by the weighting scheme had little change in their weighting (~5%). Sites that had a significant difference based on how we weighted each gauge were less than 10 km away from the Buckinghamhorse SNOTEL site, while the sites with little difference were greater than 20 km away. When the Buckinghamhorse site was removed from the estimates of precipitation with PRISM (using IDW), the simulation of snow depth and SWE decreased on average by 7% and 9%, respectively. Since results will vary depending on the gauge distribution and where precipitation is being predicted, we encourage future PRISM users to explore the effect of their weighting scheme.

Furthermore, we also examined using the SNOTEL sites in combination with only the low-elevation RAWs and USCRN precipitation gauges because RAWs gauges use unheated tipping buckets and are subject to freezing. We found that the mean percent difference in SWE increased from –1% to 6%. Again, because of the gauge distribution, this decision gave more weight to the Buckinghamhorse gauge, providing similar results to using IDSW.

d. Model sensitivity unique to warm maritime snow environments

Throughout this study many nontrivial model-forcing decisions were made to improve model skill. One of the most sensitive modeling decisions was our choice of longwave radiation. We found significant differences in model performance based on the empirical method chosen for longwave radiation. Of all possible empirical longwave decisions from Flerchinger et al. (2009), the Dilley and O'Brien (1998) clear sky method with the Unsworth and Monteith (1975) cloud correction method performed best at Snoqualmie Pass and was also one of the most transferrable equations in Flerchinger et al. (2009). Other suggested transferrable longwave parameterizations simulated either too much or too little incoming longwave radiation, corresponding to too much or not enough melt in our model simulations. Furthermore, we evaluated WRF incoming radiation at the Snoqualmie Pass energy balance tower and found biases in WRF shortwave and longwave radiation to be greater than empirical estimates and to be in opposite directions. See appendix B for more details.

Another nontrivial decision was the choice in wind speed. In this region, we found modeled turbulent fluxes to be a significant energy input to the snowpack, because throughout the season both the latent and sensible heat fluxes could be oriented in the same direction.

Therefore, we found that using the only wind speed measurement within our study domain (Waterhole SNOTEL), despite being located far away from snow monitoring sites, offered better model performance than using wind speed from WRF.

WRF's 10-m wind speed at Waterhole was on average 4 m s^{-1} greater than the Waterhole observations. Therefore, we found that the overall low bias with WRF_{Full} compared to WRF_{MPP} was the result of too much melt from turbulent heat fluxes directed toward the snowpack rather than errors with incoming radiation (which tended to offset) or with precipitation and partitioning. Independent simulations (not shown) illustrated that when air temperature, specific humidity, and shortwave and longwave radiation were provided from WRF, but wind speed was the same as in the WRF_{MPP} simulations (Waterhole observed wind speed), the overall mean percent difference in SWE was small, –2% compared to –7%. When the model was run with all meteorological forcing data (WRF_{Full}), including WRF's wind speed, the mean percent difference in near-peak SWE was significantly more biased, changing from –2% (WRF_{MPP}) to –19%. We recommend other models set up in similar environments pay close attention to choices in incoming longwave radiation and wind speed.

7. Conclusions

When output from WRF's microphysical scheme was used to partition WRF precipitation into rain versus snow (WRF_{MPP}), simulations of snow depth and SWE were relatively unbiased. However, errors in near-peak SWE were outside the model's 95% confidence interval, at 5 out of 12 sites, exhibiting that the unbiased mean difference was the result of compensating errors. Simulations that used PRISM-derived precipitation with the linear partitioning method also had compensating errors leading to an overall unbiased mean difference. However, PRISM simulations benefited from using SNOTEL precipitation data, which were used to calibrate the temperature-based partitioning method. WRF_{LP}, which also used the calibrated partitioning scheme but had independent estimates of precipitation, was biased low. Since WRF_{MPP} was relatively unbiased, this suggests that WRF_{LP} was biased low due to partitioning errors, rather than errors with the total precipitation from WRF. Furthermore, the temperature-based threshold also exhibited partitioning errors during individual snowfall events with both WRF_{LP} and PRISM-based simulations.

When WRF used all meteorological data to run the hydrologic model (WRF_{Full}), we found that WRF's wind speeds, rather than radiation, precipitation, or partitioning errors, resulted in too much midwinter melt and

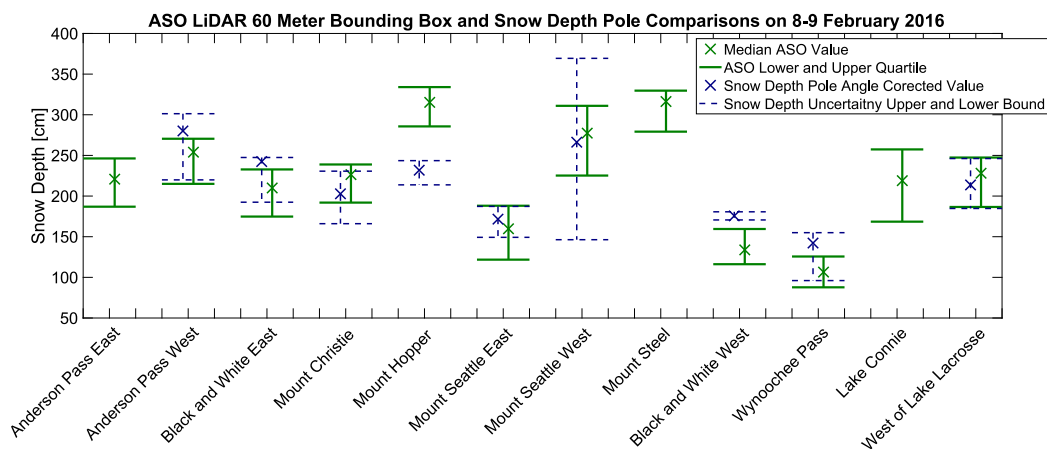


FIG. A1. The ASO lidar distribution of snow depth values within a 60-m bounding box and the snow depth pole measurements, with their associated uncertainty, on 8–9 Feb 2016. Sites that do not have blue bars (dotted lines) shown are sites with no snow depth measurement at this time because the poles were buried under snow.

an overall low bias (−19%). Therefore, the best simulations of snow depth and SWE with WRF precipitation resulted from using nearby observations of wind speed with a partitioning method based on output from the microphysical scheme (WRF_{MPP}).

Microphysical schemes in atmospheric models are an active area of research (Jankov et al. 2009; Liu et al. 2011; Minder and Kingsmill 2013), but using the microphysical scheme output from WRF for precipitation and to partition rain versus snow is an attractive path going forward in hydrology for four reasons: 1) the rain-versus-snow threshold T_{crit} , which is a highly sensitive model parameter in this environment does not have to be known or calibrated; 2) the microphysical scheme provides a more realistic approach to simulating the precipitation phase than simple temperature thresholds; 3) the lapse rate during precipitation events ($-4.5^{\circ}\text{C km}^{-1}$) differs from the commonly assumed $-6.5^{\circ}\text{C km}^{-1}$ and does not have to be known a priori; and 4) WRF precipitation and rain-versus-snow partitioning can be derived anywhere, even in watersheds with no observations.

Acknowledgments. We gratefully acknowledge funding support from NSF (EAR-1215771) and NASA (Grant NNX13AO58G and NNX14AJ72G). We also thank Olympic National Park for providing us with the permission to install snow depth poles within park boundaries, and Bill Baccus for obtaining manual snow course observations. Furthermore, we thank Clifford Mass and Neal Johnson for providing access to the WRF data archive, and Tom Painter and Kathryn Bormann for providing ASO data. We also

thank Derek Beal, Colin Butler, Max Mozer, Justin Pflug, Adam Massmann, and Brad Gaylor for the tremendous effort they gave to help install the snow depth monitoring sites in remote regions of Olympic National Park and with help processing these data. We thank Justin Minder for providing Quillayute Sounding analysis, Joe Zagrodnik for analysis of the primary storm directions during December snowfall events, Nicholas Wayand for his help with using WRF data and setting up SUMMA, Bart Nijssen for advice on an early version of this manuscript, and the remainder of the Mountain Hydrology Research Group for helpful feedback and support. Lastly, we thank three anonymous reviewers for helpful commentary that improved this manuscript.

All meteorological and snow depth data collected during the OLYMPEX campaign are archived at the Global Hydrology Resource Center Distributed Active Archive Center (GHRC DAAC) and are publicly available. SUMMA model code is available at <https://github.com/NCAR/summa/> along with more information at <http://www.ral.ucar.edu/projects/summa>. A description of how to derive shortwave radiation data using MTCLIM can be found at <https://vic.readthedocs.io/en/vic.4.2.c/Documentation/ForcingData/>. The PRISM 800 meter, 30-year (1981–2010) annual and monthly climate normals were downloaded from <http://www.prism.oregonstate.edu/normals/>. The RAWS precipitation data were downloaded from <http://www.raws.dri.edu/>. All NRCS SNOTEL data were downloaded from <http://www.wcc.nrcs.usda.gov/snow/>. Quinault USCRN data were downloaded from <ftp://ftp.ncdc.noaa.gov/pub/data/uscrn/products/hourly02/>.

APPENDIX A

Lidar Snow Depth Comparison to Time-Lapse Snow Depth

We compared ASO snow depth values from 60-m bounding boxes around our snow depth pole measurements. We found that both measurements generally agreed or were within each other's interquartile (ASO) or uncertainty range (snow depth poles; Fig. A1). At many sites, such as Mount Seattle East, Mount Seattle West, and West of Lake Lacrosse, the difference from the snow depth pole and the median ASO value was less than 15 cm. At other sites, more significant differences appeared. For instance, at Mount Hopper, camera images showed a significant snowdrift, which was formed as a result of preferential deposition of precipitation. Since the snow depth poles were located outside the snowdrift, the median ASO value was higher than the snow pole measurement by 84 cm in February and 173 cm in March. In contrast to this, at Black and White West the ASO snow depth maps indicated that our snow depth poles were located within a snowdrift, and therefore the median ASO value was 42 cm lower in February and 60 cm lower in March when compared to the snow depth pole measurement.

At these two sites, we found that the median ASO value from a larger spatial area (>60 -m bounding box) was more similar to the median ASO value from the 60-m bounding box than it was to the snow depth pole measurement. This indicated to us that the median ASO value within a 60-m bounding box was a better representation of the snow within this region than the snow depth poles alone. Furthermore, by March many snow depth poles contained significant uncertainty in their measurements because they became bent or buried with snow. Therefore, we chose to compare our model simulations with different sources of precipitation, on an annual basis, to the median March ASO value within a 60-m bounding box, rather than directly with the snow depth pole measurements.

APPENDIX B

Model Forcing Variables

PRISM, WRF_{LP}, and WRF_{MPP} simulations all use specific humidity q from WRF and from observations of RH (converted to q using observed temperature and WRF pressure), along with different density decisions, to form a snow depth/SWE ensemble. The specific humidity decisions in the PRISM, WRF_{LP},

TABLE B1. SUMMA model forcing data used in the calibration and precipitation evaluation phase. The dash indicates that the source of the model forcing data is the same as that in the column to the left.

Forcing variable	Sources of model forcing data				
	SNOTEL calibration	PRISM	WRF _{LP}	WRF _{MPP}	WRF _{full}
Precipitation ($\text{kg m}^{-2} \text{s}^{-1}$)	Uniformly distributed, observed, daily values	Uniformly distributed, observed, daily values using PRISM weights	WRF precipitation aggregated to hourly and then uniformly distributed	Hourly precipitation from WRF	Hourly precipitation from WRF
Temperature ($^{\circ}\text{K}$)	Observed, corrected for using (1)	Observed	—	—	WRF 2-m temperature
Specific humidity q (g g^{-1})	Observations and WRF	—	—	—	WRF 2-m specific humidity
Shortwave radiation (SW) (W m^{-2})	MTCLIM v. 4.2 $f(T_{\text{obs}}, \text{RH}_{\text{obs}}, \text{WRF}, \text{Ppt}_{\text{obs}})$	—	—	—	WRF
Longwave radiation (W m^{-2})	Dilley and O'Brien (1998) with Unsworth and Monteith (1975) cloud correction $f(T_{\text{obs}}, \text{RH}_{\text{obs}}, \text{WRF}, \text{SW})$	—	—	—	WRF
Pressure (Pa)	WRF	—	—	—	—
Wind speed (m s^{-1})	Waterhole observed hourly averaged wind speed	—	—	—	WRF

TABLE C1. Adjusted model parameter values, from default (see supplemental material). These model parameter values are held constant in each run. The albedo decay rate was fit to observations of albedo at Snoqualmie Pass. The default maximum snow albedo parameter value (0.84) fit the observations at Snoqualmie Pass. Vegetation parameters were adjusted to simulate an open, nonvegetated area.

Model parameter name	Model parameter description	Adjusted value
tempCritRain T_{crit} (°C)	Critical temperature where precipitation is rain/snow	−1.0
albedoDecayRate (s)	Albedo decay rate	500 000
heightCanopyTop (m)	Height of top of the vegetation canopy above the ground surface	0.01
heightCanopyBottom (m)	Height of bottom of the vegetation canopy above the ground surface	0.001
winterSAI ($\text{m}^2 \text{m}^{-2}$)	Stem area index prior to the start of the growing season	0.01
summerLAI ($\text{m}^2 \text{m}^{-2}$)	Maximum leaf area index at the peak of the growing season	0.5

and WRF_{MPP} model configurations were carried into the derivation of shortwave and longwave radiation. Precipitation was uniformly distributed over each day in both PRISM and WRF_{LP} to be consistent with the calibration methodology, which relied on daily resolution precipitation data. PRISM, WRF_{LP}, and WRF_{MPP} model simulations used empirical derivations for shortwave and longwave radiation, as these performed better than WRF at the nearby Snoqualmie Pass, Washington, energy balance tower. For instance, WRF was biased $+60 \text{ W m}^{-2}$ in shortwave radiation, while the mountain microclimate simulation model (MTCLIM Thornton and Running 1999; Bohn et al. 2013) was biased only $+23 \text{ W m}^{-2}$. These results were consistent with Lapo et al. (2017), which showed that WRF was biased higher than MTCLIM over sites in California. Similarly, WRF was biased greater in longwave radiation (-17 W m^{-2}) than our empirical derivation (-7 W m^{-2}); however, these biases were in the opposite direction of the shortwave radiation biases. Furthermore, wind speed was generally taken from the nearest observational site (Waterhole SNOTEL site), except when wind speed was taken from WRF (WRF_{Full}). In WRF_{Full} we scaled the wind speeds U using the

Prandtl–von Kármán universal velocity distribution equation (Dingman 2008):

$$U = \frac{1}{k} \times U^* \times \ln\left(\frac{z - z_d}{z_o}\right), \quad (\text{B1})$$

where k is the von Kármán constant, U^* is the frictional velocity taken from WRF, z is the measurement height at which the hydrologic model has forcing data prescribed, z_o is the time-varying surface roughness length taken from WRF, and z_d is the zero-plane displacement height. In this manuscript, we took z_d to be the simulated snow depth from WRF's embedded land surface model (Noah).

All model forcing data for SUMMA is summarized in Table B1.

APPENDIX C

Adjusted Model Parameters

Table C1 shows adjusted model parameter values. Table C2 shows adjusted new snow density and compaction model parameter values.

TABLE C2. Adjusted new snow density and compaction model parameter values. These model parameter values vary within each ensemble, along with different sources of q for the PRISM, WRF_{LP}, WRF_{MPP}, and model calibration simulations. In WRF_{Full} these snow density model parameters also vary but q is only taken from WRF.

Model parameter	Value	
	Wayand et al. (2016b) (ρ_w)	Hedstrom and Pomeroy (1998) and Anderson (1976) (ρ_{HPA})
newSnowDenMin (minimum new snow density; kg m^{-3})	100	67.92
newSnowDenMult (multiplier for new snow density; kg m^{-3})	50	51.25
newSnowDenScal (scaling factor for new snow density; K)	1	2.59
denScalOvrbdn (density scaling factor for overburden pressure; $\text{kg}^{-1} \text{m}^3$)	0.02	0.023
tempScalOvrbdn (temperature scaling factor for overburden pressure; K^{-1})	0.06	0.08

REFERENCES

- Anders, A. M., G. H. Roe, D. R. Durran, and J. M. Minder, 2007: Small-scale spatial gradients in climatological precipitation on the Olympic Peninsula. *J. Hydrometeorol.*, **8**, 1068–1081, doi:10.1175/JHM610.1.
- Anderson, E. A., 1976: A point energy and mass balance model of a snow cover. NOAA Tech. Rep. NWS 19, 150 pp., http://amazon.nws.noaa.gov/articles/HRL_Pubs_PDF_May12_2009/HRL_PUBS_51-100/81_A_POINT_ENERGY_AND_MASS.pdf.
- Bohn, T. J., B. Livneh, J. W. Oyster, S. W. Running, B. Nijssen, and D. P. Lettenmaier, 2013: Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. *Agric. For. Meteorol.*, **176**, 38–49, doi:10.1016/j.agrformet.2013.03.003.
- Clark, M. P., and Coauthors, 2015a: A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resour. Res.*, **51**, 2498–2514, doi:10.1002/2015WR017198.
- , and Coauthors, 2015b: A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resour. Res.*, **51**, 2515–2542, doi:10.1002/2015WR017200.
- , and Coauthors, 2015c: The structure for unifying multiple modeling alternatives (SUMMA), version 1.0: Technical description. NCAR Tech. Note NCAR/TN-514+STR, 50 pp., doi:10.5065/D6WQ01TD.
- Colle, B. A., and C. F. Mass, 2000: The 5–9 February 1996 flooding event over the Pacific Northwest: Sensitivity studies and evaluation of the MM5 precipitation forecasts. *Mon. Wea. Rev.*, **128**, 593–617, doi:10.1175/1520-0493(2000)128<0593:TFFEOT>2.0.CO;2.
- , K. J. Westrick, and C. F. Mass, 1999: Evaluation of MM5 and Eta-10 precipitation forecasts over the Pacific Northwest during the cool season. *Wea. Forecasting*, **14**, 137–154, doi:10.1175/1520-0434(1999)014<0137:EOMAEP>2.0.CO;2.
- Currier, W. R., 2016: An independent evaluation of frozen precipitation from the WRF model and PRISM in the Olympic Mountains for WY 2015 and 2016. Master's thesis, Department of Civil and Environmental Engineering, University of Washington, 58 pp., <https://digital.lib.washington.edu/researchworks/handle/1773/38604>.
- Daly, C., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris, 2002: A knowledge-based approach to the statistical mapping of climate. *Climate Res.*, **22**, 99–113, doi:10.3354/cr022099.
- , M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.*, **28**, 2031–2064, doi:10.1002/joc.1688.
- Dilley, A. C., and D. M. O'Brien, 1998: Estimating downward clear sky long-wave irradiance at the surface from screen temperature and precipitable water. *Quart. J. Roy. Meteor. Soc.*, **124**, 1391–1401, doi:10.1002/qj.49712454903.
- Dingman, S. L., 2008: *Physical Hydrology*. Waveland Press, 646 pp.
- Flerchinger, G. N., W. Xao, D. Marks, T. J. Sauer, and Q. Yu, 2009: Comparison of algorithms for incoming atmospheric long-wave radiation. *Water Resour. Res.*, **45**, W03423, doi:10.1029/2008WR007394.
- Frick, C., A. Seifert, and H. Wernli, 2013: A bulk parametrization of melting snowflakes with explicit liquid water fraction for the COSMO model. *Geosci. Model Dev.*, **6**, 1925–1939, doi:10.5194/gmd-6-1925-2013.
- Goodison, B. E., P. Y. T. Louie, and D. Yang, 1998: WMO solid precipitation intercomparison. Instruments and Observing Methods Rep. 67, WMO/TD-872, 212 pp., <https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-67-solid-precip/WMOtd872.pdf>.
- Gutmann, E. D., R. M. Rasmussen, C. Liu, K. Ikeda, D. J. Gochis, M. P. Clark, J. Dudhia, and G. Thompson, 2012: A comparison of statistical and dynamical downscaling of winter precipitation over complex terrain. *J. Climate*, **25**, 262–281, doi:10.1175/2011JCLI4109.1.
- Hamlet, A. F., and D. P. Lettenmaier, 2005: Production of temporally consistent gridded precipitation and temperature fields for the continental United States. *J. Hydrometeorol.*, **6**, 330–336, doi:10.1175/JHM420.1.
- , and Coauthors, 2010: Final report for the Columbia Basin Climate Change Scenarios Project. PNW Hydroclimate Scenarios Project 2860, Climate Impacts Group, University of Washington, <http://warm.atmos.washington.edu/2860/report/>.
- Harder, P., and J. Pomeroy, 2013: Estimating precipitation phase using a psychrometric energy balance method. *Hydrol. Processes*, **27**, 1901–1914, doi:10.1002/hyp.9799.
- Hedstrom, N. R., and J. W. Pomeroy, 1998: Measurements and modelling of snow interception in the boreal forest. *Hydrol. Processes*, **12**, 1611–1625, doi:10.1002/(SICI)1099-1085(199808/09)12:10<1611::AID-HYP684>3.0.CO;2-4.
- Henn, B., M. P. Clark, D. Kavetski, A. J. Newman, M. Hughes, B. McGurk, and J. D. Lundquist, 2016: Spatiotemporal patterns of precipitation inferred from streamflow observations across the Sierra Nevada mountain range. *J. Hydrol.*, doi:10.1016/j.jhydrol.2016.08.009, in press.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis, 2005: Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.*, **25**, 1965–1978, doi:10.1002/joc.1276.
- Houze, R. A., Jr., and Coauthors, 2017: The Olympic Mountains Experiment (OLYMPEX). *Bull. Amer. Meteor. Soc.*, doi:10.1175/BAMS-D-16-0182.1, in press.
- Iribarne, J. V., and W. L. Godson, 1981: *Atmospheric Thermodynamics*. D. Reidel, 259 pp.
- Jankov, I., J. W. Bao, P. J. Neiman, P. J. Schultz, H. L. Yuan, and A. B. White, 2009: Evaluation and comparison of microphysical algorithms in ARW-WRF model simulations of atmospheric river events affecting the California coast. *J. Hydrometeorol.*, **10**, 847–870, doi:10.1175/2009JHM1059.1.
- Judson, A., and N. Doesken, 2000: Density of freshly fallen snow in the central Rocky Mountains. *Bull. Amer. Meteor. Soc.*, **81**, 1577–1587, doi:10.1175/1520-0477(2000)081<1577:DOFFSI>2.3.CO;2.
- Julander, R. P., J. Curtis, and A. Beard, 2007: The SNOTEL temperature dataset. *Mountain Views Newsletter*, Vol. 1, No. 2, CIRMOUNT, USDA Forest Service, Albany, CA, 4–7, https://www.fs.fed.us/psw/cirmount/publications/pdf/Mtn_Views_aug_07.pdf.
- LaChapelle, E. R., 1958: Winter snow observation at Mt. Olympus. *Proc. 26th Annual Western Snow Conf.*, Bozeman, MT, Western Snow Conference, 59–63, <https://westernsnowconference.org/node/1174>.
- Lapo, K. E., L. M. Hinkelman, E. Sumargo, M. Hughes, and J. D. Lundquist, 2017: A critical evaluation of modeled solar irradiance over California for hydrologic and land-surface modeling. *J. Geophys. Res. Atmos.*, **121**, 299–317, doi:10.1002/2016JD025527.
- Liu, C., K. Ikeda, G. Thompson, R. Rasmussen, and J. Dudhia, 2011: High-resolution simulations of wintertime precipitation in the Colorado Headwaters Region: Sensitivity to physics parameterizations. *Mon. Wea. Rev.*, **139**, 3533–3553, doi:10.1175/MWR-D-11-00009.1.

- Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. M. Andreadis, E. P. Maurer, and D. P. Lettenmaier, 2013: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *J. Climate*, **26**, 9384–9392, doi:[10.1175/JCLI-D-12-00508.1](https://doi.org/10.1175/JCLI-D-12-00508.1).
- , J. S. Deems, D. Schneider, J. Barsugli, and N. Molotch, 2014: Filling in the gaps: Inferring spatially distributed precipitation from gauge observations over complex terrain. *Water Resour. Res.*, **50**, 8589–8610, doi:[10.1002/2014WR015442](https://doi.org/10.1002/2014WR015442).
- Lundquist, J. D., and B. Huggett, 2008: Evergreen trees as inexpensive radiation shields for temperature sensors. *Water Resour. Res.*, **44**, W00D04, doi:[10.1029/2008WR006979](https://doi.org/10.1029/2008WR006979).
- , D. Cayan, and M. Dettinger, 2003: Meteorology and hydrology in Yosemite National Park: A sensor network application. *Information Processing in Sensor Networks*, F. Zhao and L. Guibas, Eds., Lecture Notes in Computer Science, Vol. 2634, Springer, 518–528.
- , M. Hughes, B. Henn, E. D. Gutmann, B. Livneh, J. Dozier, and P. Neiman, 2015: High-elevation precipitation patterns: Using snow measurements to assess daily gridded datasets across the Sierra Nevada, California. *J. Hydrometeorol.*, **16**, 1773–1792, doi:[10.1175/JHM-D-15-0019.1](https://doi.org/10.1175/JHM-D-15-0019.1).
- Marks, D., A. Winstral, M. Reba, J. Pomeroy, and M. Kumar, 2013: An evaluation of methods for determining during-storm precipitation phase and the rain/snow transition elevation at the surface in a mountain basin. *Adv. Water Resour.*, **55**, 98–110, doi:[10.1016/j.advwatres.2012.11.012](https://doi.org/10.1016/j.advwatres.2012.11.012).
- Mass, C. F., and Coauthors, 2003: Regional environmental prediction over the Pacific Northwest. *Bull. Amer. Meteor. Soc.*, **84**, 1353–1366, doi:[10.1175/BAMS-84-10-1353](https://doi.org/10.1175/BAMS-84-10-1353).
- Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based data set of land surface fluxes and states for the conterminous United States. *J. Climate*, **15**, 3237–3251, doi:[10.1175/1520-0442\(2002\)015<3237:ALTHBD>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2).
- Minder, J. R., and D. E. Kingsmill, 2013: Mesoscale variations of the atmospheric snow line over the northern Sierra Nevada: Multiyear statistics, case study, and mechanisms. *J. Atmos. Sci.*, **70**, 916–938, doi:[10.1175/JAS-D-12-0194.1](https://doi.org/10.1175/JAS-D-12-0194.1).
- , D. R. Durran, G. H. Roe, and A. M. Anders, 2008: The climatology of small-scale orographic precipitation over the Olympic Mountains: Patterns and processes. *Quart. J. Roy. Meteor. Soc.*, **134**, 817–839, doi:[10.1002/qj.258](https://doi.org/10.1002/qj.258).
- , P. W. Mote, and J. D. Lundquist, 2010: Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains. *J. Geophys. Res.*, **115**, D14122, doi:[10.1029/2009JD013493](https://doi.org/10.1029/2009JD013493).
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, doi:[10.1029/97JD00237](https://doi.org/10.1029/97JD00237).
- Oyler, J. W., S. Z. Dobrowski, A. P. Ballantyne, A. E. Klene, and S. W. Running, 2015: Artificial amplification of warming trends across the mountains of the western United States. *Geophys. Res. Lett.*, **42**, 153–161, doi:[10.1002/2014GL062803](https://doi.org/10.1002/2014GL062803).
- Painter, T. H., and Coauthors, 2016: The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sens. Environ.*, **184**, 139–152, doi:[10.1016/j.rse.2016.06.018](https://doi.org/10.1016/j.rse.2016.06.018).
- Raleigh, M. S., J. D. Lundquist, and M. P. Clark, 2015: Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. *Hydrol. Earth Syst. Sci.*, **19**, 3153–3179, doi:[10.5194/hess-19-3153-2015](https://doi.org/10.5194/hess-19-3153-2015).
- Rasmussen, R., and Coauthors, 2012: How well are we measuring snow: The NOAA/FAA/NCAR winter precipitation test bed. *Bull. Amer. Meteor. Soc.*, **93**, 811–829, doi:[10.1175/BAMS-D-11-00052.1](https://doi.org/10.1175/BAMS-D-11-00052.1).
- Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., doi:[10.5065/D68S4MVH](https://doi.org/10.5065/D68S4MVH).
- Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519–542, doi:[10.1175/1520-0493\(2004\)132<0519:EFOWPU>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2).
- , P. R. Field, W. R. Hall, and R. M. Rasmussen, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, doi:[10.1175/2008MWR2387.1](https://doi.org/10.1175/2008MWR2387.1).
- Thornton, P. E., and S. W. Running, 1999: An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agric. For. Meteorol.*, **93**, 211–228, doi:[10.1016/S0168-1923\(98\)00126-9](https://doi.org/10.1016/S0168-1923(98)00126-9).
- , —, and M. A. White, 1997: Generating surfaces of daily meteorological variables over large regions of complex terrain. *J. Hydrol.*, **190**, 214–251, doi:[10.1016/S0022-1694\(96\)03128-9](https://doi.org/10.1016/S0022-1694(96)03128-9).
- Unsworth, M. H., and J. L. Monteith, 1975: Long-wave radiation at the ground. I. Angular distribution of incoming radiation. *Quart. J. Roy. Meteor. Soc.*, **101**, 13–24, doi:[10.1002/qj.49710142703](https://doi.org/10.1002/qj.49710142703).
- USACE, 1956: Snow hydrology: Summary report of the snow investigations. North Pacific Division, U.S. Army Corps of Engineers, 437 pp.
- Wayand, N. E., A. Massmann, C. Butler, E. Keenan, and J. D. Lundquist, 2015: A meteorological and snow observational data set from Snoqualmie Pass (921 m), Washington Cascades, USA. *Water Resour. Res.*, **51**, 10 092–10 103, doi:[10.1002/2015WR017773](https://doi.org/10.1002/2015WR017773).
- , J. Stimberis, J. P. Zagrodnik, C. F. Mass, and J. D. Lundquist, 2016a: Improving simulations of precipitation phase and snowpack at a site subject to cold air intrusions: Snoqualmie Pass, WA. *J. Geophys. Res. Atmos.*, **121**, 9929–9942, doi:[10.1002/2016JD025387](https://doi.org/10.1002/2016JD025387).
- , M. P. Clark, and J. D. Lundquist, 2016b: Diagnosing snow accumulation errors in a rain-snow transitional environment with snow board observations. *Hydrol. Processes*, **31**, 349–363, doi:[10.1002/hyp.11002](https://doi.org/10.1002/hyp.11002).
- Yang, H. W., B. Wang, and B. Wang, 2012: Reduction of systematic biases in regional climate downscaling through ensemble forcing. *Climate Dyn.*, **38**, 655–665, doi:[10.1007/s00382-011-1006-4](https://doi.org/10.1007/s00382-011-1006-4).