Approximating Input Data to a Snowmelt Model Using Weather Research and Forecasting Model Outputs in Lieu of Meteorological Measurements

SCOTT HAVENS AND DANNY MARKS

Northwest Watershed Research Center, Agricultural Research Service, USDA, Boise, Idaho

KATELYN FITZGERALD, MATT MASARIK, AND ALEJANDRO N. FLORES

Department of Geosciences, Boise State University, Boise, Idaho

PATRICK KORMOS

Colorado Basin River Forecast Center, National Weather Service, Salt Lake City, Utah

ANDREW HEDRICK

Northwest Watershed Research Center, Agricultural Research Service, USDA, and Department of Geosciences, Boise State University, Boise, Idaho

(Manuscript received 19 July 2019, in final form 26 February 2019)

ABSTRACT

Forecasting the timing and magnitude of snowmelt and runoff is critical to managing mountain water resources. Warming temperatures are increasing the rain–snow transition elevation and are limiting the forecasting skill of statistical models relating historical snow water equivalent to streamflow. While physically based methods are available, they require accurate estimations of the spatial and temporal distribution of meteorological variables in complex terrain. Across many mountainous areas, measurements of precipitation and other meteorological variables are limited to a few reference stations and are not adequate to resolve the complex interactions between topography and atmospheric flow. In this paper, we evaluate the ability of the Weather Research and Forecasting (WRF) Model to approximate the inputs required for a physics-based snow model, iSnobal, instead of using meteorological measurements, for the Boise River Basin (BRB) in Idaho, United States. An iSnobal simulation using station data from 40 locations in and around the BRB resulted in an average root-mean-square error (RMSE) of 4.5 mm compared with 12 SNOTEL measurements. Applying WRF forcings alone was associated with an RMSE of 10.5 mm, while including a simple bias correction to the WRF outputs of temperature and precipitation reduced the RMSE to 6.5 mm. The results highlight the utility of using WRF outputs as input to snowmelt models, as all required input variables are spatiotemporally complete. This will have important benefits in areas with sparse measurement networks and will aid snowmelt and runoff forecasting in mountainous basins.

1. Introduction

Rapidly warming climate in the western mountains of North America is changing the partitioning of rain and snow and increasing the rain–snow transition elevation (Nolin and Daly 2006; Nayak et al. 2010; Klos et al. 2014; Lute et al. 2015; Trujillo and Molotch 2014). Rain-on-snow flooding events are becoming more common (Surfleet and Tullos 2013; Tohver et al. 2014; Freudiger et al. 2014), can cause significant damage to downstream areas (Kattelmann 1997; Rössler et al. 2014; Pomeroy et al. 2016), and pose a significant challenge to reservoir planning and operations that are typically designed around the premise of a single, large spring melt pulse. The increasing frequency and uncertain timing of such occurrences reduces flexibility to retain or release reservoir water to optimize availability for agricultural use. On the opposite side of the spectrum, droughts require complex operations to ensure that the water releases are optimized to not waste the limited available water.

Traditional statistical forecasting methods are based on historical data relating streamflow and field measurements

Corresponding author: Scott Havens, scott.havens@ars.usda.gov

DOI: 10.1175/JHM-D-18-0146.1

of snow depth and snow water equivalent (SWE) at a relatively small number of reference sites in a given watershed. These historical models are unable to account for the impacts of winter rain or mixed phase events, suggesting the need for less statistical, and more mechanistic, physics-based snow and hydrologic forecasting models (Frei et al. 1999; Groisman et al. 2001, 2004, Mote 2003, 2006; Mote et al. 2005; Regonda et al. 2005; Pierce et al. 2008). Such models, however, require extensive meteorological forcing data, which are typically limited in remote mountain regions (Marks and Frew 1999; Garen and Marks 2005). Typical meteorological data networks in mountainous regions are relatively sparse, have variable periods of record, and contain frequent data gaps of varying length.

The Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008) has previously been used to dynamically downscale output from global weather forecasts, global climate models (GCMs), or reanalysis data with an initial spatial resolution of 50–200 km to a more management-relevant resolution of between 1 and 12 km. These dynamically downscaled WRF outputs are evenly distributed with relatively high spatial resolution (when ran at high resolution), and are temporally complete which makes them particularly efficient as forcing data for hydrologic modeling applications. In mountainous regions with complex topography and highly heterogeneous land cover, however, further downscaling may be required to produce WRF meteorological outputs at the finescale grid resolution (10–250 m) required to accurately capture small-scale snowpack variability in mountainous terrain (Winstral et al. 2014). It is critical in this process that we understand how to account for any scale differences between WRF output and local variation in slope, aspect, elevation and land cover.

Numerical weather prediction models are commonly used to provide meteorological inputs for hydrology models both for predicting streamflow from a short-term events, or for entire water years (from 1 October through 30 September). Miller and Kim (1996) investigated a 5-day flood event on the Russian River in California by using the Mesoscale Atmospheric Simulation (MAS) model, at 20-km resolution, as input to the hydrology model TOPMODEL. Results showed a 10% difference between modeled and observed streamflow.

Westrick and Mass (2001) investigated a short-term rain on snow event for the Snoqualmie River, Washington, using the Fifth-Generation PSU–NCAR Mesoscale Model (MM5) to synthesize inputs to the Distributed Hydrology Soil and Vegetation Model (DHSVM) at 4-, 12-, and 36-km spatial resolution. The authors used simple interpolation schemes to downscale the MM5 outputs to the 150-m DHSVM domain and applied a bias adjustment to precipitation and wind speed to better estimate streamflow. Westrick et al. (2002) then extended this analysis to evaluate real-time applications over multiple basins in Washington State though an entire winter. While some basins simulated streamflow better than others, the average error was 38% for the entire winter season, compared with 31% using meteorological-station-based observations.

Wayand et al. (2013) investigated three scenarios over 10 years in the American River basin, California. The first scenario utilized one measurement station and distributed the observations with the Parameter-Elevation Regressions on Independent Slopes Model (PRISM; Daly et al. 1994) for precipitation and temperature. The second case used a 6-km WRF simulation downscaled to the 150-m DHSVM grid using static lapse rates, with the third case combining WRF and empirical models for radiation. The resulting WRF outputs for precipitation were biased high, with smaller diurnal temperature ranges than observations. Modeled streamflow performance was poor for all scenarios due to differences in precipitation timing and application of an uncalibrated hydrology model.

The physically based snow model SNOWPACK (Lehning et al. 2002; Bartelt and Lehning 2002) was forced at a point with output from the Canadian regional model GEM15 to evaluate the potential of simulating critical weak layers for avalanche forecasting (Bellaire et al. 2011; Bellaire and Jamieson 2013). Chen et al. (2014a) used 2-km WRF Model outputs to drive 6 land surface models and compare how closely they matched SWE at 112 Natural Resources Conservation Service (NRCS) Snowpack Telemetry (SNOTEL) stations across the Colorado River headwaters. To do this they used output from the 4 nearest WRF cells to estimate weather parameters at each SNOTEL site. They found that WRF Model output resulted in a reasonable agreement with the 112 SNOTEL site-averaged SWE measurements, but that there was large variation from site to site. Much of this variation was due to elevation differences between the sites and the elevations assigned to the 2-km WRF grid cells.

The WRF Model could provide a significant advancement over current geostatistical methods for deriving meteorological forcing data for hydrologic models where sparse measurement networks exist. Current geostatistical methods interpolate station measurements of meteorological data (Havens et al. 2017; Garen et al. 1994; Goovaerts 2000; Livneh et al. 2014; Luo et al. 2008), however, these methods are severely limited in mountainous regions and may not accurately capture the true spatial distribution of parameters (Hedrick et al. 2018). Regional-scale models, such as WRF, provide a new

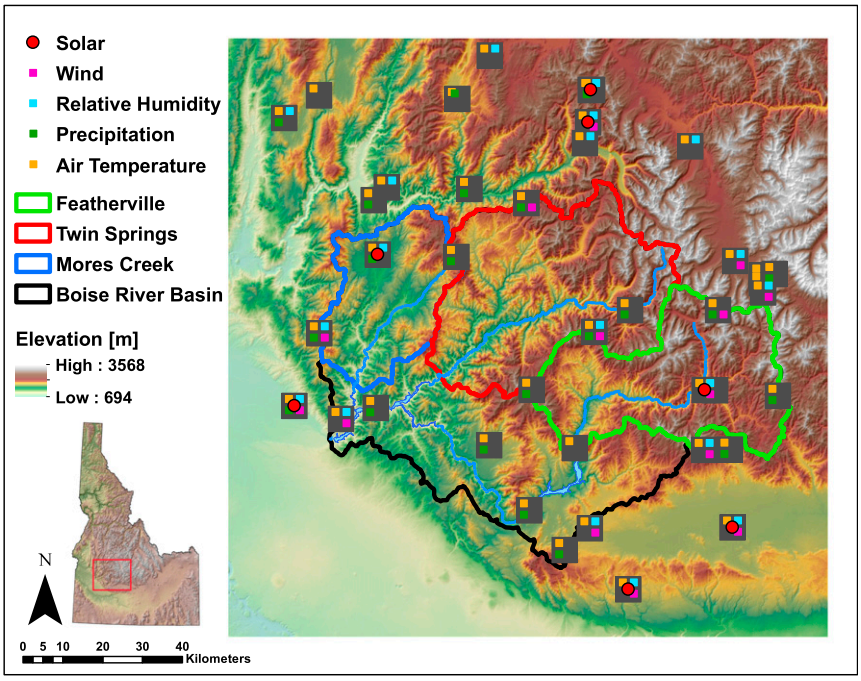

FIG. 1. Boise River Basin modeling domain with major subbasins. Meteorological stations used in modeling with measurements used at each site.

opportunity to simulate historic forcing data when ground measurements are unreliable or at a coarse temporal resolution. However, the downscaling from the regional-scale domain to the snow-modeling domain must be tested and WRF Model outputs evaluated for bias and sensitivity to downscaling.

In this study we used output from a WRF simulation to force a physically based snowmelt model, iSnobal (Marks et al. 1999), in lieu of using meteorological station measurements. The study area has a relatively dense measurement networks for a large watershed allowing for comparison with an atmospheric model, as opposed to the typically sparse and inconsistent measurements found in most large watersheds. The objectives for this study were twofold:

1) Utilize atmospheric model outputs from WRF to produce iSnobal results that are consistent with iSnobal results from measured meteorological conditions.
2) Improve iSnobal results with WRF through additional bias adjustments to precipitation and air temperature, along with applying elevation dependent downscaling.

The utility of the methodology will be determined by comparison of baseline iSnobal results from meteorological stations with two iSnobal results using WRF. The first was simple interpolation of WRF output. The second applied bias adjustment then elevation dependency to the WRF output.

2. Model configurations

a. *Boise River Basin*

The Boise River Basin (BRB) in southwest Idaho, United States, defined in this study as the watershed above Lucky Peak Dam, is located just east of Boise, Idaho, and encompasses roughly 7000 km^2 (Fig. 1). The BRB contains three large subbasins: Mores, Twin Springs, and Featherville, each draining one of the three main forks of the Boise River. The BRB ranges in elevation from 858 to 3249 m, with Mores ranging between 939 and 2470 m, Twin Springs between 997 and 3249 m, and Featherville between 1289 and 3125 m. The majority of winter precipitation occurs as snow, with average annual precipitation of 500 mm at the lowest elevations to 1500 mm at higher elevation (Garen and Marks 2005). The BRB was selected for this study because it is one of the most thoroughly instrumented large mountain basins in the western United States, with more than 40 measurement sites, including 14 SNOTEL sites monitoring both precipitation and SWE.

The BRB is composed of forest (43%), shrub land (35%), herbaceous (21%), and other land covers (1%) based on analysis of the 2011 National Land Cover Database (NLCD; Homer et al. 2015). Each subbasin has its own dominant land cover with Mores mainly forest (62%), Twin Springs land cover similar to the basin average, and Featherville 53% forest and 24% shrub.

Vegetation maps used for modeling were aggregated from 30-m NLCD classifications to a 100-m classification.

b. *iSnobal*

The physically based, distributed snowmelt model, iSnobal (Marks et al. 1999), is a topographic-grid-based energy and water balance model. The iSnobal model simulates the snowpack as two layers, with the active surface layer exchanging energy and mass with the atmosphere and the lower layer transferring energy and mass between the snow surface layer and soil. The snow temperature, density, and liquid water content are calculated for each layer. The iSnobal forcing data inputs are raster surfaces over a digital elevation model (DEM) of incoming thermal (longwave) radiation, air temperature, vapor pressure, wind speed, soil temperature, net solar radiation, and precipitation (including precipitation temperature, phase, and percent snow). Given the forcing data inputs, the energy balance, snow temperature, depth, mass, and cold content (energy required to bring the snowpack to 0°C) are computed for each grid cell. Melt cannot occur until the temperature of the snow cover rises to 0°C, under which conditions the cold content equals 0 J m^{-2}. Liquid water drainage from the snow does not occur until the liquid water holding capacity of the snow is exceeded.

The BRB modeling was conducted within a 1500 by 1500 pixel domain (Fig. 1) at 100-m resolution and at an hourly time step. This resolution allows for a grid size that captures topographically controlled snowmelt processes but does not require an unreasonable computation time. A 100-m DEM was coarsened from the publicly available 10-m DEM to match the required modeling resolution.

c. *Weather Research and Forecasting Model*

The WRF Model (Skamarock et al. 2008) is an atmospheric modeling system used in both research and operations for a variety of applications (e.g., weather forecasting, seasonal forecasting, atmospheric chemistry and regional climate modeling). The code is open source with community development and support provided by the National Center for Atmospheric Research (NCAR). WRF is applied at a range of scales from horizontal resolutions of less than 1 km for large-eddy simulations, to tens of kilometers for coarse resolution forecasts and climate simulations. The model has two dynamical cores one used primarily for operational applications, the Nonhydrostatic Mesoscale Model (NMM), and another, the Advanced Research WRF (WRF-ARW), which is geared toward research and was used to generate the dataset used for this study. The WRF-ARW model solves the governing equations of atmospheric dynamics and uses selected modularized physics packages to account for processes such as cloud microphysics, shortwave and longwave radiation, and land surface processes.

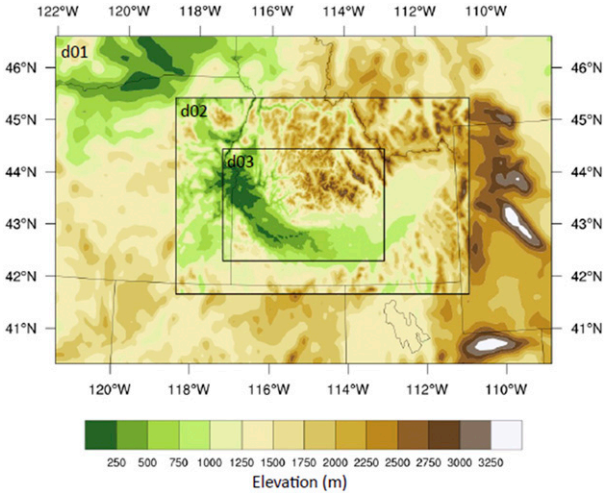

FIG. 2. Nested domains used for the WRF Model simulation. Domains d01, d02, and d03 have horizontal grid resolutions of 9, 3, and 1 km, respectively. Topography is shown as resolved by the model within each domain.

The WRF output used in this paper is from a single simulation generated using WRF version 3.5.1 over nested model domains (Fig. 2) in the interior Pacific Northwest region of the United States. The horizontal resolution of the model domains d01, d02, and d03 are 9, 3, and 1 km, respectively. The simulation extends from 1 October 2009 through 1 June 2010 and uses initial and boundary atmospheric conditions from the North American Regional Reanalysis. Based upon Ikeda et al. (2010), the model configuration included the following physics parameterizations:

- Community Noah land surface model (Chen and Dudhia 2001),
- Thompson et al. (2008) microphysics scheme,
- Mellor–Yamada–Janjić (MYJ) planetary boundary layer scheme (Janjić 2002),
- Community Atmosphere Model (CAM) shortwave and longwave radiation schemes (Collins et al. 2004), and
- Kain–Fritsch convective parameterization scheme (Kain 2004) on the outer domain only.

3. Methods

Three iSnobal model runs were performed at 100-m spatial resolution and at an hourly time step. The first run used only meteorological station measurements, serving as the control simulation. The second model run used bilinear interpolation to regrid WRF output to the

iSnobal domain. The third model run used bias-adjusted WRF air temperature and precipitation from stations in a larger area in south Idaho, then incorporated detrended interpolation of WRF outputs to the iSnobal BRB domain.

a. Spatial distribution of station measurements

Data from a total of 40 meteorological stations in and around the BRB were used as baseline data to derive distributed model inputs as a control model run (Fig. 1). Seven stations are operated by the Bureau of Land Management (BLM), nine are operated by the Boise office of the U.S. Bureau of Reclamation (USBR), seven are operated by the Idaho Transportation Department (ITD), fourteen are part of the SNOTEL network, and three are operated by the Sawtooth National Forest Avalanche Center (SNFAC). Due to multiple station operators, station instrumentation and configuration varied between meteorological stations.

For this study, we used the Spatial Modeling for Resources Framework (SMRF; Havens et al. 2017) version 0.5.3 (Havens et al. 2019b) to interpolate the point stations measurements to the gridded modeling domain. In depth methodologies within SMRF can be found in Havens et al. (2017) and the configuration file for complete replication of the distributed forcing input is available (Havens et al. 2019a). Certain forcing parameters (air temperature, precipitation, and vapor pressure) used elevation-dependent detrended kriging (Garen et al. 1994; Garen 1995), which has been successfully applied in other modeling studies (Susong et al. 1999; Garen and Marks 2005). Detrended kriging first removes the elevation trend from the measurement data by calculating the least squares fit to the measured values as a function of elevation. The residuals are distributed to the gridded model domain using ordinary kriging with an assumed linear semivariogram. The elevation trend is then added back to the kriged residuals based on the gridcell elevation.

Meteorological station measurements of air temperature from 38 stations were distributed using detrended kriging. Air temperature typically has a negative trend with elevation and the detrended kriging was constrained to keep this negative trend.

Dewpoint and vapor pressure rely on coincident measurements of air temperature and relative humidity, which were measured at 19 of the meteorological stations. At each of these stations, the vapor pressure was calculated using the measured air temperature and relative humidity based on Clausius–Clapeyron empirical relationship. The vapor pressure was then distributed using detrended kriging with a negative elevation trend. From the distributed vapor pressure, the dewpoint temperature was calculated.

Twenty-one stations measured precipitation, fourteen of which were SNOTEL stations. The measured precipitation data were filtered to remove both high- and low-amplitude noise while ensuring that the accumulated precipitation was always increasing using the Automated Precipitation Correct Program (APCP; Nayak et al. 2008). The instantaneous precipitation for each station was calculated from the difference in accumulated precipitation, then distributed at an hourly time step using detrended kriging and ensuring a positive elevation trend. Undercatch correction was not performed due to limited wind data at the measurement sites.

Snowmelt models are sensitive to the precipitation phase, as this can either build the snowpack or change the snowpack energy balance and potentially induce melt (Marks et al. 1998; Kormos et al. 2014). The precipitation phase was determined from the precipitation temperature, set as the dewpoint temperature, which allows for a reliable estimate of precipitation phase, especially near the rain–snow transition elevation (Marks et al. 2013). From the precipitation temperature, the percent snow and snow density are estimated using procedures described by Susong et al. (1999).

Fourteen meteorological stations measured wind speed and direction, with three at sheltered SNOTEL locations and the rest in more exposed locations along roadways or in large clearings. With the abundance of snowfall in the BRB, the Idaho Transportation Department and the Sawtooth National Forest Avalanche Center have installed three wind measurement stations on the exposed mountain peaks and have measured wind speeds in excess of $35\,m\,s^{-1}$. These wind stations provide a unique challenge when distributing wind speed using the methods developed by Winstral and Marks (2002) and Winstral et al. (2009). This distribution method uses the maximum upwind slope to parameterize how upwind terrain affects whether or not the pixel is determined to be either sheltered or exposed. The maximum upwind slope parameter was calculated for a 30-m DEM and averaged up to the 100-m model DEM for 72 wind directions. The exposed-peak stations required a special case to set the maximum upwind slope to the minimum calculated value to reduce the potential of artificially high wind speeds in the surrounding areas. Maximum and minimum wind speeds were set at 0.47 and $35\,m\,s^{-1}$, respectively, to ensure turbulent transfer calculation stability.

Estimating solar and thermal radiation over the BRB requires multiple steps, all of which were performed within SMRF. Both solar and thermal radiation require cloud factor calculated as the measured solar radiation over the clear sky radiation. Incoming solar radiation

measured at 7 sites, 3 at BLM stations, 2 at USBR stations, and 2 at ITD stations, was used to estimate the cloud factor. Measured solar radiation was integrated over an entire day to calculate a daily cloud factor. The cloud factor at each station was distributed to the model domain using inverse distance weighting.

The soil temperature was set to a constant −2.5°C for the entire simulation. This ensured that as the snowpack began to form, the ground heat flux did not overcome the energy balance and melt the snowpack prematurely. Once the snowpack was deep enough, the ground heat flux term was small, providing a limited contribution to the calculation of the energy balance.

b. *Simple interpolation of WRF output*

Each WRF output variable was regridded using bilinear interpolation. The variables were 2-m air temperature, 2-m relative humidity, 10-m wind speed, downwelling longwave radiation at the surface, precipitation accumulated, and cloud fraction.

Dewpoint temperature and vapor pressure were calculated from the interpolated air temperature and relative humidity using the Clausius–Clapeyron empirical relationship.

The 10-m wind speed was not adjusted for sheltering due to terrain as with the station measurements. Downscaling wind from a gridded output over complex topography is not a simple solution and will require more research to analyze the best methods.

The solar radiation output from WRF included effects from clouds but was not corrected for terrain shading or split into beam and diffuse radiation. To compensate for this, we followed the same methods as the station measurements by first calculating the clear sky radiation. We then estimated the cloud fraction (CLDFRA) from WRF outputs. CLDFRA is a four dimensional variable (time, latitude, longitude, atmospheric layer), and the distributed cloud cover at the surface was calculated as the average cloud fraction for all atmospheric layers. This method is a crude approximation and further studies will investigate the proper method to account for the cloud's effect on radiation. The final step adjusted for vegetation and corrected the beam and diffuse radiation for the canopy and cloud factor.

Downwelling longwave radiation from WRF was used as cloud corrected thermal radiation. Further adjustment for topographic and canopy effects were performed in the same manner as the station measurements.

The albedo was estimated by keeping track of the time since last storm for each pixel, using the same methodology as for the station measurements. However, the precipitation output from WRF has much finer scale spatial patterns than can be achieved from detrended kriging, making the albedo decay more prevalent and variable than those calculated from station measurements.

c. *Bias-adjusted WRF output*

Two additional modifications were made to improve the WRF outputs for the iSnobal application. First, to further improve the input data distribution, relative to the bilinear technique, we borrowed from the concept of detrended kriging. Each WRF cell was used as if it were a measurement station (i.e., a virtual WRF station) that has a latitude, longitude, and elevation. Stations that were within the BRB boundary were used to get the elevational trend in air temperature. The trend was subtracted out from all the WRF stations to get the residuals, which were distributed to the iSnobal domain using bilinear interpolation. With the residuals on the iSnobal domain, the trend could be added back in, producing an elevation-dependent air temperature within the iSnobal modeling domain.

The second modification was a simple bias correction function applied to each WRF grid cell based on the particular cells' value of temperature and precipitation. Known biases exist in WRF for air temperature when used with the Noah land surface model (Niu et al. 2011; Yang et al. 2011; Chen et al. 2014b). The following sections attempt to apply simple bias corrections to air temperature and precipitation in order to provide more robust inputs to iSnobal. These bias corrections were not meant to address the physics behind WRF but rather to solely adjust WRF output. The most recent WRF version now includes the Noah-MP land surface model which may reduce the temperature bias.

Thirty-nine stations were used in the WRF 1-km inner modeling domain to compare air temperature and precipitation (Fig. 3). Additional stations were utilized from outside of the BRB from the much larger WRF modeling domain. Seven stations were from the Reynolds Creek Experimental Watershed (RCEW) in SW Idaho maintained by the USDA-ARS. The remaining 32 were SNOTEL stations within the modeling domain. The closest WRF pixel to the measurement station was used in the comparisons.

1) AIR TEMPERATURE

To address the temperature bias, the hourly WRF air temperature was compared to the hourly measured air temperature (Fig. 4). A nonlinear cold bias was observed when the air temperature was below about 10°C with a warm bias above 10°C. The histogram shows most of the values centered between −10° and +5°C. The large range in measurements can be attributed to two main factors. The first, comparing hourly measurements directly to hourly WRF output, where a time shift of even

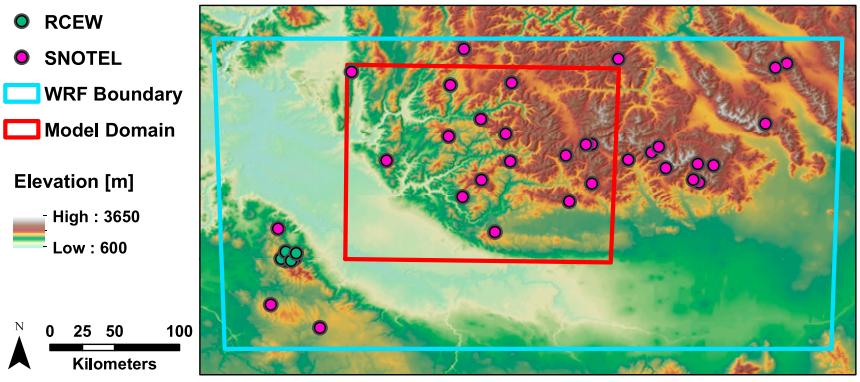

FIG. 3. WRF inner modeling domain d03 in blue with the BRB model domain in red. Stations used for bias analysis from RCEW and SNOTEL shown.

an hour would cause differences between modeled and measured air temperature. Second, the WRF output did not have the diurnal range in maximum and minimum temperatures that were measured at station locations as previously documented by Wayand et al. (2013).

A goodness-of-fit to the desired line (1:1 line), in red, showed a R^2 value of 0.73. To adjust the air temperatures, the offset of the measured bias to the desired line was calculated for the range in WRF air temperatures. This provided a continuous bias correction offset function to apply at each WRF air temperature. Once the correction was applied, the goodness of fit to the desired line improved with a R^2 of 0.79.

To address how well the bias correction did, we compared the daily minimum, average, and maximum air temperatures (Fig. 5). The differences between measured and modeled minimum and maximum air temperature were more apparent, with minimum temperatures showing a nonlinear bias about 0°C (R^2 of 0.74), the average temperature had a similar bias as the hourly measurements (R^2 of 0.80), and the maximum had a cold bias at all temperatures (R^2 of 0.63). After the bias correction was applied, the minimum temperature cold bias was not as strong (R^2 of 0.73), the average temperature improved (R^2 of 0.88), and the maximum still had a cold bias, but the nonlinearity was removed (R^2 of 0.78).

2) DEWPOINT TEMPERATURE AND VAPOR PRESSURE

Applying a bias correction to temperature will have an effect on the modeled relative humidity. To combat this potential change, the original WRF modeled air temperature and relative humidity were used to calculate the dewpoint temperature at each WRF grid cell.

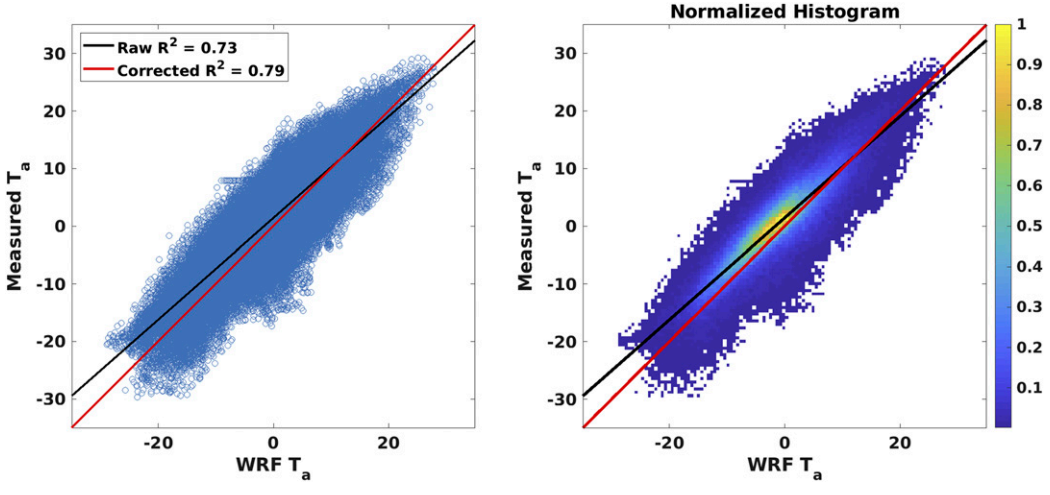

FIG. 4. Hourly measured air temperature vs hourly WRF air temperature output. A cold bias exists at lower temperatures. The wide spread in hourly measurements was due to greater temperature fluctuations at station locations or differences in model timing. Original trend in black with bias corrected in red.

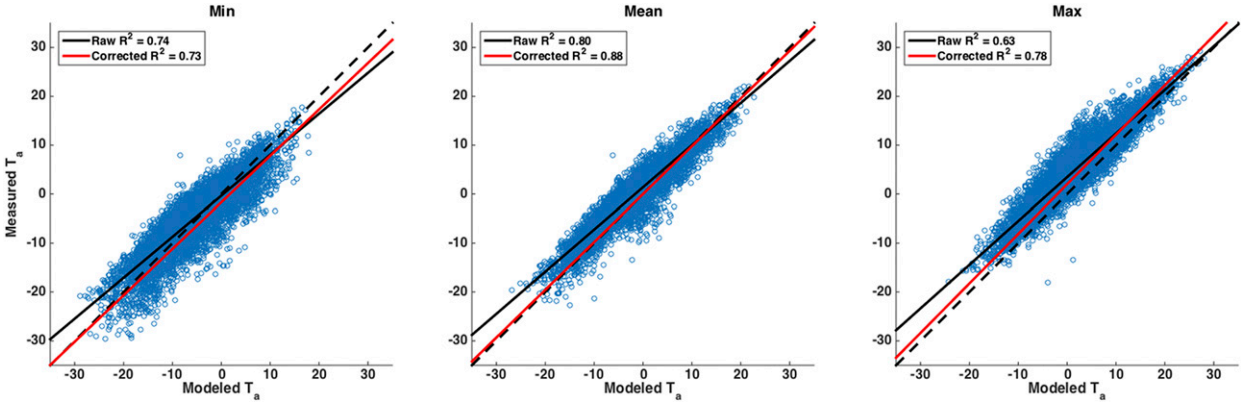

FIG. 5. Daily minimum, average, and maximum air temperatures before correction (black) and after bias correction (red).

The distribution of the dewpoint temperature was performed in a similar manner to the air temperature by assuming each WRF cell is a virtual meteorological station, calculating and removing the elevation trend, interpolating the residuals and finally retrending to the iSnobal gridded domain. The final step ensures that the dewpoint temperature cannot exceed the distributed air temperature. Vapor pressure was then calculated from the distributed dewpoint temperature.

3) PRECIPITATION

The measured accumulated precipitation was calculated for all 39 measurement sites up to the WRF Model end time. The modeled accumulation from WRF compared to measured accumulation showed a striking linear trend (Fig. 6) that was biased to underestimated modeled accumulation with a poor fit to the desired line (R^2 of 0.43). A bias correction ratio was calculated as an enhancement factor between the original linear trend to the desired line as a function of the modeled accumulated precipitation. Applying the enhancement factor improved the goodness of fit to an R^2 value of 0.82. The enhancement factor was applied based on a cell's modeled accumulation to get a distributed enhancement factor. The distributed enhancement factor was applied to the hourly precipitation prior to downscaling to the iSnobal domain. Similar to the air temperature and dewpoint temperature, the hourly precipitation was detrended, interpolated, and retrended based on WRF cells in the BRB boundary. The bias corrected precipitation and dewpoint temperature changed the precipitation phase as well.

4) OTHERS

Wind, thermal radiation, and soil temperature were all calculated in the same manner as the simple bilinear interpolation. Albedo was calculated using the detrended precipitation which changed the calculation of net solar radiation.

4. Results

The objectives for the study were to use WRF outputs as inputs to iSnobal and compare with iSnobal simulations from measured meteorological conditions. Two iSnobal simulations were performed with simple interpolation of WRF output and a more complex, elevation-dependent interpolation with bias adjustment of WRF output. Comparison of model results was feasible due to the extensive and relatively dense measurement network in the BRB, as most

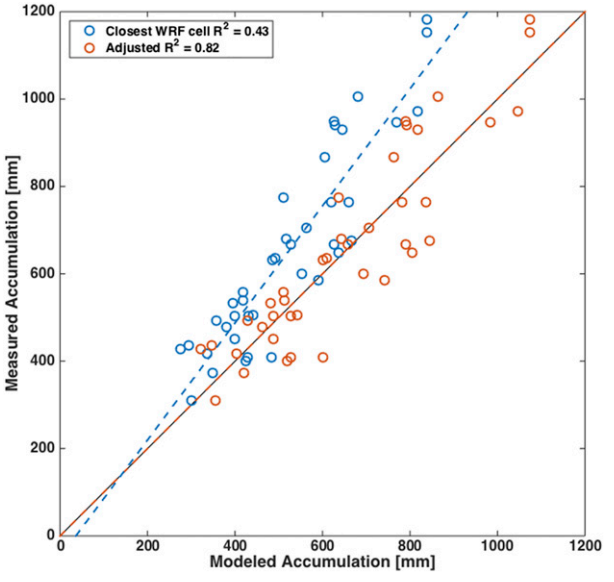

FIG. 6. Measured accumulated precipitation vs modeled accumulation shows the underestimation of precipitation from WRF (blue) and after simple ratio based on accumulation (orange).

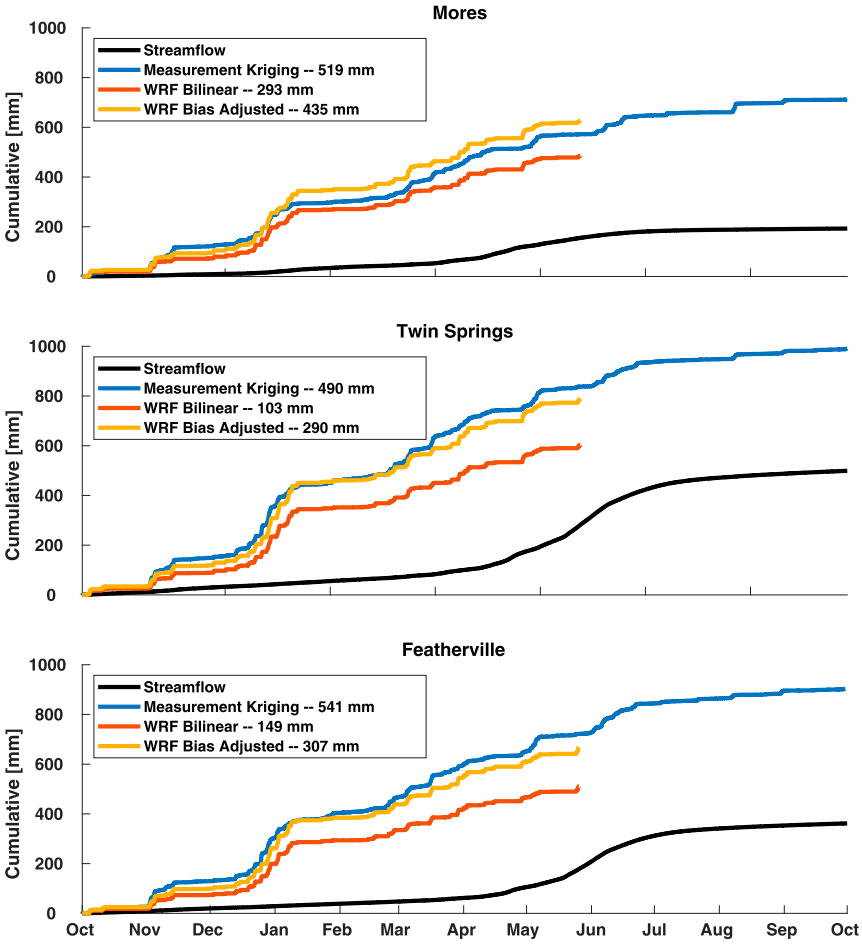

FIG. 7. Simple water balance showing that WRF bilinear did not have enough precipitation to account for streamflow and ET. Bias correcting the WRF precipitation falls more in line with station measurements and will account for streamflow and ET. Values in legend are the difference of the last value and the maximum accumulated streamflow.

watersheds of this size have a sparse and incomplete measurement network.

a. Precipitation bias correction

A simple water balance was used to ensure that enough precipitation volume was distributed to account for basin outflow and evapotranspiration (ET) demands. ET demands for various canopy cover can be estimated using eddy covariance towers. At nearby RCEW, ET estimates were 300 mm yr^{-1} for sagebrush and 700 mm yr^{-1} for aspen (Flerchinger et al. 2010). ET estimates from the semiarid southwestern United States (Ha et al. 2015) were between 415 and 510 mm yr^{-1} and between 400 and 800 mm yr^{-1} in the Sierra Nevada in California (Goulden et al. 2012) for ponderosa pine, a major vegetation species in the Boise River Basin. A rough estimate of ET based on the previous ET values and percent vegetation cover for each basin is between 400 and 490 mm yr^{-1} for Mores, from 360 to 425 mm yr^{-1} for Twin Springs, and between 360 and 440 mm yr^{-1} for Featherville. Mores subbasin has an additional complexity due to a large population drawing up to 7.6 $m^3 s^{-1}$ from surface water for residential and irrigation water rights (Idaho Department of Water Resources, https://research.idwr.idaho.gov/apps/Hydrologic/ccounting/). The water rights have the potential to draw up to 230 mm yr^{-1} for the entire year, however the actual water used is unknown and most likely lower.

Comparing the distributed accumulated precipitation and measured streamflows, normalized to basin area, provides a rough estimate of the available water left for ET (Fig. 7). Based on the ET estimates above, the distributed precipitation from station measurements provided enough water volume to account for streamflow and ET in all the basins. Even though the WRF runs stopped prior to the end of the year, there was only one significant storm during the summer, potentially adding

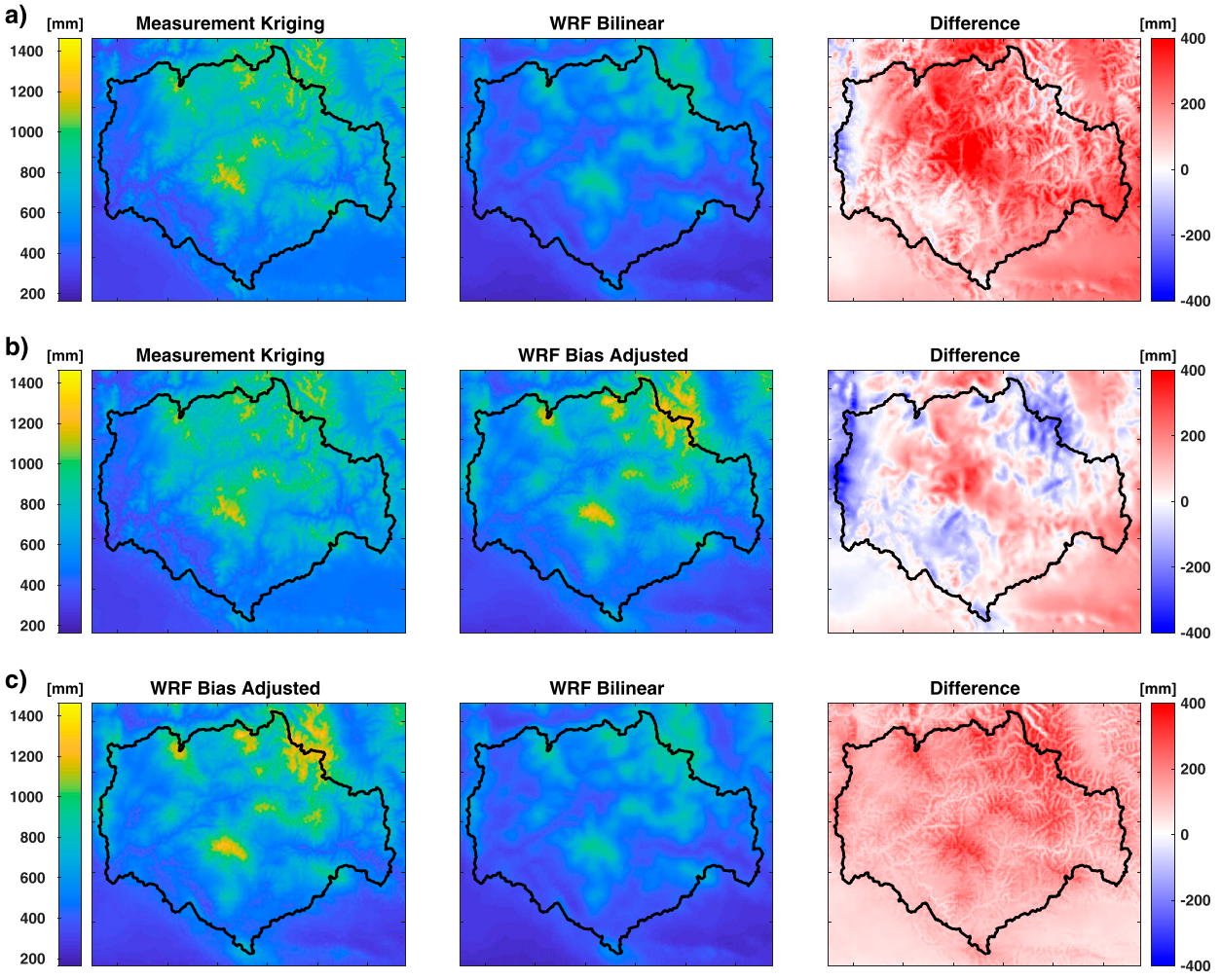

FIG. 8. Distributed accumulated precipitation compared. Measurement kriging tended to smooth the station measurements, whereas WRF had a larger elevation gradient.

up to an additional 170 mm to the distributed precipitation. Even with this addition, the WRF precipitation would not have enough precipitation to account for both streamflow and ET, indicating that the original WRF output did not produce enough precipitation. However, the simple ratio correction created enough precipitation to account for ET and streamflow.

The distributed accumulated precipitation shows a striking difference between simple WRF interpolation and bias adjusted WRF (Fig. 8). Large differences exist everywhere when comparing measurements to simple interpolation, with up to 600 mm of difference in the middle of the basin (Fig. 8a). After bias adjusting the precipitation, WRF had a much stronger elevation gradient, placing a higher concentration of precipitation at higher elevations (Fig. 8b). This differs from the measurement kriging, which tends to smooth the measurements over the domain and does not have as strong of an elevation gradient.

b. *SWE at a point*

Fourteen SNOTEL stations exist in the Boise River Basin with which to compare model results. However, two aspects must be taken into consideration when comparing a spatial model to point measurements, scale differences and topography. A large-scale difference exists between a SNOTEL pillow (7 m^2) that is essentially a point measurement and the 100-m modeling pixel (10 000 m^2). While there are large differences in scale making direct comparison difficult, the importance of the comparison was to ensure that the proper trends in SWE accumulation and melt were captured. In addition to the scale differences, SNOTEL sites are located in sheltered flat areas. However, the underlying DEM of the model will likely

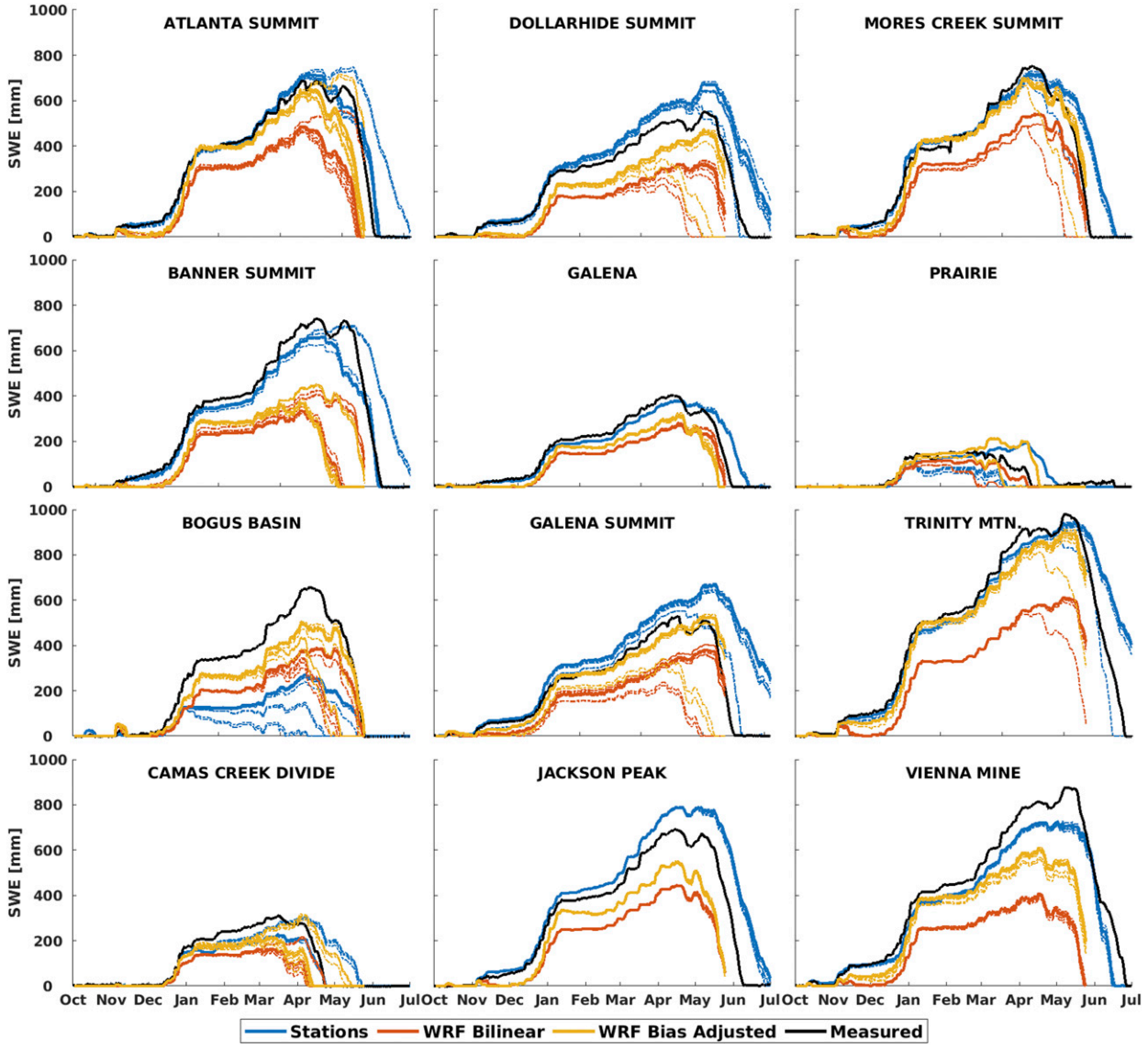

FIG. 9. SWE for station measurements performed well at some sites but not at others. WRF bilinear underestimated precipitation at all sites. WRF bias adjusted performed better but still underpredicted at some. Bogus basin either has an issue with precipitation undercatch or the relative humidity sensor leading to higher dewpoint temperatures and more rain.

not be representative and will have some slope and aspect to the pixel.

The differences between scale and topography can be addressed by not only looking at the particular pixel that the SNOTEL site occupies, but also looking at those pixels directly adjacent. Looking at the pixels adjacent to the site encompasses a 300 m × 300 m area and the local topography becomes apparent. Figure 9 compares 9 out of the 14 sites that had reliable SWE measurements. The model result at the pixel location is represented by a solid line with the adjacent pixels as dashed lines.

The model results with stations performed well at some sites but not at others (Table 1). The topography differences are quite apparent with Jackson Peak having little variation due to the relatively flat topography and significant tree cover. Other sites had a large variation in SWE, for example Banner Summit, where an adjacent pixel matches more closely to the SNOTEL measurements. The model results at Bogus basin were extremely low due to either an under catch of precipitation, low relative humidity measurement which will affect the precipitation phase, or the site being extremely sheltered next to a ski run. The Nash–Sutcliffe efficiency (Nash and Sutcliffe 1970) varies from 0.5 to 0.93 with an average of 0.76. The average root-mean-square error (RMSE) is

TABLE 1. RMSE, MSE, and NSE for each SNOTEL site. Average is the average for all sites and all nine pixels. SM = station measurements, WB = WRF bilinear, and WC = WRF bias corrected.

	RMSE			MSE			NSE		
Station	SM	WB	WC	SM	WB	WC	SM	WB	WC
Atlanta Summit	2.0	10.7	5.1	17	128	51	0.93	0.50	0.80
Banner Summit	2.8	19.5	19.1	30	220	204	0.89	0.18	0.24
Bogus Basin	9.3	9.5	5.5	112	117	66	0.50	0.48	0.71
Cozy Cove	1.9	5.8	4.2	21	62	43	0.81	0.45	0.62
Dollarhide Summit	5.2	9.6	5.0	56	127	70	0.71	0.35	0.64
Galena	2.1	5.4	4.8	21	71	60	0.85	0.50	0.58
Galena Summit	8.2	6.7	2.5	88	86	27	0.52	0.53	0.85
Jackson Peak	5.3	11.1	7.8	59	139	94	0.77	0.44	0.63
Mores Creek Summit	4.4	7.8	2.7	36	100	33	0.86	0.62	0.87
Prairie	2.0	1.8	2.7	19	20	23	0.67	0.65	0.61
Trinity Mountain	6.4	15.3	3.5	63	200	45	0.82	0.43	0.87
Vienna Mine	3.5	19.1	10.5	43	231	116	0.86	0.24	0.62
Average	4.5	10.5	6.5	48	128	73	0.76	0.42	0.65

4.5 mm and an average mean squared error (MSE) of 48 mm.

The underestimation of precipitation is quite evident in the simple interpolation WRF Model results (Fig. 9). The modeled SWE is underestimated at all sites, with up to 400 mm of difference between measured and modeled SWE. The Nash–Sutcliffe efficiency varies from 0.18 to 0.65 with an average of 0.42. The average RMSE is 10.5 mm and an average MSE of 128 mm.

After bias correcting and detrending the precipitation and air temperature, the WRF bias adjusted model results perform much better. The WRF bias adjusted results were closer to the SNOTEL measurements but some sites were still underestimated. The Nash–Sutcliffe efficiency varies from 0.24 to 0.87 with an average of 0.65. The average RMSE is 6.5 mm and the average MSE is 73 mm. The WRF bias correction outperforms the WRF bilinear and is close to the station measurements.

c. *Snow-covered area*

The snow-covered area (SCA) for the Boise River Basin and each sub basin were calculated for the station measurements, simple interpolation, and bias corrected WRF. The SCA from the MODIS product MODSCAG (Painter et al. 2009) was obtained for scenes that had less than 10% cloud cover and mainly occurred during the spring. Comparing the model runs to MODSCAG (Fig. 10) show that the snow cover modeled with station measurements contain snow in more pixels due to interpolating point measurements to the model domain. The simple interpolation of WRF outputs showed the closest to MODSCAG SCA and is attributed to the underestimation of precipitation. The bias corrected WRF outputs have a lower SCA than station measurements and are close to the MODSCAG SCA for all subbasins. The results show that the WRF outputs are providing a different spatial coherency than can be estimated from station interpolation.

5. Discussion and conclusions

The results show that WRF output could be used to provide input to iSnobal in lieu of meteorological station measurements and could replicate results in a large, well instrumented basin. WRF outputs provide spatially distributed forcing data at a much finer scale that cannot be replicated with geostatistical methods to interpolate measured meteorological data in a large basin such as the BRB. The ability to create all the necessary forcing inputs to iSnobal makes using WRF outputs enticing, especially where sparse measurement networks exist or have unreliable data. In some sense, regional weather and climate models like WRF are platforms for interpolating large-scale atmospheric motions in accordance with the physics and thermodynamics as they are represented within these models to much finer spatial resolutions. However, caution must be used and the WRF outputs should be carefully checked to ensure that the model outputs are realistic as compared with meteorological station measurements. Without correcting for bias, the modeled snowpack was underestimated at all SNOTEL locations, mainly due to the underestimation of precipitation, and produced larger errors than using measured station data. After applying simple bias corrections to the air temperature and precipitation, the modeled snowpack performed almost as well as the measured station data at SNOTEL locations. The NSE and RMSE show that using WRF outputs as inputs can simulate SWE almost as well as with meteorological measurements. The bias correction however, will most

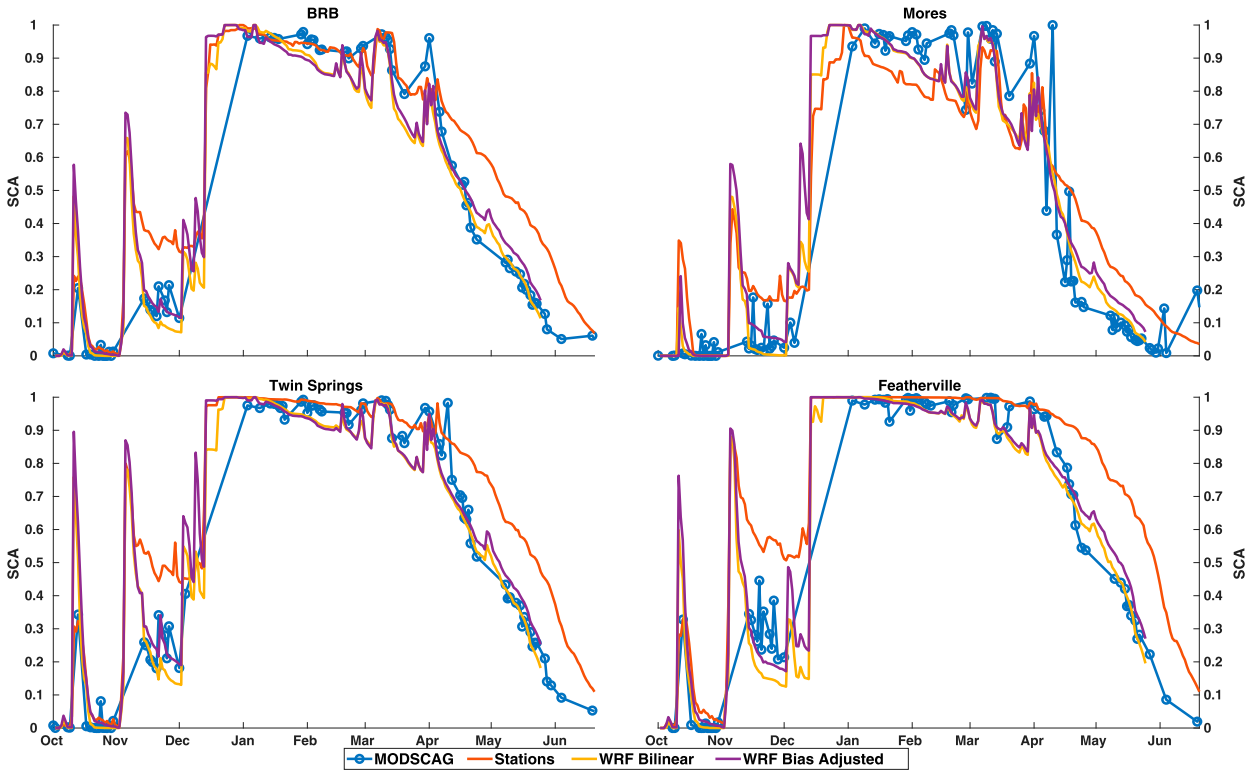

FIG. 10. Comparison of SCA between model runs. Station measurements show the highest SCA during the spring.

likely not be applicable to other WRF simulations from the complicated uncertainty in the model outputs which may vary between seasons, simulations, space, and model configuration. Therefore, each model run must be examined independently and prior to use, to ensure that WRF outputs are adequate as inputs for snowmelt modeling. To adequately evaluate WRF outputs, there must be enough station meteorological measurements to capture the measurement gradient, which can be difficult in regions with sparse measurements. WRF simulations that span a longer time frame (i.e., multiple years) will need a more sophisticated bias adjustment to address interannual variability and potential climate change effects.

In applying techniques to regions where the sparseness of surface observations precludes the bias correction technique used here, it may be necessary to rely on ancillary sources of information to correct hydrometeorological fields that are then used as input to hydrologic models. Because WRF is also associated with a land surface model (Noah, in this case), a number of land surface variables influenced by precipitation can be compared to observational data and used to infer bias in the associated precipitation. Two variables, in particular, that may be of interest in correcting biases in WRF precipitation include soil moisture and runoff derived from the land surface model and for which remote sensing (soil moisture) and surface observations (discharge) may be available.

With the increased spatial information provided, new methods were developed to handle downscaling from the WRF domain to the iSnobal domain. The WRF grid cell locations were taken as if they were virtual meteorological station measurements, with a latitude, longitude, and elevation, in order to take advantage of detrended interpolation for WRF outputs that have some elevation dependence, like air temperature, vapor pressure, and precipitation. Additionally, the precipitation distribution from WRF was highly event dependent and can spatially resolve precipitation at a much finer scale. Because a storm could affect only part of the basin, the day since last storm was tracked for each pixel in the iSnobal domain. This produced a distributed albedo estimate and each pixel could decay at a different rate.

The methods described in this study to downscale output from meso- or regional-scale atmospheric models were developed using a reanalysis dataset but can be applied for other mesoscale model outputs. Based on this work, USDA-ARS worked to implement the operational National Weather Service model High Resolution Rapid Refresh (HRRR; Benjamin et al. 2016) as input to iSnobal in near real time in support of water

supply forecasts in California. Additionally, USDA-ARS provided near-real-time forecasts for the USBR and NRCS in the Boise River Basin (Havens et al. 2015). For water year 2016, the methods developed in the study were applied to a short-term 72-h forecast to help quantify how the forecasted precipitation and temperature will affect snow accumulation or melt to provide additional information, from physically based models, to operational water managers.

Acknowledgments. The data and analysis presented were funded in part by USDA-ARS CRIS Snow and Hydrologic Processes in the Intermountain West (5362-13610-008-00D), USDA-NRCS Water and Climate Center-Portland, Oregon (60-5362-4-003), NASA-JPL Airborne Snow Observatory (58-2052-5-006), USGS (60-5362-4-002), USBR (60-5362-3-002), and NSF Reynolds Creek CZO Project (58-5362-4-004). Any reference to specific equipment types or manufacturers is for information purposes and does not represent a product endorsement or recommendation. USDA is an equal opportunity provider and employer.

REFERENCES

Bartelt, P., and M. Lehning, 2002: A physical SNOWPACK model for the Swiss avalanche warning Part I: numerical model. *Cold Reg. Sci. Technol.*, **35**, 123–145, https://doi.org/10.1016/S0165-232X(02)00074-5.

Bellaire, S., and B. Jamieson, 2013: Forecasting the formation of critical snow layers using a coupled snow cover and weather model. *Cold Reg. Sci. Technol.*, **94**, 37–44, https://doi.org/10.1016/j.coldregions.2013.06.007.

——, J. B. Jamieson, and C. Fierz, 2011: Forcing the snow-cover model SNOWPACK with forecasted weather data. *Cryosphere*, **5**, 1115–1125, https://doi.org/10.5194/tc-5-1115-2011.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 587–604, https://doi.org/10.1175/1520-0493(2001)129<0587:CAALSH>2.0.CO;2.

——, and Coauthors, 2014a: Modeling seasonal snowpack evolution in the complex terrain and forested Colorado Headwaters region: A model intercomparison study. *J. Geophys. Res. Atmos.*, **119**, 13 795–13 819, https://doi.org/10.1002/2014JD022167.

——, C. Liu, J. Dudhia, and M. Chen, 2014b: A sensitivity study of high-resolution regional climate simulations to three land surface models over the western United States. *J. Geophys. Res. Atmos.*, **119**, 7271–7291, https://doi.org/10.1002/2014JD021827.

Collins, W. D., and Coauthors, 2004: Description of the NCAR Community Atmosphere Model (CAM 3.0). NCAR Tech. Note NCAR/TN-464+STR, 214 pp., https://doi.org/10.5065/D63N21CH.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.*, **33**, 140–158, https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.

Flerchinger, G. N., D. Marks, M. L. Reba, Q. Yu, and M. S. Seyfried, 2010: Surface fluxes and water balance of spatially varying vegetation within a small mountainous headwater catchment. *Hydrol. Earth Syst. Sci.*, **14**, 965–978, https://doi.org/10.5194/hess-14-965-2010.

Frei, A., D. A. Robinson, and M. G. Hughes, 1999: North American Snow Extent : 1900–1994. *Int. J. Climatol.*, **19**, 1517–1534, https://doi.org/10.1002/(SICI)1097-0088(19991130)19:14<1517::AID-JOC437>3.0.CO;2-I.

Freudiger, D., I. Kohn, K. Stahl, and M. Weiler, 2014: Large-scale analysis of changing frequencies of rain-on-snow events with flood-generation potential. *Hydrol. Earth Syst. Sci.*, **18**, https://doi.org/10.5194/hess-18-2695-2014.

Garen, D. C., 1995: Estimation of spatially distributed values of daily precipitation in mountainous areas. *Mountain Hydrology: Peaks and Valleys in Research and Applications*, Canadian Water Resources Association, 237–242.

——, and D. Marks, 2005: Spatially distributed energy balance snowmelt modelling in a mountainous river basin: Estimation of meteorological inputs and verification of model results. *J. Hydrol.*, **315**, 126–153, https://doi.org/10.1016/j.jhydrol.2005.03.026.

——, G. L. Johnson, and C. L. Hanson, 1994: Mean areal precipitation for daily hydrologic modeling in mountainous regions. *J. Amer. Water Resour. Assoc.*, **30**, 481–491, https://doi.org/10.1111/j.1752-1688.1994.tb03307.x.

Goovaerts, P., 2000: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.*, **228**, 113–129, https://doi.org/10.1016/S0022-1694(00)00144-X.

Goulden, M. L., R. G. Anderson, R. C. Bales, A. E. Kelly, M. Meadows, and G. C. Winston, 2012: Evapotranspiration along an elevation gradient in California's Sierra Nevada. *J. Geophys. Res.*, **117**, G03028, https://doi.org/10.1029/2012JG002027.

Groisman, P. Ya., R. W. Knight, and T. R. Karl, 2001: Heavy precipitation and high streamflow in the contiguous United States: Trends in the twentieth century. *Bull. Amer. Meteor. Soc.*, **82**, 219–246, https://doi.org/10.1175/1520-0477(2001)082<0219:HPAHSI>2.3.CO;2.

——, R.W. Knight, T.R. Karl, D.R. Easterling, B. Sun, and J.H. Lawrimore, 2004: Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations. *J. Hydrometeor.*, **5**, 64–85, https://doi.org/10.1175/1525-7541(2004)005<0064:CCOTHC>2.0.CO;2.

Ha, W., T. E. Kolb, A. E. Springer, S. Dore, F. C. O'Donnell, R. Martinez Morales, S. Masek Lopez, and G. W. Koch, 2015: Evapotranspiration comparisons between eddy covariance measurements and meteorological and remote-sensing-based models in disturbed ponderosa pine forests. *Ecohydrology*, **8**, 1335–1350, https://doi.org/10.1002/eco.1586.

Havens, S., D. Marks, A. Winstral, and E. Rothwell, 2015: Application of a physically-based distributed snowmelt model in support of reservoir operations and water management. Final Rep. ST-2015-2264, Bureau of Reclamation, 39 pp https://www.usbr.gov/research/projects/detail.cfm?id=2157.

——, ——, P. Kormos, and A. Hedrick, 2017: Spatial Modeling for Resources Framework (SMRF): A modular framework for developing spatial forcing data for snow modeling in mountain basins. *Comput. Geosci.*, **109**, 295–304, https://doi.org/10.1016/j.cageo.2017.08.016.

——, ——, K. FitzGerald, M. Masarik, A. N. Flores, P. Kormos, and A. Hedrick, 2019a: Dataset: Approximating input data to a snowmelt model using Weather Research and Forecasting model outputs in lieu of meteorological measurements. Zenodo, accessed 19 February 2019, https://doi.org/10.5281/ZENODO.2565321.

——, ——, P. Kormos, A. Hedrick, M. Johnson, M. Sandusky, and M. Robertson, 2019b: USDA-ARS-NWRC/smrf: v0.5.3. Zenodo, accessed 12 February 2019, https://doi.org/10.5281/ZENODO.2563226.

Hedrick, A. R., and Coauthors, 2018: Direct insertion of NASA Airborne Snow Observatory-derived snow depth time-series into the *iSnobal* energy balance snow model. *Water Resour. Res.*, **54**, 8045–8063, https://doi.org/10.1029/2018WR023190.

Homer, C. G., and Coauthors, 2015: Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogramm. Eng. Remote Sensing*, **81**, 345–354.

Ikeda, K., and Coauthors, 2010: Simulation of seasonal snowfall over Colorado. *Atmos. Res.*, **97**, 462–477, https://doi.org/10.1016/j.atmosres.2010.04.010.

Janjić, Z. I., 2002: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. NCEP Office Note 437, 61 pp., http://www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Kattelmann, R., 1997: Flooding from rain-on-snow events in the Sierra Nevada. *North American Water and Environment Congress & Destructive Water*, C. Bathala, Ed., ASCE, 59–65.

Klos, P. Z., T. E. Link, and J. T. Abatzoglou, 2014: Extent of the rain-snow transition zone in the western U.S. under historic and projected climate. *Geophys. Res. Lett.*, **41**, 4560–4568, https://doi.org/10.1002/2014GL060500.

Kormos, P. R., D. Marks, J. P. McNamara, H. P. Marshall, A. Winstral, and A. N. Flores, 2014: Snow distribution, melt and surface water inputs to the soil in the mountain rain-snow transition zone. *J. Hydrol.*, **519**, 190–204, https://doi.org/10.1016/j.jhydrol.2014.06.051.

Lehning, M., P. Bartelt, B. Brown, C. Fierz, and P. Satyawali, 2002: A physical SNOWPACK model for the Swiss avalanche warning Part II. Snow microstructure. *Cold Reg. Sci. Technol.*, **35**, 147–167, https://doi.org/10.1016/S0165-232X(02)00073-3.

Livneh, B., J. S. Deems, D. Schneider, J. Barsugli, and N. P. Molotch, 2014: Filling in the gaps: Inferring spatially distributed precipitation from gauge observations over complex terrain. *Water Resour. Res.*, **50**, 8589–8610, https://doi.org/10.1002/2014WR015442.

Luo, W., M. C. Taylor, and S. R. Parker, 2008: A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales. *Int. J. Climatol.*, **28**, 947–959, https://doi.org/10.1002/joc.1583.

Lute, A. C., J. T. Abatzoglou, and K. C. Hegewisch, 2015: Projected changes in snowfall extremes and interannual variability of snowfall in the western United States. *Water Resour. Res.*, **51**, 960–972, https://doi.org/10.1002/2014WR016267.

Marks, D., and J. Frew, 1999: Software tools for hydro-climatic modeling and analysis: Image Processing Workbench, ARS - USGS Version 2. https://github.com/USDA-ARS-NWRC/ipw.

——, J. Kimball, D. Tingey, and T. Link, 1998: The sensitivity of snowmelt processes to climate conditions and forest cover during rain on snow: A case study of the 1996 Pacific Northwest flood. *Hydrol. Processes*, **12**, 1569–1587, https://doi.org/10.1002/(SICI)1099-1085(199808/09)12:10/11<1569::AID-HYP682>3.0.CO;2-L.

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

——, 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, https://doi.org/10.1016/j.advwatres.2012.11.012.

Miller, N. L., and J. Kim, 1996: Numerical prediction of precipitation and river flow over the Russian River watershed during the January 1995 California storms. *Bull. Amer. Meteor. Soc.*, **77**, 101–105, https://doi.org/10.1175/1520-0477(1996)077<0101:NPOPAR>2.0.CO;2.

Mote, P. W., 2003: Trends in snow water equivalent in the Pacific Northwest and their climatic causes. *Geophys. Res. Lett.*, **30**, 1601, https://doi.org/10.1029/2003GL017258.

——, 2006: Climate-driven variability and trends in mountain snowpack in western North America. *J. Climate*, **19**, 6209–6220, https://doi.org/10.1175/JCLI3971.1.

——, A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier, 2005: Declining mountain snowpack in western North America. *Bull. Amer. Meteor. Soc.*, **86**, 39–49, https://doi.org/10.1175/BAMS-86-1-39.

Nash, J. E., and J. V. Sutcliffe, 1970: River flow forecasting through conceptual models Part I—A discussion of principles. *J. Hydrol.*, **10**, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6.

Nayak, A., D. G. Chandler, D. Marks, J. P. McNamara, and M. Seyfried, 2008: Correction of electronic record for weighing bucket precipitation gauge measurements. *Water Resour. Res.*, **44**, W00D11, https://doi.org/10.1029/2008WR006875.

——, D. Marks, D. G. Chandler, and M. Seyfried, 2010: Long-term snow, climate and streamflow trends from at the Reynolds Creek Experimental Watershed, Owyhee Mountains, Idaho, United States. *Water Resour. Res.*, **46**, W06519, https://doi.org/10.1029/2008WR007525.

Niu, G. Y., and Coauthors, 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res.*, **116**, D12109, https://doi.org/10.1029/2010JD015139.

Nolin, A. W., and C. Daly, 2006: Mapping ''at risk'' snow in the Pacific Northwest. *J. Hydrometeor.*, **7**, 1164–1171, https://doi.org/10.1175/JHM543.1.

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

Pierce, D. W., and Coauthors, 2008: Attribution of declining western U.S. snowpack to human effects. *J. Climate*, **21**, 6425–6444, https://doi.org/10.1175/2008JCLI2405.1.

Pomeroy, J., R. Stewart, and P. Whitfield, 2016: The 2013 flood event in the South Saskatchewan and Elk River basins: Causes, assessment and damages. *Can. Water Resour. J.*, **41**, 105–117, https://doi.org/10.1080/07011784.2015.1089190.

Regonda, S. K., B. Rajagopalan, M. Clark, and J. Pitlick, 2005: Seasonal cycle shifts in hydroclimatology over the western

United States. *J. Climate*, **18**, 372–384, https://doi.org/10.1175/JCLI-3272.1.

Rössler, O., P. Froidevaux, U. Börst, R. Rickli, O. Martius, and R. Weingartner, 2014: Retrospective analysis of a nonforecasted rain-on-snow flood in the Alps—A matter of model limitations or unpredictable nature? *Hydrol. Earth Syst. Sci.*, **18**, 2265–2285, https://doi.org/10.5194/hess-18-2265-2014.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Surfleet, C. G., and D. Tullos, 2013: Variability in effect of climate change on rain-on-snow peak flow events in a temperate climate. *J. Hydrol.*, **479**, 24–34, https://doi.org/10.1016/j.jhydrol.2012.11.021.

Susong, D., D. Marks, and D. Garen, 1999: Methods for developing time-series climate surfaces to drive topographically distributed energy- and water-balance models. *Hydrol. Processes*, **13**, 2003–2021, https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<2003::AID-HYP884>3.0.CO;2-K.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 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, https://doi.org/10.1175/2008MWR2387.1.

Tohver, I. M., A. F. Hamlet, and S. Y. Lee, 2014: Impacts of 21st-century climate change on hydrologic extremes in the Pacific Northwest region of North America. *J. Amer. Water Resour. Assoc.*, **50**, 1461–1476, https://doi.org/10.1111/jawr.12199.

Trujillo, E., and N. P. Molotch, 2014: Snowpack regimes of the western United States. *Water Resour. Res.*, **50**, 5611–5623, https://doi.org/10.1002/2013WR014753.

Wayand, N. E., A. F. Hamlet, M. Hughes, S. I. Feld, and J. D. Lundquist, 2013: Intercomparison of meteorological forcing data from empirical and mesoscale model sources in the North Fork American River basin in northern Sierra Nevada, California. *J. Hydrometeor.*, **14**, 677–699, https://doi.org/10.1175/JHM-D-12-0102.1.

Westrick, K. J., and C. F. Mass, 2001: An evaluation of a high-resolution hydrometeorological modeling system for prediction of a cool-season flood event in a coastal mountainous watershed. *J. Hydrometeor.*, **2**, 161–180, https://doi.org/10.1175/1525-7541(2001)002<0161:AEOAHR>2.0.CO;2.

——, P. Storck, and C. F. Mass, 2002: Description and evaluation of a hydrometeorological forecast system for mountainous watersheds. *Wea. Forecasting*, **17**, 250–262, https://doi.org/10.1175/1520-0434(2002)017<0250:DAEOAH>2.0.CO;2.

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

——, ——, and R. Gurney, 2009: An efficient method for distributing wind speeds over heterogeneous terrain. *Hydrol. Processes*, **23**, 2526–2535, https://doi.org/10.1002/hyp.7141.

——, ——, and ——, 2014: Assessing the sensitivities of a distributed snow model to forcing data resolution. *J. Hydrometeor.*, **15**, 1366–1383, https://doi.org/10.1175/JHM-D-13-0169.1.

Yang, Z. L., and Coauthors, 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. *J. Geophys. Res.*, **116**, D12110, https://doi.org/10.1029/2010JD015140.