

Accumulation and melt dynamics of snowpack from a multiresolution regional climate model in the central Sierra Nevada, California

Tamlin M. Pavelsky,¹ Sarah Kapnick,^{2,3} and Alex Hall²

Received 9 December 2010; revised 23 May 2011; accepted 2 June 2011; published 25 August 2011.

[1] The depth and timing of snowpack in the Sierra Nevada Mountains are of fundamental importance to California water resource availability, and recent studies indicate a shift toward earlier snowmelt consistent with projected impacts of anthropogenic climate change. In order for future studies to assess snowpack variability on seasonal to centennial time scales, physically based models of snowpack evolution at high spatial resolution must be improved. Here we evaluate modeled snowpack accuracy for the central Sierra Nevada in the Weather Research and Forecasting regional climate model coupled to the Noah land surface model. A simulation with nested domains at 27, 9, and 3 km grid spacings is presented for November 2001 to July 2002. Model outputs are compared with daily snowpack observations at 41 locations, air temperature at 31 locations, and precipitation at 10 locations. Comparison of snowpack at different resolutions suggests that 27 km simulations substantially underestimate snowpack, while 9 and 3 km simulations are closer to observations. Regional snowpack accumulation is accurately simulated at these high resolutions, but model snowmelt occurs an average of 22–25 days early. Some error can be traced to differences in elevation and observation scale between point-based measurements and model grid cells, but these factors cannot explain the persistent bias toward early snowmelt. A high correlation between snowmelt and error in modeled surface air temperature is found, with melt coinciding systematically with excessively cold air temperatures. One possible source of bias is an imbalance in turbulent heat fluxes, erroneously warming the snowpack while cooling the surface atmosphere.

Citation: Pavelsky, T. M., S. Kapnick, and A. Hall (2011), Accumulation and melt dynamics of snowpack from a multiresolution regional climate model in the central Sierra Nevada, California, *J. Geophys. Res.*, *116*, D16115, doi:10.1029/2010JD015479.

1. Introduction

[2] Water resource availability in much of the Northern Hemisphere is strongly dependent on wintertime storage and springtime melt of mountain snowpacks. Seasonal snowpacks and glaciers provide water resources to more than a sixth of the global population [Barnett *et al.*, 2005], and global modeling studies suggest that warming temperatures and changing atmospheric circulation patterns associated with anthropogenic climate change will likely result in altered snowpack timing and magnitude in many areas [Arnell, 1999; Vicuna *et al.*, 2007]. Among the regions most dependent on mountain snowpack is the American West, especially the state of California, where snowmelt releases

stored water resources during seasons with little precipitation [Mote *et al.*, 2005].

[3] Recent studies of historical snowpack in California suggest that a move toward earlier snowmelt and peak streamflow is already apparent [Kapnick and Hall, 2009, 2010; Hidalgo *et al.*, 2009; Barnett *et al.*, 2008; Weare and Du, 2008; Maurer *et al.*, 2007; Howat and Tulaczyk, 2005a, 2005b; Barnett *et al.*, 2004; Cayan *et al.*, 2001; Dettinger and Cayan, 1995]. Most projections of future change in mountain snowpack in the American West have been based on global climate model (GCM) output, but GCMs alone often fail to correctly simulate mountain snowpacks because the highly complex orography associated with mountain belts cannot be adequately captured with coarse GCM grid spacings [Cayan *et al.*, 2008]. As a result, most studies downscale GCM output via dynamical methods such as a regional climate model (RCM) [Salathé *et al.*, 2010; Qian *et al.*, 2009], through distributed hydrologic models such as the Variable Infiltration Capacity (VIC) model [VanRheenen *et al.*, 2004; Vicuna *et al.*, 2007; Maurer, 2007; Maurer *et al.*, 2007; Hamlet *et al.*, 2005; Barnett *et al.*, 2008], or using other statistical techniques [Miller *et al.*, 2003; Dettinger *et al.*, 2004; Maurer and Duffy, 2005; Stewart *et al.*,

¹Department of Geological Sciences, University of North Carolina, Chapel Hill, North Carolina, USA.

²Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, California, USA.

³NASA Goddard Institute for Space Studies, New York, New York, USA.

2004]. Results of these recent downscaling studies suggest that as temperatures in the western United States warm, snowmelt is likely to occur earlier in the spring, more precipitation will fall as rain instead of snow, and overall snowpack will decline in thickness and duration. Beyond this general qualitative agreement, however, there is substantial quantitative uncertainty in future projections of the spatial and temporal extent of snowpack loss. The improvement of future snowpack projections may occur through improved statistical downscaling methods or increased accuracy of physically based models, but each approach has challenges. The accuracy of statistical downscaling approaches depends on the assumed stationarity of statistical relationships developed using existing data sets, an assumption that remains true under a changing climate to an unknown degree [Milly *et al.*, 2008]. Meanwhile, greater accuracy in physically based modeling efforts depends on the more accurate simulation of snowpack physics in current and future climates.

[4] Among the most promising avenues for improving physically based simulation of current and future snowpacks is through regional climate models such as the Weather Research and Forecasting (WRF) model developed by the National Center for Atmospheric Research and a number of partner agencies [Skamarock *et al.*, 2008]. WRF has been successfully coupled to several land surface models that incorporate snowpack physics, including the Noah model used here [Chen and Dudhia, 2001]. Evaluations of snowpack in earlier versions of WRF using a variety of land surface schemes and in other regional climate models suggest generally low skill in tracking the accumulation and ablation of snowpack water content over a snow season [e.g., Wang *et al.*, 2010; Caldwell *et al.*, 2009; Duffy *et al.*, 2006]. Recent studies evaluating snowpack in Noah suggest that changes to albedo parameterizations and other model components may result in substantial increases in model skill [Livneh *et al.*, 2010; Wang and Zeng, 2010; Barlage *et al.*, 2010]. Some of the suggested changes presented in these studies have been incorporated into the version of WRF-Noah used here (version 3.1.1).

[5] Despite recent advances in WRF-Noah, detailed evaluation of snowpack simulations has been limited. Some studies have largely focused on qualitative comparisons of snowpack extent and snow water equivalent (SWE) across large areas [Qian *et al.*, 2009], while others have conducted more extensive statistical testing at a handful of specific stations [Wang *et al.*, 2010; Livneh *et al.*, 2010; Feng *et al.*, 2008]. Among recent studies, only the work of Barlage *et al.* [2010] and Ikeda *et al.* [2010] has conducted a quantitative comparison between WRF-Noah snowpack and observations across a large number of observation sites in Colorado. These studies utilize version WRF-Noah v. 3.0, and their principal recommendation is to add an improved time-varying albedo term. This recommendation has been followed in WRF-Noah v 3.1.1, which is used here. In part, validation of snowpack in WRF-Noah has been limited because evaluation of gridded model output against point-based snowpack measurements presents inherent difficulties. Mountain snowpacks exhibit substantial spatial heterogeneity on scales even smaller than that of high-resolution model grid cells (e.g., 3 km) [Liston, 2004; Anderton *et al.*, 2004], so comparison of observed and modeled snowpack may result in apparent error simply due to this scaling mismatch, even

where model simulations of conditions averaged at the grid scale are correct. However, all comparisons of gridded data with point observations rest on the assumption that this type of error is randomly distributed [Barlage *et al.*, 2010; Ikeda *et al.*, 2010]. Absent any sources of persistent bias, comparison of observed and modeled snowpack at a sufficiently large number of points should yield a distribution with a mean error close to zero if the model snowpack physics are correct and sufficiently detailed. An additional source of potential error is derived from the effects of vegetation on snowpack. Versions of WRF-Noah from v 3.1 onward include consideration of vegetation impacts on surface albedo in snow covered areas but do not account for snowpack interception or effects of the canopy on longwave radiation, both of which can be important in forested environments [Andreadis *et al.*, 2009; Pomeroy *et al.*, 2008, 2009]. However, because snowpack monitoring stations are generally sited in clearings rather than under forest canopy, it is likely that these processes do not substantially affect either the observations or model output used here.

[6] In this study, we simulate snowpack and climate using the WRF-Noah model over the central portion of the Sierra Nevada Mountains for the period November 2001 through July 2002. This time period was selected to have approximately mean April snowpack values (96% of normal from 1950 to 2008) across 60 monthly snow stations used by Kapnick and Hall [2010]. We compare model output against daily SWE observations from 41 stations. We hypothesize that by comparing observed and modeled snowpack at such a large number of geographically proximate locations we will be able to determine whether errors are due to a systematic model bias or are simply the result of different scales of observation. While errors at any individual station may result from the mismatched comparison of point-based to gridded data sets, bias in the distribution of errors at many stations can likely be attributed to problems with the model. Specifically, we address: (1) the impact of model grid spacing on SWE simulation, (2) the accuracy of WRF-Noah SWE simulation, and (3) sources of error in WRF-Noah SWE output.

2. Data and Methods

2.1. The WRF-Noah Modeling System

[7] The WRF model has been developed by the National Center for Atmospheric Research (NCAR) and a number of partner agencies [Skamarock *et al.*, 2008]. WRF has been used extensively for both operational meteorological forecasts and regional studies of climate and meteorology. Here, we use WRF version 3.1.1 with the North American Regional Reanalysis (NARR) product [Mesinger *et al.*, 2006] as lateral boundary conditions over the period November 2001 to July 2002 with output every 3 h. Simulations are conducted in three one-way nested domains at 27, 9, and 3 km horizontal resolution (see Figure 1). These domains are referred to in the text as D27, D9, and D3, respectively. The simulations are run for 9 months without reinitialization by updating the large-scale forcing along the lateral boundaries at 3 h intervals. NARR provides initial (at time step 0) and the regularly updated lateral boundary conditions for D27. The boundary conditions for D9 and D3 are obtained from the successively coarser model domain. As such, simulations of all three

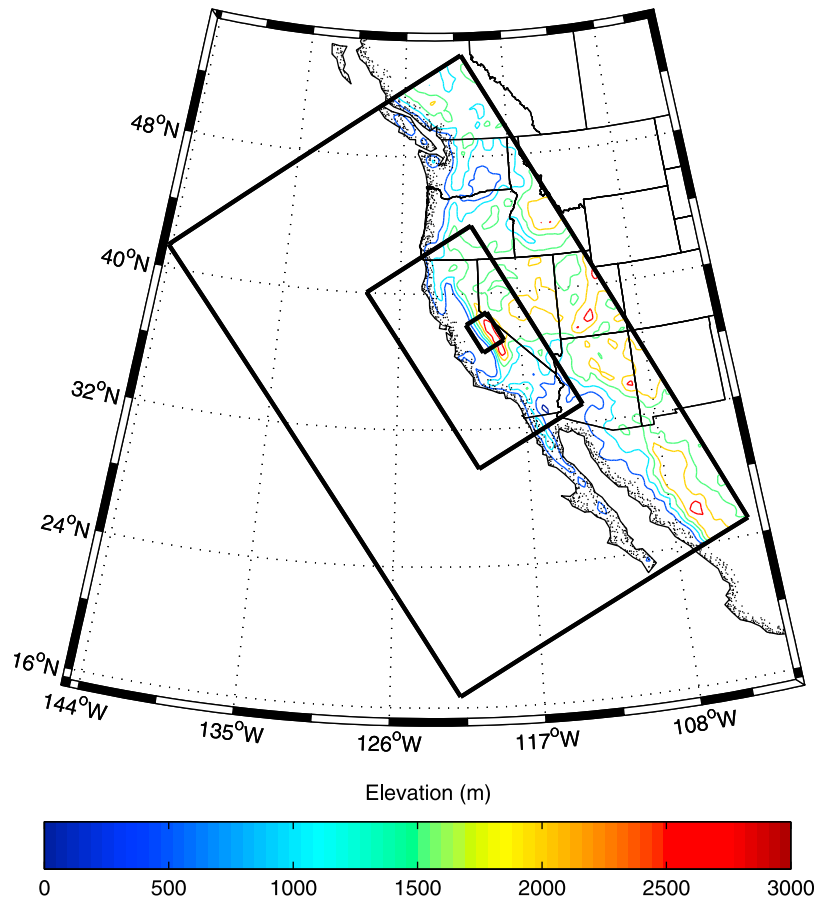


Figure 1. WRF simulation domains at 27, 9, and 3 km (from the outer to inner grids). Elevation is given in meters at the resolution of the outermost domain.

domains are run concurrently, with larger domains processed before smaller domains at each individual time step. There is no feedback of conditions from high-resolution domains back to coarser domains. The outermost domain (D27) covers the entire western coast of the U.S. and a large portion of the adjacent Northeast Pacific to capture moisture flow into the region. Elevation at the resolution of this domain is shown as contours in Figure 1. The gross features of the regional topography are captured at this resolution, but the finer definition of individual mountain ridges is not resolved. The middle domain (D9) covers the entire California Sierra Nevada. The innermost domain (D3) covers the central portion of the Sierra Nevada, including Yosemite National Park. Figure 2, showing the region covered in D3, demonstrates the substantially more detailed topography captured at 3 km compared to the D27 topography shown in Figure 1. This is unsurprising, since the entire D3 domain (3888 3 km grid cells) is covered by just 48 D27 grid cells.

[8] We use the following physics options in WRF: the Thompson *et al.* [2008] cloud microphysics scheme, the rapid radiative transfer model longwave scheme [Mlawer *et al.*, 1997], the Dudhia shortwave scheme [Dudhia, 1989], the Yonsei University planetary boundary layer scheme [Hong *et al.*, 2006], and the modified Kain-Fritsch convection parameterization for the two outer domains [Kain, 2004; Kain and Fritsch, 1990, 1993].

[9] We utilized the Noah land surface model [Chen and Dudhia, 2001; Ek *et al.*, 2003], which has been fully coupled to the WRF atmosphere such that land surface conditions can feed back into the atmosphere. The Noah model uses a simplistic canopy model and a multilayer soil model, in which the topsoil layer is used to simulate surface soil conditions, snowpack, and the vegetation canopy. As such, Noah simulates snowpack as a single layer and does not separately consider canopy snowpack. All liquid water within the snowpack is immediately routed into the soil. However, Noah does simulate snowpack accumulation, snowpack ablation via sublimation and melting, and heat exchange between the snowpack, soil, and atmosphere [Koren *et al.*, 1999]. Noah was selected over other land surface models coupled to WRF because of its simple physics and extensive use in the recent literature [e.g., Livneh *et al.*, 2010; Barlage *et al.*, 2010; Wang *et al.*, 2010].

2.2. Observational Data

[10] Daily observations of SWE were acquired at 41 locations within the innermost (D3) WRF domain from the California Department of Water Resources Data Exchange Center (CDEC) at <http://cdec.water.ca.gov/>. These stations are mapped in Figure 2 and locations and elevations are summarized in Table 1. Twelve stations are operated by the U.S. Bureau of Reclamations, 19 by the California Department of Water Resources, 9 by the Natural Resources Conservation

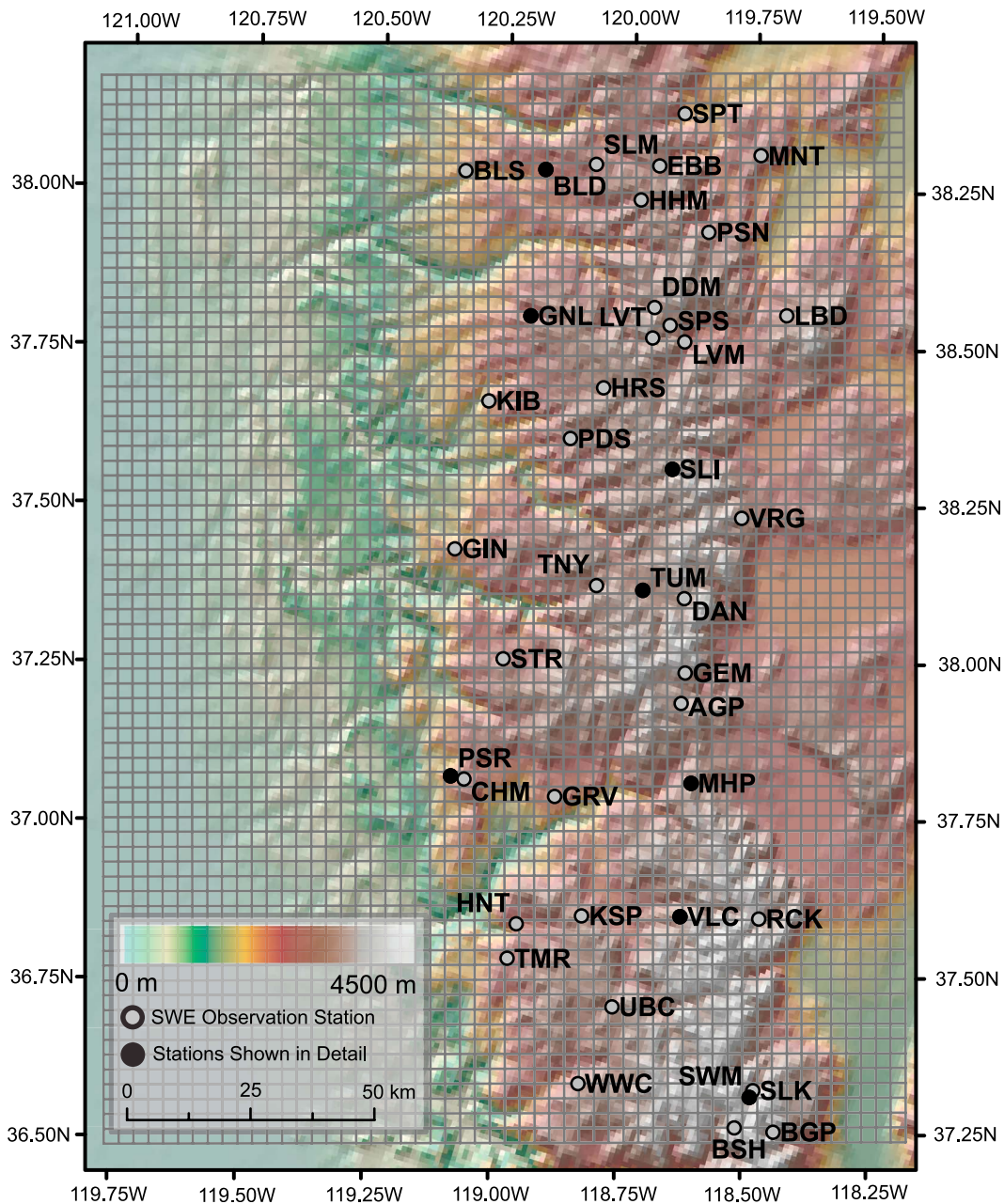


Figure 2. Map of the study area covered by the 3, 9, and 27 km model domains. The grid shows the spacing of the 3 km domain. Labeled dots indicate the locations of observation measurement stations, with those representative stations used here in Figures 5, 7, and 8 in black.

Service (NRCS), and one by the U.S. Army Corps of Engineers. Those operated by the NRCS are a part of the official network of SNOTEL stations that provide long-term snowpack measurements across the American West. At all stations, SWE observations were collected using pressure-sensing snow pillows. Observations were scanned for obvious data entry or sensor failure errors and in three cases single data points were removed and replaced with the mean of the two adjacent days. At 10 of the 41 SWE observation stations, total daily precipitation values are available for at least 95% of days during the study period (Table 1). Additional stations with intermittent precipitation data were not used here because of concerns that missing data could bias results. Finally, at 31 of

the 41 SWE observation stations, hourly measures of surface air temperature are also available (Table 1). Temperature observations were averaged to provide mean daily temperature values. Station elevations are also provided in the CDEC archive and are included in Table 1.

2.3. Analytical Methods

[11] To assess the effect of spatial resolution on modeled snowpack, we calculate the mean daily SWE for all three domains at all elevations over 1500 m within the boundaries of D3 (Figure 1). To ensure that identical areas are incorporated, D27 and D9 SWE grids are interpolated to the D3 grid using a nearest neighbor approach. Mean daily SWE is

Table 1. Summary Statistics for Observed and Modeled Snowpack^a

| Full Name | Latitude | Longitude | Elev _{Obs} (m) | Precipitation Data | Temperature Data | SCD _{Obs} ^b (days) | SCD _{D27} (days) | SCD _{D9} (days) | SCD _{D3} (days) | Accum _{Obs} ^c (mm) | Accum _{D27} (mm) | Accum _{D9} (mm) | Accum _{D3} (mm) | R _{D3} ^d | R _{Accum} | R _{Slope} |
|------------|---------------|-----------------|-------------------------|--------------------|------------------|--|---------------------------|--------------------------|--------------------------|--|---------------------------|--------------------------|--------------------------|------------------------------|--------------------|--------------------|
| SPT | 38.667 | -119.818 | 1875 | - | - | 69 | 60 | 56 | 72 | 218 | 249 | 271 | 388 | 0.97 | 0.73 | 0.79 |
| BLS | 38.375 | -120.192 | 1982 | Y | Y | 126 | 51 | 60 | 76 | 859 | 468 | 671 | 764 | 0.36 | 0.77 | -0.14 |
| KIB | 38.032 | -119.877 | 2043 | - | Y | 108 | 50 | 54 | 79 | 782 | 350 | 490 | 788 | 0.69 | 0.74 | 0.08 |
| GRV | 37.465 | -119.29 | 2104 | Y | Y | 106 | 96 | 80 | 74 | 558 | 827 | 688 | 670 | 0.66 | 0.72 | -0.02 |
| PSR | 37.403 | -119.52 | 2104 | Y | Y | 110 | 48 | 51 | 58 | 832 | 393 | 395 | 630 | 0.34 | 0.80 | -0.01 |
| HNT | 37.228 | -119.221 | 2134 | Y | Y | 110 | 88 | 57 | 72 | 610 | 672 | 461 | 562 | 0.53 | 0.84 | -0.08 |
| GIN | 37.767 | -119.773 | 2149 | - | Y | 121 | 49 | 74 | 82 | 742 | 379 | 725 | 792 | 0.55 | 0.70 | -0.03 |
| CHM | 37.41 | -119.49 | 2180 | Y | Y | 119 | 48 | 82 | 93 | 930 | 393 | 747 | 906 | 0.75 | 0.85 | 0.31 |
| BLD | 38.45 | -120.033 | 2195 | Y | Y | 121 | 97 | 111 | 105 | 878 | 839 | 1234 | 997 | 0.88 | 0.78 | 0.42 |
| LVM | 38.305 | -119.552 | 2195 | - | - | 95 | 92 | 83 | 88 | 284 | 410 | 323 | 456 | 0.97 | 0.74 | 0.89 |
| TMR | 37.165 | -119.2 | 2302 | Y | Y | 108 | 89 | 85 | 64 | 709 | 691 | 670 | 497 | 0.43 | 0.78 | -0.14 |
| PDS | 38.047 | -119.67 | 2332 | - | Y | 124 | 101 | 99 | 120 | 1257 | 909 | 960 | 1492 | 0.99 | 0.78 | 0.89 |
| SLM | 38.505 | -119.937 | 2363 | - | Y | 129 | 97 | 116 | 114 | 1536 | 839 | 1301 | 1270 | 0.92 | 0.80 | 0.67 |
| PSN | 38.501 | -119.631 | 2409 | - | Y | 107 | 75 | 106 | 91 | 498 | 270 | 745 | 485 | 0.86 | 0.80 | 0.33 |
| TNY | 37.838 | -119.448 | 2485 | - | Y | 117 | 101 | 106 | 111 | 894 | 903 | 889 | 1193 | 0.97 | 0.85 | 0.80 |
| STR | 37.637 | -119.55 | 2500 | Y | Y | 120 | 101 | 74 | 90 | 896 | 903 | 501 | 740 | 0.66 | 0.76 | -0.03 |
| MNT | 38.67 | -119.615 | 2546 | - | - | 108 | 75 | 52 | 58 | 409 | 270 | 152 | 176 | 0.38 | 0.81 | -0.19 |
| GNL | 38.205 | -119.892 | 2561 | Y | Y | 128 | 101 | 105 | 115 | 1086 | 923 | 1169 | 1297 | 0.90 | 0.82 | 0.44 |
| HRS | 38.158 | -119.662 | 2561 | - | Y | 128 | 101 | 121 | 117 | 1230 | 909 | 1309 | 1207 | 0.94 | 0.77 | 0.71 |
| TUM | 37.873 | -119.35 | 2622 | - | Y | 106 | 90 | 114 | 109 | 509 | 372 | 832 | 920 | 0.97 | 0.88 | 0.63 |
| EBB | 38.561 | -119.808 | 2652 | - | - | 124 | 75 | 108 | 106 | 1090 | 270 | 966 | 945 | 0.85 | 0.85 | 0.53 |
| HHM | 38.49 | -119.805 | 2652 | - | Y | 129 | 75 | 116 | 120 | 1400 | 270 | 1141 | 1390 | 0.95 | 0.80 | 0.78 |
| SPS | 38.318 | -119.601 | 2668 | - | - | 124 | 92 | 105 | 113 | 653 | 410 | 754 | 871 | 0.95 | 0.73 | 0.82 |
| WVC | 37.03 | -118.918 | 2774 | - | Y | 117 | 83 | 91 | 95 | 876 | 639 | 803 | 826 | 0.83 | 0.88 | 0.40 |
| KSP | 37.30 | -119.10 | 2805 | Y | Y | 111 | 88 | 83 | 100 | 746 | 672 | 555 | 844 | 0.95 | 0.88 | 0.69 |
| LBD | 38.44 | -119.377 | 2805 | - | - | 110 | 53 | 72 | 78 | 396 | 73 | 157 | 236 | 0.62 | 0.77 | -0.03 |
| SLI | 38.092 | -119.43 | 2805 | - | Y | 129 | 94 | 124 | 121 | 1148 | 455 | 1314 | 1205 | 0.97 | 0.89 | 0.81 |
| DDM | 38.332 | -119.653 | 2820 | - | Y | 128 | 92 | 105 | 120 | 790 | 410 | 754 | 1050 | 0.93 | 0.69 | 0.84 |
| MHP | 37.61 | -119.033 | 2835 | Y | Y | 127 | 89 | 109 | 114 | 975 | 330 | 810 | 1055 | 0.92 | 0.87 | 0.66 |
| VRG | 38.077 | -119.234 | 2835 | - | - | 120 | 94 | 80 | 95 | 414 | 455 | 273 | 542 | 0.73 | 0.83 | 0.08 |
| AGP | 37.728 | -119.143 | 2881 | - | Y | 111 | 89 | 91 | 100 | 757 | 330 | 458 | 679 | 0.93 | 0.78 | 0.54 |
| LVT | 38.282 | -119.621 | 2927 | - | - | 139 | 92 | 105 | 113 | 1549 | 410 | 754 | 905 | 0.68 | 0.87 | 0.06 |
| SLK | 37.176 | -118.562 | 2927 | - | Y | 111 | 88 | 97 | 90 | 398 | 322 | 389 | 313 | 0.85 | 0.78 | 0.60 |
| UBC | 37.183 | -118.937 | 2957 | - | Y | 120 | 89 | 98 | 99 | 976 | 691 | 693 | 720 | 0.79 | 0.87 | 0.17 |
| BGP | 37.128 | -118.475 | 2988 | - | Y | 106 | 31 | 62 | 84 | 379 | 19 | 151 | 271 | 0.81 | 0.84 | 0.10 |
| DAN | 37.897 | -119.257 | 2988 | - | - | 118 | 90 | 103 | 116 | 701 | 372 | 584 | 790 | 0.97 | 0.93 | 0.59 |
| RCK | 37.455 | -118.743 | 3049 | - | Y | 108 | 34 | 104 | 97 | 315 | 27 | 537 | 447 | 0.94 | 0.70 | 0.79 |
| VLC | 37.388 | -118.903 | 3064 | - | Y | 124 | 85 | 110 | 96 | 764 | 306 | 795 | 556 | 0.62 | 0.89 | -0.07 |
| SWM | 37.162 | -118.562 | 3110 | - | Y | 109 | 88 | 97 | 102 | 389 | 322 | 389 | 452 | 0.96 | 0.82 | 0.64 |
| GEM | 37.78 | -119.17 | 3277 | - | Y | 126 | 89 | 91 | 108 | 813 | 330 | 458 | 703 | 0.80 | 0.74 | 0.24 |
| BSH | 37.1 | -118.557 | 3415 | - | Y | 130 | 88 | 110 | 111 | 614 | 322 | 539 | 668 | 0.85 | 0.66 | 0.42 |
| Mean | | | 2607 | - | - | 117 | 81 | 91 | 97 | 777 | 473 | 678 | 773 | 0.80 | 0.80 | 0.38 |

^aBold lines denote the eight representative stations used in Figures 5, 9, and 10.

^bSCD refers to the snowpack centroid date.

^cAccum refers to total snowpack accumulation.

^dR refers to correlation coefficients between observed and modeled snowpack, accumulation, and melt.

also computed for all model domains in four elevation classes (1500–2000 m, 2000–2500 m, 2500–3000 m, and 3000+ m). D3 elevations are used to determine the elevation class for each grid cell in all model domains to ensure consistency. Comparison among elevation classes allows us to determine whether the influence of spatial resolution on modeled snowpack varies with terrain height.

[12] To compare observations and modeled outputs, it is first necessary to develop a method for comparison of point-based and gridded data sets. In this case, we simply compare each observation station to the model grid cell in which it falls. An inverse distance weighting approach was also tested, in which values for the four nearest model grid cells were weighted according to the distance between the observation point and the grid cell center. This approach provided no discernible advantage over our simpler approach and was abandoned. At individual observation stations, model SWE error may increase with increasing model grid spacing as grid cell parameters such as elevation become less and less representative of observation locations.

[13] At each station, we compare modeled and observed snowpack using six distinct metrics:

[14] 1. To gauge model skill in determining snowpack timing we calculate the SWE centroid date (SCD) of accumulated SWE through the entire snowpack season in the observations and in each model domain. The SCD is defined as the date corresponding to center of mass of the annual snowpack and is measured in days from 1 November 2001, the beginning of our simulation period. Observational analysis has determined that SCD anomalies are roughly proportional to anomalies in the peak date of SWE (see *Kapnick and Hall* [2010] for more details). SCD is also similar in principal to center timing, a commonly used metric of streamflow seasonality [*Maurer*, 2007; *Maurer et al.*, 2007; *Stewart et al.*, 2005, 2004].

[15] 2. To directly compare skill in simulating snowpack magnitude, we determine total observed and modeled SWE accumulation for each model domain by summing daily SWE increases throughout the snow season. These values may exhibit a slight low bias because melt occurring on the same day as accumulation would reduce apparent accumulation. However, this effect is expected to be small since melt and accumulation events are generally independent of one another.

[16] 3. We calculate the Pearson's correlation coefficient (r) for the paired time series of observed and D3 SWE. Correlations are calculated for each station over the time period between the first day on which either observed or modeled SWE is greater than 1.0 mm and the last day on which SWE is greater than 1.0 mm. This metric provides an indication of model skill, but some caution is warranted because both observed and modeled time series exhibit a high degree of serial autocorrelation and because r is insensitive to differences in time series magnitude.

[17] 4. We compare observed and WRF-derived precipitation in order to determine whether any biases in modeled snowpack are related to an overabundance or under abundance of precipitation. Daily precipitation totals are available at 10 of the stations used here, and we use a paired Student's t test to determine whether WRF-derived annual precipitation totals differ significantly from observations at these locations. In addition, we separately compare D3, D9,

and D27 precipitation at low elevations (<2000 m) and high elevations (>2000 m) to determine whether model resolution has any effect on total precipitation via greater topographic detail at higher resolutions. If so, then we expect to observe similar levels of precipitation at low altitudes in all model domains and substantially greater precipitation at high altitudes at finer grid spacings due to the effects of orography. If not, then either lower precipitation in D27 at high altitudes will be offset by greater precipitation at low altitudes, or all domains will generate similar amounts of precipitation at all altitudes.

[18] 5. We compare time series of snowpack accumulation (i.e., snowfall) and ablation (i.e., melt and sublimation) from WRF-Noah and observations. Despite the fact that they are both sensitive to temperature, the physics associated with accumulation and ablation processes are largely distinct. In WRF version 3.1.1, precipitation, the principal form of accumulation, is simulated by WRF directly while ablation processes are handled in Noah. As a result, we divide modeled and observed time series into periods of accumulation and ablation in order to assess model skill in tracking each process separately. Accumulation days are defined as those where measured SWE is greater than the previous day, while ablation days are those where accumulated SWE has declined relative to the previous day. Days on which there is both accumulation and ablation are not captured, but these are likely to be a small subset of accumulation and ablation days. We calculate Pearson's correlations between observed and simulated accumulation and ablation time series over the highest-resolution domain (D3) to determine model skill in these variables.

[19] 6. Because Noah predicts snowpack evolution using an energy-mass balance scheme [*Ek et al.*, 2003; *Feng et al.*, 2008; *Wang et al.*, 2010], it is important to evaluate the model's ability to correctly track the observed energy balance. Unfortunately, a full suite of daily energy balance observations is unavailable at most snow pillow observation sites in the Sierra Nevada, with the exception of the Mammoth Mountain Energy Balance Monitoring Site (<http://www.snow.ucsb.edu/cues/description.html>). Still, the extent and timing of differences between observed and modeled surface air temperature may provide indicators of strengths and weaknesses of energy balance calculations in WRF-Noah. As we show in section 3.2, a comparison of the time series of modeled and observed surface air temperature allows us to conduct a preliminary assessment of errors in the snowpack energy balance in WRF-Noah. Because this metric differs substantially from the first five, we will consider it separately in section 3.2.

3. Results

3.1. Comparing Simulated and Observed Snowpack Quantity and Timing

[20] A comparison of simulated area-averaged daily mean SWE values over the course of the 2001–2002 snow season at 27, 9, and 3 km within the boundaries of D3 (Figure 3a) indicates a substantial difference between D27 and the two higher-resolution grids but little difference between D9 and D3. Figure 3b shows time series of SWE averaged across 41 observation stations listed in Table 1 as well as simulated SWE in the corresponding model grid cells at 3, 9, and

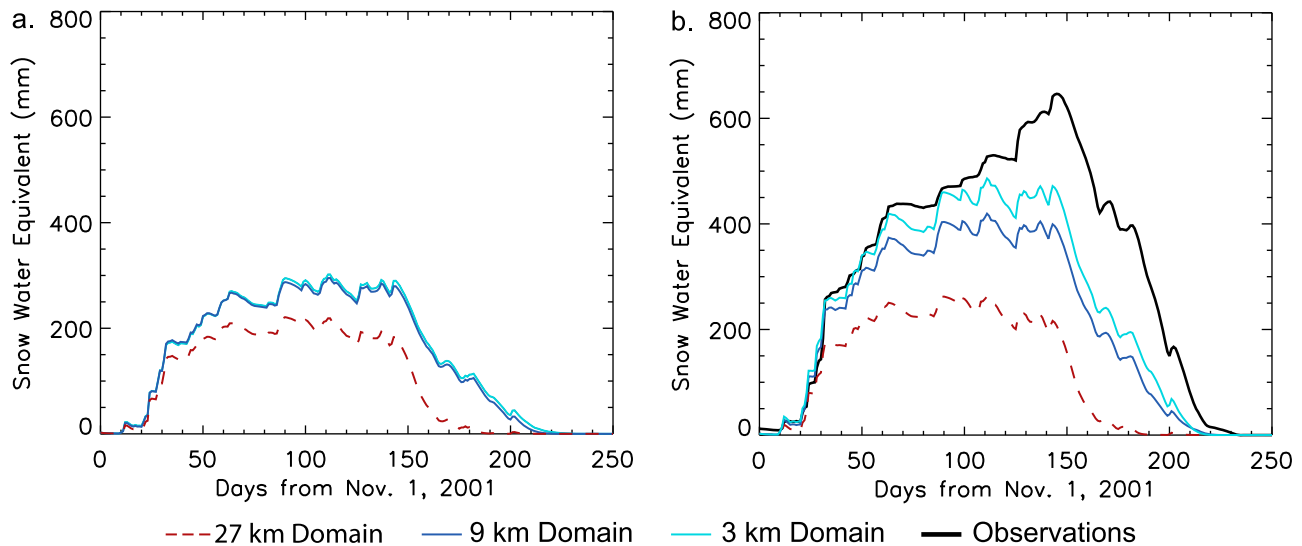


Figure 3. (a) Mean WRF simulated snowpack for all areas above 1500 m at 3, 9, and 27 km grid spacings. (b) Mean snowpack across 41 snowpack observation stations shown in Table 1 and the corresponding model grid cells at 3, 9, and 27 km grid spacings. Differences between 9 and 3 km are comparatively small, while snowpack is substantially lower in the 27 km simulation. Snowpack at all model resolutions is lower than observations, especially during the latter half of the snow season.

27 km resolutions. Figure 3 should be interpreted with caution, since observed and modeled snowpack may represent variability at different spatial scales, as noted in section 2.3, and are not directly comparable. Still, the comparison suggests that all model domains generally underestimate overall SWE, with D3 being the most realistic. During the early accumulation portion of the season, simulated SWE in the 3 km and 9 km solutions tracks the observations well. At about day 50, melt events begin to occur in the simulated time series that are either weaker or nonexistent in the observations. As melting becomes more common later in the season, simulated and observed time series diverge further, with the 3 km and 9 km solutions continuing to deliver better performance than the 27 km solution.

[21] Examination of SWE values within 500 m elevation bands (Figure 4) provides additional information about the performance of the three simulations as a function of elevation. Values of D27 SWE are substantially smaller than D3 and D9 SWE at high elevations but are comparable to, or even somewhat greater than, D9 and D3 SWE at lower elevations. This pattern is likely due to the fact that in some areas adjacent to high-elevation regions D27 elevations are higher than D3 or D9 elevations due to greater smoothing of topography at lower spatial resolution. By contrast, D9 and D3 SWE are relatively similar at all elevations, with the greatest differences occurring at elevations greater than 3000 m, where D3 SWE is about 10% greater than D9 SWE.

[22] A suite of statistics quantifying model skill in simulating snowpack evolution for the 2001–2002 snow season is provided in Table 1, and the complete snowpack time series for the D27, D9, and D3 are shown in Figure 5 for eight representative stations (mapped in Figure 2). While the fidelity of simulations exhibits substantial variability from station to station, at most stations D3 and D9 SWE are closer to observations than is D27. Additionally, at most stations (TUM excepted) simulated snowmelt generally

occurs earlier in the season than does observed snowmelt. Mean observed SCD for all stations occurs on day 117 (24 February), while mean modeled SCD occurs on days 81, 91, and 97 for D27, D9, and D3, respectively (20 and 30 January and 6 February). However, at some individual stations observed and modeled SCD match closely, especially for D9 and D3, while at other stations modeled and observed SCD differ by more than 50 days. Simulation of SCD generally improves with increasing model resolution, but even the 3 km solution exhibits an approximately 21 day bias toward early SCD.

[23] Precipitation is among the most plausible sources of bias in WRF-Noah snowpack. A substantial underestimate of total precipitation could potentially explain observed negative snowpack biases. Figure 6a shows the differences between observed and modeled total precipitation during the study period at 10 observation stations. The mean value of modeled minus observed precipitation is 123 mm (11.9%) for D3, 28 mm (2.7%) for D9, and -169 mm (-16.4%) for D27. None of these differences is statistically significant at $p = 0.05$, which may be due to a small sample size and substantial spread in the data. However, D3 and D27 observations are significantly different from each other $p = 0.05$. Mean correlations between daily observed and D3 ($r = 0.79$), D9 ($r = 0.79$), and D27 ($r = 0.77$) precipitation at all ten observation stations suggest that WRF successfully reproduces observed precipitation timing, which is likely due to the accuracy of the NARR boundary conditions and WRF physics parameterizations. Figures 6b and 6c show mean WRF precipitation for each model domains in portions of the D3 region below 2000 m (Figure 6b) and above 2000 m (Figure 6c). At low elevations, model resolution has almost no impact on precipitation, while at high elevations D27 precipitation is substantially lower than D9 and, especially, D3 precipitation. This contrast suggests that greater topographic detail available in mountainous landscapes in

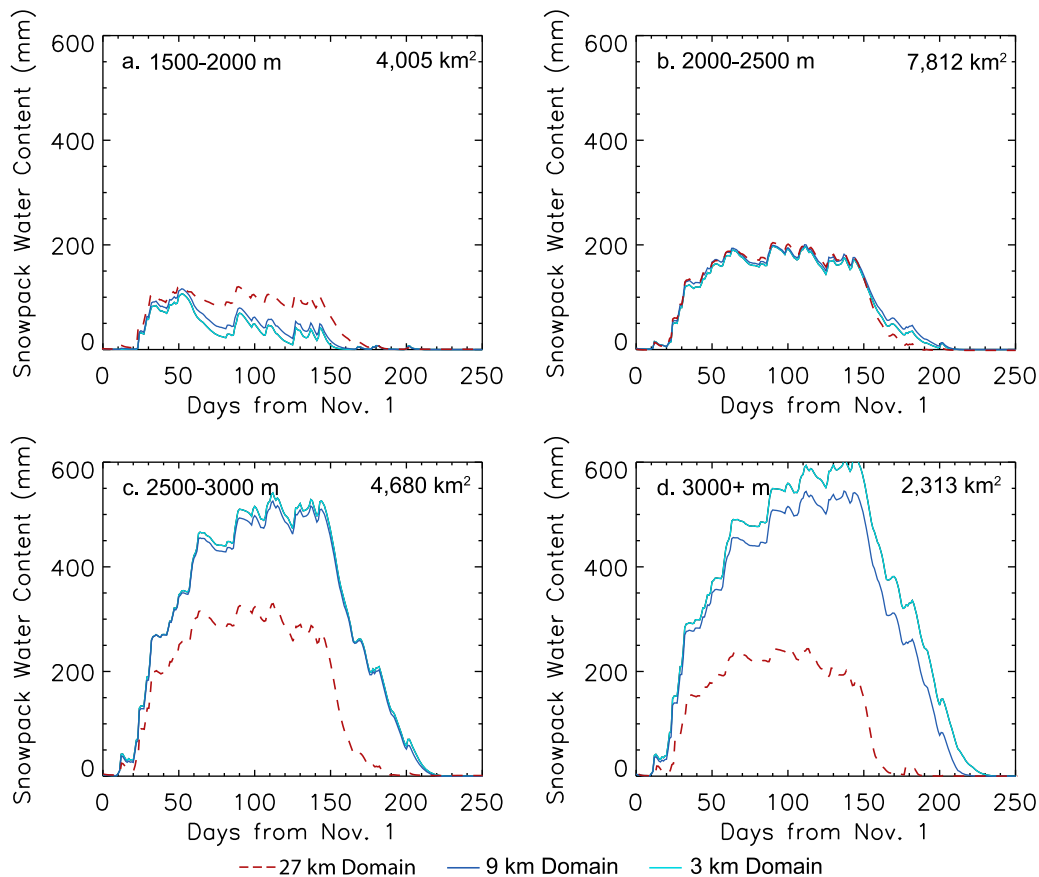


Figure 4. (a–d) Mean WRF simulated snowpack for four 500 m elevation windows at 3, 9, and 27 km grid spacings. Differences between 9 and 3 km are small for all windows, while snowpack is substantially smaller in the 27 km simulation only at high elevations. Areas for each elevation band are calculated on the basis of 3 km domain elevation values.

D3 and D9 results in greater orographic precipitation, while at low elevations precipitation is largely independent of model resolution.

[24] Comparison of observed and modeled total snow accumulation also suggests a substantial resolution dependence of model snowpack accuracy (Table 1). Mean observed SWE accumulation across the 41 test sites is 777 mm, and D3 SWE in the corresponding model grid cells matches this very closely at 773 mm. By contrast, D9 (678 mm) and especially D27 (473 mm) each generally underestimate total accumulation. At the highest resolution, WRF simulates accumulation timing with a high degree of fidelity as well. The mean correlation coefficient between observed and D3 accumulation is 0.80, with a range of 0.66 to 0.93 (Table 1). In most cases, simulations at lower resolutions also reproduce accumulation timing accurately, though somewhat less so than at D3 resolution. The mean correlations for D9 and D27 are 0.76 and 0.73, with ranges of 0.57 to 0.90 and 0.13 to 0.92, respectively.

[25] In contrast to snow accumulation, even D3 simulates snowmelt quite poorly at many observation stations. The mean correlation coefficient between observed and D3 ablation events is 0.38, with a range of -0.19 to 0.89 (Table 1). D9 and D27 simulations are even poorer, with mean values of 0.23 and 0.05, and ranges of -0.26 to 0.86 and -0.29 to 0.79, respectively. This wide disparity between the fidelity of

modeled ablation and accumulation suggests that the primary source of error in WRF simulations of Sierra Nevada snowpack is simulation of ablation rather than accumulation. Correlations between observed and D3 total SWE range from 0.34 to 0.99, with a mean of 0.80 (Table 1), suggesting that despite major problems with simulation of ablation, high-resolution WRF-Noah simulations match observed variations in total SWE at many stations. Correlation coefficients for total SWE at individual stations are themselves highly correlated with the difference between observed and modeled snowpack SCD ($r > 0.93$ for all domains). Thus the correlation coefficient is essentially an additional measure of how well the simulation reproduces snowpack timing. Consequently, we do not include D9 and D27 snowpack correlations here and include D3 correlations principally to demonstrate the similarity to SCD differences.

[26] One potential source of error in WRF simulation of ablation is the effect of smoothing model elevations at low resolutions. If WRF elevations in mountain environments are, on average, substantially lower than observed elevations, it is likely that snowmelt will occur too early. Figure 7 shows scatterplots of model errors in SCD against model errors in elevation. In each domain, calculation of Pearson's correlation coefficients reveals a statistically significant ($p = 0.05$) and approximately linear relationship between SCD error and elevation error. Regression slopes are nearly identical in all

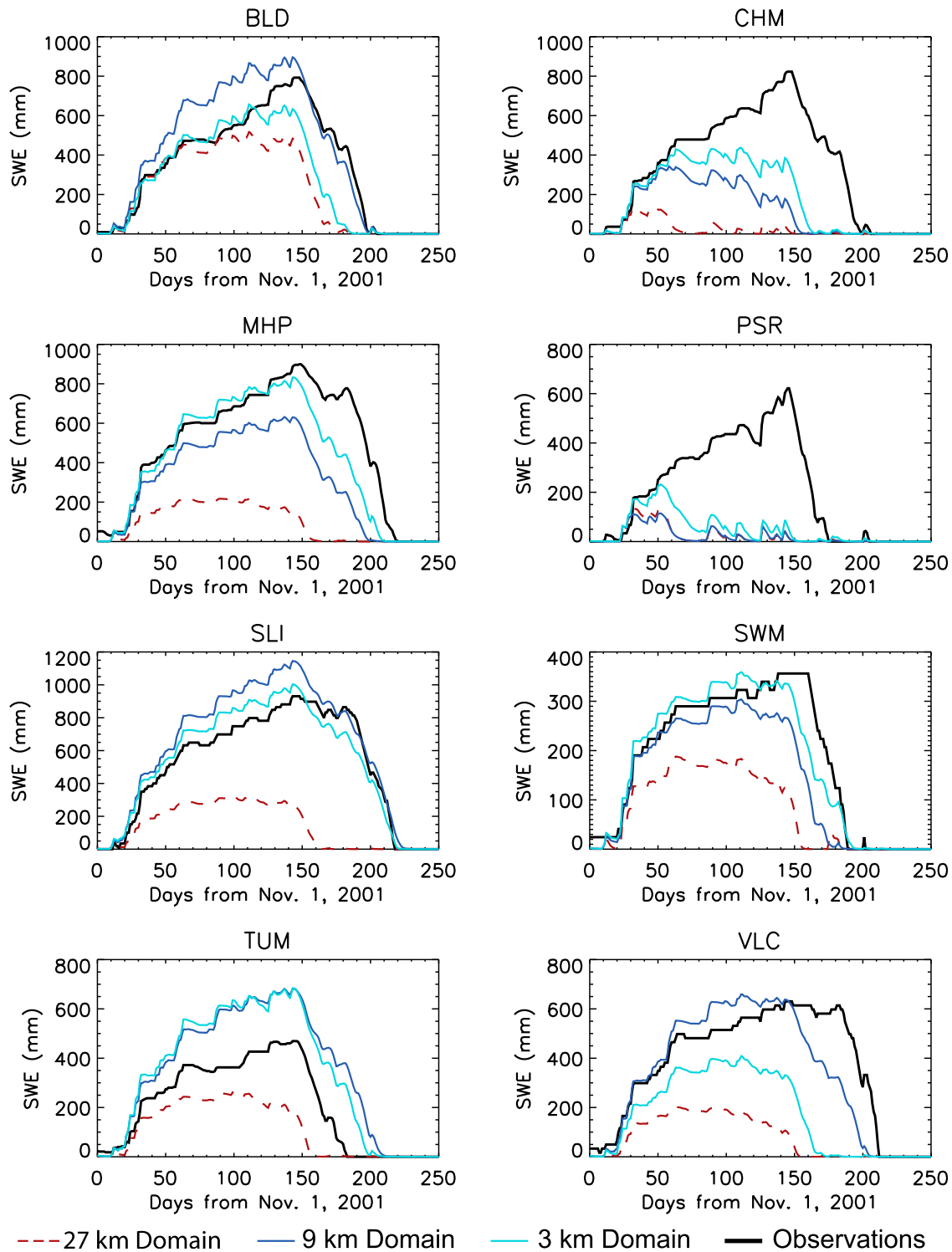


Figure 5. Daily SWE measurements for three WRF model domains and observations at eight representative stations (see Table 1 for summary statistics for all stations). In general, 3 and 9 km domains better match observations than does the 27 km domain. At most stations, all domains exhibit a bias toward early snowmelt relative to observations.

domains. In D27 there is a clear systematic bias in model elevation (mean model elevation is 206 m lower than observed elevation). Given the slope of the elevation error/SCD error relationship, this bias can account for approxi-

mately 11 days of the 36 day early bias in SCD in the 27 km solution. In contrast, D9 and D3 have much smaller mean elevation biases of 11 and -44 m, respectively. Thus, the elevation bias can account for very little of the SCD bias in

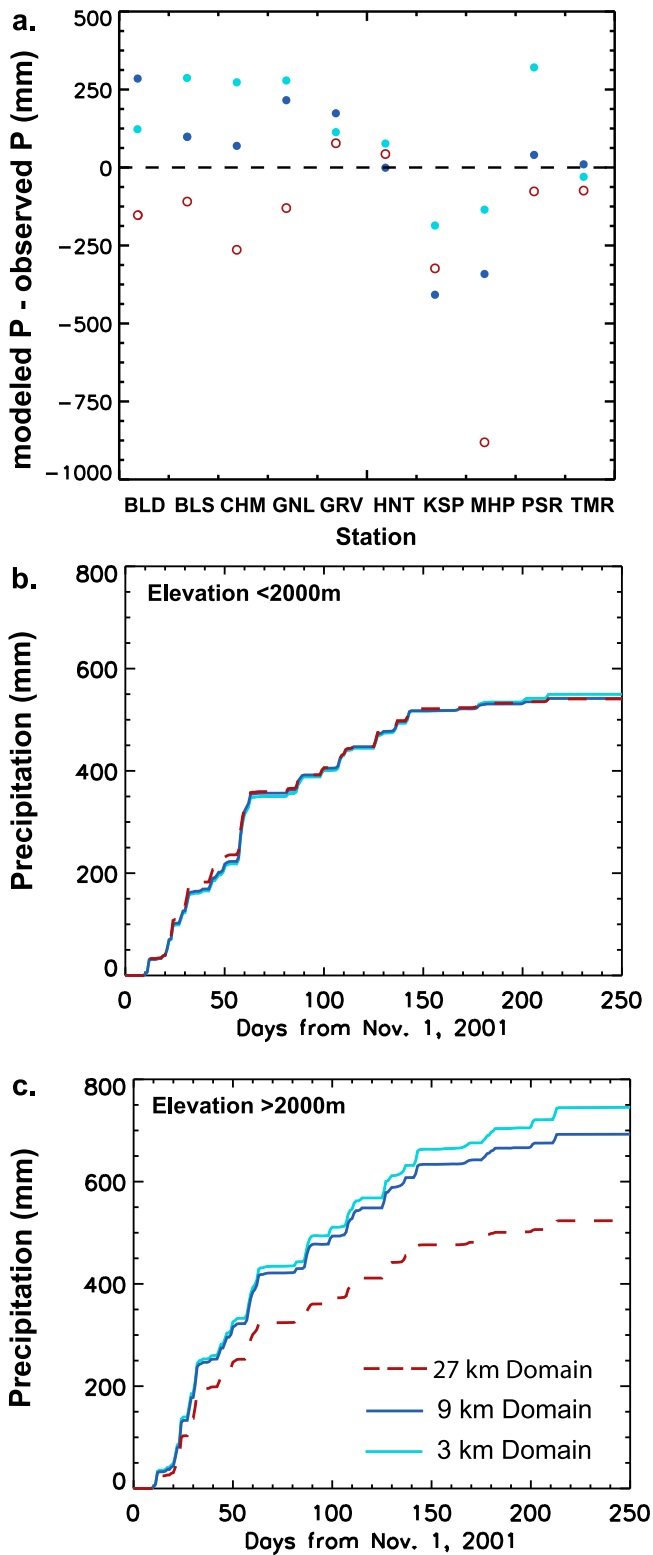


Figure 6. (a) Total WRF precipitation minus observed precipitation at 10 stations with complete precipitation records. (b) Mean cumulative precipitation in portions of the 3 km WRF domain below 2000 m. (c) Mean cumulative precipitation in portions of the 3 km WRF domain above 2000 m.

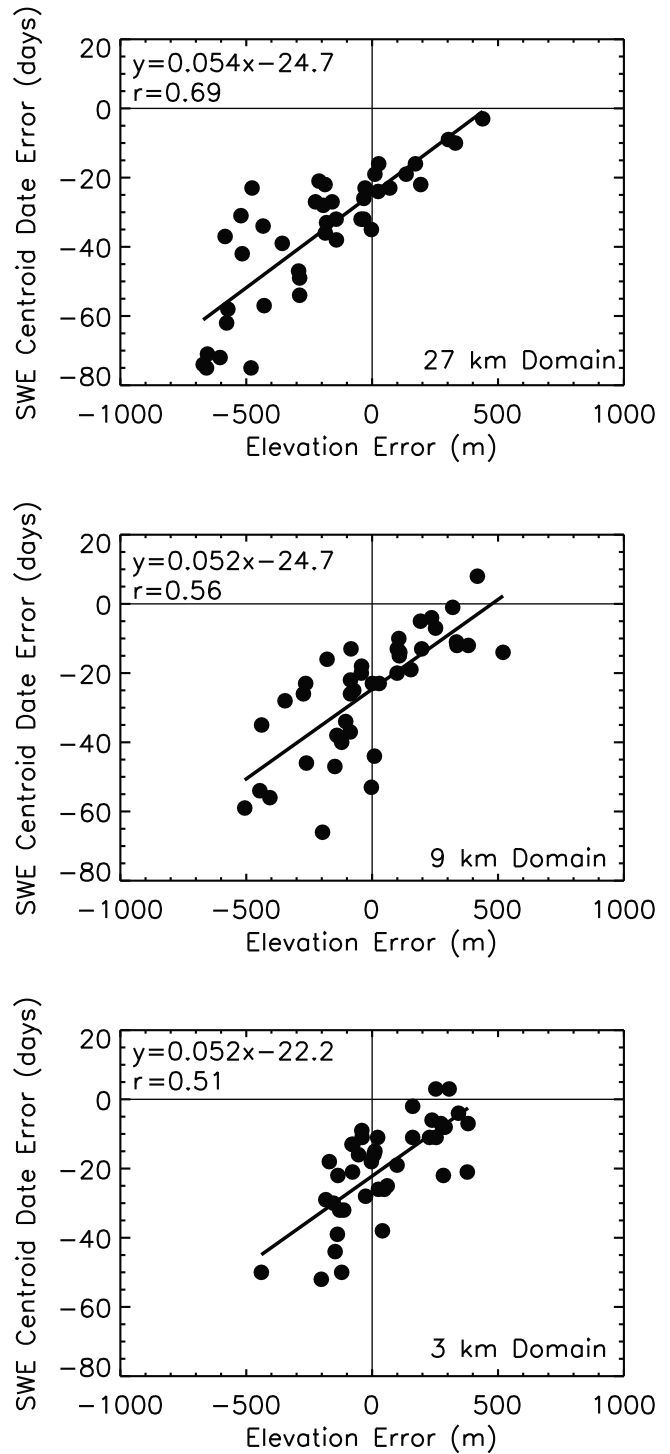


Figure 7. Scatterplots comparing the error in simulated SWE centroid date (SCD) and the error in simulated elevations relative to station observations. Each point represents one observation station (Table 1). Scatterplots are for three different model grid spacings: (top) 27 km, (middle) 9 km, and (bottom) 3 km. In all domains, there is a substantial relationship between elevation error and snowpack timing error. Also, each domain exhibits a significant bias toward early SCD.

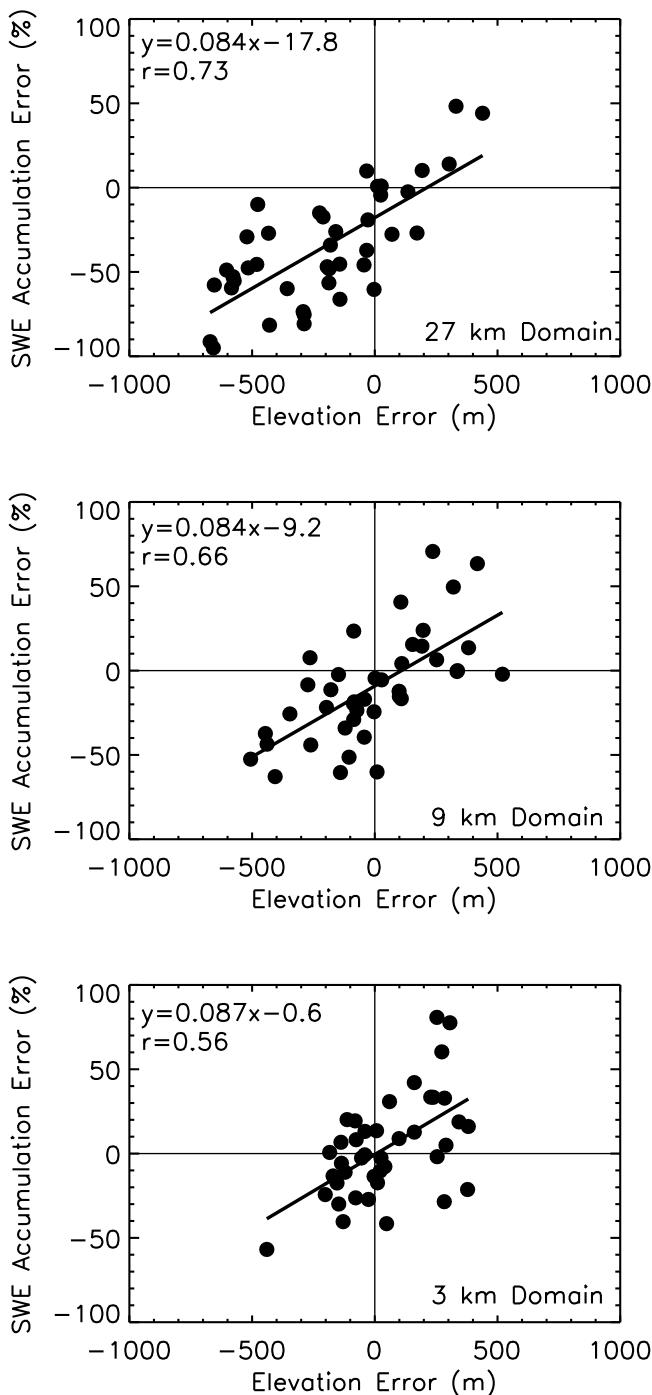


Figure 8. Scatterplots comparing the error in simulated SWE accumulation and the error in simulated elevations relative to station observations. Each point represents one observation station (Table 1). Scatterplots are for three different model grid spacings: (top) 27 km, (middle) 9 km, and (bottom) 3 km. In all domains, there is a substantial relationship between SWE accumulation and elevation errors. However, the low bias in simulated accumulation is diminished as model resolution increases.

these simulations. For all three domains, the y intercept value of the linear regression equations in Figure 7 is strikingly similar, and indicates that even with no elevation error

WRF will underestimate SCD by, on average, between 22 and 25 days. There is therefore a systematic bias in the WRF-Noah framework that would not disappear even if topography were resolved perfectly.

[27] We perform a similar analysis for error in SWE accumulation to understand the impact of topographic fidelity on simulated snowpack. Figure 8 shows scatterplots of model errors in SWE accumulation (expressed as a percent of observed accumulation) against model elevation errors. In all domains, there is a strong relationship ($r > 0.56$) between accumulation errors and elevation errors, though the strength of the relationship decreases with increasing model resolution. D27 and D9 show systematic underestimates of SWE accumulation of -17.8% and -9.2% , respectively, while D3 shows virtually no systematic bias (-0.6%) in accumulation. This suggests that improved model resolution substantially improves simulation of accumulation up to 3 km. Further resolution increases are unlikely to be necessary to regional SWE accumulation accurately, though they may improve SWE accumulation at individual sites as elevation errors and simulation-scale discrepancies are reduced further.

3.2. Evaluation of the Snowpack Energy Balance in WRF-Noah

[28] Because Noah computes snowpack ablation on the basis of an energy-mass balance scheme, a likely source of the systematic bias toward early snowmelt is in its simulation of the energy balance during melt events. Snowmelt in the Sierra Nevada is most responsive to increases in downwelling solar radiation associated with lengthening days during the spring season [Marks and Dozier, 1992]. This process is partially mediated by decreases in snowpack albedo associated with aging and partial melting of snowpack surfaces [Flanner and Zender, 2006]. Versions of WRF-Noah prior to 3.1 contained very simplistic snowpack albedo schemes, and analysis by Livneh *et al.* [2010] and Wang *et al.* [2010] suggested that improved simulation of snowpack albedo was necessary to improve model snowmelt. A new time-varying albedo scheme developed by Livneh *et al.* [2010] is included in WRF-Noah v.3.1, and subsequent analysis including this modification shows substantial improvements in snowpack simulation [Barlage *et al.*, 2010]. In addition, however, recent evaluations of WRF-Noah snowpack also suggest that modification of other elements of the Noah snowpack-energy balance scheme can also lead to substantial improvements in snowpack simulation, including surface-atmosphere energy exchange, effects of vegetation on surface roughness, and simulation of wind speed [Livneh *et al.*, 2010; Wang *et al.*, 2010; Barlage *et al.*, 2010]. Many weaknesses in WRF-Noah associated with these proposed changes are related to accuracy of turbulent heat fluxes between the snowpack and the atmosphere.

[29] The most widely available observational metric providing information related to snowpack energy balance is surface air temperature, which we use here to infer linkages between the energy balance and snowmelt in WRF-Noah. Evaluation of modeled albedo using point-scale measurements is far more problematic, and we do not attempt to do so here. A comparison of D3 and observed daily mean 2 m air temperatures shows that day-to-day and seasonal variations in WRF temperature closely match observations in most cases (Figure 9). The mean correlation between observed and

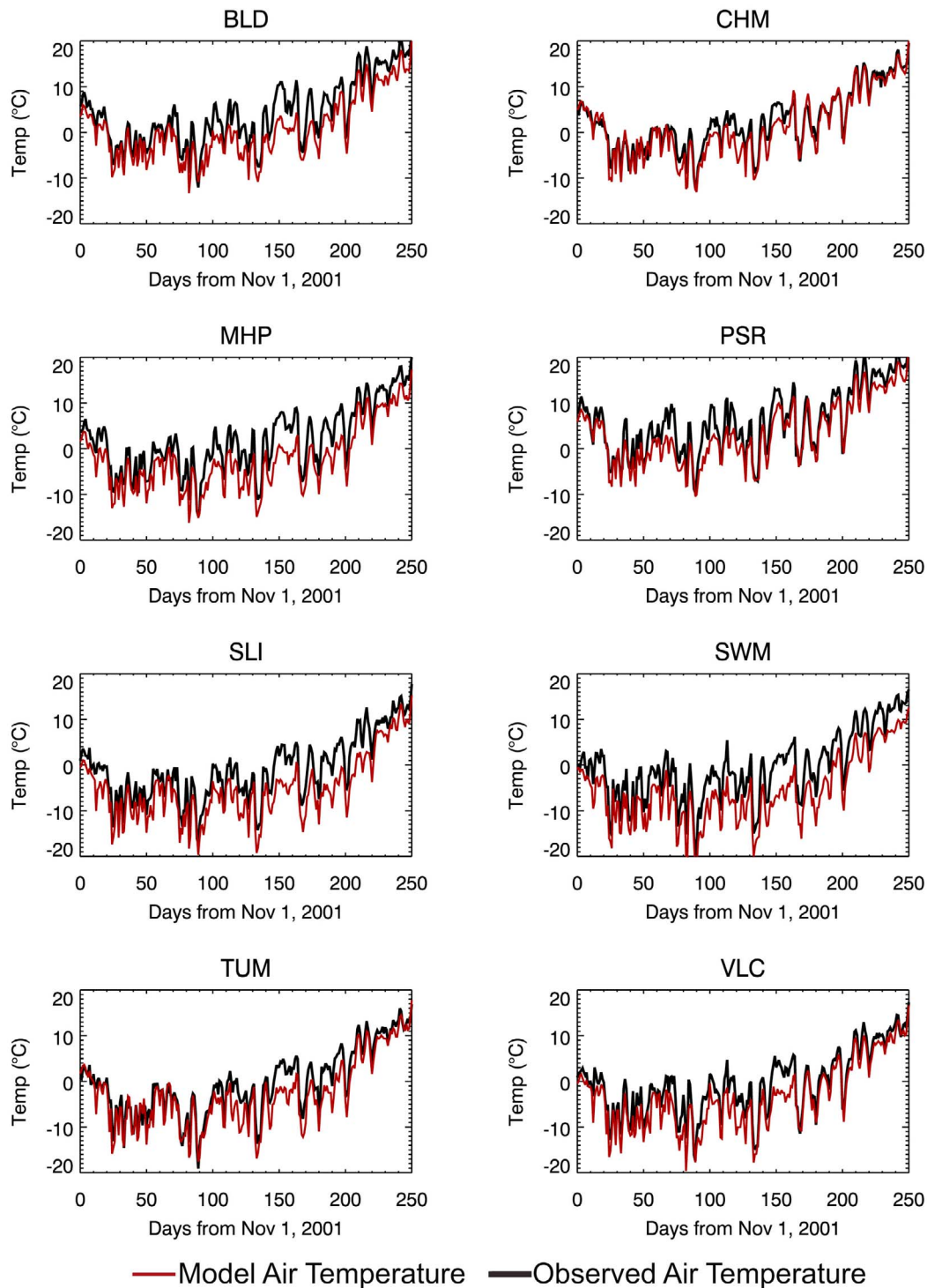


Figure 9. Observed mean daily surface air temperature ($^{\circ}\text{C}$) at eight representative stations shown in black and 3 km WRF mean daily surface air temperature shown in red. Time series are highly correlated, but WRF temperatures are sometimes substantially lower than observations.

D3 mean daily 2 m temperatures at all 31 stations with temperature data is $r = 0.88$ (individual station correlations not shown). However, D3 temperatures exhibit a negative bias relative to observations of between -1.1 and -5.6 K at 28 of 31 stations. Two stations showed substantial positive biases (2.9, 4.5 K), while one station showed a negative bias of less

than -1 K (Table 2). The overall mean bias is -2.6 K. However, this bias is not evenly distributed throughout the snowpack season. A comparison of temperature error in D3 and modeled snowmelt timing shows a strong correspondence between melt events and strongly negative temperature errors (Figure 10). Indeed, the mean temperature error during

Table 2. Observed and 3 km Modeled Elevations and Mean Errors in Model Temperature

| | Elev _{obs} (m) | Elev Diff ^a D3 (m) | Elev Diff ^a D9 (m) | Elev Diff ^a D27 (m) | Mean D3 Temperature Error (K) | R _{mte} ^b |
|------|-------------------------|-------------------------------|-------------------------------|--------------------------------|-------------------------------|-------------------------------|
| BLS | 1982 | -121 | -197 | -481 | -1.86 | 0.75 |
| KIB | 2043 | -184 | -446 | -573 | -1.52 | 0.90 |
| GRV | 2104 | -113 | -86 | 331 | -1.94 | 0.88 |
| PSR | 2104 | -202 | -506 | -578 | -2.50 | 0.68 |
| HNT | 2134 | 41 | -3 | 193 | -1.91 | 0.80 |
| GIN | 2149 | -138 | -148 | -604 | -1.95 | 0.93 |
| CHM | 2180 | 25 | -87 | -654 | -1.31 | 0.66 |
| BLD | 2195 | 7 | 106 | 24 | -3.50 | 0.67 |
| TMR | 2302 | -147 | 28 | 135 | -2.49 | 0.72 |
| PDS | 2332 | 343 | -73 | 70 | -4.52 | 0.62 |
| SLM | 2363 | 11 | 100 | -143 | -1.48 | 0.80 |
| TNY | 2485 | 238 | 335 | 26 | -2.51 | 0.67 |
| STR | 2500 | -154 | -261 | 11 | -0.38 | 0.88 |
| GNL | 2561 | -80 | -264 | -225 | -3.30 | 0.69 |
| HRS | 2561 | 254 | 252 | -159 | -2.83 | 0.57 |
| TUM | 2622 | 253 | 418 | 172 | -2.47 | 0.66 |
| HHM | 2652 | -41 | -84 | -288 | -2.78 | 0.64 |
| WWC | 2774 | -136 | -273 | -433 | -2.37 | 0.60 |
| KSP | 2805 | -40 | -346 | -477 | -4.41 | 0.59 |
| SLI | 2805 | 290 | 192 | -3 | -4.21 | 0.58 |
| DDM | 2820 | 284 | 1 | -186 | -4.75 | 0.69 |
| MHP | 2835 | -76 | -42 | -142 | -4.11 | 0.59 |
| AGP | 2881 | 22 | -43 | -187 | -5.55 | 0.60 |
| SLK | 2927 | 378 | 520 | -28 | 4.53 | 0.96 |
| UBC | 2957 | -78 | -86 | -521 | -1.14 | 0.70 |
| BGP | 2988 | 283 | 9 | -658 | 2.91 | 0.75 |
| RCK | 3049 | 161 | 236 | -671 | -4.37 | 0.87 |
| VLC | 3064 | -26 | 110 | -357 | -2.80 | 0.65 |
| SWM | 3110 | 381 | 337 | -211 | -5.14 | 0.40 |
| GEM | 3277 | -3 | -439 | -584 | -3.91 | 0.72 |
| BSH | 3415 | 99 | 99 | -516 | -4.34 | 0.65 |
| Mean | 2612 | 49 | -21 | -249 | -2.55 | 0.71 |

^aElev Diff refers to the difference in elevation between the observation station and the corresponding model grid cell.

^bR_{mte} is the Pearson's correlation between modeled snowmelt and temperature error.

modeled ablation events across all stations (-3.2 K) is nearly twice that for nonmelt periods (-1.7 K), and temperature error and modeled daily ablation are highly correlated, on average, at $r = -0.71$, with systematically high anticorrelations between these two quantities at nearly all stations (Table 2).

[30] The fact that errors in modeled surface air temperature coincide so closely with melt events indicates a possible physical link between the two variables. Rather than erroneously high temperatures leading to excessive melt, simulated air temperatures are too low during melt events. Zhang *et al.* [2009] found a similar cold bias in an earlier version of WRF during springtime, especially for maximum daily temperatures. In addition, Duffy *et al.* [2006] observed both cold biases and anomalously early snowmelt in at least two other regional climate models incorporating earlier versions of the Noah snowpack scheme. The source of this persistent low bias is unlikely to be pervasively low albedo values, since increased absorption of incoming solar radiation warms both the snowpack and the adjacent atmosphere. On the other hand, this finding is consistent with an erroneously positive net energy flux from the surface atmosphere into the snowpack, which would cool the atmosphere and warm the snowpack. This finding also reaffirms recent findings [Wang *et al.*, 2010; Livneh *et al.*, 2010; Barlage *et al.*, 2010] that components of the energy balance other than albedo likely play a role in observed errors in WRF-Noah snowpack ablation. However, determination of the precise source of this error in the WRF-Noah snowpack

energy balance scheme requires additional energy balance data not available at the observation sites used here.

4. Discussion and Conclusions

4.1. Comparing Simulated and Observed Snowpack Quantity and Timing

[31] There are three principal conclusions to be drawn from this analysis. The first is that model resolution substantially affects the quantity and timing of SWE accumulation and ablation in WRF simulation of mountain snowpacks. When run at a 27 km spacing, WRF appears to substantially underestimate snowpack persistence and magnitude at most stations (Figure 5). By contrast, model grid spacings of 9 km and 3 km come much closer to accurately simulating both accumulation and ablation processes (Table 1). Examination of Figures 6, 7, and 8 suggests that much of this discrepancy is related to the limited ability of WRF to capture the effects of elevation and orography on precipitation and ablation at 27 km in high-altitude environments (i.e., spatially diverse high-elevation mountain ridges are reduced to lower-elevation plateaus). This is apparent in Figure 4, where at relatively low elevations SWE at 27 km is comparable to that at higher resolutions, while above approximately 2500 m 27 km SWE estimates are dramatically lower. It is also possible that elevation differences among the domains result in differences in precipitation phase, which could also play a role in comparison of snowpack among the three domains. The importance

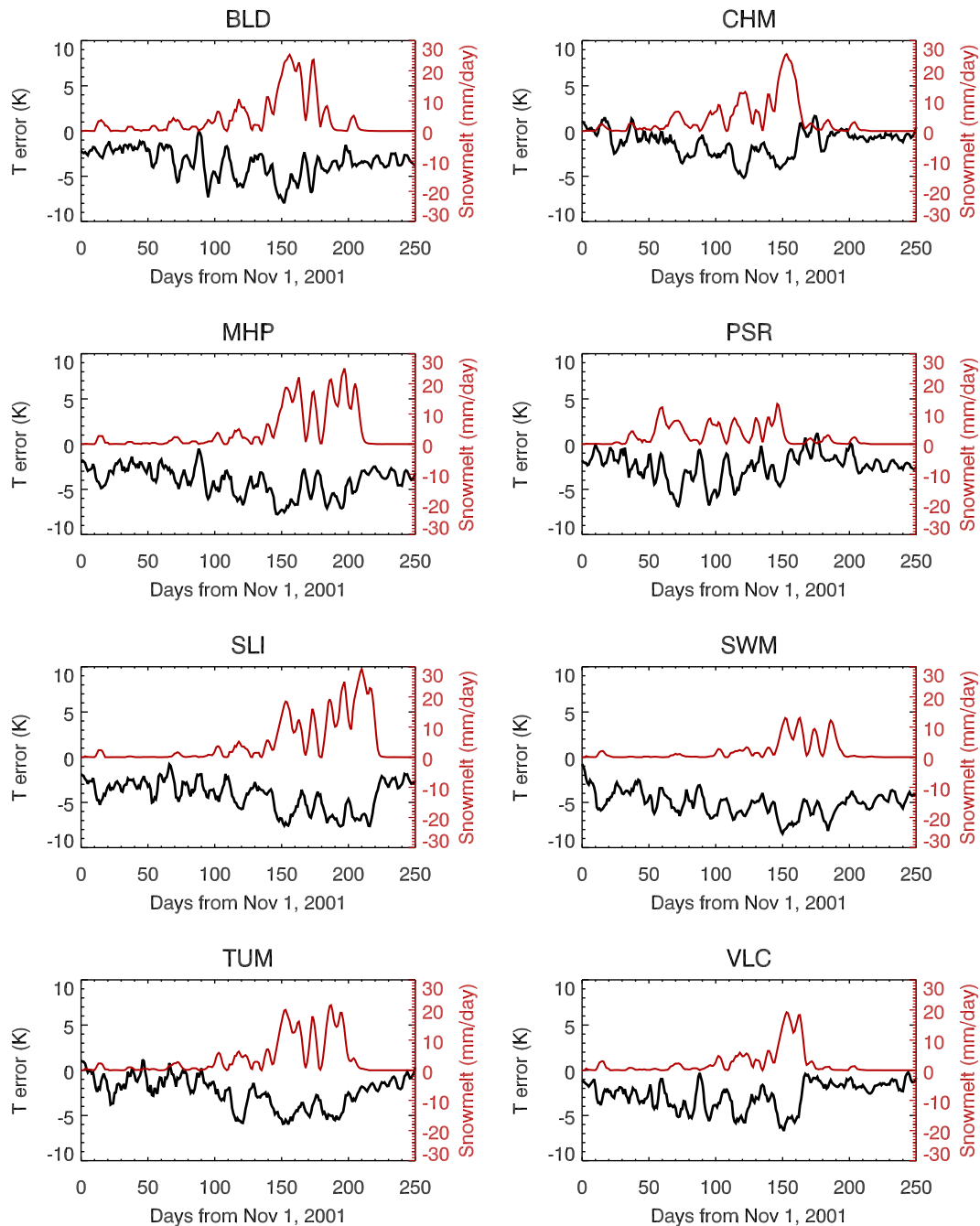


Figure 10. Errors in modeled surface air temperature relative to observations are shown in black, while modeled snowmelt is shown in red for eight representative observation locations. As discussed in section 2.3, snowmelt is inferred from day-to-day changes in total snowpack and is not directly measured. At all stations, there is a statistically significant relationship between simulated snowmelt and temperature error, suggesting a relationship between the timing of model snowmelt and the surface energy balance in WRF.

of accurate topography in driving SWE simulations is also clear in Figures 7 and 8, which show clear relationships between model skill and elevation accuracy regardless of the grid spacing selected. However, the fact that observed station elevations are substantially higher than model elevations at 27 km (206 m), while at 9 and 3 km spacings mean elevation errors are much lower (44 and -12 m, respectively) supports the idea that a grid spacing sufficiently fine to accurately reflect topographic variations is critical to robust simulation

of snowpack. For the Sierra Nevada, it appears that the optimal grid spacing is less than 27 km, which supports previous conclusions by *Jin and Miller [2007]*. A comparison of SWE in 9 km and 3 km WRF simulations with station data does show generally higher correspondence at 3 km (Table 1 and Figure 3b), especially in simulation of overall SWE accumulation. Moreover, the presence of a -9.2% systematic bias in D9 SWE accumulation (relative to -0.6% for D3) shown in Figure 8 suggest that snowpack may be better

simulated in D3 than in D9. However, the close correspondence between 9 km and 3 km area-averaged SWE in Figure 3a (as opposed to SWE averaged over the observation locations in Figure 3b) suggests that this difference may be largely a result of improved correspondence between model grid cells and observations at higher resolutions as opposed to simulation of overall snowpack. This finding supports previous work by Ikeda *et al.* [2010], who found little difference in accumulation accuracy between 6 km and 2 km WRF-Noah simulations over the Colorado Plateau. Even though mean elevation errors in both D9 and D3 are close to 0 m, D3 likely better simulates the spatial extent of precipitation shadows and other orographic features. Overall, for climate-focused studies that do not require great geographic precision a 9 km grid spacing may be sufficient in mountainous environments such as the Sierra Nevada. A 3 km grid appears to provide further incremental improvements in simulated snowpack but only at the cost of substantially increased computational resources. A 9 km grid spacing is close to those used in some recent RCM studies of mountain climates [Suklitsch *et al.*, 2010; Caldwell *et al.*, 2009] but somewhat finer than those used in other studies [Qian *et al.*, 2009; Weare and Du, 2008].

[32] The second principal conclusion of this work is that WRF-Noah simulates snowpack accumulation without major systematic bias, provided model resolution is sufficiently high, but fails to realistically capture melt processes. A comparison of observed and modeled precipitation at 10 stations yields no statistically significant differences and nonsignificant high biases at 3 km and 9 km resolutions, suggesting that erroneous precipitation is likely not the source of observed snowpack biases in WRF. The deficiency of WRF in simulating ablation processes has been observed by other recent studies [e.g., Livneh *et al.*, 2010; Wang *et al.*, 2010; Barlage *et al.*, 2010], but in these studies accumulation and melt events have not been explicitly separated as they are here. The dichotomy between simulation of accumulation and ablation is first evident in Figure 3b, where D3 and D9 SWE closely match observations during the accumulation season but diverge during the melt season. At individual stations, timing and magnitude of accumulation events match closely between WRF and observations (mean $r = 0.80$), while most modeled ablation events occur earlier than observed melt (Table 1 and Figure 5). At some individual stations shown in Figure 5 this early melt is readily apparent (e.g., GIN and PSR), while at others there is a much closer match between observed and modeled melt timing (e.g., SLI and SWM). Some portion of the variation among stations is likely caused by differences between the grid cell and the point-scale observation in the spatial scale sampled. For example, past research suggests that a point-based snowpack observation (on the order of 1 m²) does not accurately reflect spatial variability at larger scales such as the 9 to 729 km² model grid cells used here [Anderton *et al.*, 2004]. However, this source of error likely has a mean near zero when averaged over a large number of observation locations [Ikeda *et al.*, 2010; Barlage *et al.*, 2010] and is thus more important from a validation perspective than for regional hydroclimate studies. Of greater concern are the systematic biases toward early snowmelt apparent when comparing observations and model output. At low resolutions (e.g., D27), a portion of this melt bias (approximately 11 days observed

bias in SCD) is almost certainly related to differences in elevation between the model grid cell and observation location, as is apparent in Figure 7. Only a few hundred meters of difference in elevation can result in major differences in the timing of snowmelt. However, even with discrepancies in elevation taken into account, the model exhibits a systematic bias of 22–25 days toward earlier snowmelt. This bias can most easily be explained by problems arising from snowpack physics in WRF-Noah.

4.2. Evaluation of the Snowpack Energy Balance in WRF-Noah

[33] The third principal conclusion of this work is that an imbalance in heat fluxes is a likely contributor to the observed early snowmelt bias in WRF-Noah. At first glance, the cold bias in WRF-Noah surface air temperatures during melt events seems incompatible with the erroneously early simulated snowmelt. When viewed from an energy balance perspective, though, these seemingly contradictory observations suggest an erroneously positive heat flux from the atmosphere into the snowpack as one possible source of early melt bias. This hypothesis requires further testing, however, as currently available energy balance data at observation locations used here is insufficient to fully evaluate modeled fluxes. We recommend this aspect of the WRF-Noah system as a strong candidate for future study.

[34] In absence of validation data, examination of how snowmelt is simulated in Noah can provide guidance for future efforts to improve WRF-Noah. The primary energy balance equation for snowmelt in Noah is

$$Q_{melt} = Q_{down} - Q_p - Q_{fr} - Q_{up} - Q_{soil} - Q_{lh} - Q_{sh} \quad (1)$$

All fluxes are in W/m². Q_{Melt} is the snowmelt heat flux, Q_{down} is downwelling radiation, Q_p is the heat flux between new precipitation and the snow surface, Q_{fr} is the heat flux associated with freezing rain, Q_{up} is upwelling longwave radiation, Q_{soil} is the soil heat flux, Q_{lh} is the latent heat flux (limited to sublimation over snowpack), and Q_{sh} is the sensible heat flux between the atmosphere and snowpack [Koren *et al.*, 1999]. Several components of the snowmelt equation cannot adequately explain the observed low-temperature biases during snowmelt events. Q_p and Q_{fr} are associated with precipitation events and are thus unlikely to affect melt except in unusual cases (e.g., large rain-on-snow events). Problems with Q_{down} and Q_{soil} are unlikely to result in the negative air temperature observed specifically during melt events. This leaves Q_{up} , Q_{sh} , and Q_{lh} as potential sources of error.

[35] The upwelling longwave flux is calculated in Noah using the Stefan-Boltzmann equation:

$$Q_{up} = \varepsilon \sigma T^4 \quad (2)$$

where ε is the snowpack emissivity, σ is the Stefan-Boltzmann constant, and T is the snowpack temperature in K. One potentially large source of error is the emissivity value selected. The range of snowpack emissivity values in the published literature is between approximately 0.94 and 0.99, with lower values occurring only in very limited cases such as bare ice [Hori *et al.*, 2006]. While the maximum snowpack emissivity value in WRF-Noah is 0.95, at the low end of this range, model snowpack proved nearly identical when the

maximum emissivity was increased to 1.0. As such, selection of the maximum snowpack emissivity is unlikely to be the source of observed error in surface air temperature and snowmelt timing.

[36] The other possible sources of error are the sensible and latent heat flux terms. In Noah, sensible heat is calculated as

$$Q_{sh} = RCh(T_s - \theta) \quad (3)$$

where RCh is a surface heat exchange coefficient multiplied by the wind speed, air density, and the specific heat of water; T_s is the surface skin temperature; and θ is the air potential temperature. During major melt events, the air temperature is warmer than the surface skin temperature, leading to heating of the surface, or negative sensible heat flux from the surface to the atmosphere. The temperature component of equation (3), $(T_s - \theta)$, is unlikely to be the source of anomalously high snowmelt because while T_s is fixed near 0°C during melt events, simulated θ is too low. This scenario results in less heating of the snowpack by the atmosphere than if there were no temperature bias, thus decreasing melt. The other term in equation (3), RCh, is highly dependent on wind speed, and in order to achieve the excess snowmelt and negative bias in modeled air temperature, WRF wind speeds would likely have to be substantially too high. Although accurate in situ wind speed data is not available at the snowpack observation stations used here, past studies of wind speed in WRF in the western United States do suggest a bias toward anomalously high wind speeds [e.g., Cheng and Steeburgh, 2005]. Over snow covered areas, latent heat fluxes in Noah are limited to sublimation and frost, and there are no terms in the energy balance equations to account for evaporation of liquid water from the snowpack [Koren et al., 1999]. As melting snowpacks often contain substantial amounts of liquid water [Jordan, 1983], the inclusion of a term to explicitly account for evaporative heat flux from snow could also improve simulation of snowmelt in Noah.

[37] Ultimately, the source of the observed bias toward early snowmelt remains uncertain. Attempts to improve simulation of snowpack in earlier versions of WRF-Noah by adjusting the turbulent heat flux equations have met with some success, however. For example, Wang et al. [2010] show substantially increased snowpack when the effect of under-canopy resistance on wind speed is included. In addition, RCh is estimated in Noah using an iterative approach, and with a small number of iterations the model can at times fail to converge. Adding a larger number of iterations (thirty, increased from five) results in improved snowpack estimates [Wang et al., 2010]. In future studies, comparison of WRF-Noah and observations for a full suite of snowpack mass and energy terms should be conducted to determine the precise source of errors in both snowmelt timing and surface air temperature and to assess the best means of improving snowpack simulation. Regardless, this study has resolved some of the uncertainty in past studies [e.g., Duffy et al., 2006] regarding the nature of persistently early snowmelt combined with low biases in temperature. Duffy et al. [2006] speculate that either insufficiently fine model resolution or positive temperature errors on days with high precipitation totals could be the cause of observed snowpack biases. We have shown that neither of these is a likely source.

[38] Overall, comparison of our analysis with other recent simulations of snowpack in earlier versions of WRF and/or Noah [e.g., Livneh et al., 2010; Barlage et al., 2010; Wang et al., 2010] suggests that version 3.1.1 of the WRF-Noah modeling system represents a substantial improvement over earlier versions in simulating mountain snowpack. Skill in simulation of accumulation can largely be attributed to accurate simulation of precipitation processes in WRF. Examination of the recent literature suggests that increased fidelity in simulation of ablation processes, on the other hand, is likely due to improvements in handling of snowpack albedo in Noah [Livneh et al., 2010; Barlage et al., 2010]. Some apparent error will always be present when comparing model output with point-based snowpack observations because of a mismatch in observation scale, but recent studies suggest that this type of error is randomly distributed with a mean near zero when comparing model results against stations that are not directly influenced by a vegetation canopy [Ikeda et al., 2010; Barlage et al., 2010]. Our results show that at 3 km and 9 km resolutions there is little systematic bias in snowpack accumulation, but that systematic bias in snowpack ablation is still present in the WRF-Noah configuration used here. With additional improvements to the simulation of snowpack energy balance in Noah, it may be possible to substantially reduce systematic bias in WRF-Noah simulation of mountain snowpack. The result would be a more robust modeling system for simulation of changing snowpack and climate in mountainous regions such as the Sierra Nevada.

[39] **Acknowledgments.** Funding for this research was provided by the National Science Foundation grant AGS-0735056 in part by the NSF through TeraGrid resources provided by the Pittsburgh Supercomputing Center. This research also used resources of the National Energy Research Scientific Computing Center under contract m995, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231. Sarah Kapnick is also supported by a NASA Earth and Space Science Graduate Fellowship (07-Earth07F-0232) and a 2010 Switzer Environmental Fellowship. Finally, we thank Michael Durand and two anonymous reviewers for their useful comments.

References

- Anderton, S. P., S. M. White, and B. Alvera (2004), Evaluation of spatial variability in snow water equivalent for a high mountain catchment, *Hydrol. Processes*, *18*, 435–453, doi:10.1002/hyp.1319.
- Andreadis, K. M., P. Storck, and D. P. Lettenmaier (2009), Modeling snow accumulation and ablation processes in forested environments, *Water Resour. Res.*, *45*, W05429, doi:10.1029/2008WR007042.
- Arnell, N. W. (1999), Climate change and global water resources, *Global Environ. Change*, *9*, S31–S49, doi:10.1016/S0959-3780(99)00017-5.
- Barlage, M., F. Chen, M. Tewari, K. Ikeda, D. Gochis, J. Dudhia, R. Rasmussen, B. Livneh, M. Ek, and K. Mitchell (2010), Noah land surface model modifications to improve snowpack prediction in the Colorado Rocky Mountains, *J. Geophys. Res.*, *115*, D22101, doi:10.1029/2009JD013470.
- Barnett, T., R. Malone, W. Pennell, D. Stammer, B. Stentner, and W. Washington (2004), The effects of climate change on water resources in the west: Introduction and overview, *Clim. Change*, *62*, 1–11, doi:10.1023/B:CLIM.0000013695.21726.b8.
- Barnett, T. P., J. C. Adam, and D. P. Lettenmaier (2005), Potential impacts of a warming climate on water availability in snow-dominated region, *Nature*, *438*, 303–309, doi:10.1038/nature04141.
- Barnett, T. P., et al. (2008), Human-induced changes in the hydrology of the western United States, *Science*, *319*, 1080–1083, doi:10.1126/science.1152538.
- Caldwell, P., H.-N. S. Chin, D. C. Bader, and G. Bala (2009), Evaluation of a WRF dynamical downscaling simulation over California, *Clim. Change*, *95*, 499–521, doi:10.1007/s10584-009-9583-5.

- Cayan, D., S. Kammerdiener, M. Dettinger, J. Caprio, and D. Peterson (2001), Changes in the onset of spring in the western United States, *Bull. Am. Meteorol. Soc.*, *82*, 399–415, doi:10.1175/1520-0477(2001)082<0399:CITOOS>2.3.CO;2.
- Cayan, D. R., E. P. Maurer, M. D. Dettinger, M. Tyree, and K. Hayhoe (2008), Climate change scenarios for the California region, *Clim. Change*, *87*, S21–S42, doi:10.1007/s10584-007-9377-6.
- Chen, F., and J. Dudhia (2001), Coupling an advanced land-surface/hydrology model with the Penn State/NCAR MM5 modeling system. Part I: Model description and implementation, *Mon. Weather Rev.*, *129*, 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.
- Cheng, W. Y. Y., and W. J. Steeburgh (2005), Evaluation of surface sensible weather forecasts by the WRF and the Eta models over the western United States, *Weather Forecast.*, *20*, 812–821, doi:10.1175/WAF885.1.
- Dettinger, M. D., and D. R. Cayan (1995), Large-scale atmospheric forcing of recent trends toward early snowmelt runoff in California, *J. Clim.*, *8*, 606–623, doi:10.1175/1520-0442(1995)008<0606:LSAFOR>2.0.CO;2.
- Dettinger, M. D., D. R. Cayan, M. K. Meyer, and A. E. Jeton (2004), Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River basin, Sierra Nevada, California, 1900–2099, *Clim. Change*, *62*, 283–317, doi:10.1023/B:CLIM.0000013683.13346.4f.
- Dudhia, J. (1989), Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model, *J. Atmos. Sci.*, *46*, 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.
- Duffy, P. B., R. W. Arritt, J. Coquard, W. Gutowski, J. Han, J. Iorio, J. Kim, L.-R. Leung, J. Roads, and E. Zeddon (2006), Simulations of present and future climates in the western United States with four nested regional climate models, *J. Clim.*, *19*, 873–895, doi:10.1175/JCLI3669.1.
- Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, *J. Geophys. Res.*, *108*(D22), 8851, doi:10.1029/2002JD003296.
- Feng, X., A. Sahoo, K. Arsenault, P. Houser, Y. Luo, and T. J. Troy (2008), The impact of snow model complexity at three CLPX sites, *J. Hydrometeorol.*, *9*, 1464–1481, doi:10.1175/2008JHM860.1.
- Flanner, M. G., and C. S. Zender (2006), Linking snowpack microphysics and albedo evolution, *J. Geophys. Res.*, *111*, D12208, doi:10.1029/2005JD006834.
- Hamlet, A. F., P. W. Mote, M. P. Clark, and D. P. Lettenmaier (2005), Effects of temperature and precipitation variability on snowpack trends in the western United States, *J. Clim.*, *18*, 4545–4561, doi:10.1175/JCLI3538.1.
- Hidalgo, H. G., et al. (2009), Detection and attribution of streamflow timing changes to climate change in the western United States, *J. Clim.*, *22*, 3838–3855, doi:10.1175/2009JCLI2470.1.
- Hong, S.-Y., Y. Noh, and J. Dudhia (2006), A new vertical diffusion package with an explicit treatment of entrainment processes, *Mon. Weather Rev.*, *134*, 2318–2341, doi:10.1175/MWR3199.1.
- Hori, M., et al. (2006), In-situ measured spectral directional emissivity of snow and ice in the 8–14 μm atmospheric window, *Remote Sens. Environ.*, *100*, 486–502, doi:10.1016/j.rse.2005.11.001.
- Howat, I. M., and S. Tulaczyk (2005a), Climate sensitivity of spring snowpack in the Sierra Nevada, *J. Geophys. Res.*, *110*, F04021, doi:10.1029/2005JF000356.
- Howat, I. M., and S. Tulaczyk (2005b), Trends in spring snowpack over a half-century of climate warming in California, USA, *Ann. Glaciol.*, *40*, 151–156, doi:10.3189/172756405781813816.
- Ikeda, K., et al. (2010), Simulation of seasonal snowfall over Colorado, *Atmos. Res.*, *97*, 462–477, doi:10.1016/j.atmosres.2010.04.010.
- Jin, J., and N. L. Miller (2007), Analysis of the impact of snow on daily weather variability in mountainous regions using MM5, *J. Hydrometeorol.*, *8*, 245–258, doi:10.1175/JHM565.1.
- Jordan, P. (1983), Meltwater movement in a deep snowpack: 1. Field observations, *Water Resour. Res.*, *19*(4), 971–978, doi:10.1029/WR019i004p00971.
- Kain, J. (2004), The Kain-Fritsch convective parameterization: An update, *J. Appl. Meteorol.*, *43*, 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.
- Kain, J. S., and J. M. Fritsch (1990), A one-dimensional entraining/detraining plume model and its application in convective parameterization, *J. Atmos. Sci.*, *47*, 2784–2802, doi:10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2.
- Kain, J. S., and J. M. Fritsch (1993), Convective parameterization for mesoscale models: The Kain-Fritsch scheme, in *The Representation of Cumulus Convection in Numerical Models*, Meteor. Monogr., vol. 46, edited by K. A. Emanuel and D. J. Raymond, pp. 165–170, Am. Meteorol. Soc., Reston, Va.
- Kapnick, S., and A. Hall (2009), Observed changes in the Sierra Nevada snowpack: Potential causes and concerns, *Rep. CEC-500-2009-016-F*, Calif. Environ. Prot. Agency, Sacramento.
- Kapnick, S., and A. Hall (2010), Observed climate-snowpack relationships in California and their implications for the future, *J. Clim.*, *23*, 3446–3456, doi:10.1175/2010JCLI2903.1.
- Koren, V., J. Schaake, K. Mitchell, Q.-Y. Duan, F. Chen, and J. M. Baker (1999), A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, *J. Geophys. Res.*, *104*(D16), 19,569–19,585, doi:10.1029/1999JD900232.
- Liston, G. E. (2004), Representing subgrid snow cover heterogeneities in regional and global models, *J. Clim.*, *17*, 1381–1397, doi:10.1175/1520-0442(2004)017<1381:RSSCHI>2.0.CO;2.
- Livneh, B., Y. Xia, K. E. Mitchell, M. B. Ek, and D. P. Lettenmaier (2010), Noah LSM snow model diagnostics and enhancements, *J. Hydrometeorol.*, *11*, 721–738, doi:10.1175/2009JHM1174.1.
- Marks, D., and J. Dozier (1992), Climate and energy exchange at the snow surface, *Water Resour. Res.*, *28*(11), 3043–3054, doi:10.1029/92WR01483.
- Maurer, E. P. (2007), Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California, under two emissions scenarios, *Clim. Change*, *82*, 309–325, doi:10.1007/s10584-006-9180-9.
- Maurer, E. P., and P. B. Duffy (2005), Uncertainty in projections of streamflow changes due to climate change in California, *Geophys. Res. Lett.*, *32*, L03704, doi:10.1029/2004GL021462.
- Maurer, E. P., I. T. Stewart, C. Bonfils, P. B. Duffy, and D. Cayan (2007), Detection, attribution, and sensitivity of trends toward earlier streamflow in the Sierra Nevada, *J. Geophys. Res.*, *112*, D11118, doi:10.1029/2006JD008088.
- Mesinger, F., et al. (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.*, *87*(3), 343–360, doi:10.1175/BAMS-87-3-343.
- Miller, N. L., K. E. Bashford, and E. Strem (2003), Potential impacts of climate change on California hydrology, *J. Am. Water Resour. Assoc.*, *39*, 771–784, doi:10.1111/j.1752-1688.2003.tb04404.x.
- Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008), Stationarity is dead: Whither water management?, *Science*, *319*, 573–574, doi:10.1126/science.1151915.
- 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 long wave, *J. Geophys. Res.*, *102*(D14), 16,663–16,682.
- Mote, P. W., A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier (2005), Declining mountain snowpack in western North America, *Bull. Am. Meteorol. Soc.*, *86*(1), 39–49, doi:10.1175/BAMS-86-1-39.
- Pomeroy, J., A. Rowlands, J. Hardy, T. Link, D. Marks, R. Essery, J. E. Sicart, and C. Ellis (2008), Spatial variability of shortwave irradiance for snowmelt in forests, *J. Hydrometeorol.*, *9*, 1482–1490, doi:10.1175/2008JHM867.1.
- Pomeroy, J. W., D. Marks, T. Link, C. Ellis, J. Hardy, A. Rowlands, and R. Granger (2009), The impact of coniferous forest temperature on incoming longwave radiation to melting snow, *Hydrol. Processes*, *23*, 2513–2525, doi:10.1002/hyp.7325.
- Qian, Y., S. J. Ghan, and L. R. Leung (2009), Downscaling hydroclimatic changes over the western US based on CAM subgrid scheme and WRF regional climate simulations, *Int. J. Climatol.*, *30*(5), 675–693.
- Salathé, E., L. Leung, Y. Qian, and Y. Zhang (2010), Regional climate model projections for the state of Washington, *Clim. Change*, *102*, 51–75, doi:10.1007/s10584-010-9849-y.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. Duda, X. Huang, W. Wang, and J. G. Powers (2008), A description of the advanced research WRF version 3, *Tech. Note 475*, Natl. Cent. for Atmos. Res., Boulder, Colo.
- Stewart, I. T., D. R. Cayan, and M. D. Dettinger (2004), Changes in snowmelt runoff timing in western North America under a ‘business as usual’ climate change scenario, *Clim. Change*, *62*, 217–232, doi:10.1023/B:CLIM.0000013702.22656.e8.
- Stewart, I. T., D. R. Cayan, and M. D. Dettinger (2005), Changes toward earlier streamflow timing across western North America, *J. Clim.*, *18*, 1136–1155, doi:10.1175/JCLI3321.1.
- Suklitsch, M., A. Gobiet, H. Truhetz, N. K. Awan, H. Götzel, and D. Jacob (2010), Error characteristics of high resolution regional climate models over the Alpine area, *Clim. Dyn.*, *17*, 377–390, doi:10.1007/s00382-010-0848-5.
- Thompson, G., P. Field, R. M. Rasmussen, and W. Hall (2008), Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization, *Mon. Weather Rev.*, *136*, 5095–5115, doi:10.1175/2008MWR2387.1.
- VanRheenen, N. T., A. W. Wood, R. N. Palmer, and D. P. Lettenmaier (2004), Potential implications of PCM climate change scenarios for

- Sacramento–San Joaquin River Basin hydrology and water resources, *Clim. Change*, 62, 257–281, doi:10.1023/B:CLIM.0000013686.97342.55.
- Vicuna, S., E. P. Maurer, B. Joyce, J. A. Dracup, and D. Purkey (2007), The sensitivity of California water resources to climate change scenarios, *J. Am. Water Resour. Assoc.*, 43(2), 482–498, doi:10.1111/j.1752-1688.2007.00038.x.
- Wang, Z., and X. Zeng (2010), Evaluation of snow albedo in land models for weather and climate studies, *J. Appl. Meteorol. Climatol.*, 49, 363–380, doi:10.1175/2009JAMC2134.1.
- Wang, Z., X. Zeng, and M. Decker (2010), Improving snow processes in the Noah land model, *J. Geophys. Res.*, 115, D20108, doi:10.1029/2009JD013761.
- Weare, B. C., and H. Du (2008), Modelling regional climate changes: Influences of recent global warming and irrigation in California, *Int. J. Climatol.*, 28, 1201–1212, doi:10.1002/joc.1622.
- Zhang, Y., V. Duliere, P. W. Mote, and E. P. Salathe (2009), Evaluation of WRF and HadRM mesoscale climate simulations over the U.S. Pacific Northwest, *J. Clim.*, 22, 5511–5526, doi:10.1175/2009JCLI2875.1.

A. Hall and S. Kapnick, Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA 90095, USA.

T. M. Pavelsky, Department of Geological Sciences, University of North Carolina, Chapel Hill, NC 27599, USA. (pavelsky@unc.edu)