Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent

Ming Pan,[1] Justin Sheffield,[1] Eric F. Wood,[1] Kenneth E. Mitchell,[2] Paul R. Houser,[3] John C. Schaake,[4] Alan Robock,[5] Dag Lohmann,[2] Brian Cosgrove,[3] Qingyun Duan,[4] Lifeng Luo,[5,6] R. Wayne Higgins,[7] Rachel T. Pinker,[8] and J. Dan Tarpley[9]

Received 21 July 2003; revised 10 September 2003; accepted 24 September 2003; published 26 November 2003.

[1] This is the second part of a study on the cold season process modeling in the North American Land Data Assimilation System (NLDAS). The first part concentrates on the assessment of model simulated snow cover extent. In this second part, the focus is on the evaluation of simulated snow water equivalent (SWE) from the four land surface models (Noah, MOSAIC, SAC and VIC) in the NLDAS. Comparisons are made with observational data from the Natural Resources Conservation Service's Snowpack Telemetry (SNOTEL) network for a 3-year retrospective period at selected sites in the mountainous regions of the western United States. All models show systematic low bias in the maximum annual simulated SWE that is most notable in the Cascade and Sierra Nevada regions where differences can approach 1000 mm. Comparison of NLDAS precipitation forcing with SNOTEL measurements revealed a large bias in the NLDAS annual precipitation which may be lower than the SNOTEL record by up to 2000 mm at certain stations. Experiments with the VIC model indicated that most of the bias in SWE is removed by scaling the precipitation by a regional factor based on the regression of the NLDAS and SNOTEL precipitation. Individual station errors may be reduced further still using precipitation scaled to the local station SNOTEL record. Furthermore, the NLDAS air temperature is shown to be generally colder in winter months and biased warmer in spring and summer when compared to the SNOTEL record, although the level of bias is regionally dependent. Detailed analysis at a selected station indicate that errors in the air temperature forcing may cause the partitioning of precipitation into snowfall and rainfall by the models to be incorrect and thus may explain some of the remaining errors in the simulated SWE. *INDEX TERMS:* 1833 Hydrology: Hydroclimatology; 1863 Hydrology: Snow and ice (1827); 1878 Hydrology: Water/energy interactions; 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; *KEYWORDS:* snow water equivalent, land surface modeling

Citation: Pan, M., et al., Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent, *J. Geophys. Res.*, *108*(D22), 8850, doi:10.1029/2003JD003994, 2003.

1. Introduction

[2] The terrestrial hydrologic system has two types of state variables; one related to temperature (skin, ground and snow temperatures) and one related to moisture (soil or snow water). The ability of land surface models to accurately determine these state variables is critical for their ability to provide information to atmospheric weather prediction models [*Groisman et al.*, 1994a; *Entekhabi et al.*, 1996]. For coupled models, these state variables are prognostic but within a land data assimilation system they would be variables assimilated by a numerical weather prediction system. Within the North American Land Data Assimilation System (NLDAS) [*Mitchell et al.*, 1999, 2000, 2003],

[1]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.
[2]National Centers for Environmental Prediction Environmental Modeling Center, NOAA/NWS, Camp Springs, Maryland, USA.
[3]Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[4]Office of Hydrologic Development, NOAA/NWS, Silver Spring, Maryland, USA.
[5]Department of Environmental Sciences, Rutgers University, New Brunswick, New Jersey, USA.
[6]Now at Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.
[7]National Centers for Environmental Protection Climate Prediction Center, NOAA/NWS, Camp Springs, Maryland, USA.
[8]Department of Meteorology, University of Maryland, College Park, Maryland, USA.
[9]Office of Research and Applications, NOAA/NESDIS, Camp Springs, Maryland, USA.

Copyright 2003 by the American Geophysical Union.
0148-0227/03/2003JD003994

predictions of land surface moisture and temperature states could be used to improve forecasts from weather prediction models in a real time operational framework [*Mitchell et al.*, 2003]. Cold season processes play a major role in defining these land surface states, being the dominant regime over much of North America during the winter and spring periods [*Groisman et al.*, 1994b; *Brown*, 2000]. Through the storage of moisture and influence on incoming energy fluxes, snow and ice are not only important in determining current hydrological conditions but also in shaping future states via snow accumulation and spring melt [*Yeh et al.*, 1983; *Namias*, 1985]. The subsequent effects on flooding and water resources are of considerable interest to the social and agricultural communities.

[3] This paper is the second part of two-part study on the modeling of cold season processes within the NLDAS project. The first part focuses on the evaluation of model simulated snow cover extent [*Sheffield et al.*, 2003]. This second part concentrates on the evaluation of model-derived snow water equivalent (SWE). Snow water equivalent is defined as the depth of water that would be obtained if a column of snow were completely melted. It quantifies the amount of frozen moisture storage and determines the amount of spring melt and subsequent flooding.

[4] The goal of this paper is to evaluate the components of the NLDAS in terms of cold season modeling and how these components interact to provide a final product that may be used to improve numerical weather prediction model forecasts. These components include the input forcings to the models, the performance of the models themselves relative to observational data and to each other, and the characteristics of the NDLAS modeling framework, which includes the spatial resolution of the modeling grid. The NLDAS project has completed a 3-year, retrospective simulation over the conterminous United States using the four participating land surface models (LSM): MOSAIC [*Koster and Suarez*, 1996], Noah [*Betts et al.*, 1997; *Chen et al.*, 1996, 1997; *Koren et al.*, 1999], SAC [*Burnash et al.*, 1973; *Burnash*, 1995], and VIC [*Liang et al.*, 1994, 1996, 1999; *Cherkauer and Lettenmaier*, 1999]. To carry out the evaluation, the forcings and the modeled snow water equivalent from the NLDAS simulations are compared with measurements from the SNOTEL station based observing network [*Crook*, 1977; *Serreze et al.*, 1999]. The computational and operational limitations of the NLDAS modeling framework mean that the models are implemented at a coarse resolution relative to the scale of cold season processes in complex and mountainous terrain. Therefore this comparison also provides the opportunity to assess the impact of the NLDAS spatial resolution on the ability of the models to represent such processes. Furthermore, an assessment can be made of the various sub-grid parameterizations that the different land surface models implement to resolve the inherent sub-grid variability in the processes and land surface characteristics that make this type of modeling so challenging.

2. Cold Season Process Modeling

[5] The four land surface models that contribute to the NLDAS modeling effort (MOSAIC, Noah, SAC and VIC) simulate cold season processes with varying degrees of complexity. In general, all models simulate the physical processes of changes of moisture states and the related partitioning of energy fluxes (except the SAC model which does not simulate the land surface energy balance) but the parameterizations used may differ between models. In addition, each model handles sub-grid variability of vegetation and elevation at different levels of complexity, which effects snow cover predictions through sub-grid variations in precipitation, temperature and radiation budgets.

[6] All of the snow modules used in the different models are based on balances of mass and energy in the snowpack. The change in snowpack SWE is balanced by the input snowfall, and output snowmelt and snow sublimation. The heat flux through the snowpack (sum of net radiation, sensible/latent heat, ground heat fluxes) is used to change the temperature, phase composition, and amount of snowpack. MOSAIC, Noah, and VIC run at full energy mode, which means that the snow energy process is coupled into the energy transfer processes of the entire LSM. Thus in one time step, temperatures of soil layers, soil surface, and snowpack layers (if any) will be solved from heat transfer/balance equations for the entire system (soil, snowpack, vegetation, and air) together with the corresponding water balance equations. Each individual model may have different simplifying assumptions, e.g., linearization of the heat transfer equation (MOSAIC) or constant temperature boundary conditions in the deep layer (VIC). Noah, uniquely, addresses the change of snow density due to compaction in time, and assumes the maximum liquid water storage capacity in the snowpack to be 13%, above which it is removed from the snowpack [*Koren et al.*, 1999]. Noah also accounts for effects from frozen soil, e.g., reduction of soil infiltration capacity. VIC accounts for snow aging by decreasing its albedo with time, and assumes the maximum liquid water storage capacity to be 6% [*Wigmosta et al.*, 1994]. SAC is different from the other three in that it only calculates the water balance, and the snow calculation is done separately by an independent snow model developed by the Hydrologic Research Laboratory of the Office of Hydrology [*Anderson*, 1973], which calculates snowmelt as a function of air temperature. In the case of rainfall falling on snow, the snowmelt rate in SAC is controlled by heat exchange between the liquid water and the snowpack. The Noah, SAC and VIC models all include liquid water content in the snowpack but the MOSAIC model routes rainfall falling onto a snowpack directly to the soil surface for subsequent infiltration or runoff [*Koster and Suarez*, 1996]. Details of the model parameterizations of snow cover extent are given in part one of this study [*Sheffield et al.*, 2003].

3. Data

3.1. Snow Water Equivalent Measurements

[7] Available validation data for snow water equivalent is limited in terms of the length of record and area of coverage. In comparison to snow cover extent (SCE), which is generally measured at large scales via satellite remote sensing techniques, for example the IMS [*Ramsay*, 1998] and MODIS [*Hall et al.*, 2002] products, SWE is relatively difficult to retrieve, especially from satellite imagery. The NWS National Operational Hydrologic Remote Sensing Center (NOHRSC) provide daily maps of snow cover, derived from the NOAA GOES and AVHRR satellites, for

the conterminous United States and Alaska [*Hartman et al.*, 1995]. As part of their remote sensing analysis, the NOHRSC product also includes rough estimates of snow water equivalent based on ground and airborne observations combined with snow cover information from the satellite maps [*Carroll et al.*, 1999]. In addition, estimates of SWE based solely on airborne gamma survey have been available for since 1980 [*Carroll and Carroll*, 1989], but the spatial coverage of airborne flight lines are limited and change from time to time.

[8] The lack of reliable, large-scale measurements of SWE, especially in mountainous regions that are most important for flood and water supply information, hampers an evaluation of modeled estimates. As a result, this paper uses point observations based on the Natural Resources Conservation Service's (NRCS) SNOTEL network (see http://www.wcc.nrcs.usda.gov/snotel/SNOTEL_Info/snotel_info.html). Data from ground-based point observations are a good choice for model validation if the number of observing points is large and if the stations are well spread geographically. SNOTEL has been successfully used for large-scale analysis of snowpack estimation and water supplies. *Serreze et al.* [1999] analyzed regional snowpack and precipitation for the mountainous western US using SNOTEL data. *Simpson et al.* [1998] used measurements from SNOTEL stations to validate a remote sensing derived snow cover product and concluded that, although there are scales issues, the SNOTEL are the only practical options for validation. The SNOTEL network, which has been in operation since 1980 and is operated by NRCS's Western Regional Climate Center (WRCC), provides measurements of SWE and basic meteorology using a pressure sensing snow pillow, a storage precipitation gage and air temperature sensor. Data are recorded every 15 min and reported daily. Currently there are approximately 600 stations across the western U.S. and Alaska. In this study we use the SNOTEL daily observations for SWE, air temperature and precipitation.

[9] The geographic region for the comparison between the SNOTEL station observations and the NLDAS model predictions is west of $-104°$ longitude, which includes the Rockies, Cascades, and the Sierra Nevada mountains. In this region there are a total of 560 SNOTEL stations with good quality records which are well distributed throughout this area. The elevation range of these stations is 627.9 to 3535.7 m, with the majority of the stations at elevations above 1000 m, mean elevation around 2500 m and a mean annual temperature of about 4°C. The high elevation locations of the SNOTEL stations provide a good opportunity to evaluate the cold season parameterizations of the NLDAS LSMs.

3.2. North American Land Data Assimilation System (NLDAS) Land Surface Model Simulations

[10] The land surface models participating in the NLDAS operate within a framework that consists of a common 1/8 degree geographic grid over the conterminous United States, using common soil and vegetation parameters and distributions. The meteorological forcings are derived from observations and from the National Centers for Environmental Prediction (NCEP) Eta weather forecast model and its data assimilation system (EDAS). Precipitation is taken from the observation based analysis of *Higgins et al.* [2000] and solar radiation is derived from the satellite retrieval of *Pinker et al.* [2003]. Other forcing variables are provided by the EDAS. Further details of the NLDAS forcings are given by *Cosgrove et al.* [2003]. Simulations were run retrospectively for the period October 1996 to September 1999. Model outputs include predictions of grid average snow water equivalent as well as standard water and energy states and fluxes. Details of the NLDAS modeling framework and the retrospective simulations are given in the NLDAS overview paper of *Mitchell et al.* [2003]. The models are currently also running in near real time, as would be the case in an operational implementation of the NLDAS, but these data are not evaluated in this study.

3.3. Comparison of NLDAS Data With SNOTEL Point Data

[11] The lack of reliable large-scale observations of SWE means that the point scale observations from the SNOTEL network are the only available data set for evaluation of the NLDAS simulations. However, issues of scale must be addressed when comparing grid and point data. The NLDAS models use a 1/8 degree (approximately 12 km) computational grid and so how representative the SNOTEL site is of the grid average is somewhat questionable, especially in relation to elevation and temperature effects. The SNOTEL stations were originally set up to measure snowpack for long-term streamflow forecasting and water supply management (SNOTEL website: http://www.wcc.nrcs.usda.gov/snow) and the sites were collocated with snow course sites that correlated well with streamflow volumes over long periods [*Serreze et al.*, 1999]. This may indicate that they are representative of larger regions in terms of streamflow.

[12] This problem is addressed in a number of ways. Firstly, the SNOTEL network consists of a large number of stations, which are well distributed over the mountainous regions of the western US and encompass a diversity of terrain and climate. The geographic coverage extends from Washington State to New Mexico and includes the Northern and Southern Rockies, the Cascades and the Sierra Nevada mountain ranges. It can be argued that by using a large number of stations that are well spread geographically, the variability that exists in cold season processes over such large scales can be accounted for. Secondly, the general problem of point to grid scale comparisons can be overcome by aggregating the data up to larger time and space scales before doing the comparisons. This, in effect, removes the discrepancies that are introduced by the effect of scale by removing the small-scale variability in snow processes and retaining the large-scale variability in the climate forcings. Therefore comparisons are carried out for annual or seasonal totals and the results are shown for all stations together. Conclusions are then based on the average performance of the models over all stations in the region. Thirdly, comparison of the elevation of SNOTEL stations with that at the equivalent grid box (not shown) indicate no systematic bias but significant differences at a number of individual locations. Therefore it was decided to screen from the comparisons all SNOTEL stations where the absolute difference in elevation was greater than 50 m. This resulted in 110 SNOTEL stations remaining, with an elevation range of 856.6 m to 3474.7 m, which is not dissimilar to the

elevation range of the original 560 stations. Although this is a reduced number of stations, it is still sufficient to encompass the variability of cold season processes given that the geographic diversity is retained.

[13] A geostatistical analysis of SWE is useful to understand its variation in space and determine whether significant spatial correlation exists at scales comparable to that of the NLDAS modeling grid. If such correlation exists, then one can argue that the point observations, such as those from SNOTEL stations, contain sufficient information about the larger scale that they can be considered as representative of the grid mean and thus can be potentially used for comparison against grid scale data. Of course, the lack of large-scale SWE observations makes this analysis impossible, but data sets do exist that combine ground observations, remote sensing and modeling to provide estimates of SWE over large scales. As mentioned before in Section 3.1, the NOHRSC remote sensing based analysis of daily snow cover for the coterminous United States and Alaska [*Carroll et al.*, 1999] also includes rough estimates of snow water equivalent based on ground and airborne observations combined with the snow cover information. Daily maps of SWE for the US are available from October 2001 to the present. To estimate the spatial correlation structure of SWE over the vicinity of SNOTEL stations, semivariograms of SWE were calculated around 100 randomly selected stations for randomly selected days. These semivariograms were then averaged to get an estimate of the regional semivariogram. The results indicate that SWE is spatially correlated up to 1/4 degree lag distance. Therefore it can be concluded that small-scale measurements contain sufficient information about the surrounding grid cell to be considered as representative of the grid mean. Although this does not show that any individual station is representative of the grid scale, it does show that on average, over the study region, the SNOTEL stations are representative of the grid scale.

4. Analysis

4.1. Comparison of Model-Simulated and SNOTEL-Measured SWE

[14] Figure 1 shows the mean annual maximum SWE for the model simulations and the SNOTEL measurements for the 110 stations remaining after elevation screening. It can be seen that all models underestimate maximum SWE over all regions. The bias is most prominent over the Sierra Nevada and the Cascade Mountains where differences between model simulations and observed data approach 1000 mm. The differences generally reduce as we move eastward into the Rocky Mountains where a few stations on the eastern edge are within 100 mm of the observations.. Although direct comparison of data at individual stations with the corresponding grid cell simulated data is questionable because of the inherent variability in cold season processes within a grid cell, these plots do provide an indication of the general bias of the simulations over this region. Figure 2 depicts the same information as a scatterplot for the four models, which highlights the consistent low bias (average bias: MOSAIC = −59.4%, Noah = −77.6%, SAC = −51.0%, VIC = −59.9%). Despite this bias, the correlation coefficients are significant at the 99% confidence level for all models, indicating that there is an association between the simulated and observed data. It should also be noted that the largest discrepancies are in the Cascades (diamond symbols) and in the Sierra Nevada mountains (square symbols), although the bias in this latter region is less obvious because of the smaller amounts of accumulated snow. An analysis of the performance of the models as a function of elevation (not shown) showed that the results are independent of elevation.

4.2. Comparison of NLDAS and SNOTEL Precipitation

[15] The low bias in the predicted mean annual maximum SWE is most likely explained by either deficiencies in the model physics or errors in the input meteorological forcings, although there may be other contributing factors. Figures 3 and 4 compare the mean annual precipitation from the NLDAS with the measurements at the 110 SNOTEL stations. It can be seen that the NLDAS precipitation data are generally low for the whole region, the average bias being −57.9%. One should also note that the stations with the highest precipitation and the largest bias are those in the Cascade Mountains and those with the lowest bias are located on the eastern edge of the Northern Rockies.

[16] The differences in precipitation are consistent with the differences seen in the model simulated SWE (Figures 1 and 2) in which all models under predict the annual maximum SWE. Errors in the precipitation may therefore explain some of the errors in the model simulation. To test this, the VIC model was run for the 110 stations using precipitation forcing scaled by a regional factor based on the regression fit of the NLDAS and SNOTEL mean annual precipitation (Figure 4). The regression between the local SNOTEL precipitation and the NLDAS precipitation yields the following relationship:

$$P_{SNOTEL} = 2.1693\ P_{NLDAS} \tag{1}$$

with an R^2 value of 0.64. This relationship provides an indication of the average under-estimation of mountainous precipitation by the NLDAS. As explained by *Mitchell et al.* [2003], NLDAS precipitation amounts are based on NWS precipitation gauges, and it is well known that gauges located in valleys underestimate higher elevation precipitation [*Schultz et al.*, 2002].

[17] The VIC model simulated annual maximum SWE using the adjusted precipitation is shown in Figure 5. The low bias has been removed with the values now clustered around the 1:1 regression line (average bias is 4.0%), although the level of scatter in the values remains. This is to be expected as the precipitation scaling is carried out on a regional basis and so individual stations may still be biased. To address this, a second experiment was carried out in which the VIC model was forced with locally adjusted precipitation such that the NLDAS precipitation at each station was scaled so that the annual total matched that of the SNOTEL measured record. This is, in effect, equivalent to running VIC as a point model, using an estimate of the local precipitation. Other meteorological forcings, such as temperature and radiation, have a considerable influence on the development of the snowpack, but these are not adjusted to the local values in this experiment. It is known that the accuracy of the NLDAS solar radiation forcing is reduced

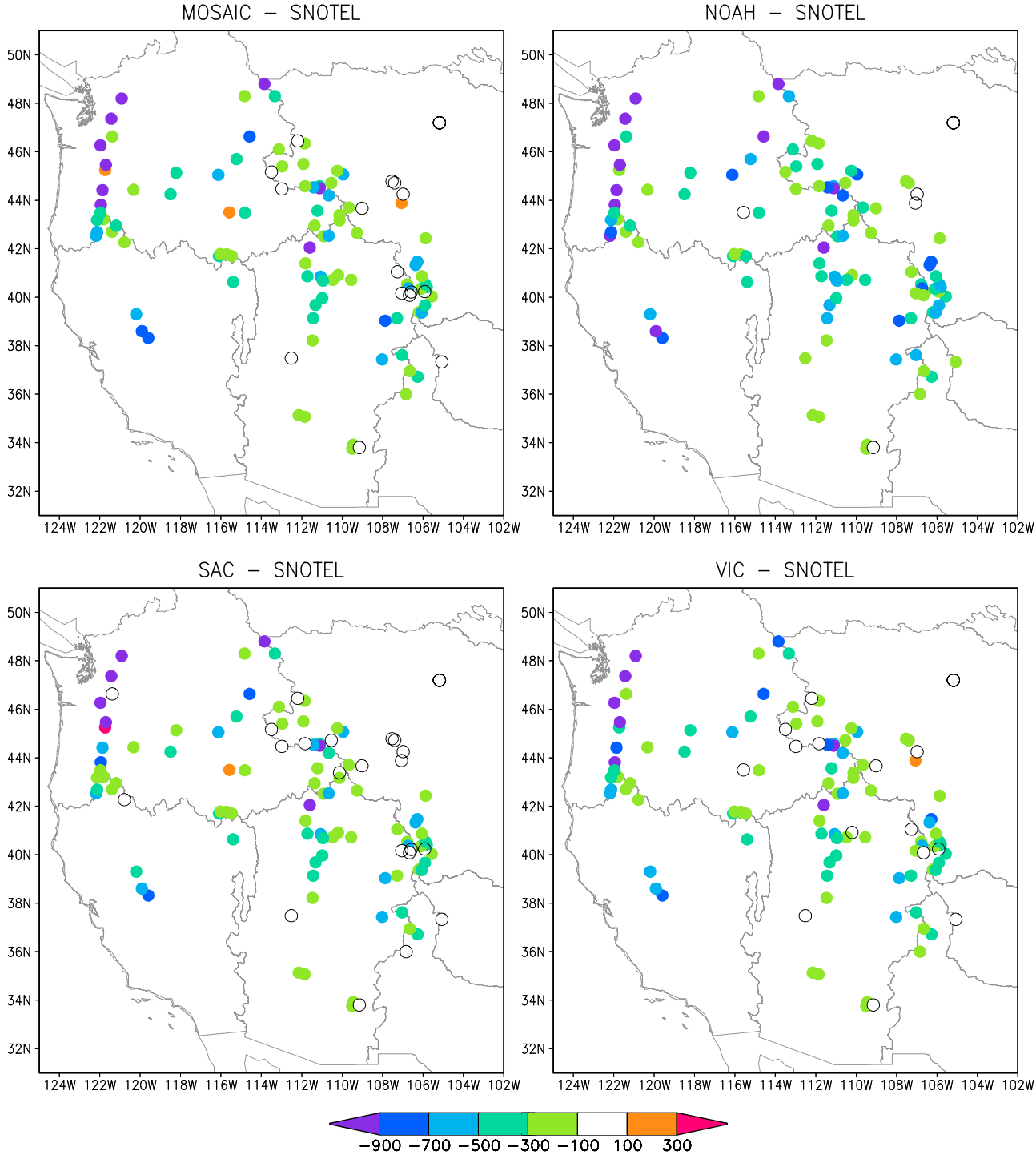

Figure 1. Difference in mean annual maximum SWE between model simulations and SNOTEL measurements (mm).

over snow covered regions due to the inability to distinguish between clouds and snow cover [*Pinker et al.*, 2003]. Radiation measurements are not taken at the SNOTEL sites and so no adjustment is possible. Air temperature measurements are taken at the SNOTEL sites and an analysis of the temperature bias and its effect on the results are discussed in detail in the next section. The resulting simulated SWE is shown as a scatterplot in Figure 6 (the label for "Lone Pine" is referred to in section 3.4). Note that the bias is still low (average bias is 7.8%) but the scatter has been reduced significantly as shown by the improved R^2 of 0.82. The few stations that still show significant errors are all from the Cascades region.

4.3. Comparison of NLDAS and SNOTEL Air Temperature

[18] Despite the improved model predictions using the corrected precipitation, differences still remain between the

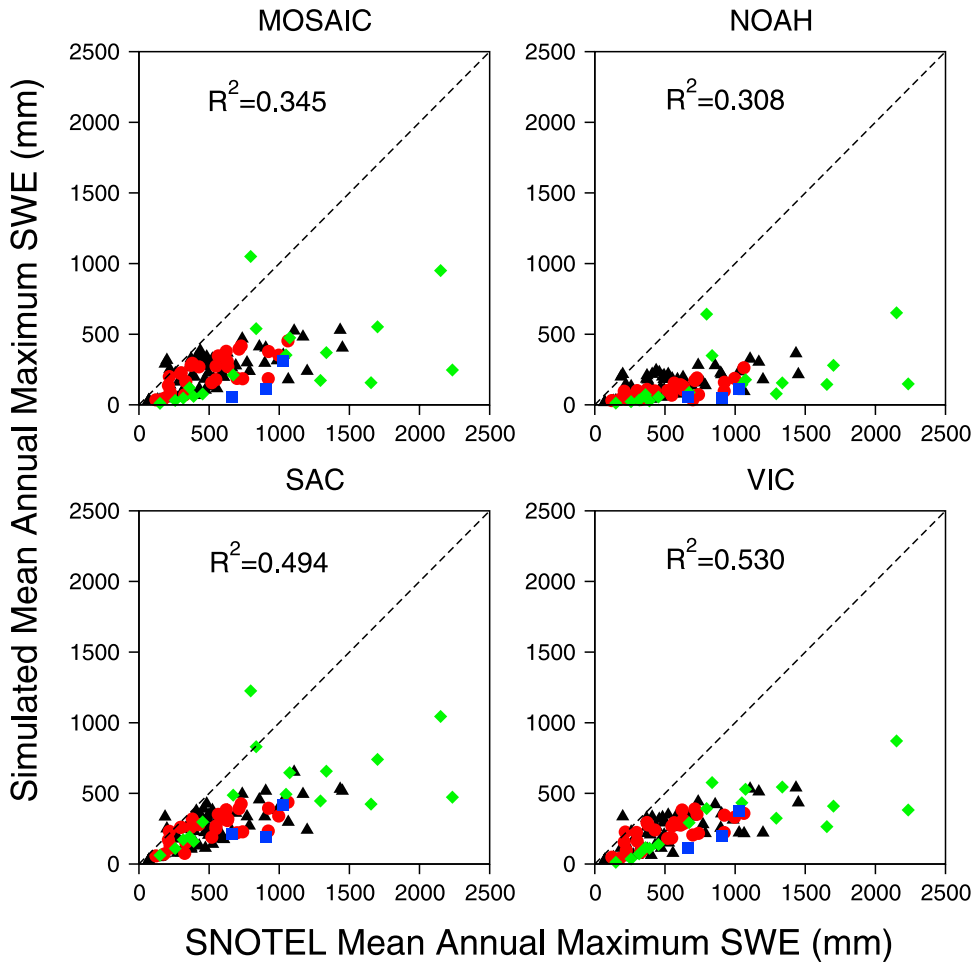

Figure 2. Comparison of mean annual maximum SWE between model simulations and SNOTEL measurements (mm) (legend as in Figure 4).

model predicted and measured SWE data. In addition to precipitation, air temperature also has a large influence on the dynamics of the snowpack and thus errors in the given NLDAS air temperature may explain some of the remaining errors in the simulated SWE. An analysis of the difference in mean cold season (Nov–Apr) air temperature between the NLDAS and the SNOTEL is shown in Figures 7 and 8. These indicate biases of less than 1°C in the mean cold season temperature, with the NLDAS temperature being too cold in the Rockies and being warmer but essentially unbiased in the Cascades and Sierra Nevada regions.

[19] The differences in mean cold season temperature as shown in Figure 8 indicate that the bias is constant with elevation. This suggests that the elevation lapse rate of the NLDAS system, as the analysis fields are downscaled to 1/8 degree computational grid are, on average, correct [*Cosgrove et al.*, 2003]. Average monthly temperatures for all stations and variability amongst stations are shown in Figure 9. It is quite apparent that, in general, the NLDAS temperature is biased colder in winter months and biased warmer in spring and summer. The level of average bias and variability amongst individual stations is regionally dependent however. In the Cascades the variability is quite large but is relatively less in the Rockies, although the maximum bias can reach as much as 6°C colder during the winter. The data for the Sierra Nevada region are not shown because only 3 stations remained after elevation screening.

[20] The biases seen in the temperature comparisons, which although in general are relatively small compared with those seen in the precipitation analysis, are still of the order of 2°C colder during the winter and 1°C warmer in the spring in the Rockies, for example, and may have a significant impact on the temporal evolution of the snowpack. This may be especially true at specific sites where average winter and spring temperatures may be biased from 3 to 6°C. The effect that these biases have on the snowpack and the SWE depends on the timing of the switch between cold and warm bias and the number of days on which the NLDAS temperatures and the SNOTEL measurements are of opposite sign, in the sense that one is below freezing while the other is above freezing. This would have implications on the partitioning of precipitation into snowfall and rainfall by the models and subsequent effects on the accumulation and melt of the snowpack.

[21] Figure 10 shows a scatterplot of the first day of snow and last day of snow plotted as the difference between the model simulation and the SNOTEL measurements. The first day of snow was chosen to be the day at which the SWE first exceeds 10 mm to avoid days with missing data in the SNOTEL record in the early part of the cold season and

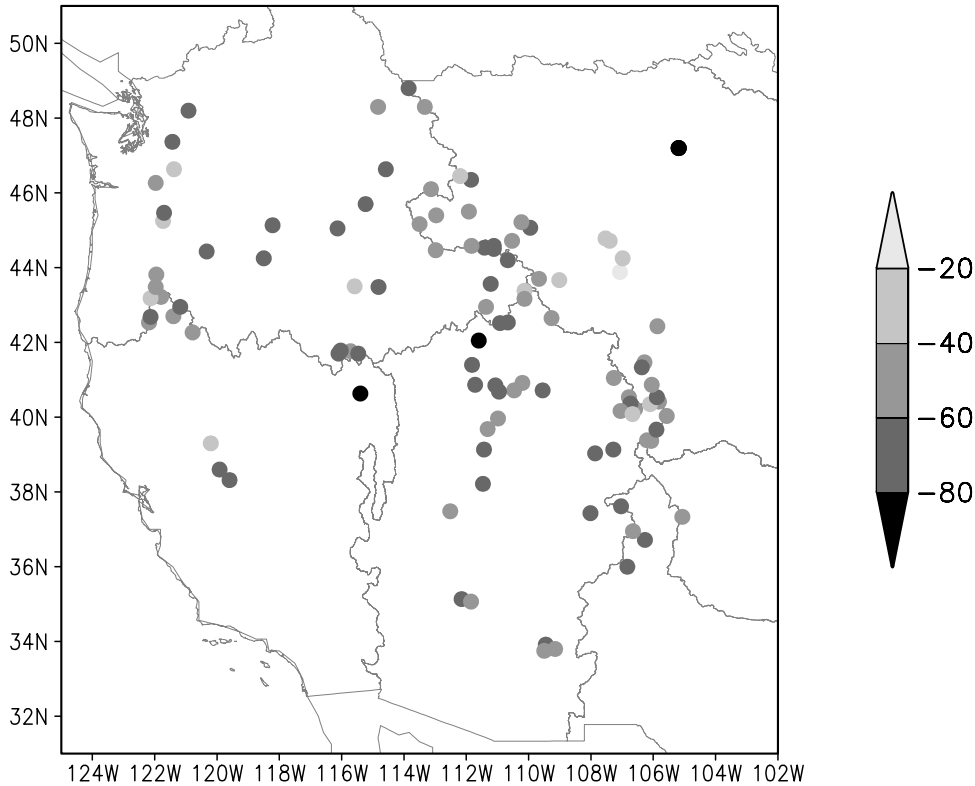

Figure 3. Difference in mean annual precipitation between model simulations and SNOTEL measurements as a percentage of the SNOTEL value (%).

those days when model simulated SWE periodically melts and reappears due to fluctuations around freezing temperatures. It can be seen that all models simulate the onset of snow accumulation later than the measured data. The consistency between models is to be expected as they are all forced by the same precipitation and air temperature and employ the same temperature criteria for partitioning the precipitation into snowfall and rainfall. Small differences between the models are due to how the models handle the accumulation/melt process and the effect of temperature lapsing over sub-grid elevation banding in the VIC model. In terms of the last day of snow, the Noah and SAC models tend to under-estimate the last day of snow and are thus melting the snowpack too early. The same is true of the MOSAIC model but to a lesser extent. Noah tends to have a shorter snow period (later start and earlier end), and this is consistent with the lower maximum SWE for the Noah model that was indicated previously. The analysis of snow cover extent in the first part of this evaluation of the NLDAS cold season modeling [*Sheffield et al.*, 2003] showed that the way in which the models calculate the snow albedo may be key in explaining the differences in simulated snow cover. In that study, the Noah model was found to estimate low snow albedo values relative to the other models. Lower albedo leads to greater absorption of downward solar radiation by the snowpack and thus more available energy for snowmelt. This may explain why the Noah model tends to melt the snowpack faster than the other models. Some of the differences may also be due, in part, to the location of SNOTEL stations which are generally found in clearings in which the ground snowpack is much easier to initiate at small snow events than under heavy vegetation covers.

4.4. Detailed Comparison at a Selected Site

[22] The previous analysis indicated large biases in the model simulated annual maximum SWE and showed that

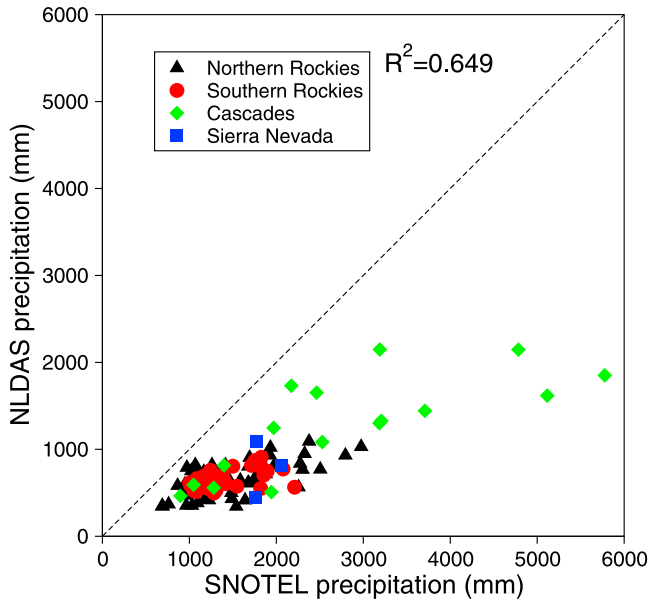

Figure 4. Comparison of NLDAS and SNOTEL mean annual precipitation (mm).

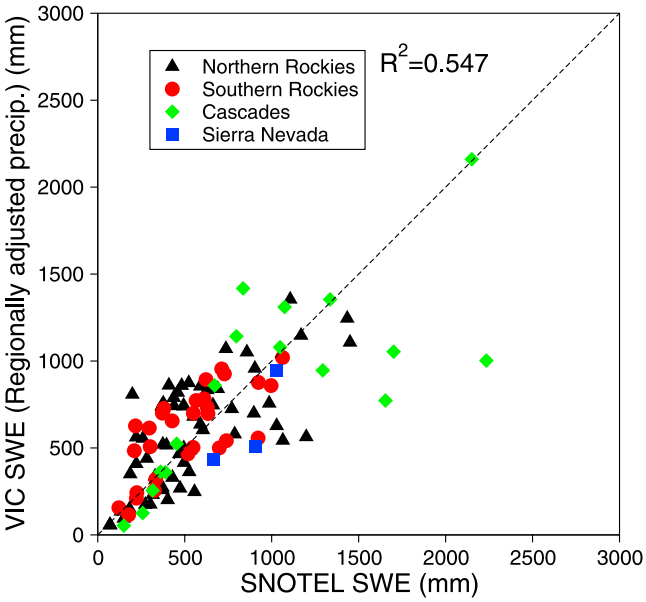

Figure 5. Comparison of mean annual maximum SWE between model simulations forced with regionally adjusted precipitation and SNOTEL measurements (mm).

these biases can be explained largely by errors in the NLDAS precipitation forcing. Furthermore, seasonal biases in air temperature exist and may in turn account for some of the remaining errors. To understand the seasonal dynamics of the snowpack and how temperature biases may affect its evolution, a single site was selected for detailed analysis. The Lone Pine site (station ID = LPSW1, latitude = 46.267N, longitude = 121.967W, elevation = 1158.24 m) was chosen as it had one of the largest errors in maximum SWE of all the stations. It is located in the Cascade Mountains and receives over 5000 mm of precipitation annually during the simulation period. The relatively complex topography and the spatially variable meteorology associated with this terrain ensure that modeling cold season processes in such a region is challenging.

[23] After removing the bias in the precipitation by scaling the model precipitation forcing to match the annual precipitation total at the SNOTEL site, the simulated SWE still showed large errors when compared with the SNOTEL measurements (see Figure 6, Lone Pine lies to the lower right of the cloud of points). Figure 11 shows the 3-year accumulation time series of several water balance components for the SNOTEL measurements and the VIC model simulation using the locally adjusted precipitation forcing. The underestimation of SWE in the VIC model simulation can be clearly seen. The total precipitation is the sum of the snowfall and rainfall, i.e., the solid and liquid precipitation, and the time series of snowfall and rainfall shown in Figure 11 indicate that the total amount of precipitation is sufficient to produce the measured SWE. This is to be expected as the model precipitation forcing was adjusted to match the measured annual total. As accumulation of SWE is governed by snowfall inputs, it appears that the snowfall component is underestimated, especially in the second and third years. Therefore the underestimation of SWE may be attributed, in part, to the partitioning of precipitation into snowfall and rainfall. All models within the NLDAS use a threshold air temperature of 0°C to partition the precipitation inputs, such that if the air temperature is above this value then the precipitation is considered to be rainfall and conversely, it is considered to be snowfall if the air temperature is below this value. Figure 12 shows the average monthly air temperature bias of the NLDAS compared to the SNOTEL measurements and indicates a consistent positive bias in the NLDAS data that is most prominent in the summer. Figure 13 shows the monthly number of freezing days (below 0°C) and non-freezing days (above 0°C) for the NLDAS air temperature and the SNOTEL measurements. Over the 3-year period the NLDAS data had 241 freezing days compared to the 324 freezing days in the SNOTEL measurement record (43 days on which the SNOTEL record was missing were excluded). By using the model precipitation partitioning scheme, this bias would lead to a general underestimation of snowfall. Furthermore, warm rain falling onto an existing snowpack causes an increase in the rate of snowmelt as more heat energy is available for transfer to the underlying snowpack [*Harr*, 1981; *Kattelmann*, 1987; *Berg et al.*, 1991]. These factors and the direct effect of warm-biased air temperatures, combine to produce an underestimation of SWE.

5. Discussion and Conclusions

[24] Simulated snow water equivalent from the four models within the NLDAS was compared with measured data from 110 SNOTEL sites situated in the western United States for a 3-year retrospective period. All models showed consistent underestimation of maximum annual SWE with an average regional bias of −62.0% that can be explained to

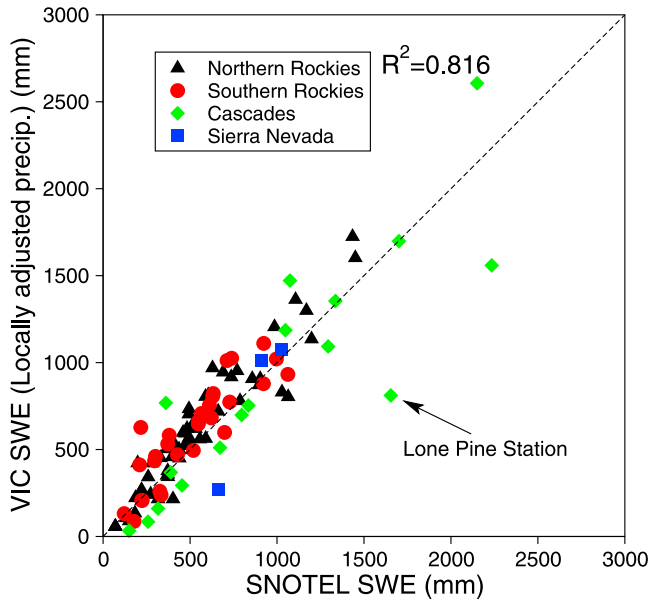

Figure 6. Comparison of mean annual maximum SWE between model simulations forced with locally adjusted precipitation and SNOTEL measurements (mm).

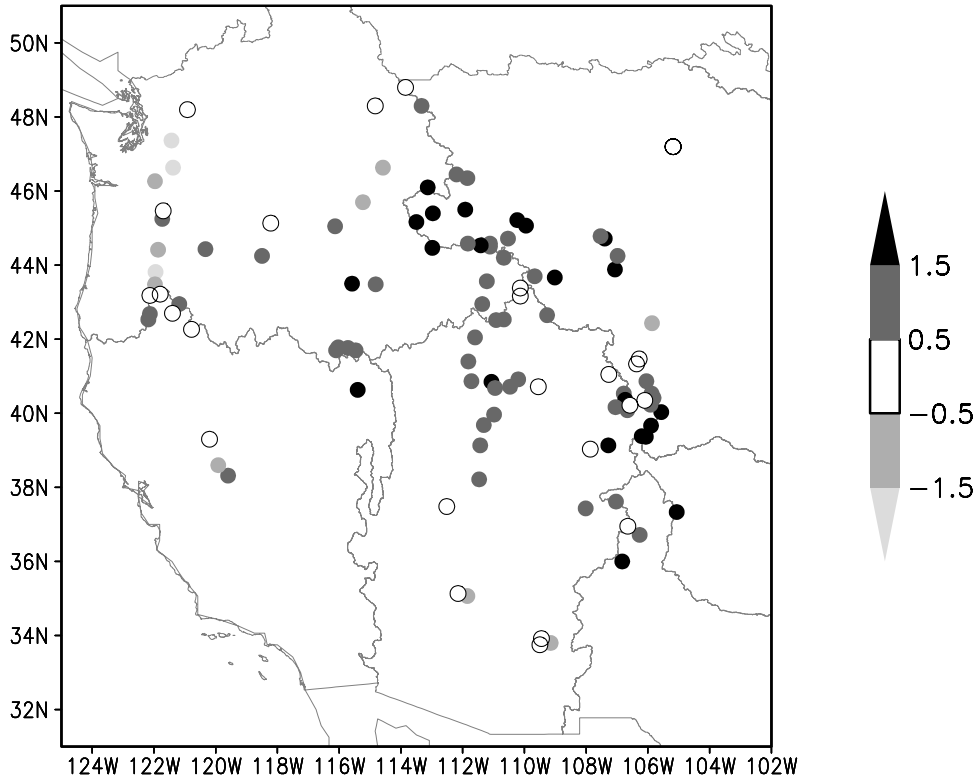

Figure 7. Difference between NLDAS and SNOTEL mean cold season (Nov–Apr) air temperature (°C).

a large extent by biases in the prescribed NLDAS precipitation model forcing (average regional bias = −57.9%). Experiments with the VIC model using regionally adjusted precipitation based on the SNOTEL measurements reduced the bias to 4.0%, although significant differences remained at individual stations as would be expected. Using locally adjusted precipitation removed much of the scatter at individual stations and increased the correlation R^2 from 0.530 to 0.855. Additionally, relatively smaller biases (but potentially significant in terms of cold season processes) were also identified in the air temperature forcing. The smaller number of freezing days in the NLDAS temperature series suggest that the partitioning by the models of precipitation into snowfall and rainfall may result in too much rain falling onto the existing snowpack and causing an overestimation of snowpack melt and thus an underestimation of SWE.

[25] The NLDAS runs over the coterminous USA at a fixed resolution of 1/8 degree and as such is restricted in how it can represent the variability in terrain and meteorology that is found in mountainous regions. This is a recognized limitation, but it is also a fixed parameter of the modeling framework because of the potential assimilation activities with NWP models and the computational and operational limitations that are imposed on a system that is, in the end, intended to run in near real time. The NLDAS models are designed to run at these spatial scales, which are, in general, larger than the scale of many cold season processes. The evolution of the snowpack in mountainous regions is affected by a wide variety of processes and land surface charateristics, including elevation and shading effects, redistribution by wind and the effect of variability in vegetation cover.

[26] Running models at finer resolutions over large domains is challenging due to the lack of reliable fine scale data for vegetation, soils and meteorological forcings and the limitation of computational power. In practice, operational modeling over such large regions can only be done using the types of land surface models and resolu-

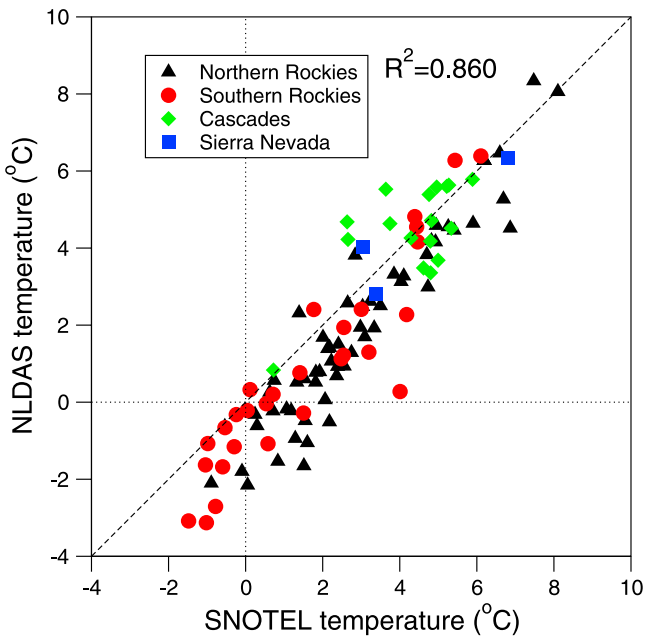

Figure 8. Comparison of NLDAS and SNOTEL mean cold season (Nov–Apr) air temperature (°C).

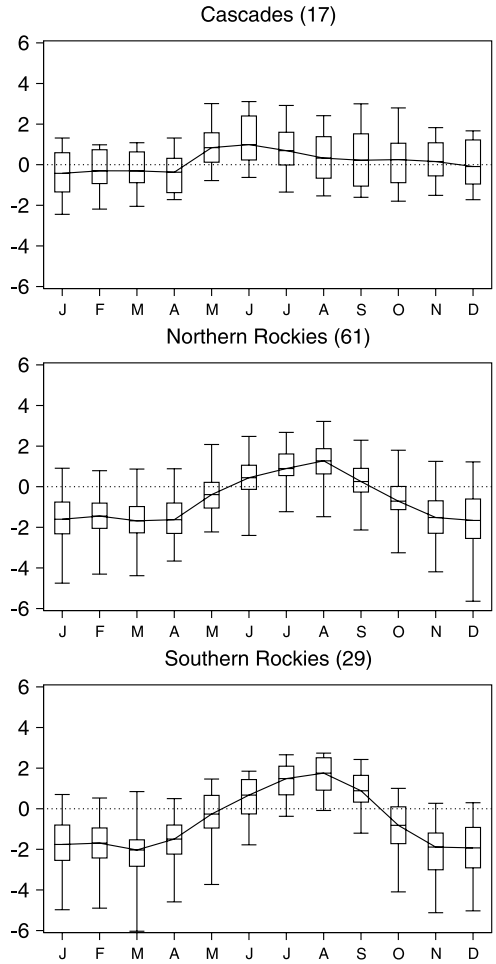

Figure 9. Box and whisker plots of the NLDAS temperature bias for the Cascades, Northern Rockies and Southern Rockies regions. Data for the Sierra Nevada region are not included as only 3 stations remained after elevation screening. The numbers in parenthesis indicate the number of stations within each region. The solid line represents the mean temperature differences between the NLDAS and SNOTEL data over all the sites in the region. The upper and lower error bars represent the maximum and minimum bias whilst the upper and lower limits of the boxes are the 25% and 75% quartiles.

tions used in the NLDAS. Large-scale models by their nature do not directly model the micro-physics of the small scale. Instead a variety of modeling methods are employed to represent sub-grid variability such as calibration, effective parameters, parameterizations of variability, and sub-grid tiling [*Bloschl*, 1999; *Luce et al.*, 1998]. Although it is acknowledged that this scale of modeling, even with the most advanced sub-grid representations, will not capture the true sub-grid variability of cold season mountain hydrology, it provides reasonable estimates by capturing the grid scale gross effect of the small-scale processes.

[27] The NLDAS models implement parameterizations of sub-grid variability, but to different degrees of complexity, from multiple parameterizations to the use of no parameterizations: VIC models sub-grid variability of vegetation, storm coverage and elevation, including the effects on precipitation and temperature; the MOSAIC and Noah models use sub-grid vegetation tiling only; the SAC model uses no sub-grid variability representations. To some extent, the effect of using different levels of sub-grid complexity is indicated by the inter-comparison of the four models, which use different representations of sub-grid variability, ranging from no representation (SAC) to multiple parameterizations (VIC). However, at the annual and regional scale, the inter-comparison reveals that no one model does particularly better or worse than another. An inter-comparison of all the models forced with corrected precipitation may reveal more details but this is beyond the scope of this study. The comparison of the simulation of the onset and disappearance of the snowpack (Figure 10) shows larger inter-model differences, with the VIC model performing the best and the SAC and Noah models tending to melt the snowpack too early. Whether this is a result of the differing treatment of sub-grid variability is unclear and may actually be a result of other factors. The analysis of NLDAS simulated snow cover extent by *Sheffield et al.* [2003] showed that the SAC model performed particularly well despite having no sub-grid variability representation in its model structure and this could be attributable to its simplistic use of air temperature as the sole driving force for the development of the snowpack. The study also showed that the under estimation of snow cover by the Noah model was likely attributable to its under estimation of snow albedo and not necessarily its lack of complex sub-grid parameterizations.

[28] In general, the use of point data for the evaluation of gridded data is problematic because of scaling issues. Given the lack of reliable measurements of SWE over large spatial scales, one has to resort to point data to be able to carry out any evaluation of the NLDAS model performance and the modeling framework. The NLDAS forcings and model simulated data are representative of a relatively coarse grid cell, and as such, cannot give a true representation of the behaviour at the SNOTEL sites. However, a number of factors instill confidence in the use of SNOTEL for the evaluation of grid-based data. Firstly the SNOTEL sites are located so as to be representative of the larger scale in terms of streamflow. Secondly, the stations were screened to remove those that are at elevations more than 50 m different from the corresponding grid cell, therefore removing any biases that would be a result of the effect of elevation on the meteorological forcings. Thirdly, geostatistical analysis of a remote sensing based estimate of SWE showed that the NLDAS grid size is within the spatial correlation length indicating that the small scale is representative of the grid mean value. Lastly, the analysis was carried out over large time and space scales, thus removing the small-scale variability in the data that underlies the scale-dependent differences.

[29] The results presented in this paper complement and are consistent with the findings of related studies of the NLDAS. The first part of this paper, which discusses the evaluation of the NLDAS simulated snow cover extent [*Sheffield et al.*, 2003], found that, in general, the MOSAIC and Noah models underestimated snow cover extent, the VIC model overestimated and the SAC model

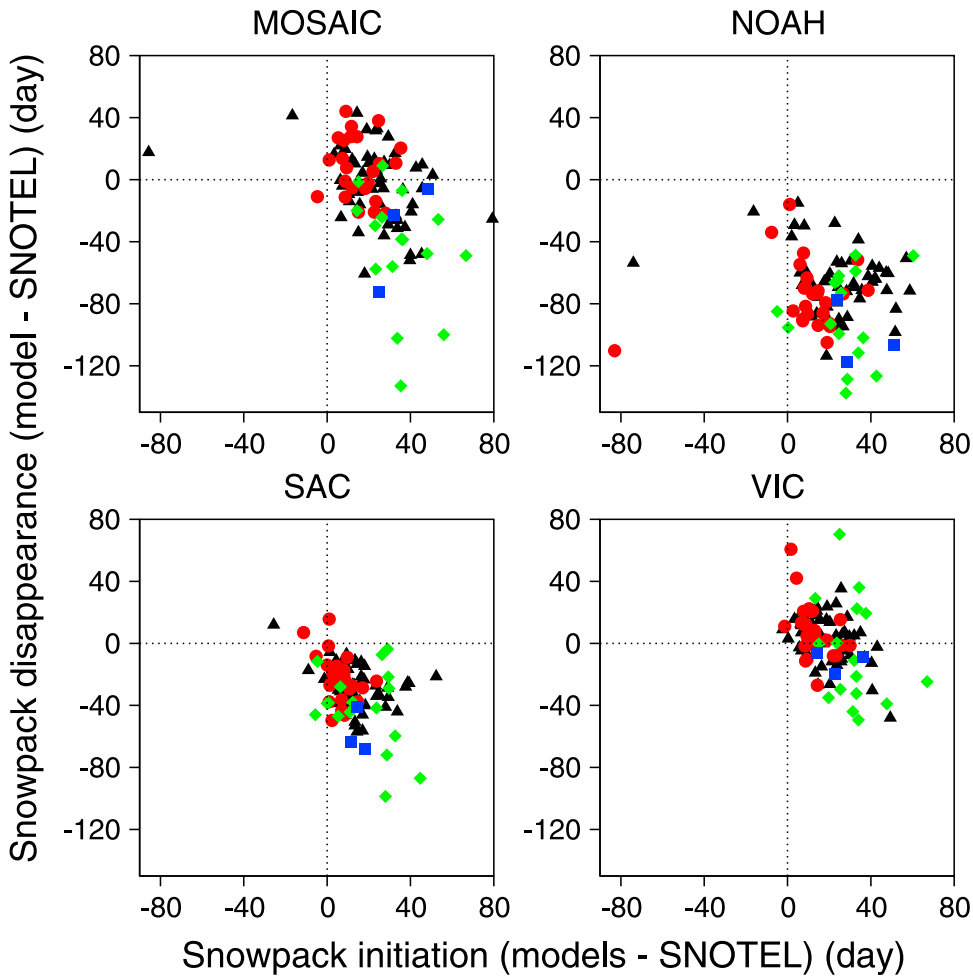

Figure 10. Difference in model simulated and SNOTEL measuremed first day and last day of snow.

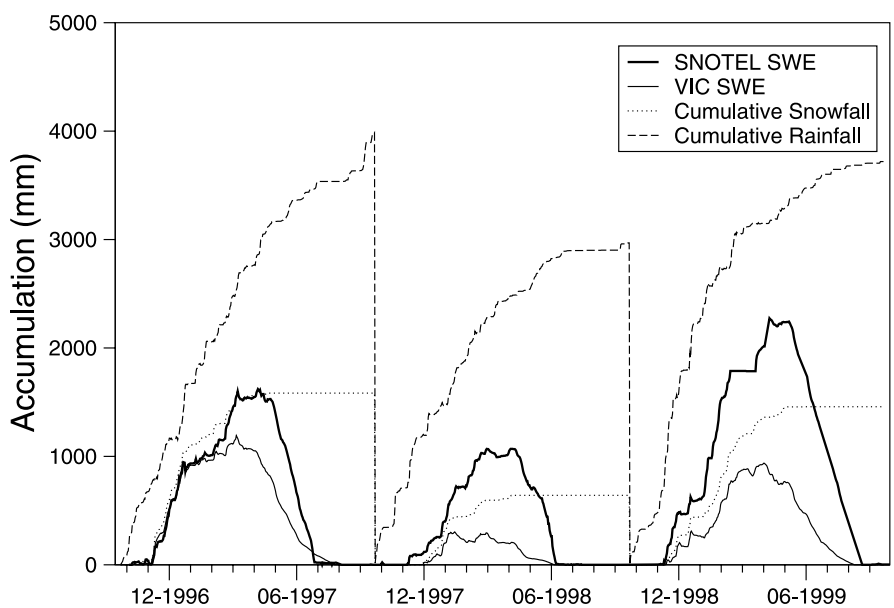

Figure 11. Time series of SNOTEL SWE, VIC simulated SWE and NLDAS accumulated snowfall and rainfall for the Lone Pine station (latitude = 46.267, longitude = −121.967, elevation = 1158.24 m, Cascades region).

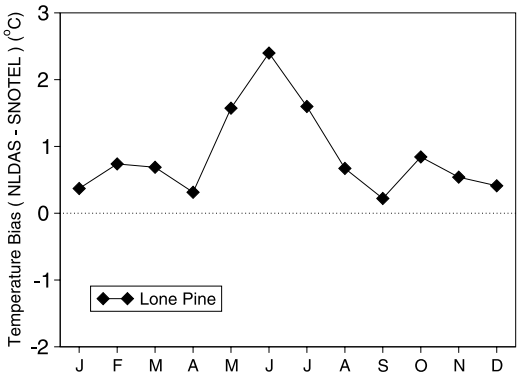

Figure 12. Difference between SNOTEL and NLDAS average monthly mean air temperature at the Lone Pine station.

was essentially unbiased. Over the mountainous western part of the US all models tended to under estimate snow cover extent, supporting the conclusion that there is a regional low bias in the NLDAS precipitation. *Lohmann et al.* [2003] analyzed simulated streamflow and water balance of the NLDAS simulations. They found that all models underestimate runoff in the NW quadrant of the US, which contains the majority of the SNOTEL stations used in the present study. More specifically, the mean annual runoff for most basins in the northern Rockies was underestimated by 20–80%. This is consistent with the conclusion that the NLDAS precipitation is underestimated in the mountainous western US. In terms of peak streamflow timing, it was found that over most of the country the models simulate the streamflow peak within ±3 days. However, in the Rockies and the northeast region, the Noah model simulated many peaks more than 2 months early, the Mosaic and SAC model had errors of about 1 month and the VIC model simulated the timing reasonably well. The dominance of cold season processes in this region indicates that 3 out of 4 of the models tended to underestimate the timing of snowmelt. This ties in with the findings in this study that Noah and SAC, and to some extent Mosaic, tend to melt the snowpack too early (see Figure 10). The warm spring bias found in the NLDAS air temperature when compared to the SNOTEL sites is a likely contributing factor to early onset of snowmelt and the findings of *Lohmann et al.* [2003] indicate that the bias may apply over regional scales.

[30] This study has shown that reasonable estimates of SWE can be obtained through the use of these models in the NLDAS modeling framework but only if the accuracy of meteorological forcings is improved. Although the experiments with the adjusted precipitation were only carried out with the VIC model, there is no reason to believe that the other models in the NLDAS cannot do a similar job. The analysis highlights the problems in determining meteorological variables over mountainous and complex terrain, where small numbers of measurement stations generally located at lower elevations, do not capture the full variability of the meteorology. Further work is required to remove the biases from the model forcings, notably precipitation and temperature, so that model inter-comparisons may reveal more detailed information about their ability to represent cold season processes and provide accurate simulations of snow states within data assimilation systems. At present, this may be beyond the capabilities of an operational implementation of NLDAS. However, planned future work intends to build on the findings described in this and related NLDAS evaluation papers to address the biases in the forcings. For example, since February 2002, the real time version of the NLDAS forcing data set, has used the Parameter-elevation Regressions on Independent Slopes Model (PRISM) climatology product [*Daly et al.*, 1994] to account for topographical influences on precipitation

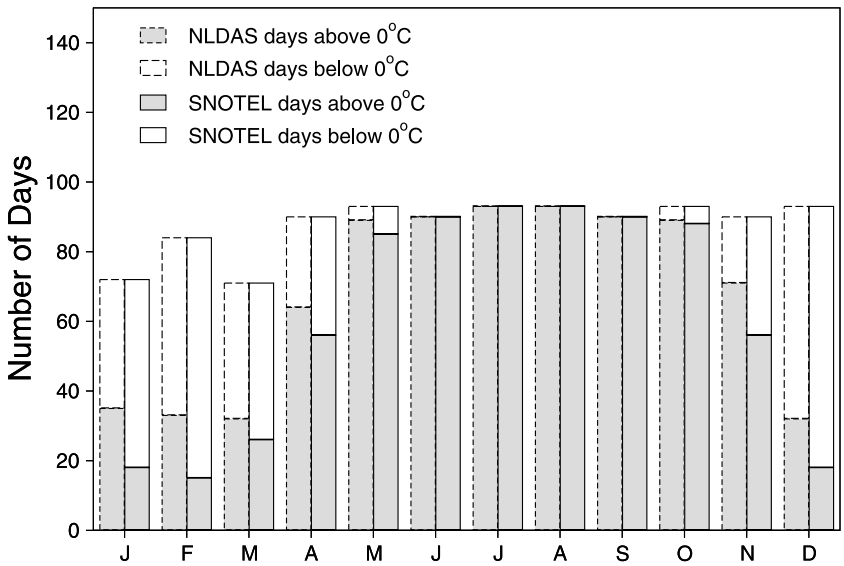

Figure 13. Total monthly number of days with mean daily air temperature above and below 0°C for the period October 1996 to September 1999 at the Lone Pine station for the NLDAS and SNOTEL. The 43 days with missing SNOTEL records (21 days in Jan 1997 and 21 days in Mar 1997) are not included.

[*Cosgrove et al.*, 2003]. The main effect of this is to significantly increase the precipitation amounts at high elevations, a change that would likely benefit the accuracy of the NLDAS simulations when compared against SNOTEL measurements.

[31] In the end the analysis presented in this paper may be a compromise due to the lack of large-scale SWE measurements. By resorting to the use of point measurements, a number of caveats arise that prevent a definitive evaluation of the SWE simulations in the NLDAS. Although this work goes some way to dealing with these issues, they can never be fully addressed. More importantly, this shows the pressing need for large-scale measurements of SWE over mountainous regions that are required for the evaluation of land surface and climate models. Current research into the assimilation of ground observations and remote sensing into detailed distributed snow models [*Wilson et al.*, 1999; *Carroll et al.*, 1999; *Carroll et al.*, 2001] may help alleviate this situation.

[32] **Acknowledgments.** This work was supported by NOAA grant NA86GP0258 "Development of a Hydrologically-Based Land Data Assimilation System for the U.S." (Eric F. Wood, PI). The work on this project by NCEP/EMC, NWS/OHD, and NESDIS/ORA was supported by the NOAA OGP grant for the NOAA Core Project for GCIP/GAPP (co-PIs K. Mitchell, J. Schaake, D. Tarpley). The work by NASA/GSFC/HSB was supported by NASA's Terrestrial Hydrology Program (P. Houser, PI). The work by Rutgers University was supported by NOAA OGP GAPP grant GC99-443b (A. Robock, PI), the Cook College Center for Environmental Prediction, and the New Jersey Agricultural Experiment Station, and additionally, figures were drawn with GrADS, created by Brian Doty. The work by NCEP/CPC was supported by NOAA/NASA GAPP Project 8R1DA114 (R. W. Higgins, PI). The work by University of Maryland was supported by grants NA56GPO233, NA86GPO202 and NA06GPO404 from NOAA/OGP and by NOAA grant NA57WC0340 to University of Maryland's Cooperative Institute for Climate Studies (R. Pinker, PI). The SNOTEL data were downloaded from Western Regional Climate Center website (http://www.wrcc.dri.edu). We appreciate and acknowledge the efforts of the Western Regional Climate Center to make the data available to researchers free of charge via the Web.

References

Anderson, E., National Weather Service river forecast system: Snow accumulation and ablation model, *NOAA Tech. Memo.*, *NWS HYDRO-17*, U.S. Dep. Of Comm., Natl. Oceanic and Atmos. Admin., Washington, D. C., 1973.

Berg, N., R. Osterhuber, and J. Bergman, Rain-induced outflow from deep snowpacks in the central Sierra Nevada, California, *Hydrol. Sci. J.*, *36*, 611–629, 1991.

Betts, A. K., F. Chen, K. Mitchell, and Z. Janjic, Assessment of the land surface and boundary layer models in two operational versions of NCEP Eta model using FIFE data, *Mon. Weather. Rev.*, *125*, 2896–2916, 1997.

Bloschl, G., Scaling issues in snow hydrology, *Hydrol. Proc.*, *13*, 2149–2175, 1999.

Brown, R. D., Northern Hemisphere snow cover variability and change, 1915–1997, *J. Clim.*, *13*, 2339–2355, 2000.

Burnash, R. J. C., The NWS river forecast system: Catchment modeling, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 311–366, Water Resour. Publ., Boulder, Colo., 1995.

Burnash, R. J. C., R. L. Ferral, and R. A. McGuire, A Generalized streamflow simulation system: Conceptual modeling for digital computers, U. S. Dept. of Commerce, National Weather Service and State of Calif., Dept. of Water Resources, Sacramento, Calif., 1973.

Carroll, S. S., and T. R. Carroll, Effect of uneven snow cover on airborne snow water equivalent estimates obtained by measuring terrestrial gamma radiation, *Water Resour. Res.*, *25*(7), 1505–1510, 1989.

Carroll, S. S., T. R. Carroll, and R. W. Poston, Spatial modeling and prediction of snow water equivalent using ground-based, airborne, and satellite snow data, *J. Geophys. Res.*, *104*(D16), 19,623–19,629, 1999.

Carroll, T. R., D. W. Cline, G. Fall, A. Nilsson, L. Li, and A. Rost, NOHRSC operations and the simulation of snow cover properties for the conterminous U.S., Presented at the 69th Annual Meeting of the Western Snow Conference, Sun Valley, Idaho, 16–19 April 2001.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H. Pan, V. Koren, Q. Duan, M. Ek, and A. Betts, Modeling of land surface evaporation by four schemes and comparison with FIFE observations, *J. Geophys. Res.*, *101*(D3), 7251–7268, 1996.

Chen, F., Z. Janjic, and K. Mitchell, Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model, *Boundary Layer Meteorol.*, *85*, 391–421, 1997.

Cherkauer, K., and D. P. Lettenmaier, Hydrologic effects of frozen soils in the upper Mississippi river basin, *J. Geophys. Res.*, *104*(D16), 19,599–19,610, 1999.

Cosgrove, B. A., et al., Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, *J. Geophys. Res.*, *108*(D22), 8842, doi:10.1029/2002JD003118, 2003.

Crook, A. G., SNOTEL: Monitoring climatic factors to predict water-supplies, *J. Soil Water Conserv.*, *32*(6), 294–295, 1977.

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

Entekhabi, D., I. Rodriguez-Iturbe, and F. Castelli, Mutual interaction of soil moisture state and atmospheric processes, *J. Hydrol.*, *184*(1)–(2), 3–17, 1996.

Groisman, P. Y., T. R. Karl, and R. W. Knight, Observed impact of snow cover on the heat balance and rise of continental spring temperatures, *Science*, *263*, 198–200, 1994a.

Groisman, P. Y., T. R. Karl, R. W. Knightand, and G. L. Stenchikov, Changes of snow cover, temperature, and radiative heat balance over the Northern Hemisphere, *J. Climate*, *7*, 1633–1656, 1994b.

Hall, D. K., G. A. Riggs, V. V. Salomonson, N. E. DiGirolamo, and K. J. Bayr, MODIS snow-cover products, *Remote Sens. Environ.*, *83*(1)–(2), 181–194, 2002.

Harr, R. D., Some characteristics and consequences of snowmelt during rainfall in western Oregon, *J. Hydrol.*, *53*, 277–304, 1981.

Hartman, R. K., A. A. Rost, and D. M. Anderson, Operational processing of multi-source snow data, in *Proceedings of the Western Snow Conference*, pp. 147–151, Sparks, Nevada, 1995.

Higgins, R. W., W. Shi, and E. Yarosh, Improved United States precipitation quality control system and analysis, NCEP/Climate Prediction Center Atlas, 7, 40 pp., 2000.

Kattelmann, R. C., Water release from a forest snowpack during rainfall, *IAHS Publ.*, *167*, 265–272, 1987.

Koren, V., J. Schaake, K. Mitchell, Q. Duan, F. Chen, and J. M. Baker, A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, *J. Geophys. Res.*, *101*, 7251–7268, 1999.

Koster, R. D., and M. J. Suarez, Energy and water balance calculations in the MOSAIC LSM, *NASA Tech. Memo.*, *104606*, *9*, NASA Goddard Space Flight Center, Greenbelt, Md., 1996.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, A Simple hydrologically based model of land surface water and energy fluxes for GSMs, *J. Geophys. Res.*, *99*(D7), 14,415–14,428, 1994.

Liang, X., D. P. Lettenmaier, and E. F. Wood, One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model, *J. Geophys. Res.*, *101*(D16), 21,403–21,422, 1996.

Liang, X., E. F. Wood, and D. P. Lettenmaier, Modeling ground heat flux in land surface parameterization schemes, *J. Geophys. Res.*, *104*(D8), 9581–9600, 1999.

Lohmann, D., et al., Streamflow and water balance intercomparisons of four land-surface models in the North American Land Data Assimilation System project, *J. Geophys. Res.*, *108*, in press, 2003.

Luce, C. H., D. G. Tarboton, and K. R. Cooley, The influence of the spatial distribution of snow on basin-averaged snowmelt, *Hydrol. Proc.*, *12*, 1671–1683, 1998.

Mitchell, K., et al., The GCIP Land Data Assimilation System (LDAS) project: Now underway, *GEWEX News*, *9*(4), 1, 7–11, 1999.

Mitchell, K., et al., The collaborative GCIP land data assimilation (LDAS) project and supportive NCEP uncoupled land-surface modeling initiatives, presented at the 15th AMS Conference on Hydrology, Am. Meteorol. Soc., Long Beach, Calif., 2000.

Mitchell, K. E., et al., The Multi-institution North American Land Data Assimilation System (NLDAS) project: Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res.*, *108*, in press, 2003.

Namias, J., Some empirical evidence for the influence of snow cover on temperature and precipitation, *Mon. Weather Rev.*, *113*(9), 1542–1553, 1985.

Pinker, R. T., et al., Surface radiation budgets in support of the GEWEX Continental-Scale International Project (GCIP) and the GEWEX Americas Prediction Project (GAPP), including the North American Land Data

Assimilation System (NLDAS) project, *J. Geophys. Res.*, *108*(D22), 8844, doi:10.1029/2002JD003301, in press, 2003.

Ramsay, B., The interactive Multisensor snow and ice mapping system, *Hydrol. Processes*, *12*, 1537–1546, 1998.

Schultz, D. M., et al., Understanding Utah winter storms: The Intermountain Precipitation Experiment, *Bull. Am. Meteorol. Soc.*, *83*, 189–210, 2002.

Serreze, M. C., M. P. Clark, R. L. Armstrong, D. A. McGinnis, and R. S. Pulwarty, Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, *Water Resour. Res.*, *35*(7), 2145–2160, 1999.

Sheffield, J., et al., Snow process modeling in the North American Land Data Assimilation System (NLDAS): 1. Evaluation of model-simulated snow cover extent, *J. Geophys. Res.*, *108*(D22), 8849, doi:10.1029/2002JD003274, in press, 2003.

Simpson, J. J., J. R. Stitt, and M. Sienko, Improved estimates of the areal extent of snow cover from AVHRR data, *J. Hydrol.*, *204*, 1–23, 1998.

Wigmosta, M. S., L. W. Vail, and D. P. Lettenmaier, A distributed hydrology-vegetation model for complex terrain, *Water Resour. Res.*, *30*, 1665–1679, 1994.

Wilson, L. L., L. Tsang, J.-N. Hwang, and C.-T. Chen, Mapping snow water equivalent by combining a spatially distributed snow hydrology model with passive microwave remote-sensing data, *IEEE Trans. Geosci. Remote Sens.*, *37*, 690–704, 1999.

Yeh, T-C., R. T. Wetherald, and S. Manabe, A model study of the short-term climatic and hydrologic effects of sudden snow-cover removal, *Mon. Weather Rev.*, *111*(5), 1013–1024, 1983.

B. Cosgrove and P. R. Houser, Hydrological Sciences Branch, NASA Goddard Space Flight Center, Mail Code 974, Greenbelt, MD 20771, USA.

Q. Duan and J. C. Schaake, Office of Hydrologic Development, NOAA/NWS, Silver Spring, MD 20910, USA.

R. W. Higgins, NCEP Climate Prediction Center (NOAA/NWS), Camp Springs, MD, USA.

D. Lohmann and K. E. Mitchell, NCEP Environmental Modeling Center, NOAA/NWS, Camp Springs, MD 20746-4304, USA.

M. Pan, L. Luo, J. Sheffield, and E. F. Wood, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, USA. (justin@princeton.edu)

R. T. Pinker, Department of Meteorology, Space Sciences Building, University of Maryland, College Park, MD 20742, USA.

A. Robock, Department of Environmental Sciences, Rutgers University, New Brunswick, NJ 08901-8551, USA.

J. D. Tarpley, Office of Research and Applications, World Weather Building, NOAA/NESDIS, Camp Springs, MD 20746-4304, USA.