Assessment of Forecasts during Persistent Valley Cold Pools in the Bonneville Basin by the North American Mesoscale Model

HEATHER DAWN REEVES AND KIMBERLY L. ELMORE

National Severe Storms Laboratory, and Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

GEOFFREY S. MANIKIN

National Centers for Environmental Prediction/Environmental Modeling Center, Camp Springs, Maryland

DAVID J. STENSRUD

National Severe Storms Laboratory, Norman, Oklahoma

(Manuscript received 12 August 2010, in final form 16 December 2010)

ABSTRACT

North American Mesoscale Model (NAM) forecasts of low-level temperature and dewpoint during persistent valley cold pools in the Bonneville Basin of Utah are assessed. Stations near the east sidewall have a daytime cold and nighttime warm bias. This is due to a poor representation of the steep slopes on this side of the basin. Basin stations where the terrain is better represented by the model have a distinct warm, moist bias at night. Stations in snow-covered areas have a cold bias for both day and night. Biases are not dependent on forecast lead or validation time. Several potential causes for the various errors are considered in a series of sensitivity experiments. An experiment with 4-km grid spacing, which better resolves the gradient of the slopes on the east side of the basin, yields smaller errors along the east corridor of the basin. The NAM assumes all soil water freezes at a temperature of 273 K. This is likely not representative of the freezing temperature in the salt flats in the western part of the basin, since salt reduces the freezing point of water. An experiment testing this hypothesis shows that reducing the freezing point of soil water in the salt flats leads to an average error reduction between 1.5 and 4 K, depending on the station and time of day. Using a planetary boundary layer scheme that has greater mixing alleviates the cold bias over snow somewhat, but the exact source of this bias could not be determined.

1. Introduction

Valley cold pools (VCPs), which are shallow layers of cold air trapped in a valley or basin (Whiteman et al. 2001), are common in the western United States during winter. Numerical forecasts of basic variables, such as temperature and humidity, are known to be problematic during VCPs, which makes for difficulty in anticipating the various forms of hazardous weather, such as fog or freezing rain that can occur. In this study, the North American Mesoscale Model (NAM) forecasts for three winter seasons are assessed to determine the typical errors observed during persistent VCP events, and sensitivity experiments are performed to detect possible sources of error.

There is no exact definition of or agreed-upon way to define or identify a VCP. Usually, low-level static stability must exceed a certain threshold, such as the presence of an inversion (Wolyn and McKee 1989; Whiteman et al. 2001; Reeves and Stensrud 2009) and some studies require low-level winds to be less than a prescribed value [generally 5 m s^{-1} or less; Whiteman et al. (2001); Reeves and Stensrud (2009)]. Valley cold pools may be classified as either diurnal or persistent. Diurnal VCPs form at night from radiational cooling and are removed the next day by insolation (e.g., Lenschow et al. 1979; Banta and Cotton 1981; Whiteman 1982; Whiteman and McKee 1982; Bader and McKee 1985; Vrhovec 1991; Fast et al. 1996). Persistent VCPs, which are those that last longer than one

Corresponding author address: Heather Dawn Reeves, DOC/NOAA/OAR, National Severe Storms Laboratory, 120 David L. Boren Blvd., Ste. 2401, Norman, OK 73072-7319.
E-mail: heather.reeves@noaa.gov

DOI: 10.1175/WAF-D-10-05014.1

diurnal cycle, may by partially enhanced by radiative effects, but are predominantly forced by large-scale subsidence and/or midlevel warm-air advection (Wolyn and McKee 1989; Whiteman et al. 1999, 2001; Zhong et al. 2001; Zängl 2005; Hoggarth et al. 2006; Reeves and Stensrud 2009). Persistent VCPs can occur in any valley at any time of year, but in the intermountain United States they are most common in winter and have been observed to last several days or even up to 3 weeks (Wolyn and McKee 1989; Whiteman et al. 1999; Reeves and Stensrud 2009).

Because the air within a VCP is decoupled from the free atmosphere by a capping stable layer, VCP winds are generally quite weak. Very long-lived VCPs are often associated with severe air pollution and dense fog due to the lack of vertical mixing associated with the weak low-level winds (e.g., Hill 1993; Smith et al. 1997; Pataki et al. 2005; Struthwolf 2005). Warm-cloud precipitation systems moving over a subfreezing VCP can result in freezing rain or drizzle while extended periods of cold can delay snow and ice melt. Accurate forecasting of these forms of hazardous weather requires reliable forecasting of the low-level temperature and humidity. Forecasts for several persistent VCP events are presented in the literature. These studies indicate that the low-level temperatures are generally too high, leading to vertical mixing that is too strong and, consequently, VCPs that are removed too early (Hart et al. 2004, 2005; Billings et al. 2006; Myrick and Horel 2006; Reeves and Lin 2006; Cheng and Steenburgh 2007). However, there have been no long-term multi-episode evaluations of operational model forecasts during VCPs, so characteristic model errors are not directly known.

The types of errors (i.e., magnitude, diurnal variations, etc.) one encounters may be endemic to the numerical model design and setup. For example, comparisons of the local closure Mellor–Yamada–Janjić (MYJ; Mellor and Yamada 1982; Janjić 2002) and nonlocal closure Yonsei University (YSU; Hong et al. 2006) planetary boundary layer (PBL) schemes show the MYJ scheme has weaker mixing and lower low-level temperatures during relatively quiescent flow, such as is dominant during VCPs (e.g., Pagowski 2004; Hu et al. 2010). Berg and Zhong (2005) also note low-level temperatures and humidity are dependent on the choice of PBL scheme. Improper specification of the land use (Zehnder 2002), vegetation, and/or soil (Rife et al. 2002; Kurkowski et al. 2003; Knievel et al. 2007), and/or incorrect initial estimates of soil temperature and/or moisture (Dirmeyer et al. 2000; Godfrey and Stensrud 2008), can affect low-level temperature and humidity forecasts via changes in the latent heat flux. Zängl (2002) found that the two-sided horizontal diffusion used in some numerical models may act to unrealistically mix warmer air above a VCP with the colder air in the VCP, leading to low-level temperatures that are too high. He proposed using a one-sided scheme. Billings et al. (2006) tested this scheme for a single event and noted an improved forecast (temperatures were between 4 and 16 K closer to the observations depending on the time of day and location in the valley). Although they did not specifically consider VCPs, Müller and Scherer (2005) and Liu et al. (2008) found that low-level temperature forecasts in winter are improved if the shading of insolation by terrain is accounted for in the shortwave radiation scheme. Finally, changes in the horizontal resolution are known to affect low-level temperature and humidity, with decreased grid spacing generally being associated with better forecasts (Chen et al. 2004; Billings et al. 2006).

In this study, the operational NAM (Janjić et al. 2005) forecasts are assessed to determine its characteristic errors during persistent VCPs. This model has been run operationally over the continental United States since June 2006. Given the 12-km grid spacing of the NAM, there are a limited number of valleys and basins that are resolved and that have sufficient surface observations with which to validate model forecasts. In this study, the Bonneville Basin of Utah (represented by the green shading in Fig. 1a) is used. The basin is resolved by the NAM (Fig. 1b), but the Salt Lake Valley (shown by the black outline in the southeast corner of the Bonneville Basin in Fig. 1a) and many of the narrow north-to-south ridges surrounding the basin are not. The National Weather Service is currently running experimental forecasts with a 4-km grid spacing. This grid spacing better resolves the Salt Lake Valley and some of the north–south ridges (Fig. 1c), but a long time series of these forecasts is not available.

This paper is organized as follows. First, we assess the characteristic errors from a statistical point of view using forecasts from the 2006/07, 2007/08, and 2008/09 winter months (October–March). A description of the NAM model forecasts considered and the results of the statistical analysis are provided in section 2. This is followed in section 3 by a discussion of the probable sources of model error and a presentation of numerical sensitivity experiments that isolate various potential causes of errors. Concluding thoughts are given in section 4.

2. Comparison of forecasts to observations

a. The NAM and identification of VCP events

The NAM forecasts are produced using the Weather Research and Forecasting Nonhydrostatic Mesoscale Model (WRF-NMM; Janjić et al. 2005). The NAM has

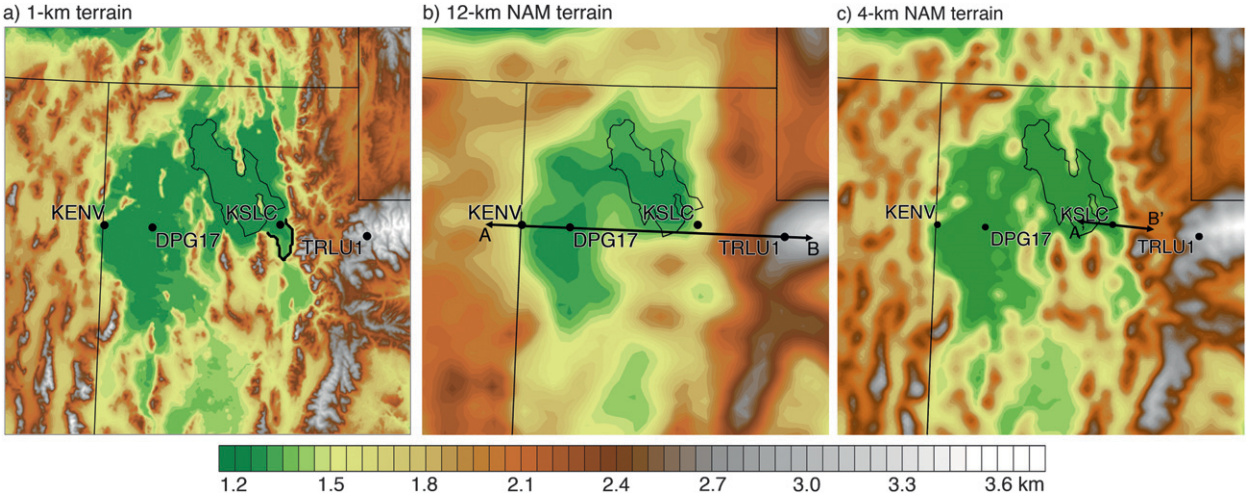

FIG. 1. The (a) 30-s terrain height according to the U.S. Geological Survey; and (b) 12- and (c) 4-km NAM terrain heights. The locations of the four stations considered are indicated. The black outline in (a) indicates the location of the Salt Lake Valley.

a horizontal grid spacing of 12 km, 60 terrain-following vertical levels, and is initialized at 0000, 0600, 1200, and 1800 UTC daily. Its initial conditions are generated by blending the 12-h forecast from the global forecast model with assimilated variables including surface, aircraft, satellite, and upper-air observations using a three-dimensional variational-analysis scheme. This model uses the Ferrier microphysical (Ferrier 1994), Betts–Miller–Janjić cumulus (Betts and Miller 1986; Janjić 1994), Geophysical Fluid Dynamics Laboratory short- and longwave radiation (Fels and Schwarzkopf 1975; Ramaswamy and Freidenreich 1998), and MYJ boundary layer parameterization schemes and the Noah land surface model (LSM; Ek et al. 2003). The NAM also accounts for orographic shading in the shortwave radiation scheme and uses a one-sided horizontal diffusion along terrain slopes. No major alterations were made to the NAM code or choice of parameterization schemes during the time period considered. There were minor changes to some parameterization schemes during the time period considered; however, the trend of errors with time was considered and we found there was no change in the mean or standard deviation of errors at any station after a change to the model source code.

Valley cold pools are defined as in Reeves and Stensrud (2009). Namely, a VCP is identified when the 12-hourly observed Salt Lake City sounding (KSLC; Fig. 1a) has a temperature inversion below 3.4 km (which is close to the maximum ridge crest height on the east side of the basin) and a layer-average wind speed below the inversion top of less than 5 m s^{-1}. A persistent VCP is defined as having three or more consecutive soundings that meet these criteria. This three-consecutive minimum removed from the collection any cases of shallow, short-lived, or diurnal-type VCPs. During the three winter seasons considered, there are 38 persistent VCPs. Only forecasts and analyses that are valid during these VCPs are considered. Also, only forecasts from model integrations started at 0000 and 1200 UTC and only the 0-, 12-, 36-, and 72-h lead times are assessed. (Forecasts beyond a 60-h lead time are not archived prior to December 2006, so the 72-h assessments only include data after December 2006.) If the observation at a given time is flagged as being suspect, that time is not included in the analyses.

Observed data from select surface stations in Utah are compared to the NAM forecasts. The stations used are Wendover Air Force Auxiliary Field (KENV), Dugway Proving Ground number 17 (DPG17), Salt Lake City International Airport (KSLC), and Trial Lake (TRLU1; see Fig. 1). These sites are chosen because they provide the most continuous measurements throughout the time period and have very infrequent episodes of data loss, they are located in different parts of the basin (near the west sidewall, near the middle of the basin, near the east sidewall, and outside of the basin), and, with the exception of KSLC (see below), the four model grid points surrounding each station are at comparable altitudes to the stations' altitudes. KENV, DPG17, and KSLC use the Automated Surface Observing System (ASOS; National Oceanic and Atmospheric Administration 1998). Air temperature and humidity are measured 2 m above ground. The TRLU1 station is a snow telemetry (SNOTEL) site operated by the Natural Resources Conservation Service. The height of the thermometer is unspecified, but the SNOTEL sites generally have heights ranging from 2 to 10 m above ground in order to prevent them from being buried by snow. Pepin

et al. (2005) note that differences between SNOTEL-observed temperatures and other sources are $O(0.1\ \text{K})$.

The model data are interpolated to the latitude and longitude of the observation sites using the inverse-distance Cressman method:

$$T = \frac{\sum_{n=1,4} W_n T_n}{\sum_{n=1,4} W_n}, \tag{1}$$

where T_n is the 2-m temperature or dewpoint at the four model grid points surrounding the observation site and W_n is a weight given by

$$W_n = \frac{R^2 - D_n^2}{R^2 + D_n^2}. \tag{2}$$

In (2), R is the horizontal grid distance of the model (12 km) and D is the distance from the grid point to the observation site (Cressman 1959). The four nearest grid points to KENV, DPG17, and TRLU1 have elevations that are within 150 m of the stations' elevations (with no directional trend at any station). However, two of the nearest grid points for KSLC have elevations that are greater than 450 m higher than the observation point. Cheng and Steenburgh (2005) note that applying a correction to the temperature that accounts for differences in elevation can lead to erroneous results. Therefore, these two grid points are not used in the interpolation (the remaining two are within 150 m). In all discussion and figures, the error is defined as the forecast value minus the observed value.

The reader may question whether validation of wind speeds is of value. A comparison of model-forecast to observed wind speeds shows the average differences is quite small (usually ranging between 2 and 3 m s^{-1}) and so this is not assumed to be a cause of early VCP removal or related to other forecast problems.

b. The 2-m temperature and dewpoint temperature errors

The mean errors of forecast 2-m temperature at each station are shown in Fig. 2. At KENV and DPG17, forecasts valid at 1200 UTC have a distinct warm bias exceeding 2 K while those valid at 0000 UTC have a bias close to 0 K (Figs. 2a,b). The 95% confidence interval (Efron and Tibshirani 1993, chapter 4), calculated using the bootstrap method of Davison et al. (1986), indicates the warm biases at 1200 UTC are statistically significant at the 95% level with errors typically ranging from 1.5 to 4 K. Forecasts valid at 0000 UTC have biases that are generally not statistically different from 0 K. There is also a 1200 UTC warm bias at KSLC that exceeds 1 K (Fig. 2c). Here, though, forecasts valid at 0000 UTC have a well-defined cold bias. At both times, the biases are statistically different from 0 K. At TRLU1, all forecasts have a cold bias, with the 0000 UTC forecasts having a greater cold bias (with biases ranging from -2 to -4 K) than those at 1200 UTC (Fig. 2d). This cold bias is noteworthy: even if the basin temperatures were well forecast, a poor forecast along the basin rim could obscure the presence of a cold pool if one were solely relying on 2-m temperature patterns as a means of detection.

An identical comparison is performed for the 2-m dewpoint temperature, except at TRLU1, which does not have moisture observations. Both KENV and DPG17 have well-defined moist biases above 2 K for most forecast lead and validation times (Figs. 3a,b). At KSLC, forecasts valid at 0000 UTC tend to have small moist biases while those valid at 1200 UTC have comparatively large dry biases (Fig. 3c).

Examination of errors on a case-by-case basis reveals some important trends in the model forecasts. Consider the timeline of observed and forecast 2-m temperatures available every 3 h from 18 to 20 January 2009, which is during the midst of a particularly long-lived VCP that lasted from 15 to 23 January 2009 (Fig. 4). The nighttime warm bias at KENV and DPG17 is clearly depicted (Figs. 4a,c), as is the daytime cold bias at TRLU1 (Fig. 4g), consistent with the statistical analysis above. The forecast temperatures at 1200 UTC at KENV and DPG17 are between 4 and 8 K too high. These values are beyond the 95% confidence interval in Figs. 2a,b. Also note that the dewpoint temperature has a strong nighttime moist bias at KENV and DPG17 while the daytime dewpoint temperature is comparatively well forecast (Figs. 4b,d). Examination of Figs. 3a,b indicates that similar magnitudes of moist biases should occur at 0000 and 1200 UTC. A case-by-case analysis shows that the forecast patterns and large magnitudes of the errors (>5 K as in Figs. 4a–d) only emerge after a VCP has been in place for three or more days. Of the 38 persistent VCPs in this study, 26 are shorter than 3 days. Hence, the statistical analysis of Fig. 3 is weighted toward the short-lived events and the biases are not likely to be representative of very long-lived events.

The biases at KENV, DPG17, and TRLU1 exist regardless of how well the model's initial temperature agrees with the observations and have no obvious monotone growth or reduction with increasing lead time. As an example, consider the 0000 UTC 18 January 2009 forecast cycle at KENV. The initial temperature in the NAM model is only about 1.5 K higher than the observations, but by 1200 UTC 18 January, the forecast temperature is nearly 7 K warmer than the observations.

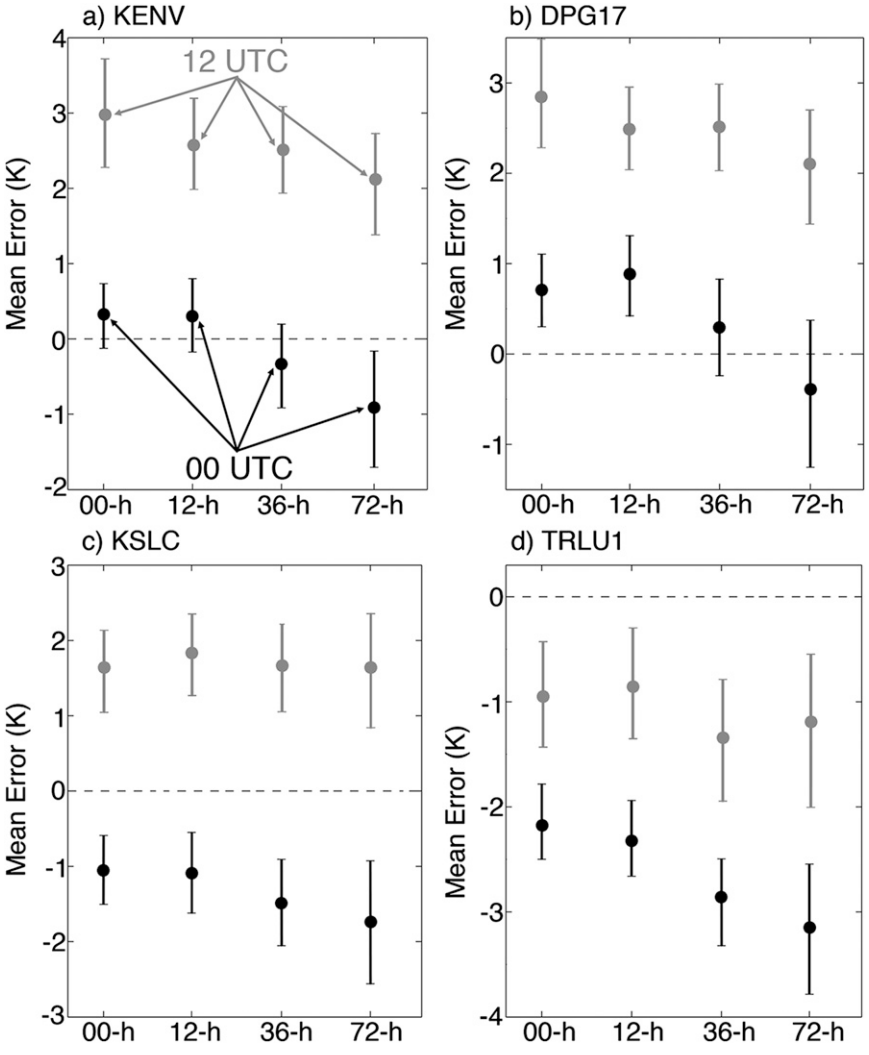

FIG. 2. The mean 2-m temperature errors (filled circles) and the 95% confidence interval (vertical bars) partitioned by the time of day the forecast is valid (0000 UTC, black; 1200 UTC, gray) and by the forecast lead time (0, 12, 36, or 72 h). Note that the vertical axes are different in each panel.

Similarly, the 1200 UTC 18 January forecast at KENV has an initial temperature that is about 6 K warmer than the observations, but by 0000 UTC 19 January, its temperature is only about 1 K warmer than the observations. Similar patterns of behavior occur at DPG17 and TRLU1. In fact, the forecasts at KENV, DPG17, and TRLU1 generally have better agreement with one another than with the observations. This is consistent with the statistical analysis in Fig. 2 in which the 95% confidence intervals and biases do not increase with increasing lead time. Rather, the uncertainty appears to be static regardless of the lead time. Such a finding indicates that the flow phenomenon is predictable (in that the errors do not grow monotonically with time) and that some systematic bias (or biases) may exist in the model that could possibly be corrected by tuning appropriate empirical relations used in the parameterization schemes.

At KSLC, the picture is quite different and the curves do not show as pronounced a daytime cold and nighttime warm bias as in Fig. 2c. Forecasts initialized at 0000 UTC have a pronounced nighttime warm bias while those initialized at other times have a slight daytime cold bias (Fig. 4e). Other long-lived VCP events have a more obvious nighttime warm bias and daytime cold bias for all forecast cycles (not shown). The 2-m dewpoint temperature at KSLC has a large nighttime dry bias, similar to the statistical analysis (cf. Figs. 4f, 3c). It is unclear why the patterns at KSLC are different from those at KENV and DPG17. It is possible that the Salt Lake Valley behaves as somewhat of a closed system (meaning

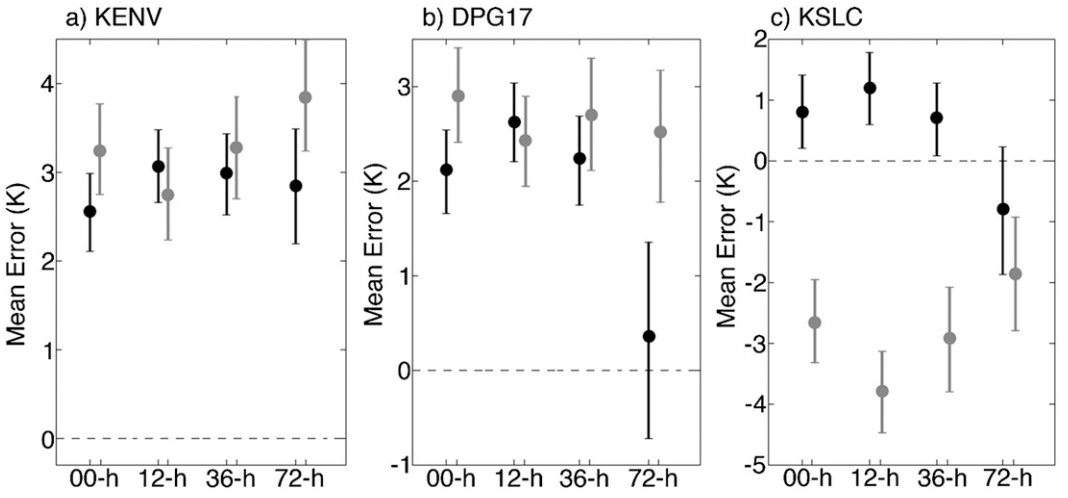

FIG. 3. As in Fig. 2, but for dewpoint temperature.

that air is not freely exchanged between the Salt Lake Valley and larger part of the basin) and, therefore, is not representative of the temperature and flow at other locations in the basin. However, a more likely explanation appears to be the inaccurate positioning of KSLC on the east slope of the basin in the NAM (Fig. 1b). The daytime cold bias (Fig. 2c) is consistent with daytime upslope ascent and its associated adiabatic cooling. The nighttime warm and dry biases (Figs. 2c, 3c) are consistent with nighttime downslope flow and its associated adiabatic heating and drying. Decreasing the horizontal grid spacing so that the Salt Lake Valley is better resolved may improve forecasts at this location.

The 2-m temperature and dewpoint biases have been calculated for eight other stations in and around the Bonneville Basin that are not discussed herein. All stations near the center or west side of the basin have a nighttime warm bias, similar to that noted for KENV and DPG17, stations along the east sidewall of the basin that are, in reality, located on the basin floor, but in the model are located on the slope, have a daytime cold and nighttime warm bias similar to that at KSLC, and all stations considered at higher elevations, outside of the basin and in snow-covered areas, have a daytime cold bias, similar to that at TRLU1. Thus, the patterns of model behavior at the four selected stations are representative of the model behavior across the basin.

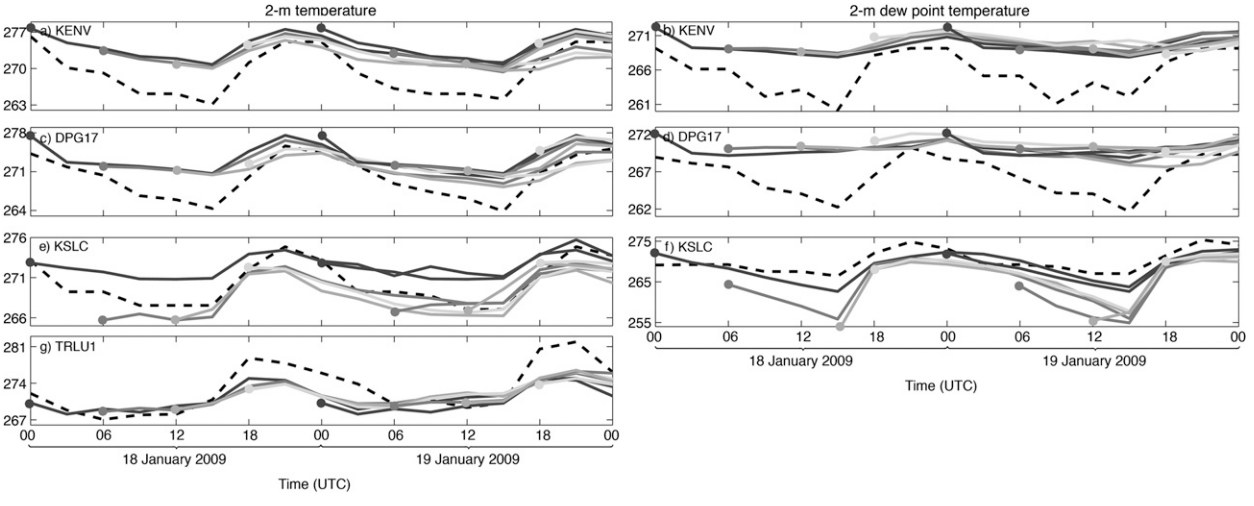

FIG. 4. A timeline of the observed (black dashed curves) and forecast (gray curves) (left) 2-m and (right) dewpoint temperatures for all forecast cycles started between 0000 UTC 18 and 20 Jan 2009. Different cycles are given in differing shades of gray and the start times are indicated by the large filled circles.

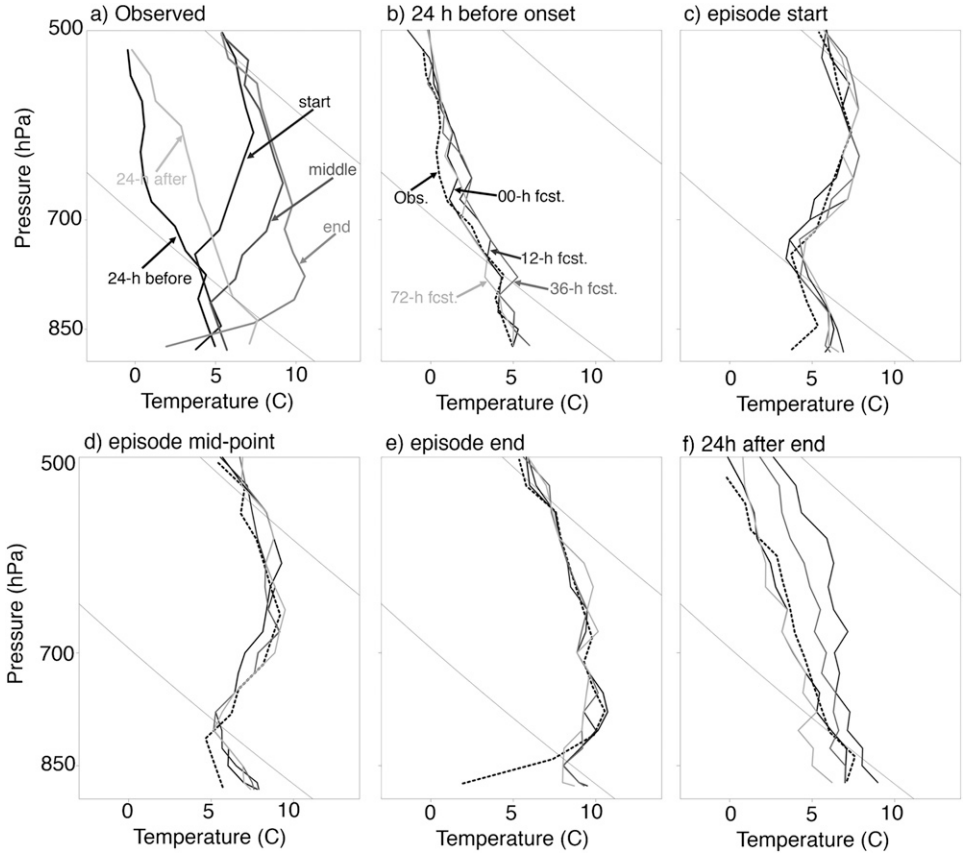

FIG. 5. Composites of (a) observed and (b)–(f) observed and forecast soundings at different times in the VCP life cycle. The line styles and colors for (c)–(f) are the same as in (b) and the thin diagonal lines are lines of constant potential temperature.

c. *Temperature and dewpoint forecasts above the ground*

Persistent VCPs have a characteristic synoptic-scale evolution that manifests itself in the temperature profiles (Reeves and Stensrud 2009). This is demonstrated using composites of observed soundings from KSLC taken at various times in the VCPs' life cycles. These are 24 h before the start of a VCP, at the start of the VCP, at the VCP midpoint, at the end of the VCP, and 24 h after the end of the VCP. Before VCP onset, the temperature profiles are near moist neutral (Fig. 5a). A gradual warming and descent of the inversion occurs between the start and end of the events, so that by the end of the events, the inversions are usually resting on the ground and are quite strong. The events typically end with the passage of an upper-level short-wave trough or surface cold front (Reeves and Stensrud 2009) and the temperature profiles return to a near moist-neutral state.

Temperature profiles from the NAM forecasts are partitioned and composited as above and compared to the observations. Before the onset, the forecasts have a slight warm bias, but are near moist neutral, in agreement with the observations (Fig. 5b). At the starts, midpoints, and ends, there is a well-defined low-level warm bias whose magnitude increases through the evolution of the events (Figs. 5c–e). At the starts and midpoints, the low-level warm bias has very little effect on the forecast strength of the capping inversions near 700 hPa, but at the ends, the inversions are typically much weaker than in the observations, primarily because the near-surface temperatures are typically much higher than in the observations and the observed inversions are much closer to the ground. Above the inversions, there is no clear temperature bias for any forecast lead time, indicating the model reasonably forecasts the midtropospheric temperature most of the time. After the events, there is a midlevel warm bias for the 0- and 12-h forecasts, but the forecast cold pools do not usually outlive those in the real world (Fig. 5f).

Observed sounding data from KSLC are also compared to NAM forecasts at 500, 700, and 850 hPa at the

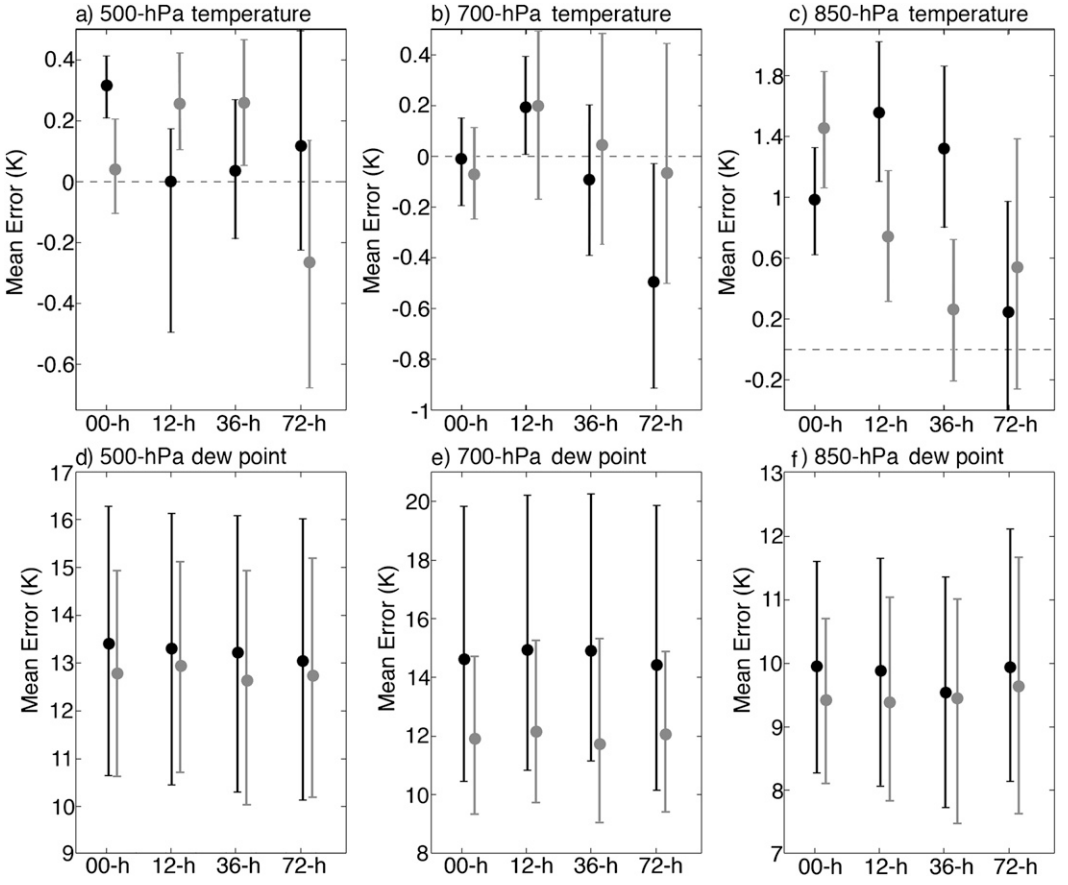

FIG. 6. The mean (filled circles) of the (top) temperature and (bottom) dewpoint and the 95% confidence interval (vertical bars) at different levels at KSLC partitioned by the time of day the forecast is valid (0000 UTC, black; 1200 UTC, gray) and by the forecast lead time (0, 12, 36, or 72 h). Note that the vertical axes are different in each panel.

nearest neighboring grid point for all VCP times. Errors are partitioned by the forecast lead time and validation time, as in section 2b. Temperature forecasts at 500 and 700 hPa are, on average, quite good (Figs. 6a,b). Biases have absolute values less than 1 K and the 95% confidence intervals indicate that for most lead and validation times, the biases are not statistically different from 0 K. The 95% confidence interval increases as lead time is increased. However, note that even for forecasts with a 72-h lead time, the 95% confidence interval is smaller than 1 K. This is consistent with Fig. 5 and indicates the NAM model provides reliable forecasts of the midlevel temperature evolution. In contrast to the 500- and 700-hPa levels, the 850-hPa level is, for much of the VCP life cycles, below the inversion base (Figs. 5c–e). Longer lead times have slight warm biases (between 0.3 and 0.6 K), but are not statistically different from 0 K (Fig. 6c). Shorter lead times have a clear warm bias that is statistically different from 0 K. [The reader may note that Fig. 6c is not entirely consistent with Fig. 2c in that no clear day–night couplet of cold–warm biases exist. Rather the 850-hPa biases are more consistent with those at KENV and DPG17 (Figs. 2a,b). It appears that 850 hPa is sufficiently far enough above the ground in the NAM to not suffer strong effects from the underlying sloped terrain.]

The forecast and observed dewpoints are also compared. At all levels, there is an apparent strong moist bias, with means ranging from 9 to 15 K (Figs. 6d–f). This may be partly due to a known dry bias on the part of radiosonde instrumentation (Häberli 2006), but the mean errors are sufficiently large to point to errors with the model analyses and forecasts. Additionally, the 95% confidence intervals are quite large, especially at 700 and 500 hPa (ranging from 5 to 10 K). This underscores the difference in uncertainty for temperature versus moisture in numerical models, with the latter generally being associated with a higher degree of uncertainty. Given that the relative humidities usually are in the range from 30% to 50%, it is unlikely that these differences lead to major problems in other forecast quantities. Finally, note that at 700 hPa there is a clear diurnal trend, with

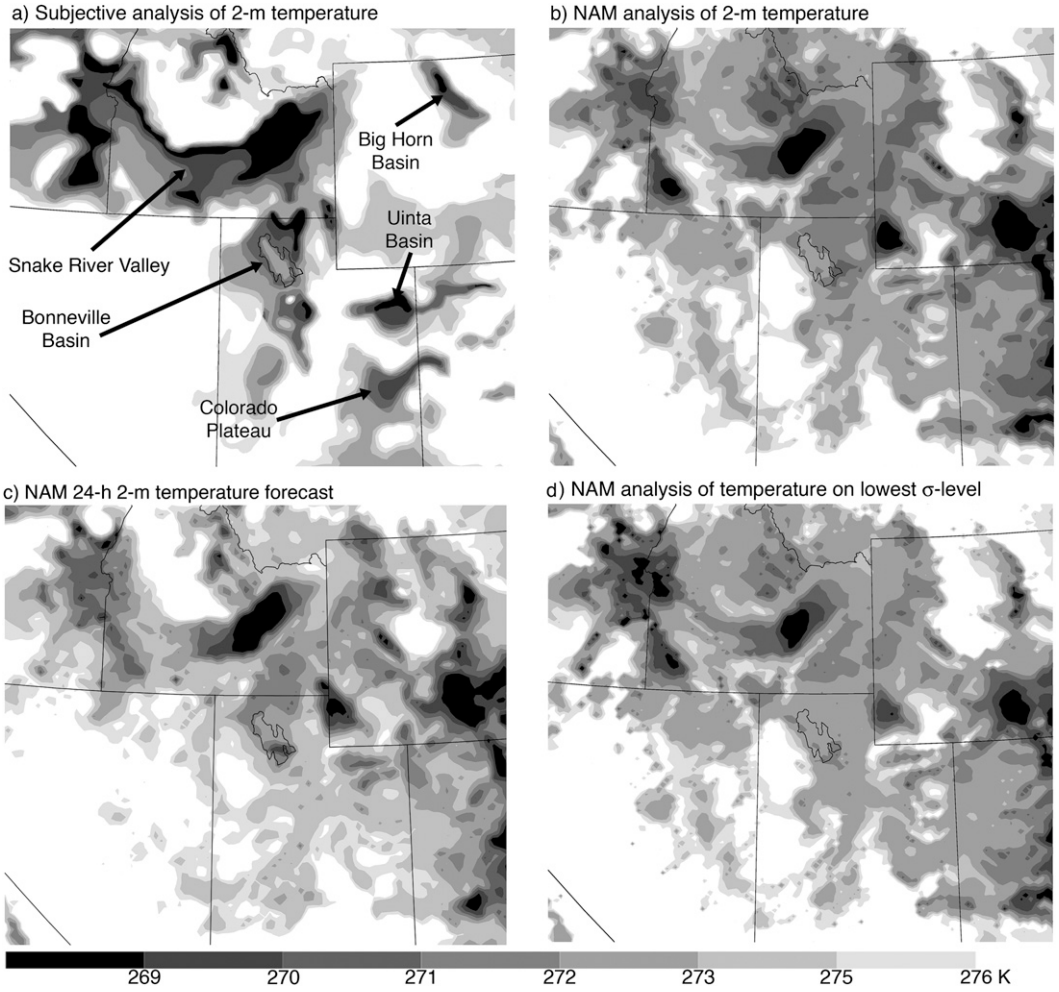

FIG. 7. (a) Subjective analysis of the observed 2-m temperature; the NAM (b) analysis and (c) 24-h forecast of 2-m temperature; and (d) temperature on the lowest model sigma level at 0000 UTC 19 Jan 2009.

the forecasts valid at 0000 UTC having a greater degree of uncertainty (Fig. 6e). The 95% confidence interval is also larger for the forecast valid at 0000 UTC at 500 hPa. There is not a corollary trend in temperature (Fig. 6b) and a cause for this enhanced uncertainty during the day is not known.

These comparisons suggest that the mid- and upper-level temperatures are usually well forecast. However, low-level temperature forecasts are problematic. Thus, any diagnosis of the causes of errors should focus on those parameterization schemes that directly affect the surface. We note that it is likely that the errors in temperature are related to the errors in dewpoint. Temperatures that are too high may lead to excessive snowmelt and evaporation, which can cause the dewpoint to be too high. Excessively high dewpoints, in turn, may lead to excessive downwelling longwave radiation and, hence, nighttime temperatures that are too high.

3. Diagnosis of model errors

a. Plausible sources of errors

The first suspicion one may have is that the errors in 2-m temperature and dewpoint are due to poor assumptions used in the interpolation from sigma coordinates to a 2-m height. Simple comparisons show this is not the case. Consider the subjective analysis of observed 2-m temperature at 0000 UTC 19 January 2009[1] (Fig. 7a). At this time, there are cold pools in many of the basins and valleys of the western United States. Such a pattern

[1] The subjective analysis is performed using all available observational data from the Mesowest network (information online at http://mesowest.utah.edu) excepting observations from the Union Pacific railroad and observations from valleys that are too narrow to be resolved by the NAM.

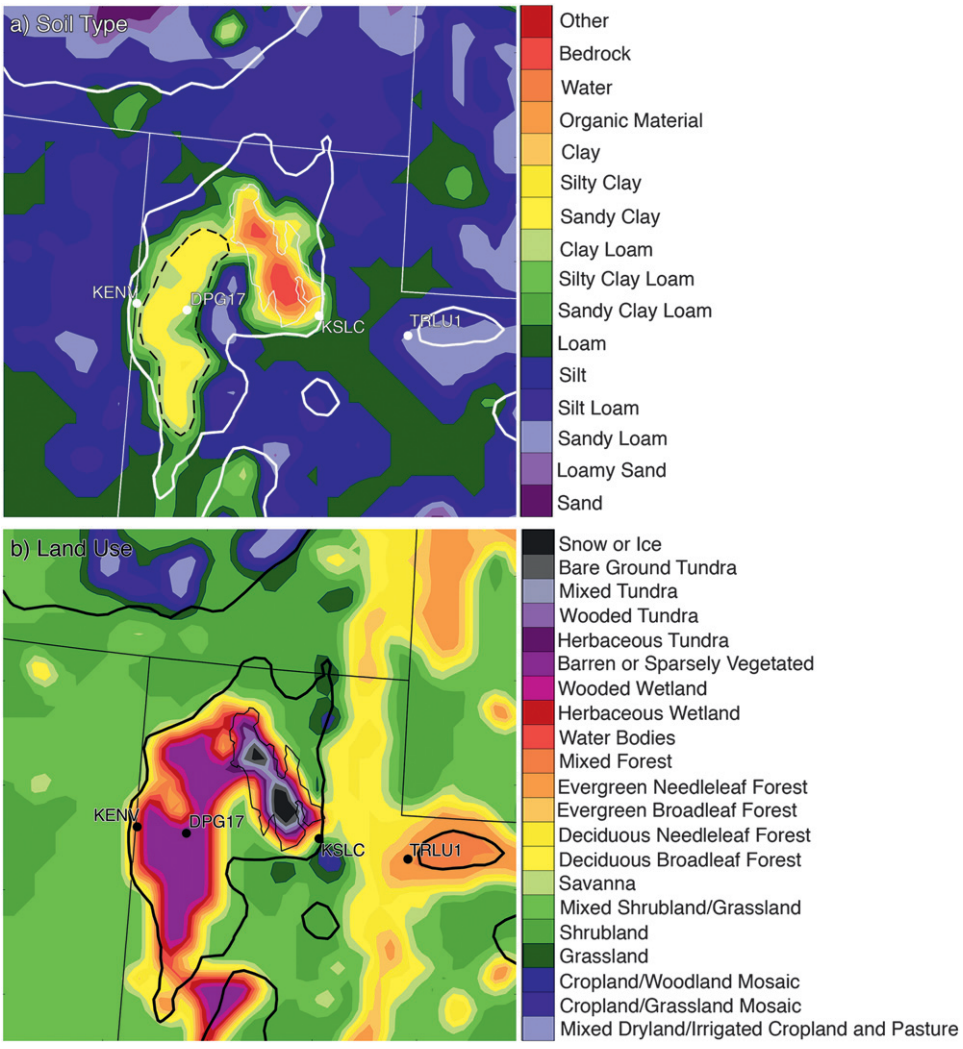

FIG. 8. The (a) soil and (b) land-use classification in the NAM. The white contour in (a) and the black contour in (b) indicate the 1550- and 3100-m terrain height, respectively. The dashed line in (a) indicates the region where the land-use and soil properties are modified in experiment LUSE.

is typical during persistent VCPs (Reeves and Stensrud 2009). While the corresponding NAM 0- and 24-h forecasts of 2-m temperature also show an accumulation of cold air in the northeast part of the Snake River Valley, there are no cold pools in the Bonneville, Big Horn, and Uinta Basins or over the Colorado Plateau (according to the 2-m temperature and using model sounding data) and, again, the forecasts show better agreement with each other than with the observations (Figs. 7b,c, respectively). A comparison of the NAM-analyzed 2-m temperature to the analyzed temperature at the lowest model level (Fig. 7d) shows the patterns are nearly identical. Similar patterns of agreement exist at other times and for the other cases (not shown), indicating the temperature biases are not the result of poor assumptions made when interpolating to 2 m.

The temperature and dewpoint errors could be due to incorrect specifications of the soil type and land use. The NAM classifies the soil in the western part of the basin as sandy clay and clay loam (Fig. 8a) and the land use as being barren or sparsely vegetated (Fig. 8b). However, much of the western part of the basin is a salt flat or playa, a category not used in the NAM. Changes in the land use and/or soil type are known to affect local forecasts (Rife et al. 2002; Zehnder 2002). This notion is expanded upon in section 3b.

Errors may also be related to the treatment of freezing water in soil by the LSM. The Noah LSM assumes all soil water freezes at a temperature of 273 K. Yet, salt reduces the freezing point of water by up to 20 K. An overestimate of the freezing point leads to an overestimate of the soil heat capacity and, hence, a more sluggish response of

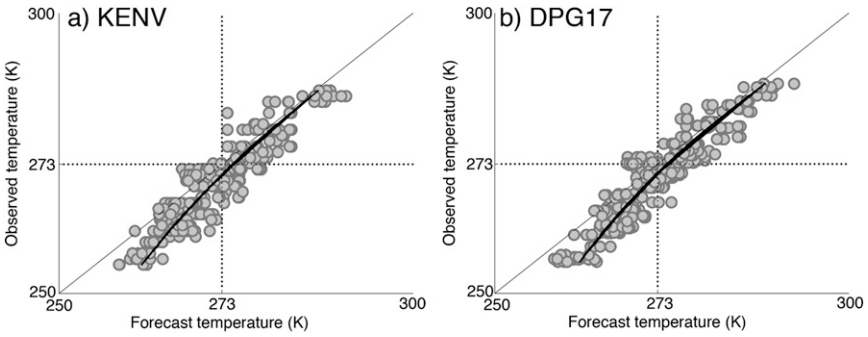

FIG. 9. The observed vs the forecast 2-m temperatures at KENV and DPG17. The diagonal line represents a perfect forecast and the thick curve is a 2nd-order polynomial fitted using the least squares method.

low-level temperatures to changes in the surface energy. Indeed, there is a correlation between the magnitude of the temperature errors and the forecast 2-m temperatures at KENV and DPG17. Scatterplots of observed and forecast 2-m temperatures for all forecast lead and validation times at these sites show that for forecast temperatures greater than 273 K, the forecast bias approaches 0 K (Fig. 9). However, as the forecast temperature is decreased below 273 K, a warm bias is found and becomes progressively larger as the temperature decreases. (The above was also repeated using a local least squares linear fit and permutation tests to confirm that errors are dependent on the forecast 2-m temperature, and both tests confirm the above, although permutation tests are somewhat inconclusive for 0-h forecasts at KENV.) Such a relation implies that midwinter VCP events, which are more likely to have subfreezing temperatures, are also more likely to have larger errors. A case-by-case consideration of all events reveals this hypothesis to be true (not shown).

Model treatment of snow effects on temperature could also explain some of the errors. Vertical cross sections of equivalent potential temperature (θ_e) through the basin show that at 0000 UTC 19 January there is a strong inversion over the east rim and sidewall (Fig. 10a). Close inspection of this inversion shows that the temperature decreases about 10 K in the lowest 50–100 m. Inversions of similar strength have been observed in the Salt Lake Valley at night (Whiteman and Zhong 2008), but during the day, such an inversion seems unlikely. In fact, the inversion only exists over snow-covered areas and the snow depth patterns at 0000 UTC 19 January (Fig. 10b) bear a remarkable similarity to regions of cold air in the NAM-analyzed 2-m temperature (Fig. 7b).

All of the above potential sources of error can be diagnosed using numerical sensitivity experiments.

b. Sensitivity experiment description

To diagnose the source of low-level temperature and humidity errors, a series of sensitivity experiments is performed. All experiments are integrated from 0000 UTC 18 January to 0000 UTC 20 January 2009. The initial and boundary conditions are from the NAM 0000 UTC 18 January 2009 forecast cycle. Model data

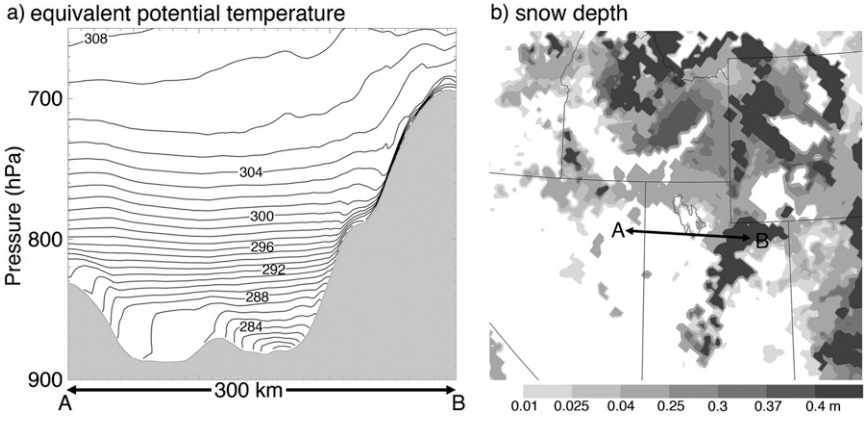

FIG. 10. (a) Vertical cross section of θ_e [cross section indicated in (b)] and (b) snow depth (m; shaded) from the NAM analysis at 0000 UTC 19 Jan 2009.

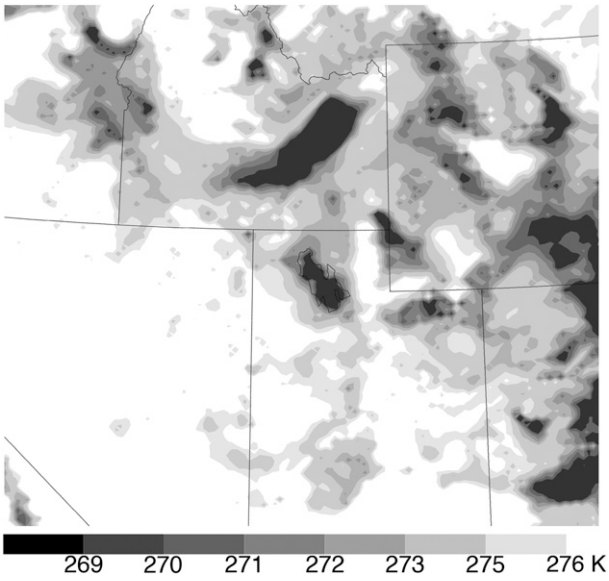

FIG. 11. The 24-h forecast 2-m temperature from the CNTL experiment.

are output every hour. The WRF-NMM is used and set up in exactly the same way as the NAM forecasts from this time period except for the domain, which is smaller than that used operationally (3300 km × 4512 km compared to roughly 16 044 km × 13 980 km), and the initial temperature of the Great Salt Lake is derived from climatological analyses while the NAM uses a climatology of point observations to initialize the lake temperature. This latter difference leads to lake temperatures that are between 3 and 4 K colder than those used operationally. The data were examined to ensure the following results and conclusions are not dependent on this difference.

The first experiment, referred to as CNTL, is in every way identical to the NAM forecast, except for the differences noted above. A plan view of the 24-h forecast 2-m temperature from CNTL (Fig. 11) looks very similar to the NAM forecast, exhibiting a well-defined cold pool in the Snake River Valley and a lack of cold pools in the Bonneville, Big Horn, and Uinta Basins as well as the Colorado Plateau (cf. Figs. 7c, 11). Time sequences of 2-m temperature and dewpoint at the four stations also show reasonable agreement with the operational NAM forecasts. At KENV and DPG17, there are nighttime warm and moist biases in CNTL that are similar to the NAM forecasts (Figs. 12a–d). At KSLC, there is a nighttime warm and dry bias in CNTL (Figs. 12e,f). However, the minimum temperature and dewpoint in CNTL are somewhat lower than those in the NAM forecasts. Finally, the daytime cold bias at TRLU1 is similar in both CNTL and in the NAM forecasts (Fig. 12g). Since the error characteristics in CNTL are similar to those in the NAM forecasts, we will use the CNTL experiment as a baseline against which all sensitivity experiments are compared.

Several potential causes of the forecast errors have already been identified (section 3a). First, as noted above, there is a marked cold bias over snow-covered regions (Fig. 10). In the NAM, snow effects are handled by the Noah LSM, which makes adjustments to the albedo,

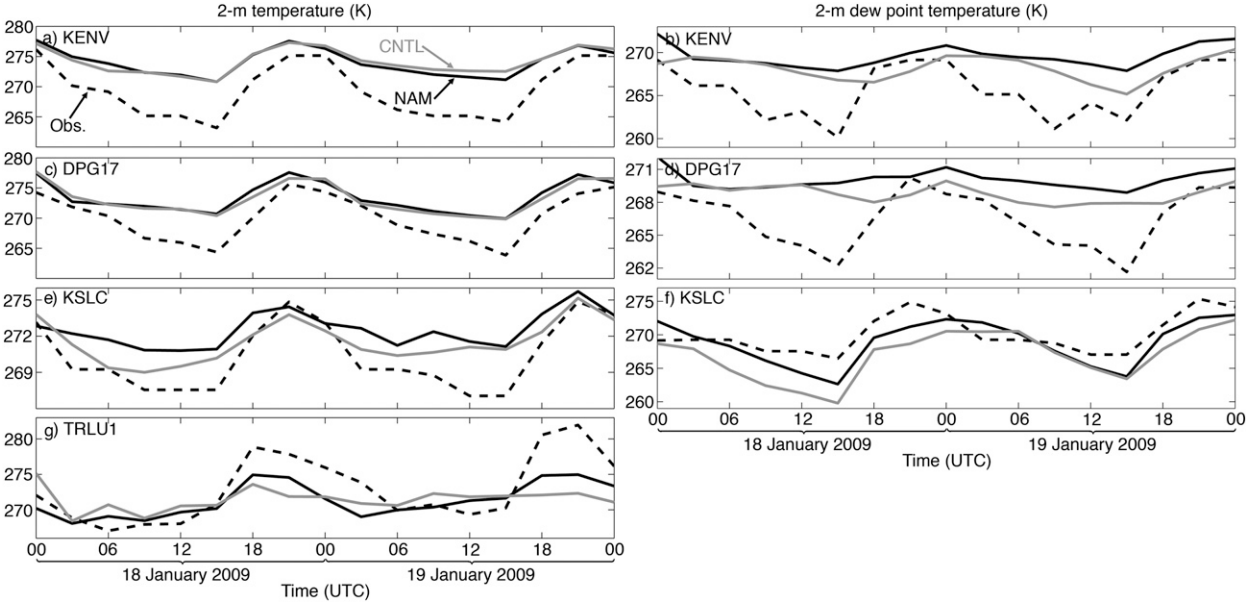

FIG. 12. A timeline of observed (dashed) and forecast (solid) (left) 2-m temperature and (right) 2-m dewpoint temperature from 0000 UTC 18 Jan to 0000 UTC 20 Jan 2009. The CNTL (NAM) forecasts are given by the solid gray (black) contours.

TABLE 1. The albedo (α, %), surface emissivity (ε, % at 9 μm), roughness length (Z_0, m), saturation soil conductivity (κ), saturation soil diffusivity (D), and the dry soil diffusivity–conductivity coefficient (ϱ).

Category	α	ε	Z_0	κ	D	ϱ
Sandy clay	40	0.98	0.01	1.34×10^{-6}	0.964×10^{-5}	−1.916
Playa	40	0.90	1.0	9.74×10^{-7}	0.112×10^{-4}	−10.472
Water	8	0.98	0.00	0.0	0.0	0.0
Salt–water mix	24	0.94	0.505	4.87×10^{-7}	0.056×10^{-4}	−5.236

emissivity, conductivity, and roughness length depending on the depth and water equivalent of the snow. To test whether snow effects as handled by the Noah LSM are responsible for any of the errors, experiment NSNO is performed in which snow effects are neglected in the LSM.

As noted above, the dominant soil type in the western part of the basin is sandy clay, according to the NAM. However, much of the western basin is actually a playa. A playa category is available in both the land-use and soil parameter tables, and as Table 1 indicates, these values can be quite different from the sandy clay category. However, the playa variables in the soil tables for the NAM are consistent with dry playa measurements, to the extent that data for this category ar available (McCurdy 1989; Tapper 1991). The Bonneville Salt Flats have a high water table in winter such that the surface resembles a slurry of salt and water and so the dry playa parameters are not appropriate. The exact ratio of salt and water is not known and may not be constant across the playa. Moreover, reasonable estimates for the soil variables in Table 1 for a salt–water mix are not known. For the sake of testing the sensitivity to soil moisture and type, the area enclosed by the dashed line in Fig. 8a is modified so that the soil properties are halfway between those of a dry playa and a water body (Table 1) in experiment LUSE. We caution that exact measurements of surface emissivity, soil conductivity, diffusivity, etc. are not known for the Bonneville Basin in winter and we make no recommendation for the values used herein to be used in future work.

To test the sensitivity to the soil-water freezing point, the LSM was modified in experiment LFRZ so that soil water has a freezing point of 258 K in the area enclosed by the dashed line in Fig. 8a. To the best of our knowledge, the freezing point in the Bonneville Salt Flats has not been rigorously measured and the above value is chosen as a conservative estimate.

Given the lack of synoptic and mesoscale forcing to produce precipitation and/or convection, the only parameterizations that are likely to affect the low-level temperature and moisture are the radiation and planetary boundary layer schemes. To isolate whether the cause of errors rests with either of these, sensitivity experiments are conducted with the nonlocal closure YSU planetary boundary layer scheme (the YSU experiment) and with the Rapid Radiative Transfer Model radiation scheme (the RRTM experiment). Finally, to see if decreasing the grid spacing has an effect on forecasts, an experiment with a 4-km grid spacing that is otherwise identical to CNTL is performed. The terrain for this experiment, referred to as NAM4, is shown in Fig. 1c. To ensure a more direct comparison of forecast temperatures between NAM4 and CNTL, the inverse Cressman method is applied to the 16 closest grid points to the latitudes and longitudes of the stations. Table 2 gives a summary of all experiments performed.

c. *Results of the sensitivity experiments*

The hourly model output is averaged from 0100 UTC 18 January to 0000 UTC 20 January and the 95% confidence intervals are calculated. Results for the 2-m temperature forecasts show that the different experiments have different results at the different stations (Fig. 13). At KENV, the only experiment that is statistically different from CNTL is LFRZ, which has a reduced mean

TABLE 2. The numerical experiments performed to diagnose possible sources of model errors.

Expt	Expt description
CNTL	Control experiment, designed to be nearly identical to the operational forecasts
NSNO	Snow effects are neglected in the land surface model
LUSE	Soil type in the west basin (see dashed outline in Fig. 8a) is changed to a 50–50 mix of salt and water
LFRZ	Soil water freezing point is reduced by 15 K to 258 K
YSU	YSU planetary boundary layer scheme is used in place of MYJ
RRTM	RRTM radiation scheme is used in place of GFDL
NAM4	Grid spacing is reduced from 12 to 4 km

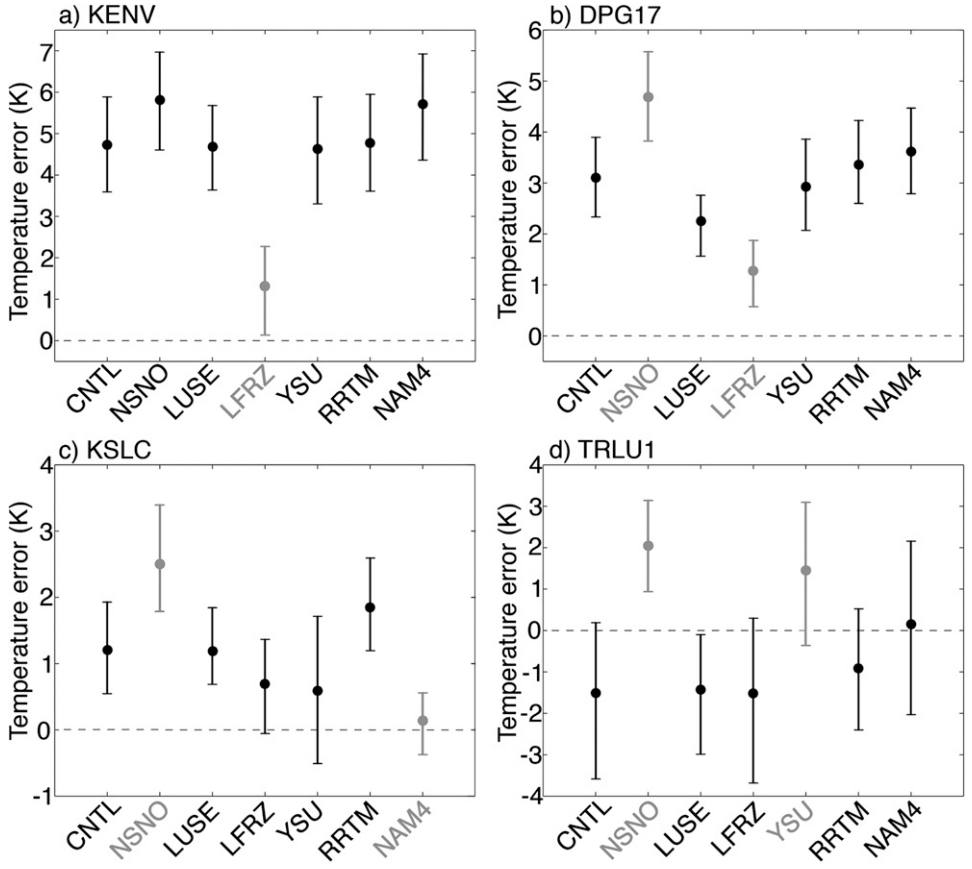

FIG. 13. The mean temperature (filled circles) and the 95% confidence interval (vertical bars) for the different sensitivity experiments. Note that the vertical axis is different in each panel; and, unlike previous similar figures, the data are not partitioned according to time of day.

error (Fig. 13a). The same is also true at DPG17 (Fig. 13b). The NSNO experiment has 2-m temperatures that are consistently higher than CNTL at DPG17, KSLC, and TRLU1 (Figs. 13b–d). At DPG17 and KSLC, this is a detrimental effect. At TRLU1, the mean absolute error in NSNO is higher than in CNTL, but the confidence interval is smaller, suggesting snow introduces additional uncertainty into the forecast. The NAM4 experiment is a notable improvement over CNTL at KSLC and has a mean error that is not statistically different from 0 K (Fig. 13c). Finally, like NSNO, YSU produces temperatures that are, on average, warmer than CNTL at TRLU1 (Fig. 13d).

Timelines of the forecast 2-m temperature from those experiments that have the greatest differences from CNTL are compared to CNTL and the observations from 18 to 20 January 2009. At KENV and DPG17, LFRZ offers a much improved level of agreement with the observations at night and better captures the amount of nocturnal cooling, particularly at DPG17 (Figs. 14a,c). At KSLC, NAM4 gives a very good forecast (Fig. 14e), and at TRLU1, YSU and NSNO give generally higher temperatures, but fail to capture the full diurnal range (Fig. 14f). A similar comparison for dewpoints at KENV and DPG17 shows LFRZ has lower dewpoint values than CNTL, which results in a better forecast at night, but the daytime maximum is underestimated by several degrees (Figs. 14b,d). At KSLC, the NAM4 also has better dewpoint forecasts at night, but underestimates the daytime maximum (Fig. 14f).

A comparison of observed and modeled soundings at KSLC shows that above about 800 hPa the various experiments compare very well to the observations (Fig. 15). Potentially important differences are only manifest at low levels and only in the NAM4 experiment, which better captures the near-surface temperatures, particularly during the day (Figs. 15b,d). Although there is no observed sounding at DPG17, model sounding comparisons at this site are elucidating. During the day, only NSNO is significantly different from CNTL (Fig. 15e). It has a shallower, surface-based cold layer and warmer low-level temperatures. At night, NSNO is also somewhat warmer than CNTL (Fig. 15f). However, note that

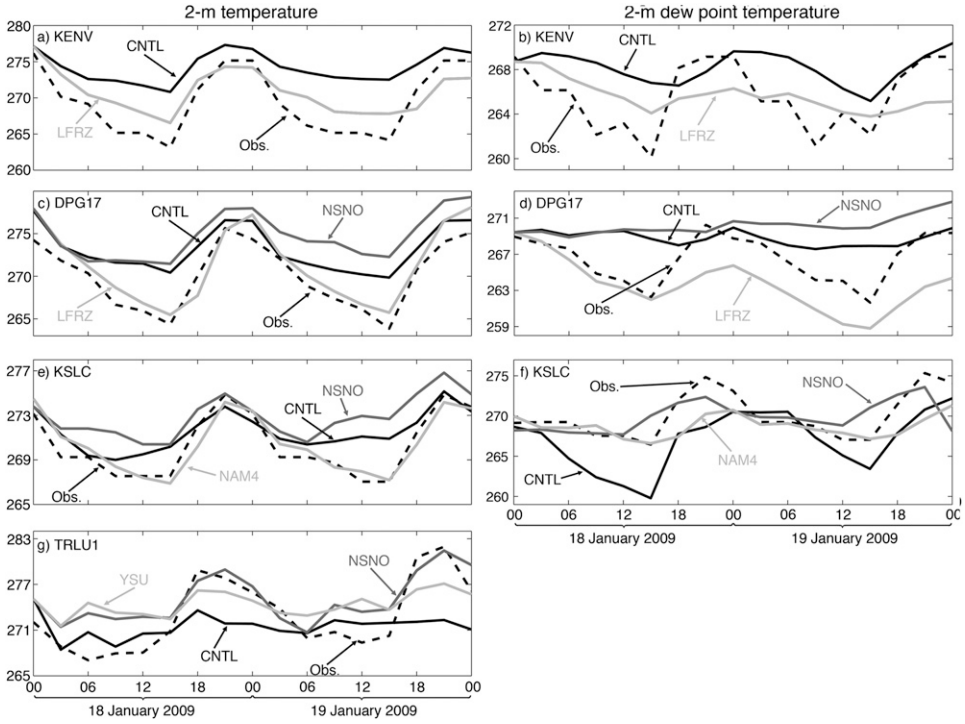

FIG. 14. A timeline of observed (black dashed curve) and forecast (solid) (left) 2-m temperature and (right) 2-m dewpoint temperature from 0000 UTC 18 Jan to 0000 UTC 20 Jan 2009. The experiments that go with each curve are indicated in the panels.

LFRZ has colder near-surface temperatures. This is consistent with Figs. 13b and 14c.

The most easily explained results are those from YSU. The YSU scheme is associated with stronger boundary layer mixing (Pagowski 2004; Hu et al. 2010), which results in a more vigorous mixing of warmer air aloft toward the surface and generally warmer near-surface temperatures, consistent with results shown for TRLU1. There is not a similar effect at the three basin stations, however. At these locations, YSU is not statistically different from CNTL (Fig. 13), a result that appears to be linked to the very strong surface-based inversion and rather shallow boundary layers in both YSU and CNTL (not shown).

Reducing the freezing point for soil water, as in LFRZ, affects the low-level temperature in multiple ways via latent heat effects. As compared to CNTL, LFRZ has a lower latent heat flux at the surface (Figs. 16a,b) and a considerably lower soil heat flux (Figs. 16c,d) in the western part of the basin. Consequently, the temperatures over the modified soil are several degrees cooler in LFRZ (Figs. 16e,f). The soil heat flux is strongly controlled by the latent heat of fusion from freezing soil water. The effects of this were tested in an experiment where the latent heat of fusion was set to zero everywhere in the model domain. The results from this experiment agree closely with LFRZ (not shown).

The NAM4 experiment merits a deeper investigation. The improvement in NAM4 is not due to differences in elevation at KSLC, as only grid points that have elevations within 150 m of the station elevation are used in the inverse-distance Cressman method. Rather, differences appear to be due to the differing wave patterns that form near the slope at night. Vertical cross sections of potential temperature and vertical velocity at 1200 UTC 18 January 2009 and along line A′B′ (from Fig. 1c) show that while both the CNTL and NAM4 have downslope winds on the basin sidewall, the downward-directed flow penetrates deeper downward and is closer to KSLC in CNTL (cf. Fig. 17a,b). In fact, the flow at KSLC in NAM4 is directed upward at this time due to the shorter wavelength. Although there is not ascending flow in this region throughout the night, NAM4 consistently exhibits a stronger wave pattern near the east sidewall that advects cooler temperatures from the middle of the basin toward KSLC at low levels. The stronger wave in the NAM4 experiment may be related to the more abrupt changes in elevation in this experiment (Smith 1979). These results are contradictory to those of Zhong and Whiteman (2008). They found that steeper slopes are associated with weaker downslope winds due to the reduced accumulation of cold air along the steeper slope. The change in slope angle between NAM4 and

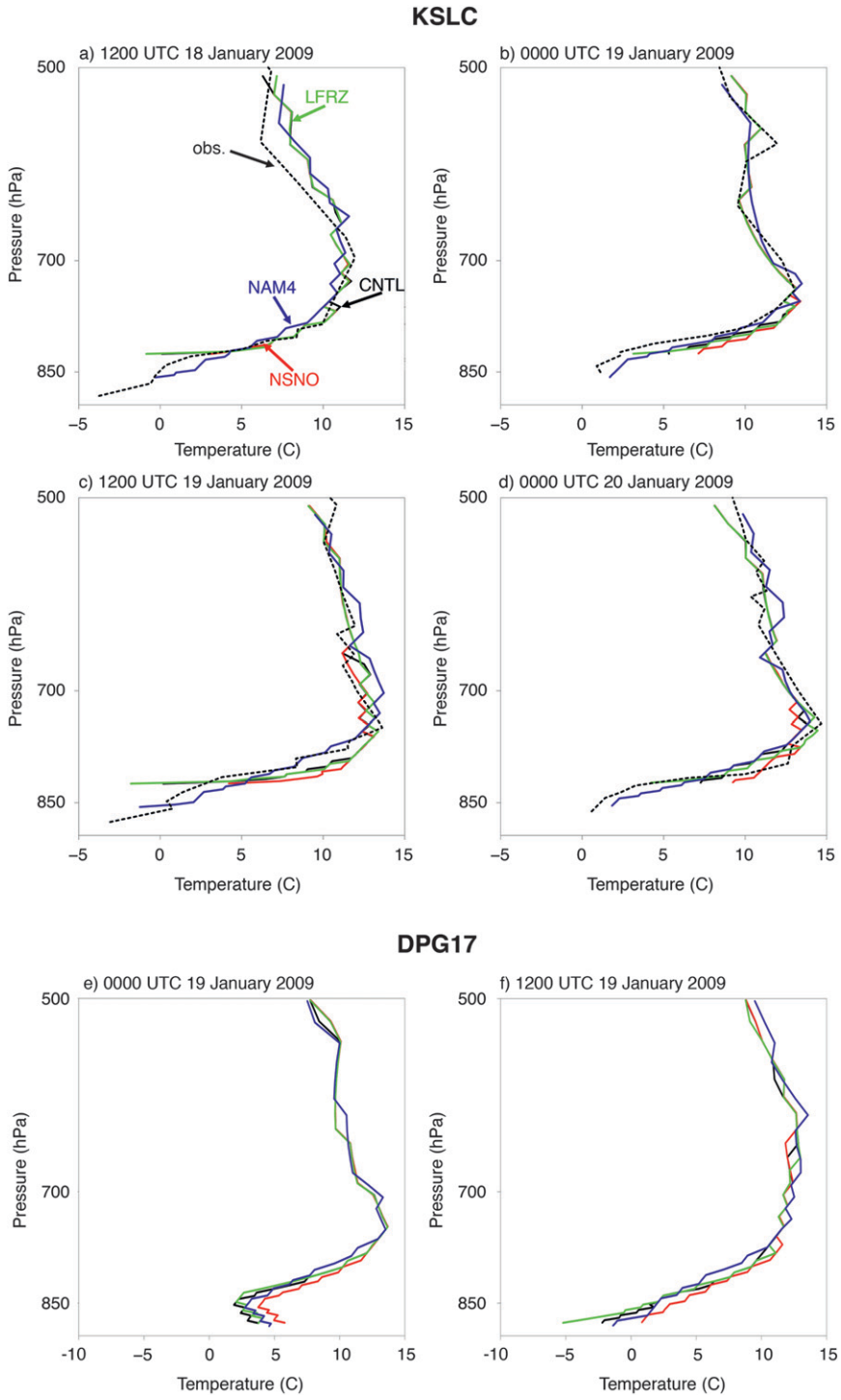

FIG. 15. Skew T–log$p$ profiles of observed (dashed) and model forecast [solid contours, with colors corresponding to the different experiments indicated in (a)] temperatures at KSLC (top four panels) and DPG17 (bottom two panels).

CNTL is rather modest compared to that in Zhong and Whiteman (2008), which may account for the different results, but the cold-air accumulation in NAM4 and CNTL may also be modulated by the presence of snow and the fact that the air at higher elevations is potentially colder than that in the basin.

The results from NSNO indicate that DPG17 and KSLC experience a statistically significant change in the

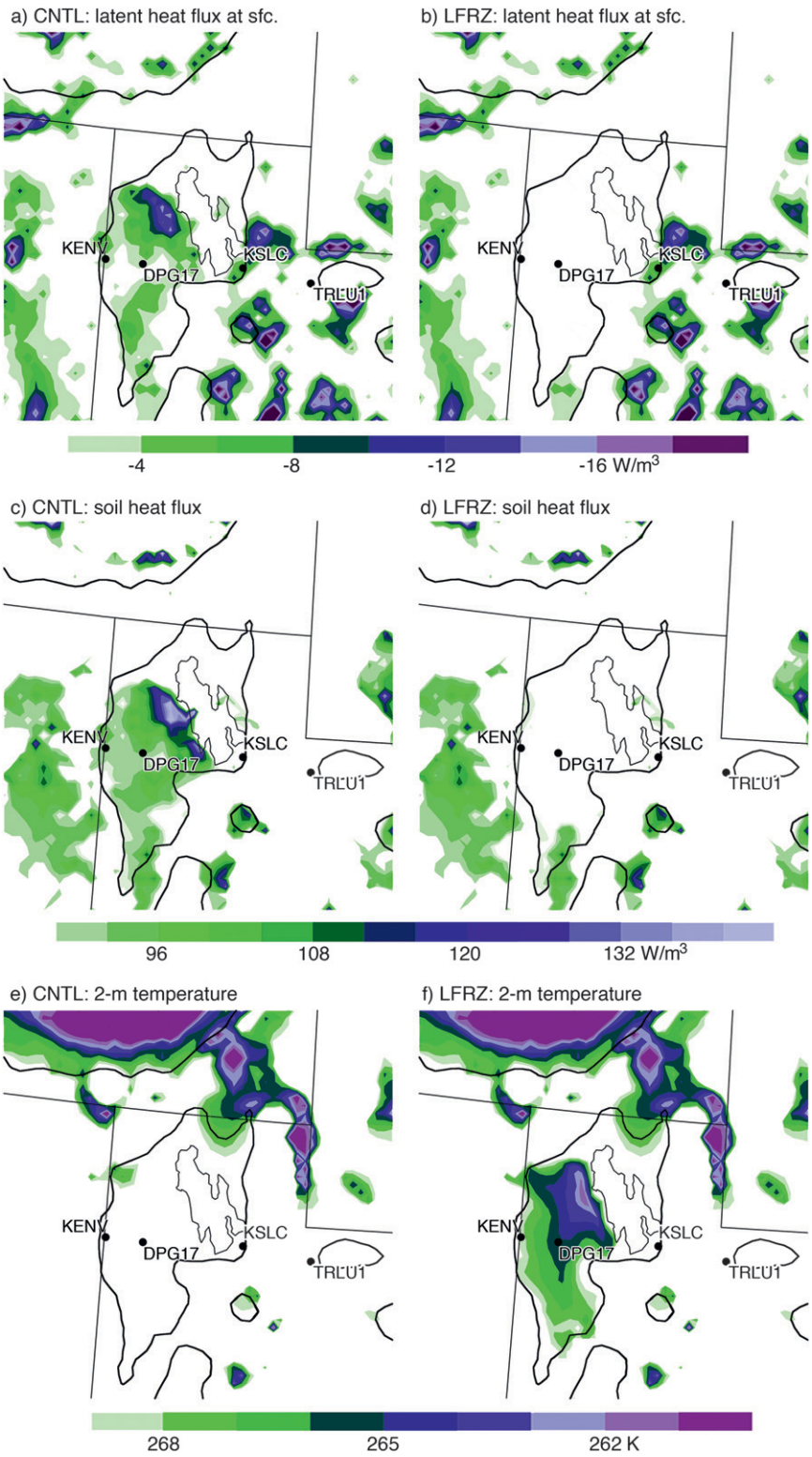

FIG. 16. (a),(b) Surface latent heat flux, (c),(d) soil sensible heat flux, and (e),(f) 2-m temperature at 1200 UTC 18 Jan 2009 for the (left) CNTL and (right) LFRZ experiments.

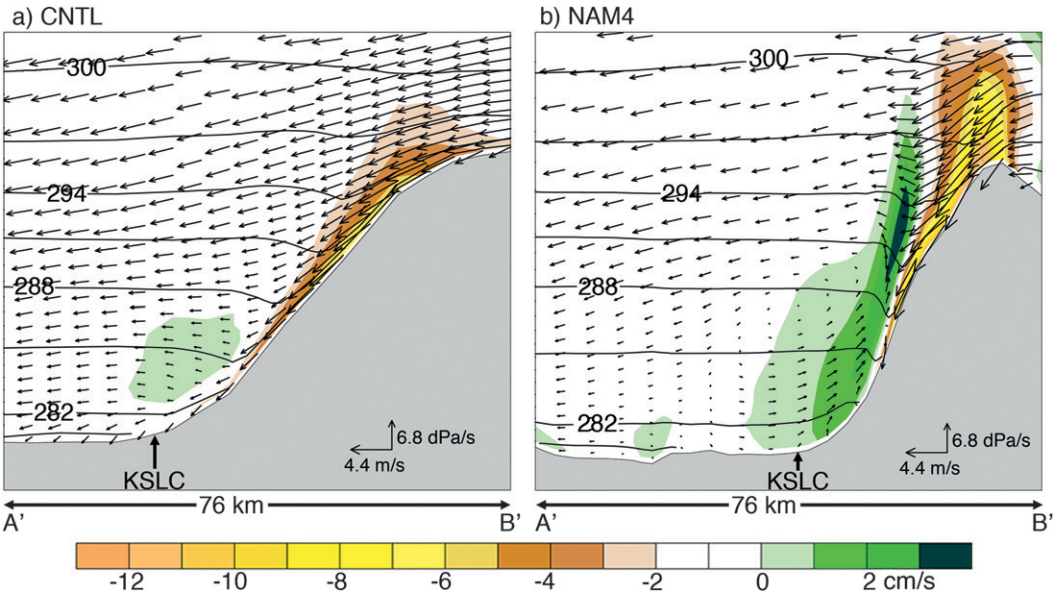

FIG. 17. Vertical cross sections along line A′B′ (from Fig. 1c) of potential temperature (K; contoured) vertical velocity (cm s^{-1}; shaded), and wind vectors perpendicular to the plane at 1200 UTC 18 Jan 2009 for (a) CNTL and (b) NAM4. The center location for the grid points used in determining the 2-m temperature is indicated as KSLC.

temperature forecast when the snow effects are neglected in the LSM (Fig. 13), even though there is no snow at these stations. The potential temperature and winds from the lowest model level in CNTL at 1200 UTC 19 January 2009 (when the differences between CNTL and NSNO are the most pronounced at KSLC and DPG17) reveal that potentially colder air over the higher terrain drains into the basin from the north and spreads across the basin floor (Fig. 18a). In NSNO, this drainage flow pattern is still present, but the air is potentially warmer than the air at the bottom of the basin (Fig. 18b). Overall, the snow effects in CNTL have some benefit. The drainage of cold air into the basin helps to reduce the nighttime temperature at DPG17 and KSLC, but observational data are too sparse to confirm whether such an effect occurs in the real atmosphere. Several additional experiments were conducted wherein the LSM was modified so that various snow effects are neglected, including removal of the dependency on snow of albedo, surface emissivity, and latent cooling from melting. None of these experiments yield results that are statistically different from CNTL (not shown). Snow effects were also neglected in the planetary boundary layer scheme in one experiment, but this did not give statistically different results from CNTL, consistent with the findings of Billings et al. (2006).

4. Conclusions

The North American Mesoscale Model forecasts of low-level temperature and dewpoint during persistent valley cold pools in the Bonneville Basin of Utah are assessed over three winter seasons from 2007 to 2009. Bias calculations show that the forecast 2-m temperatures in the western part of the basin have a pronounced warm bias at night but are usually good during the day. Outside of the basin, at stations where there is snow cover, there is a cold bias that is strongest during the day. This combination of cold and warm biases interferes with one's ability to use the 2-m temperature as a means of identifying VCPs. Bias calculations for 2-m dewpoint show that basin stations have a well-defined moist bias that is likely related to biases in temperature. Consideration of different forecast lead times shows there is no significant difference in forecast quality as the lead time is increased, suggesting that forecast problems are not due to rapid error growth but, rather, are due to some inherent fault or assumption in the model. A case-by-case evaluation reveals that the longer the VCP is in residence, the greater the nocturnal warm bias in the basin. In addition, the biases are dependent on the forecast temperature. When the forecast temperature is below 273 K, the magnitude of biases increases substantially. Thus, midseason events, which tend to have longer durations (Reeves and Stensrud 2009) and lower low-level temperatures, are likely to have much larger biases than are indicated by the statistical analysis.

Above the surface, results indicate that the 700- and 500-hPa temperature errors are negligible, suggesting that the model generally does a good job of handling the large-scale flow patterns associated with cold pools. In

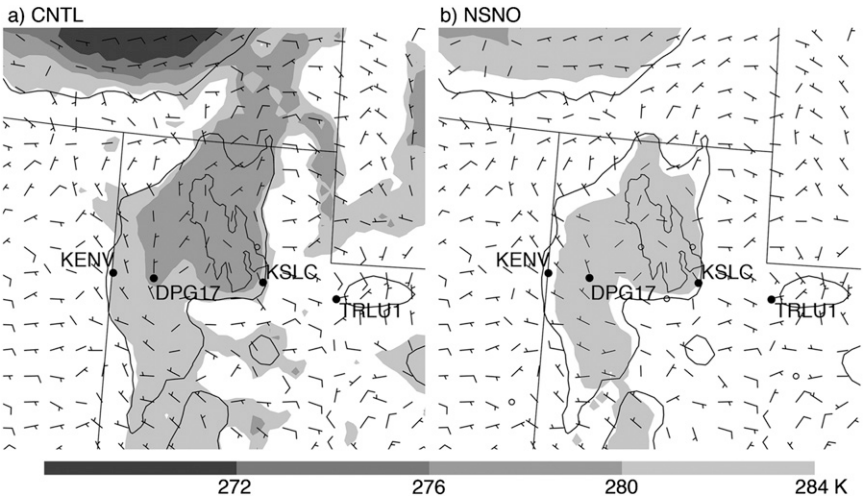

FIG. 18. Potential temperature and wind barbs (1 full barb is 5 m s^{-1}) on the lowest model layer for the (left) CNTL and (right) NSNO experiments.

contrast, the 850-hPa temperature errors show a statistically significant warm bias for some lead times. A comparison of observed and forecast soundings also shows that the primary forecast problems are at low levels, where the temperature forecast is too high.

A series of numerical experiments are used to explore potential causes of the observed errors. These include a control experiment that is nearly identical to the operational NAM configuration (CNTL), an experiment with a different planetary boundary layer scheme (YSU), an experiment where the radiation parameterization scheme is changed (RRTM), an experiment where the effects of snow are neglected in the land surface model (NSNO), an experiment where the land use in the western part of the basin is changed to better reflect the salt–water mix typical of that region (LUSE), an experiment where the soil-water freezing point is reduced from 273 to 258 K (LFRZ), and an experiment where the grid spacing is reduced from 12 to 4 km (NAM4).

The different experiments have different results that are dependent on where the stations are located and the types of biases noted. KENV and DPG17, which are adjacent to the salt flats, show marked improvement in LFRZ due to effects associated with the latent heat of fusion. The NAM4 experiment showed the greatest improvement at KSLC. This appears to be due to the steeper slope and terrain-induced wave along the east sidewall. At TRLU1, the cold bias is found to be due to the treatment of snow by the land surface model. Using the YSU planetary boundary layer scheme, which has stronger mixing than in the MYJ scheme, helps to alleviate the daytime cold bias, but leads to excessively warm temperatures at night. The cause of the cold bias over snow within the land surface model remains unknown. Observational work measuring snow effects on low-level temperature in the vicinity of the Bonneville Basin may prove useful for identifying why the Noah LSM produces a cold bias in this area.

At KENV and DPG17, the LFRZ experiment yielded a mean temperature error that is between 3.5 and 5 K closer to the observations than in CNTL. At KSLC, NAM4 shows a mean improvement of about 2.5 K. Yet, despite the fact that the LFRZ and NAM4 experiments have improved the 2-m temperature forecasts, they did not have greatly improved 2-m dewpoint forecasts. No permutation of the model significantly improves dewpoint forecasts beyond that in LFRZ and NAM4. Further work in this area is recommended. While LUSE did not yield statistically different forecasts from CNTL, others have found that a more accurate representation of land-use and soil type yields improved forecasts (e.g., Rife et al. 2002; Zehnder 2002). Perhaps tests over a longer forecast time than the 48 h considered herein would reveal a greater dependence.

Acknowledgments. This study was made possible in part due to the data made available by the governmental agencies, commercial firms, and educational institutions participating in MesoWest. Partial funding for H. Reeves was provided by the National Research Council. Comments from S. Zhong and two anonymous reviewers were helpful for improving the manuscript. Funding was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA17RJ1227, U.S. Department of Commerce.

REFERENCES

Bader, D. C., and T. B. McKee, 1985: Effect of shear, stability and valley characteristics on the destruction of temperature inversions. *J. Climate Appl. Meteor.,* **24,** 822–832.

Banta, R., and W. R. Cotton, 1981: An analysis of the structure of local wind systems in a broad mountain basin. *J. Appl. Meteor.,* **20,** 1255–1266.

Berg, L. K., and S. Zhong, 2005: Sensitivity of MM5-simulated boundary layer characteristics to turbulence parameterization. *J. Appl. Meteor.,* **44,** 1467–1483.

Betts, A. K., and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and Arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.,* **112,** 693–709.

Billings, B. J., V. Grubisic, and R. Borys, 2006: Maintenance of a mountain valley cold pool: A numerical study. *Mon. Wea. Rev.,* **134,** 2266–2278.

Chen, Y., F. L. Ludwig, and R. L. Street, 2004: Stably stratified flows near a notched transverse ridge across the Salt Lake Valley. *J. Appl. Meteor.,* **43,** 1308–1328.

Cheng, W. Y. Y., and W. J. Steenburgh, 2005: Evaluation of surface sensible weather forecasts by the WRF and the Eta Models over the western United States. *Wea. Forecasting,* **20,** 812–821.

——, and ——, 2007: Strengths and weaknesses of MOS, running-mean bias removal, and Kalman filter techniques for improving model forecasts over the western United States. *Wea. Forecasting,* **22,** 1304–1318.

Cressman, G., 1959: An operational objective analysis system. *Mon. Wea. Rev.,* **87,** 367–374.

Davison, A. C., D. V. Hinkley, and E. Schechtman, 1986: Efficient bootstrap simulation. *Biometrika,* **73,** 555–566.

Dirmeyer, P. A., F. J. Zeng, A. Ducharne, J. C. Morrill, and R. D. Koster, 2000: The sensitivity of surface fluxes to soil water content in three land surface schemes. *J. Hydrometeor.,* **1,** 121–134.

Efron, B., and R. J. Tibshirani, 1993: *An Introduction to the Bootstrap*. Chapman and Hall, 436 pp.

Ek, M., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of the Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta Model. *J. Geophys. Res.,* **108,** 8851, doi:10.1029/2002JD003296.

Fast, J. D., S. Zhong, and C. D. Whiteman, 1996: Boundary layer evolution within a canyonland basin. Part II: Numerical simulations of nocturnal flows and heat budgets. *J. Appl. Meteor.,* **35,** 2162–2178.

Fels, S. B., and M. D. Schwarzkopf, 1975: The simplified exchange approximation—A new method for radiative transfer calculations. *J. Atmos. Sci.,* **32,** 1475–1488.

Ferrier, B. S., 1994: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description. *J. Atmos. Sci.,* **51,** 249–280.

Godfrey, C. M., and D. J. Stensrud, 2008: Soil temperature and moisture errors in operational Eta Model analyses. *J. Hydrometeor.,* **9,** 367–387.

Häberli, C., 2006: Assessment, correction and impact of the dry bias in radiosonde humidity data during MAP SOP. *Quart. J. Roy. Meteor. Soc.,* **132,** 2827–2852.

Hart, K. A., W. J. Steenburgh, D. J. Onton, and A. J. Siffert, 2004: An evaluation of the mesoscale-model-based model output statistics (MOS) during the 2002 Olympic and Paralympic Winter Games. *Wea. Forecasting,* **19,** 200–218.

——, ——, and ——, 2005: Model forecast improvements with decreased horizontal grid spacing over finescale intermountain orography during the 2002 Olympic Winter Games. *Wea. Forecasting,* **20,** 558–576.

Hill, C. D., 1993: Forecast problems in the western region of the National Weather Service: An overview. *Wea. Forecasting,* **8,** 158–165.

Hoggarth, A. M., H. D. Reeves, and Y.-L. Lin, 2006: Formation and maintenance mechanisms of the stable layer over the Po Valley during MAP IOP-8. *Mon. Wea. Rev.,* **134,** 3336–3354.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.,* **134,** 2318–2341.

Hu, X., J. W. Nielsen-Gammon, and F. Zhang, 2010: Evaluation of three planetary boundary layer schemes in the WRF model. *J. Appl. Meteor. Climatol.,* **49,** 1831–1844.

Janjić, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.,* **122,** 927–945.

——, 2002: Nonsingular implementaion of the Mello–Yamada level 2.5 scheme in the NCEP Meso Model. NCEP Office Note 437, 61 pp.

——, T. L. Black, M. E. Pyle, H.-Y. Chuang, E. Rogers, and G. J. DiMego, 2005: The NCEP WRF-NMM core. Preprints, *Joint WRF/MM5 User's Workshop,* Boulder, CO, NCAR, 2.9.

Knievel, J. C., G. H. Bryan, and J. P. Hacker, 2007: Explicit diffusion in the WRF model. *Mon. Wea. Rev.,* **135,** 3808–3824.

Kurkowski, N. P., D. J. Stensrud, and M. E. Baldwin, 2003: Assessment of implementing satellite-derived land cover data in the Eta Model. *Wea. Forecasting,* **18,** 404–416.

Lenschow, D. H., B. B. Stankov, and L. Mahrt, 1979: The rapid morning boundary-layer transition. *J. Atmos. Sci.,* **36,** 2108–2124.

Liu, Y., and Coauthors, 2008: The operational mesogamma-scale analysis and forecast system of the U. S. Army Test and Evaluation Command. Part II: Interchange comparison of the accuracy of model analyses and forecasts. *J. Appl. Meteor. Climatol.,* **47,** 1093–1104.

McCurdy, G. D., 1989: Radiation balance of a desert salt playa. M.S. thesis, Dept. of Soil Science and Biometeorology, Utah State University, 103 pp.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.,* **20,** 851–875.

Müller, M. D., and D. Scherer, 2005: A grid- and subgrid-scale radiation parameterization of topographic effects for mesoscale weather forecast models. *Mon. Wea. Rev.,* **133,** 1431–1442.

Myrick, D. T., and J. D. Horel, 2006: Verification of surface temperature forecasts from the National Digital Forecast Database over the western United States. *Wea. Forecasting,* **21,** 869–892.

National Oceanic and Atmospheric Administration, 1998: Automated Surface Observing System (ASOS) user's guide. National Weather Service, 61 pp. [Available from ASOS Program Office, NWS, 1325 East-West Highway, Silver Spring, MD 20910.]

Pagowski, M., 2004: Some comments on PBL parameterizations in WRF. *Proc. Joint WRF/MM5 Users' Workshop,* Boulder, CO, NCAR, 1.13. [Available online at http://www.mmm.ucar.edu/mm5/worshop/ws04/Session1/Pagowski.Mariusz_web.pdf.]

Pataki, D. E., B. J. Tyler, R. E. Peterson, A. P. Nair, W. J. Steenburgh, and E. R. Pardyjak, 2005: Can carbon dioxide be used as a tracer

of urban atmospheric transport? *J. Geophys. Res.,* **110,** D15102, doi:10.1029/2004JD005723.

Pepin, N. C., M. Losleben, M. Hartman, and K. Chowanski, 2005: A comparison of SNOTEL and GHCN/CRU surface temperatures with free-air temperatures at high elevations in the western United States: Data compatibility and trends. *J. Climate,* **18,** 1967–1985.

Ramaswamy, V., and S. M. Freidenreich, 1998: A high-spectral resolution study of the near-infrared solar flux disposition in clear and overcast atmospheres. *J. Geophys. Res.,* **103,** 23 255–23 273.

Reeves, H. D., and Y.-L. Lin, 2006: Effect of stable layer formation over the Po Valley on the development of convection during MAP IOP-8. *J. Atmos. Sci.,* **63,** 2567–2584.

——, and D. J. Stensrud, 2009: Synoptic-scale flow and valley cold pool evolution in the western United States. *Wea. Forecasting,* **24,** 1625–1643.

Rife, D. L., T. T. Warner, F. Chen, and E. G. Astling, 2002: Mechanisms for diurnal boundary layer circulations in the Great Basin Desert. *Mon. Wea. Rev.,* **130,** 921–938.

Smith, R. B., 1979: The influence of mountains on the atmosphere. *Advances in Geophysics,* Vol. 21, Academic Press, 87–230.

——, and Coauthors, 1997: Local and remote effects of mountains on weather: Research needs and opportunities. *Bull. Amer. Meteor. Soc.,* **78,** 877–892.

Struthwolf, M., 2005: An evaluation of fog forecasting tools for a fog event and non-event at Salt Lake City International Airport. NWS Tech. Attachment 05-05, 24 pp. [Available online at http://www.wrh.noaa.gov/wrh/05TAs/ta0505.pdf.]

Tapper, N. J., 1991: Evidence for a mesoscale thermal circulation over dry salt lakes. *Palaeogeogr. Palaeoclimatol. Palaeoecol.,* **84,** 259–269.

Vrhovec, T., 1991: A cold air lake formation in a basin: A simulation with a mesoscale numerical model. *Meteor. Atmos. Phys.,* **8,** 91–99.

Whiteman, C. D., 1982: Breakup of temperature inversions in deep mountain valleys: Part I. Observations. *J. Appl. Meteor.,* **21,** 270–289.

——, and T. B. McKee, 1982: Breakup of temperature inversions in deep mountain valleys. Part II: Thermodynamic model. *J. Appl. Meteor.,* **21,** 290–302.

——, and S. Zhong, 2008: Downslope flows on a low-angle slope and their interactions with valley inversions. Part I: Observations. *J. Appl. Meteor. Climatol.,* **47,** 2023–2038.

——, X. Bian, and S. Zhong, 1999: Wintertime evolution of the temperature inversion in the Colorado Plateau Basin. *J. Appl. Meteor.,* **38,** 1103–1117.

——, S. Zhong, W. J. Shaw, J. M. Hubbe, X. Bian, and J. Mittelstadt, 2001: Cold pools in the Columbia basin. *Wea. Forecasting,* **16,** 432–447.

Wolyn, P. G., and T. B. McKee, 1989: Deep stable layers in the intermountain western United States. *Mon. Wea. Rev.,* **117,** 461–472.

Zängl, G., 2002: Improved method for computing horizontal diffusion in a sigma-coordinate model and its application to simulations over mountainous topography. *Mon. Wea. Rev.,* **130,** 1423–1432.

——, 2005: Wintertime cold-air pools in the Bavarian Danube Valley basin: Data analysis and idealized numerical simulations. *J. Appl. Meteor.,* **44,** 1950–1971.

Zehnder, J. A., 2002: Simple modifications to improve fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model performance for the Phoenix, Arizona, metropolitan area. *J. Appl. Meteor.,* **41,** 971–979.

Zhong, S., and C. D. Whiteman, 2008: Downslope flows in a low-angle slope and their interactions with valley inversions. Part II: Numerical modeling. *J. Appl. Meteor. Climatol.,* **47,** 2039–2057.

——, ——, X. Bian, W. J. Shaw, and J. M. Hubbe, 2001: Meteorological processes affecting the evolution of a wintertime cold air pool in the Columbia basin. *Mon. Wea. Rev.,* **129,** 2600–2613.