Examination of Errors in Near-Surface Temperature and Wind from WRF Numerical Simulations in Regions of Complex Terrain

Hailing Zhang and Zhaoxia Pu

Department of Atmospheric Sciences, University of Utah, Salt Lake City, Utah

Xuebo Zhang

Computational Engineering and Science, University of Utah, Salt Lake City, Utah

ABSTRACT

The performance of an advanced research version of the Weather Research and Forecasting Model (WRF) in predicting near-surface atmospheric temperature and wind conditions under various terrain and weather regimes is examined. Verification of 2-m temperature and 10-m wind speed and direction against surface Mesonet observations is conducted. Three individual events under strong synoptic forcings (i.e., a frontal system, a low-level jet, and a persistent inversion) are first evaluated. It is found that the WRF model is able to reproduce these weather phenomena reasonably well. Forecasts of near-surface variables in flat terrain generally agree well with observations, but errors also occur, depending on the predictability of the lower-atmospheric boundary layer. In complex terrain, forecasts not only suffer from the model's inability to re-produce accurate atmospheric conditions in the lower atmosphere but also struggle with representative issues due to mismatches between the model and the actual terrain. In addition, surface forecasts at finer resolutions do not always outperform those at coarser resolutions. Increasing the vertical resolution may not help predict the near-surface variables, although it does improve the forecasts of the structure of mesoscale weather phenomena. A statistical analysis is also performed for 120 forecasts during a 1-month period to further investigate forecast error characteristics in complex terrain. Results illustrate that forecast errors in near-surface variables depend strongly on the diurnal variation in surface conditions, especially when synoptic forcing is weak. Under strong synoptic forcing, the diurnal patterns in the errors break down, while the flow-dependent errors are clearly shown.

1. Introduction

The near-surface atmosphere, namely, the bottom 10% of the atmospheric boundary layer, is unique due to its direct interaction with the earth's surface (Stull 1988). For instance, near-surface temperature is characterized by diurnal variation, with a maximum at local afternoon and a minimum at local midnight. This is very different from the free atmosphere in which temperature shows little diurnal variation. Turbulence causes the wind field in the atmospheric boundary layer (ABL) and the near-surface to behave differently from that in the free atmosphere because the ABL transports momentum, heat, and moisture between the earth's surface and the air above. Due to its unique features, accurate forecasts of near-surface atmospheric conditions are very important in many applications such as wind energy, agriculture, aviation, and fire weather forecasts. However, difficulties in forecasting near-surface variables such as temperature and wind have long been recognized and studied (Hanna and Yang 2001; Zhang and Zheng 2004).

To accurately simulate near-surface atmospheric conditions, several factors must be represented properly in numerical models. These include land use, topography, surface heat flux transport, and various characteristics of the lower atmosphere (Lee et al. 1989; Wolyn and McKee 1989; Shafran et al. 2000; Cheng and Steenburgh 2005). Thus, the accurate simulation of near-surface atmospheric diurnal variation is one of the most important and difficult tasks in numerical weather

Corresponding author address: Dr. Zhaoxia Pu, Dept. of Atmospheric Sciences, Rm. 819, University of Utah, 135 S 1460 E, Salt Lake City, UT 84112.
E-mail: zhaoxia.pu@utah.edu

prediction (NWP). Owing to our limited understanding of near-surface atmospheric processes and the uncertainties in model physics parameterizations, a comprehensive verification of the NWP models' performance in forecasting near-surface variables becomes a necessary step for model improvement.

Hanna and Yang (2001) found that the uncertainties regarding wind speed and direction in the lower atmosphere are primarily due to random turbulent processes that were not appropriately represented in the models, as well as errors in subgrid terrain and land use. They also argued that the models tend to underestimate the vertical temperature gradients in the lowest 100 m during the nighttime. Thus, the simulated boundary layer stability is not as strong as the observed. Considering the different capabilities of planetary boundary layer (PBL) parameterization schemes to reproduce atmospheric structures in the lowest few kilometers, Zhang and Zheng (2004) tested the performances of different PBL schemes in simulating near-surface temperature and wind speed and direction. Their results revealed that the model could reproduce diurnal variations in surface temperature and wind direction. However, all the boundary layer schemes underestimated (overestimated) wind speeds during the daytime (nighttime). Their study was conducted over the central United States during the summer, where little organized convection and topographical forcing was present.

The problem becomes more complicated in complex terrain. Liu et al. (2008) conducted an interrange comparison of the model analyses and forecasts of five U.S. Army test and evaluation command ranges over a 5-yr period. They concluded that forecast errors vary from range to range and season to season. They also found that larger errors are typically associated with complex terrain. Zhong and Fast (2003) compared three mesoscale numerical models and evaluated the simulations over the Salt Lake Valley for cases influenced by both weak and strong synoptic scenarios. They found a cold bias in the valley extending from the surface to the top of the atmosphere. The simulated nocturnal inversion was much weaker than the observed. There were significant errors in wind forecasts even under strong synoptic forcing. Hart et al. (2005) validated surface forecasts over mountainous terrain during wintertime. They employed the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) at high resolution with three nested domains at 36-, 12-, and 4-km horizontal grid spacing. Simulation results did show improved wind and precipitation forecasts in the 4-km horizontal grid-spacing domain, compared with those at the 12- and 36-km domains. However, temperature forecasts did not benefit from the high-resolution simulation. It is noteworthy that although the model properly simulated the persistent nocturnal cold-air pool along with the better-resolved orography at higher resolution, it still did not improve the 2-m temperature forecasts. Apparently, forecasting surface conditions in complex terrain is a challenging problem.

Mass et al. (2002) presented an objective multiyear verification of the University of Washington real-time MM5 forecasts. In their study, triple one-way nested domains at 36-, 12-, and 4-km horizontal grid spacings were used and the forecasts of near-surface atmospheric conditions (i.e., 2-m temperature, 10-m wind direction and speed, precipitation, and sea level pressure) from all three domains were verified at observation locations in western Washington State. Their results suggested that the forecasts benefited significantly from decreasing the grid spacing from 36 to 12 km. However, little improvement was found with further reduction in grid spacing from 12 to 4 km. These results are in contrast with the conclusions of some other studies. For instance, Rife and Davis (2005) suggested that the gains in forecast accuracy from finer grid spacing are generally incremental. With a regional climate simulation over complex terrain, Leung and Qian (2003) found that a higher-resolution simulation improves not only the spatial distribution and regional mean precipitation during summer but also snowpack during winter. However, they also commented that the accuracy of snow simulation is limited by factors such as deficiencies in the land surface model or biases in other model variables. The disagreement between these different studies further indicates the complexity of numerical prediction over complex terrain. Nevertheless, most of these previous studies emphasized the verification of synoptic cases at large and mesoscales. Little attention has been paid to the assessment of near-surface atmospheric conditions.

In this study, we attempt to assess the accuracy of the near-surface atmospheric conditions, specifically the 2-m temperature and 10-m wind speed and direction, in numerical simulations produced by the Weather Research and Forecasting Model (WRF). In particular, version 3.3 of an Advanced Research version of the WRF (ARW; Skamarock et al. 2008) is used for three typical severe weather events (i.e., a low-level jet, a cold front, and a wintertime persistent inversion) over the southern Great Plains (SGP) and the Intermountain West of the United States. Our purposes are not only to examine the ability of the ARW to predict near-surface atmospheric conditions, but also to compare the predictability of near-surface conditions in flat and

TABLE 1. Configurations of numerical simulations.

Case	1–3 Jun 2008	1–3 Dec 2010	Fall 2011 (15 Sep–14 Oct)
No. of domains	3	3	4
Horizontal grid spacing (km)	27, 9, 3	12, 4, 1.33	30, 10, 3.33, 1.11
Microphysics scheme	WSM 6	WSM 6	Purdue Lin
PBL scheme	YSU	MYJ	YSU
Cumulus scheme		Kain–Fritsch (not applied for grid spacing 9 km)	
Land surface scheme		Noah	
Longwave radiation		Rapid Radiative Transfer Model	
Shortwave radiation		Dudhia	

complex terrain. The sensitivity of numerical forecasts of near-surface atmospheric conditions to various PBL schemes and model resolutions is also investigated. In addition, forecasts during a 1-month period are evaluated to further reveal the characteristics of the forecast errors of near-surface variables under different synoptic forcings in complex terrain.

This paper is organized as follows. Section 2 briefly describes three synoptic events in two individual cases, as well as the numerical simulations and verification methods. Sections 3 and 4 detail the simulation and verification results for the three synoptic events. Error characteristics of the near-surface variables are also evaluated. Section 5 examines the sensitivity of numerical simulations of near-surface atmospheric conditions to various PBL parameterization schemes and model vertical resolutions. Section 6 characterizes the errors in near-surface variables statistically with forecasts during a 1-month period. Section 7 summarizes the results and offers several concluding remarks.

2. Description of cases, numerical simulations, and verification methods

a. Cases

1) 1–3 JUNE 2008: A FRONTAL SYSTEM AND A LOW-LEVEL JET

There are two events of interest during 1–3 June 2008: a front evolved over the north-central United States and a nocturnal low-level jet occurred over the SGP. Surface maps (not shown) show that a cold front initially located north of North Dakota at 1200 UTC 1 June entered North Dakota at 1500 UTC 1 June. It arrived in South Dakota at 0100 UTC 2 June, then changed to a stationary front at 1200 UTC 2 June and evolved into a cold front again as it moved southward. A temperature gradient of 7°C along with wind direction changes were found between the two stations closest to the front on both sides, with a southwest wind on the south side and a northeast wind on the north side.

A low-level jet was dominant from the surface up to 1800 m above ground level (AGL) over the entire SGP. It influenced near-surface conditions by interacting with the surface and the lower atmosphere. Radar wind profiles in Jayton, Texas, from the National Oceanic and Atmospheric Administration (NOAA) Profiler Network indicated two periods of evidently greater wind speed during 1–3 June (see Fig. 4a in section 3a). One was between 0200 and 1400 UTC 1 June and the other was between 0200 and 1400 UTC 2 June.

2) 1–3 DECEMBER 2010: A PERSISTENT INVERSION

A persistent inversion began on 29 November 2010 and was maintained over the Salt Lake Valley, Utah, for seven successive days. During this period, extremely strong surface cooling occurred during the night of 1–2 December, accompanied by a clear sky and a strong temperature inversion layer aloft. Very low temperatures were observed in the Salt Lake Valley and the adjacent mountains. A high pressure system controlled the area and helped build and maintain the persistent cold air pools during this period through downward air motion. The strong inversion layer extended from the surface up to 2000 m during this time.

b. A brief description of numerical simulations

Numerical experiments are conducted to simulate the aforementioned cases using the ARW with one-way nested domains. The initial and boundary conditions are derived from the National Centers for Environmental Prediction's (NCEP's) Northern American Mesoscale (NAM) model analysis by WRF preprocessing. A topography dataset at 30 arc-second (about 1000 m) resolution and an update land-use dataset with 27 land-use categories (instead of the 24 land-use categories provided by the WRF version 3.3 release) from the U.S. Geological Survey (USGS) are used in order to ensure more accurate surface conditions, especially for playa and desert regions in the western United States. The Noah land surface model is used because it predicts the land states, such as
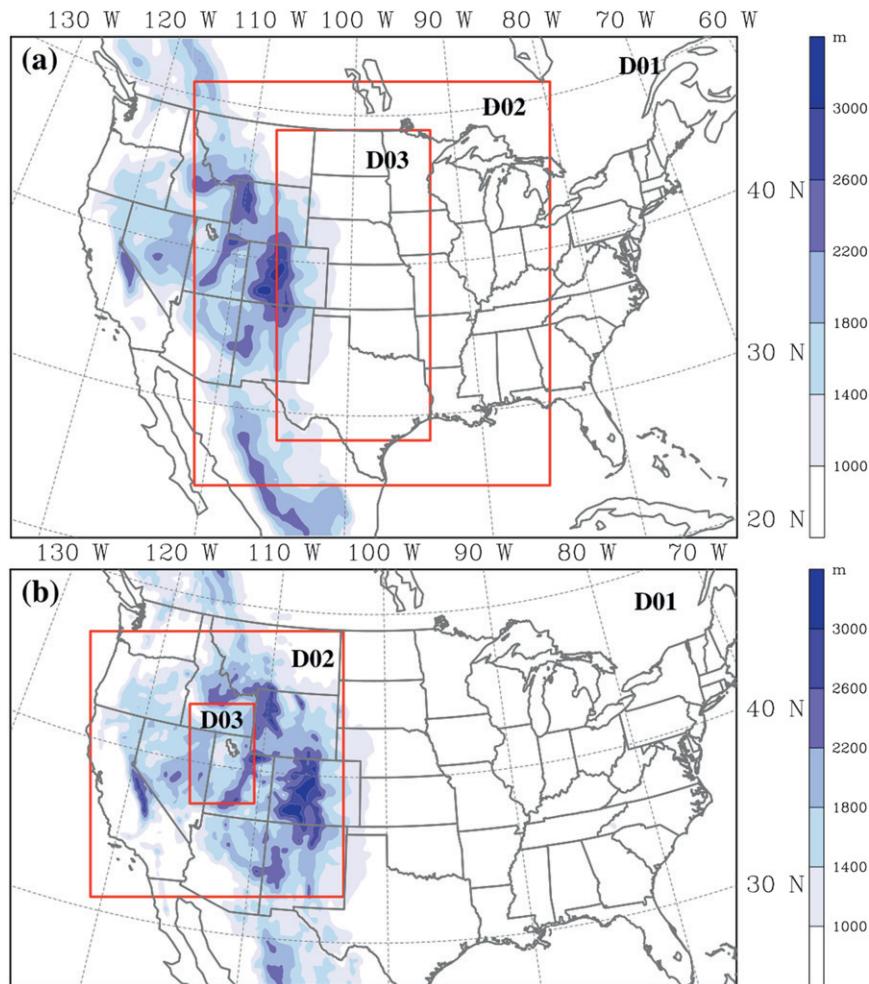
FIG. 1. Locations of model domains for numerical simulations: (a) 0000 UTC 1 Jun–
0000 UTC 3 Jun 2008 and (b) 0000 UTC 1 Dec–0000 UTC 3 Dec 2010. Shaded contours denote
the terrain heights.

surface temperature and soil moisture and temperature, in each layer with time. Table 1 lists the configurations of horizontal resolutions, model domains, and physical schemes used for each simulation in this study. Since PBL parameterization schemes contain key physical factors that strongly influence the predictability of near-surface atmospheric conditions, the PBL scheme used for each individual simulation (i.e., the control simulation) is chosen from the sensitivity studies (as described in section 5).

c. Verification methods

1) SYNOPTIC VERIFICATION

Verification is first conducted to evaluate the accuracy of the numerical simulation of each synoptic event. Simulation results are compared with available observations and analyses.

2) VERIFICATION OF NEAR-SURFACE ATMOSPHERIC CONDITIONS

The major emphasis of this study is on characterizing errors in the near-surface atmosphere. To quantify these errors, we use surface Mesonet observations (Horel et al. 2002) to verify the model's performance in terms of the near-surface variables, namely, 2-m temperature and 10-m wind speed and direction. According to Horel et al. (2002), quality control algorithms and data monitoring programs are performed for all available data. The quality-controlled data are then made available hourly with quality flags. In this study, only those observations with a quality flag of "OK" (the highest quality) are used for verification.

Since there is case-by-case variation in near-surface atmospheric conditions due to various synoptic systems and terrain, verification of near-surface atmospheric
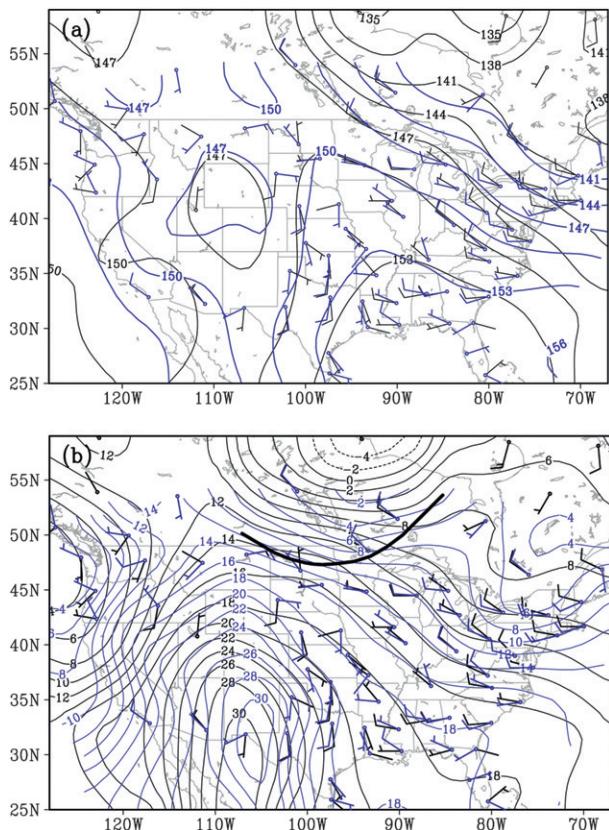
FIG. 2. Weather maps at 850 hPa valid at 0000 UTC 2 Jun 2008. (a) Geopotential heights (contour interval is 30 m) and wind barbs, and (b) temperature (contour interval is 2°C) with wind barbs. Black contour lines and wind barbs represent observations from the upper-level observation network. Blue contour lines and wind barbs denote model simulations (48-h forecast). The thick black curve in (b) marks the cold front.

conditions is performed for each synoptic event. Representation errors due to discrepancies between the model and actual terrain heights are commonly present in the forecasts of surface variables. Therefore, model terrain heights are compared against actual terrain heights for each case to examine the representative errors. Model performance is then checked for each case for each variable over time. In this study, we use variable mean, mean absolute error (MAE), and bias error (BE) of 2-m temperature and 10-m wind speed and direction against observations to characterize the errors in numerical simulations. We also calculate time-averaged mean absolute errors (TMAEs or cumulative MAEs) to average the MAEs over the whole simulation period. Because observational errors in wind direction are usually larger at lower wind speeds, only those observations with wind speeds greater than 1.5 m s^{-1} are used to verify wind direction. To verify the model simulation, simulation results at model grid points are interpolated to observation locations using a bilinear method. The statistical calculations are as follow:

$$\mathrm{MAE} = \frac{1}{n} \sum_{i=1}^{n} |F_i - O_i|$$

$$\mathrm{BE} = \frac{1}{n} \sum_{i=1}^{n} (F_i - O_i)$$

$$\mathrm{TMAE} = \frac{1}{n}\frac{1}{m} \sum_{t=1}^{m} \sum_{i=1}^{n} |F_{it} - O_{it}|,$$

where i denotes the ith observation, t denotes the observation time, O_i represents the value of the observation at the ith location, F_i denotes the forecast value interpolated to that observation location, n is the total number of stations, and m represents the total number of times used to calculate TMAE.

3. Simulation and verification: 1–3 June 2008

Triple-level, one-way nested domains (D01, D02, and D03 in Fig. 1a) with 27-, 9-, and 3-km horizontal grid
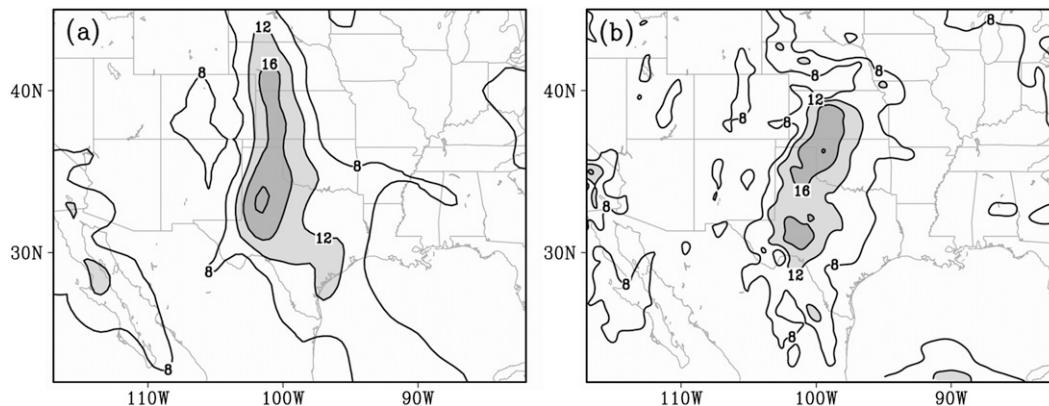
FIG. 3. Horizontal wind speeds (contour interval is 4 m s^{-1}) valid at 0900 UTC 2 Jun 2008 at 850 hPa: (a) NARR reanalysis and (b) 33-h model forecast from the 3-km domain. Wind speeds greater than 12 m s^{-1} are shaded.
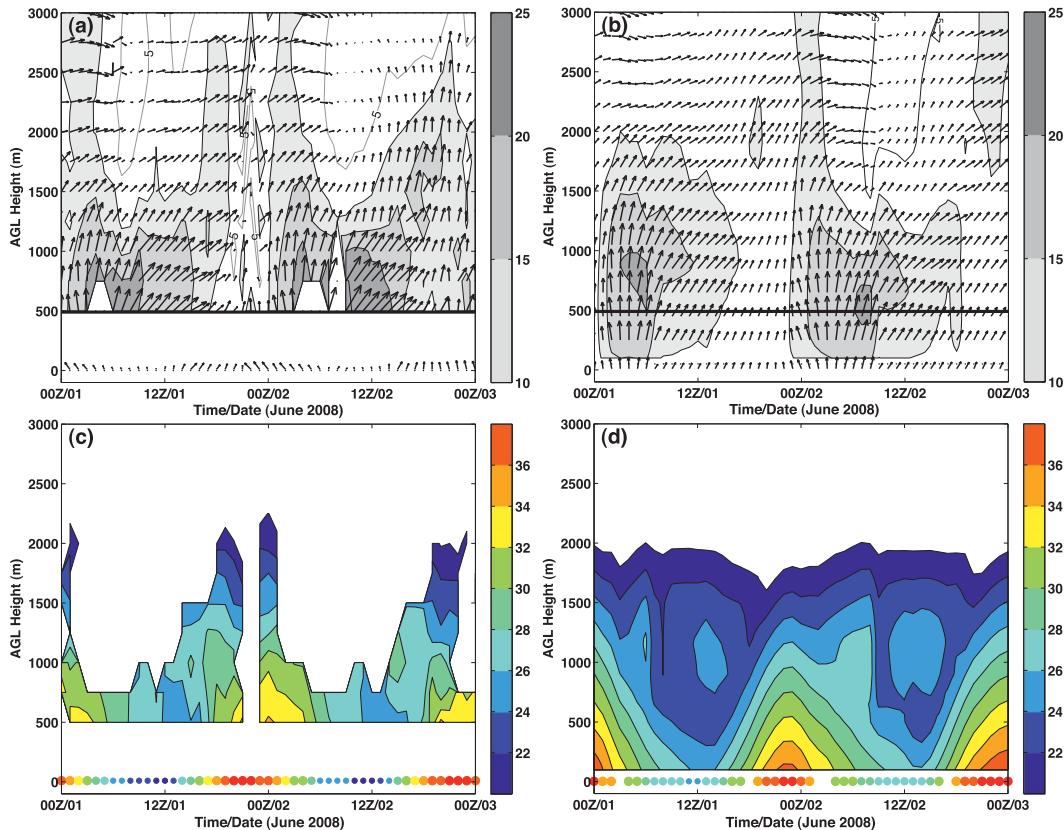
FIG. 4. Time series of vertical profiles from 0000 UTC 1 Jun to 0000 UTC 3 Jun 2008 at Jayton (33.01°N, 100.98°W; elevation: 707 m): (a) wind speeds (contour interval is 5 m s^{-1}) and vectors obtained from NOAA radar profiler and near-surface (10 m) winds from surface Mesonet observations, (b) wind speeds and vectors of 24–72-h model forecast from the 3-km domain, (c) temperatures obtained from NOAA radar profiler and near-surface (2 m) temperature from surface Mesonet, and (d) temperatures of 24–72-h model forecast from the 3-km domain. In (a) and (b), wind speeds greater than 10 m s^{-1} are shaded. NOAA profiler data are not available below 490 m AGL.

spacings (hereafter referred to as the 27-, 9-, and 3-km domains, respectively) are utilized in this simulation. There are 37 vertical levels from the surface up to 50 hPa. The innermost domain (i.e., the 3-km domain) focuses on the two weather systems of interest over the north-central United States and the SGP. The model is initialized at 0000 UTC 31 May, 24 h ahead of the verification.

Physical parameterization options, as listed in Table 1, include the WRF single-moment six-class microphysics scheme (WSM6; Hong and Lim 2006), the Yonsei University (YSU) PBL scheme (Hong and Pan 1996), the Kain–Fritsch cumulus parameterization scheme (Kain and Fritsch 1993), the Noah land surface model (Chen and Dudhia 2001), the Rapid Radiative Transfer Model for longwave radiation (RRTM; Mlawer et al. 1997), and the Dudhia shortwave radiation scheme (Dudhia 1989). The cumulus scheme is used only in the 27- and 9-km domains.

a. Synoptic verification

1) THE FRONTAL SYSTEM

Sounding observations from the National Weather Service (NWS) are compared with the simulated temperature, geopotential height, and wind barbs on a weather map at the 850-hPa pressure level at 0000 UTC 2 June 2008 (Fig. 2). The simulated geopotential heights almost overlap with the observations (Fig. 2a), indicating the front is well simulated. Over the frontal region, the observed and simulated temperature fields are almost identical. The larger temperature gradient over North Dakota and the wind barbs representing a realistic frontal system were reproduced by the model simulation (Fig. 2b).

2) THE LOW-LEVEL JET

To verify the forecast of the low-level jet, a comparison is first made using NCEP Northern American Regional
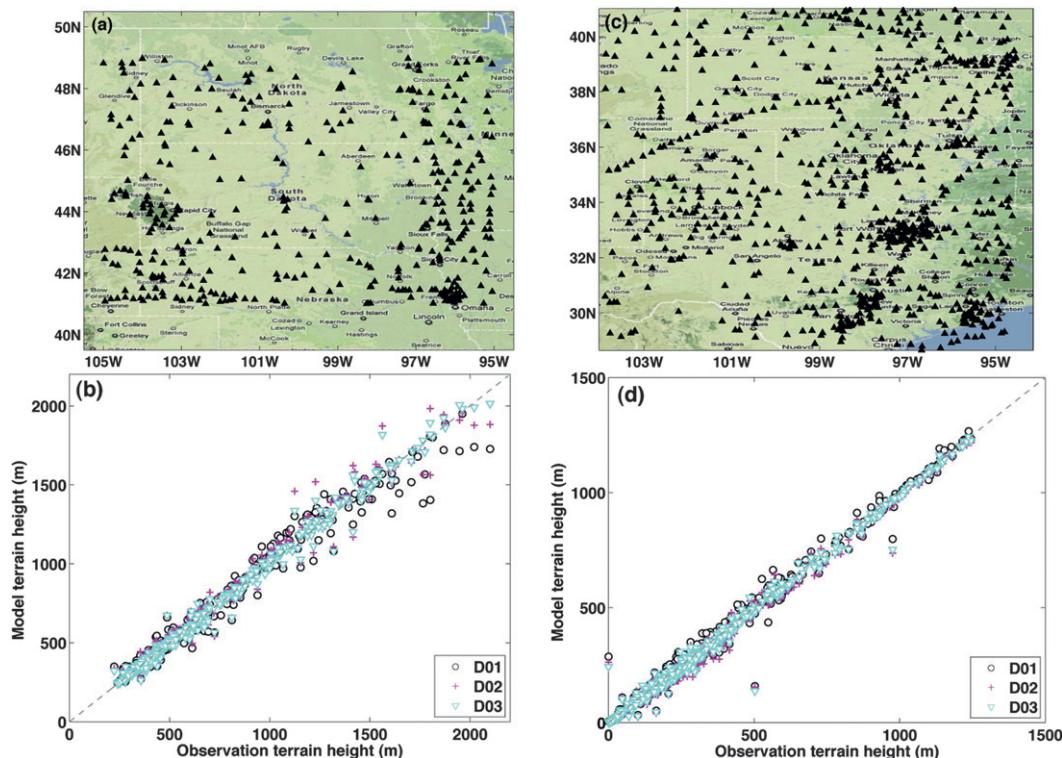
FIG. 5. The area and Mesonet observation stations used for verification: (a) for the front case and (c) for the low-level jet case. (b),(d) Comparison of the actual and model terrain heights for stations in (a) and (c), respectively. D01, D02, and D03 represent the domains at 27-, 9-, and 3-km horizontal resolutions, respectively. The straight dashed lines in (b) and (d) denote $Y = X$.

Reanalysis (NARR; Mesinger et al. 2006) data products. Figure 3 compares wind speeds from the NARR data and the model simulations at 850 hPa at 0900 UTC 2 June 2008. The similarities between the wind fields in the NARR and the simulation, in terms of jet coverage and intensity, prove that the model has successfully simulated the low-level jet.

Wind profile observations from the NOAA Profiler Network in Jayton clearly reveal the nocturnal jet. A time series of vertical profiles with wind speeds and vectors from 0000 UTC 1 June to 0000 UTC 3 June 2008 are displayed in Fig. 4a. Compared with the observed wind speeds and vectors, the simulated wind (Fig. 4b) reasonably presents the structure of the low-level jets, although the simulations underestimate the wind speed and intensity of the low-level jet.

Meanwhile, surface winds (10-m wind from Mesonet surface stations) are relatively calm during the night because the turbulence ceases after sunset. However, the simulated 10-m wind speeds during the night are much higher than the observations, indicating the model has not captured the decoupling between the surface and higher-level air after sunset. Figure 4c shows that

the near-surface air cools quickly after sunset due to radiative cooling. The air at 2 m is 3°–5°C colder than the air above since radiative cooling begins at the surface. The near-surface air temperature decreases immediately after sunset and then decouples with the air in the residual layer above that stays relatively warmer during the night. The simulated temperature (Fig. 4d) in the boundary layer captures the observed temperature inversions on both nights in the higher-level air. However, the simulated 2-m temperatures during both nights are much warmer than the observations. As a result, the model reproduces only a weak decoupling between the surface and the air above.

b. Verification of near-surface atmospheric conditions

1) THE FRONTAL CASE OVER FLAT TERRAIN

Figure 5a shows the area and Mesonet observation stations used for verifying the front. There are over 400 observations available hourly. Figure 5b compares the realistic and model terrains in these stations. The model terrain at all resolutions generally matches the actual
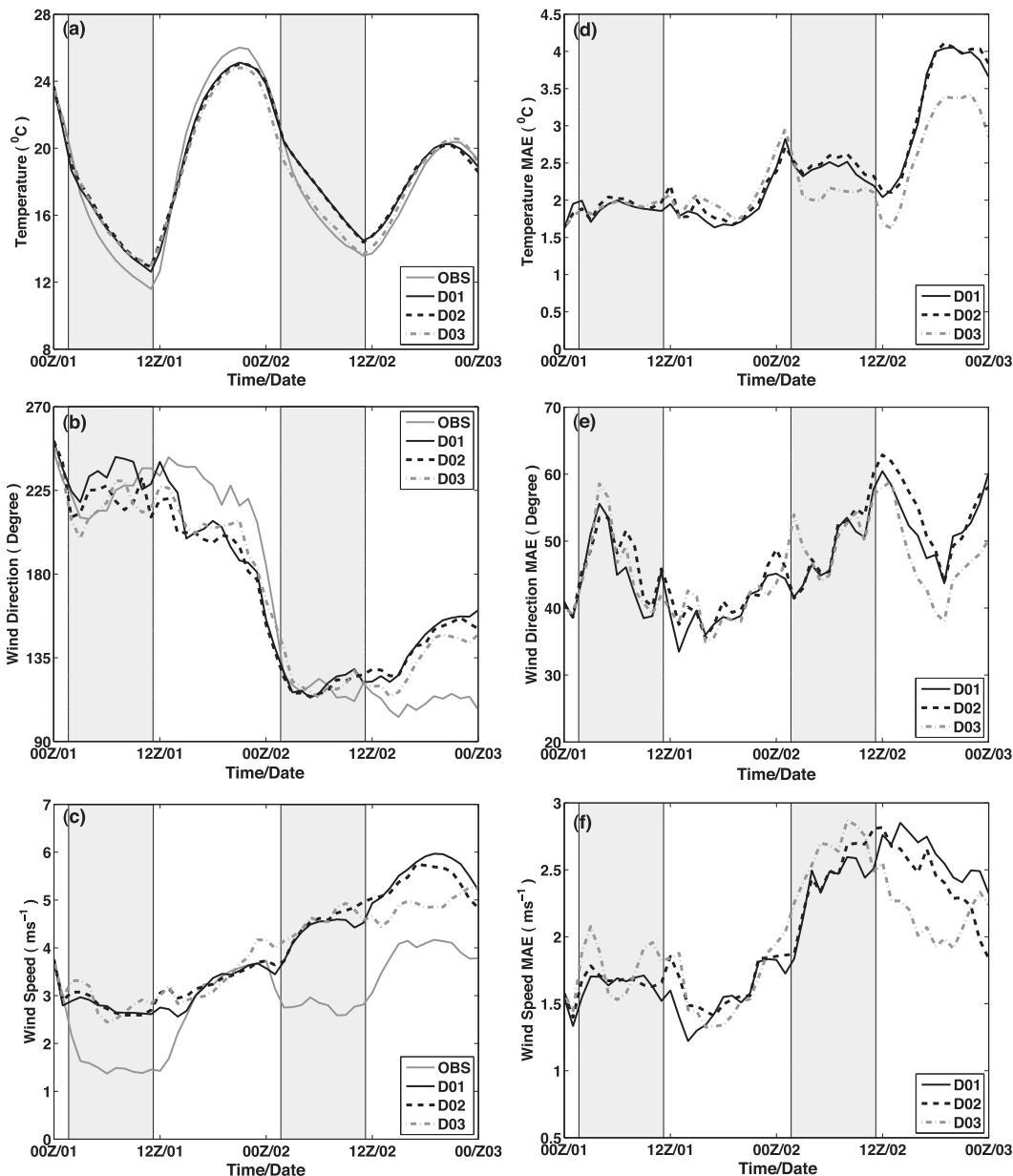
FIG. 6. Comparison of (a) 2-m temperature, (b) 10-m wind direction, and (c) 10-m wind speed between observations and simulations averaged over the area of the frontal system (e.g., the domain covered in Fig. 5a) and mean absolute errors of simulated (d) 2-m temperature, (e) 10-m wind direction, and (f) 10-m wind speed. D01, D02, and D03 represent results from model domains at horizontal resolutions of 27, 9, and 3 km, respectively. The shaded areas indicate the nighttime, hereafter.

terrain, while the 3-km domain (D03) makes the best match.

The simulated 2-m temperatures in the frontal area generally agree well with observations in all domains (Fig. 6a), showing no systematic bias. However, large MAEs occur in 2-m temperature at the end of the simulation when the stationary front changes to a cold front (Fig. 6d). The MAEs of 2-m temperature are

smallest in the 3-km domain, specifically during the second day.

The model captures the southwest-to-northeast wind direction change accompanying the frontal passage (Fig. 6b). Particularly, wind directions in the simulation agree well with the observations from 1200 UTC 1 June to 1200 UTC 2 June 2008. Relatively larger errors in wind direction occur near the end of the simulation period
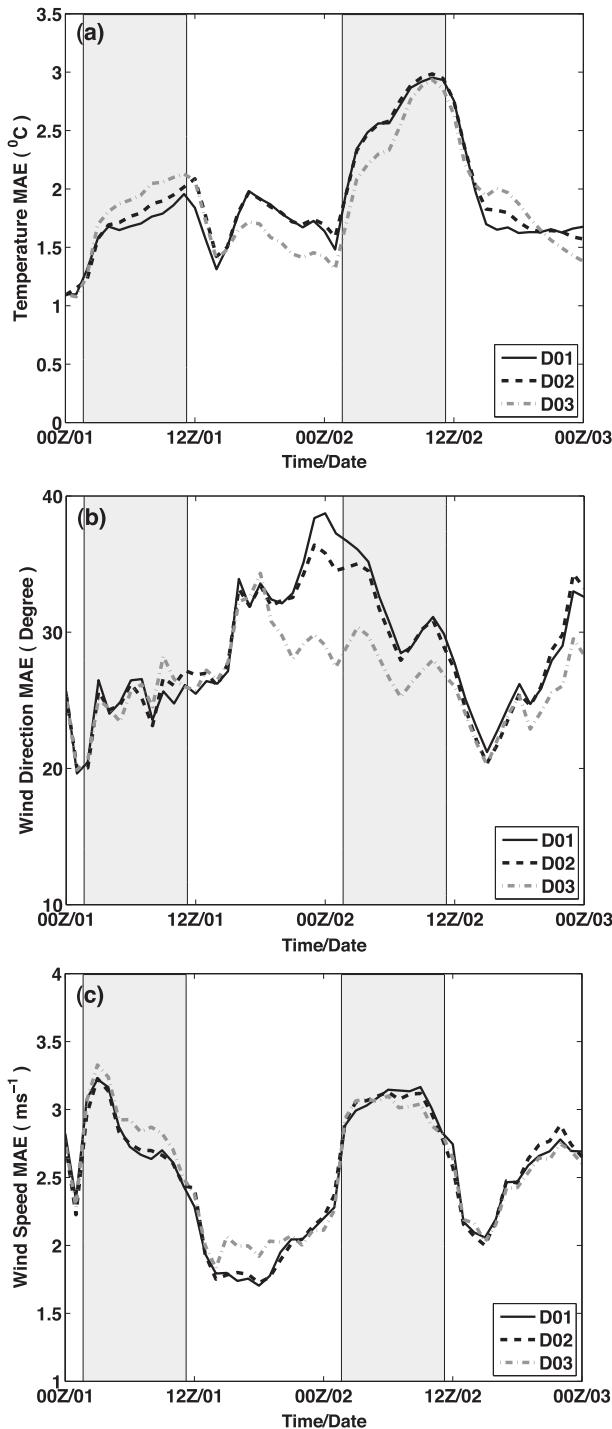
FIG. 7. Mean absolute errors over the low-level jet area of simulated (a) 2-m temperature, (b) 10-m wind direction, and (c) 10-m wind speed. D01, D02, and D03 represent results from model domains at horizontal resolutions of 27, 9, and 3 km, respectively.

(Fig. 6e) and are caused mainly by the rapid transitions from a cold front to a stationary front and then back again to a cold front. These changes complicate the

frontal event and present additional difficulties for the numerical simulations. Overall, the errors in 10-m wind direction are similar in all three domains in the first 36 h of forecasts. The 3-km domain performs better during the last 12 h, when rapid transitions take place. A diurnal feature of the errors in wind direction, characterized by larger errors during nighttime and smaller errors during daytime, can also be seen in Fig. 6e.

The observations clearly depict diurnal variations in wind speeds, with higher speeds during the daytime and lower speeds at night. The model well simulates the diurnal signals but generates larger errors at night (Figs. 6c and 6f) during the nocturnal jet. In particular, the model does well in simulating wind speeds between 0000 and 0200 UTC 1 June [corresponding to 1800–2000 central standard time (CST) 31 May] and 1500 UTC 1 June and 0100 UTC 2 June (corresponding to 0900–1900 CST 1 June), both of these periods are during the daytime. However, larger errors (Fig. 6f), characterized by positive biases (Fig. 6c), are found for nocturnal wind speeds. These positive bias errors in 10-m wind speed can be attributed to the incomplete representation of the decoupling between the higher-layer air and the near-surface atmosphere in the simulation (similar to that shown in Figs. 4b and 4d). In addition, the forecasts in the 3-km domain outperform the 9- and 27-km domains during the daytime, but not at night. This is mainly because the coarser-resolution domains do not resolve the intensity of the low-level jets as well as the higher-resolution domain does during the nighttime, and thus, the lower wind speeds produced by the coarser-resolution domains have less impact on the near-surface wind speeds. In other words, due to the model's inability to represent the decoupling between the near-surface layer and the boundary layer above (as mentioned in section 3a), the coarser-resolution domains outperform the high-resolution domain during nighttime.

2) THE LOW-LEVEL JET OVER FLAT TERRAIN

Figure 5c shows the Mesonet observation stations used for verifying the low-level jet. Figure 5d compares the actual and model terrain heights over these stations. The model terrain matches the actual terrain very well at most stations. The terrain heights in the 3-km domain (the innermost domain), again, best match the actual terrain heights.

The 3-km domain results in the smallest errors in 2-m temperature during 1500 UTC 1 June–1200 UTC 2 June (Fig. 7a). However, it produces the largest errors during the first night and errors that are comparable to the other domains in the early morning for both days. The
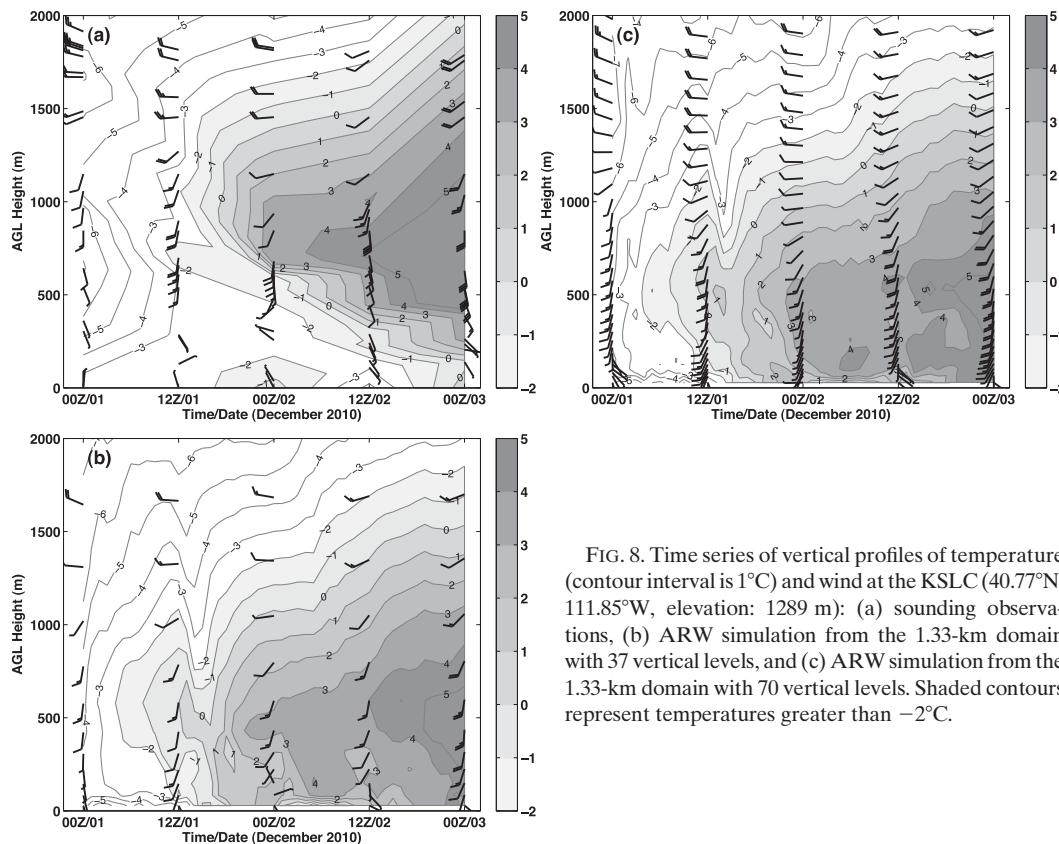
FIG. 8. Time series of vertical profiles of temperature (contour interval is 1°C) and wind at the KSLC (40.77°N, 111.85°W, elevation: 1289 m): (a) sounding observations, (b) ARW simulation from the 1.33-km domain with 37 vertical levels, and (c) ARW simulation from the 1.33-km domain with 70 vertical levels. Shaded contours represent temperatures greater than −2°C.

errors peak at 1200 UTC 1 June and 1200 UTC 2 June, when the boundary layers are most stable. Southerly flow dominates during the simulation period. The errors in 10-m wind direction are relatively small due to the strong southerly forcing (Fig. 7b). Owing to the influence of the low-level jet and the inaccurate representation of the nocturnal decoupling and radiative cooling as mentioned above, there are relatively larger errors in 10-m wind speed during the nighttime (Fig. 7c).

Accurate simulation of the transition boundary layer and the typical stable boundary layer near the ground is still one of the challenges in numerical simulation. The method of parameterizing the stable boundary layer has also been an active research area in recent studies (Brown and Wood 2003; Teixeira et al. 2008). A discussion of the best way to overcome forecast errors in a stable boundary layer is beyond the scope of this study. However, accurate forecasts of near-surface conditions depend on the model's ability to simulate the stable boundary layer.

4. Simulation and verification: 1–3 December 2010

The simulation is initialized at 0000 UTC 31 November, and the results from 1 to 3 December 2010 are used to verify the persistent inversion. Three-level, one-way

nested domains (D01, D02, and D03) at 12-, 4-, and 1.33-km horizontal grid spacings (hereafter referred to as the 12-, 4-, and 1.33-km domains, respectively) are used. The innermost domain (i.e., the 1.33-km domain) focuses on the Salt Lake Valley and its surrounding mountains. The model also includes 37 vertical levels from the surface up to 50 hPa.

Physical parameterization configurations for this case (see Table 1) are the same as for the 1–3 June 2008 case, except for the Mellor–Yamada–Janjić (MYJ) turbulent kinetic energy (TKE) PBL scheme (Mellor and Yamada 1982). The cumulus scheme is used only in the 12-km domain.

a. Synoptic verification

Sounding observations, obtained every 12 h from the station at the Salt Lake City International Airport (KSLC) are available to examine the structure of the atmospheric boundary layer. Model-simulated winds and temperatures are interpolated to the sounding locations for comparison. Figure 8 shows the evaluation of the temperature and wind fields for both soundings and simulations throughout the 2-day period. Observations show warmer air (relative to near-surface air) above
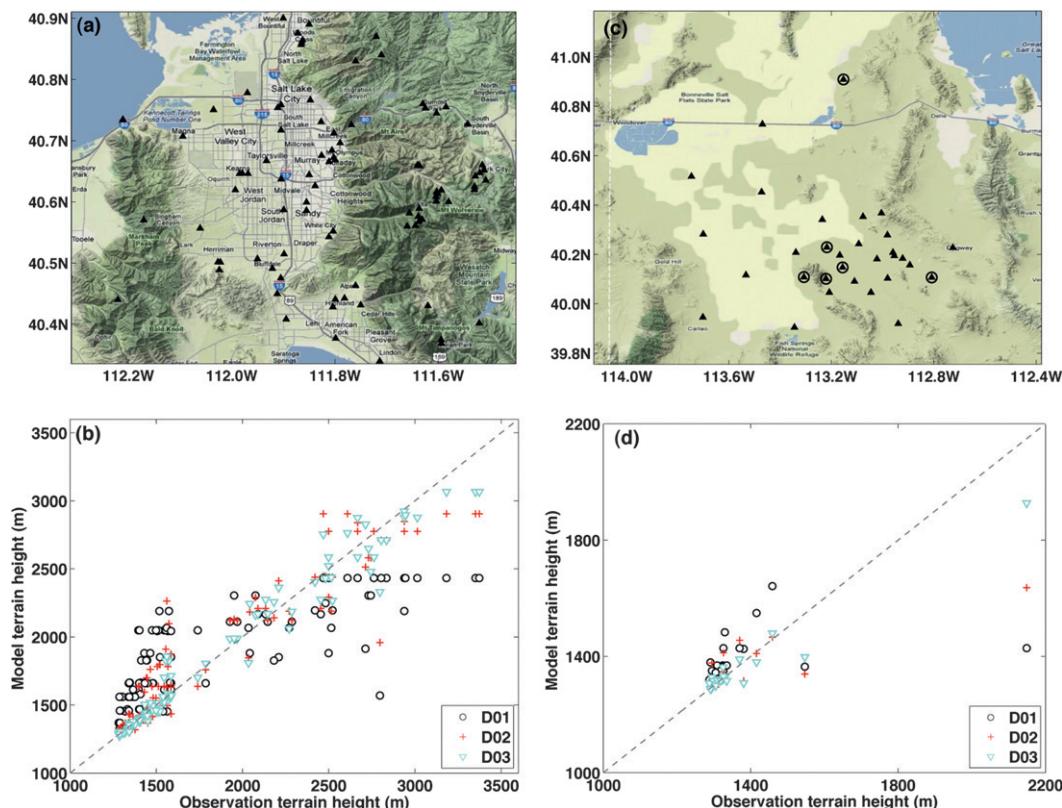
FIG. 9. As in Fig. 5, but for the persistent inversion case over the (a),(b) Salt Lake Valley and (c),(d) DPG. D01, D02, and D03 represent the domains at 12-, 4-, and 1.33-km horizontal resolutions, respectively.

the surface; namely, an inversion layer was present throughout the entire period, although the inversion was more intense in the late stages. The top of the inversion varies from 700 to 1200 m AGL during this period. The simulation reproduces the persistent inversion for the entire period. The wind shears, which were present as southeasterly to southerly beneath the inversion layer and southwesterly above it at 1200 UTC 2 December and 0000 UTC 3 December 2010, are well captured in the simulations, although the transition heights are slightly different from the observations. The simulated heights of the inversion layer are lower than those in the sounding observations. The simulated temperature gradient at the bottom of the boundary layer is not as strong as the observed. These results are similar to those of Hanna and Yang (2001). Apparently, discrepancies between simulations and observations can be attributed mainly to errors in the simulation of the near-surface atmospheric conditions.

b. Verification of near-surface atmospheric conditions

Figure 9a shows the distribution of Mesonet observation stations used for verifying this case. In complex

terrain, observations are distributed unevenly in the Salt Lake Valley and the surrounding mountains. Figure 9b compares the model and actual terrain heights for all three domains. The 1.33-km domain represents the actual terrain substantially better than the 12- and 4-km domains. The 12-km domain misrepresents lower (higher) terrain [less (greater) than 2000 m in the valley (mountains)] with higher (lower) heights. Consequently, the coarser-resolution domain does not resolve the deep valley and sharp mountains accordingly. The 4-km domain has an intermediate ability to represent the terrain compared with the 1.33- and 12-km domains.

Because of the large differences in terrain height between the stations in the valley and those in the mountains, the stations are separated into two groups during the verification: stations inside the Salt Lake Valley and those in the surrounding mountains (valley stations and mountain stations hereafter). A similar separation was used in Hart et al. (2005).

1) SALT LAKE VALLEY: VALLEY STATIONS

Consistent with synoptic verification results, large forecast errors are found for 2-m temperature at the valley stations, especially during the night of 2 December
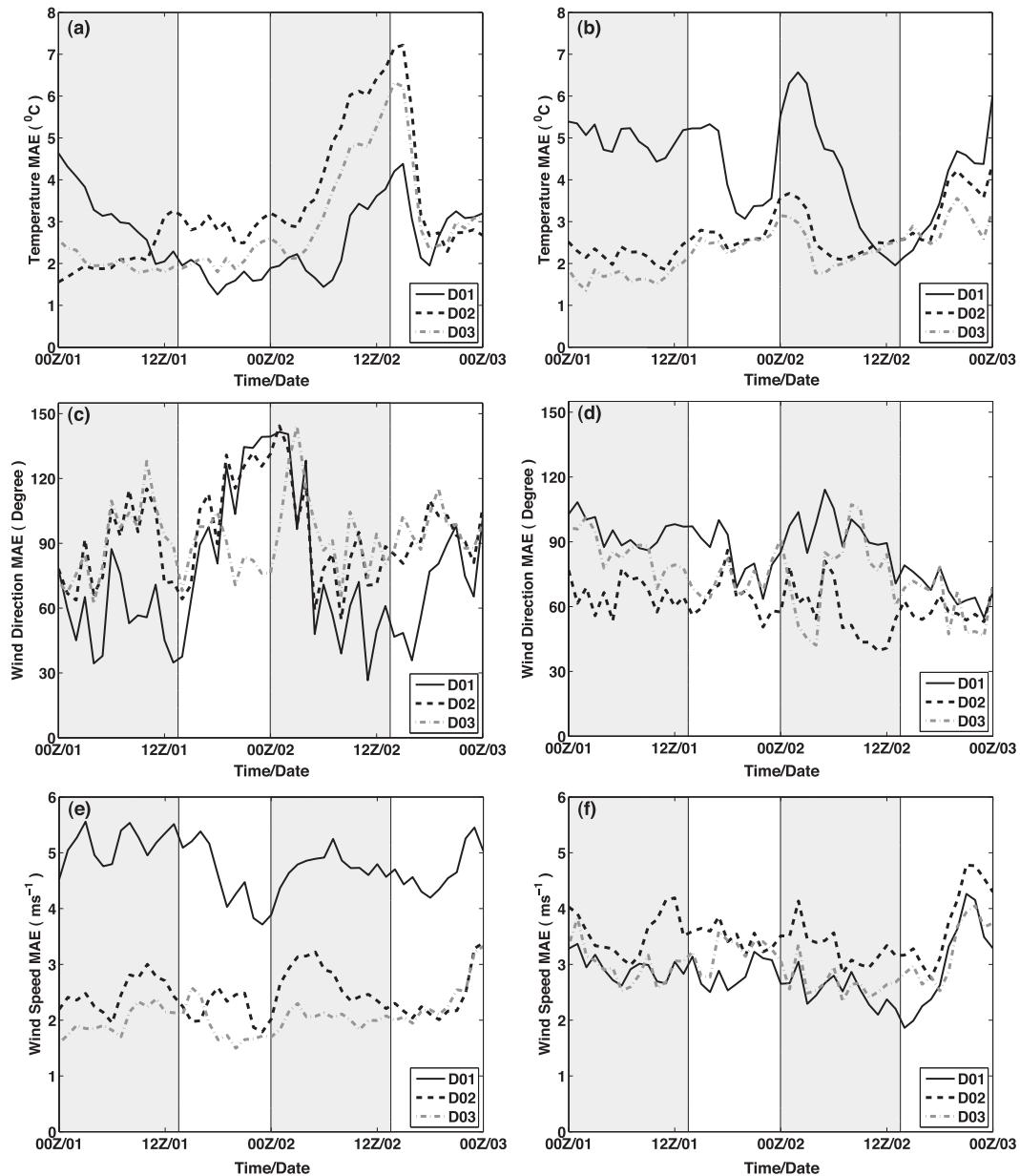
FIG. 10. MAEs by station type for numerical simulations of (a),(b) 2-m temperature; (c),(d) 10-m wind direction; and (e),(f) speed. Figures in the left column [(a), (c), and (e)] represent MAEs for valley stations, and these in the right column [(b), (d), and (f)] represent mountain stations. D01, D02, and D03 represent the domains at 12-, 4-, and 1.33-km horizontal resolutions, respectively.

2010 (Fig. 10a), when the valley was undergoing extreme nocturnal cooling induced by the persistent inversion. The 12-km domain produces the smallest MAEs during that night, while the 4-km domain produces the largest MAEs. This degradation accompanying the increase in resolution from 12 to 4 km can be attributed to the predictability of the intense inversion and the terrain representation in the ARW model. Specifically, the

terrain heights in the 4-km domain are typically lower than those in the 12-km domain at the valley stations. During the inversions, the temperature usually increases with height inside the valley. However, the model cannot fully capture the intense cold pools, resulting in a temperature profile that decreases with height in the near-surface layer. Therefore, the 4-km domain produces warmer surface temperatures than the 12-km
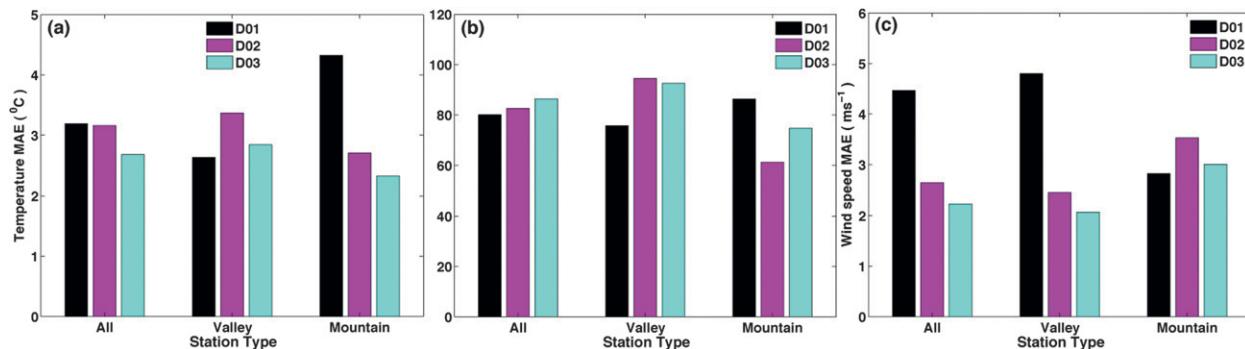
FIG. 11. Time-averaged MAEs of (a) 2-m temperature, (b) 10-m wind direction (degrees), and (c) 10-m wind speed by station type. D01, D02, and D03 represent the simulations from the domains at 12-, 4-, and 1.33-km horizontal resolutions, respectively.

domain due to its deeper valley representation, causing even larger warm biases. A similar argument was presented by Hart et al. (2005) using the MM5 model.

It is interesting that the MAEs of the 1.33-km domain are intermediate between those of the 12- and 4-km domains. As seen in Fig. 8, the 1.33-km domain captures the inversion in the lower atmosphere and properly represents the temperature lapse rate (i.e., increase with height). With a deeper topography and proper lapse rate representation, the 1.33-km domain produces better 2-m temperature forecasts than the 4-km domain. However, MAEs of the 2-m temperature are still larger in the 1.33-km domain than in the 12-km domain because the temperature lapse rate in the lower atmosphere (namely, the inversion intensity) resolved by the 1.33-km domain is much weaker than that of the sounding observations (as shown in Fig. 8).

The TMAEs in wind direction over all the stations are about 80° (Fig. 11b). Both the 4- and 1.33-km domains produce degraded wind direction forecasts (Figs. 10c and 11b) because of the contradiction between the model's failure to fully capture the strong inversion and the better terrain representation in the higher-resolution domains, as discussed above.

The MAEs of 10-m wind speed at the valley stations are significantly reduced in the 4-km domain (Figs. 10e and 11c). Much greater wind speeds are produced in the 12-km domain because it has a shallower valley floor (at a higher elevation than those of the 4- and 1.33-km domains). There are larger MAEs in wind speed since the observed winds are relatively calm at the surface.

2) SALT LAKE VALLEY: MOUNTAIN STATIONS

Meanwhile, increasing the model horizontal resolution from 12 to 4 and then 1.33 km substantially improves the temperature forecasts over the mountain stations (Figs. 10b and 11a). These forecasts benefit from

the sharper and better representation of mountain terrain at the higher resolution while the mountain stations are above the cold pools.

The wind direction forecasts are improved in the 4-km domain from the 12-km domain but are degraded in the 1.33-km domain from the 4-km domain (Fig. 10d). The improvement in the 4-km domain is mainly because the mountain stations are more connected to the free atmosphere in the forecasts at higher resolution. The reasons for the degradation in the 1.33-km domain are unknown, reflecting the challenges in producing accurate forecasts of near-surface atmospheric conditions. In reality, the accurate forecasting of near-surface wind direction is extremely difficult, as it depends not only on the terrain representation in the model, but also on the accurate prediction of wind speed, and the thermal and dynamical forcings. In addition, low wind speeds at the surface layer pose extra difficulties in forecasting wind
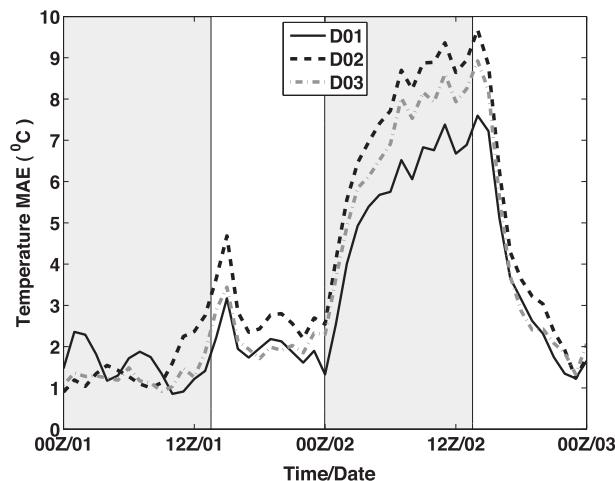
FIG. 12. MAEs of 2-m temperature at DPG. D01, D02, and D03 represent the simulations from the domains at 12-, 4-, and 1.33-km horizontal resolutions, respectively.
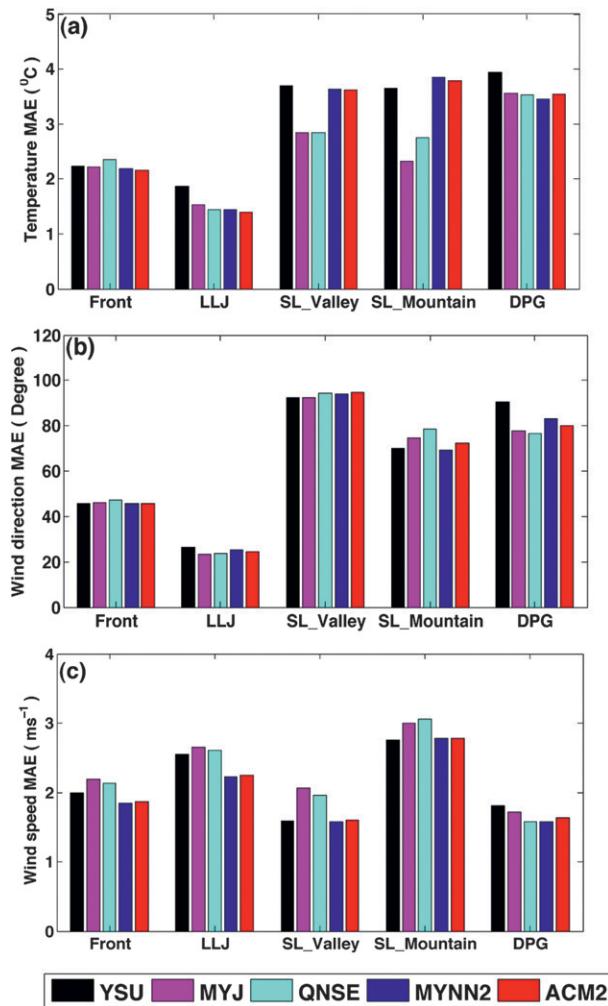
FIG. 13. TMAEs of (a) 2-m temperature, (b) 10-m wind direction, and (c) 10-m wind speed for various cases and locations with various PBL schemes. All results are from the innermost domain.

direction, although observations with speeds less than 1.5 m s^{-1} are already excluded in the comparison.

The simulation in the 12-km domain produces the smallest MAEs in wind speed. Slight degradations are found in the forecasts of the 1.33-km domain and relatively larger degradations are found in the 4-km domain (Figs. 10f and 11c).

3) DUGWAY PROVING GROUND

The great complexity of the terrain in the Salt Lake Valley presents a significant challenge to forecast verification. To further examine error characteristics in near-surface variables in complex terrain, an additional verification is conducted in the Dugway Proving Ground (DPG) area. DPG is located approximately 80 mi southwest of Salt Lake City. It is characterized by complex

terrain and is surrounded on three sides by mountain ranges. Considering the height of its mountains relative to its flat regions, the DPG area is more representative of common complex terrain features.

Figure 9c shows the DPG area map and stations used for verification. Currently, there are a total of 31 automatic surface stations in the DPG area. However, six stations (denoted by circles) are not used for verification in this case since they did not begin providing observations until 2011. Figure 9d compares the actual and model terrain heights. Similar to the previous section, the 1.33-km domain has the best terrain representation. There is evidently a large error in both the 12- and 4-km domains in representing the actual terrain at station DPG16, located on a mountaintop at an elevation of 2149 m. For instance, the terrain height at DPG16 in the 12-km domain is only 1450 m.

The effect of nocturnal cooling in the DPG area is shown in Fig. 12. The temperature error is as high as 9°C during the cooling night. The 12- and 4-km domains produce the smallest and the largest MAEs, respectively, while the 1.33-km domain produces intermediate MAEs. The 2-m temperature errors display similar features to those of the valley stations, as seen in Figs. 10a and 11a, since they are under the same weather regime and synoptic environment. The MAEs of wind direction and speed also show patterns similar to those of the valley stations (not shown). To eliminate the impact of the mismatched terrain at DPG16, statistical analyses are rerun but with this station excluded. The results remain almost the same. This implies that the representative error caused by terrain mismatch is not the sole reason for the errors in simulated near-surface variables in complex terrain.

5. Sensitivity to PBL schemes and vertical resolution

It has been recognized that PBL parameterization schemes have a substantial influence on the simulation of surface variables (Hu et al. 2010; Shin and Hong 2011). It is also commonly believed that simulations could be improved with increased model vertical resolution, especially in the boundary layer, in which the model can better resolve small-scale processes. To examine the impact of PBL schemes and model vertical resolution on the simulation of near-surface atmospheric conditions, additional experiments are conducted and discussed in this section.

a. Sensitivity to various PBL schemes

The ARW has multiple PBL scheme options, characterized by different closure methods, prognostic
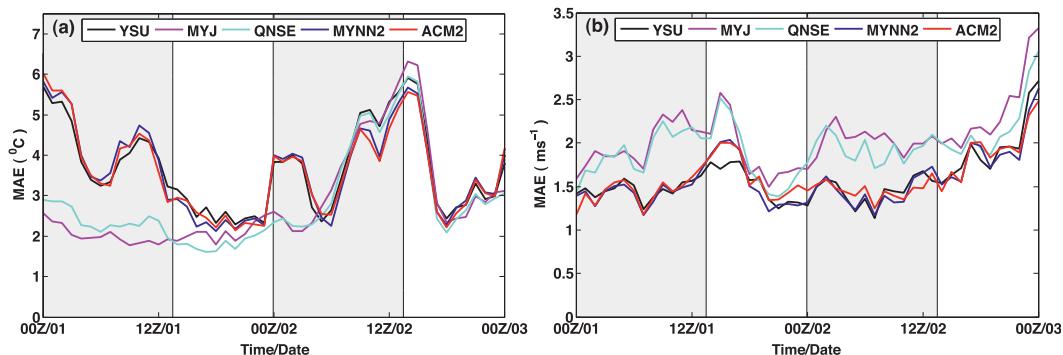
FIG. 14. MAEs of (a) 2-m temperature and (b) 10-m wind speed with various PBL schemes from the 1.33-km domain for Salt Lake Valley stations.

variables, cloud mixing, and other aspects. The first set of sensitivity experiments uses various PBL schemes while keeping other configurations the same, as specified in the control simulations (Table 1). These experiments are designed to evaluate the sensitivity of numerical simulation of near-surface temperature and wind fields to different PBL schemes.

Five PBL schemes in the ARW are tested and compared—YSU, MYJ, quasi-normal scale elimination (QNSE; Sukoriansky et al. 2005), Mellor–Yamada–Nakanishi–Niino level 2.5 (MYNN2; Nakanishi and Niino 2004), and the Asymmetric Convective Model version 2 (ACM2; Pleim 2007). Among these schemes, the YSU and ACM2 are first-order, nonlocal schemes. They do not require additional prognostic equations to describe the effects of turbulence on mean variables. They are both based on the K profile in determining the diffusivity in the boundary layer and consider nonlocal mixing by convective large eddies. The YSU scheme expresses nonlocal mixing by simply adding a nonlocal gradient term in the eddy diffusion equation. The YSU scheme is modified in ARW version 3 (Hong and Kim 2008) to enhance mixing in the stable boundary layer. The ACM2 scheme explicitly expresses the nonlocal upward flux transport. The MYJ, QNSE, and MYNN2 schemes are classified as 1.5-order TKE closure schemes. They calculate eddy diffusion coefficients by predicting TKE. They differ in how they define the coefficients in the diffusion equation. More details about these PBL schemes can be found on the ARW Web site (http://www.mmm.ucar.edu/wrf/users), as well as in Shin and Hong (2011) and Hu et al. (2010).

Figure 13 depicts the sensitivity of near-surface variable forecasts to model PBL schemes by comparing the TMAEs of surface variables over the whole simulation period in each case. For temperature (Fig. 13a), the TMAEs of the cold front and the low-level jet in

summer 2008 are smaller than those of the inversion case in winter 2010. For the frontal and low-level jet cases, all schemes generate similar TMAEs. For the inversion case, however, the TMAEs of temperature are very sensitive to the choice of PBL scheme. The MYJ and QNSE schemes produce smaller TMAEs than the others at both valley and mountain stations. All PBL schemes generate similar errors in wind direction for all cases (Fig. 13b), except that the MYJ and QNSE schemes produce relatively larger TMAEs at the mountain stations. In terms of wind speed, the ACM2 and MYNN2 schemes perform best in the frontal and low-level jet events (Fig. 13c). For the inversion, the YSU, MYNN2, and ACM2 schemes perform equally well at both valley and mountain stations. The MYJ and QNSE schemes, which produce the best temperature forecasts for the inversion (as seen in Fig. 13a), are the least accurate in wind speed forecasts.

Overall, all schemes perform similarly in the frontal and low-level jet cases in terms of temperature and wind direction forecasts. The MYJ and QNSE schemes produce relatively larger TMAEs in wind speed in the frontal and low-level jet cases. They improve the temperature forecasts at both valley and mountain stations for the winter 2010 inversion but at the same time produce the least accurate wind speed and direction forecasts at the mountain stations.

Figure 14a further examines the sensitivity of 2-m temperature simulation to model PBL schemes as a function of forecast leading time in the Salt Lake Valley for the December 2010 simulation. The five schemes split into two groups. The group with the MYJ and QNSE schemes performs much better than the group with the three other schemes, especially during the first day of the simulation. The MYJ and QNSE schemes, however, have larger MAEs in the 10-m wind speed forecast, as shown in Fig. 14b.
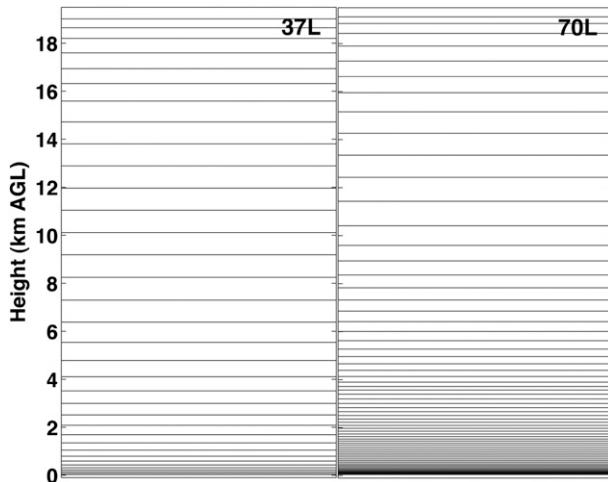
FIG. 15. Sketches of vertical models with (left) 37 and (right) 70 levels in AGL heights.

The above results indicate that no single PBL scheme leads to overall improvement in the forecasts of near-surface wind and temperature fields for all cases and regions. The model is more sensitive to the choice of PBL scheme in the inversion case in complex terrain than in the frontal and low-level jet cases in flat terrain. In addition, one PBL scheme that results in better wind (temperature) forecasts does not reproduce improved temperature (wind) forecasts. Because the MYJ and YSU schemes have been frequently used in recent applications and also because of their respective performance in accurately forecasting wind and temperature fields, they were chosen for the control experiments for the low-level jet in flat terrain and the inversions in complex terrain, as listed in Table 1.

b. Sensitivity to model vertical resolution

The other set of experiments is used to examine the sensitivity of model simulations to vertical resolution. In these experiments, the vertical levels are increased to 70 (70L) instead of 37 (37L) as in the control simulations. Figure 15 shows sketches of vertical model levels in both configurations. The increase in the vertical levels is more obvious below 6 km AGL, especially the lowest 2 km AGL. As seen in Fig. 8c, simulations with increased vertical resolution can better resolve the structure of the persistent inversion in the boundary layer and produce more reasonable inversion depth and intensity. However, the higher vertical resolution does not improve the forecasts of near-surface variables. Figure 16 shows that both experiments (with 37 and 70 vertical levels) perform almost identically in terms of the TMAEs for both 2-m temperature and 10-m wind speed and
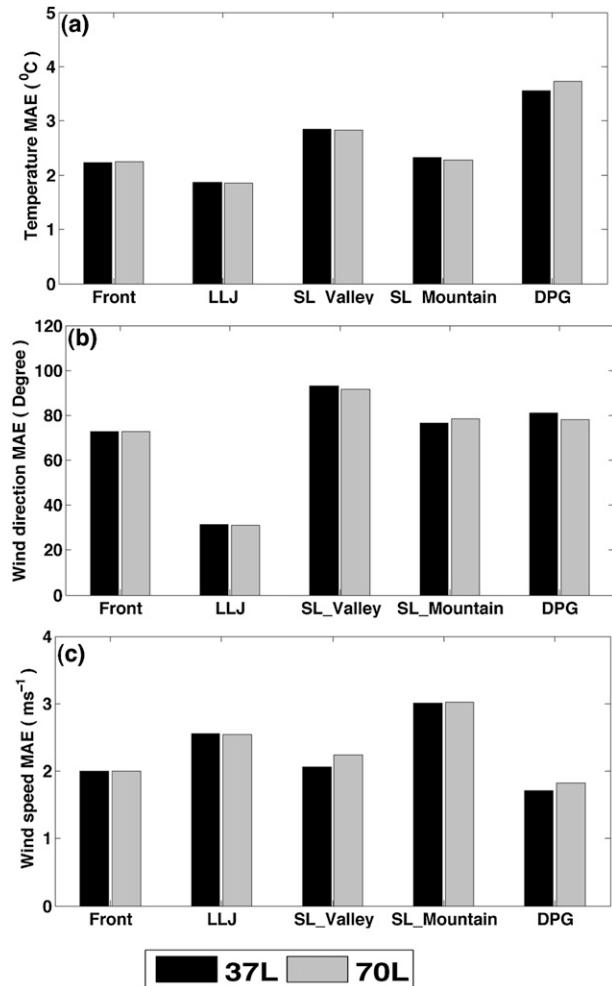
FIG. 16. As in Fig. 13, but for different model vertical resolutions.

direction for almost all cases and over all regions. Overall, improvement in surface forecasts is not ensured by increased vertical resolution, although it can help to better resolve the structures of the mesoscale phenomena in the boundary layer.

6. Characteristics of flow-dependent errors: Statistics over a 1-month period in the fall of 2011

The above results from the three typical weather patterns indicate the case-by-case variability of the errors in forecasts of near-surface variables. To further understand the general characteristics of the errors in near-surface forecasts in complex terrain, additional verification is conducted for forecasts over a 1-month period in the DPG area. We chose the DPG area as the focus for multiple case statistics for two reasons: 1) as mentioned above, the DPG area better represents
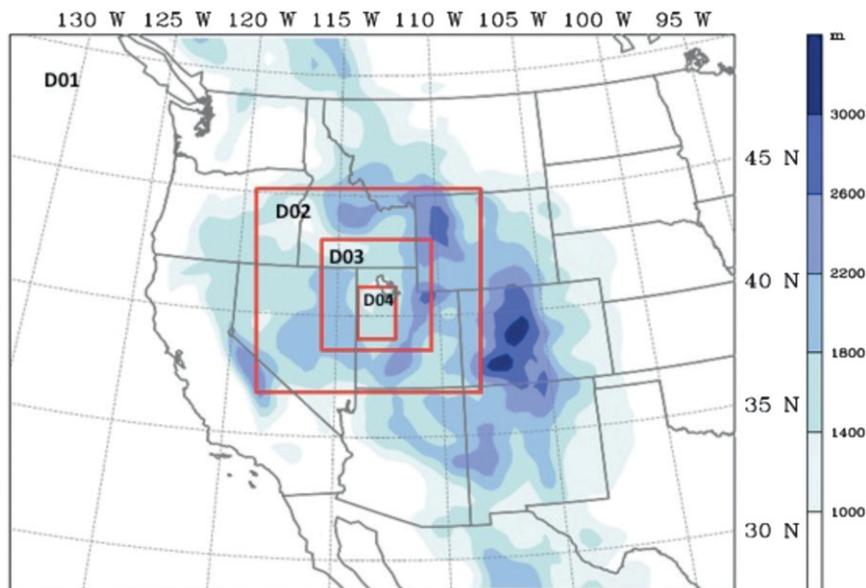
FIG. 17. Locations of model domains for near-real-time forecasting from 15 Sep to 14 Oct 2011.

common complex terrain features and 2) forecasts of near-surface variables behave similarly over the DPG area and the Salt Lake Valley (as seen in Fig. 12 and described in section 4b), making it easier to examine the forecast errors associated with synoptic forcings.

A near-real-time forecasting system was built using version 3.3 of the ARW. From 15 September to 14 October 2011, near-real-time forecasts were performed 4 times daily (at 0000, 0600, 1200, and 1800 UTC) to produce a 48-h forecast each time. Four one-way nested domains, with 30-, 10-, 3.33-, and 1.11-km horizontal grid spacings, are used (Fig. 17). The innermost domain (1.11 km) focuses on the DPG area. Initial and boundary conditions are derived from the analyses and forecasts produced by the NAM forecast system at NCEP at 0000, 0600, 1200, and 1800 UTC. Over a 1-month period, a total of 120 forecasts are generated. As shown in Table 1, in addition to using physical schemes similar to those in the aforementioned case study, the Purdue Lin microphysics scheme (Lin et al. 1983, Chen and Sun 2002) and the YSU PBL scheme are used.

a. Overall evaluation

MAEs and BEs are employed to characterize the forecast errors in near-surface variables. A statistical calculation is done for each of the four initialization times. For example, all forecasts initialized at 0000 UTC during the month are averaged over all stations to calculate the MAEs and BEs as a function of forecast leading time.

Figures 18a–c show the MAEs calculated in the DPG area in the 10-, 3.33-, and 1.11-km domains for the forecasts initialized at 0000 UTC. Results confirm that simulations at high resolution do not always outperform those at coarser resolution. However, a clear diurnal pattern is found in the errors of all variables produced by all model domains. Specifically, the temperature error peaks twice per day, around 0300 mountain standard time (MST) and 1500 MST (corresponding to 1000 and 2200 UTC). There are also two error minima for temperature at around 0700 and 1900 MST (corresponding to 1400 and 0200 UTC). Wind speed and direction follow the same error trends, with a maximum in the early evening or before sunrise and a minimum in the afternoon.

Using results from the 1.11-km domain, the dependence of the surface forecasts on initialization time is examined. Figures 18d–f show that the error trends are independent of initialization time and forecast leading time and follow the same diurnal variation. However, compared with the forecasts initialized during the daytime (0000 and 1800 UTC), relatively large errors occur in the first 2–3 h in 2-m temperature for the forecasts initialized at night (0600 and 1200 UTC). The large errors in the nighttime-initiated forecasts could be caused by the erroneous soil temperature initialization in the NAM analysis. Apparently, the large errors associated with initial conditions in the nighttime-initiated forecasts do not persist beyond a few hours. This may indicate that a better local-scale initialization using data assimilation can help reduce forecast errors within the
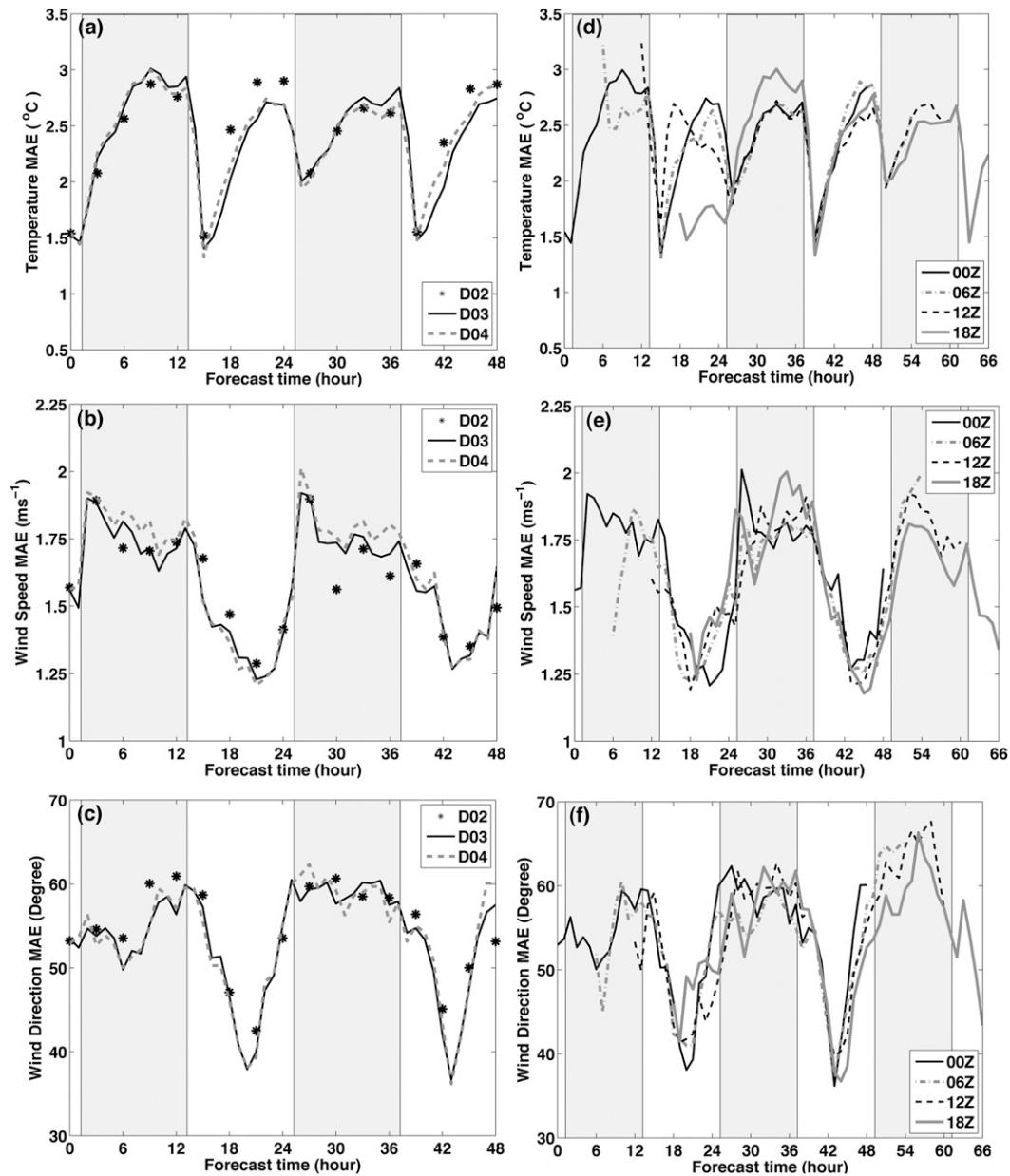
FIG. 18. MAEs of simulated near-surface variables for (a)–(c) different model domains and (d)–(f) various initialization times. Shown are the (a),(d) 2-m temperature; (b),(e)10-m wind speed; and (c),(f) 10-m wind direction. Results in (a)–(c) are from 0000 UTC forecasts and D02, D03, and D04 represent results from model domains at horizontal resolutions of 10-, 3.33-, and 1.11 km, respectively. Results in (d)–(f) are from the 1.11-km domain and various curves represent forecasts initialized at different times. The forecasting period for all forecasts is 48 h. The forecasts are output every 3 h in the 10-km domain (D02), and every hour in the 3.33- and 1.11-km domains.

first few hours, but its impact may subsequently vanish. Therefore, cycled data assimilation may be required to mitigate the problem.

Overall, compared with the previous study by Liu et al. (2008), the statistical errors are moderate for the whole period, with maximum errors of 3°C in 2-m temperature and 2 m s^{-1} in 10-m wind speed. Figure 19 further demonstrates the diurnal patterns of the bias errors in 2-m temperature over the whole month. Positive (warm) biases are found at night and negative (cold) biases are present during the daytime. No systematic biases are found in wind direction and speed (figures not shown).
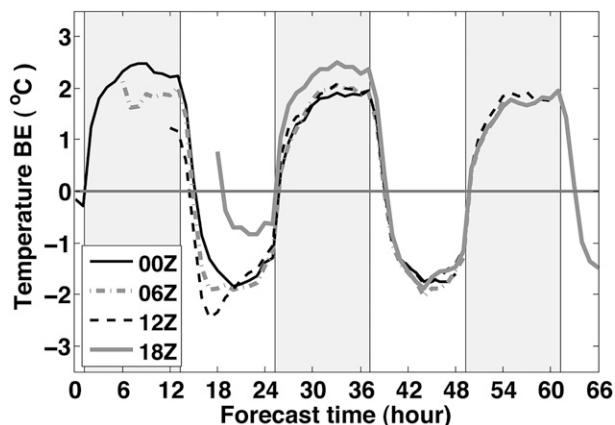
FIG. 19. Bias error of simulated 2-m temperature from the 1.11-km domain with various initialization times. The forecasting period for all forecasts is 48 h.

b. Strong versus weak synoptic forcing cases

In sections 3 and 4, it was concluded that the errors in 2-m temperature are significant under strong synoptic forcing. With 1-month forecasts that include both quiescent periods and strong synoptic forcing cases, we categorize the forecasts into strong and weak synoptic forcing cases by checking the weather maps at the surface and at the 700- and 850-hPa pressure levels. A strong forcing case is identified when a cold front, a closed low, a trough, a low pressure system, or large wind speeds (greater than 5 m s^{-1}) is present at the surface or on 700- or 850-hPa weather maps. In contrast, a high pressure system, a ridge, low wind speeds (less than 5 m s^{-1}) at 700 hPa or 850 hPa is identified as a weak forcing case. Overall, three weak synoptic forcing cases (i.e., 0000 UTC 21 September–1800 UTC 23 September, 0000 UTC 27 September–1800 UTC 29 September, and 0000 UTC 13 October–1800 UTC 15 October) and three strong forcing cases (i.e., 0000 UTC 16 September–1800 UTC 18 September, 0000 UTC 3 October–1800 UTC 5 October, and 0000 UTC 5 October–1800 UTC 7 October) are identified. Figure 20 compares the errors between the weak and strong forcing cases. For each case, four forecasts with different initial times are compared. It is apparent that diurnal patterns are present in the forecast errors for the weak forcing cases. The errors are independent of initialization time and forecast leading time (Figs. 20a,c,e). The strong forcing cases, however, show flow-dependent features (Figs. 20b,d,f). The forecast errors do not follow a diurnal pattern, implying that the errors become more closely related to the influence of the weather systems. In addition, the magnitude of the errors is generally greater in the strong forcing than in the weak forcing cases.

7. Summary and concluding remarks

In this study, the performance of version 3.3 of the ARW in predicting near-surface atmospheric temperature and wind conditions under various terrain and weather regimes is evaluated. Three individual events under strong synoptic forcing, namely, a frontal system, a low-level jet, and a persistent cold air pool, are first verified against observations over both flat and complex terrain. It is found that the ARW is able to produce reasonable simulations of weather phenomena. Verification of near-surface conditions (i.e., 2-m temperature and 10-m wind) indicates the complexity in forecasting these surface variables. For the frontal case and low-level jet case over the central United States, the model terrain matches the actual terrain and thus mitigates representative errors. The forecasts of surface variables generally agree well with the observations. However, errors still occur, depending on the model's ability in forecasting the structures in the lower-atmospheric boundary layer. For the inversion case over the Salt Lake Valley, different error characteristics are found over the mountain and valley stations. Terrain mismatch and the ARW's limited ability to simulate near-surface atmospheric conditions make the forecasting errors even more complicated.

Overall, forecast errors in near-surface atmospheric variables show flow-dependent features in all three of the individual cases when strong synoptic forcings are present. To better understand the characteristics of flow-dependent errors in complex terrain as they relate to near-surface forecasting, 1-month forecasts (from 15 September to 14 October 2011) are conducted over complex terrain in the western United States. It is found that the forecast errors of surface variables depend to a large degree on the diurnal cycle of the surface variables themselves, especially when the synoptic forcing is weak. The forecast errors for 2-m temperature reach two daily maxima at 0300 and 1500 local time, and two daily minima at 0700 and 1900 local time. Errors in wind speed and direction follow the same trends, with a maximum at night and a minimum in afternoon. Forecast errors follow the same trends regardless of the initialization time, showing that forecast errors are independent of the initialization time and forecast leading time. Further analyses reveal positive (warm) temperature biases at night and negative (cold) biases during the daytime. In contrast to the 2-m temperature, wind direction and speed have no systematic biases from a long-term perspective. Under strong synoptic forcing, diurnal patterns in forecast errors are broken, while flow-dependent errors are clearly shown.
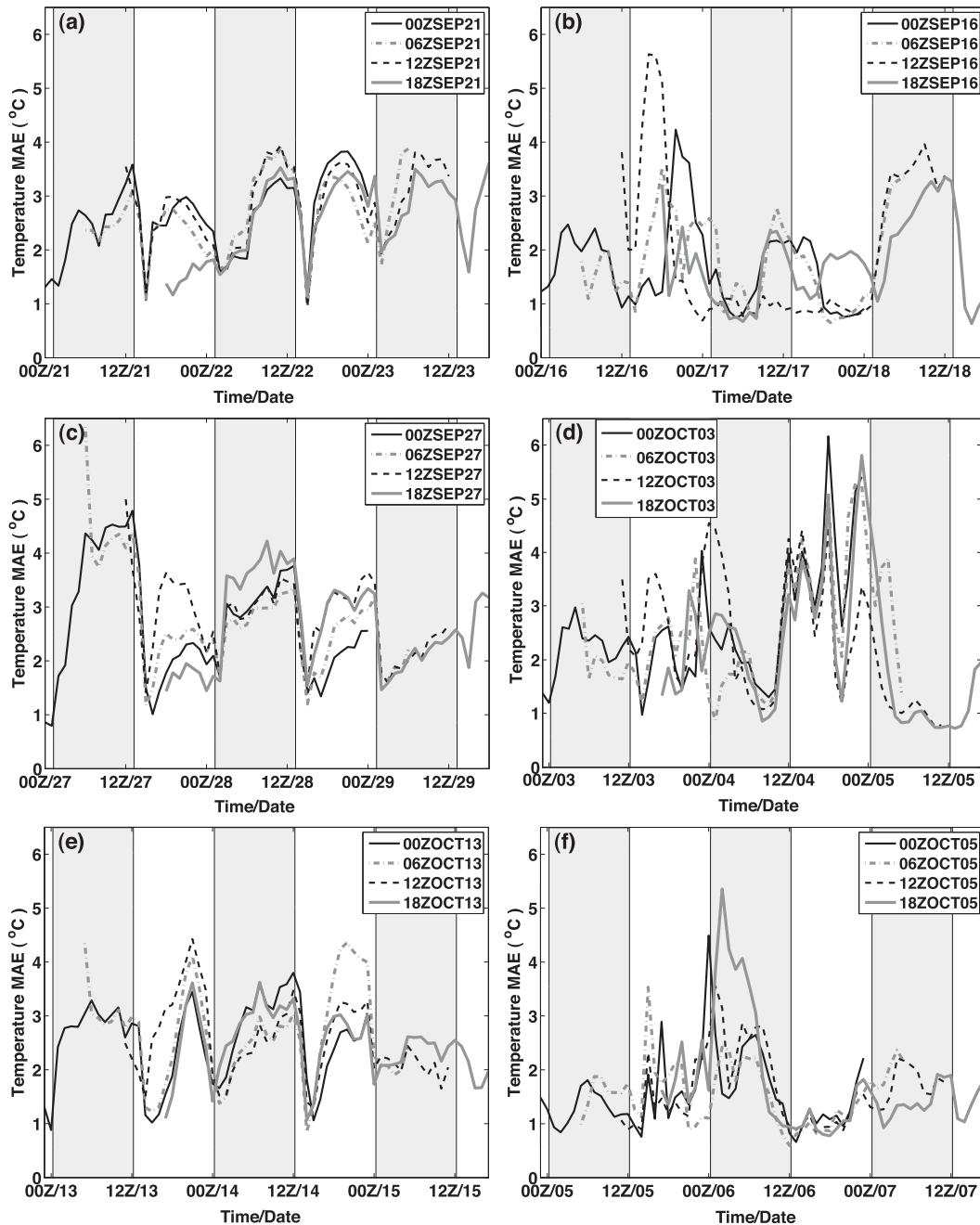
FIG. 20. MAEs of simulated 2-m temperature from the 1.11-km domain for various cases: (a) 0000 UTC 21 Sep–1800 UTC 23 Sep, (b) 0000 UTC 16 Sep–1800 UTC 18 Sep, (c) 0000 UTC 27 Sep–1800 UTC 29 Sep, (d) 0000 UTC 3 Oct–1800 UTC 5 Oct, (e) 0000 UTC 13 Oct–1800 UTC 15 Oct, and (f) 0000 UTC 5 Oct–1800 UTC 7 Oct. Four forecasts with different initial times are available for each case. The forecasting period for all forecasts is 48 h. Specifically, left panels [(a), (c), and (e)] represent weak synoptic forcing cases, and right panels [(b), (d), and (f)] denote strong synoptic forcing cases.

Finally, it is apparent that simulations at finer resolutions do not outperform those at coarser resolutions in most cases. This was explained well in Hart et al. (2005), who found that the inability of the numerical model to depict near-surface structures (such as strong temperature inversions) results in worse forecasts, even with better terrain representation. Meanwhile, increasing the model's vertical resolution

does not help the predictability of near-surface variation, although it improves the forecasts of mesoscale weather phenomena. Numerical forecasts of near-surface atmospheric conditions are also sensitive to the PBL scheme in the ARW model, but there is no single PBL scheme that performs better than the others. These factors illustrate the complexity and challenges involved in near-surface simulation over complex terrain. Future work should emphasize investigating the decoupling between near-surface variables and the atmospheric boundary layer and its impact on the predictability of near-surface variables. The sensitivity of numerical predictions of near-surface variables to terrain representation, land surface parameters, and model errors will also need to be examined.

REFERENCES

Brown, A. R., and N. Wood, 2003: Properties and parameterization of the stable boundary layer over moderate topography. *J. Atmos. Sci.,* **60,** 2797–2808.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.,* **129,** 569–585.

Chen, S.-H., and W.-Y. Sun, 2002: A one-dimensional time dependence cloud model. *J. Meteor. Soc. Japan,* **80,** 99–118.

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.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.,* **46,** 3077–3107.

Hanna, S. R., and R. Yang, 2001: Evaluations of mesoscale models' simulations of near-surface winds, temperature gradients, and mixing depths. *J. Appl. Meteor.,* **40,** 1095–1104.

Hart, K. A., W. J. Steenburgh, and D. J. Onton, 2005: Model forecast improvements with decreased horizontal grid spacing over finescale intermountain orography during the 2002 Olympic Winter Games. *Wea. Forecasting,* **20,** 558–576.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.,* **124,** 2322–2339.

——, and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.,* **42,** 129–151.

——, S.-W. Kim, J. Dudhia, M.-S. Koo, and K.-H. Seol, 2008: Stable boundary layer mixing in a vertical diffusion packagee. *Ninth Annual WRF User's Workshop,* Boulder, CO, NCAR, 3.3. [Available online at http://www.mmm.ucar.edu/wrf/users/workshops/WS2008/presentations/3-3.pdf.]

Horel, J., and Coauthors, 2002: Mesowest: Cooperative mesonets in the western United States. *Bull. Amer. Meteor. Soc.,* **83,** 211–226.

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

Kain, J. S., and J. M. Fritsch, 1993: The role of the convective "trigger function" in numerical prediction of mesoscale convective systems. *Meteor. Atmos. Phys.,* **49,** 93–106.

Lee, T. J., R. A. Pielke, R. C. Kessler, and J. Weaver, 1989: Influence of cold pools downstream of mountain barriers on downslope winds and flushing. *Mon. Wea. Rev.,* **117,** 2041–2058.

Leung, L. R., and Y. Qian, 2003: The sensitivity of precipitation and snowpack simulations to model resolution via nesting in regions of complex terrain. *J. Hydrometeor.,* **4,** 1025–1043.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.,* **22,** 1065–1092.

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

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.,* **83,** 407–430.

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

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.,* **87,** 343–360.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.,* **102** (D14), 16 663–16 682.

Nakanishi, M., and H. Niino, 2004: An improved Mellor–Yamada level-3 model with condensation physics: Its design and verification. *Bound.-Layer Meteor.,* **112,** 1–31.

Pleim, J. E., 2007: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *J. Appl. Meteor. Climatol.,* **46,** 1383–1395.

Rife, D. L., and C. A. Davis, 2005: Verification of temporal variations in mesoscale numerical wind forecasts. *Mon. Wea. Rev.,* **133,** 3368–3381.

Shafran, P. C., N. L. Seaman, and G. A. Gayno, 2000: Evaluation of numerical predictions of boundary layer structure during the Lake Michigan Ozone Study (LMOS). *J. Appl. Meteor.,* **39,** 412–426.

Shin, H. H., and S.-Y. Hong, 2011: Intercomparison of planetary boundary-layer parameterizations in the WRF model for a single day from CASES-99. *Bound.-Layer Meteor.,* **139,** 261–281.

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

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology.* Springer, 666 pp.

Sukoriansky, S., B. Galperin, and V. Perov, 2005: Application of a new spectral theory of stably stratified turbulence to atmospheric boundary layers over sea ice. *Bound.-Layer Meteor.,* **117,** 231–257.

Teixeira, J., and Coauthors, 2008: Parameterization of the atmospheric boundary layer: A view from just above the inversion. *Bull. Amer. Meteor. Soc.,* **89,** 453–458.

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

Zhang, D.-L., and W. Zheng, 2004: Diurnal cycles of surface winds and temperatures as simulated by five boundary-layer parameterizations. *J. Appl. Meteor.,* **43,** 157–169.

Zhong, S., and J. Fast, 2003: An evaluation of the MM5, RAMS, and Meso-Eta models at subkilometer resolution using VTMX field campaign data in the Salt Lake Valley. *Mon. Wea. Rev.,* **131,** 1301–1322.