Performance Assessment of New Land Surface and Planetary Boundary Layer Physics in the WRF-ARW

ROBERT C. GILLIAM AND JONATHAN E. PLEIM

Atmospheric Modeling and Analysis Division, National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina

(Manuscript received 23 September 2008, in final form 29 October 2009)

ABSTRACT

The Pleim–Xiu land surface model, Pleim surface layer scheme, and Asymmetric Convective Model (version 2) are now options in version 3.0 of the Weather Research and Forecasting model (WRF) Advanced Research WRF (ARW) core. These physics parameterizations were developed for the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) and have been used extensively by the air quality modeling community, so there was a need based on several factors to extend these parameterizations to WRF. Simulations executed with the new WRF physics are compared with simulations produced with the MM5 and another WRF configuration with a focus on the replication of near-surface meteorological conditions and key planetary boundary layer features. The new physics in WRF is recommended for retrospective simulations, in particular, those used to drive air quality simulations. In the summer, the error of all variables analyzed was slightly lower across the domain in the WRF simulation that used the new physics than in the similar MM5 configuration. This simulation had an even lower error than the other more common WRF configuration. For the cold season case, the model simulation was not as accurate as the other simulations overall, but did well in terms of lower 2-m temperature error in the western part of the model domain (plains and Rocky Mountains) and most of the Northeast. Both MM5 and the other WRF configuration had lower errors across much of the southern and eastern United States in the winter. The 2-m water vapor mixing ratio and 10-m wind were generally well simulated by the new physics suite in WRF when contrasted with the other simulations and modeling studies. Simulated planetary boundary layer features were compared with both wind profiler and aircraft observations, and the new WRF physics results in a more precise wind and temperature structure not only in the stable boundary layer, but also within most of the convective boundary layer. These results suggest that the WRF performance is now at or above the level of MM5. It is thus recommended to drive future air quality applications.

1. Introduction

Mesoscale models require land surface, surface layer, and planetary boundary layer (PBL) parameterizations to represent the transfer of heat, moisture, and momentum between the surface and atmosphere. A new land surface and PBL physical parameterization have been implemented in version 3.0 of the Weather Research and Forecasting model (WRF), Advanced Research WRF (ARW) core (Skamarock et al. 2008). The Pleim–Xiu land surface model (PX LSM; Xiu and Pleim 2001; Pleim and Xiu 2003), Pleim surface layer scheme (Pleim 2006), and Asymmetric Convective Model, version 2 (ACM2), (Pleim 2007a,b) for the planetary boundary layer have been used extensively as physics options in the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5; Grell et al. 1995). Many users of the Community Multiscale Air Quality (CMAQ) modeling system (Byun and Schere 2006) have traditionally employed MM5 with this physics configuration for air quality modeling applications. The ACM2 PBL is preferred because it allows for consistent turbulent mixing in the meteorological (WRF) and air quality model (CMAQ), which also uses the ACM2 for subgrid vertical transport of pollutants. Additionally, the soil moisture and soil temperature nudging (Pleim and Xiu 2003; Pleim and Gilliam 2009) algorithms in the PX LSM, along with four-dimensional data assimilation (FDDA) using model analyses or reanalyses (Stauffer and Seaman 1990, 1994;

Corresponding author address: Robert C. Gilliam, U.S. EPA, 109 TW Alexander Dr., Research Triangle Park, NC 27709.
E-mail: gilliam.robert@epa.gov

DOI: 10.1175/2009JAMC2126.1

Stauffer et al. 1991), result in high-quality meteorological fields (Gilliam et al. 2006) for retrospective applications.

A set of month-long simulations that employs the PX LSM, surface layer, and ACM2 PBL options of WRF-ARW, version 3.0, (collectively referred to as WRF PXACM herein) has been evaluated. This paper describes the performance of the WRF PXACM relative to observations, a similar MM5 simulation (MM5 PXACM), and one of the most commonly used WRF configurations. The goal is not necessarily to rank the models or physics options according to performance, since some model configurations are suited for certain applications or regions and may require special configuration methods. The main objective of the assessment is to ensure that the retrospective WRF PXACM simulation is of similar skill as MM5 and comparable to another common WRF solution, all of which employ FDDA. The evaluation mainly focuses on the ability to simulate 2-m temperature with a secondary focus on the 2-m water vapor mixing ratio and 10-m wind. Also examined is the error of simulated potential temperature and wind in the PBL using wind profiler and aircraft observations. These comparisons will provide an assessment of the WRF PXACM's ability to represent key characteristics of the atmosphere that are known to have an impact on air quality.

2. Methodology

a. Overview of PX LSM, surface layer scheme, and ACM2 PBL

The PX LSM coupled with the ACM2 PBL and Pleim surface layer scheme, historically available only in MM5, have been found to be well suited for extended (weeks, months, or even years) retrospective simulations where the indirect soil moisture and temperature nudging scheme lead to more accurate near-surface meteorology (Xiu and Pleim 2001; Pleim and Xiu 2003; Gilliam et al. 2006). The PX LSM simulates the evolution of soil moisture and temperature in two layers (0–1 and 1–100 cm), canopy moisture, and aerodynamic and stomatal resistance. There are three pathways for evaporation in the PX LSM: soil surface, canopy, and evapotranspiration. Gridcell representative values of surface and vegetative parameters, such as roughness length, leaf area index, vegetation coverage, albedo, and minimum stomatal resistance, are computed by the PX LSM using the fractional land use data. Similarly, soil parameters are computed from fractional soil texture data. The Asymmetric Convective Model, version 2, is a hybrid of the original nonlocal closure model (Pleim and Chang 1992) and a local closure eddy diffusion scheme (Pleim 2007a,b). These physics options from the MM5 have been implemented as separate LSM, surface layer, and PBL physics parameterizations in WRF version 3.0. Each of these new physics options has been successfully tested with the other WRF–ARW physics options for compatibility, or what is commonly referred to as ''plug-and-play'' capability. *Note that the PX LSM has not been tested in a forecast mode or with soil nudging deactivated.*

The PX LSM currently does not contain a process to account for the accumulation, sublimation, or melting of snow. Rather, it uses 3-hourly gridded snow-water equivalent from the National Centers for Environmental Prediction (NCEP) North American Mesoscale model (NAM) analysis for snow cover. It has been noted in past evaluations (i.e., Gilliam et al. 2004, 2006) that the PX LSM does not perform as well as some other land surface models over snow. Several improvements for snow cover were made in the PX LSM including an updated volumetric heat capacity for snow and a fractional snow coverage that is a function of land use and snow depth, following the method used in the Noah land surface model (Ek et al. 2003). The fractional snow coverage is then used by the PX LSM to compute a weighted surface heat capacity and albedo, which have a significant impact on ground heating and cooling rates.

b. Model configuration

The simulations were executed using 5.5-day overlapping run segments on an eastern U.S. grid (Fig. 1) with 12-km horizontal spacing and 34 vertical levels, extending from the surface to the 50-hPA level. All WRF–ARW, version 3.0, and MM5, version 3.7, simulations were configured on a similar (MM5 has an Arakawa B grid while WRF is on an Arakawa C grid) horizontal grid and the same vertical sigma levels. Simulations were completed for cold and warm season cases (January and August 2006). All model runs utilized the NAM analysis (0000, 0600, 1200, and 1800 UTC) and the 3-h NAM (WRF–NMM) forecast (0300, 0900, 1500, and 2100 UTC) for the first guess fields for initial and boundary conditions as well as the four-dimensional data assimilation. To clarify, surface and upper-air observations (12-hourly radiosonde) were reintroduced once the NAM analysis was interpolated to the model grid to generate an objective reanalysis with a tighter fit to the observations. For the MM5 simulations the objective reanalysis tool RAWINS (NCAR 2009a) was used, and for WRF a newly developed tool named Obsgrid (NCAR 2009b) was employed. These tools provide a three-dimensional temperature, water vapor mixing ratio, and wind analyses for the FDDA, and a separate 2-m temperature and 2-m mixing ratio analysis that is used by the PX LSM's indirect soil nudging algorithm (Pleim and Gilliam 2009; Pleim and Xiu 2003).

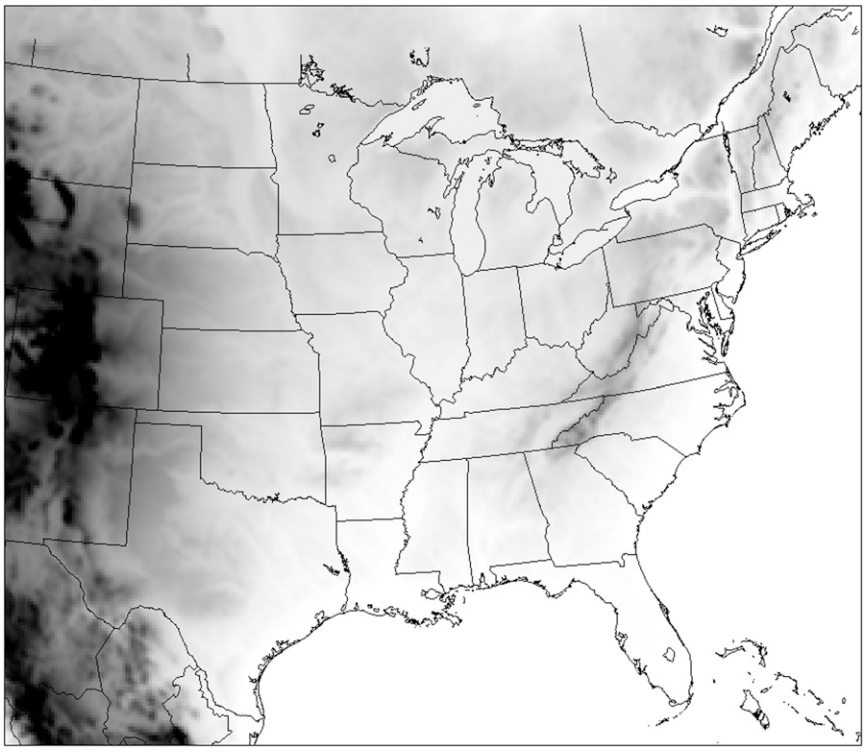

FIG. 1. The eastern U.S. WRF and MM5 model domain. Shading represents terrain height.

The FDDA was configured similar to Stauffer et al. (1991) and Otte (2008a), where no temperature or moisture nudging was done within the ACM2 diagnosed PBL (bulk Richardson number approach; Pleim 2007a), but the nudging of wind was applied at all levels of the model. Also, following Otte (2008a), who noted dramatic improvements in the meteorological and air quality simulation that employed FDDA, the nudging strength applied was greater for temperature and wind (3.0×10^{-4} s^{-1}) than moisture (1.0×10^{-5} s^{-1}). It should be noted that although the surface observations used to evaluate the model are used by RAWINS and Obsgrid, which produce the PX LSM's soil nudging analyses, these observations are not used to directly nudge the 2-m temperature and 2-m water vapor mixing ratio directly.

Each simulation that used the PX LSM (WRF and MM5) was spun up for 10 days prior to the first day of January and August 2006. The spinup period allows the indirect soil nudging algorithms in the PX LSM to adjust soil moisture and temperature, which minimizes the bias of 2-m temperature and 2-m water vapor mixing ratio. The initial deep soil temperature at the start of the spinup period (21 December 2005 and 20 July 2006) was set to the 5.5-day run-segment-averaged 2-m temperature of the NAM data using the Intermediate Processor for Pleim–Xiu for WRF (IPXWRF) utility (available for download at http://www.wrf-model.org). After the first run segment, IPXWRF was used to pass the soil moisture and temperature in both soil layers from one run segment to the next. Each run segment overlapped the end of the previous one by 12 h to allow the spinup of the atmospheric moisture fields (cloud water, ice, etc.).

The simulations will be referred to by the run identification in Table 1 throughout the discussion of the results. Table 1 also provides the physics configuration of each simulation. Note that the physics options, other than PBL, LSM, and surface layer schemes, are essentially the same. All simulations used the Rapid Radiative Transfer Model (RRTM) longwave radiation scheme (Mlawer et al. 1997), Dudhia (1989) shortwave radiation scheme, and Kain–Fritsch 2 (Kain 2004) scheme for convective precipitation. The microphysics schemes available in MM5 and WRF are different; the Thompson scheme (Thompson et al. 2004) was used by both WRF simulations and Reisner 2 (Reisner et al. 1998) by the MM5 simulation.

The other WRF simulation employed the Noah LSM (Ek et al. 2003), the Yonsei University (YSU) PBL model (Noh et al. 2003), and the Monin–Obukhov (M–O) surface layer scheme ported from MM5 (WRF NoahYSU; Dyer and Hicks 1970). This configuration represents one of the most common land surface, surface-layer, and PBL configurations used by the WRF community. It should be stressed that the Noah LSM does not have an internal mechanism like the soil moisture and temperature

TABLE 1. Configuration and physics options for each model simulation. Also provided are the identification tags used throughout the paper for each model simulation.

	Model simulation		
Configuration	WRF PXACM	MM5 PXACM	WRF NoahYSU
Land surface	PX	PX	Noah
PBL	ACM2	ACM2	YSU
Surface layer	Pleim	Pleim	M-O
Microphysics	Thompson	Reisner 2	Thompson
Convective	Kain–Fritsch 2	Kain–Fritsch 2	Kain–Fritsch 2
Shortwave	Dudhia	Dudhia	Dudhia
Longwave	RRTM	RRTM	RRTM
FDDA driver	Obsgrid	RAWINS	Obsgrid
PX soil nudging	Obsgrid	RAWINS	—

nudging of the PX LSM, so a run segment of 5.5 days may not be the best run strategy; the accuracy of these simulations might have been improved with a shorter simulation length, but such sensitivity tests are outside the scope of this paper.

c. *Model assessment techniques*

The evaluation of the three simulations was primarily done using the Atmospheric Model Evaluation Tool (AMET; Gilliam et al. 2005), which pairs surface observations from the Meteorological Assimilation Data Ingest System (MADIS) database with the corresponding model simulations in space (bilinear interpolation of model to observation location) and time (hourly). The highest level of quality control was applied using the MADIS surface observation extraction program sfcdump.exe. For this study, the main focus of the near-surface evaluation is on model–observation comparisons of 2-m temperature with a less comprehensive examination of 2-m mixing ratio and 10-m wind. Another main focus is the wind and temperature structure of the PBL using wind profiler and aircraft observations.

The model performance statistics, computed using MADIS surface observations, are assessed collectively on a monthly time scale and parsed by time of day. Of most interest is the model error expressed as mean absolute error (MAE) and root-mean-square error (RMSE) using the standard calculations in Wilks (1995), although we also present the modeled and observed mean diurnal temperature, which provides information on model bias. These model statistics are also computed at each observation site and presented spatially. The RMSE of the 2-m temperature analyses are also presented, in part to explain differences in model performance, since these RAWINS–Obsgrid reanalyses are used to drive the PX LSM's soil nudging.

An examination of the simulated PBL structure is conducted. This assessment employs an AMET utility that matches vertical profiles of potential temperature and wind from wind profilers (Barth et al. 1994) and aircraft observations (Moninger et al. 2003; Daniels et al. 2004) with the model simulations. Since the vertical levels of the observations vary among sites and platforms, the observed profiles are interpolated to the physical height above ground of the model sigma levels. For the first assessment, the evaluation tool was used to pair the observations with the model for a group of 19 wind profilers in the central United States that is part of the National Oceanic and Atmospheric Administration (NOAA) wind profiler network. A second assessment is provided using aircraft data and corresponding model values at 119 airports in the eastern United States. The mean absolute error of the profiles is examined to determine how well the new model configuration is replicating the boundary layer structure, and how these values compare with the other model simulations. This will provide a sense of the uncertainty in the simulated boundary layer structure using a rather large dataset that includes hourly data at multiple sites over a period of a month.

3. Assessment of model performance

a. *Surface-based meteorological conditions*

A convenient and useful method to assess the general performance of a model is to use all observations to compute the domainwide RMSE. In Table 2, the RMSE values are provided for the near-surface variables including 2-m temperature, 2-m mixing ratio, and 10-m wind speed and direction. For reference, the RMSE of the 2-m temperature analyses are also provided, which again are used by the PX LSM's indirect soil nudging scheme and can be thought of as the lower bound of error one can expect to achieve. In addition, a two-sample t test with a 95% confidence interval was applied to each collection of model and observation samples used to compute the errors in Table 2, and the model solutions were significantly different from the observations. Additionally, tests of the WRF PXACM versus the other models' 2-m temperature simulations revealed that in all cases the simulations were significantly different when all data in the model domain were considered.

For 2-m temperature in January 2006 (Table 2), the WRF PXACM has a slightly smaller domainwide RMSE (2.48 K) than the MM5 PXACM (2.52 K), but a larger RMSE than the WRF NoahYSU (2.33 K). The Obsgrid- and RAWINS-derived analyses have RMSE values in January of 1.29 and 1.47 K, respectively. The lower RMSE achieved through the 2-m (temperature and mixing ratio) reanalysis by these utilities is a key factor in improved

TABLE 2. Summary of surface-based model performance statistics for each simulation. Also provided is the RMSE (2-m temperature only) of the analysis dataset that was used for the indirect soil moisture and temperature nudging of the PX LSM.

RMSE	WRF PXACM	MM5 PXACM	WRF NoahYSU	Obsgrid analysis	RAWINS analysis
2-m temperature (K)					
January	2.48	2.52	2.33	1.29	1.47
August	1.94	2.00	2.31	1.22	1.31
2-m mixing ratio (g kg^{-1})					
January	0.92	0.84	0.78		
August	1.86	1.92	2.11		
10-m wind speed (m s^{-1})					
January	1.64	1.79	1.78		
August	1.47	1.49	1.60		
10-m wind direction (°)					
January (MAE)	21	25	23		
August (MAE)	30	33	32		

model simulations since these analyses are used to drive the soil nudging that indirectly impacts the soil moisture and temperature, surface flux partitioning, and, thus, near-surface meteorology. To support this statement, the RMSE of the NAM analysis prior to Obsgrid for this same period was 2.39 K. The resulting WRF simulation that used the same PXACM configuration had a 2.79-K RMSE of 2-m temperature versus 2.48 K when Obsgrid is used. This is a clear indication that Obsgrid is a valuable tool for lowering the RMSE of the analysis, and that the indirect soil nudging algorithms of the PX LSM are effective in reducing model error.

The 2-m temperature RMSE in August (Table 2) associated with the WRF PXACM simulation is the lowest of all models. WRF PXACM (1.94 K) has a slightly lower RMSE than the MM5 PXACM (2.00 K) and much lower error than the WRF NoahYSU (2.31 K). The Obsgrid analyses, for instance in January, have a lower RMSE (1.22 K) or better fit than RAWINS analyses when compared to the surface observations (1.31 K). The RMSE of the NAM analysis was 1.89 K and the resulting model simulation, which is not shown here, had a 2-m temperature RMSE of 2.18 K. This is more supporting evidence that by lowering the RMSE of the 2-m temperature analysis used by the PX LSM, the model simulations improve.

A review of previous modeling studies was done to place these RMSE values in perspective. And the values revealed here are at the lower range of what has been published in the literature. Deng et al. (2004), for example, found an RMSE of 2.52 K for a similar model configuration (eastern U.S. MM5 using FDDA) whereas the RMSE of the WRF PXACM presented here was less than 2.0 K. While the RMSE was the only error metric presented in Table 2, the MAE values were also computed. Our WRF PXACM simulation had an MAE of 1.89 K in January and 1.46 K in August, which compares very favorably to past studies that present MAE metrics. Emery (2001) examined annual simulations that used FDDA and suggested that MAE of 2-m temperature that were less than 2.0 K were reasonable. Gilliam et al. (2006) found MAE values of 2.38 K in the winter and 1.67 K in the summer for the same domain but for an annual 2001 simulation. Baker (2004) found that MAE, computed for an entire eastern U.S. domain on a daily interval, ranged from 1.5 K during warmer months to 3.0 K in the winter. McNally (2009) presented a comprehensive matrix of MAE values for several annual MM5 simulations (2004–06) with an almost identical domain and model configuration. The MAE of temperature ranged from around 2.00–2.25 K in January and 1.70–2.00 K in August for the eastern United States, and around 2.5 K in both January and August for the western United States. Otte (2008a,b) found an MAE of temperature of 1.7 K for a July case using MM5 that employed a similar FDDA strategy.

The mean and RMSE of 2-m temperature as a function of time of day are presented in Fig. 2. Included in the figure is not only the model performance in terms of RMSE, but also the RMSE of the analyses. For January, the RAWINS and Obsgrid reanalyses have a similar RMSE, but the Obsgrid has a slightly better fit to the observations. Also note, since the reanalyses are available on a 3-hourly interval, linear interpolation is done within the PX LSM to provide an analysis value at each time step. This interpolation leads to noticeable errors. As a result, the RMSE is lower at 0000, 0300, 0600, 0900, 1200, 1500, 1800, and 2100 UTC when interpolation is not done. Focusing on the model simulations, the January WRF PXACM has a lower RMSE of 2-m temperature than the MM5 PXACM at night, but larger error during the middle part of the day. The WRF NoahYSU has a much lower RMSE than the other two simulations at night and during the middle part of the day, but similar

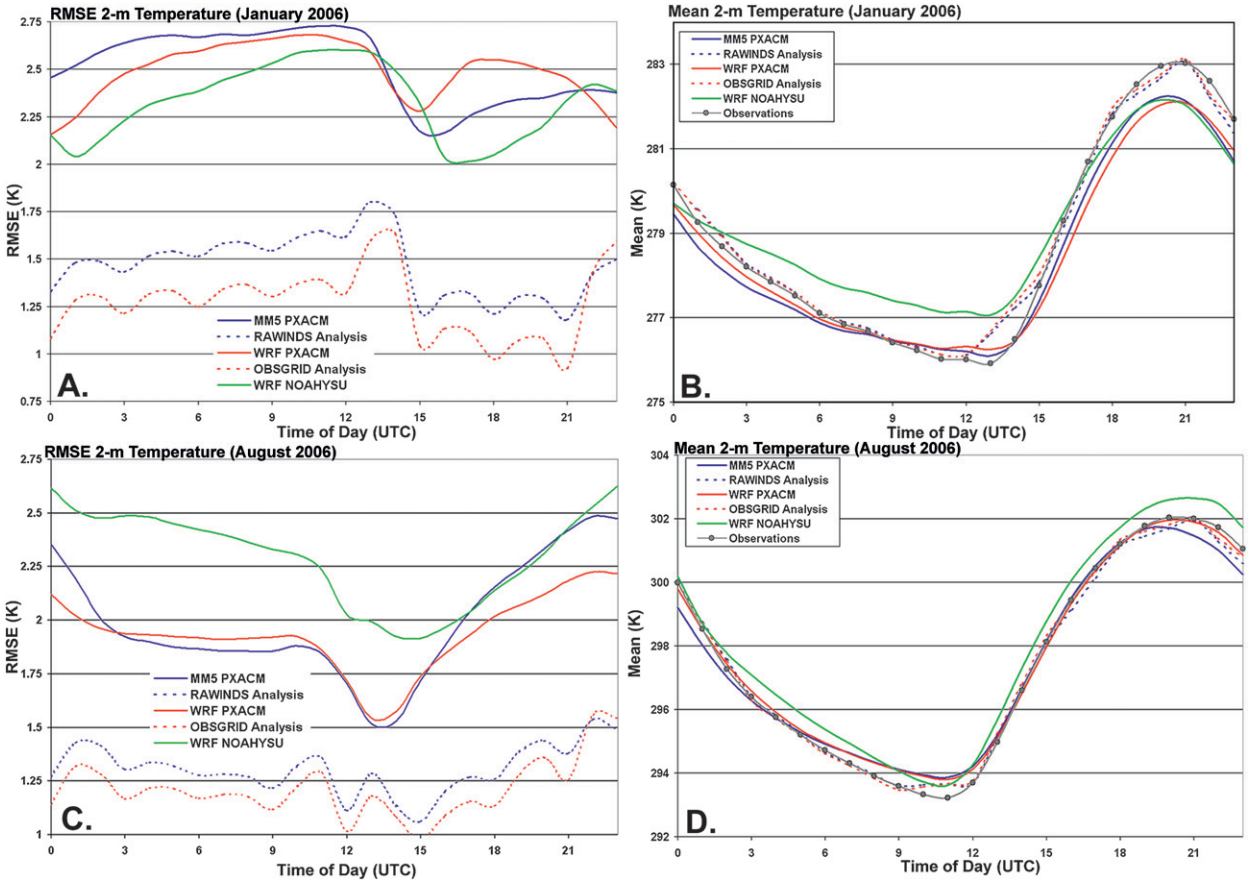

FIG. 2. Diurnal (UTC) (left) RMSE and (right) mean 2-m temperature (K) for (top) January and (bottom) August 2006. Also included is the RMSE of the RAWINS and Obsgrid analyses. The statistics include data from all observations sites in the model domain.

error during the PBL transition periods in the morning and evening.

Mean 2-m temperature is also presented for January 2006 in Fig. 2b. It is interesting that the WRF NoahYSU has a lower RMSE than the other simulations, but the diurnal bias is relatively large. This implies a systematic widespread warm bias at night and cold bias during the day. Spatial bias plots from the WRF NoahYSU were examined (not shown) revealing that the model did have a widespread warm bias (most of the domain excluding the Northeast) at night of 1–3 K with higher values (3–4 K) across the central plains. During the day the WRF NoahYSU had only a small warm bias over the plains, but a large cold bias of 1–3 K over the Northeast and little bias elsewhere. The mean temperature of the WRF PXACM and MM5 PXACM follows the observations closely at night, but both have a large cold bias during the middle of the day. Both the RAWINS and Obsgrid analyses have very little diurnal bias as expected.

For August 2006, the diurnal RMSE in Fig. 2c indicates that the Obsgrid analysis fits the observations slightly better than RAWINS. The WRF PXACM has a lower RMSE, from early afternoon (~1800 UTC) through early evening, than the MM5 PXACM and WRF NoahYSU, but a slightly higher error than the MM5 PXACM during the early morning hours from 0300 to 0900 UTC. Both the WRF and MM5 PXACM simulations have a much lower RMSE than the WRF NoahYSU at night.

The mean diurnal 2-m temperature of each simulation and analysis dataset for August 2006 is presented in Fig. 2d. The WRF PXACM is nearly identical to the mean observed temperature from 1200 to 0000 UTC while the MM5 PXACM is colder and the WRF NoahYSU is warmer. There is less difference of temperature when compared with the observations at night; all are slightly warmer than observations of temperature.

For January 2006, it was shown that the overall RMSE of temperature of the WRF PXACM (Table 2) is greater than the WRF NoahYSU, but lower than the MM5 PXACM. The spatial distribution of this RMSE provides more insight into the domainwide numbers. Figure 3a indicates a low RMSE in the Great Lakes states (<1.5 K), along the New England coast (<2.0 K), and over water

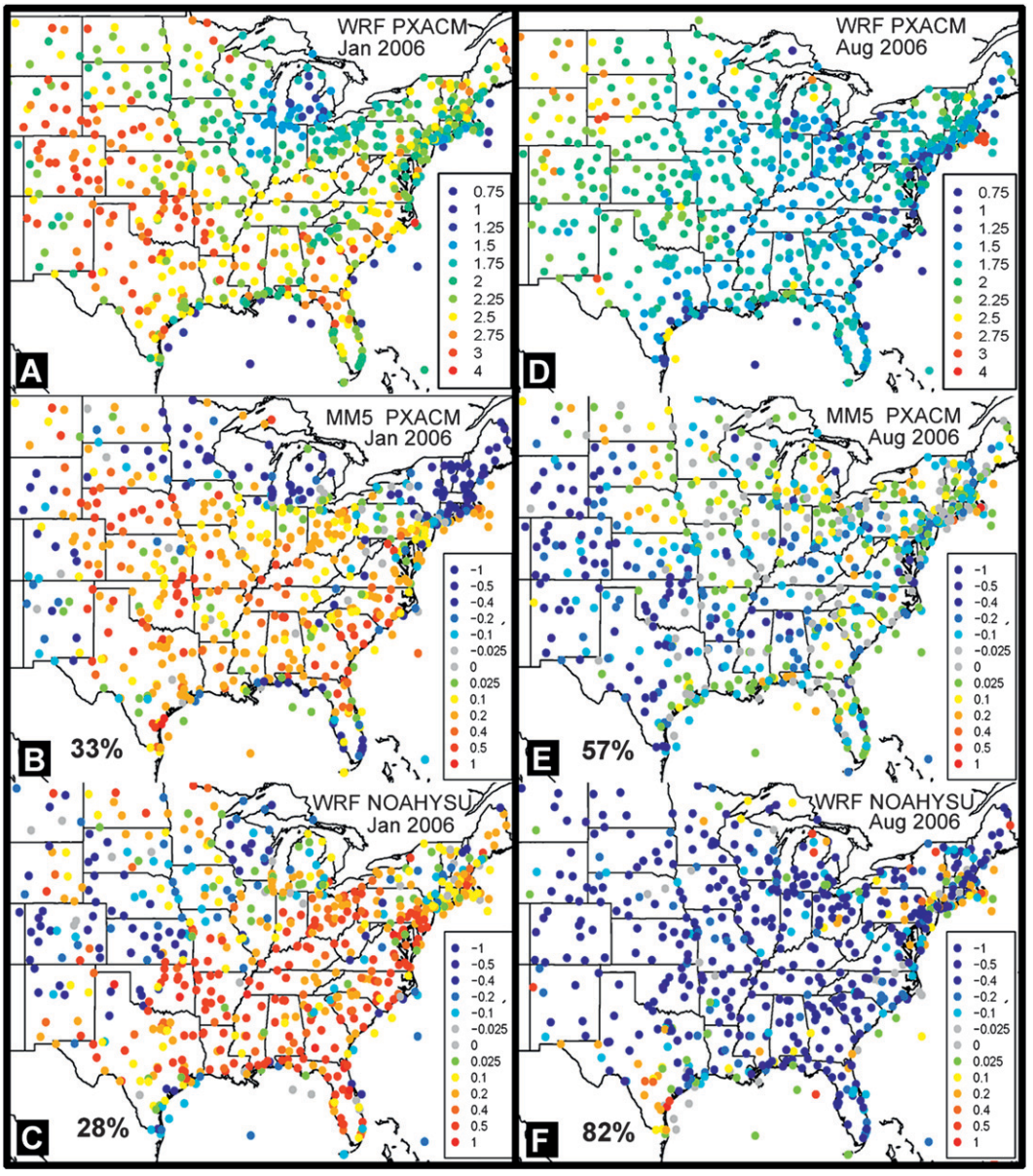

FIG. 3. Spatially distributed RMSE of simulated 2-m temperature (K) for the (a),(d) WRF PXACM for each month. The difference between the RMSE of the WRF PXACM and (b),(e) MM5 PXACM and the (c),(f) WRF NoahYSU. Cold (warm) colors or negative (positive) values indicate WRF PXACM has a lower (larger) RMSE.

bodies (~1.0 K). The model simulates temperature more poorly across much of the western part of the model domain and many areas of the southern United States (RMSE > 2.5 K). The larger errors across the western part of the domain are common because of the complex landscape. McNally (2009) documented similarly configured MM5 simulations over the western United States and found MAE that averaged from 2.8 to 3.0 K for the winters of 2004–06.

Figures 3b and 3c provide the RMSE differences to identify which WRF PXACM sites have lower or higher errors than the other simulations. In this case negative values (cool colors) indicate that WRF PXACM has lower RMSE, and positive values indicate a higher RMSE than the compared simulation. Although the domainwide RMSE of the WRF PXACM is lower than the MM5 configuration, the MM5 has a lower error at 66% of sites, which leads to the conclusion that the reduction of RMSE of the WRF PXACM over MM5 PXACM at 33% of sites is much greater than the increase in error at the other 67% of sites. A histogram of these differences was examined (not shown) and at most of the sites where

the WRF PXACM had lower RMSE, the error was between 0.5 and 1.0 K less than the MM5 PXACM. At a majority of sites where MM5 had lower RMSE, the difference in error was around 0.25 K or less. The WRF NoahYSU has a lower-temperature RMSE than the WRF PXACM across most of the eastern United States, while the WRF PXACM has much lower error relative to the WRF NoahYSU across much of the upper plains, Rocky Mountains, and Appalachian Mountains.

Figure 3d provides the RMSE of 2-m temperature associated with the WRF PXACM for the month of August 2006. As discussed earlier, past studies show that MAE values below 2.0 K indicate reasonable model performance. The WRF PXACM has very low RMSE over water (~1.0 K), which is expected, as the 2-m temperature is mainly controlled by the accurate sea surface temperatures that are satellite derived. Most areas east of the Mississippi River and eastern Texas are well simulated by the WRF PXACM where RMSE ranges from 1.25 to 2.00 K. The RMSE of the WRF PXACM range from 2 to 3 K over the central and western United States. For contrast, McNally (2009) found MAE values of similar MM5 simulations that ranged from 2.7 to around 3.0 K over the western United States during the 2004–06 summers. These MAE values equate to RMSE values well above 3.0 K.

The unusually large errors along coastlines like Cape Cod, Massachusetts, are mostly an issue of model resolution. These observations sites are very near the coastline so the model land use fraction of these grid cells is often primarily water, while the site is actually over land and in reality has a larger diurnal temperature fluctuation than the model simulates.

Figures 3e and 3f provide the difference of RMSE between the WRF PXACM and the other two simulations, which again, provides easy identification of where the WRF PXACM has less (or more) error than the simulation being compared. The most dramatic improvement of the WRF PXACM relative to the MM5 PXACM is clearly over the western part of the domain. In some locations, like Colorado, the RMSE is reduced by anywhere from 0.2 to 1.0 K. Otherwise, there are scattered reductions as well as some increases, generally less than 0.40 K, across much of the eastern United States. The WRF PXACM has a lower RMSE at 57% of observation sites.

A comparison of WRF PXACM to the WRF NoahYSU configuration (Fig. 3f) indicates a more definitive difference in model performance. The WRF PXACM has a much lower RMSE than the WRF NoahYSU at 82% of sites in August. The WRF NoahYSU has a lower RMSE at a few sites (southern Texas and around Cape Cod), however, the WRF PXACM has an RMSE that is around 0.50 K lower than the WRF NoahYSU at almost all other sites. This demonstrates the usefulness of the soil nudging done in the PX LSM for retrospective simulations. The NoahYSU combination may work well in forecast application, but the PX LSM has much lower errors, as it was specifically designed for warm season retrospective simulations. Note that work is ongoing to improve the cold season performance, such as the incorporation of a dynamic snow/ice model.

Water vapor mixing ratio is an observed variable that is, like temperature, not nudged in the PBL. A slight difference exists between MM5 and WRF. The 2-m mixing ratio is diagnosed in WRF from the lowest model level value, surface moisture, and the stability. In MM5, the first-level mixing ratio is used directly to compare to the 2-m mixing ratio observations. In January, the WRF NoahYSU has the lowest RMSE (0.78 g kg^{-1}), while the WRF PXACM has the largest error (0.92 g kg^{-1}) of the model configurations, as indicated in Table 2. In August, however, the WRF PXACM has the lowest RMSE of the three configurations (1.86 versus 1.92 and 2.11 g kg^{-1}).

While most studies examine model performance using other moisture-related variables like relative humidity and dewpoint temperature, a few have used water vapor mixing ratio and the WRF PXACM errors presented in Table 2 compare favorably. Deng et al. (2004) conducted summer simulations for the eastern United States that utilized FDDA and found RMSE of mixing ratio near 2.0 g kg^{-1}. McNally (2009) presented mixing ratio MAE values for similar MM5 simulations and found values that ranged from 1.30 g kg^{-1} (the Northeast) to 2.00 g kg^{-1} (southern United States) in August, and from as low as 0.33 g kg^{-1} (midwestern United States) to 0.90 g kg^{-1} (southern United States) in January. Otte (2008a,b) found mixing ratio MAE that ranged from about 1.8 to 2.2 g kg^{-1} in the summer for a 36-km MM5 conterminous United States (CONUS) simulation.

The wind components are nudged in the PBL by the FDDA algorithm; however, the 10-m wind is a diagnosed variable that uses the first-level wind and micrometeorological parameters, so the 10-m wind is not a variable that is directly nudged. Thus, there is some use in at least mentioning the model performance with regard to simulating the wind speed and direction at 10 m because the wind at this level is used directly by many applications, and in some sense this is an assessment of the parameters involved in the diagnosed values (stability, roughness, and first-level wind speed). In both January and August the WRF PXACM has a lower RMSE of wind speed (1.64 and 1.47 m s^{-1}) than the MM5 PXACM (1.79 and 1.49 m s^{-1}) and the WRF NoahYSU (1.78 and 1.60 m s^{-1}). Emery (2001) suggested that an MAE for

10-m wind speed below 2.0 m s^{-1} is reasonable. The MAE of the WRF PXACM was 1.24 m s^{-1} in January and 1.11 m s^{-1} in August. Other studies that use similar modeling techniques and domains found wind speed MAE that ranged from 1.2 to 1.5 m s^{-1} (Otte 2008a) and 1.15 to 1.29 m s^{-1} (Gilliam et al. 2006). Studies that have published the RMSE of 10-m wind speed found values of 1.6–1.9 m s^{-1} (Hanna and Yang 2001) and 3.06 m s^{-1} (Deng et al. 2004).

For wind direction the MAE was chosen over RMSE because of the greater sensitivity to large difference between the model and observations (Wilks 1995), which often occurs with wind direction, especially during light wind conditions. For a simple example, if the difference between the modeled and observed wind direction is 1°, 10°, and 100°, the RMSE would be 58° while the MAE is 37°.

In both months the WRF PXACM has a lower MAE (20° and 31°) of wind direction, by a few degrees, than the other simulations. These numbers compare favorably with Otte (2008a,b) and Deng et al. (2004), who found MAE of wind direction in the 35°–40° range for similar simulations over the eastern United States.

In summary, near-surface observations, other models and configurations, as well as past studies used to assess the new WRF physics show that, in general, the implementation provides a state-of-the-science representation of the atmosphere, at least near the surface. The new physics parameterizations performed best during the warm season where near-surface meteorology had very low uncertainty across the central and eastern United States. However, improvements in the simulated near-surface temperature across the eastern United States are necessary for the colder times of the year, although wind speed and direction were well simulated relative to the other simulations.

b. PBL wind and potential temperature

Many model evaluation studies focus primarily on model performance based on observations near the surface (e.g., 2- and 10-m levels) or other surface measurements like precipitation. Although most air pollution is emitted near the surface, it is quickly mixed throughout the PBL, so it is also important to understand model uncertainty throughout other parts of the PBL.

To examine this ability of the model to replicate key features within the PBL, wind profiler and aircraft observations were compared to the model simulations. The mean simulated and observed wind speed and direction as a function of height and time of day are presented in Fig. 4. The mean wind speed is computed using observations at 19 NOAA wind profiler sites in the central United States from Texas northward to Nebraska and Iowa. The simulated wind speed is nudged at all model levels; however, these wind profile observations were not directly used by Obsgrid or RAWINS, but were used by the NAM data assimilation system. With this said, a test was conducted in which these wind profilers were considered by Obsgrid, and then used to drive the FDDA in WRF. The resulting wind speed errors were significantly less, so the NAM data assimilation system may more loosely fit the analysis to these observations than Obsgrid, or the impact of the profilers in the NAM data assimilation system is diminished to an extent by other observations. Furthermore, only the 0000, 0600, 1200, and 1800 UTC analyses used for the FDDA are NAM analyses; the 0300, 0900, 1500, and 2100 UTC FDDA data are NAM 3-h forecasts that do not have assimilated profiler data.

Figure 4 indicates that in January the WRF PXACM does replicate some of the key PBL features that were observed. In particular, the nocturnal jet located at approximately 600 m is well simulated by the model. The dots on both the observation and model panels represent the WRF PXACM–simulated PBL height. Observed PBL height is not a standard product from the wind profiler, so it is not provided. However, the model-simulated PBL does generally agree with the PBL top as suggested by the collection of 19 wind profilers, which is often considered the level just below the core of the nocturnal jet (Stull 1997). Also of note, the strength of the nocturnal jet (~13 m s^{-1}) is well simulated by the WRF PXACM. During the daytime hours, the model seems to maintain the low-level jet into the morning hours and overestimates the wind speed below 1000 m in the afternoon.

Figure 4 indicates that in August the WRF PXACM replicates the PBL wind structure well. According to the wind profilers the nocturnal jet has a magnitude of 10–11 m s^{-1}, while the model slightly underestimates this feature (9–10 m s^{-1}). Similar to January, the model simulates the height and evolution of the nocturnal jet core, which is around 500 m. During the day, the WRF PXACM reproduces the changes in the wind speed profile during the transition from the stable to convective PBL (1300–1800 UTC). At around 2100 UTC, the observed mean wind speed profile has a near-constant wind speed layer from the surface to 1800–1900 m, a result of strong vertical mixing in the convective boundary layer. The modeled mean wind profile is also well mixed by the midafternoon with a mean PBL height maximum around 1900 m. Overall, the wind profiler observations imply that the WRF PXACM reasonably simulates the evolution of wind speed in the PBL across a large part of the central United States.

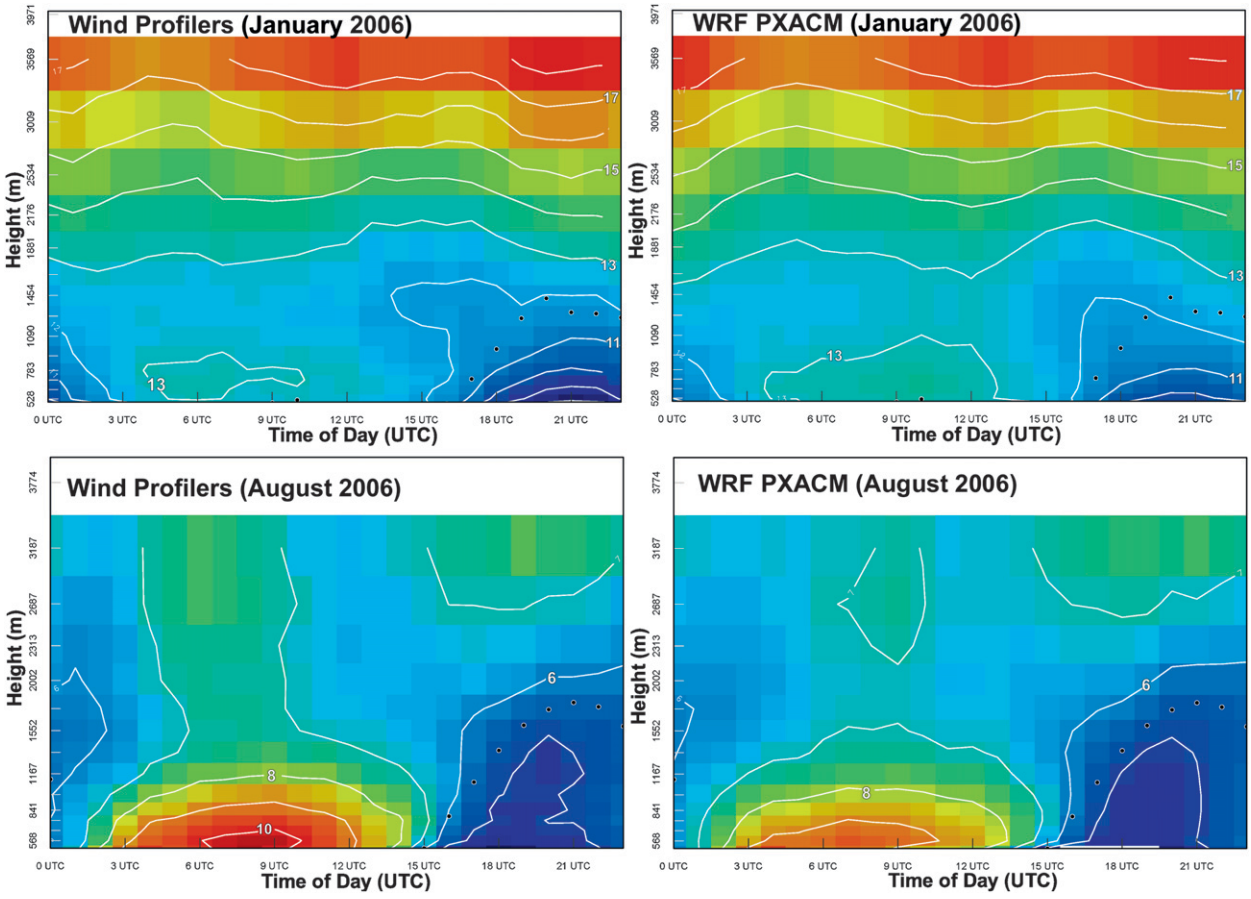

FIG. 4. Diurnal mean wind speed profiles (m s^{-1}, height above ground level) for (top) January and (bottom) August 2006. (left) The mean observed wind speed computed using 19 NOAA wind profilers located in the central United States and (right) the corresponding model-simulated mean wind speed using the grid points closest to the wind profiler sites.

To examine the performance of the WRF PXACM relative to the other model simulations, each configuration was paired with observations from the central U.S. wind profilers at each model level and then the MAE profile was computed. Figure 5 presents the MAE profiles of simulated wind speed and direction for January and August, and Fig. 6 presents the MAE at three specific model levels as a function of time of day. For January, the WRF PXACM has a lower wind speed and direction error than the other simulations, especially below 1000 m. All models have a larger MAE of wind speed near 500 m that decreases at 1000 m and increases again to a maximum at around 1300 m, which is about the mean depth of the PBL. It is also interesting to note that the overall wind speed MAE decreases above the mean height of the PBL where the wind speed increases with height. For example, the mean wind speed increases from 11–12 m s^{-1} on average around 1000 m to around 17 m s^{-1} above 3000 m, while the MAE decreases slightly, so the error decreases in the free atmosphere from 13% to around 9%. In January, the WRF PXACM has consistently less wind error (speed and direction) than the MM5 PXACM. Figure 6 indicates that the WRF PXACM has almost identical diurnal error to the WRF NoahYSU at 1300 and 2000 m. However, the WRF PXACM has noticeably lower error around the 700-m level relative to the WRF NoahYSU at night (0100–1000 UTC), but similar error during the day.

The WRF PXACM also compares favorably with the other two simulations in August 2006 (Fig. 5). The WRF PXACM has a clear advantage over the MM5 PXACM at all levels in terms of MAE. The WRF PXACM has a slightly lower wind speed error between 500 and 1000 m, and slightly more near the 2000-m level than the WRF NoahYSU. This may be an indication that the WRF PXACM more accurately represents the nocturnal jet, but that the WRF NoahYSU better represents the height and wind distribution near the top of the convective PBL. Diurnally averaged MAE data in Fig. 6 do provide some evidence to support this thought. The MAE at these

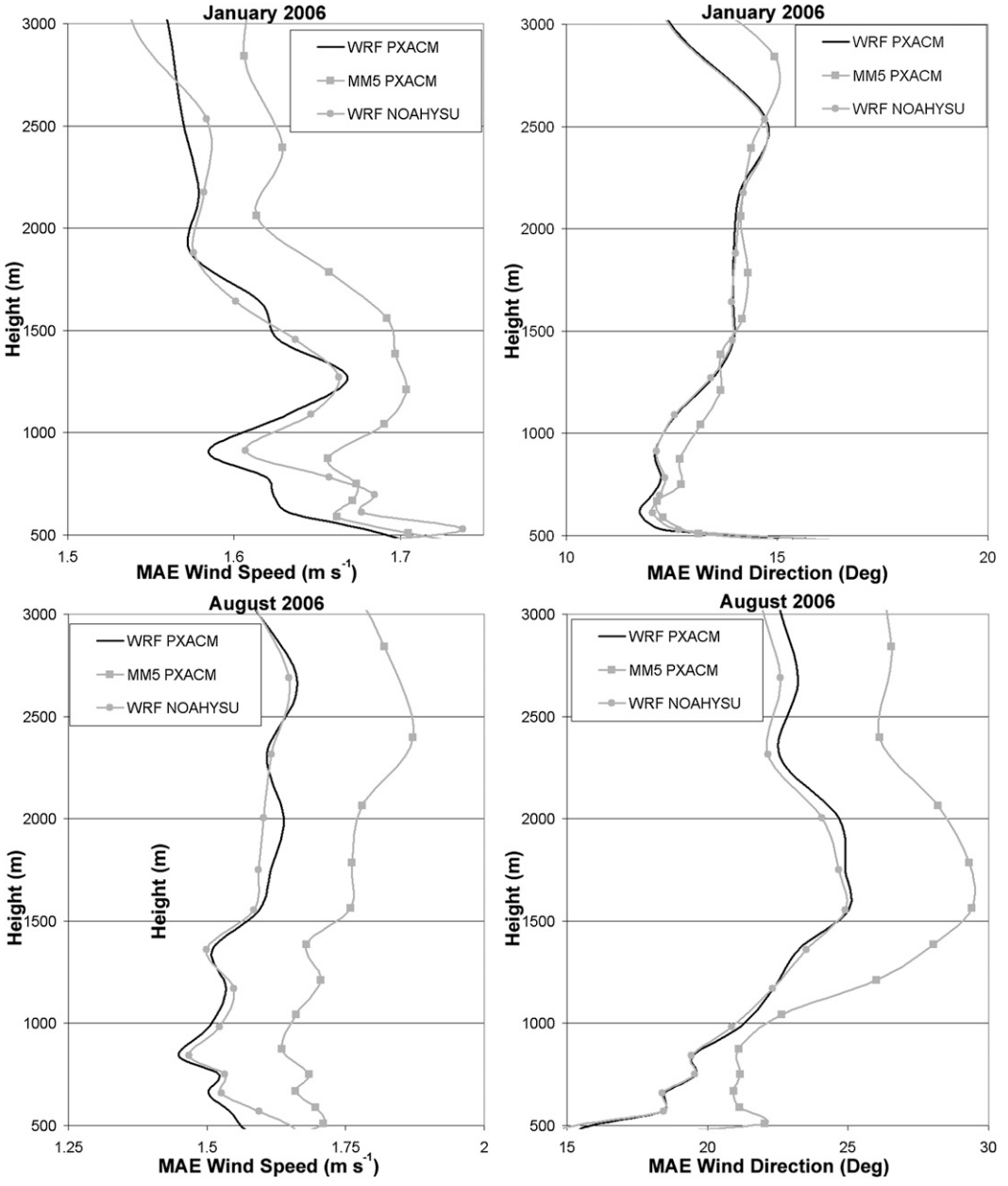

FIG. 5. MAE profiles of model-simulated wind (left) speed and (right) direction for (top) January and (bottom) August 2006. The observations used to compute MAE include 19 NOAA wind profilers located in the central United States.

levels indicates that the WRF PXACM does have a slightly lower error around 700 m from night through the midmorning (0400–1500 UTC), and a slightly larger MAE than the WRF NoahYSU around 2000 m during the day (1700–2300 UTC). There is little systematic difference between the WRF PXACM and WRF NoahYSU at the 1300-m level.

The MM5 MAE is visibly different than both WRF simulations. This raises the question of the validity of the data extraction and interpolation of observations to the slightly different vertical grids. This potential issue was examined and no errors could be found. Also, the temperature and wind speed profiles were extracted and compared independently. The mean MM5 profiles were different than the WRF. The amount of MAE in both WRF simulations does converge above the PBL, which is expected since both used the same Obsgrid analyses to nudge the wind. The difference of MAE within the PBL

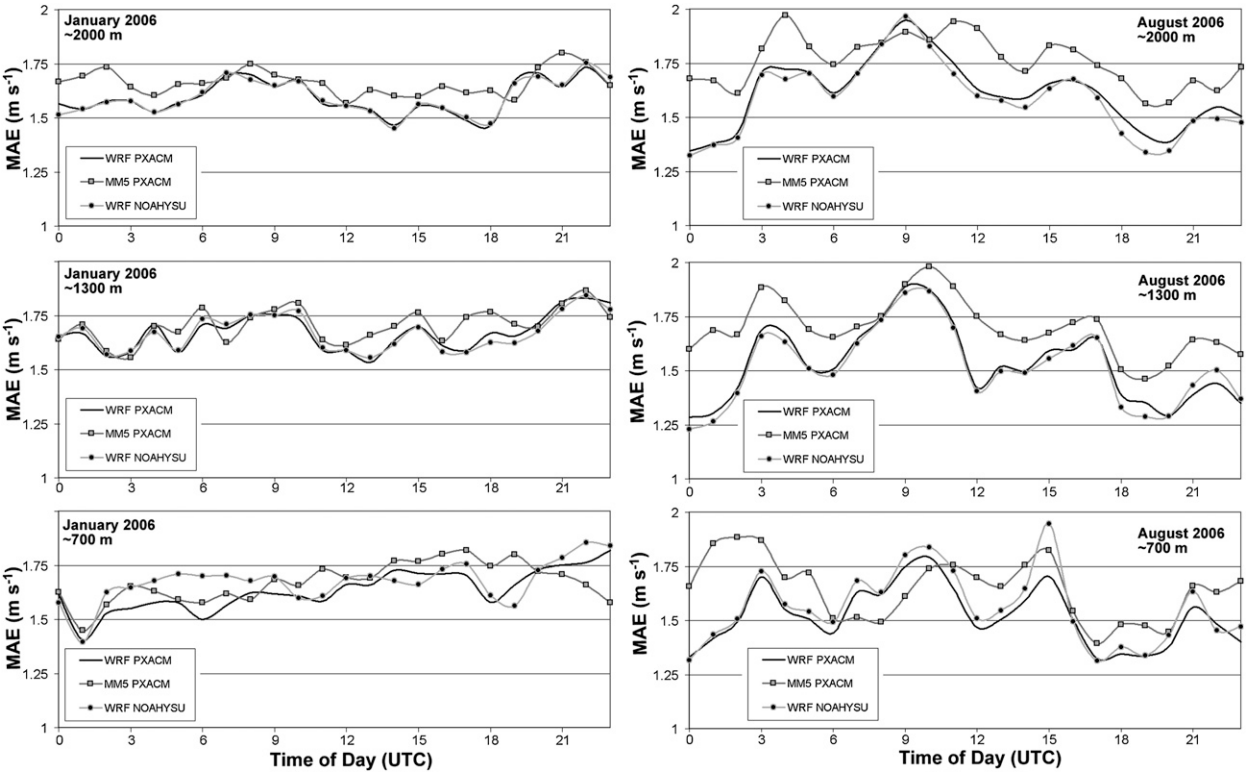

FIG. 6. Comparison of diurnally averaged MAE of wind speed at approximately (top) 2000, (middle) 1300, and (bottom) 700 m for the (left) January and (right) August 2006 simulations.

does signify that despite the same nudging strategy, the nudging strength is low enough to allow some solution diversity among PBL schemes.

Aircraft profiles were also used to examine the model performance in the PBL. In this comparison, observations at 119 airports in the eastern United States were collected and paired with each model. It should be mentioned that these aircraft data are used by the NAM data assimilation system to generate the 0000, 0600, 1200, and 1800 UTC NAM analyses that are used for FDDA, so these aircraft data are not completely independent in this analysis. The MAE profiles are provided in Fig. 7 for each month and in this case not only for wind, but also potential temperature. The wind speed MAE is generally lower in the WRF PXACM than the other simulations in January, especially between 600 and 1500 m. Wind direction errors are very similar between all simulations. The temperature MAE of the WRF PXACM is larger at the surface, which is in agreement with the RMSE of 2-m temperature across the eastern United States shown in Fig. 3. However, above the surface between 300 and 1500 m, the WRF PXACM does have a much lower MAE.

Figure 7 indicates that the WRF PXACM has the lowest wind and temperature MAE within the PBL in August 2006. The difference of wind speed MAE between the WRF PXACM and WRF NoahYSU is small. Both have larger errors just above the surface that decrease within the PBL but increase again near the top of the mean convective PBL (~2000 m). It is clear that the MM5 PXACM has a larger MAE than both WRF simulations. In particular, the potential temperature MAE is as much as 0.5–1.0 K larger in the 500–3000-m layer.

Similar to Fig. 5, the WRF NoahYSU has a slightly lower MAE around the 1800-m level, which is additional evidence the NoahYSU is capturing the wind structure around the top of the PBL, or the PBL height more accurately. The diurnal MAE for the WRF PXACM and WRF NoahYSU were examined (not shown) and the MAE of the WRF NoahYSU was lower at the 1800–2000-m level during the 1800–2100 UTC part of the day.

The WRF PXACM has the smallest potential temperature MAE in the lowest 500 m of the profile, and an error of 1.0 K or less between 500 and 1500 m. This level of MAE in the PBL (~1.0 K) where temperature nudging does not take place is much lower than the errors seen at 2 m and on the order of the NAM analyses that use an error minimization technique.

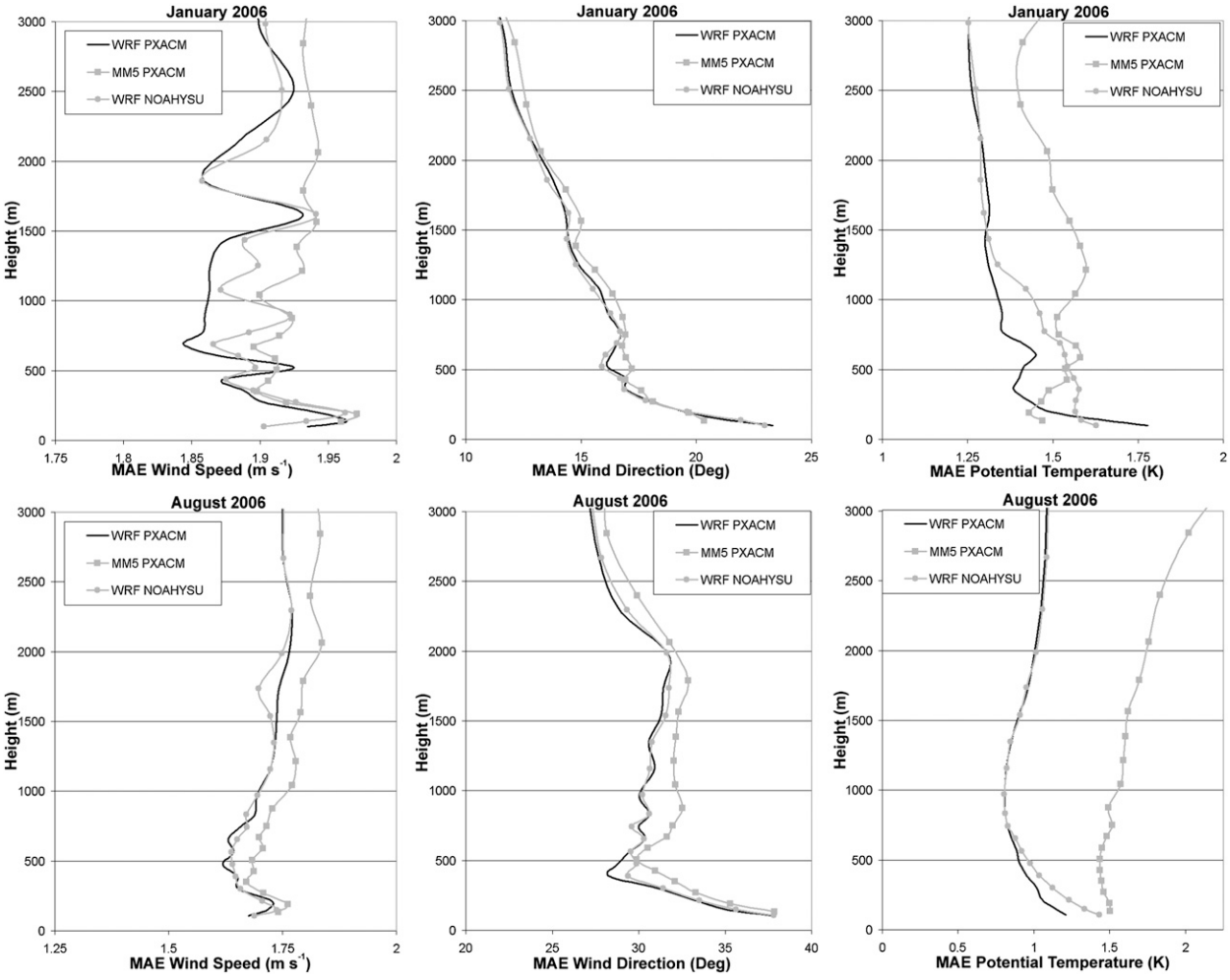

FIG. 7. MAE profiles of model-simulated (left) wind speed (m s^{-1}), (center) wind direction (°), and (right) potential temperature (K) for (top) January and (bottom) August 2006. The observations that were used to compute the MAE include hourly aircraft profiles at 109 airports located in the eastern United States.

4. Summary

The Pleim–Xiu LSM, Pleim surface layer scheme, and Asymmetric Convective Model, version 2, were implemented in version 3 of the Advance Research WRF. This new physics suite is specifically designed for retrospective simulations, in particular, those used to drive air quality models. Several upgrades were made to the PX LSM in WRF. The enhancements include a deep soil temperature nudging scheme (Pleim and Gilliam 2009) and a new method to compute the surface heat capacity that accounts for a fractional land use–dependent snow cover. Also, FDDA capability was added to WRF just prior to the release of version 3.0, and an objective reanalysis tool Obsgrid was developed. These enhancements bring the retrospective simulation capabilities of WRF to the level of MM5.

Model simulations using the new WRF implementation were conducted for January and August 2006. For benchmark purposes, a similar PX LSM and ACM2 MM5 simulation and another WRF simulation with a commonly used physics configuration were also executed. A comparison of simulated near-surface meteorology was conducted that focused on 2-m temperature using point observations, but a brief examination of 2-m mixing ratio and 10-m wind was also performed. Domainwide RMSE statistics indicate that the new physics implementation in WRF, considering 2-m temperature, 2-m mixing ratio, and 10-m wind, has lower overall RMSE than the MM5 counterpart and the other WRF configuration. The exception is that the WRF NoahYSU simulation did have a lower 2-m temperature and mixing ratio error in winter. Also, the MM5 PXACM did have a lower 2-m temperature RMSE across much of the

domain in winter, but the WRF PXACM had a much lower error across the Northeast, Great Lakes region, and Rocky Mountains. The amount of error in the WRF PXACM is low, considering past modeling studies that use similar model configurations. RMSE errors in the eastern United States are on the order of 1.5–2.0 K and MAE are often 1.0–1.5 K.

It was discovered that the analyses, 2-m temperature and mixing ratio, that drive the soil moisture and temperature nudging of the PX LSM have different precision when compared to point observations. The analyses used by MM5 were generated from the RAWINS utility that reintroduces the point observations to a base model analysis (NAM), while the WRF simulations used a similar utility Obsgrid. It was found that as the RMSE of the analysis decreases there is clear improvement in the simulated 2-m temperature and mixing ratio. This fact clearly demonstrates the effectiveness of the soil nudging algorithm in the PX LSM to indirectly nudge soil temperature and moisture according to differences between the model projection and the analysis. This soil nudging then impacts surface sensible and latent heat flux partitioning that in turn influences the diagnosed 2-m temperature and water vapor mixing ratio.

Comparisons of the simulated PBL wind and temperature with wind profiler and aircraft measurements indicated that the new physics suite in WRF performs as well or better than MM5 PXACM and WRF NoahYSU. The WRF PXACM replicates key PBL features like the nocturnal jet and convective boundary layer. Mean absolute errors of wind speed, wind direction, and temperature in the PBL are much lower than the values found at the 2- and 10-m levels. Temperature errors (MAE) were below 1.0 K in parts of the PBL during August. The evaluation of the simulated PBL using wind profiler and aircraft observations is particularly relevant for air quality applications since pollution transport and dispersion occurs not just at the surface, but throughout the PBL.

Considering the evaluation presented in this study, the evidence indicates that WRF PXACM's RMSE and MAE values are at, or in most cases below, the level of MM5, especially during the warmer parts of the year. However, several improvements are necessary. A snow model for the PX LSM to improve the temperature and associated PBL properties over snow covered surfaces is under development. Also, we see evidence that during the winter the implementation of the PXACM in WRF generates excessive cloud cover at the top of the PBL, which impacts the daytime high temperature in some areas. This cloud issue needs to be examined in more detail. Finally, it is recommended that model analyses be used more regularly in model evaluation studies, as presented here, since they are used in both FDDA and soil nudging, thus representing the lowest error one can expect.

Acknowledgments. The U.S. Environmental Protection Agency through its Office of Research and Development funded and managed the research described here. It has been subjected to Agency review and approved for publication.

The authors acknowledge NCAR and collaborators who have invested much time and effort in the development of the WRF, and who also helped with technical aspects of implementing the physics parameterizations, particularly Jimy Dudhia and Wie Wang; A. J. Deng and David Stauffer at Penn State University, Tanya Otte at the U.S. EPA, and Cindy Bruyère at NCAR are acknowledged for their contributions to the development of FDDA in WRF, version 3.0, and the Obsgrid reanalysis utility that was critical for a fair evaluation of WRF. Aijun Xiu and the University of North Carolina at Chapel Hill's Institute for the Environment are thanked for their contributions, especially the initial effort to port the MM5 code to the WRF framework. NOAA ESRL/GSD is recognized for the development of the MADIS observational archive, which was critical for our evaluation. Finally, the National Climatic Data Center is thanked for the model analysis archive used to drive the WRF simulations.

REFERENCES

Baker, K., 2004: Meteorological modeling protocol for application to PM2.5/haze/ozone modeling projects. Lake Michigan Air Directors Consortium, Des Plaines, IL, 9 pp.

Barth, M. F., R. B. Chadwick, and D. W. van de Kamp, 1994: Data processing algorithms used by NOAA's Wind Profiler Demonstration Network. *Ann. Geophys.,* **12,** 518–528.

Byun, D., and K. L. Schere, 2006: Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. *Appl. Mech. Rev.,* **59,** 51–77.

Daniels, T. S., G. Tsoucalas, M. Anderson, D. Mulally, W. Moninger, and R. Mamrosh, 2004: Tropospheric Airborne Meteorological Data Reporting (TAMDAR) sensor development. Preprints, *11th Conf. on Aviation, Range, and Aerospace,* Hyannis, MA, Amer. Meteor. Soc., 7.6. [Available online at http://ams.confex.com/ams/pdfpapers/81841.pdf.]

Deng, A., N. L. Seaman, G. K. Hunter, and D. R. Stauffer, 2004: Evaluation of interregional transport using the MM5–SCIPUFF System. *J. Appl. Meteor.,* **43,** 1864–1886.

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.

Dyer, A. J., and B. B. Hicks, 1970: Flux-gradient relationships in the constant flux layer. *Quart. J. Roy. Meteor. Soc.,* **96,** 715–721.

Ek, M. B., K. E. Mitchell, Y. Lin, P. Grunmann, E. Rogers, G. Gayno, and V. Koren, 2003: Implementation of the upgraded Noah land-surface model in the NCEP operational mesoscale Eta model. *J. Geophys. Res.,* **108,** 8851, doi:10.1029/2002JD003296.

Emery, C. A., 2001: Enhanced meteorological modeling and performance evaluation for two Texas ozone episodes. Texas Natural Resource Conservation Commission, ENVIRON International Corporation, 235 pp. [Available online at http://www.tceq.state.tx.us/assets/public/implementation/air/am/contracts/reports/mm/EnhancedMetModelingAndPerformanceEvaluation.pdf.]

Gilliam, R. C., P. V. Bhave, J. E. Pleim, and T. L. Otte, 2004: A year-long MM5 evaluation using a model evaluation toolkit. *Extended Abstracts, Third Annual Community Analysis and Modeling System Conf.,* Chapel Hill, NC, University of North Carolina Institute for the Environment, 4 pp. [Available online at http://www.cmascenter.org/conference/2004/abstracts/poster/gilliam_abstract.pdf.]

——, W. Appel, and S. Philips, 2005: The Atmospheric Evaluation Tool (AMET): Meteorology module. *Extended Abstracts, Fourth Annual Community Analysis and Modeling System Conf.,* Chapel Hill, NC, University of North Carolina Institute for the Environment, 6 pp. [Available online at http://www.cmascenter.org/conference/2005/abstracts/6_1.pdf.]

——, C. Hogrefe, and S. T. Rao, 2006: New methods for evaluating meteorological models used in air quality applications. *Atmos. Environ.,* **40,** 5073–5086.

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1995: A description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398+STR, 117 pp. [Available online at http://www.mmm.ucar.edu/mm5/doc1.html.]

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.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.,* **43,** 170–181.

McNally, D., 2009: 12km MM5 performance goals. *10th Annual Ad-Hoc Meteorological Modelers Meeting,* Boulder, CO, U.S. Environmental Protection Agency, 46 pp. [Available online at http://www.epa.gov/scram001/adhoc/mcnally2009.pdf.]

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

Moninger, W. R., R. D. Mamrosh, and P. M. Pauley, 2003: Automated meteorological reports from commercial aircraft. *Bull. Amer. Meteor. Soc.,* **84,** 203–216.

NCAR, cited 2009a: RAWINS tutorial. [Available online at http://www.mmm.ucar.edu/mm5/mm5v3/tutorial/rawins/rawins.html.]

——, cited 2009b: Objective analysis (*Obsgrid*). [Available online at http://www.mmm.ucar.edu/wrf/users/docs/user_guide_V3/users_guide_chap7.htm.]

Noh, Y., W. G. Cheon, S.-Y. Hong, and S. Raasch, 2003: Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Bound.-Layer Meteor.,* **107,** 401–427.

Otte, T. L., 2008a: The impact of nudging in the meteorological model for retrospective air quality simulations. Part I: Evaluation against national observations networks. *J. Appl. Meteor. Climatol.,* **47,** 1853–1867.

——, 2008b: The impact of nudging in the meteorological model for retrospective air quality simulations. Part II: Evaluating collocated meteorological and air quality observations. *J. Appl. Meteor. Climatol.,* **47,** 1868–1887.

Pleim, J. E., 2006: A simple, efficient solution of flux–profile relationships in the atmospheric surface layer. *J. Appl. Meteor. Climatol.,* **45,** 341–347.

——, 2007a: 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.

——, 2007b: A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: Application and evaluation in a mesoscale meteorological model. *J. Appl. Meteor. Climatol.,* **46,** 1396–1409.

——, and J. S. Chang, 1992: A non-local closure model for vertical mixing in the convective boundary layer. *Atmos. Environ.,* **26A,** 965–981.

——, and A. Xiu, 2003: Development of a land-surface model. Part II: Data assimilation. *J. Appl. Meteor.,* **42,** 1811–1822.

——, and R. Gilliam, 2009: An indirect data assimilation scheme for deep soil temperature in the Pleim–Xiu land surface model. *J. Appl. Meteor. Climatol.,* **48,** 1362–1376.

Reisner, J., R. M. Rasmussen, and R. T. Bruintjes, 1998: Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model. *Quart. J. Roy. Meteor. Soc.,* **124,** 1071–1107.

Skamarock, W. C., and Coauthors, 2008: A description of the advanced research WRF version 3. NCAR Tech Note NCAR/TN-475+STR, 125 pp. [Available from UCAR Communications, P.O. Box 3000, Boulder, CO 80307.]

Stauffer, D. R., and N. L. Seaman, 1990: Use of four-dimensional data assimilation in a limited area mesoscale model. Part I: Experiments with synoptic-scale data. *Mon. Wea. Rev.,* **118,** 1250–1277.

——, and ——, 1994: Multiscale four-dimensional data assimilation. *J. Appl. Meteor.,* **33,** 416–434.

——, ——, and F. S. Binkowski, 1991: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part II: Effects of data assimilation within the planetary boundary layer. *Mon. Wea. Rev.,* **119,** 734–754.

Stull, R. B., 1997: *An Introduction to Boundary Layer Meteorology*. 6th ed. Kluwer, 670 pp.

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.,* **132,** 519–542.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Xiu, A., and J. E. Pleim, 2001: Development of a land surface model. Part I: Application in a mesoscale meteorological model. *J. Appl. Meteor.,* **40,** 192–209.