The Impact of Nudging in the Meteorological Model for Retrospective Air Quality Simulations. Part II: Evaluating Collocated Meteorological and Air Quality Observations

TANYA L. OTTE

Atmospheric Sciences Modeling Division, NOAA/Air Resources Laboratory, Research Triangle Park, North Carolina*

(Manuscript received 29 May 2007, in final form 6 November 2007)

ABSTRACT

For air quality modeling, it is important that the meteorological fields that are derived from meteorological models reflect the best characterization of the atmosphere. It is well known that the accuracy and overall representation of the modeled meteorological fields can be improved for retrospective simulations by creating dynamic analyses in which Newtonian relaxation, or "nudging," is used throughout the simulation period. This article, the second of two parts, provides additional insight into the value of using nudging-based data assimilation for dynamic analysis in the meteorological fields for air quality modeling. Meteorological simulations are generated by the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) using both the traditional dynamic analysis approach and forecasts for a summertime period. The resultant meteorological fields are then used for emissions processing and air quality simulations using the Community Multiscale Air Quality Modeling System (CMAQ). The predictions of surface and near-surface meteorological fields and ozone are compared with a small network of collocated meteorological and air quality observations. Comparisons of 2-m temperature, 10-m wind speed, and surface shortwave radiation show a significant degradation over time when nudging is not used, whereas the dynamic analyses maintain consistent statistical scores over time for those fields. Using nudging in MM5 to generate dynamic analyses, on average, leads to a CMAQ simulation of hourly ozone with smaller error. Domainwide error patterns in specific meteorological fields do not directly or systematically translate into error patterns in ozone prediction at these sites, regardless of whether nudging is used in MM5, but large broad-scale errors in shortwave radiation prediction by MM5 directly affect ozone prediction by CMAQ at specific sites.

1. Introduction

Complex, limited-area, multipollutant air quality models typically obtain meteorological fields from Eulerian (gridded) meteorological models. It is well known that the meteorological conditions exert a significant influence on air quality (e.g., Flaum et al. 1996; Milanchus et al. 1998). For nearly two decades, there has been a great interest in assessing the impact that meteorological modeled fields have on air quality model simulations. Pielke and Uliasz (1998) and Seaman (2000) provide overviews of meteorological modeling approaches that are used for air quality modeling applications and present advantages and challenges for air quality modeling. Focused evaluations of various aspects of planetary boundary layer evolution within meteorological models have been conducted because of its impact in air quality modeling applications (e.g., Berman et al. 1999; Ku et al. 2001; Hanna and Yang 2001; Angevine and Mitchell 2001). Various components of meteorological models are being developed specifically to address the needs of the air quality modeling community (e.g., Jacobson 2001; Pleim and Xiu 2003; Grell et al. 2005; Pleim 2007). Several evaluation studies have examined meteorological and air quality models together as a system (e.g., Biswas and Rao 2001; Hogrefe et al. 2001a,b, 2006; Lee et al. 2004). There have also been advances in model evaluation techniques specifically for meteorological fields used in

* In partnership with the U.S. EPA National Exposure Research Laboratory.

Corresponding author address: Tanya L. Otte, NOAA/Atmospheric Sciences Modeling Division, U.S. EPA Mail Drop E243-03, 109 T. W. Alexander Dr., Research Triangle Park, NC 27711. E-mail: tanya.otte@noaa.gov

DOI: 10.1175/2007JAMC1791.1

air quality modeling applications (e.g., Liu et al. 2003; Gilliam et al. 2006).

Newtonian relaxation, or "nudging," is commonly used throughout meteorological modeling simulations to create "dynamic analyses" of the evolving meteorological state. Nudging involves adding an artificial forcing term to the governing equations that reflects the difference between the best estimate of the observed state and the model state at a given location and time (Stauffer and Seaman 1990, 1994; Stauffer et al. 1991). Dynamic analyses from sophisticated meteorological models, such as the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5), are typically used to generate multiday meteorology simulations for the Community Multiscale Air Quality Modeling System (CMAQ; Byun and Schere 2006) and other Eulerian air quality models for retrospective research and regulatory applications.

Thus far, little emphasis has been placed on evaluating the impact to the air quality model of the dynamic analyses that are typically used by air quality models such as CMAQ. Using air quality models other than CMAQ, Barna and Lamb (2000) and Tanrikulu et al. (2000) present air quality simulation results from meteorological models with and without nudging and show that nudging can have a positive benefit on ozone prediction in episodic cases (i.e., less than 1-week modeling periods). Using the same episode as Tanrikulu et al. (2000), Umeda and Martien (2002), however, showed that nudged meteorological fields did not improve ozone predictions and that the impact of the nudged meteorological fields may be overwhelmed by the impact of other forcing from emissions and chemical boundary conditions. This research provides additional insights into the value of using nudging-based data assimilation for dynamic analysis in the meteorological fields on air quality simulations by evaluating a longer modeling period, which should lead to more robust and conclusive results. This paper is the second of two parts that quantifies the impact of nudging on the air quality simulation. In Otte 2008 (hereinafter Part I), which focuses on broad evaluation against independent meteorological and air quality observation networks that have dense national coverage, it is shown that nudging in the meteorological model improves the daily maximum 1-h ozone simulation by CMAQ, on average, throughout the simulation domain. Part I also illustrates that the relationship and trends between statistical skill scores for daily maximum 1-h ozone do not parallel those obtained for near-surface meteorological variables, regardless of whether nudging is used in the meteorological model. Here the focus is on the evaluation of meteorological fields and air quality at collocated observation sites. This paper also offers perhaps the first look at the impact of nudging in the meteorological model on shortwave radiation by making comparisons with a network of observations. Meteorological fields and air quality (specifically, ozone) are evaluated by comparing simulations for a 5-week summer period with and without nudging. Section 2 briefly describes the model configurations. Section 3 includes analysis of the MM5 and CMAQ simulations to illustrate the impact of nudging in MM5 on CMAQ. The final section includes a summary of the conclusions.

2. Model configuration

The meteorological, emissions, and air quality modeling suite is run for two configurations of input meteorological conditions: one that uses analysis nudging throughout the simulation (i.e., a dynamic analysis) and one that does not (i.e., effectively a series of forecasts). The simulations are performed on a domain with 36-km horizontal grid spacing that includes the continental United States and parts of Canada and Mexico (see Part I). Thirty-four terrain-following layers are used for both the meteorological and air quality simulations; there is no "collapsing of layers" or reduction of vertical resolution in the air quality model, as is commonly done (e.g., Eder and Yu 2006; Hogrefe et al. 2006). There are 18 layers in the lowest 2 km of the atmosphere for the meteorological and air quality simulations.

MM5 (Grell et al. 1994), version 3.6, is used for the meteorological simulations. In the MM5 simulation that includes nudging, the assimilation is based on 3-h 3D analyses of temperature, water vapor mixing ratio, horizontal wind components, and 3-h surface analyses of horizontal wind components; there is no nudging of mass fields within the PBL following Stauffer et al. (1991). The emissions are processed using the Sparse Matrix Operator Kernel Emissions (SMOKE) modeling system (Houyoux et al. 2000) version 2.2. The biogenic and point-source emissions sectors are modulated by the MM5 fields; all other emissions sectors are independent of MM5. The chemistry transport is modeled using CMAQ (Byun and Schere 2006), version 4.6. The MM5, SMOKE, and CMAQ model configurations, input data, and settings are thoroughly described in Part I.

The MM5 fields are generated for the period 1200 UTC 19 June–0000 UTC 4 August 2001. The period is broken into nine overlapping 5.5-day run segments. The first 12 h of each MM5 segment are a "spinup" period for cloud processes, and they are not used for emissions or chemistry processing; the remaining 5 days

TABLE 1. Dates in 2001 used for analysis as given by time elapsed within each MM5 run segment. MM5 segments 1 and 2 are part of the CMAQ spinup period, and they are not included in the analysis.

MM5 segment	Day 1	Day 2	Day 3	Day 4	Day 5
3	30 Jun	1 Jul	2 Jul	3 Jul	4 Jul
4	5 Jul	6 Jul	7 Jul	8 Jul	9 Jul
5	10 Jul	11 Jul	12 Jul	13 Jul	14 Jul
6	15 Jul	16 Jul	17 Jul	18 Jul	19 Jul
7	20 Jul	21 Jul	22 Jul	23 Jul	24 Jul
8	25 Jul	26 Jul	27 Jul	28 Jul	29 Jul
9	30 Jul	31 Jul	1 Aug	2 Aug	3 Aug

are input for the air quality model. All fields except soil moisture are reinitialized in each MM5 segment, as is typically done for retrospective modeling applications using CMAQ. The CMAQ simulations cover the period 0000 UTC 20 June–0000 UTC 4 August 2001, but the first 10 days are considered spinup to allow the chemistry to come into equilibrium, and they are not used in the analysis.

3. Analysis

Two sets of MM5 and CMAQ simulations for the 5-week period 30 June–3 August 2001 are analyzed to assess the impact of using nudging in MM5 on the CMAQ simulation. This time period is selected because it is in the middle of the ozone season, which is typically May through September in most areas of the United States, when high levels of pollutants are typically observed. The first set of simulations ("NONUDGE") comprises overlapping 5.5-day forecasts in which temperature, water vapor mixing ratio, and horizontal winds in MM5 are not nudged toward analyses. The second set of simulations ("NUDGE") includes overlapping 5.5-day dynamic analyses that are prepared in the traditional manner as input to CMAQ for retrospective studies. The 5-week period includes seven MM5 run segments (see Table 1). As the first 12 h of each MM5 run segment are not used, "day 1" refers to an aggregate of hours 13–36 of the seven MM5 run segments (i.e., the 24-h period from 0100–0000 UTC the following day), "day 2" refers to hours 37–60, and so on. The CMAQ performance is binned in time using the same temporal structure to determine the impact on the chemical transport model as it corresponds to increased simulation run time in MM5. Statistics are computed for surface and near-surface meteorological fields and the near-surface ozone mixing ratio using statistical metrics suggested by Willmott (1982).

The meteorological and air quality observations in this study are from the Clean Air Status and Trends Network (CASTNET; http://www.epa.gov/castnet), which is jointly maintained by the U.S. Environmental Protection Agency (EPA) and the National Park Service. There are more than 125 CASTNET sites throughout the United States (see Fig. 1), and they are mostly located in rural areas (i.e., in national parks, at national monuments, and near land-grant universities). The purpose of CASTNET is to provide atmospheric data related to the dry deposition component of total acid deposition, near-surface ozone, and other forms of atmospheric pollution to identify long-term pollutant deposition trends and spatial patterns. CASTNET records hourly measurements of surface and near-surface meteorological fields (i.e., temperature, dewpoint, wind speed and direction, shortwave radiation, and precipitation, among others) and ozone together, which provides a unique opportunity to examine meteorological fields and air quality observed at the same location using a nationwide network of sites. CASTNET also records average weekly concentrations of several photochemical species, which are not evaluated herein because daily totals cannot be derived from those measurements.

The CASTNET sites in the western United States are often located in complex terrain in U.S. national parks and national monuments, and those data are typically less representative of the average conditions within a 36 km $\times$ 36 km grid cell. Therefore, the western stations are omitted, and this analysis is limited to the 55 CASTNET stations east of 100°W longitude. All of the CASTNET sites in the eastern United States are included, although four of the sites are at mountaintops and eight others are located in complex terrain, and exhibit the same complications as sites in the western United States. It is also noteworthy that the predominantly rural CASTNET sites may be far removed from the source of the ozone production. Thus the dominant effect at these sites is likely to be the transport of ozone from urban areas.

There are inherent limitations with using point-based observational data for the evaluation of grid-based meteorological and air quality models. Obviously there are spatial representativeness issues with comparing point measurements with volume-average model predictions. However, both model simulations should experience the same disadvantages for this evaluation because the same measurements are used consistently.

Only a cursory evaluation of the near-surface meteorological state variables is done here; Part I provides a more thorough comparison of the near-surface fields against the broader network of National Weather Service (NWS) observations. Although precipitation is

FIG. 1. Map of CASTNET sites in 2001 (from Lavery et al. 2002).

measured by CASTNET, it is not used for evaluation here because summertime precipitation is generally convective (i.e., subgrid scale at 36-km horizontal grid spacing) and matching precipitation in space and time is not appropriate in this context. The CASTNET data used here are an entirely independent data source for evaluation. The meteorological observations in CASTNET are not reported to the NWS, and they are not included in the analyses toward which the MM5 simulations are forced in NUDGE.

Figure 2 shows the mean absolute error (MAE) for hourly 2-m temperature computed "by day" within the MM5 simulation at each of the CASTNET sites in the eastern United States (Fig. 1) for NONUDGE and NUDGE. Here and throughout the paper, day 1, for example, refers to an average for 30 June, 5 July, and every fifth day through 30 July 2001 (see Table 1). The MAE for 2-m temperature is generally lower, on average, in NUDGE than in NONUDGE, as expected. Not surprisingly, the 2-m temperature is consistently predicted poorly in both NUDGE and NONUDGE at three mountaintop sites, which stand out as outliers, particularly in NUDGE: from south to north, Cranberry, North Carolina (PNF126 in Fig. 1), Shenandoah National Park, Virginia (SHN418 in Fig. 1), and Lye Brook, Vermont (LYE145 in Fig. 1). The MAE is greater than 2 K at 20 sites on day 1 in NONUDGE, and it exceeds 2.5 K at 6 sites (including the 3 aforementioned mountaintop sites). By contrast, the day 1 MAE in NUDGE is greater than 2 K at only 10 sites, and it exceeds 2.5 K at 3 sites (including 2 of the 3 mountaintop sites). The magnitude of the average MAE grows progressively with increased MM5 simulation run time in NONUDGE. By day 5, only seven of the CASTNET sites have an average MAE that is less than 2 K, and approximately one-third of the sites have an average MAE that exceeds 2.5 K. In NUDGE, the MAE for 2-m temperature does not show any appreciable change with increased MM5 simulation run time, as also shown in Part I by comparing with the denser network of NWS observations. On each of the days through day 5, several of the CASTNET sites have an average MAE of less than 1.5 K, and the average MAE is generally less than 2.5 K, other than at the mountaintop sites.

The CASTNET offers a unique opportunity to examine the effects of nudging on global shortwave radiation, for which there are few routine measurements in the United States. Evaluation of the shortwave radiation can be interesting because it is not directly forced by nudging in MM5, although it can include some indirect effects of nudging through the vertical column. Cloud attenuation of shortwave radiation is important because it can directly affect the ozone production in CMAQ through the photolysis rates. Figure 3 shows the MAE for hourly shortwave radiation computed by day within the MM5 simulation at each of the CASTNET sites in the eastern United States (Fig. 1)

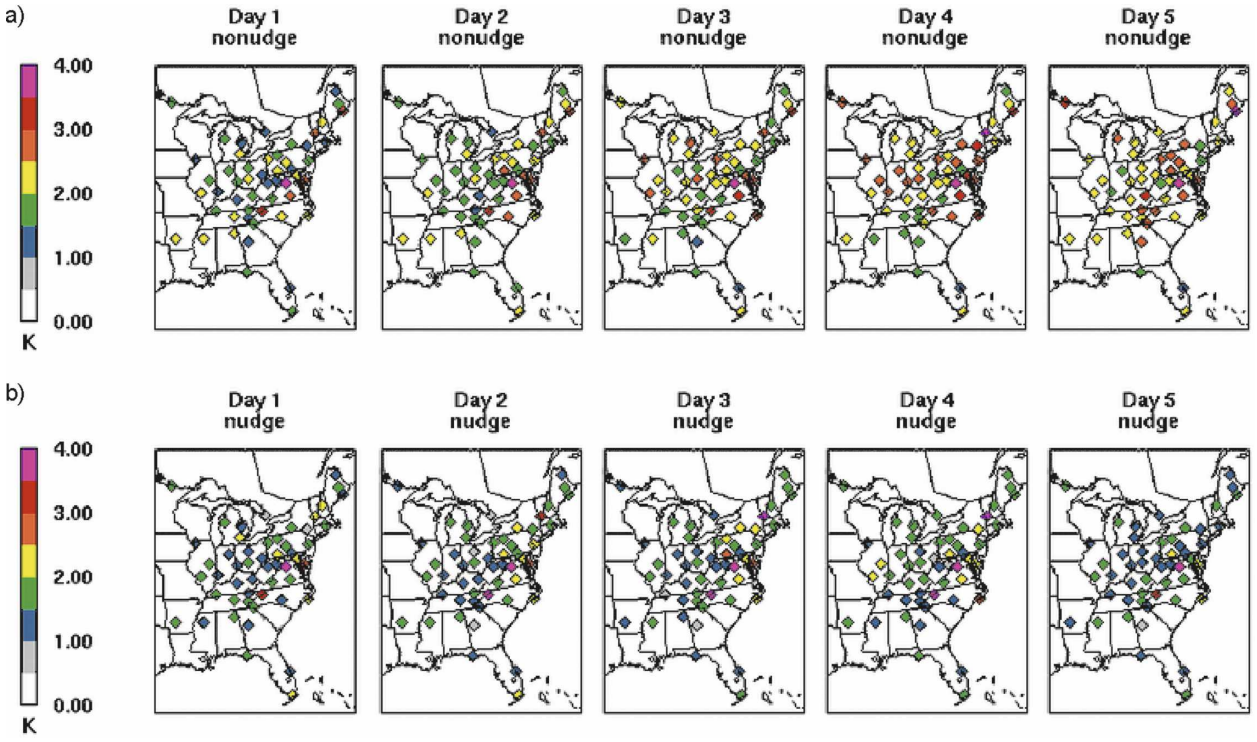

FIG. 2. MAE averaged by day within the MM5 simulation (see Table 1) for hourly 2-m temperature in K at CASTNET sites east of 100°W longitude. (a) Computed for simulation NONUDGE. (b) Computed for simulation NUDGE.

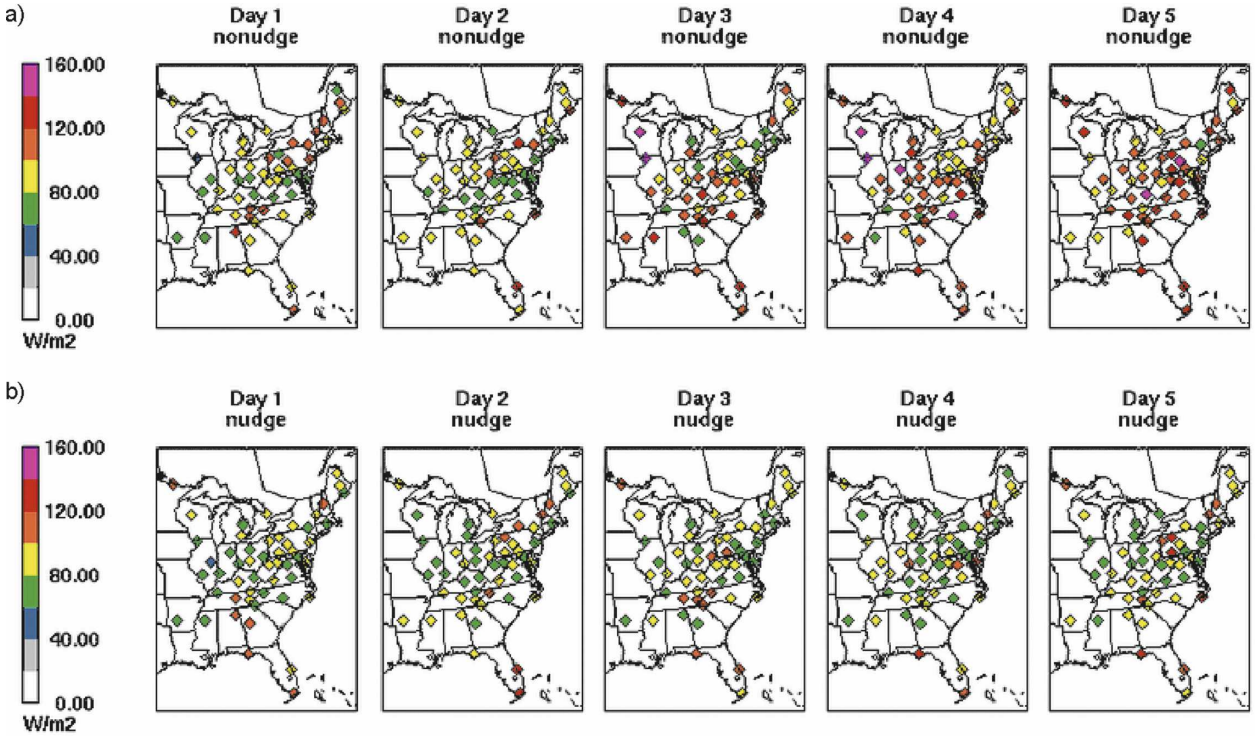

FIG. 3. As in Fig. 2, but for hourly shortwave radiation (W m^{-2}).

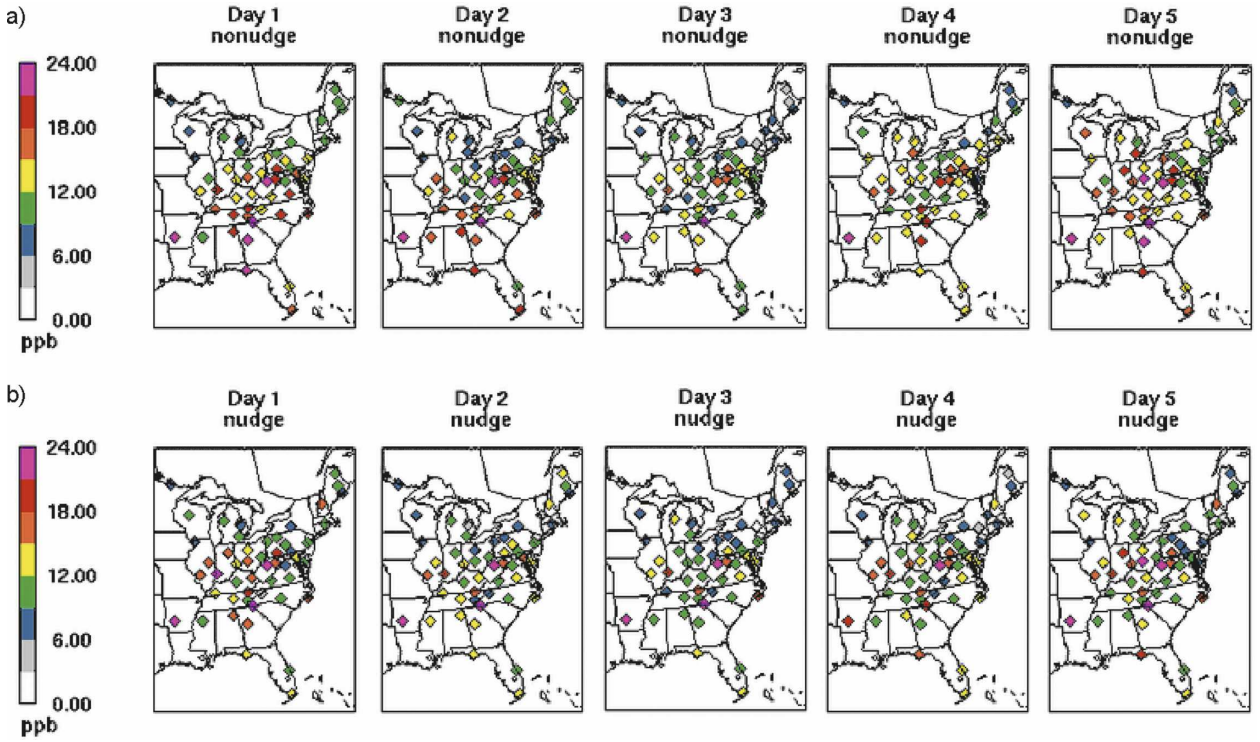

FIG. 4. As in Fig. 2, but for hourly ozone mixing ratio (ppb).

for NONUDGE and NUDGE. There are widespread differences in the average MAE as early as day 1 between NUDGE and NONUDGE. NUDGE generally has at least 20 W m^{-2} lower MAE than NONUDGE (i.e., one color category in Fig. 3) at most CASTNET sites on day 1. In NONUDGE, the average MAE for shortwave radiation grows steadily as MM5 simulation run time increases. By day 5, the average MAE for shortwave radiation is greater than 100 W m^{-2} at more than two-thirds of the CASTNET sites, and it is less than 80 W m^{-2} at only one site. The increased MAE for shortwave radiation in NONUDGE compared with NUDGE suggests that the cloud predictions in MM5 are greatly increasing in error as simulation run time increases in the absence of nudging. However, in NUDGE, the average MAE does not appear to grow or diminish spatially with increased MM5 simulation run time. For each of the 5 days in NUDGE, the average MAE appears to be about 80 W m^{-2} (i.e., roughly the same number of sites shown in green and in yellow each day in Fig. 3), with very few sites (less than 15%) that are more than 20 W m^{-2} (one color category in Fig. 3) from the average MAE. The consistency in the MAE for shortwave radiation in NUDGE as the MM5 runtime increases suggests that the use of analysis nudging for temperature, moisture, and momentum fields can have indirect benefits for simulating clouds and solar radiation, which are not directly forced by analyzed fields in the dynamic analyses.

Subjectively, for shortwave radiation there are four outlier sites that consistently have a high-average MAE relative to the other CASTNET stations in NUDGE and, to a lesser degree, NONUDGE (Fig. 3). Each of the sites is not far from the coast: counterclockwise from the Gulf Coast is Sumatra, Florida (SUM156 on Fig. 1), Everglades National Park, Florida (EVE419 on Fig. 1), Indian River Lagoon, Florida (IRL141 on Fig. 1), and Beaufort, North Carolina (BFT142 on Fig. 1). One possible source of this error is perhaps a poor simulation of shallow cumulus clouds associated with the sea-breeze front formation in the 36-km simulation. Although shallow convection is considered within the cumulus parameterization scheme used for these MM5 simulations, the horizontal grid spacing is too coarse to adequately resolve the sea-breeze front. The local-scale fluctuations in cloud cover with sea breezes (i.e., at the CASTNET site) are unlikely to be reflected in the MM5 simulation at 36 km, which could be a considerable source of error in shortwave radiation, particularly in summertime, at these sites.

Comparisons of average MAE for hourly ozone at the CASTNET sites are shown in Fig. 4. The intraday variations in ozone prediction can be difficult to simulate (e.g., Hogrefe et al. 2000, 2001b, 2006), but hourly

ozone is presented to determine if trends emerge that parallel the meteorological predictions with and without nudging. Overall, the average MAE for hourly ozone is lower by about 2 parts per billion (ppb) in NUDGE than in NONUDGE when compared by day within the MM5 simulation. Unlike the 2-m temperature and shortwave radiation (Figs. 2, 3), the average MAE for hourly ozone does not grow steadily with increased MM5 simulation run time (i.e., from day 1 to day 5) in NONUDGE. However, there is a noticeable decrease in average MAE (i.e., gain in skill) between day 1 and day 3 in NONUDGE, followed by a rise in average MAE (i.e., loss of skill) between day 3 and day 5. On day 1, many of the CASTNET sites have average MAE between 9 and 21 ppb; only four sites are lower and five sites are higher than that range. By day 3, only seven sites have an average MAE that is greater than 15 ppb, and 15 sites have an average MAE that is <9 ppb. The day 5 numerical distribution of average MAE at CASTNET sites in NONUDGE is very similar to day 1, and only seven sites have average MAE outside the range of 9–21 ppb. A similar trend in average MAE from day 1 to day 3 to day 5 is seen in NUDGE, but with a more subtle decrease in average MAE at day 3 and with overall lower average MAE than NONUDGE. Using visual inspection of Fig. 4, there are six CASTNET sites that have comparatively high MAE for hourly ozone in both NUDGE and NONUDGE. Four of the sites are in forested, complex terrain: from southwest to northeast there is Caddo Valley, Arkansas (CAD150 on Fig. 1), Coweeta, North Carolina (COW137 on Fig. 1), Cedar Creek, West Virginia (CDR119 on Fig. 1), and Laurel Hill, Pennsylvania (LRL117 on Fig. 1). The other two (Vincennes, Indiana, VIN140 on Fig. 1; and Speedwell, Tennessee, SPD111 on Fig. 1) are agricultural sites in rolling terrain. At each of these six sites there is a persistent, large overprediction, often by 30–40 ppb, of the nocturnal ozone mixing ratio (not shown), which frequently approaches zero at these sites.

Figures 2–4 suggest that CMAQ's skill at predicting hourly (i.e., intraday) ozone is not well correlated with the skill of MM5 at predicting 2-m temperature or solar radiation at individual CASTNET sites. Interestingly, the outlier sites with relatively high MAE for each of the fields considered in Figs. 2–4 have different physical and chemical issues. For 2-m temperature (Fig. 2), the outlier sites are located at mountaintops. These sites are likely to be unrepresentative of the surrounding areas because they are located near the peak elevation (rather than the average elevation) within the grid cell and the temperature can vary significantly with elevation depending on the atmospheric lapse rate. The outliers for shortwave radiation (Fig. 3) are in coastal zones that are likely to experience subgrid-scale phenomena (i.e., the development of shallow cumulus clouds associated with the sea-breeze front) that are unlikely to be replicated by the MM5 simulation at 36-km horizontal grid spacing, which will directly impact the prediction of shortwave radiation. For hourly ozone, the outlier sites all have a persistent, large overprediction of the nocturnal ozone mixing ratio that appears to be unrelated to the shortcomings in the meteorological fields. Thus, some challenges related to model physics, model chemistry, and scale can overwhelm the influence that nudging alone has on the MM5–CMAQ simulation.

Figures 5–9 examine the behavior of the MM5–CMAQ system at four geographically diverse CASTNET sites. Abington, Connecticut (ABT147 in Fig. 1), which is described by CASTNET as an urban agricultural site in rolling terrain, is selected because it is the only site with at least a partial urban designation for land use, although other sites such as Beltsville, Maryland (BEL116 in Fig. 1), and Washington Crossing, New Jersey (WSP144 in Fig. 1), are often strongly influenced by nearby metropolitan centers. Quaker City, Ohio (QAK172 in Fig. 1), is an agricultural site in rolling terrain. Quaker City is located east-northeast (often downwind) of a high concentration of power plants (emissions sources) in the Ohio Valley. Candor, North Carolina (CND125 in Fig. 1), is a forested site in rolling terrain. Georgia Station, Georgia (GAS153 in Fig. 1), is an agricultural site in rolling terrain.

Comparisons of modeled and observed hourly ozone are shown in Fig. 5 for the four sites listed above. As expected, the predictions for hourly ozone do not match well with the 1:1 line for either NUDGE or NONUDGE at any of the sites. At each of the four sites, there is considerably more scatter in NONUDGE than in NUDGE, suggesting that NONUDGE is a less skillful simulation for this 5-week period. At Abington, there are several hourly ozone measurements that exceed 120 ppb, and all of them are underpredicted by the MM5–CMAQ system in both NUDGE and NONUDGE. At Quaker City, Candor, and Georgia Station, low observations of ozone (i.e., less than 40 ppb) are generally overpredicted in both NUDGE and NONUDGE, often by more than 20 ppb. Appel et al. (2007), using an MM5–CMAQ system with 12-km horizontal grid spacing and an MM5 configuration similar to NUDGE, also show that CMAQ generally underpredicts high values of ozone and overpredicts low values of ozone.

Figures 6–9 show time series of surface and near-surface parameters for NUDGE and NONUDGE through one MM5 run segment (20–24 July 2001) at the

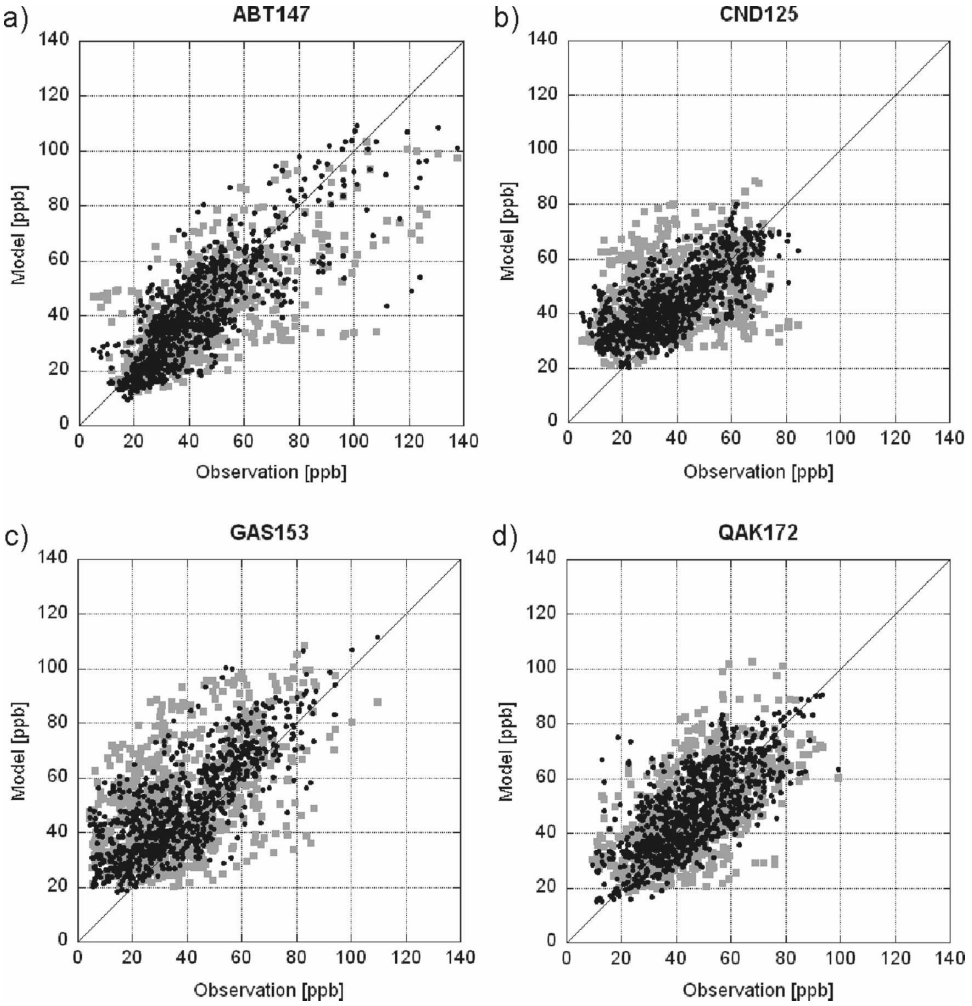

FIG. 5. Comparisons of modeled and observed hourly ozone mixing ratio (ppb) at four CASTNET sites. NUDGE is shown in black circles, and NONUDGE is shown in gray squares. (a) Abington, (b) Candor, (c) Georgia Station, and (d) Quaker City.

four CASTNET sites shown in Fig. 5. At Abington, the 2-m temperature (Fig. 6a) is fairly well predicted by NUDGE through all 5 days, while there are larger errors in NONUDGE, particularly on 23–24 July where there are large underpredictions of the daytime amplitude. In NUDGE, the daily maxima and minima are generally overpredicted by ~1 K, and the maxima are slightly inferior to the predictions by NONUDGE on 20–21 July. The shortwave radiation at Abington (Fig. 6b) is also fairly well predicted in NUDGE. The peak shortwave radiation in NUDGE is well predicted on 21 and 24 July, but is overpredicted by ~50 W m^{-2} on the remaining 3 days, which is consistent with the overprediction in 2-m temperature at this site. The large underprediction of the daytime shortwave radiation (by 400 W m^{-2}) on 23–24 July in NONUDGE is consistent with the underprediction of temperature, suggesting MM5-generated excessive cloud cover on those days. The observed 10-m wind speed (Fig. 6c) is very light throughout the period (generally less than 2 m s^{-1}), but it is substantially overpredicted in both NUDGE and NONUDGE, and it is equally poor in both simulations. The 10-m wind direction (Fig. 6d) is reasonably well characterized by NUDGE throughout the simulation period, whereas it is simulated fairly well by NONUDGE on the first 3 days. The PBL height in NUDGE (Fig. 6e) is typically deeper in NUDGE than in NONUDGE, and the relative differences in PBL height are related to the differences in the development of the daytime 2-m temperature (Fig. 6a). The diurnal ozone cycle at Abington is well simulated throughout the 5-day run segment in NUDGE, with the exception of overprediction during the night and early morning of 21 July (Fig. 6f). NUDGE captured the gradual in-

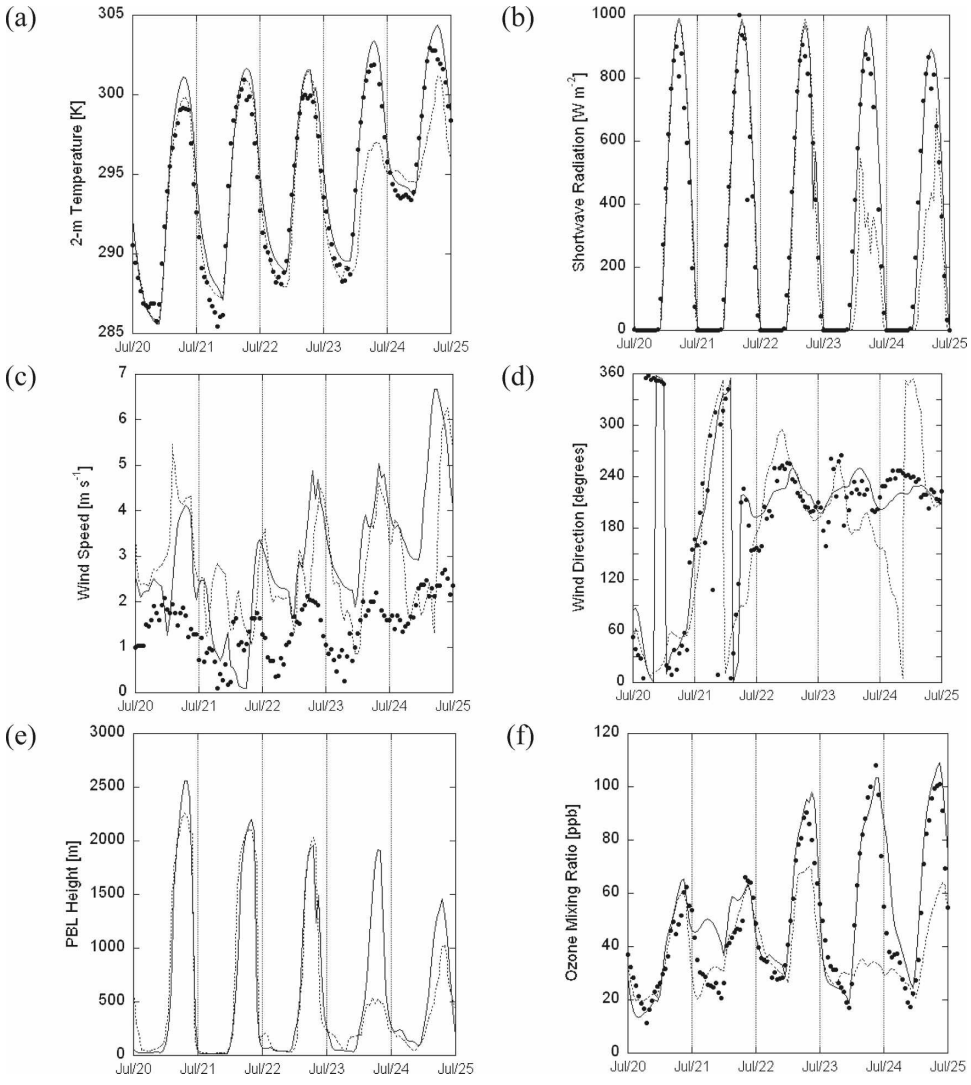

FIG. 6. Time series of hourly observations (circles) and predictions by NUDGE (solid lines) and NONUDGE (dashed lines) at Abington for 0000 UTC 20 Jul–0000 UTC 25 Jul 2001: (a) 2-m temperature (K), (b) shortwave radiation (W m^{-2}), (c) 10-m wind speed (m s^{-1}), (d) 10-m wind direction (°), (e) PBL height (m), and (f) surface ozone mixing ratio (ppb).

crease in peak observed ozone through this run segment, particularly as the peak hourly ozone levels reached or exceeded 85 ppb on the final 3 days. The maximum PBL height in NUDGE gradually decreased by more than 1000 m from 20 to 24 July (Fig. 6e) under largely clear-sky conditions (Fig. 6b), which led to a gradual increase in the ozone mixing ratio (Fig. 6f) as the depth of the vertical column in which the pollutants were mixed became increasingly more constrained. The trend in PBL depth, in conjunction with high near-surface temperatures and clear skies, may have contributed to the high ozone predictions toward the end of the run segment. After the first 2 days of the run segment, the diurnal cycle of the ozone at Abington is not well simulated in NONUDGE, with progressively poorer simulations through the run segment. The substantial underprediction of shortwave radiation in NONUDGE on 23–24 July (Fig. 6b) and the underdeveloped PBL depths (Fig. 6e) are well correlated with the underprediction of ozone on those days (Fig. 6f). Figure 6 shows that for this period, using nudging in MM5 (i.e., NUDGE) can generate a better representation of the meteorological fields, which also resulted in improved ozone predictions when compared with an inferior representation of the meteorological conditions (i.e., NONUDGE) most likely due to erroneous cloud prediction. At this site and for this period, the considerable overprediction of 10-m wind speed did not di-

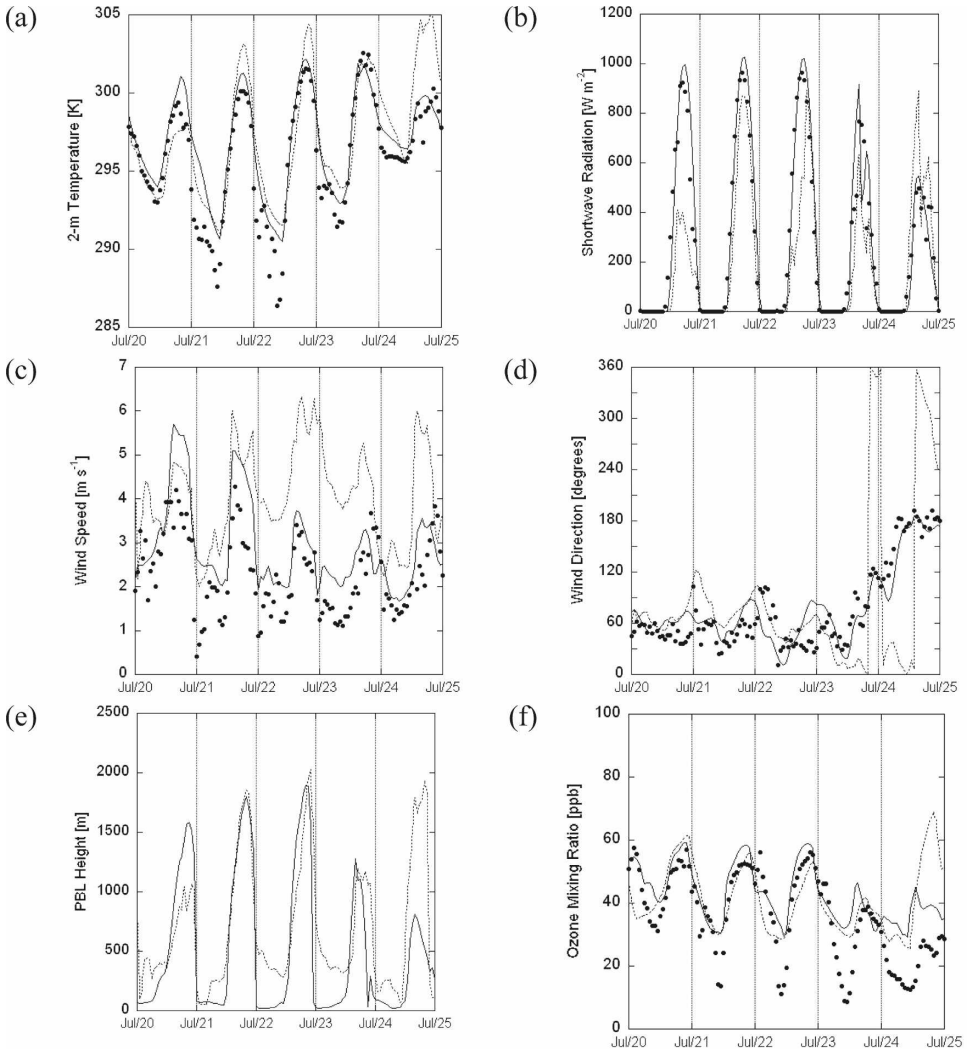

FIG. 7. As in Fig. 6, but for Candor.

rectly or adversely impact the ozone predictions in NUDGE. It is possible that Abington, which is in the urban corridor in the northeastern United States, is less sensitive to errors in wind speed during high ozone periods unlike other (rural) CASTNET sites, or that there are compensating errors in the modeling system.

The observed parameters for Candor are presented in Fig. 7. The 2-m temperature (Fig. 7a) is more erroneous at Candor than at Abington (Fig. 6a). At Candor, the daily minimum 2-m temperature is consistently overpredicted by NUDGE. As expected, the 2-m temperature simulation in NONUDGE is less skillful than in NUDGE. The daily trend in solar radiation at Candor (Fig. 7b) is well simulated by NUDGE, and it is poorly simulated by NONUDGE, which has little resemblance to the observed values on any day. The daily maximum solar radiation is overpredicted at Candor on the clear days (20–22 July) by ~40 W m^{-2} in NUDGE, and MM5 simulated the partial cloudiness on 23 July and cloudiness on 24 July. The good match in the solar radiation prediction in NUDGE on 24 July results in a reasonable prediction of the 2-m temperature on that day, which is 5 days into the dynamic analysis. The 10-m wind speed (Fig. 7c) is overpredicted at Candor, but often by a lesser magnitude for NUDGE than for NONUDGE or at Abington (Fig. 6c). The 10-m wind direction (Fig. 7d) is well simulated throughout the MM5 run segment in NUDGE, while there are large errors in NONUDGE, particularly on 23–24 July. The relative differences in the PBL evolution between NUDGE and NONUDGE (Fig. 6e) are consistent with the simulation of 2-m temperature (Fig. 6a) and shortwave radiation (Fig. 6b). The observed ozone at Candor (Fig. 7f) is somewhat low (peak observations less

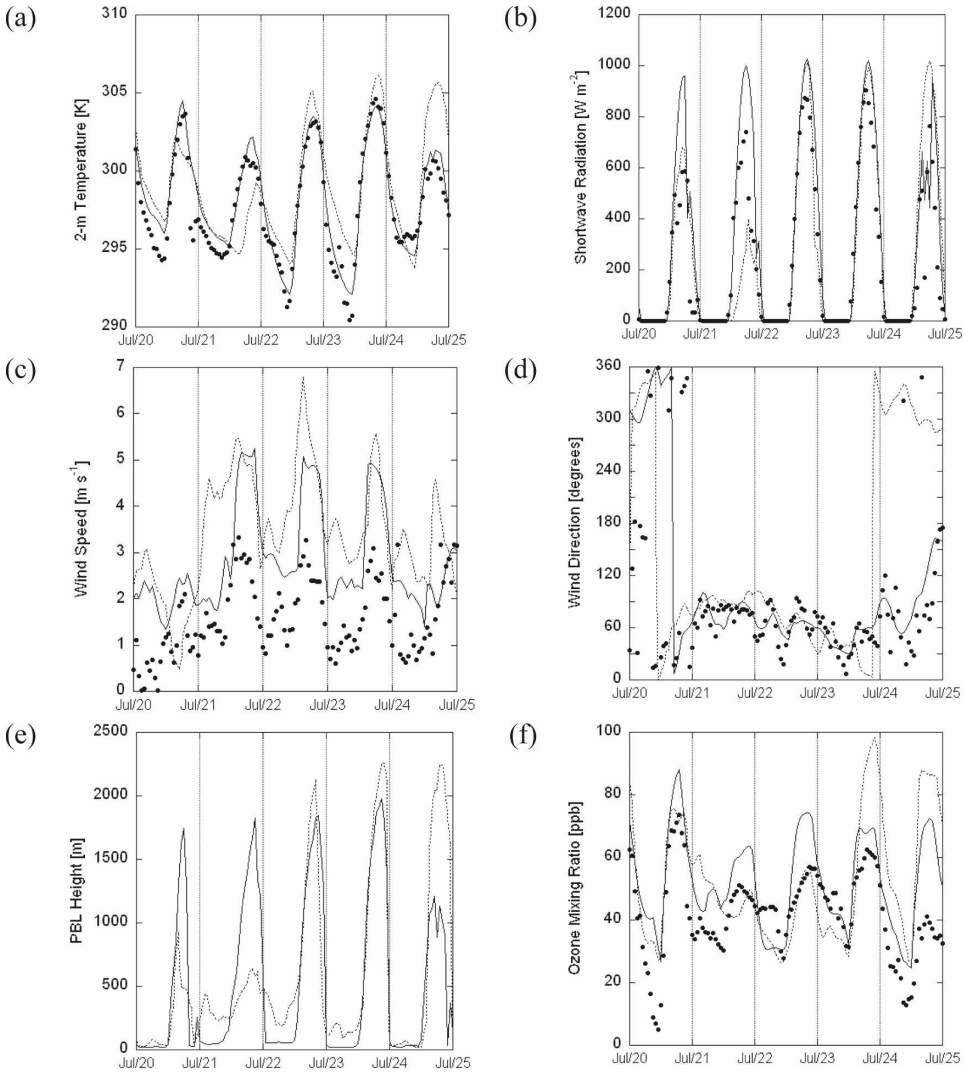

FIG. 8. As in Fig. 6, but for Georgia Station.

than 60 ppb on each day) throughout this period. Neither NUDGE nor NONUDGE adequately simulates the diurnal minimum (less than 20 ppb on the last 4 days), and both overpredict the minimum by ~20 ppb. Although NUDGE correctly simulates the trend toward decreasing peak ozone by 24 July, the diurnal simulation of ozone by NUDGE on 24 July is still nearly 20 ppb in error. The overprediction of 2-m temperature (Fig. 7a) and shortwave radiation (Fig. 7b), and the deep PBL growth on 24 July in NONUDGE, are consistent with the overprediction of ozone (Fig. 7f) on that day. The 2-m temperature, solar radiation, and 10-m wind speed and direction are generally predicted better at the forested site, Candor, by NUDGE than NONUDGE, but neither of the ozone predictions that are forced by those simulations is clearly better for the first 4 days of the MM5 run segment. Again, it should be noted that the ozone observed at rural sites, like Candor, is likely to be transported rather than formed there, so errors in ozone prediction there are sensitive to large-scale errors in the prediction of meteorological fields.

Figure 8 displays the time series at Georgia Station, an agricultural site. As in the predictions at Candor, the daily minimum 2-m temperature (Fig. 8a) is frequently overpredicted by NUDGE and NONUDGE. The diurnal cycle for 2-m temperature is otherwise well simulated by NUDGE through the MM5 run segment, where it is clearly in larger error in NONUDGE. Although the 2-m temperature on 20 July is better simulated by NUDGE than NONUDGE, the solar radiation (Fig. 8b) in NONUDGE is overpredicted by ~50 W m^{-2} while it is substantially overpredicted in NUDGE by 350 W m^{-2}. As with the simulation at

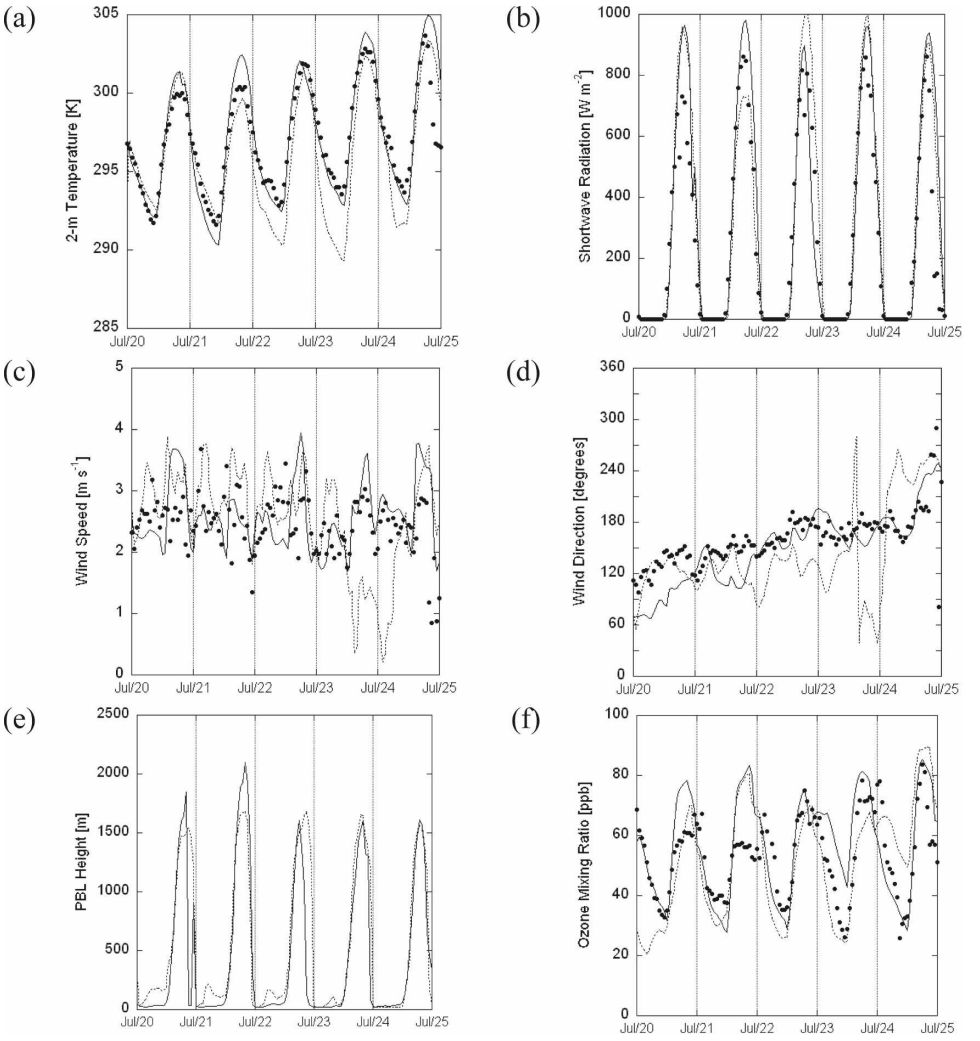

FIG. 9. As in Fig. 6, but for Quaker City.

Candor, the solar radiation prediction suggests that NUDGE simulates well the cloudy period on 24 July. The 10-m wind speed (Fig. 8c) is generally overpredicted in both simulations, but the use of nudging appears to mitigate the error in NUDGE, as at Candor. The 10-m wind direction (Fig. 8d) is also reasonably well captured by NUDGE throughout the simulation, where there are significant errors (approaching 90°) by 24 July in NONUDGE. As expected, the evolution of the PBL (Fig. 8e) is closely related to the 2-m temperature and the solar radiation predictions. For example, the gross overprediction of solar radiation in NUDGE on 20 July resulted in a deeper PBL and an overprediction of ozone (Fig. 8f) on that day. In addition, the missed cloudy period on 24 July in NONUDGE and the overprediction of 2-m temperature led to a deeper PBL and an overprediction of maximum ozone on that day (Fig. 8f). Although there are overall advantages in the meteorological fields in NUDGE as compared with NONUDGE, the ozone prediction at Georgia Station (Fig. 8f) is arguably better in NONUDGE than in NUDGE for the first 3 days of this period (20–22 July). As at Candor, neither simulation correctly simulates the very low (less than 20 ppb) minima that occur at Georgia Station on 20 and 24 July. On 20 July the peak ozone is overpredicted by ~15 ppb in NUDGE, while it is slightly overpredicted (~2 ppb) in NONUDGE, which is consistent with the solar radiation predictions, and thus the deep PBL, rather than the 2-m temperature predictions at this site on this day.

Similar time series comparisons for another agricultural site, Quaker City, are presented in Fig. 9. Surprisingly, the daytime 2-m temperatures (Fig. 9a) are consistently better in NONUDGE than in NUDGE throughout this MM5 run segment; the daily minima are better simulated by NUDGE, as the error in the

minimum 2-m temperature is noticeably larger (~2–4 K) in NONUDGE during the last 3 days of the run segment (22–24 July). Other than on the second and third days (21–22 July), the solar radiation predictions (Fig. 9b) are comparable in both NUDGE and NONUDGE through the MM5 run segment. The 10-m wind speed (Fig. 9c) is simulated comparably through most of the run segment, although there is a clear advantage in NUDGE on 23–24 July. The 10-m wind direction (Fig. 9d) is clearly superior in NONUDGE than in NUDGE on 20 July, but it becomes quite well simulated by NUDGE on the last 3 days (22–24 July) while the error growth through the run segment is noticeable in NONUDGE. The PBL evolution (Fig. 9e) shows that there is a 300-m spike in the growth of the PBL in NUDGE on 20 July as compared with NONUDGE, even though both simulations have comparable 2-m temperatures and shortwave radiation. The maximum PBL depth on 22–24 July is comparable in NUDGE and NONUDGE, but there are differences in the 2-m temperature, shortwave radiation, and ozone (Fig. 9f) between NUDGE and NONUDGE on those days. The ozone prediction (Fig. 9f) on 20 July, the first day of the run segment, is shown to be better in NONUDGE than in NUDGE; this is interesting because the daytime 2-m temperature and shortwave radiation predictions are quite similar and are both overpredicted in NUDGE and NONUDGE, while there is a large difference in PBL height between the two simulations. As noted above, the wind direction is better predicted in NONUDGE through the preceding evening and daytime leading to the peak observed value. If the modest peak ozone (~65 ppb) is observed at Quaker City because of transport on 20 July, it is plausible that the wind direction simulation is critical here.

Figure 10 presents the daily mean bias error (MBE; cf. Willmott 1982) averaged for all hourly observations at the CASTNET sites. MBE is used here because it, as well as MAE, naturally lends itself to model intercomparison, and the statistics are not influenced by the number of and variability of observations (e.g., Willmott and Matsuura 2005, 2006). The 2-m temperature (Fig. 10a) suggests that there is almost always an overprediction (typically ~0.5–1.5 K) in the dynamic analyses (i.e., NUDGE) for this period. In the absence of nudging (i.e., NONUDGE), the magnitude of the daily bias in 2-m temperature fluctuates more erratically (generally between ~−1.0 and 2.0 K) during this summer period. Part I reports a cold (rather than warm) MBE for this period when comparing with the larger network of NWS observations, and a cold MBE (−0.24 K) is shown in Gilliam et al. (2006) for the full summer period using an MM5 simulation that is similar to NUDGE and comparing with NWS observations in the eastern United States. The warm MBE in comparison with CASTNET sites may be generally related to the tendency for MM5 to overpredict the daily minimum temperature at the predominantly rural CASTNET locations (see Figs. 6–9).

Figure 10b shows the MBE in hourly shortwave radiation averaged for all CASTNET sites for each day during the period of interest. The influence of nudging on shortwave radiation predictions in MM5 has not been extensively evaluated elsewhere. When nudging is used to generate dynamic analyses, shortwave radiation is overpredicted, and the daily average MBE is often between 20 and 40 W m^{-2}. In the absence of nudging, the daily MBE is much more variable on a day-to-day basis, ranging from −60 to 50 W m^{-2} in NONUDGE as compared with 15 to 55 W m^{-2} in NUDGE. The MBE (but not necessarily its magnitude) is usually lower in NONUDGE, as nearly one-half of the days have an average MBE that is negative. The tendency to underpredict shortwave radiation suggests that this configuration of MM5 tends to generate too much cloud cover, and that tendency is exacerbated when nudging is not used in MM5. Overprediction of cloud cover in daytime suppresses the model's prediction of ozone, as shown in Figs. 6b,f for NONUDGE. Other configurations of MM5 have also shown a tendency to overpredict cloud cover and cloud lifetime (Chiriaco et al. 2006), and precipitation, particularly in the summertime (O. R. Bullock 2007, personal communication).

The daily bias in hourly 10-m wind speed, averaged at CASTNET sites, is shown in Fig. 10c. Both NUDGE and NONUDGE tend to overpredict wind speed. In NUDGE, the MBE is often between 0.8 and 1.5 m s^{-1}, whereas it is typically higher (~1.3–2.1 m s^{-1}) and more variable between days in NONUDGE. As with 2-m temperature, the 10-m wind speed biases for CASTNET sites do not match well with the summertime MBEs at NWS sites in the eastern United States reported in Gilliam et al. (2006) and Part I, which are ~−0.25 m s^{-1} for nudged simulations. The overprediction of 10-m wind speed can be clearly seen in Figs. 6–8, but it is less prevalent in Fig. 9. Further analysis of the MM5 simulations is needed to determine the source of the persistent overprediction of winds at rural sites.

The daily MBE for hourly ozone, averaged at CASTNET sites, is shown in Fig. 10d. There is an average overprediction of hourly ozone during this period in both NUDGE and NONUDGE. For both simulations, the MBE for hourly ozone typically ranges from 2 to 12 ppb throughout this period. The overprediction of ozone is also shown consistently at three of the four sites in Fig. 5 (Quaker City, Candor, and Georgia Sta-

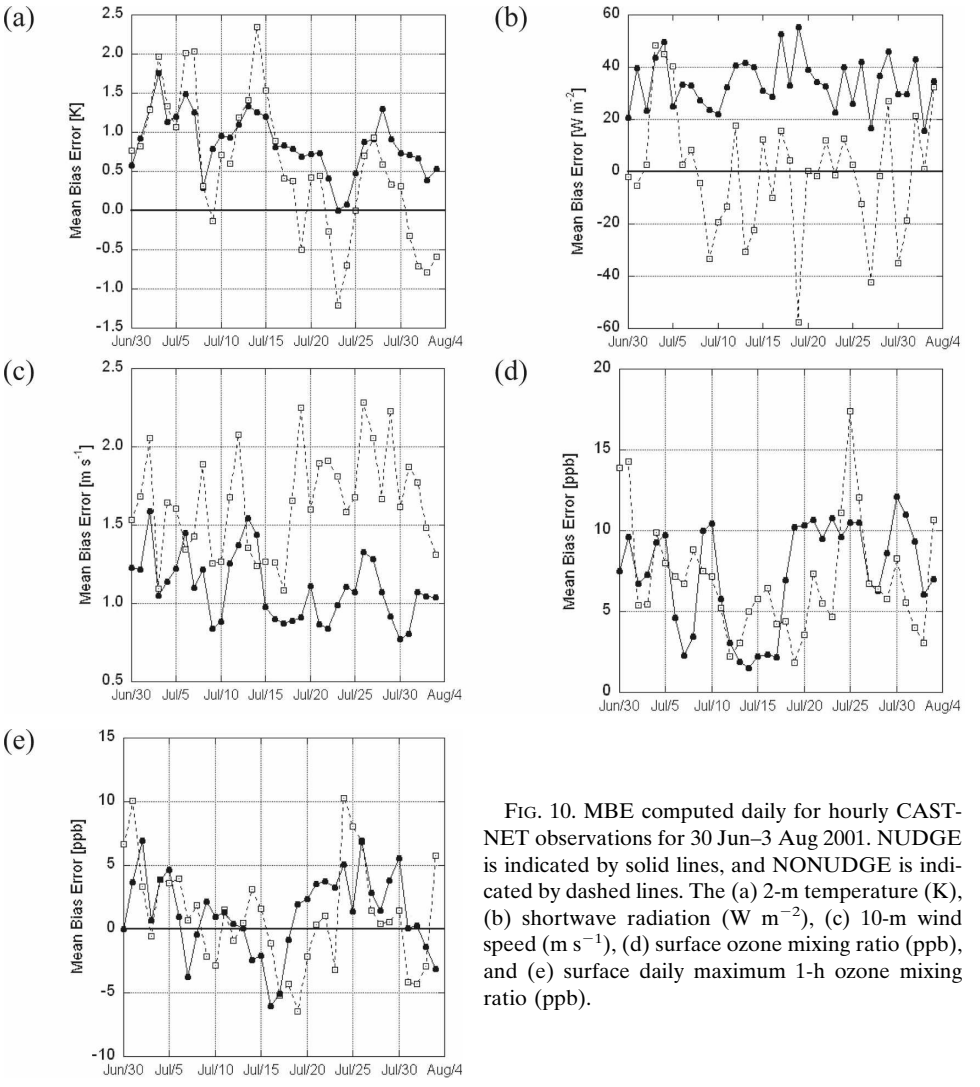

FIG. 10. MBE computed daily for hourly CASTNET observations for 30 Jun–3 Aug 2001. NUDGE is indicated by solid lines, and NONUDGE is indicated by dashed lines. The (a) 2-m temperature (K), (b) shortwave radiation (W m^{-2}), (c) 10-m wind speed (m s^{-1}), (d) surface ozone mixing ratio (ppb), and (e) surface daily maximum 1-h ozone mixing ratio (ppb).

tion) for both NUDGE and NONUDGE. Unlike hourly shortwave radiation and 10-m wind speed, but somewhat like hourly 2-m temperature, NUDGE is not clearly advantageous over NONUDGE for hourly (Fig. 10d) and daily maximum 1-h (Fig. 10e) predictions of ozone. Figures 10d,e, along with Fig. 5, also show that CMAQ does not capture the intraday variability in ozone well [as suggested by Hogrefe et al. (2006)] regardless of whether or not nudging is used. However, using nudging in MM5 improves the day-to-day variability in the ozone predictions by CMAQ, as suggested by the smaller range of MAE in NUDGE compared with NONUDGE (Figs. 10d,e). Overprediction of ozone by CMAQ (Fig. 10d) has also been shown in comparisons against ozone measurements from the EPA's Air Quality System (AQS) network. The positive bias in ozone has been demonstrated at 36-km horizontal grid spacing (Eder and Yu 2006) with input from MM5 simulations similar to NUDGE and in forecasts with 12-km horizontal grid spacing (Eder et al. 2006) that used meteorological fields from the NWS operational meteorological model. Most commonly, the overprediction of ozone reflected by the MBE is from the tendency to overpredict at the low ozone range including the nighttime minimum values (see Figs. 6–8).

Figure 11 shows the MAE averaged by day within the MM5 run segment (see Table 1) for several surface and near-surface parameters for NUDGE and NONUDGE. The shortwave radiation (Fig. 11a) suggests that the MAE is fairly constant (~80 W m^{-2}) with increasing simulation run time in NUDGE, but it increases steadily from ~85 to 110 W m^{-2} in NONUDGE (see also Fig. 3). The average MAE for daily maximum shortwave radiation (Fig. 11a) is fairly

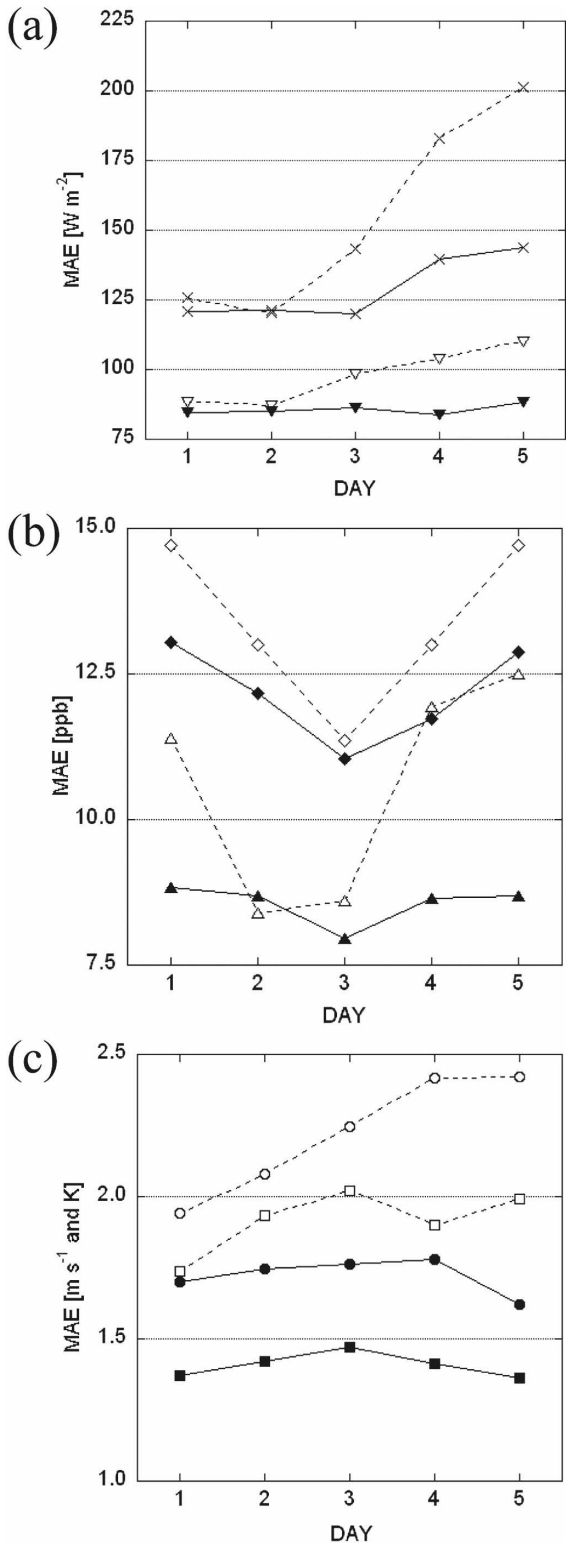

FIG. 11. MAE computed "by day" within the MM5 simulation for 30 Jun–3 Aug 2001. NUDGE is indicated by solid lines, and NONUDGE is indicated by dashed lines. (a) Hourly shortwave radiation (inverted triangles) and maximum shortwave radiation (cross hatches) (W m^{-2}), (b) surface hourly ozone (diamonds) and daily maximum 1-h ozone (triangles) mixing ratio (ppb), and (c) 2-m temperature (circles) (K) and 10-m wind speed (squares) (m s^{-1}).

constant during the first 3 days of the MM5 run segment in NUDGE (~120 W m^{-2}) before rising to ~140 W m^{-2}. In NONUDGE, the MAE increases slightly from ~125 to 145 W m^{-2} from day 1 to day 3, and it climbs dramatically to ~200 W m^{-2} by day 5. The MAE trends for hourly shortwave radiation by day for NUDGE and NONUDGE are similar to the trends reported for 2-m temperature and 10-m wind speed in Part I. For most meteorological fields, the use of nudging to generate dynamic analyses can, on average, generate stable error patterns throughout the 5.5-day simulation period for surface shortwave radiation, while the error growth is substantial in the absence of nudging.

The MAE for hourly ozone and daily maximum 1-h ozone in NUDGE and NONUDGE is shown by day in Fig. 11b. In NUDGE, the MAE drops steadily from 13.0 to 11.0 ppb from day 1 to day 3, and then it rises steadily to 12.9 ppb on day 5. In NONUDGE, the trend is more pronounced; the MAE decreases from 14.7 to 11.4 ppb from day 1 to day 3, and then it increases to 14.7 ppb by day 5. The average MAE for daily maximum 1-h ozone is also fairly constant in NUDGE (~8.0–8.7 ppb) through the simulation period, but the average pattern also mimics the by day MAE for hourly ozone with the minimum MAE on day 3. There is a larger variation in the average MAE for daily maximum 1-h ozone in NONUDGE, with day 1 MAE of 10.6 ppb, followed by a substantial decrease in error to 9.2 ppb on day 2, a subtle decline to 8.9 ppb on day 3, and larger errors on days 4 and 5 (11.5 and 12.3 ppb, respectively).

Each of the 35 days in the study period has an MAE for hourly ozone that is greater than 5 ppb in both NUDGE and NONUDGE (not shown). Collectively day 3 is much more consistent through the study period (NUDGE ranges between 6 and 9.5 ppb, and NONUDGE ranges between 6 and 11 ppb). The spread of MAE in both NUDGE and NONUDGE for the other 4 "days" is greater; the MAE is as much as 12 ppb for NUDGE and as high as 17 ppb for NONUDGE. There are at least two dates with higher MAE for the other 4 days than the worst date in day 3 for NUDGE and NONUDGE, except day 2 in NUDGE, which only has one date higher than the worst day 3. The relative minimum MAE for hourly and daily maximum 1-h ozone on day 3 (Fig. 11b) is not necessarily related to isolated cases of poor performance on other days or excessively good performance on day 3, but rather it is due to the consistency and smaller spread of MAE on day 3 (as compared with the other days) through the study period.

It is speculated that the poorer skill for ozone prediction, on average, at the beginning of the MM5 run segment is related to the accumulation of error in the ozone prediction through the previous run segment, which requires about one ozone time scale (~1–2.5 days; Rao et al. 1997) to be corrected. The error in the ozone prediction may be related to errors in other meteorological fields (e.g., clouds and precipitation) that are not directly evaluated in this study. The ozone predictions are also influenced by uncertainties in other parameters (e.g., emissions) and processes (e.g., chemistry), and their effects have not been analyzed here. Nevertheless, Fig. 11b suggests that the MAE for ozone is lower in NUDGE than in NONUDGE throughout the MM5 run segment.

The MAE for 2-m temperature and 10-m wind speed is shown in Fig. 11c. The MAE for 2-m temperature and 10-m wind speed for NUDGE and NONUDGE at CASTNET sites is comparable to the MAE against NWS sites in Part I. The MAE for 2-m temperature (Fig. 11c) in NONUDGE climbs from 1.94 to 2.42 K. There is also a subtle and steady climb in MAE in 2-m temperature in NUDGE from 1.70 to 1.78 K through day 4, followed by a decrease to 1.62 K on day 5. Part I shows an MAE of 2.1–2.7 K in NONUDGE and a fairly constant MAE of ~1.9 K in NUDGE at the NWS sites. At the CASTNET sites, the MAE for 10-m wind speed varies gently between 1.37 and 1.47 m s^{-1} in NUDGE, while it climbs from 1.73 m s^{-1} on day 1 to 2.00 m s^{-1} on day 5 in NONUDGE. The average MAE for 10-m wind speed at NWS sites is 1.31–1.36 m s^{-1} in NUDGE and 1.59–1.84 m s^{-1} in NONUDGE (Part I). The MAE for daily maximum and minimum 2-m wind speed (not shown) also indicates that NUDGE is clearly superior to NONUDGE. The index of agreement (e.g., Willmott 1982) computed at the CASTNET sites for the parameters shown in Fig. 11 also suggests that NUDGE is consistently a better model than NONUDGE (not shown).

Because there is a fairly small sample of CASNET sites used here (as compared with larger national networks, e.g., NWS and AQS used in Part I), there may be a question of statistical significance in the results shown here. For example, Fig. 11b shows that while NUDGE has a lower average MAE than NONUDGE on each of the 5 days, the statistical values are very close on day 3. One test of statistical robustness, the bootstrap (e.g., Willmott et al. 1985), is used to infer details related to the reliability of the statistics shown here and to infer the variability from those values. In the bootstrapping tests, 55 CASTNET sites (i.e., one for each of the actual sites used in the analysis) are randomly selected with replacement (i.e., some sites may be randomly selected multiple times while others may not be selected at all), and the statistics are computed for the new "sample" of CASTNET sites for both NUDGE and NONUDGE. Here 100 bootstrap samples of 55 sites each are selected from the available CASTNET sites for the 5-week period, and the differences in MAE between NUDGE and NONUDGE are computed for each bootstrap sample and presented as a box plot distribution for 2-m temperature, surface shortwave radiation, 10-m wind speed, and surface ozone (Fig. 12). The median values for MAE difference (Fig. 12) match well with the differences in by-day averages computed at the CASTNET sites (Fig. 11). The trends for the meteorological variables clearly indicate the smallest separation of MAE on day 1 and the greatest separation of MAE on day 5, indicating that NUDGE has, on average, an increasingly lower error than NONUDGE with increased MM5 simulation run time, which is due to the average gain in error in NONUDGE. For 2-m temperature and 10-m wind speed, none of the bootstrap samples has an average error on any of the days that cross the zero line, which indicates that NUDGE has unanimously lower average MAE for those fields than NONUDGE. In addition, for 2-m temperature (Fig. 12a), there is no overlap in the interquartile range (25th–75th percentiles, shown by the boxes) on consecutive aggregated days, so this suggests that there is a greater than 50% probability that the MAE separation between NUDGE and NONUDGE incrementally grows larger with increased MM5 simulation run time. For shortwave radiation (Fig. 12b), there are very small differences in MAE (less than 5 W m^{-2}) between NUDGE and NONUDGE on days 1 and 2 in the bootstrap samples, which is consistent with Fig. 11a. Although the NUDGE is usually more skillful than NONUDGE throughout the simulation period, the bootstrap samples suggest that there is a small (less than 25%) probability that the day 2 shortwave radiation in NONUDGE will be more skillful than in NUDGE. However, there is a more dramatic difference between NUDGE and NONUDGE in shortwave radiation after day 2 that is consistent with the other near-surface meteorological fields. Although the separation in MAE differences tends to grow with increased simulation run time for 10-m wind speed (Fig. 12c), there is an anomaly on day 4, which is consistent with the trend shown in Fig. 11c.

The bootstrap difference in MAE for NUDGE and NONUDGE is quite different for hourly ozone (Fig. 12d) than for the meteorological fields. As with the meteorological fields, NUDGE almost always has a lower average MAE than NONUDGE when computed

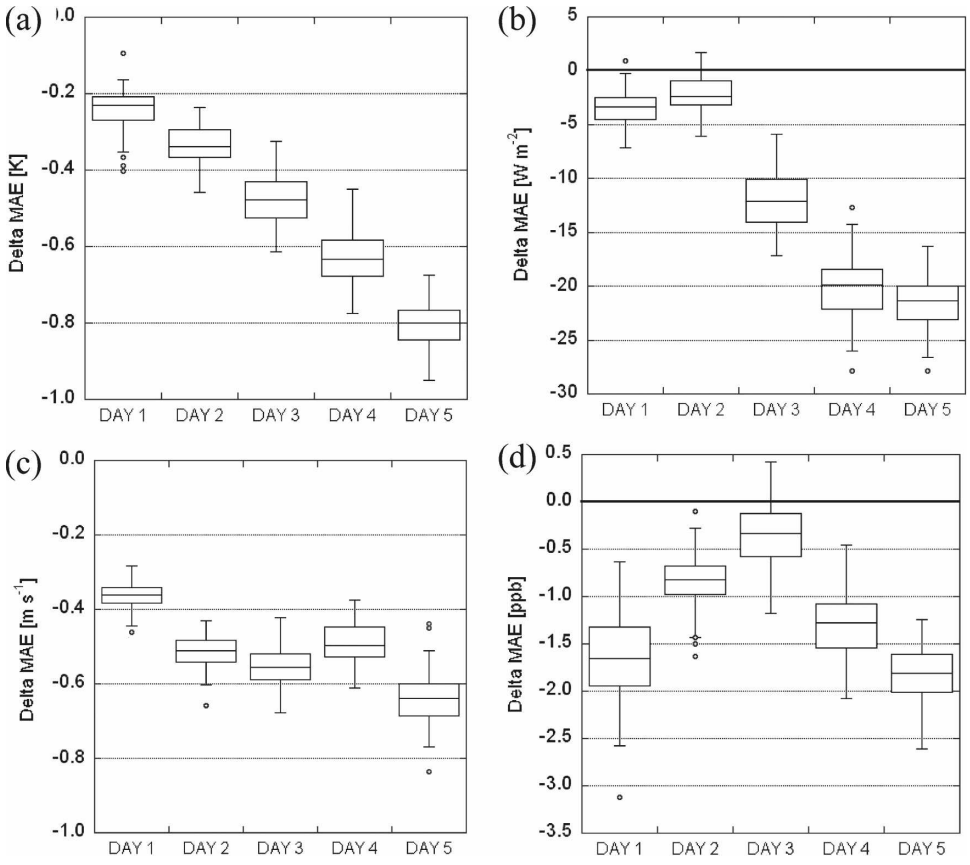

FIG. 12. Box plot representation of differences in MAE (NUDGE minus NONUDGE) "by day" for 100 bootstrap samples of CASTNET data with replacement at 55 CASTNET sites. Boxes indicate the interquartile range (between 25th and 75th percentile), and the median value is shown within the box. The lines that extend from the boxes indicate the maximum and minimum values within a statistically acceptable range. Outliers are shown as values that are outside the box by at least 150% of the interquartile range. The (a) 2-m temperature (K), (b) shortwave radiation (W m^{-2}), (c) 10-m wind speed (m s^{-1}), and (d) surface ozone mixing ratio (ppb).

by day for hourly ozone. But the bootstrap samples indicate that NONUDGE has a lower MAE on day 3 than NUDGE in some (less than 25%) of the cases. In addition, the meteorological fields all suggest that the lowest variability in the MAE difference (i.e., shown by the lowest interquartile range) occurs on day 1. The lowest variability in MAE difference for hourly ozone is on day 2, and day 1 has the greatest variability (largest interquartile range). The bootstrap samples generated similar patterns in the differences in the index of agreement between NUDGE and NONUDGE for the same fields in Fig. 12 (not shown). The MAE trends for hourly ozone shown in Figs. 11 and 12 (and the index of agreement trends) are unrelated to the trends in the meteorological variables that are shown to be generally in lower error with increased simulation run time when nudging is used in MM5. This finding is consistent with Part I for the same simulations when compared with larger observing networks.

4. Summary

This paper provides a further demonstration of the impact of using nudging in the meteorological model on the retrospective air quality simulations using MM5 and CMAQ. A 5-week period is examined by binning the MM5 and CMAQ simulation days according to time elapsed in each of the overlapping 5.5-day MM5 simulation segments. Comparison of hourly surface and near-surface meteorological variables (2-m temperature, surface shortwave radiation, and 10-m wind speed) with measurements from the small, independent CASTNET in the eastern United States shows that, on average, there is no discernable degradation in skill in the MM5 dynamic analyses through the 5.5-day simulation segments. By contrast, the error in the MM5 simulation without nudging grows progressively larger, on average, with increased simulation duration. The MAE for 2-m temperature and 10-m wind speed for

MM5 simulations with and without nudging computed against the CASTNET are similar (i.e., within 0.3 K and 0.2 m s^{-1}) to values shown in Part I for evaluation against NWS observations. Unlike in Part I, there is a warm bias in 2-m temperature, which is particularly evident during nighttime, and a high bias in 10-m wind speed at the CASTNET sites, which are predominantly rural. The MAE for shortwave radiation remains fairly constant (~80 W m^{-2}) when nudging is used throughout the MM5 simulation. The MAE for maximum daily shortwave radiation remains fairly constant (~120 W m^{-2}) during the first 3 days of the run segment before increasing to ~140 W m^{-2} in the last 2 days of the run segment when nudging is used. The MAE for shortwave radiation is larger (~85 and ~125 W m^{-2} for shortwave radiation and daily maximum shortwave radiation on day 1) and grows steadily with increased simulation run time (110 and 200 W m^{-2}, respectively, on day 5) when nudging is not used. The shortwave radiation is generally overpredicted by ~20–40 W m^{-2} in MM5 for this period when nudging is used, but it becomes underpredicted with more day-to-day variability in the absence of nudging. As in Part I for comparisons with the denser network of NWS observations, the MM5 dynamic analyses that are used in this study show no discernable loss of skill with increased simulation run time, on average, in comparisons with the smaller, independent CASTNET, which suggests that 36-km dynamic analyses are reasonable for at least 5.5-day simulations.

Comparisons at the CASTNET sites indicate that using nudging in MM5 to generate dynamic analyses, on average, often leads to a simulation of hourly ozone with lower error. At individual CASTNET sites, there is less scatter in the hourly ozone predictions from CMAQ when nudging is used to control the error in the meteorological fields in MM5. There is a general overprediction of hourly ozone by CMAQ, regardless of whether nudging is used in MM5, which is due to a persistent high bias in the prediction of nighttime ozone mixing ratios and the predominance of low (<40 ppb) ozone mixing ratio observations at the CASTNET sites. On average, the MAE at CASTNET sites is up to 2 ppb lower when MM5 dynamic analyses provide meteorological background for CMAQ. However, the statistical trends for hourly ozone, demonstrated by MAE and index of agreement, do not mimic the trends for the meteorological state variables, regardless of whether or not nudging is used in MM5. The error for hourly ozone on day 1 is higher than on days 2 and 3. Part I showed a similar gain in skill on day 2 when compared with day 1 for daily maximum 1-h ozone. The statistical behavior of the ozone predictions may be related to an accumulation of error in the ozone with increased MM5 run time, which requires ~1–2.5 days (i.e., the ozone time scale) to be corrected in CMAQ after MM5 is reinitialized and the cloud fields are "reset."

The errors in the hourly ozone mixing ratio predictions by CMAQ are not well correlated with individual errors in 2-m temperature, 10-m wind speed and direction, or shortwave radiation. The spatial distribution of MAE in 2-m temperature, surface shortwave radiation, and near-surface ozone is fairly consistent throughout the eastern United States (i.e., error growth is not constrained or accentuated spatially) with the exception of a few outliers for each field that are evident regardless of whether or not nudging is used in the meteorological model. It is also shown that errors in 2-m temperature, 10-m wind speed and direction, and shortwave radiation are all important in different situations to impact the ozone prediction, which is primarily due to transport rather than production, at the CASTNET sites. That is, domainwide error patterns in specific meteorological fields do not directly or systematically translate into error patterns in the prediction of hourly ozone at these sites. However, large broad-scale errors in shortwave radiation predicted by MM5, which suggest erroneous cloud coverage, directly impact the ozone prediction by CMAQ at specific sites. In addition, the use of nudging in MM5 improves the day-to-day variability in the CMAQ prediction of ozone by reducing the range of the daily bias.

The research presented in this paper substantiates the use of nudging in the meteorological model to create dynamic analyses that provide the background for Eulerian chemical-transport models. Because the dynamic analyses are imperfect (i.e., MAE of 80 W m^{-2} in shortwave radiation, 1.4 m s^{-1} in 10-m wind speed, and 1.7 K in 2-m temperature), there will still be errors in the air quality simulation that can be attributed to errors in the meteorological fields, even when nudging is used and the error growth in the meteorological fields is mitigated. As the focus for air quality modeling moves away from episodic simulations and toward annual and multiyear applications (cf. Eder and Yu 2006; Gilliam et al. 2006; Hogrefe et al. 2006; Appel et al. 2007), more emphasis should be placed on determining the best methodology for generating the meteorological fields. In addition, more research is needed to fully investigate the apparent average gain in skill on days 2 and 3 for hourly ozone.

Acknowledgments. This research could not have been performed without leveraging the past and ongoing research of my ASMD colleagues. The MM5 pre-

processing and the MM5 simulations that included the nudging were created by Lara Reynolds (CSC). The MM5-independent emissions files were created by Allan Beidler (CSC), Charles Chang (CSC), and Ryan Cleary (CSC). Guidance on meteorologically dependent emissions processing was generously provided by George Pouliot. Golam Sarwar (EPA/NERL/AMD) created the photolysis files. Harvard University provided the GEOS-CHEM simulations, and Steven Howard created the chemistry lateral boundary conditions files for June and July 2001. Shawn Roselle created the chemistry initial condition file and the CMAQ executable. Christopher Rogers (MACTEC) kindly prepared a chronological file of the CASTNET observations. Valuable discussions with Brian Eder, Jenise Swall, Kristen Foley, Jonathan Pleim, Rohit Mathur, Robert Gilliam, and Annmarie Carlton contributed to this work. The author is grateful for the technical reviews and constructive suggestions from Jonathan Pleim, Rohit Mathur, S. T. Rao, and the anonymous reviewers. The research presented here was performed under the Memorandum of Understanding between the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Commerce's National Oceanic and Atmospheric Administration (NOAA) and under Agreement DW13921548. This work constitutes a contribution to the NOAA Air Quality Program. Although it has been reviewed by EPA and NOAA and approved for publication, it does not necessarily reflect their policies or views.

REFERENCES

Angevine, W. M., and K. Mitchell, 2001: Evaluation of the NCEP mesoscale Eta Model convective boundary layer for air quality applications. *Mon. Wea. Rev.,* **129,** 2761–2775.

Appel, K. W., A. B. Gilliland, G. Sarwar, and R. C. Gilliam, 2007: Evaluation of the Community Multiscale Air Quality (CMAQ) model version 4.5: Uncertainties and sensitivities impacting model performance. Part I—Ozone. *Atmos. Environ.,* **41,** 9603–9615.

Barna, M., and B. Lamb, 2000: Improving ozone modeling in regions of complex terrain using observational nudging in a prognostic meteorological model. *Atmos. Environ.,* **34,** 4889–4906.

Berman, S., J.-Y. Ku, and S. T. Rao, 1999: Spatial and temporal variation in the mixing depth over the northeastern United States during the summer of 1995. *J. Appl. Meteor.,* **38,** 1661–1673.

Biswas, J., and S. T. Rao, 2001: Uncertainties in episodic ozone modeling stemming from uncertainties in meteorological fields. *J. Appl. Meteor.,* **40,** 117–136.

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.

Chiriaco, M., R. Vautard, H. Chepfer, H. Haeffelin, J. Dudhia, Y. Wanherdrick, Y. Morille, and A. Protat, 2006: The ability of MM5 to simulate ice clouds: Systematic comparison between simulated and measured fluxes and lidar/radar profiles at the SIRTA atmospheric observatory. *Mon. Wea. Rev.,* **134,** 897–918.

Eder, B., and S. Yu, 2006: A performance evaluation of the 2004 release of Models-3 CMAQ. *Atmos. Environ.,* **40,** 4811–4824.

——, D. Kang, R. Mathur, S. Yu, and K. Schere, 2006: An operational evaluation of the Eta–CMAQ air quality forecast model. *Atmos. Environ.,* **40,** 4894–4905.

Flaum, J. B., S. T. Rao, and I. G. Zurbenko, 1996: Moderating the influence of meteorological conditions on ambient ozone concentrations. *J. Air Waste Manage. Assoc.,* **46,** 35–46.

Gilliam, R. C., 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, 1994: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398+STR, 138 pp.

——, S. E. Peckham, R. Schmitz, S. A. McKeen, G. Frost, W. C. Skamarock, and B. Eder, 2005: Fully coupled "online" chemistry within the WRF Model. *Atmos. Environ.,* **39,** 6957–6975.

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.

Hogrefe, C., S. T. Rao, I. G. Zurbenko, and P. S. Porter, 2000: Interpreting the information in ozone observations and model predictions relevant to regulatory policies in the eastern United States. *Bull. Amer. Meteor. Soc.,* **81,** 2083–2106.

——, and Coauthors, 2001a: Evaluating the performance of regional-scale photochemical modeling systems: Part I—Meteorological predictions. *Atmos. Environ.,* **35,** 4159–4174.

——, S. T. Rao, P. Kasibhatla, W. Hao, G. Sistla, R. Mathur, and J. McHenry, 2001b: Evaluating the performance of regional-scale photochemical modeling systems: Part II—Ozone predictions. *Atmos. Environ.,* **35,** 4175–4188.

——, P. S. Porter, E. Gego, A. Gilliland, R. Gilliam, J. Swall, J. Irwin, and S. T. Rao, 2006: Temporal features in observed and simulated meteorology and air quality over the eastern United States. *Atmos. Environ.,* **40,** 5041–5055.

Houyoux, M. R., J. M. Vukovich, C. J. Coats Jr., N. M. Wheeler, and P. S. Kasibhatla, 2000: Emission inventory development and processing for the Seasonal Model for Regional Air Quality (SMRAQ) project. *J. Geophys. Res.,* **105,** 9079–9090.

Jacobson, M. Z., 2001: GATOR-GCMM: A global- through urban-scale air pollution and weather forecast model. 1. Model design and treatment of subgrid soil, vegetation, roads, rooftops, water, sea ice, and snow. *J. Geophys. Res.,* **106,** 5385–5402.

Ku, J.-Y., H. Mao, K. Zhang, K. Civerolo, S. T. Rao, C. R. Philbrick, B. Doddridge, and R. Clark, 2001: Numerical investigation of the effects of boundary-layer evolution on the predictions of ozone and the efficacy of emission control options in the northeastern United States. *Environ. Fluid Mech.,* **1,** 209–233.

Lavery, T. F., C. M. Rogers, H. K. Howell, M. C. Burnett, C. A. Wanta, and M. O. Stewart, 2002: Clean Air Status and Trends Network (CASTNet) 2001 annual report. Harding ESE Rep. to the U.S. Environmental Protection Agency, 147 pp. [Available online at http://www.epa.gov/castnet/library/annual01.html.]

Lee, S.-M., S.-C. Yoon, and D. W. Byun, 2004: The effect of mass inconsistency of the meteorological field generated by a com-

mon meteorological model on air quality modeling. *Atmos. Environ.,* **38,** 2917–2926.

Liu, G., C. Hogrefe, and S. T. Rao, 2003: Evaluating the performance of regional-scale meteorological models: Effect of clouds simulation on temperature prediction. *Atmos. Environ.,* **37,** 1425–1433.

Milanchus, M. L., S. T. Rao, and I. G. Zurbenko, 1998: Evaluating the effectiveness of ozone management efforts in the presence of meteorological variability. *J. Air Waste Manage. Assoc.,* **48,** 201–215.

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

Pielke, R. A., and M. Uliasz, 1998: Use of meteorological models as input to regional and mesoscale air quality models—Limitations and strengths. *Atmos. Environ.,* **32,** 1455–1466.

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.

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

Rao, S. T., I. G. Zurbenko, R. Neagu, P. S. Porter, J. Y. Ku, and R. F. Henry, 1997: Space and time scales in ambient ozone data. *Bull. Amer. Meteor. Soc.,* **78,** 2153–2166.

Seaman, N. L., 2000: Meteorological modeling for air-quality assessments. *Atmos. Environ.,* **34,** 2231–2259.

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.

Tanrikulu, S., D. R. Stauffer, N. L. Seaman, and A. J. Ranzieri, 2000: A field-coherence technique for meteorological field-program design for air quality studies. Part II: Evaluation in the San Joaquin Valley. *J. Appl. Meteor.,* **39,** 317–334.

Umeda, T., and P. T. Martien, 2002: Evaluation of a data assimilation technique for a mesoscale meteorological model used for air quality modeling. *J. Appl. Meteor.,* **41,** 12–29.

Willmott, C. J., 1982: Some comments on the evaluation of model performance. *Bull. Amer. Meteor. Soc.,* **63,** 1309–1313.

——, and K. Matsuura, 2005: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Res.,* **30,** 79–82.

——, and ——, 2006: On the use of dimensioned measures of error to evaluate the performance of spatial interpolators. *Int. J. Geog. Info. Sci.,* **20,** 89–102.

——, S. G. Ackleson, R. E. Davis, J. J. Feddema, K. M. Klink, D. R. Legates, J. O'Donnell, and C. M. Rowe, 1985: Statistics for the evaluation and comparison of models. *J. Geophys. Res.,* **90,** 8995–9005.