Available online at www.sciencedirect.com

ScienceDirect

Atmospheric Environment 41 (2007) 9603–9615

ATMOSPHERIC ENVIRONMENT

www.elsevier.com/locate/atmosenv

Evaluation of the Community Multiscale Air Quality (CMAQ) model version 4.5: Sensitivities impacting model performance Part I—Ozone

K. Wyat Appel[a,*,1], Alice B. Gilliland[a], Golam Sarwar[b], Robert C. Gilliam[a]

[a]*Atmospheric Sciences Modeling Division, Air Resources Laboratory, National Oceanic and Atmospheric Administration, U.S. EPA, 109 T.W. Alexander Drive, Research Triangle Park, NC 27711, USA*

[b]*Atmospheric Modeling Division, National Exposure Research Laboratory, US Environmental Protection Agency, Research Triangle Park, NC 27711, USA*

Received 23 March 2007; received in revised form 15 August 2007; accepted 16 August 2007

Abstract

This study examines ozone (O_3) predictions from the Community Multiscale Air Quality (CMAQ) model version 4.5 and discusses potential factors influencing the model results. Daily maximum 8-h average O_3 levels are largely underpredicted when observed O_3 levels are above 85 ppb and overpredicted when they are below 35 ppb. Using a clustering approach, model performance was examined separately for several different synoptic regimes. Under the most common synoptic conditions of a typical summertime Bermuda High setup, the model showed good overall performance for O_3, while associations have been identified here between other, less frequent, synoptic regimes and the O_3 overprediction and underprediction biases. A sensitivity test between the CB-IV and CB05 chemical mechanisms showed that predictions of daily maximum 8-h average O_3 using CB05 were on average 7.3% higher than those using CB-IV. Boundary condition (BC) sensitivity tests show that the overprediction biases at low O_3 levels are more sensitive to the BC O_3 levels near the surface than BC concentrations aloft. These sensitivity tests also show the model performance for O_3 improved when using the global GEOS-CHEM BCs instead of default profiles. Simulations using the newest version of the CMAQ model (v4.6) showed a small improvement in O_3 predictions, particularly when vertical layers were not collapsed. Collectively, the results suggest that key synoptic weather patterns play a leading role in the prediction biases, and more detailed study of these episodes are needed to identify further modeling improvements.
Published by Elsevier Ltd.

Keywords: Air quality model; Community Multiscale Air Quality (CMAQ) model; Model evaluation; Ozone; Synoptic cluster

1. Introduction

Air quality models are needed to assess the efficacy of emission control scenarios for policy decisions, namely, national rulemaking and State Implementation Plans (SIPs). In addition to these regulatory applications, air quality models are now

---

[*]Corresponding author. Tel.: +1 919 541 0757; fax: +1 919 541 1379.
E-mail address: Wyat.Appel@noaa.gov (K.W. Appel).

[1]In partnership with the National Exposure Resource Laboratory, US Environmental Protection Agency, Research Triangle Park, NC 27711, USA.

1352-2310/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.atmosenv.2007.08.044

being used for air quality forecasting (Otte et al., 2005; Wayland and Davidson, 2006). Operational model evaluation approaches are used to first benchmark model performance and identify performance deficiencies, while diagnostic evaluation approaches are used to then identify the particular cause of the performance deficiency with the ultimate goal of improving model performance.

The Community Multiscale Air Quality (CMAQ) modeling system is a leading 1-atm air quality model used for policy and research applications related to ozone (O_3) and aerosols (Byun and Schere, 2006). A series of recent papers present a variety of operational and diagnostic evaluations of the CMAQ version 4.4 (v4.4) model (e.g., Eder and Yu, 2006; Hogrefe et al., 2006; Phillips and Finkelstein, 2006; Gilliland et al., 2006; Swall and Davis, 2006; Tesche et al., 2006). Most relevant to this study is Eder and Yu (2006), which presented statistical performance metrics for an annual simulation (2001) of CMAQ v4.4 with a 36-km horizontal grid. They provided summary statistics for O_3, as well as SO_4^{2-}, NO_3^-, NH_4^+, EC, OC, and $PM_{2.5}$ using data from the various observation networks.

This paper provides an evaluation of the CMAQ version 4.5 (v4.5) model that was released in 2005 for O_3, while part II of this study will present performance results for the organic and inorganic particulate matter (PM) species. For the CMAQ v4.5 results presented here, results are presented from an annual 12-km CMAQ simulation for the Eastern United States. Not all results are presented here, but a full evaluation document is available at http://www.cmascenter.org/help/model_docs/cmaq/4.5/Model_Performance_Evaluation.pdf. The results presented here will show that O_3 model performance for CMAQ v4.5 has similar behavior and biases as seen in previous versions of CMAQ (e.g., Eder and Yu, 2006). An expanded detailed analysis in this study identifies key factors that influence these prediction biases, which is important for prioritizing further diagnostic evaluation for model improvement.

2. Description of CMAQ simulations

In this study, an annual (2001) CMAQ v4.5 simulation was developed with 12 km × 12 km horizontal grid spacing and a 14-layer vertical structure for the domain covering the Eastern United States. The 12-km simulation was nested within a 36 km × 36 km horizontal grid spacing simulation that used the same model configuration as the nested simulation. Boundary conditions (BCs) for the 36-km simulation were provided by a global chemical transport model (GEOS-CHEM) (Bey et al., 2001). The meteorological fields were simulated at both 36 and 12 km (nested within the 36-km simulation) by version 3.6.1 of MM5, the Fifth-Generation Pennsylvania State University/National Center for Atmospheric Research (NCAR) Mesoscale model (Grell et al., 1994). Details regarding the MM5 configuration used for these simulations are described in Gilliam et al. (2006). The MM5 fields were processed for CMAQ using version 3.0 of the Meteorology–Chemistry Interface Program (MCIP).

The CMAQ v4.5 simulations utilized the CB-IV gas-phase chemistry mechanism (Gery et al., 1989), the efficient Euler backward interactive (EBI) solver (Hertel et al., 1993), the AERO4 aerosol module which contains a preliminary treatment of sea salt emissions and chemistry (Bhave et al., 2005; Shankar et al., 2005), the eddy scheme for treatment of the planetary boundary layer (PBL) and the asymmetric convective module (ACM) for cloud treatment in the model (Pleim and Chang, 1992). In addition to the inclusion of sea salt emissions, other significant changes made to CMAQ v4.5 from the previous version of the model include an updated aerosol dry deposition algorithm, an updated version of ISORROPIA (Nenes et al., 1999), an updated calculation of the minimum eddy diffusivity based on the percent urban fraction and other minor corrections for incomplete codes. Additional details regarding the latest release of CMAQ can be found at the website of the Community Modeling and Analysis System (CMAS) center (http://www.cmascenter.org/help/documentation.cfm).

Emissions of sulfur dioxide (SO_2), carbon monoxide (CO), nitrogen oxides (NO_x), ammonia (NH_3), volatile organic compounds (VOCs) and particulate matter ($PM_{2.5}$) were based on the US Environmental Protection Agency's (EPA) 2001 National Emission Inventory (NEI) and were processed by the Sparse Matrix Operator Kernel Emission (SMOKE) processor (Houyoux et al., 2000). Since current NEI NH_3 emissions are limited to annual estimates with no intra-annual variation, the posterior estimated monthly NH_3 emission factors from Gilliland et al. (2006) were used to estimate seasonal variability. Biogenic emissions were processed using the Biogenic Emissions

Inventory System (BEIS) version 3.13 (Schwede et al., 2005), while mobile emissions of CO, VOC, NO_x, NH_3, SO_2 and $PM_{2.5}$ from vehicles were based on MOBILE6 (EPA, 2003).

3. Observational data and comparison methods for evaluation

Model performance is evaluated here against atmospheric concentrations of O_3. For this evaluation, observed concentrations are extracted from the Air Quality System (AQS) network (previously known as the Aerometric Information Retrieval System (AIRS)). Observations and model results were then paired in space (no interpolation) and time with the hourly AQS observational data and maximum daily 1- and 8-h average O_3 values were calculated.

A variety of statistical metrics are used here to compare the observed and predicted surface-level daily maximum 8-h O_3 values, which is the regulatory metric used for the National Ambient Air Quality Standard (NAAQS) for O_3. The mean bias (MB) is used as a measure of model bias. The mean error (ME), root mean square error (RMSE), systematic and unsystematic RMSE ($RMSE_s$ and $RMSE_u$, respectively) and index of agreement (IA; Willmott, 1981) are used as measures of model error. The $RMSE_s$, $RMSE_u$ and IA are defined below as

$$C = a + bC_O, \tag{1}$$

$$\mathrm{RMSE_s} = \sqrt{\frac{1}{N}\sum_1^N (C - C_O)^2}, \tag{2}$$

$$\mathrm{RMSE_u} = \sqrt{\frac{1}{N}\sum_1^N (C - C_M)^2}, \tag{3}$$

$$\mathrm{IA} = 1.0 - \frac{\sum_1^N (C_O - C_M)^2}{\sum_1^N (|C_M - \overline{C_O}| + |C_O - \overline{C_O}|)^2}, \tag{4}$$

where C_M and C_O are modeled and observed concentrations, respectively, $\overline{C_O}$ is the mean observed concentration, a and b are the least squares regression coefficients of C_M and C_O, and N is the total number of model/ob pairs. The $RMSE_s$ represents the portion of the error due to systematic model errors, and the $RMSE_u$ represents random errors in the model or model inputs that are less easily addressed (Delle Monache et al., 2006). The IA is a measure of the degree to which model predictions are free of error (Rao et al., 1985).

4. Results

In Fig. 1a, a comparison of maximum 8-h average O_3 for June through August 2001 is shown for the entire Eastern United States modeling domain. Shown are both summary statistics and categorical predictive skill statistics (Eder et al., 2006), which highlight the ability of the model to capture exceedances (observed daily maximum 8-h average $O_3 \geqslant 85$ ppb). From the domain-wide summary statistics one might conclude the model is performing well, as the MB is small (0.52 ppb), with an ME of 8.9 ppb and IA of 0.84. However, the categorical prediction skill statistics reveals that the model simulation is not capturing O_3 exceedence events very well, indicated by the low probability of detection (POD) and bias of 0.3 (bias values <1.0 indicate underprediction). The performance for the other seasons is similar to the summer, with maximum 8-h average O_3 overpredicted in the winter (December–February) and spring (March–May) and slightly underpredicted in the fall. Fig. 1b shows the median and inter-quartile ranges of hourly average O_3 calculated for each hour of the day for both observed and simulated values for July 2001, showing that the median observed and predicted O_3 correspond well between 10 a.m. and 6 p.m., which are the hours that typically makeup the daily maximum 8-h average for O_3 during the summer. Fig. 1c and d, which shows the MB and RMSE for daily maximum 8-h average O_3 for June–August binned by the observed concentration, highlight the tendency of the model to significantly overpredict low concentrations and underpredict high concentrations of O_3. These biases are consistently evident throughout the modeling domain, as illustrated in the regional time series in Fig. 2.

The following sections investigate potential factors that may be contributing to these model biases at high and low O_3 concentrations, including meteorology, BCs, chemical mechanism, and vertical-layer structure.

4.1. Sensitivity of O_3 prediction biases to synoptic conditions

It is known that meteorology plays a role in determining the O_3 concentrations (e.g., Camalier et al., 2007; Zheng et al., 2007), and the CMAQ

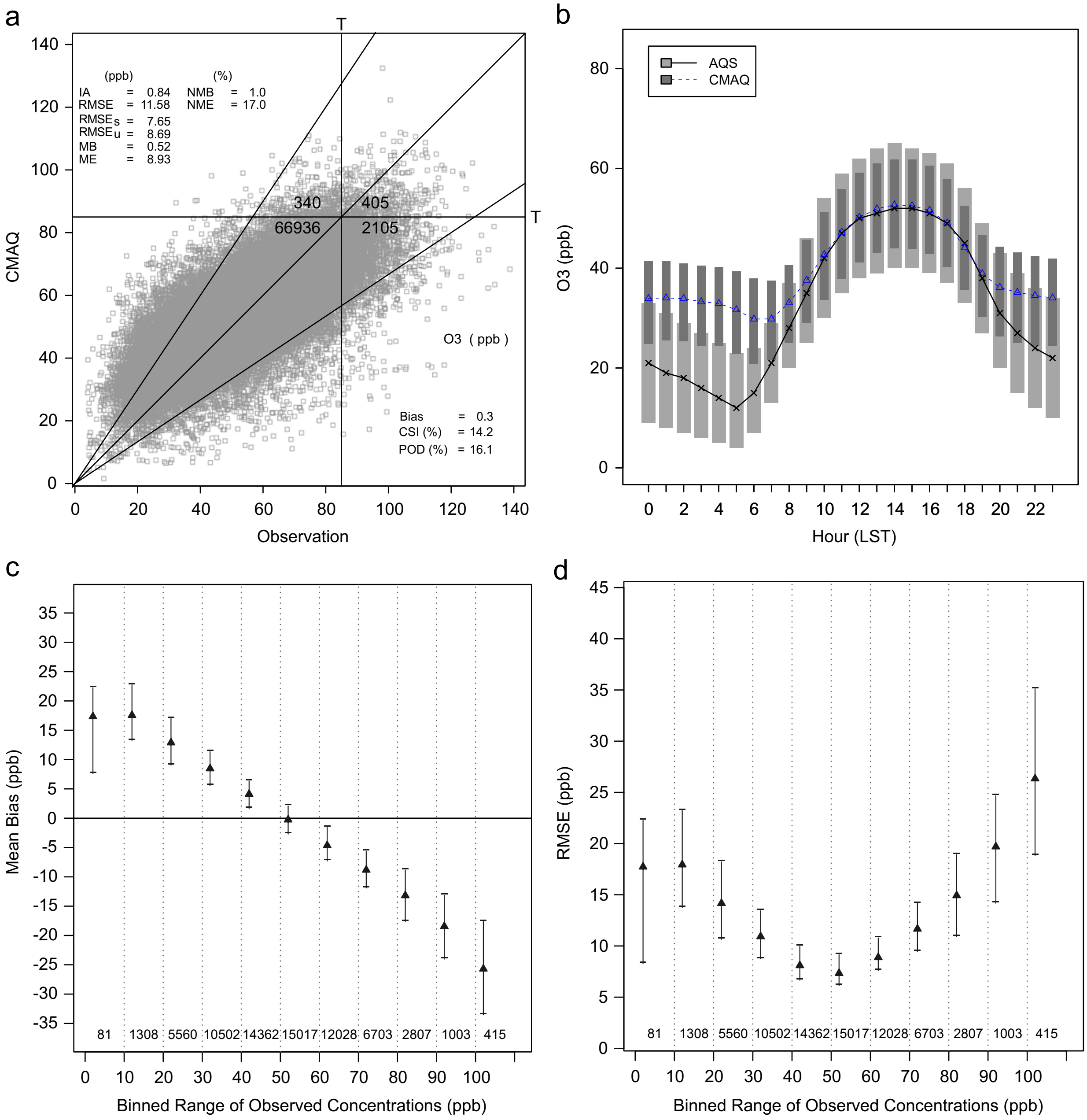


Fig. 1. (a) Scatter plot of daily maximum 8-h average O_3 for all AQS sites in the 12-km domain for June through August 2001. Included are the 1:1 and 2:1 lines, along with several summary statistics. Categorical O_3 predictive skills statistics for bias, critical success index (CSI; %) and probability of detection (POD; %) are shown in the lower right. *T* denotes the threshold value of 85 ppb. (b) Box plot of median (× —AQS; triangle—CMAQ) and inter-quartile ranges (light shading—AQS; dark shading—CMAQ) for hourly average O_3 for July 2001. (c) Median and inter-quartile range of MB binned by observed concentration for June–August 2001 for the CMAQ v4.5 12-km simulation. The number of model/ob pairs for each bin is shown above the *x*-axis. (d) As in (c), except for RMSE.

model performance may vary significantly based on the synoptic conditions. To assess whether an association is evident between different synoptic conditions and the O_3 predictions, model predictions are grouped into synoptic classes. Following Gilliam et al. (2006) and McKendry et al. (1995), synoptic clusters were based on the mean sea level pressure (MSLP) pattern for the Eastern United States (12 km domain) for the period of May through September 2001. Five synoptic clusters

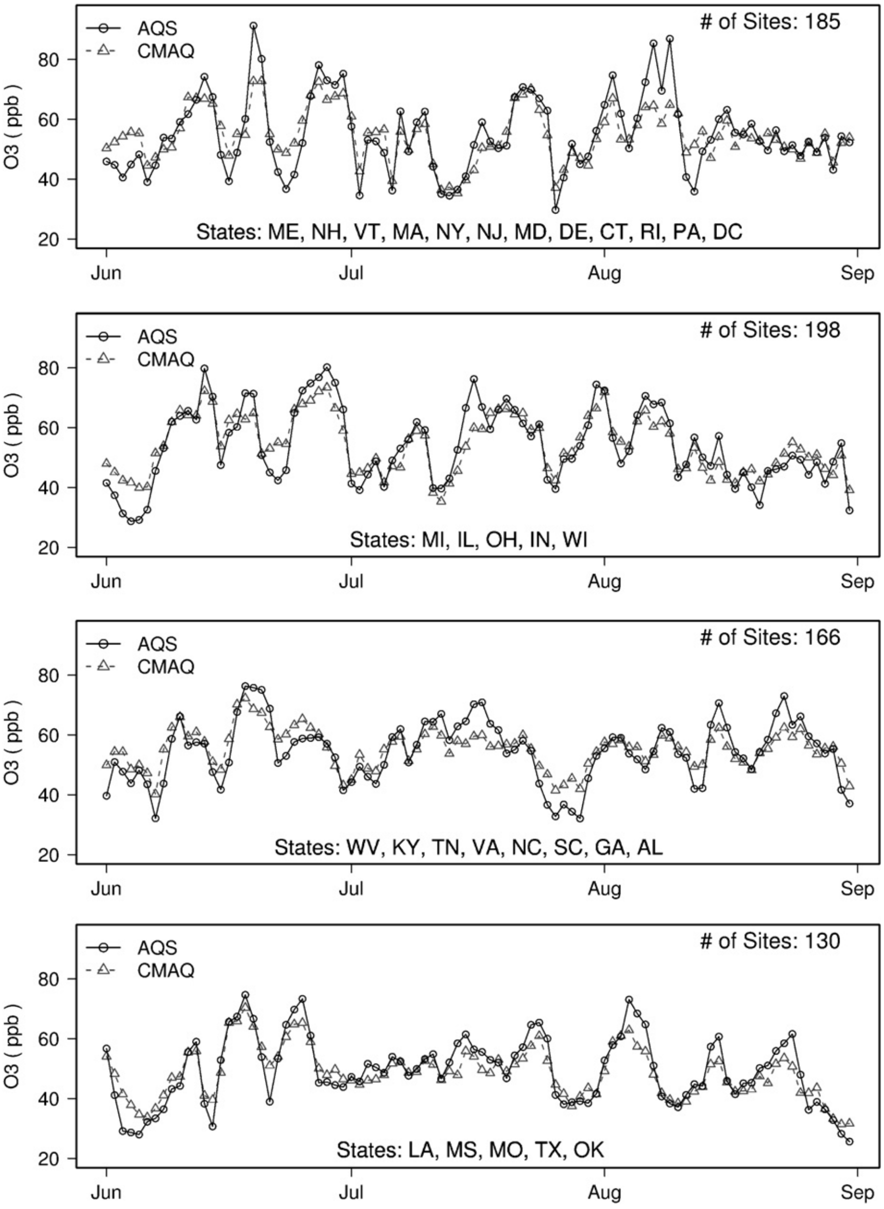


Fig. 2. Time series plots of daily AQS observed (circle, solid black line) and CMAQ predicted (triangle, gray dashed line) maximum 8-h average O_3 for four regions with the 12-km domain. The states comprising each region are shown on each figure.

resulted, ranging in size from 14 days in the smallest cluster to 51 days in the largest cluster.

The synoptic pattern in the Eastern United States during the summer is generally dominated by the Bermuda High, which is a semi-permanent area of high pressure located over the western North Atlantic Ocean (near Bermuda). Variations in the strength and western extent of the Bermuda High can have a significant impact on the weather affecting the Eastern United States during the summer. A strong (high central pressure) Bermuda High with a large western extent often leads to hot and dry (little precipitation) conditions in the Eastern United States during the summer, while a weaker Bermuda High allows for wetter conditions. The synoptic pattern for cluster 1 (Fig. 3a) shows the typical synoptic setup for a Bermuda High, with high pressure over the western North Atlantic Ocean extending over the Eastern United States. This pattern is conducive for high concentrations of

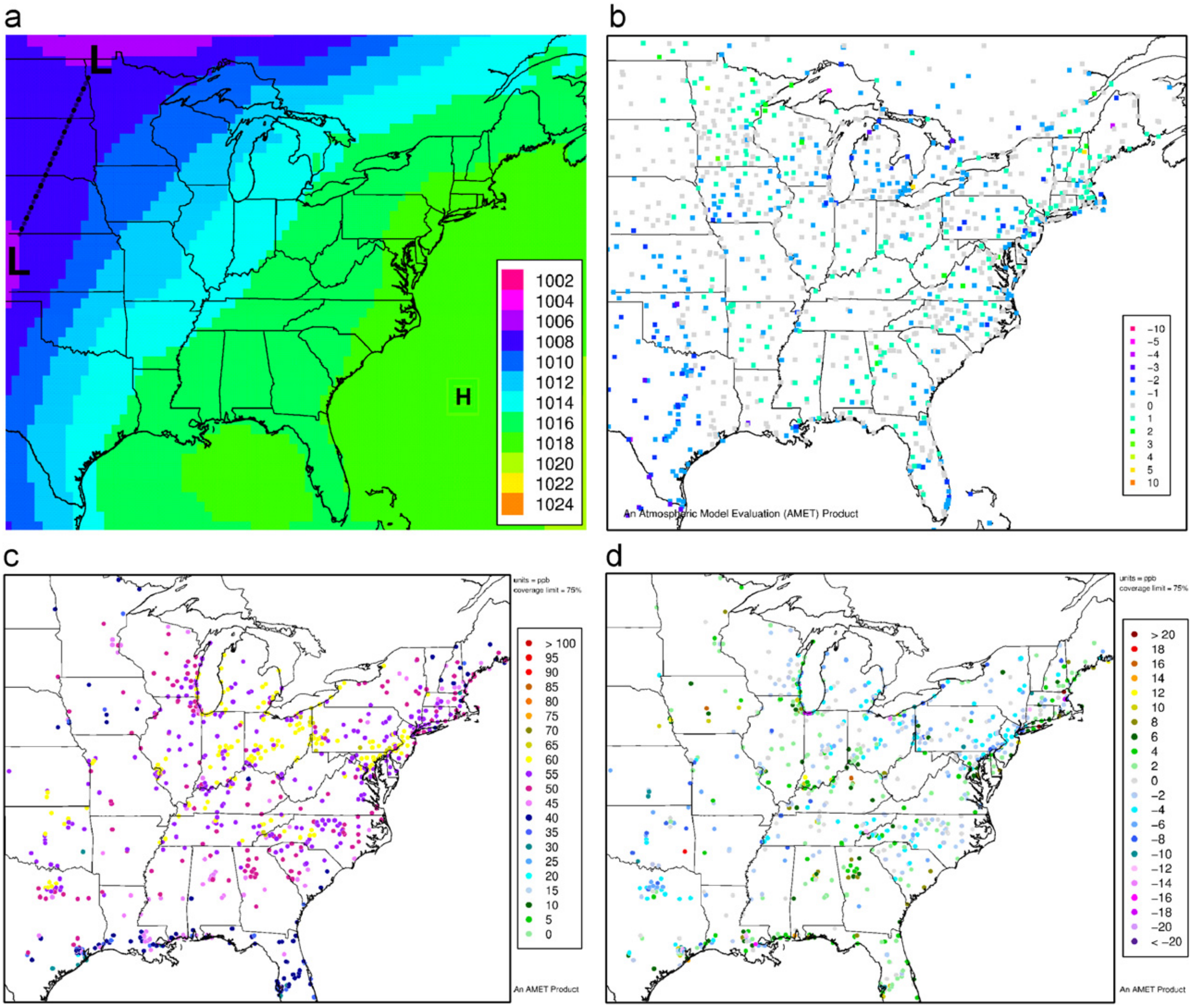


Fig. 3. (a) MSLP (hPa) pattern for those days comprising synoptic cluster 1 for the 12-km domain, limited to the May through September 2001 period. (b) Mean bias of 2-m temperature (°C). (c) Observed maximum 8-h average O_3 (ppb). (d) Mean bias of maximum 8-h average O_3 (ppb).

O_3 in the Eastern United States (Fig. 3c), especially in the northeast where conditions would be expected to be hot and dry. Not surprisingly, this cluster has the largest number of days (51), comprising 33% of days between May and September 2001 and has the highest domain-wide observed mean O_3 concentration of all the clusters (Table 1). The CMAQ O_3 performance for this cluster is relatively good ($\text{ME} = 9.3$ ppb, $\text{NME} = 17.1\%$), with a slight overprediction ($\text{MB} = 1.0$ ppb, $\text{NMB} = 1.9\%$) of daily maximum 8-h average O_3. The MB of 2-m predicted temperature (Fig. 3b) typically ranges between -1 and 1 °C across most of the domain.

The remaining four clusters are all various progressions of a synoptic system (front) moving from northwest to southeast through the domain (this system is apparent in the far northwest portion of the domain in cluster 1). Here, the focus will be on cluster 2 (Fig. 4), which has the largest overprediction of O_3 of all the clusters, and cluster 4 (Fig. 5), which has the largest underprediction of O_3 (Table 1).

Fig. 4a shows the MSLP pattern for cluster 2, in which the dominant Bermuda High present in cluster 1 has retreated to the north and east and is now located over the North Atlantic, while the low-pressure system previously northwest of the Great Lakes region has moved southeast and is now impacting the Great Lakes and midwest regions. As in cluster 1, the MB in 2-m temperature typically

Table 1
Domain-wide statistics for maximum 8-h average O_3 for May through September 2001 (all days) and each of the synoptic clusters (1–5) for the same time period

Cluster	No. of days	Mean (observed)	Mean (model)	IA	MB (ppb)	NMB (%)	ME (ppb)	NME (%)	RMSE (ppb)	$RMSE_s$ (ppb)	$RMSE_u$ (ppb)
All days	153	50.5	51.4	0.83	0.9	2.0	9.0	17.9	11.6	7.6	8.8
1	51	54.2	55.2	0.82	1.0	1.9	9.3	17.1	12.1	7.8	9.2
2	14	45.8	49.5	0.77	3.8	8.2	9.2	20.0	11.5	8.4	7.9
3	28	47.4	49.4	0.83	2.1	4.4	9.0	19.2	11.6	7.9	8.6
4	19	52.0	51.0	0.85	−1.1	−2.0	9.1	17.5	12.0	7.3	9.5
5	15	52.6	52.2	0.80	−0.4	−0.7	10.0	18.9	12.7	9.0	8.9

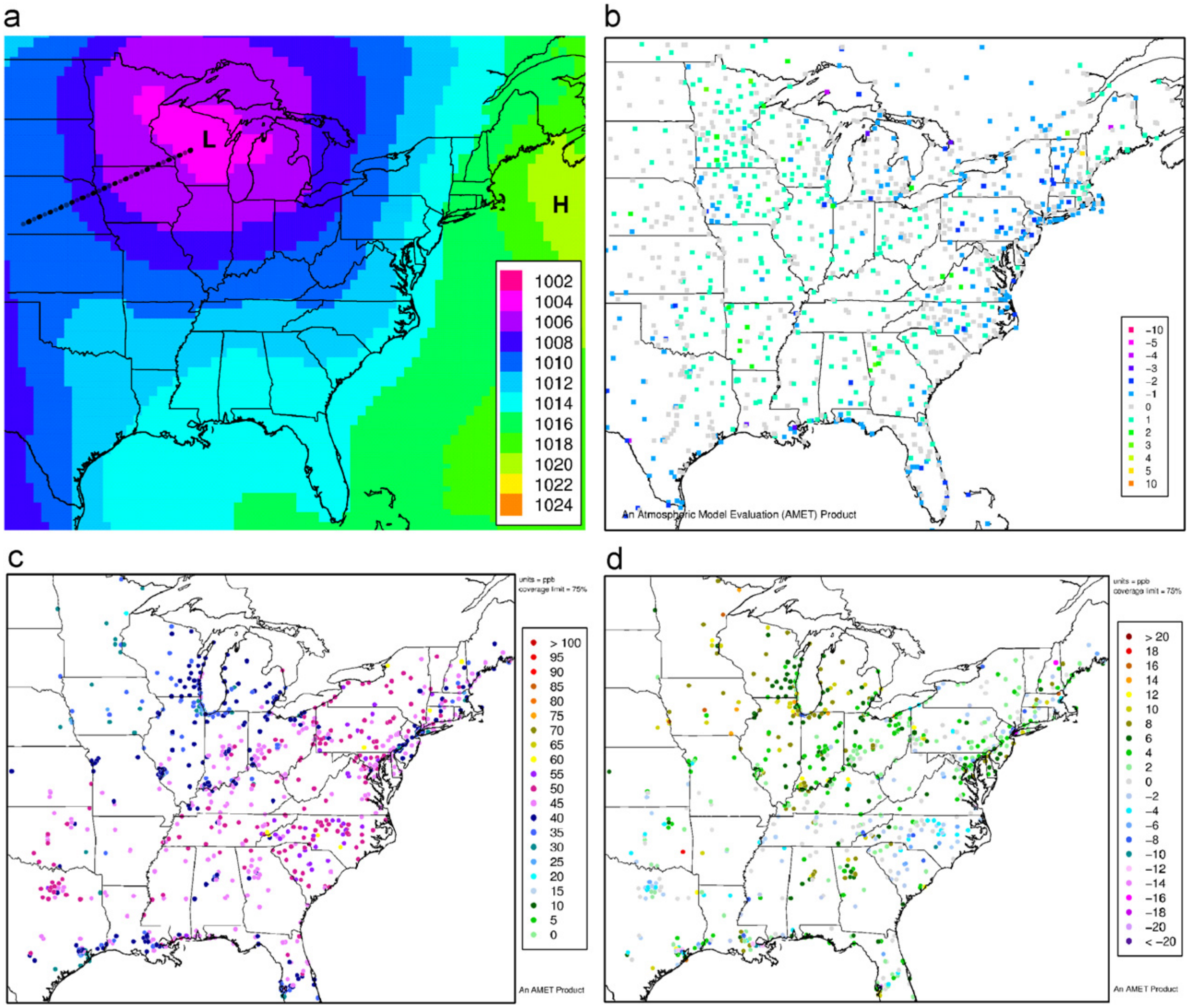


Fig. 4. As in Fig. 3, except for cluster 2.

ranges between −1 and 1 °C across much of the domain. Mean O_3 concentrations are relatively low over those regions (Fig. 4c), while over the northeast and southeast O_3 concentrations remain relatively high, although somewhat lower than in cluster 1 (Fig. 3c). CMAQ performance is poor for this

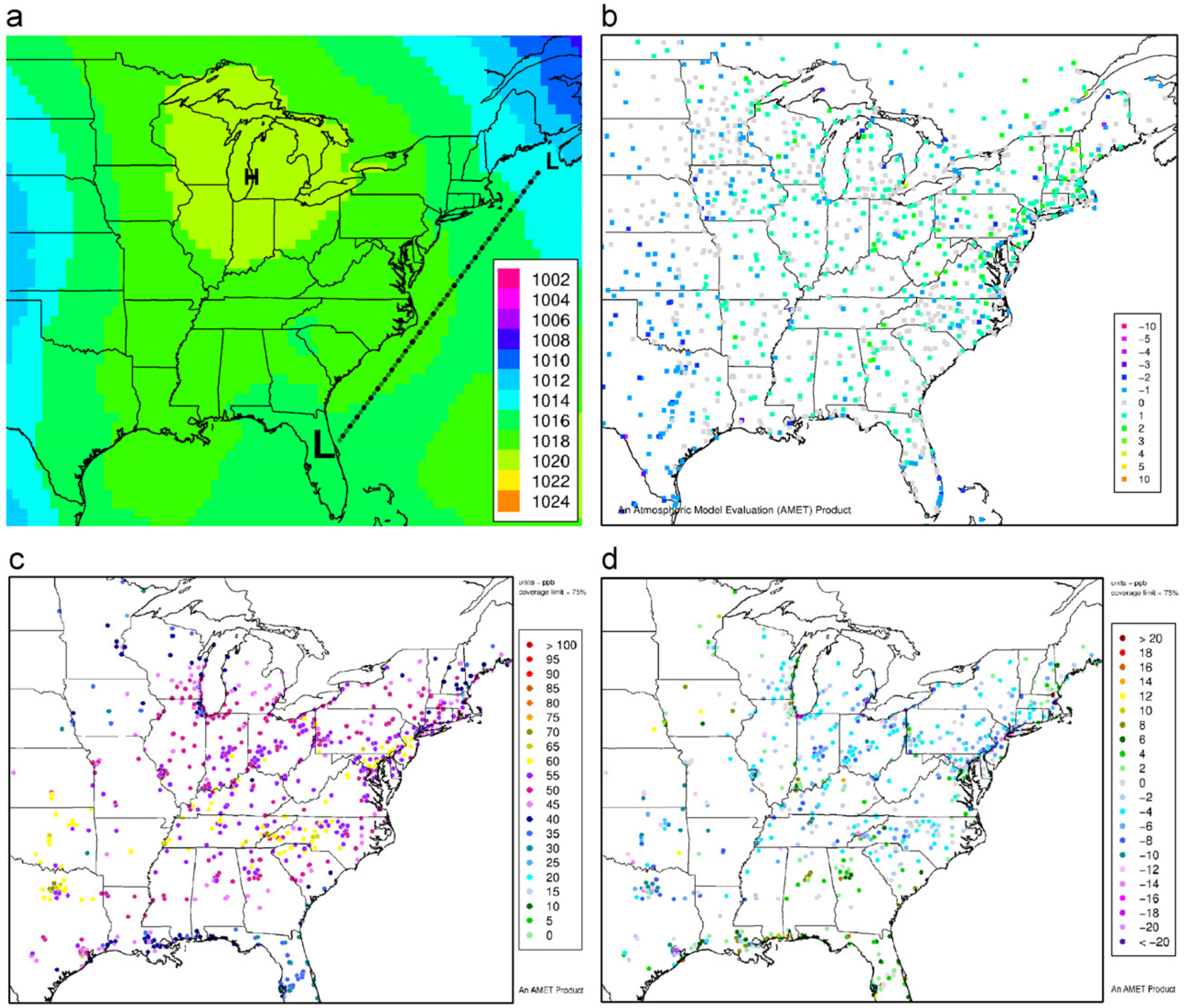


Fig. 5. As in Fig. 3, except for cluster 4.

cluster compared to the other clusters, with an IA of 0.77, MB of 3.8 ppb, NMB of 8.2%, and NME of 20.0%. O_3 is systematically overpredicted in and around the Great Lakes region (Fig. 4d), with biases at some sites >18 ppb. The relatively good performance of the meteorological predictions (based on the temperature bias) suggests that the biases in the O_3 predictions are likely related to the model's ability to simulate O_3 under these synoptic conditions, rather than poor meteorological predictions. However, errors may still exist in the meteorological predictions that are not necessarily reflected in the temperature predictions. This is the smallest cluster with only 14 days, none of which were in July, as it is generally rare to have strong low-pressure systems in the Eastern United States during the peak of the summer. Cluster 3 has a low-pressure system over the extreme northeast, with a cold front oriented northeast to southwest stretching from the northeast to Texas, while high pressure is located over south central Canada and the Caribbean. CMAQ also overpredicts O_3 for cluster 3, although the performance is improved over cluster 2 (Table 1).

Fig. 5a shows the MSLP pattern for cluster 4, in which the low-pressure system has moved off the east coast of the United States into the northern Atlantic and the trailing synoptic cold front has move off the coast into southern Florida and the Atlantic, while high pressure is beginning to build behind the front over the Great Lakes region. The highest observed O_3 concentrations are in southern New Jersey, Maryland, North Carolina, Tennessee and Texas, while the lowest concentrations are in extreme northeast, the upper midwest, and Florida (Fig. 5c). CMAQ underpredicts O_3 under this pattern (MB = −1.1 ppb, NMB = −2.0%, NME 17.5%), with the largest number of sites showing underpredictions in northeast, through the Ohio Valley and into the lower midwest (Fig. 5d). O_3 concentrations are slightly overpredicted in the upper midwest, the southeast and along the Gulf Coast. The MB in 2-m temperature generally ranges between −1 and 1 °C across most of the domain, with the exception of the northeast, where the biases are slightly larger (Fig. 5b). Cluster 5 is similar to cluster 4, except with a stronger (higher central pressure) high located farther east over the Great Lakes region. O_3 is slightly underpredicted for cluster 5 (MB = −0.4 ppb, NMB = −0.7%); however, the error is relatively large (ME = 10.0 ppb, NME = 18.9%) as compared to the other clusters. As in cluster 2, the areas with the largest biases in the O_3 predictions do not correspond to areas with large temperature biases in the meteorological predictions, suggesting again that the biases are likely related to the model's ability to simulate O_3 under these synoptic conditions.

4.2. Sensitivity of O_3 prediction biases to chemical mechanism

For CMAQ v4.5, both the CB-IV and SAPRC99 chemical mechanisms are available. The 2006 release of CMAQ (v4.6) now also includes support for the new CB05 chemical mechanism (Yarwood et al., 2005). Different chemical mechanisms can affect model predictions and performance, even when using the same version of the model, emissions and meteorology. A comparison of two model simulations for July 2001 using CMAQ v4.5 (12-km domain) with both the CB-IV and CB05 chemical mechanism show that hourly O_3 concentrations are about 7–8% higher on average with CB05 for hourly and daily maximum 1- or 8-h O_3 predictions. Fig. 6 includes a comparison of the CMAQ maximum 8-h average O_3 performance for the CB-IV and CB05 simulations, binned by observed O_3 concentration. The MB and RMSE were notably improved at concentrations >60 ppb for the CB05 simulation, while at concentrations <60 ppb the MB and RMSE increase for the simulation using CB05. Simulations using the SAPRC99 chemical mechanism have also shown higher O_3 concentrations than the carbon bond mechanism (Faraji et al., 2007; Byun, 2002). Additional analysis of the impact of the CB05 chemical mechanism on the CMAQ predictions can be found in Sarwar et al. (2007).

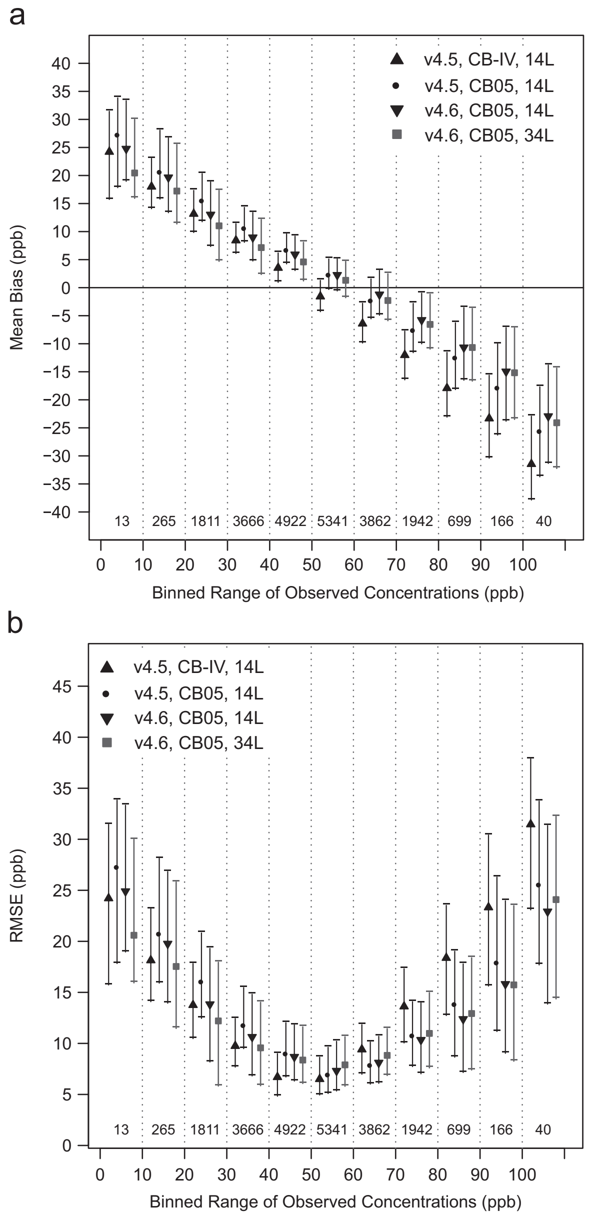


Fig. 6. Median and inter-quartile ranges of daily maximum 8-h average O_3 for (a) mean bias (ppb) and (b) RMSE (ppb) binned by the observed concentration (every 10 ppb) for four CMAQ simulations: CMAQ v4.5, CB-IV, 14 vertical layers (left, point-up triangle); CMAQ v4.5, CB05, 14 vertical layers (left-center, circle); CMAQ v4.6, CB05, 14 vertical layers (right-center, point-down triangle); CMAQ v4.6, CB05, 34 vertical layers (right, square). Shown above the x-axis is the number of model/ob pairs in each bin.

4.3. Sensitivity of O_3 prediction biases to boundary conditions

Currently, chemical fields from the GEOS-CHEM global chemical transport model are being used as BCs for the 36-km CMAQ continental-scale simulation. The 12-km CMAQ simulation then relies on this 36-km domain for BCs to the nested simulation. To test the sensitivity of the O_3 predictions to these GEOS-CHEM BCs, the 14-layer CMAQ 36- and 12-km simulations using CB-IV were repeated for July 2001 using the profile BC profiles that are provided with the CMAQ model code to estimate BCs for the 36-km simulation. O_3 concentrations in the profile BCs range between 30 and 35 ppb at the lowest layer to 70 ppb at the highest layer. In the GEOS-CHEM BCs, the O_3 concentrations range between 20 and 35 ppb at the surface to between approximately 60 ppb (southern boundary) and 200 ppb (northern boundary) at the highest layer. Results from the sensitivity tests indicate that overpredictions of daily maximum 8-h average O_3 below 35 ppb for the 12-km simulation become worse when using the profile BCs, with MB increasing from 6.1 to 8.8 ppb, NMB increasing from 19.9% to 29.2% and RMSE increasing from 13.8 to 15.1 ppb. Since the O_3 concentrations in the BCs are lower near the surface in GEOS-CHEM than the default profile, yet are higher aloft, we conclude that the reduction in the positive bias at low O_3 concentrations is due to higher O_3 concentrations near the surface in the profile BCs than in the GEOS-CHEM BCs. This is in direct contrast to the results of Tong and Mauzerall (2006), in which they concluded that the near surface BCs have a larger impact on the aloft rather than near surface tropospheric O_3 concentrations. The underprediction at O_3 concentrations >85 ppb also becomes worse using the profile BCs, with the MB increasing from -13.5 to -15.8 ppb, NMB increasing from -15.3% to -17.9% and RMSE increasing from 23.4 to 24.6 ppb. The impact from the BCs is greatest along the western boundary of the simulation (both 36 and 12-km). Generally, these results suggest that the simulation with GEOS-CHEM BCs is an improvement over the profile BCs for O_3 predictions.

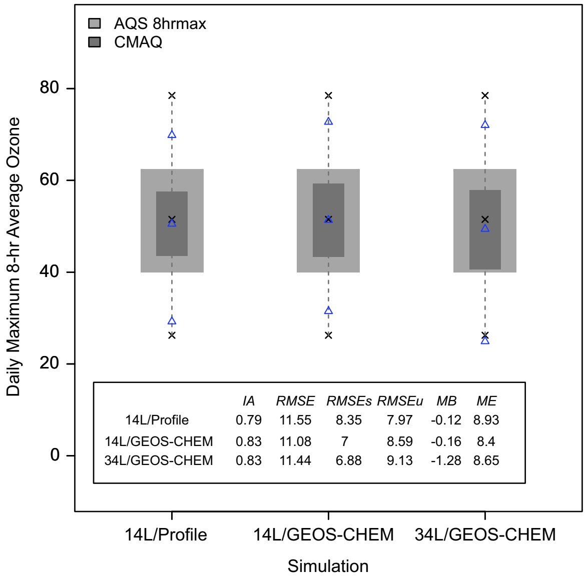
4.4. Sensitivity of O_3 prediction biases to vertical layer collapsing

While the meteorological model simulations usually have 30 or more vertical layers that extend up to approximately 100 hPa, the vertical layers are often collapsed in the middle and upper troposphere to reduce the number of layers for the air quality model simulations. To test the sensitivity of the O_3 predictions to the vertical layer resolution, the CMAQ simulations using CB-IV were again repeated for July 2001 using the full 34-layer vertical structure from the MM5 meteorological simulation, with all other specifications being exactly the same. Note that the top of the lowest vertical layer is the same height ($\sim$39 m) in both the 14- and 34-layer structures.

The results of these sensitivities show that predicted O_3 concentrations typically decrease at low concentrations (observed or predicted concentrations <35 ppb) and increase at high concentrations (observed or predicted concentrations >85 ppb) when layers are not collapsed. The result is a decrease in MB from 6.2 to 4.5 ppb for the low range and -13.2 to -10.2 ppb for the upper range. The change in RMSE for both ranges was <1 ppb. The improvement at low concentrations is likely due to less O_3 aloft reaching the surface when layers are not collapsed. Fig. 7 shows the 5th and 95th percentile, median and inter-quartile range of daily maximum 8-h average O_3 for three CMAQ simulations (July 2001) which compare the influence of 14 versus 34 vertical layers and default versus GEOS-CHEM BCs. Using the same GEOS-CHEM BCs and comparing 14 versus 34 vertical layer CMAQ results, the 34-layer simulation is better able to capture both the low and high end of the O_3 distribution. The comparison further suggests that the 34-layer simulation using GEOS-CHEM BCs has a smaller overprediction bias when O_3 is lower than about 50 ppb because less O_3 aloft is reaching the surface when the upper layers are not collapsed. Overall, the summary statistics are not substantially different among these sensitivity runs; however, the results do show that greater vertical resolution and BCs from GEOS-CHEM improve CMAQ's ability to capture low and high O_3 periods.

4.5. Sensitivity of O_3 prediction biases to CMAQ model version

While the primary goal here is to provide a reference for the CMAQ v4.5 operational performance for O_3 and identify some potential influences on systematic biases, a preliminary comparison is included here of CMAQ v4.5 to the most recent version of the model (v4.6). A new option available



	IA	RMSE	RMSEs	RMSEu	MB	ME
14L/Profile	0.79	11.55	8.35	7.97	-0.12	8.93
14L/GEOS-CHEM	0.83	11.08	7	8.59	-0.16	8.4
34L/GEOS-CHEM	0.83	11.44	6.88	9.13	-1.28	8.65

Fig. 7. Domain-wide observed and simulated 5th percentile, median, and 95th percentile daily maximum 8-h average O_3 (× —AQS; triangle—CMAQ) and inter-quartile range (light shading—AQS; dark shading—CMAQ) for July 2001 for three CMAQ v4.5 simulations: 14 vertical layers and profile BCs (14L/Profile); 14 vertical layers and GEOS-CHEM BCs (14L/GC); 34 vertical layers and GEOS-CHEM BCs (34L/GC). Also shown are the summary statistics of IA, RMSE, $RMSE_s$, $RMSE_u$, MB, and ME for each simulation.

in v4.6 that may impact O_3 predictions is the inclusion of the new ACM2 PBL scheme (Pleim, 2007). Fig. 6 presents a comparison of two simulations for July 2001 using CMAQ v4.6, with one utilizing a 14-layer vertical structure and the other using an uncollapsed 34-layer vertical structure (both simulations used the CB05 chemical mechanism). All other inputs are the same as the CMAQ v4.5 simulations presented here. Comparing the v4.6, 14-layer simulation to the v4.5 simulation using CB05, in which the only difference was the version of the model, O_3 performance (MB and RMSE) is improved with v4.6, although the differences between the two simulations are generally small. The simulation using v4.6 with 34-vertical layers is an improvement over the simulation using 14 layers, particularly at low concentrations (performance for high concentrations is relatively unchanged). The v4.6 simulation using 34 vertical layers is generally comparable to the CMAQ v4.5 using CB-IV at low concentrations, while it represents a significant improvement at high concentrations. These results suggest that the best overall performance for maximum 8-h average O_3 is obtained by using CMAQ v4.6 with the CB05 chemical mechanism and a vertical-layer structure that is not collapsed.

5. Summary

In this paper, we presented an evaluation of O_3 predictions from the Models-3 CMAQ version 4.5, utilizing several model simulations at the 12-km horizontal grid spacing. For the summer period (June–August) of 2001, overall concentrations of daily maximum 8-h average O_3 were found to be nearly unbiased for the entire Eastern United States domain. Further analysis comparing model performance reveals that O_3 is significantly overpredicted for observed concentrations below 35 ppb and underpredicted for concentrations above 85 ppb, with lower bias and error for the O_3 concentration range between 35 and 85 ppb. Consistent with previous evaluations of CMAQ, the model does well in representing the daily O_3 trend but has difficulty capturing the upper and lower ends of the observed concentration range.

Examining model performance under various synoptic regimes identified key meteorological conditions where O_3 prediction biases were most evident. Under the most common Bermuda High pattern, model performance was good, with no widespread systematic biases evident. The model had a large, domain-wide overprediction of O_3 when a low-pressure system was present over the Great Lakes region, while for the pattern where high pressure was building over the same region there was tendency for the model to underpredict O_3 concentrations. This analysis demonstrates that model performance does vary based on the meteorological conditions, and illustrates the importance of considering meteorological conditions when evaluating an air quality model. To further improve O_3 CMAQ predictions, more detailed, diagnostic analysis should focus on these synoptic conditions that are associated with systematic biases in O_3 prediction.

Sensitivity tests presented here have also shown that an updated version of the carbon bond mechanism, released along with CMAQ v4.6 (2006), may alleviate some of the problems of underpredictions at higher O_3 concentrations, as O_3 concentrations are consistently higher with the CB05 chemical mechanism than with CB-IV. However, overpredictions at low concentrations (which were already large) become worse when using CB05.

Increasing the number of vertical layers in CMAQ (by not collapsing vertical layers) and using BCs from the global GEOS-CHEM model reduce systematic biases in the upper and lower range of O_3 levels as compared to the profile BCs included with the standard CMAQ model code. Simulations using the newest version of CMAQ (v4.6) using the CB05 and a new PBL scheme show additional refinement of the maximum 8-h average O_3 predictions, particularly when the vertical layers are not collapsed. While these sensitivity tests suggest that model options can help improve O_3 predictions, the model's ability to capture O_3 responses to key synoptic conditions still appears to be the more dominant influence on model biases at the low and high end of the O_3 concentration range.

Acknowledgments

Thanks to Lucille Bender of Computer Sciences Corporation for help creating these model simulations and to Christian Hogrefe and Brian Eder for their help with the synoptic clustering. The research presented here was performed under the Memorandum of Understanding between the US Environmental Protection Agency (EPA) and the US Department of Commerce's National Oceanic and Atmospheric Administration (NOAA) and under agreement number 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

Bey, I., Jacob, D.J., Yantosca, R.M., Logan, J.A., Field, B.D., Fiore, A.M., Li, Q., Liu, H.Y., Mickley, L.J., Schultz, M.G., 2001. Global modeling of tropospheric chemistry with assimilated meteorology: model description and evaluation. Journal of Geophysical Research 106, 23073–23096.

Bhave, P., Nolte, C., Pleim, J.E., Schwede, D., Roselle, S.J., 2005. Recent developments in the CMAQ modal aerosol module. In: The 2005 Models-3 Users Workshop, Chapel Hill, NC, 26–28 September 2005. Available at: 〈http://www.cmascenter.org/conference/2005/ppt/p17.pdf〉.

Byun, D.W., 2002. A study of photochemical processes of the Houston-Galveston metropolitan airshed with EPA CMAQ. 2002 Models-3 User's Workshop, October 21-23, 2002. EPA, Research Triangle Park, NC USA.

Byun, D., Schere, K.L., 2006. Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. Applied Mechanics Reviews 55, 51–77.

Camalier, L., Cox, W., Dolwick, P., 2007. The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. Atmospheric Environment, in press, doi:10.1016/j.atmosenv.2007.04.061.

Delle Monache, L., Deng, X., Zhou, Y., Stull, R., 2006. Ozone ensemble forecasts: 1. A new ensemble design. Journal of Geophysical Research 111, D05307.

Eder, B., Yu, S., 2006. A performance evaluation of the 2004 release of Models-3 CMAQ. Atmospheric Environment 40, 4811–4824.

Eder, B., Kang, D., Mathur, R., Yu, S., Schere, K., 2006. An operational evaluation of the Eta-CMAQ air quality forecast model. Atmospheric Environment 40, 4894–4905.

EPA, 2003. User's guide to MOBILE6.1 and MOBILE6.2. EPA Office of Air and Radiation, EPA420-R-03-010, Assessment and Standards Division, Office of Transportation and Air Quality, US EPA, 262pp.

Faraji, M., Kimura, Y., McDonald-Buller, E., Allen D., 2007. Comparison of the carbon bond and SAPRC photochemical mechanisms under conditions relevant to southeast Texas. Atmospheric Environment, in press, doi:10.1016/j.atmosenv.2007.07.048.

Gery, M., Whitten, G., Killus, J., Dodge, M., 1989. A photochemical kinetics mechanism for urban and regional scale computer modeling. Journal of Geophysical Research 94, 925–956.

Gilliam, R.C., Hogrefe, C., Rao, S.T., 2006. New methods for evaluating meteorological models used in air quality applications. Atmospheric Environment 40, 5073–5086.

Gilliland, A., Appel, K.W., Pinder, R.W., Dennis, R., 2006. Seasonal NH_3 emissions for the continental United States: inverse model estimation and evaluation. Atmospheric Environment 40, 4986–4998.

Grell, G., Dudhia, J., Stauffer, D., 1994. A description of the Fifth-Generation Penn State/NCAR Mesoscale model (MM5) NCAR Tech. Note NCAR/TN-398 + STR. National Center For Atmospheric Research, Boulder, CO.

Hertel, O., Berkowicz, R., Christensen, J., 1993. Test of two numerical schemes for use in atmospheric transport-chemistry models. Atmospheric Environment 27A, 2591–2611.

Hogrefe, C., Porter, P.S., Gego, E., Gilliland, A., Gilliam, R.C., Swall, J., Irwin, J., Rao, S.T., 2006. Temporal features in observed and simulated meteorology and air quality over the Eastern United States. Atmospheric Environment 40, 5041–5055.

Houyoux, M.R., Vukovich, J.M., Coats Jr., C.J., Wheeler, N.J.M., Kasibhatla, P., 2000. Emission inventory development and processing for the seasonal model for regional air quality. Journal of Geophysical Research 105 (D7), 9079–9090.

McKendry, I.G., Steyn, D.G., McBean, G., 1995. Validation of synoptic circulation patterns simulated by the Canadian climate center general circulation model for Western North America. Atmosphere Ocean 33, 809–825.

Nenes, A., Pandis, S.N., Pilinis, C., 1999. Continued development and testing of a new thermodynamic aerosol module for urban and regional air quality models. Atmospheric Environment 33, 1553–1560.

Otte, T.L., Pouliot, G., Pleim, J.E., Young, J.O., Schere, K.L., Wong, D.C., Lee, P.C.S., Tsidulko, M., McQueen, J.T., Davidson, P., Mathur, R., Chuang, H.Y., DiMego, G., Seaman, N.L., 2005. Linking the Eta model with the

Community Multiscale Air Quality (CMAQ) modeling system to build a national air quality forecasting system. Weather and Forecasting 20, 367–384.

Phillips, S.B., Finkelstein, P.L., 2006. Comparison of spatial patterns of pollutant distribution with CMAQ predictions. Atmospheric Environment 40, 4999–5009.

Pleim, J.E., 2007. A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: Application and evaluation in a mesoscale meteorological model. Journal of Applied Meteorology and Climatology, in press.

Pleim, J.E., Chang, J.S., 1992. A non-local closure model for vertical missing in the convective boundary layer. Atmospheric Environment 26A, 965–981.

Rao, S.T., Sistla, G., Pagnotti, V., Petterson, W.B., Irwin, J.S., Turner, D.B., 1985. Evaluation of the performance of RAM with the regional air pollution study database. Atmospheric Environment 19, 229–245.

Sarwar, G., Luecken, D., Yarwood, G., Whitten, G., Carter, W.P.L., 2007. Impact of an updated Carbon Bond mechanism on predictions from the Community Multiscale Air Quality modeling system: preliminary assessment. Journal of Applied Meteorology and Climatology, in press.

Schwede, D., Pouliot, G., Pierce, T., 2005. Changes to the Biogenic Emissions Inventory System version 3 (BEIS3). In: 4th CMAS Models-3 Users' Conference, Chapel Hill, NC, 26–28 September 2005. Available at: 〈http://www.cmascenter.org/conference/2005/abstracts/2_7.pdf〉.

Shankar, U., Bhave, P.V., Vukovich, J.M., Roselle, J.S., 2005. Implementation and initial applications of sea salt aerosol emissions and chemistry algorithms in the CMAQ v4.5—AERO4 module. In: The 2005 Models-3 Users Workshop, Chapel Hill, NC, 26–28 September 2005. Available at: 〈http://www.cmascenter.org/conference/2005/abstracts/p7.pdf〉.

Swall, J.L., Davis, J.M., 2006. Atmospheric Environment 40, 4883–4893.

Tesche, T.W., Morris, R., Tonnesen, G., McNally, D., Boylan, J., Brewer, P., 2006. CMAQ/CMAx annual 2002 performance evaluation over the Eastern United States. Atmospheric Environment 40, 4906–4919.

Tong, D.Q., Mauzerall, D.L., 2006. Spatial variability of summertime tropospheric ozone over the continental United States: implications of an evaluation of the CMAQ model. Atmospheric Environment 40, 3041–3056.

Wayland, R., Davidson, P., 2006. On alert: air quality forecasting and health advisory warnings. Air and Waste Management Association. Available at: 〈http://www.awma.org/em/pdfs/2006/9/wayland.pdf〉.

Willmott, C.J., 1981. On the validation of models. Physical Geography 2, 184–194.

Yarwood, G., Roa, S., Yocke, M., Whitten, G., 2005. Updates to the carbon bond chemical mechanism: CB05. Final report to the US EPA, RT-0400675. Available at: 〈http://www.camx.com〉.

Zheng, J., Swall, J.L., Cox, W.M., Davis, J.M., 2007. Interannual variation in meteorologically adjusted ozone levels in the Eastern United States: a comparison of two approaches. Atmospheric Environment 41, 705–716.