JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 114, D22304, doi:10.1029/2009JD011898, 2009

Probing into regional ozone and particulate matter pollution in the United States: 1. A 1 year CMAQ simulation and evaluation using surface and satellite data

Yang Zhang,[1] Krish Vijayaraghavan,[2] Xin-Yu Wen,[1] Hilary E. Snell,[3] and Mark Z. Jacobson[4]

Received 11 February 2009; revised 20 September 2009; accepted 2 October 2009; published 25 November 2009.

[1] As part 1 in a series of papers describing long-term simulations using the Community Multiscale Air Quality (CMAQ) modeling system and subsequent process analyses and sensitivity simulations, this paper presents a comprehensive model evaluation for the full year of 2001 over the continental U.S. using both ground-based and satellite measurements. CMAQ is assessed for its ability to reproduce concentrations and long-term trends of major criteria pollutants such as surface ozone (O_3) and fine particulate matter ($PM_{2.5}$) and related variables such as indicator species, wet deposition fluxes, and column mass abundances of carbon monoxide (CO), nitrogen oxides (NO_2), tropospheric ozone residuals (TORs), and aerosol optical depths (AODs). The domain-wide and site-specific evaluation of surface predictions shows an overall satisfactory performance in terms of normalized mean biases for annual mean maximum 1 h and 8 h average O_3 mixing ratios (−11.6 to 0.1% and −4.6 to 3.0%, respectively), 24 h average concentrations of $PM_{2.5}$ (4.2−35.3%), sulfate (−13.0 to 43.5%), and organic carbon (OC) (−37.6 to 24.8%), and wet deposition fluxes (−13.3 to 31.6%). Larger biases, however, occur in the concentrations and wet deposition fluxes of ammonium and nitrate domain-wide and in the concentrations of $PM_{2.5}$, sulfate, black carbon, and OC at some urban/suburban sites. The reasons for such model biases may be errors in emissions, chemistry, aerosol processes, or meteorology. The evaluation of column mass predictions shows a good model performance in capturing the seasonal variations and magnitudes of column CO and NO_2, but relatively poor performance in reproducing observed spatial distributions and magnitudes of TORs for winter and spring and those of AODs in all seasons. Possible reasons for the poor column predictions include the underestimates of emissions, inaccurate upper layer boundary conditions, lack of model treatments of sea salt and dust, and limitations and uncertainties in satellite data.

Citation: Zhang, Y., K. Vijayaraghavan, X.-Y. Wen, H. E. Snell, and M. Z. Jacobson (2009), Probing into regional ozone and particulate matter pollution in the United States: 1. A 1 year CMAQ simulation and evaluation using surface and satellite data, *J. Geophys. Res.*, *114*, D22304, doi:10.1029/2009JD011898.

1. Introduction

[2] Regional ozone (O_3) and fine particulate matter ($PM_{2.5}$) air pollution and associated health effects have been one of the major environmental concerns in the United States (U.S.) and abroad in the past several decades. The effect of these radiatively active species on climate has also been widely examined and reported [e.g., *Intergovernmental Panel on Climate Change* (*IPCC*), 2007]. While significant progress has been made to reduce emissions of their precursors, and their concentrations have been declined steadily in recent years [*U.S. Environmental Protection Agency* (*U.S. EPA*), 2005; *Seinfeld*, 2004; *Seigneur*, 2005], regional O_3 and $PM_{2.5}$ pollution as well as the impacts on climate on local to global scales continue to be a pervasive problem worldwide. They have been a central focus of three-dimensional (3-D) air quality and climate modeling efforts on urban to global scales worldwide in the past half century [*Seinfeld*, 2004; *Seigneur*, 2005; *Horowitz*, 2006; *Zhang*, 2008] and will continue to be key pollutants for modeling future regional/global change into 2100 [*Hogrefe et al.*, 2004; *IPCC*, 2007; *U.S. Climate Change Science Program*,

[1]Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina, USA.
[2]Atmospheric and Environmental Research, Inc., San Francisco, California, USA.
[3]Atmospheric and Environmental Research, Inc., Lexington, Massachusetts, USA.
[4]Department of Civil and Environmental Engineering, Stanford University, Stanford, California, USA.

0148-0227/09/2009JD011898$09.00

2008; *S. Wu et al.*, 2008; *Zhang et al.*, 2008]. Numerous 3-D air quality models have been developed to simulate the formation and fate of O_3 and $PM_{2.5}$ on regional [e.g., *Carmichael et al.*, 1991; *Jacobson*, 1997; *Binkowski and Roselle*, 2003; *Griffin et al.*, 2003; *Zhang et al.*, 2004; *Grell et al.*, 2005] and global [e.g., *Bey et al.*, 2001; *Jacobson*, 2002; *Gong et al.*, 2003; *Easter et al.*, 2004; *Lohmann et al.*, 2007] scales.

[3] For this study, a full year of simulation with the U.S. EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System [*Binkowski and Roselle*, 2003; *Byun and Schere*, 2006] and the process analysis tool in CMAQ is conducted over the continental U.S. (CONUS). Simulated concentrations of gaseous and PM species are evaluated against measurements from ground-based monitoring networks and satellites. The seasonal and annual photochemical characteristics over different regions are examined and the relative contributions of controlling processes are quantified for the formation and fate of key pollutants such as O_3 and $PM_{2.5}$. The effects of regional pollution on the global atmosphere are examined by estimating the export of O_3 and $PM_{2.5}$ pollution from the urban/regional scales to the global environment. Sensitivity simulations are conducted to study the effects of model inputs (e.g., biogenic emissions) and processes (e.g., cloud processes) on major air pollutants, due to their important roles in the fate and formation of these pollutants. Additional sensitivity simulations are conducted under both current and future climate and emission conditions to examine model responses to a changing world. The objectives of this study are to evaluate the model's capability to reproduce regional air pollution, improve the current understanding of regional O_3 and $PM_{2.5}$ pollution in terms of governing chemical reactions and physical processes, and to estimate the potential impacts of the export of O_3, aerosols, and their precursors, such as NO_x and VOCs, from the urban/regional scale on global air pollution. The results are presented in two parts. Part 1 presents the model configurations, evaluation protocols, and evaluation results. Part 2 [*Zhang et al.*, 2009] presents results from process analyses, export calculation, and sensitivity simulations.

2. Model Configurations, Evaluation Protocols, and Observations

2.1. Model Configurations

[4] CMAQ version 4.4 released in October 2004 is applied for the full year of 2001 at a horizontal grid resolution of 36 km. The modeling domain covers the contiguous U.S. and a portion of southern Canada and northern Mexico with 148 × 112 horizontal grid cells. The vertical resolution includes 14 logarithmic structure layers from the surface to the tropopause (~100 hPa or 15.7 km), with a finer resolution in the planetary boundary layer (PBL). The first model layer height is set to be 35 m above the ground level. The meteorological fields, emissions, initial conditions (ICONs), and boundary conditions (BCONs) are provided by the U.S. EPA. The meteorological fields are generated from the Pennsylvania State University (PSU)/National Center for Atmospheric Research (NCAR) Mesoscale Modeling System Generation 5 (MM5) version 3.6.1 with four-dimensional data assimilation analysis nudging. The MM5 hourly output files are processed with the Meteorology-Chemistry Interface Processor (MCIP) version 2.2. The EPA's National Emissions Inventories (NEI) 2001 (also referred to as NEI 1999 Version 3) is used to generate a gridded anthropogenic emission inventory for sulfur dioxide (SO_2), carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO_2), ammonia (NH_3), volatile organic compounds (VOCs), and PM. Emissions of VOCs, CO, NO, NO_2, and PM from mobile sources are generated with the latest on-road motor vehicle emission model, MOBILE6, and biogenic emissions are generated using the Biogenic Emissions Inventory System (BEIS) Version 3.12. The emission inventory is processed with the Sparse Matrix Operator Kernel Emissions system (SMOKE v.1.4). The seasonality of the ammonia emissions is accounted for based on work by *Gilliland et al.* [2003] and *Pinder et al.* [2004]. ICONs and BCONs are generated using the results from a global chemistry model of *Bey et al.* [2001] (GEOS-CHEM). The 1 year simulation is performed continuously with a spin-up period of 10 days (December 22−31, 2000) to minimize the influence of ICONs.

[5] The gas-phase chemistry is based on the Carbon-Bond Mechanism version IV (CBM-IV). The aerosol module AERO3 in CMAQ simulates major aerosol processes including thermodynamic equilibrium for both inorganic and organic PM, binary homogeneous nucleation, coagulation, condensation, PM formation due to aqueous-phase chemistry, aerosol scavenged by cloud droplets, and dry and wet deposition. The particle size distribution is simulated with three lognormally distributed modes: Aitken, accumulation, and coarse mode (corresponding to particles with diameters up to 0.1 μm, 0.1−2.5 μm, and 2.5−10 μm, respectively, for mass distribution). While the modal approach of the PM size representation is subject to inherent limitation from a mathematical point of view and is typically less accurate than the sectional approach [*Zhang et al.*, 1999], it is computationally more efficient and thus is best suited for long-term simulations for 1 year or longer. Sea salt and dust are not treated in this version of CMAQ. CMAQ uses a modified version of the aqueous-phase chemistry of the Regional Acid Deposition Model (RADM). A more detailed description of aerosol and cloud treatments is given by *Binkowski and Roselle* [2003] and *Byun and Schere* [2006].

2.2. Observational Data Sets and Model Evaluation Methodology

[6] The U.S. EPA recommended the evaluation of surface O_3 and related gaseous pollutants such as CO, SO_2, NO_2, hydrogen peroxide (H_2O_2), nitric acid (HNO_3), peroxyacetyl nitrate (PAN), VOCs, NH_3, NO_y (the sum of NO_x and other reactive nitrogen compounds), $PM_{2.5}$ and its components such as sulfate (SO_4^{2-}), nitrate (NO_3^-), ammonium (NH_4^+), black carbon (BC), organic carbon (OC), and other unspeciated inorganic PM (e.g., crustal, heavy metals), and wet deposition fluxes of species [*U.S. EPA*, 2007]. The U.S. EPA also recommended that an evaluation include indicator species such as NO_y and ratios of H_2O_2/HNO_3, O_3/NO_x, O_3/NO_y, and O_3/NO_z (where $NO_z = NO_y - NO_x$) [*Sillman*, 1995, *Sillman et al.*, 1997; *Sillman and He*, 2002; *Lu and Chang*, 1998; *Zhang et al.*, 2002, 2005], as these indicators can indicate whether local O_3 formation is VOC- or NO_x-

Table 1a. Parameters and Associated Observational Databases Included in the Model Evaluation: Surface Databases

Database[a]	Parameter	Data Frequency	Number of Sites	Data Source
CASTNET	O_3, SO_4^{2-}, NO_3^-, NH_4^+	Hourly O_3, Weekly average SO_4^{2-}, NO_3^-, NH_4^+	83 sites, mostly in remote/rural areas in the U.S.	http://www.epa.gov/castnet/
AIRS-AQS	O_3	Hourly	1161 sites, primarily in cities and towns in the U.S.	http://www.epa.gov/air/data/index.html
IMPROVE	$PM_{2.5}$, SO_4^{2-}, NO_3^-, NH_4^+, EC, OC	24-h average, 1 in every 3 days	134 sites, mostly remote locations in the western U.S.	http://vista.cira.colostate.edu/improve/
STN	$PM_{2.5}$, SO_4^{2-}, NO_3^-, NH_4^+	24-h average, 1 in every 3 days	139 sites, all in urban areas	http://www.epa.gov/air/data/index.html
SEARCH	O_3, NO_x, NO_y, $PM_{2.5}$, SO_4^{2-}, NO_3^-, EC, OC	Hourly for gases and daily average for PM and its components	8 sites, located in the urban/suburban areas in the southeastern U.S.	http://www.atmospheric-research.com/studies/SEARCH/index.html
NADP	Wet deposition of SO_4^{2-}, NO_3^-, NH_4^+	Weekly total	Over 250 sites nationwide	http://nadp.sws.uiuc.edu

[a]CASTNET, Clean Air Status and Trends Network; AIRS-AQS, Aerometric Information Retrieval System-Air Quality Subsystem; IMPROVE, Interagency Monitoring of Protected Visual Environments; NADP, National Acid Deposition Program; SEARCH, Southeastern Aerosol Research and Characterization; and STN, Speciated Trends Network.

limited at any point in space and time and whether the model is correctly predicting the responses of O_3 to VOC and/or NO_x emission controls. They may also reveal whether the model can correctly reproduce the sensitivity of secondary $PM_{2.5}$ to changes in the emissions of gaseous precursors such as SO_2, NH_3, VOCs, and/or NO_x [*Ansari and Pandis*, 1998; *Blanchard*, 2000; *Pun and Seigneur*, 2001]. If a model accurately predicts observed indicators, the predicted change in future year O_3 or PM is more likely to be accurate.

[7] In this work, the above individual species and some indicator species such as NO_y, O_3/NO_x, and O_3/NO_y are included in the model evaluation where observational data are available (see Table 2 of *Zhang et al.* [2009] for the definitions of the indicator species). Tables 1a and 1b summarizes measured parameters and relevant data sets used for model evaluation. The maximum 1 h and 8 h average O_3 mixing ratios, the 24 h average mass concentrations of $PM_{2.5}$ and its composition, their gaseous precursors, and photochemical indicator species are evaluated using the surface observations from one special field study (i.e., the Southeastern Aerosol Research and Characterization study (SEARCH)) and four nationwide routine monitoring networks (i.e., the Clean Air Status and Trends Network (CASTNET), the Aerometric Information Retrieval System–Air Quality System (AIRS-AQS), the Interagency Monitoring of Protected Visual Environments (IMPROVE), and the Speciation Trends Network (STN)). The wet deposition fluxes of SO_4^{2-}, NO_3^-, and NH_4^+ are evaluated with observations from the National Atmospheric Deposition Program (NADP).

[8] Satellite observations have been increasingly used to support air quality modeling and forecasting to improve the accuracy of model inputs (e.g., emissions, ICONs, and BCONs), evaluate model predictions of column quantities and vertical profiles (e.g., CO, NO_2, O_3, and aerosol optical depths (AODs)), and correct model simulation/forecasting errors by assimilating satellite-derived chemical fields (e.g., CO, NO_2, O_3, and AODs) [e.g., *Roy et al.*, 2007a, 2007b; *Pouliot et al.*, 2008; *Mathur*, 2008; *Vijayaraghavan et al.*, 2008]. Here, the predicted total tropospheric column abundances of CO, NO_2, and O_3 from CMAQ and AODs derived from CMAQ $PM_{2.5}$ predictions are compared with derived quantities based on satellite products from the Measurements Of Pollution In The Troposphere (MOPITT), the Global Ozone Monitoring Experiment (GOME), the Total Ozone Mapping Spectrometer (TOMS), the Solar Backscatter Ultraviolet (SBUV) instrument, and the Moderate Resolution Imaging Spectroradiometer (MODIS). Such an evaluation will assess the model's capability in capturing total column mass, which may influence their abundance and distribution in the lower atmosphere.

[9] Model evaluation is performed using an evaluation protocol that follows the *U.S. EPA* [2001, 2007] and *Zhang et al.* [2006a]. The evaluation protocol includes spatial distribution, temporal variation, vertical profiles, column abundances, and overall statistical trends. The performance statistics are calculated separately for different networks, given differences in sampling methods, protocols, periods, and frequencies and resultant differences in measured species in terms of accuracy, biases, and precision across different networks. Such a separate evaluation will also enable an assessment of systematic biases in measurements from each network and in the model predictions. While it is not feasible to perform detailed analyses at more than one thousand individual sites in the networks shown in Table 1a for the 1 year simulation, such analyses at sites that represent air masses of different origins (e.g., urban vs. rural, inland vs. coast, NO_x-limited vs. VOC-limited regions) would be valuable and provide further insights into the model's capability in capturing trends at smaller time/spatial scales to guide air quality attainment that is targeted for both short- (i.e., 8 h or 24 h average) and long- (i.e., 1 year) terms. In this study, a total of 22 representative locations are selected for detailed temporal variation analyses and process analyses, as shown in Figure 1. These include 8 sites from SEARCH (Jefferson Street (JST), Atlanta, GA, Yorkville (YRK), GA, North Birmingham (BHM), AL, Centreville (CTR), AL, Gulfport (GFP), MS, Oak Grove (OAK), MS, Pensacola (PNS), FL, and Outlying Landing (OLF), FL), 13 sites from AIRS-AQS (i.e., Big Bend NP (BBE), TX, Great Smoky National Park (GRS), TN, Olympic National Park (OLY), WA, Yellowstone National Park (YEL), WY, YRK, Chicago (CHI), IL, Fresno (FRE), CA, Huston (HOU), TX, Los Angeles (LAX), CA, New York city (NYC), NY, Pittsburgh (PIT), PA, Riverside (RIV), CA, Tampa (TAM), FL), 5 sites from CASTNET (i.e., BBE, GRS, OLY, YEL, and Penn State (PSU), PA), and 4 sites from IMPROVE (i.e., BBE, GRS, YEL, and Grand Canyon National Park (GRC), AZ) (note

Table 1b. Parameters and Associated Observational Databases Included in the Model Evaluation: Satellite Databases

Sensor/Satellite[a]	Parameter	Spatial, Vertical, and Temporal Resolutions for Raw Measurements	Level-3 Data Spatial Grid (Latitude × Longitude)[b]	Time Resolution for Evaluation	Note and Reference
MOPITT/Terra	CO	Horizontal: 22 × 22 km, Vertical: 5 km, Swath width: ~640 km, Temporal: every 0.4 seconds, compiled as daily	1 × 1	Monthly mean	missing June and July because of sensor issues, http://terra.nasa.gov/About/MOPITT/ [*Edwards et al.*, 2004]
GOME/ERS-2	NO_2, Tropospheric Column Ozone (TCO)	Horizontal:40 × 40 km to 40 × 320 km, Vertical: 5 km, Swath width:120–960 km, Temporal: 10:30 am local time	0.5 × 0.5	Monthly mean	Raw NO_2 data were provided by Dr. Andreas Richter, the Institute of Environmental Physics, University of Bremen, Germany, GOME NO_2 data were obtained from Dr. Kelly Chance, the Harvard-Smithsonian Center for Astrophysics, Cambridge, MA http://gcmd.gsfc.nasa.gov/records/GCMD_gov.noaa.class.GOME.html, [*Heland et al.*, 2002; *Richter and Burrows*, 2002; *Liu et al.*, 2005, 2006a]
TOMS/SBUV Earth Probe	Tropospheric ozone residual (TOR)	Horizontal: 125 × 100 km Vertical: tropospheric column Temporal: 3 or 4 times per day	1.25 × 1	Daily and monthly mean	TOMS/SBUV data were provided by Jack Fishman and John K. Creilson, NASA Langley Research Center, U.S., http://asd-www.larc.nasa.gov/TOR/TOR_Data_and_Images.html [*Fishman et al.*, 2003]
MODIS/Terra	Aerosol Optical Depth (AOD)	Horizontal: 1,000-m, 500-m, and 250-m resolution spectral bands, Vertical: tropospheric column Swath width: 2,330-km Temporal: 10:30 am local time	1 × 1	Daily and monthly mean	NASA Goddard Distributed Active Archive Center (DAAC), http://ladsweb.nascom.nasa.gov/data/search.html

[a]MOPITT, Measurements Of Pollution In The Troposphere; GOME, Global Ozone Monitoring Experiment; TOMS/SBUV, Total Ozone Mapping Spectrometer/the Solar Backscatter Ultraviolet (SBUV); and MODIS, Moderate Resolution Imaging Spectroradiometer.
[b]Spatial grid refers to the level-3 data, not to the measurement spatial resolution.

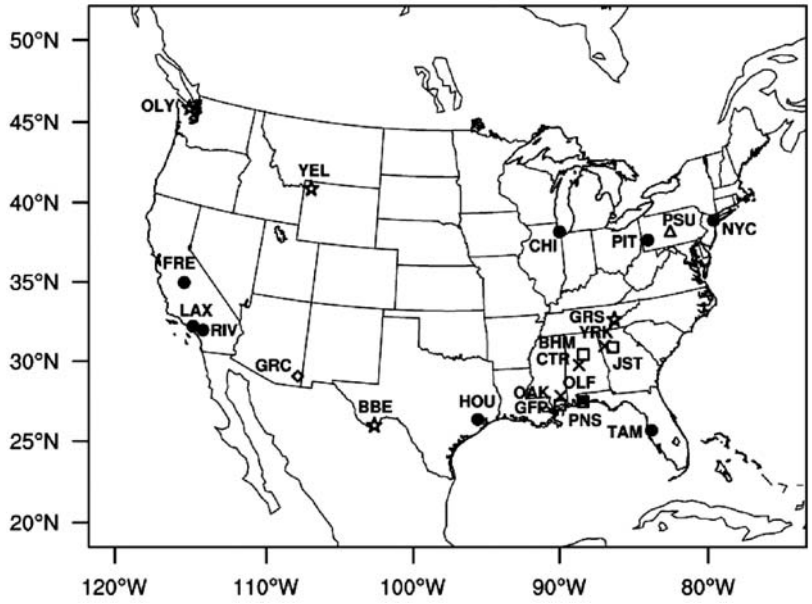

Figure 1. Map of the 22 sites locations. Dots denote 8 sites from AIRS-AQS: Chicago (CHI), IL; Fresno (FRE), CA; Houston (HOU), TX; Los Angeles (LAX), CA; New York City (NYC), NY; Pittsburgh (PIT), PA; Riverside (RIV), CA; and Tampa (TAM), FL. The squares (for urban sites) and crosses (for rural sites) denote 8 sites from SEARCH: Jefferson Street (JST), GA; Yorkville (YRK), GA; North Birmingham (BHM), AL; Centreville (CTR), AL; Gulfport (GFP), MS; Oak Grove (OAK), MS; Pensacola (PNS), FL; and Outlying Landing (OLF), FL. The triangle denotes 1 site from CASTNET: Penn State (PSU), PA. The diamond denotes 1 site from IMPROVE: Grand Canyon National Park (GRC), AZ. The star denotes 4 sites that belong to multiple networks: Big Bend NP (BBE), TX from AIRS-AQS, CASTNET, and IMPROVE; Great Smoky National Park (GRS), TN from AIRS-AQS, CASTNET, and IMPROVE; Olympic National Park (OLY), WA from AIRS-AQS and CASTNET; and Yellowstone National Park (YEL), WY from AIRS-AQS, CASTNET, and IMPROVE.

that 7 sites belong to multiple networks). The 8 SEARCH sites represent three urban-rural pairs in GA, AL, and MS, and one urban-suburban pair in FL. 13 AIRS-AQS sites represent major cities (e.g., CHI, HOU, LAX, NYC, PIT, and TAM), suburban/rural sites (e.g., FRE, RIV, and YRK), and national parks (e.g., BBE, GRS, OLY, and YEL) with different emissions/meteorological characteristics in different regions of CONUS. PSU and GRC are two additional rural and national park sites in the eastern and western U.S, respectively, from CASTNET and IMPROVE.

[10] The statistical measures calculated include correlation coefficient (R), mean bias (MB), the mean normalized bias (MNB), and the mean normalized gross error (MNE), the normalized mean bias (NMB), and the normalized mean gross error (NME), the mean fractional bias (MFB), and the mean fractional error (MFE) (see their definitions given by *Zhang et al.* [2006a]). While the values of MNB/MNE and NMB/NME are similar in many cases, they may give different results in some cases. For example, when the observed values at a specific time or location are extremely low, extremely high values may occur for MNB/MNE but not NMB/NME. In such cases, the use of NMB/NME will provide a more reasonable evaluation. The ranges of performance statistics over different networks indicate the deviation of model performance, i.e., whether the model is robust or subject to systematic biases over a group of sites having similar characteristics under each network considered in the model evaluation.

3. Satellite Data Processing and Column Mass Calculation From CMAQ Predictions

[11] The MOPITT CO data are available for all months except June and July. The GOME NO_2 and TCO, TOMS/SBUV TORs, and MODIS AODs are available for all months. Monthly products are computed by averaging daily products over each month. In addition, vertical CO profiles are derived from the MOPITT CO retrievals to evaluate simulated profiles. For profile measurements, the finite vertical spatial resolution of the sensor must be treated explicitly when comparing the measurements with predictions. The appropriate approach to treat these differences in vertical resolution is to use the averaging kernel of the sensor, as described by M. N. Deeter (Calculation and application of MOPITT averaging kernels, 2002, available from http://www.acd.ucar.edu/mopitt/data/avg_krnls_app.pdf, accessed May 2009). The averaging kernel represents the way in which the vertical structure of the atmospheric profile is mapped into the radiances measured by the sensor. It is expressed mathematically as a matrix where each row defines the averaging kernel for a particular retrieval level within the measured profile, and each element in this row represents the relative weighting of the 'true' mixing ratio value at each

level to the retrieved mixing ratio. The value of the averaging kernel is a function of sensor parameters (e.g., the field-of-view) and those parameters input to the forward radiative transfer model (e.g., the temperature profile and the species of interest). Simulated CO mixing ratios are extracted in each layer from the surface to the model top at the tropopause. The averaging kernel matrix values from MOPITT are then applied to these model output in order to compare with the finite resolution MOPITT profiles.

[12] CMAQ simulates 3-D gridded hourly average mixing ratios of CO, NO_2, O_3, and concentrations of $PM_{2.5}$ species from surface to 15.7 km. These values and vertically resolved temperature and pressure estimated from MM5 are used to calculate the total column mass and AODs. The simulated total column mass of species i, $TCOLMASS_i$, is calculated as:

$$TCOLMASS_i = \sum_{l=1}^{M} \frac{p_l \times \Delta z_l \times A \times [C_i]_{ppm,l}}{R \times T_l \times CONST_i} \quad (1)$$

where $[C_i]_{ppm,\ l}$ is the mixing ratio of species i (i.e., CO, NO_2, O_3) in ppm in layer l, Δz_l is the height of layer l, M is the total number of model layers (M = 14 in this application), p_i and T_i are the pressure in Pa and temperature in K, respectively, in layer l, R is the gas constant, = 8.34 J mole^{-1} K^{-1}, A is the Avogadro's number, = 6.02213×10^{23} molecules mole^{-1}, $CONST_i$ is the constant for unit conversion, = 10^{-18} for CO column mass, $TCOLMASS_{CO}$($\times 10^{18}$ molecules cm^{-2}), 10^{-14} for NO_2 column mass, $TCOLMASS_{NO2}$($\times 10^{14}$ molecules cm^{-2}), and 2.687×10^{28} for O_3 column mass, $TCOLMASS_{O3}$ (Dobson Unit (DU), one DU refers to the number of molecules of O_3 that would be required to create a layer of O_3 that would be 10 μm thick under standard temperature and pressure). Monthly, seasonal, and annual averages are computed for each grid cell. $TCOLMASS_{CO}$ values are compared against the tropospheric CO columns derived from the MOPITT instrument on EOS Terra [*Edwards et al.*, 2004]. The $TCOLMASS_{NO2}$ values are compared against the tropospheric NO_2 columns derived from GOME instrument on the European Remote Sensing (ERS-2) satellite [*Richter and Burrows*, 2002]. $TCOLMASS_{O3}$ values are compared against the Tropospheric O_3 Residuals (TORs) derived from TOMS and SBUV by *Fishman et al.* [2003, 2005] using the NCEP/NCAR reanalysis tropopause heights and total tropospheric column O_3 (TCO) retrieved from GOME by *Liu et al.* [2005, 2006a, 2006b]. The TOMS/SBUV TORs are based on the TOR method and obtained from the difference between the TOMS TCO and SBUV Stratospheric Column Ozone (SCO). This method makes assumptions about the distribution and variability of SCO. The GOME TCO retrieval of *Liu et al.* [2005, 2006a, 2006b] is the first directly retrieved global distribution of TCO from GOME. The advantage of direct retrievals over the residual-based approaches is that daily global distributions of tropospheric O_3 can be derived without other collocated satellite measurements of SCO, or the need to make assumptions about the spatiotemporal distribution of SCO. More detailed information is given by *Liu et al.* [2006a].

[13] The spectral AOD (i.e., τ_λ) is a strong function of aerosol mass concentration, size distribution, chemical composition, and wavelength. It is given by [*Jacobson*, 2005]:

$$\tau_\lambda = \int_0^\infty \sigma_\lambda dz \quad (2)$$

where σ_λ is the total aerosol spectral extinction coefficient, it is the sum of the aerosol absorption coefficient, $\sigma_{ap,\lambda}$, and scattering coefficient, $\sigma_{sp,\lambda}$, [*Jacobson*, 2005]:

$$\sigma_\lambda = \sigma_{ap,\lambda} + \sigma_{sp,\lambda} \quad (3)$$

The aerosol scattering and absorption coefficients can be calculated as a function of PM number concentrations, single-particle radius, and single-particle absorption and scattering efficiencies. AODs can be calculated using physically based [e.g., *Jacobson*, 2005], parameterized [e.g., *Ghan et al.*, 2001], and empirical approaches [e.g., *Chameides et al.*, 2002; *Roy et al.*, 2007a]. Since the physically based approach is computationally expensive and CMAQ does not explicitly simulate AODs, the most appropriate approach is the empirical approach that is based on CMAQ $PM_{2.5}$ predictions. In this study, an empirical method of *Chameides et al.* [2002] is used to calculate AODs at a nominal wavelength of 550 nm based on CMAQ $PM_{2.5}$ predictions. In this approach, the scattering coefficient is estimated as:

$$\begin{aligned}\sigma_{sp} &= \sigma_{sp}^{SO_4^{2-}} + \sigma_{sp}^{NO_3^-} + \sigma_{sp}^{OC} + \sigma_{sp}^{BC} \\ &= \left\{ \left[SO_4^{2-}\right] \times \alpha_{sp}^{SO_4} + \left[NO_3^-\right] \times \alpha_{sp}^{NO_3} + [OC] \right. \\ &\quad \left. \times \alpha_{sp}^{OC} + [BC] \times \alpha_{sp}^{BC} \right\} \times f(RH) / \left(1.0 \times 10^6\right)\end{aligned} \quad (4)$$

where $[SO_4^{2-}]$, $[NO_3^-]$, [OC], and [BC] are the mass concentrations of SO_4^{2-}, NO_3^-, OC, and BC in μg m^{-3} in the Aitken and accumulation modes. $\alpha_{sp}^{SO_4}$, $\alpha_{sp}^{NO_3}$, α_{sp}^{OC}, and α_{sp}^{BC} are the specific scattering coefficients in m^2 g^{-1}; their values are as follows [*Chameides et al.*, 2002]:

$$\alpha_{sp}^{SO_4} = 5.0\ m^2 g^{-1}, \alpha_{sp}^{NO_3} = 5.0\ m^2 g^{-1},$$
$$\alpha_{sp}^{OC} = 5.0\ m^2 g^{-1},\ and\ \alpha_{sp}^{BC} = 3.0\ m^2 g^{-1}$$

The function, f(RH) ($\geq$1), is the species-dependent hygroscopic growth factor for total particle scattering, it accounts for the effect of relative humidity (RH) on scattering efficiency. Some measurements exist for f(RH) for some PM species. For example, *Kotchenruther et al.* [1999] measured an overall f(RH) of 1.9–2.6 at an RH of 80% for PM consisting of primarily ammonium bisulfate and carbonaceous PM. *Kotchenruther and Hobbs* [1998] measured an overall f(RH) of 1.1–1.3 at an RH of 80% for PM emitted from biomass burning in Brazil. A bulk hygroscopic growth factor of 2.3 ± 0.4 at an RH of 80% and a wavelength of 550 nm were recommended [*Penner et al.*, 2001]. We therefore assume a constant RH of 80% and use f(80%) of 2.3. Note that NH_4^+ and other organic PM are not explicitly included in equation (4), since they have been implicitly included in the specific scattering coefficients chosen for $[SO_4^{2-}]$, $[NO_3^-]$,

and [OC]. The PM absorption coefficient is estimated as follows:

$$\sigma_{ap} = \{[BC]_{hydrophobic} \times \alpha_{ap}^{hydrophobic} + [BC]_{hydrophilic} \times \alpha_{ap}^{hydrophilic}\} / (1.0 \times 10^6) \quad (5)$$

where $\alpha_{ap}^{hydrophobic}$ and $\alpha_{ap}^{hydrophilic}$ are the specific absorption coefficients:

$$\alpha_{ap}^{hydrophobic} = 10\, m^2 g^{-1},\ \alpha_{ap}^{hydrophilic} = 20\, m^2 g^{-1}$$

$[BC]_{hydrophobic}$ and $[BC]_{hydrophilic}$ are the hydrophobic and hydrophilic fractions of BC. Externally mixed BC is generally considered to be hydrophobic, whereas well internally mixed BC is considered to be hydrophilic. Uncertainties in AODs and aerosol radiative properties and forcing exist with different assumptions of the mixing state of BC, namely, externally mixed, well internally mixed, or core treatment in which BC particles become coated with other types of aerosols such as sulfate and nitrate [*Jacobson*, 2001]. While the three PM modes simulated in CMAQ are treated as externally mixed, the PM components in each mode are assumed to be well internally mixed. The simulated BC concentration is not specified as hydrophobic or hydrophilic, and an assumption must be made about the relative fractions of each. Most BC emitted can be coated with other aerosols, and internally mixed BC may comprise a significant fraction on a global scale [e.g., *Pósfai et al.*, 1999]. An internally mixed assumption is well supported by field measurements thus seems to be more realistic than the externally mixed assumption [*Jacobson*, 2001]. In this study, 90% of BC is assumed to be internally mixed (i.e., hydrophilic), and the remaining 10% is externally mixed (i.e., hydrophobic). The total AODs predicted by CMAQ can thus be estimated as:

$$\tau_a = \sum_{l=1}^{M} (\sigma_{sp,l} + \sigma_{ap,l}) \times \Delta z_l \quad (6)$$

where τ_a is the total integrated AOD (dimensionless) from CMAQ. $\sigma_{sp,l}$ is the aerosol scattering coefficient in m^{-1} and $\sigma_{ap,l}$ is the aerosol absorption coefficient in m^{-1} in layer l.

[14] The model-estimated AODs for $PM_{2.5}$ are compared with the observed daily and monthly mean total MODIS AODs. They are from the Level-3 land-corrected AOD products at 550 nm with a grid resolution of $1^\circ \times 1^\circ$ retrieved from MOD04 level 2 aerosol products that are originally collected at a resolution of 10 km × 10 km [*Remer et al.*, 2005]. MODIS contains fine-mode AODs that are derived for fine-mode aerosols from urban/industrial and biomass burning [*Chu et al.*, 2003; *Remer et al.*, 2005], which are generally associated with aerosols with geometric mean diameters of 0.01–0.22 μm over land and 0.12–0.2 μm over ocean [*Remer et al.*, 2005; *Kaufman et al.*, 2005]. Since the MODIS fine-mode AOD can only represent a portion of $PM_{2.5}$ AODs, it is appropriate to compare the MODIS total AODs with calculated CMAQ $PM_{2.5}$ AODs which are nearly the same as the simulated total AODs because this version of CMAQ does not simulate coarse PM such as sea salt and dust. The stratospheric component of the MODIS AODs is assumed to be negligible because the only reported volcanic eruption was the eruption of Mt. Cleveland, Alaska between February 2 and April 15, 2001 (http://www.avo.alaska.edu/ and *Dean et al.* [2004]), and no other major volcanic eruptions were reported over CONUS in 2001. PM sources other than volcanoes in the stratosphere are considered to be fairly small, as compared to tropospheric sources. The $1^\circ \times 1^\circ$ AODs in latitude/longitude coordinates are mapped to the Lambert conformal projection used in CMAQ using bilinear interpolation in the NCAR Command Language (NCL) in order to compare them with CMAQ predictions. Terra orbits cross the equator at 10:30 local time. The monthly mean AODs from CMAQ are therefore calculated as an average of values during 1500–2000 UTC when the Terra satellite passes CONUS, following *Roy et al.* [2007a].

4. Evaluation With in Situ Surface Measurements

4.1. Spatial Distributions and Performance Statistics

[15] Figure 2 shows overlay plots of the simulated seasonal mean daily maximum 8 h average surface O_3 mixing ratios (ppb) and 24 h average $PM_{2.5}$ concentrations ($\mu g\ m^{-3}$) against all available observations. Table 2 summarizes the overall seasonal and annual statistical performance of CMAQ for the maximum 1 h and 8 h average surface O_3 mixing ratios and 24 h average $PM_{2.5}$ concentrations. In winter (December, January, and February (DJF)), CMAQ reproduces well peak 8 h O_3 at most sites. The ranges of NMBs and NMEs are −10.6% to 6.4% and 21.7–25.4%, respectively. In spring (March, April, and May (MAM)), underpredictions in maximum 8 h O_3 occur at a number of sites in the states of NC, SC, VA, KY, and GA in the eastern U.S. and several CASTNET sites in CA, AZ, UT, and WY in the western U.S. The ranges of NMBs and NMEs are −2.0% to 3.0%, and 13.3–16.3%, respectively. In summer (June, July, and August (JJA)), CMAQ reproduces well maximum 8 h O_3 except underpredictions at a few sites in CA, WY, UT, AZ, KS, OK, NC, VA, and PA and overpredictions at some SEARCH sites in the southeastern U.S. The ranges of NMBs and NMEs are −2.3% to 15.9%, and 17.3–23.3%, respectively. In fall (September, October, and November (SON)), CMAQ reproduces well maximum 8 h O_3 at most sites. Overpredictions occur for maximum 8 h O_3 at a number of CASTNET sites in the northeastern and western U.S. and at most AIRS-AQS sites throughout the domain. The ranges of NMBs and NMEs are −6.1% to 1.9%, and 18.9–19.6%, respectively. The spatial distributions of the seasonal mean maximum 1 h O_3 follow similar trends but with larger biases (except winter, summer, and fall at the SEARCH sites), as compared with those of maximum 8 h O_3. The NMBs, NMEs, MFBs, MFEs, and R at various networks are −11. 6% to 0.1%, 19.8–20.2%, −12.7% to 2.5%, 21.7–22.9%, and 0.7, respectively for annual mean maximum 1 h O_3 and NMBs, NMEs, MFBs, MFEs, and Rs are −4.6% to 3.0%, 17.7–22.1%, −5.0% to 5.8%, 20.3–24.2%, and 0.7–0.8, respectively, for annual mean maximum 8 h O_3. These results are consistent with those reported by other studies [e.g., *Eder and Yu*, 2006; *Zhang et al.*, 2006b; *S.-Y. Wu et al.*, 2008].

[16] The latest U.S. EPA's guidance on the attainment demonstrations for the maximum 8 h O_3 NAAQS states that it is not appropriate to assign 'bright line' (i.e., a clear cutoff value based on one statistical measure) criteria that distin-

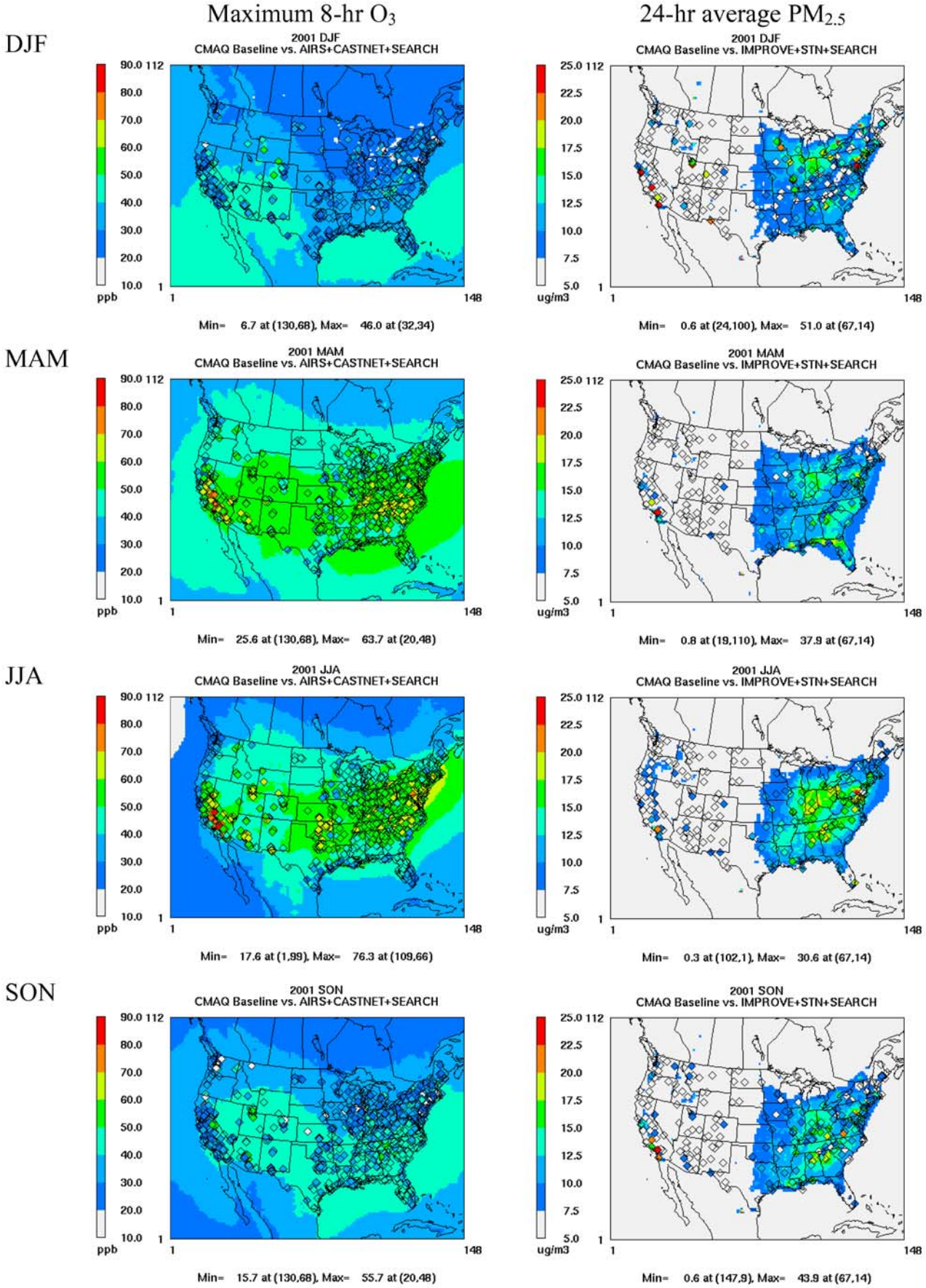
Maximum 8-hr O3
24-hr average PM2.5
DJF
MAM
JJA
SON
2001 DJF
CMAQ Baseline vs. AIRS+CASTNET+SEARCH
Min= 6.7 at (130,68), Max= 46.0 at (32,34)
2001 DJF
CMAQ Baseline vs. IMPROVE+STN+SEARCH
Min= 0.6 at (24,100), Max= 51.0 at (67,14)
2001 MAM
CMAQ Baseline vs. AIRS+CASTNET+SEARCH
Min= 25.6 at (130,68), Max= 63.7 at (20,48)
2001 MAM
CMAQ Baseline vs. IMPROVE+STN+SEARCH
Min= 0.8 at (19,110), Max= 37.9 at (67,14)
2001 JJA
CMAQ Baseline vs. AIRS+CASTNET+SEARCH
Min= 17.6 at (1,99), Max= 76.3 at (109,66)
2001 JJA
CMAQ Baseline vs. IMPROVE+STN+SEARCH
Min= 0.3 at (102,1), Max= 30.6 at (67,14)
2001 SON
CMAQ Baseline vs. AIRS+CASTNET+SEARCH
Min= 15.7 at (130,68), Max= 55.7 at (20,48)
2001 SON
CMAQ Baseline vs. IMPROVE+STN+SEARCH
Min= 0.6 at (147,9), Max= 43.9 at (67,14)
ppb
ug/m3

Figure 2

Table 2. Seasonal and Annual Normalized Mean Bias and Error of CMAQ Predictions in 2001[a]

		Winter		Spring		Summer		Fall		Annual	
Variable[b]	Network	NMB	NME	NMB	NME	NMB	NME	NMB	NME	NMB	NME
Max 1-h O_3	CASTNET	−17.5	24.1	−7.9	16.3	−10.5	19.7	−12.7	20.8	−11.6	19.8
	AIRS-AQS	−8.0	24.1	−5.0	17.0	−8.0	20.2	−8.9	20.9	−7.4	19.8
	SEARCH	3.3	22.9	−3.1	13.1	10.2	20.2	0.5	18.3	0.1	20.5
Max 8-h O_3	CASTNET	−10.6	21.7	−2.0	14.6	−2.3	17.3	−6.1	18.9	−4.6	17.7
	AIRS-AQS	1.0	25.4	3.0	16.3	2.1	18.3	−1.5	19.6	1.4	18.7
	SEARCH	6.4	24.9	−1.1	13.3	15.9	23.3	1.9	19.3	3.0	21.7
24-h avg. $PM_{2.5}$	IMPROVE	29.7	59.6	−0.7	45.2	−9.1	39.8	18.3	47.6	5.8	46.0
	STN	1.3	55.6	10.0	45.7	−1.4	37.7	8.5	47.0	4.2	45.5
	SEARCH	24.9	48.9	32.4	53.7	41.0	55.7	42.8	54.8	35.5	53.4
24-h avg. SO_4^{2-}	CASTNET	−22.8	30.2	−15.7	24.4	−0.5	21.3	22.4	35.6	−2.9	26.5
	IMPROVE	−5.0	44.6	−8.4	36.5	−7.6	38.5	15.5	47.1	−1.7	41.0
	STN	−31.2	55.6	−28.1	47.7	−7.7	45.6	0.7	55.0	−13.0	49.8
	SEARCH	28.8	58.1	−8.0	34.1	17.5	40.5	59.4	64.6	43.5	72.6
24-h avg. NO_3^-	CASTNET	3.8	51.5	13.3	56.2	−26.3	63.7	1.8	60.6	4.6	55.2
	IMPROVE	−29.2	72.5	52.9	105.1	−10.4	102.6	24.2	107.1	7.6	93.3
	STN	−33.0	60.8	−20.6	83.2	−47.8	100.2	−33.7	86.2	−33.9	82.0
	SEARCH	−37.8	85.7	130.4	143.6	94.4	126.9	143.0	173.7	46.0	120.3
24-h avg. NH_4^+	CASTNET	−3.6	28.4	11.4	29.3	9.2	24.8	24.0	40.9	9.7	30.0
	IMPROVE	17.7	50.0	27.6	46.5	25.2	47.8	29.4	44.8	26.0	47.0
	STN	−40.3	69.5	−25.5	71.7	−19.6	69.7	−9.6	88.0	−22.3	74.6
	SEARCH	−10.4	77.7	67.1	82.2	14.5	80.5	58.6	68.8	−2.9	76.7
24-h avg. BC	IMPROVE	−4.7	63.8	−8.6	54.0	8.0	70.1	−15.4	51.1	−5.4	59.5
	SEARCH	−49.6	67.8	−43.8	55.4	−54.1	59.4	−59.8	70.2	−52.6	64.0
24-h avg. OC	IMPROVE	32.7	75.1	15.5	59.0	25.9	77.8	22.7	71.8	24.8	73.1
	SEARCH	−39.2	52.7	−21.3	47.1	−44.3	49.3	−43.3	48.9	−37.6	49.5
Weekly total $SO_{4\ wet}^{2-}$	NADP	35.0	71.0	11.0	64.7	−14.4	57.6	63.3	99.0	19.6	71.3
Weekly total $NO_{3\ wet}^-$	NADP	26.8	83.9	−19.1	60.5	−58.9	69.0	11.9	73.9	−13.0	71.6
Weekly total NH_{4wet}^+	NADP	85.7	136.0	22.4	84.1	1.1	70.0	40.0	99.0	31.6	92.8
Tropospheric CO column[c]	MOPITT	−0.6	10.3	−6.7	14.8	26.0	29.1	1.9	7.9	−0.6	10.3
Tropospheric NO_2 column	GOME	18.4	53.6	5.5	39.8	−7.5	42.6	5.7	43.6	6.9	45.6
Tropospheric O_3 column	GOME	25.0	29.0	20.0	27.0	−7.0	24.0	11.0	20.0	7.0	25.0
	TOMS/SBUV	34.9	35.7	28.7	31.4	−13.0	18.1	6.3	13.8	12.0	24.1
Aerosol optical depth[d]	MODIS	−30.0	88.2	−37.1	63.6	−36.1	48.1	−17.6	62.3	−31.9	54.8

[a]Values are given as percent.

[b]Normalized mean bias (NMB) and error (NME) defined as $NMB = [\sum_{i=1}^{N}(M_i - O_i)]/\sum_{i=1}^{N}O_i = (\frac{\overline{M}}{\overline{O}} - 1)$; $NME = [\sum_{i=1}^{N}|M_i - O_i|]/\sum_{i=1}^{N}O_i = MAGE/\overline{O}$ where $\overline{M} = (1/N)\sum_{i=1}^{N}M_i$, $\overline{O} = (1/N)\sum_{i=1}^{N}O_i$, $MNGE = \frac{1}{N}\sum_{i=1}^{N}[(|M_i - O_i|)/O_i]$, M_i and O_i are values of model prediction and observation at time and location i, respectively. N is the number of samples (by time and/or location). The ranges of values are −1 to +∞ for NMB and 0 to +∞ for NME.

[c]The statistics of column CO for summer is calculated based on the data in August. No data are available for June and July 2001.

[d]The observational data are obtained for CO from MOPITT, NO_2 from GOME, TOR from TOMS/SBUV and GOME, and AOD from MODIS.

guish between adequate and inadequate model performance [*U.S. EPA*, 2007]. Various statistical measures have been used to determine O_3 performance. *Russell and Dennis* [2000] suggested using an MNB < ±15% and an MNGE < 35% for an acceptable performance. In the U.S. EPA's reports published prior to 2005, the values of MNB and MNGE recommended by EPA for a good performance of O_3 are ≤±15% and ≤30%, respectively, and O_3 peak accuracy <20%. However, MNB/MNGE are not as robust as NMB/NME for cases with extremely low observed values, NMBs ≤±15% and NMEs ≤ 30% are therefore recommended to indicate good performance [*Zhang et al.*, 2006a]. *Morris et al.* [2004] suggested using MFBs ≤±15% and MFEs ≤ 35% to determine a satisfactory O_3 performance. Compared with the criteria used by other studies, the ranges of NMBs/NMEs and MFBs/MFEs for seasonal and annual statistics obtained in this study indicate that the model performance for maximum 1 h and 8 h average O_3 is overall satisfactory.

[17] The observed $PM_{2.5}$ mass concentrations are the largest at the IMPROVE, STN, and SEARCH sites in summer and the smallest at the IMPROVE and SEARCH sites in winter and at the STN sites in spring. This seasonality is generally reproduced by CMAQ. At the STN sites, the overall agreement is good with NMBs and NMEs for seasonal mean values ranging from −1.4% to 10.0% and from 37.7% to 55.6%, respectively, except for CA where large underpredictions occur and in the eastern U.S. where underpredictions occur at a few cities in all seasons. Small overpredictions occur in winter, spring, and fall (NMBs of 1.3–10.0%) and a small underprediction occurs in JJA (−1.4%), resulting in a net NMB of 4.2% annual average overprediction. Compared with the STN sites, NMEs are similar (39.8–59.6%) but NMBs at the IMPROVE sites are larger, ranging from −9.1% to 29.7%, with moderate underpredictions occurring in the eastern U.S. and small overpredictions in the western U.S. in summer, and moder-

Figure 2. The observed and simulated seasonal mean surface maximum 8 h average O_3 mixing ratios (ppb) and 24 h average $PM_{2.5}$ concentrations ($\mu g\ m^{-3}$) in 2001. The observations are indicated by colored diamonds; they were obtained from CASTNET, AIRS-AQS, and SEARCH for O_3 and from IMPROVE, STN, and SEARCH for $PM_{2.5}$; DJF, December-January-February; MAM, March-April-May; JJA, June-July-August; SON, September-October-November.

ate overpredictions in the eastern U.S. in winter and fall. The model reproduces the observed $PM_{2.5}$ well in spring with the smallest NMB of −0.7% but the NME is 45.2%, indicating error cancelation in the NMB. At the IMPROVE sites, the overpredictions in winter and fall are compensated by underpredictions in summer and spring, resulting in only 5.8% annual average overprediction. Moderate overpredictions (with NMBs of 24.9–42.8%) occur for all seasons at the SEARCH sites, leading to a moderate overprediction of 35.5% in the annual mean value. NMEs for seasonal mean values are in the range of 35.5–55.7%. For annual mean 24 h average $PM_{2.5}$, NMBs, NMEs, MFBs, MFEs, and Rs at various network sites are 4.2–35.5%, 45.5–53.4%, 2.3–26.1%, 44.7–47.8%, and 0.5–0.7, respectively. In addition to $PM_{2.5}$, it is very important to evaluate $PM_{2.5}$ components, because apparent "good performance" for total PM2.5 does not indicate whether the good performance is achieved for "the right reasons" [*U.S. EPA*, 2007]. Table 2 summarizes the performance statistics for annual and seasonal mean 24 h average $PM_{2.5}$ major components. The annual mean NMBs and NMEs are −22.3 to 26.0% and 30.0–76.7% for NH_4^+, −13.0 to 43.5% and 26.5–72.6% for SO_4^{2-}, −33.9% to 46.0% and 55.2–120.3% for NO_3^-, −52.6% to −5.4% and 59.5–64.0% for BC, and −37.6% to 24.8% and 49.5–73.1% for OC at various network sites. Among all PM components from all networks, OC has the best overall performance in all seasons. Simulated SO_4^{2-} and BC also have relatively low biases and errors except at the SEARCH sites. NO_3^- and NH_4^+ have the worst performance, with seasonal mean NMBs, NMEs, MFBs, and MFEs of −47.8% to 143.0%, 51.5–173.7%, −84.7% to 60.0%, and 59.9–125.7%, respectively, for NO_3^- and −40.3 to 67.1%, 24.8–88.8%, −2.4% to 54.0%, and 27.6–88.3%, respectively, for NH_4^+. At the IMPROVE sites, small-to-moderate overpredictions in NH_4^+, NO_3^-, and OC are compensated by small underpredictions in BC and SO_4^{2-}, leading to a small positive bias of 5.8% for annual $PM_{2.5}$. At the SEARCH sites, large underpredictions in BC and OC cannot sufficiently compensate large overpredictions in NH_4^+, NO_3^-, and SO_4^{2-}, leading to a net moderate positive NMB of 35.5% for annual $PM_{2.5}$. These results are generally consistent with those reported by other studies [e.g., *Eder and Yu*, 2006; *Boylan and Russell*, 2006; *S.-Y. Wu et al.*, 2008].

[18] Simulating $PM_{2.5}$ is more challenging than simulating O_3, because of complexities involving multiple components and size ranges of $PM_{2.5}$ and uncertainties associated with model treatments of microphysical and chemical processes related to $PM_{2.5}$. The values of MNB and MNGE recommended by EPA as lower limits for a good performance of $PM_{2.5}$ are ≤15% and ≤30%, respectively [*U.S. EPA*, 2001]. *Seigneur* [2001] suggested MNBs ≤ 50% for $PM_{2.5}$ and SO_4^{2-}. *Zhang et al.* [2006b] used NMBs ≤ 15% and NMEs ≤ 30% for $PM_{2.5}$. *Morris et al.* [2004] suggested MFBs ≤ 50% and MFEs ≤ 75% for $PM_{2.5}$. *Boylan and Russell* [2006] proposed using MFBs ≤ ±30% and MFEs ≤ 50% for $PM_{2.5}$ and MFBs ≤ ±60% and MFEs ≤ 75% for PM major components that have concentrations ≥2.25 $\mu g\ m^{-3}$. In this study, simulated $PM_{2.5}$ has NMBs ≤ 15% at the STN sites in all seasons and at the IMPROVE sites in spring and summer, NMEs > 30% at all sites, MFBs of −9.0% to 32.9%, and MFEs of 40.0–56.2%, indicating a satisfactory or marginally satisfactory performance, depending on the criteria used. For $PM_{2.5}$ components, while MFBs and MFEs of OC and SO_4^{2-} are within the performance thresholds of *Boylan and Russell* [2006], those for NO_3^- and NH_4^+ at the STN and SEARCH sites, and BC at the SEARCH sites exceed the range of thresholds. Nevertheless, the performance of CMAQ in this study is fairly consistent with current PM model performance.

[19] Several reasons may be responsible for the discrepancies between observed and simulated surface O_3 and $PM_{2.5}$. First, uncertainties exist in the emissions of gaseous precursors of O_3 and secondary $PM_{2.5}$ as well as primary $PM_{2.5}$ such as BC and OC. In particular, underpredictions in O_3 and $PM_{2.5}$ over CA and some locations in the eastern U.S. (particularly at the SEARCH sites) may be related to underestimations in wildfire emissions and incorrect temporal profiles of these emissions in those areas [*Roy et al.*, 2007b] since wildfires provide a major source for BC and OC emissions. Second, the overpredictions in $PM_{2.5}$ and its inorganic components at the SEARCH sites may be caused by overestimation in the emissions of precursors such as SO_2, NO_x, and NH_3 [*Zhang et al.*, 2006c]. Third, biases in chemical predictions are affected by biases in the meteorological predictions. *Gilliam et al.* [2006] evaluated MM5 performance for 2001 and reported biases of −1.36, −0.37, −0.24, and −0.35°C for 2 m temperatures during winter, spring, summer, and fall, respectively. Such cold biases can help explain in part the underpredictions of O_3 at the CASTNET sites, particularly the largest underpredictions in winter. Our meteorological evaluation shows underpredictions of 2 m temperatures in all months at the SEARCH sites (figures not shown), which may contribute partially to the overpredictions in $PM_{2.5}$. The overpredictions in precipitation may also contribute in part to the underpredictions in $PM_{2.5}$ at the IMPROVE sites during summer. Finally, the comparison of grid-averaged predictions at a horizontal grid resolution of 36 km with point-wise measurements at a specific site is another source of errors.

[20] Table 2 summarizes performance statistics for annual mean weekly total wet deposition fluxes of SO_4^{2-}, NO_3^-, and NH_4^+ ($SO_{4\,wet}^{2-}$, $NO_{3\,wet}^-$, and $NH_{4\,wet}^+$, respectively). $SO_{4\,wet}^{2-}$ is overpredicted for all seasons except for summer, $NO_{3\,wet}^-$ is overpredicted in winter and fall but underpredicted in spring and summer, and $NH_{4\,wet}^+$ is overpredicted for all seasons, with annual mean NMBs and NMEs of 19.6% and 71.3% for $SO_{4\,wet}^{2-}$, −13.0% and 71.6% for $NO_{3\,wet}^-$, and 31.6% and 92.8% for $NH_{4\,wet}^+$. No performance criteria are recommended for wet deposition fluxes by the U.S. EPA. When compared to other studies on simulated wet deposition fluxes [e.g., *Gilliland et al.*, 2006; *Queen and Zhang*, 2008], these values are within the ranges of NMBs and NMEs reported and also within the ranges of NMBs and NMEs expected for $PM_{2.5}$ and its composition. Thus, they are satisfactory on an annual mean basis, although some seasonal mean NMBs and NMEs are fairly large (e.g., those for wet deposition fluxes of NH_4^+ in winter). MM5 overpredicts precipitation with an NMB of 24.4% in summer but underpredicts it with NMBs of −6.0%, −2.8%, and −13.1% for winter, spring, and fall, respectively, leading to an overall NMB of 1.6% in annual mean values. The lack of correlation between precipitation and wet deposition indicates that factors other than precipitation dominate the trends of wet deposition. For example,

SO_4^{2-} concentrations are underpredicted (NMBs of −0.5% to −7.7%) at all network sites (except at the SEARCH sites) in summer, resulting in insufficient amount for wet scavenging (with an NMB of −14.4% for wet deposition) even though precipitation is moderately overpredicted (with an NMB of 25.0%). By contrast, SO_4^{2-} concentrations are overpredicted (NMBs of 0.7−59.3%) at all network sites in fall, resulting in more than sufficient amount for wet scavenging (with an NMB of 63.3% for wet deposition) even though precipitation is moderately underpredicted (with an NMB of −13.1%). The interplay between the meteorological (i.e., precipitation) and the chemical (i.e., ambient precursor concentrations in gas and aqueous phase) forcing in controlling wet deposition fluxes was studied for August and December 2002 over NC by *Queen and Zhang* [2008]. They found that a correlation exists between wet deposition fluxes and precipitation in both months, with stronger magnitudes in August, because the summer convective precipitation events having larger intensities and therefore the meteorological forcing is expected to dominate August correlations. In this study, their correlation on domain-wide seasonal or annual basis is much weaker, largely because the correlation signals at in situ locations at a local scale may have been wiped out after averaging of values over a large domain for a long time period.

4.2. Temporal Variations

[21] Figure 3 compares simulated temporal variations of the maximum 8 h O_3 concentration against available observations at 20 sites including 8 SEARCH sites, 12 AIRS-AQS sites (one of which is also a SEARCH site) and 5 CASTNET sites (four of which are also AIRS-AQS sites). Overall, CMAQ can reproduce well the daily variations and magnitudes of maximum 8 h O_3 at most locations, with overpredictions at two sites (i.e., LAX for all months and BBE during January−May) and underproductions at a few sites (e.g., NYC during January−May, RIV during May−October, and GRS throughout the year). LAX and RIV are located in the south coast of CA, a well-known area with complex terrain, dry summer/light rainfall, and microclimate with a strong spatial temperature gradient and the effect of sea breezes (e.g., the temperatures of 18°F (10°C) warmer in the inland area than the coastal, and a temperature gain of over one degree per mile inland). NYC is the largest city in the U.S. Summers are typically hot and humid with frequent thunderstorms, winters are cold with 25 inches snowfall but sea breezes keep temperatures slightly milder than inland regions. Spring and fall are erratic, and can range from cool to hot. The BBE site is located in the Chihuahuan Desert. Because of the range in altitude from ∼1,800 ft (550 m) along the river to 7,800 ft (2,400 m) in the Chisos Mountains, BBE exhibits dramatic weather changes, e.g., dry and hot in late spring and summer with temperatures often exceeding 100 °F (38 °C) in the lower elevations, winters are normally mild throughout the park, but subfreezing temperatures occasionally occur. GRS straddles the ridgeline of the Great Smoky Mountains, as part of the Blue Ridge Mountains which are a division of the larger Appalachian Mountain chain. Elevations in the park range from 876 ft (267 m) at the mouth of Abrams Creek to 6,643 ft (2,025 m) at the summit of Clingmans Dome. Very high humidity and precipitation often occur in the park, with 95% areas of deciduous, temperate, and old growth forest. Large biases in model predictions may be attributed to the inability of MM5 in capturing major meteorological and topography characteristics such as complex terrain, land-sea air exchange, impact of sea breezes, extreme wet or dry weathers, and strong diurnal variations in temperatures and mixing heights over these areas [e.g., *Zhang et al.*, 2006a; *Queen et al.*, 2008] as well as the uncertainties in emissions used at these sites.

[22] Figure 4 compares simulated temporal variations in the concentrations of 24 h average $PM_{2.5}$ against available observations at 12 sites including 8 SEARCH sites and 4 IMPROVE sites. CMAQ reproduces higher $PM_{2.5}$ over SEARCH sites and lower values at the IMPROVE sites, with magnitudes generally comparable with observations (mostly within a factor 2). Overpredictions dominate at the SEARCH sites (particularly at JST and YRK), and slight underpredictions occur at two IMPROVE sites: GRS during May−Sept. and BBE during Jan.−Jul., largely caused by the aforementioned uncertainties associated with meteorology and emissions. For example, the cold biases in the 2 m temperature predictions at JST and YRK throughout the year can explain, in part, the overpredictions in secondary PM. Figures 5 and 6 show temporal variations of three major inorganic $PM_{2.5}$ species (NH_4^+, SO_4^{2-}, and NO_3^-) at 2 SEARCH sites (JST and YRK) and 3 CASTNET sites (PSU, GRS, and BBE) and two organic $PM_{2.5}$ species (BC and OC) at 4 SEARCH sites and 2 IMPROVE sites, respectively. The concentrations of NH_4^+ and NO_3^- are moderately-to-significantly overpredicted at JST and YRK for all seasons except for winter. While CMAQ is able to reproduce total nitrate (TNO_3), it is known to give either too much or too little nitrate predictions [*Yu et al.*, 2005; *Zhang et al.*, 2006c]. Several factors may contribute to moderate-to-significant overpredictions of NO_3^- in spring, summer, and fall at the SEARCH sites and in spring and fall at the IMPROVE sites including underestimation in vertical mixing, too fast rate of gas-phase N_2O_5 hydrolysis, too large reaction probability (γ) value of heterogeneous N_2O_5 hydrolysis, an overproduction in TNO_3, as well as underestimation in dry and wet deposition. For example, the values of γ used for heterogeneous N_2O_5 hydrolysis on aerosol surfaces used in all CMAQ versions prior to v4.6 are highly uncertain and led to overpredictions in both summer and winter, with a stronger impact on winter predictions [*Pleim et al.*, 2003; *Bhave et al.*, 2006; *Gilliland et al.*, 2006]. Changing a constant γ value of 0.1 in CMAQ v4.2.2 to account for its dependence on the concentrations of SO_4^{2-} and NO_3^- in CMAQ v4.3 reduced NO_3^- concentrations by factors of 2−5 [*Pleim et al.*, 2003]. In addition to the dependence on the concentrations of SO_4^{2-} and NO_3^- that was considered in CMAQ v4.3−4.5.1, the values of γ was modified to account for their dependence on temperature and RH in v4.6, which reduced total NO_3^- by 8−16% in winter [*Bhave et al.*, 2006]. For moderate-significant underpredictions at the STN sites for all seasons, in winter and summer at the IMPROVE sites, and in summer at the CASTNET sites, contributing factors may include the equilibrium assumption used in the partitioning of total NO_3^-, the neglect of condensational growth of NO_3^-, and the missing heterogeneous uptake of HNO_3 by aerosols and aqueous-phase kinetic reactions that lead to NO_3^- formation [*Zhang et al.*, 2006c]. Simulated SO_4^{2-} compares relatively well with observations, except on

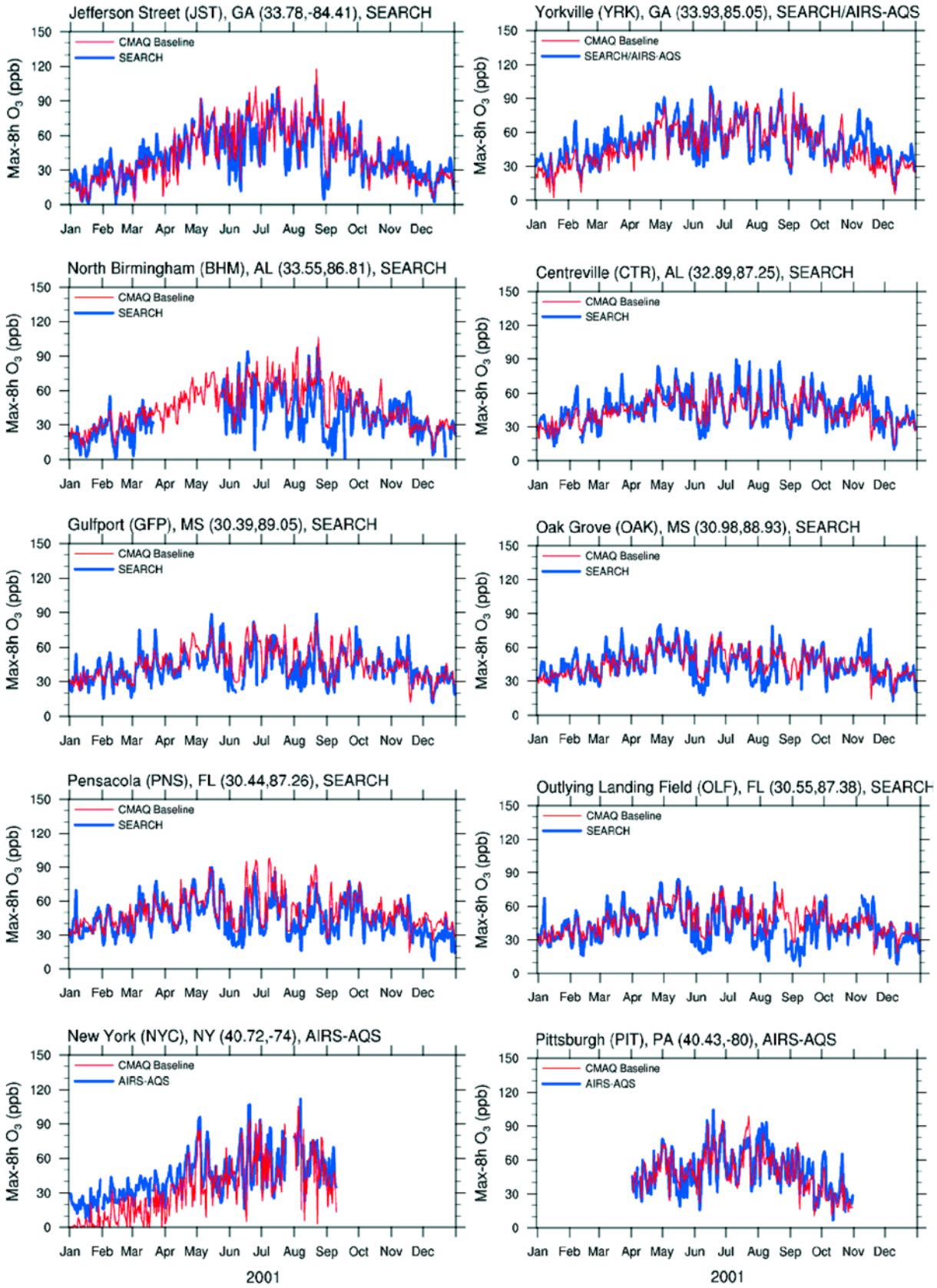

Figure 3. Daily maximum 8 h average O_3 at 20 sites including 8 SEARCH sites, 11 AIRS-AQS sites, and 1 CASTNET site in 2001. Note that 4 of the 11 AIRS-AQS sites, i.e., BBE, GRS, OLY, YEL, are also CASTNET sites.

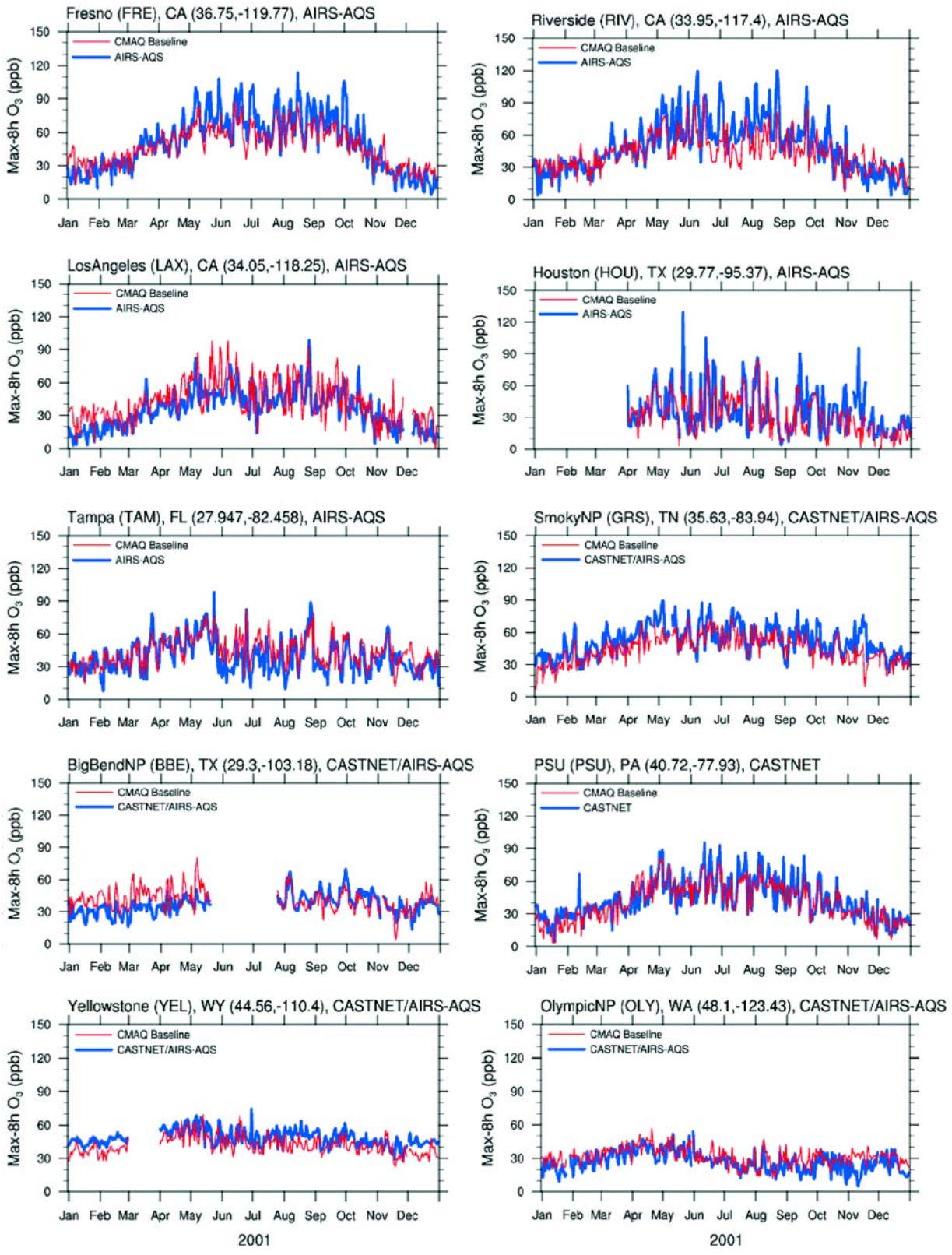

Figure 3. (continued)

Nov. 16 during which CMAQ gave a high spike. While very high $PM_{2.5}$ concentrations are observed on Nov. 16 at both sites, they are dominated by different components. At JST, OC dominates (total OM accounts for ~61% of $PM_{2.5}$) and the sum of the five observed major components can explain > 93% of observed $PM_{2.5}$. By contrast, the sum of the four observed major components (note that no BC measurements are available) can only explain ~42% of observed $PM_{2.5}$ at

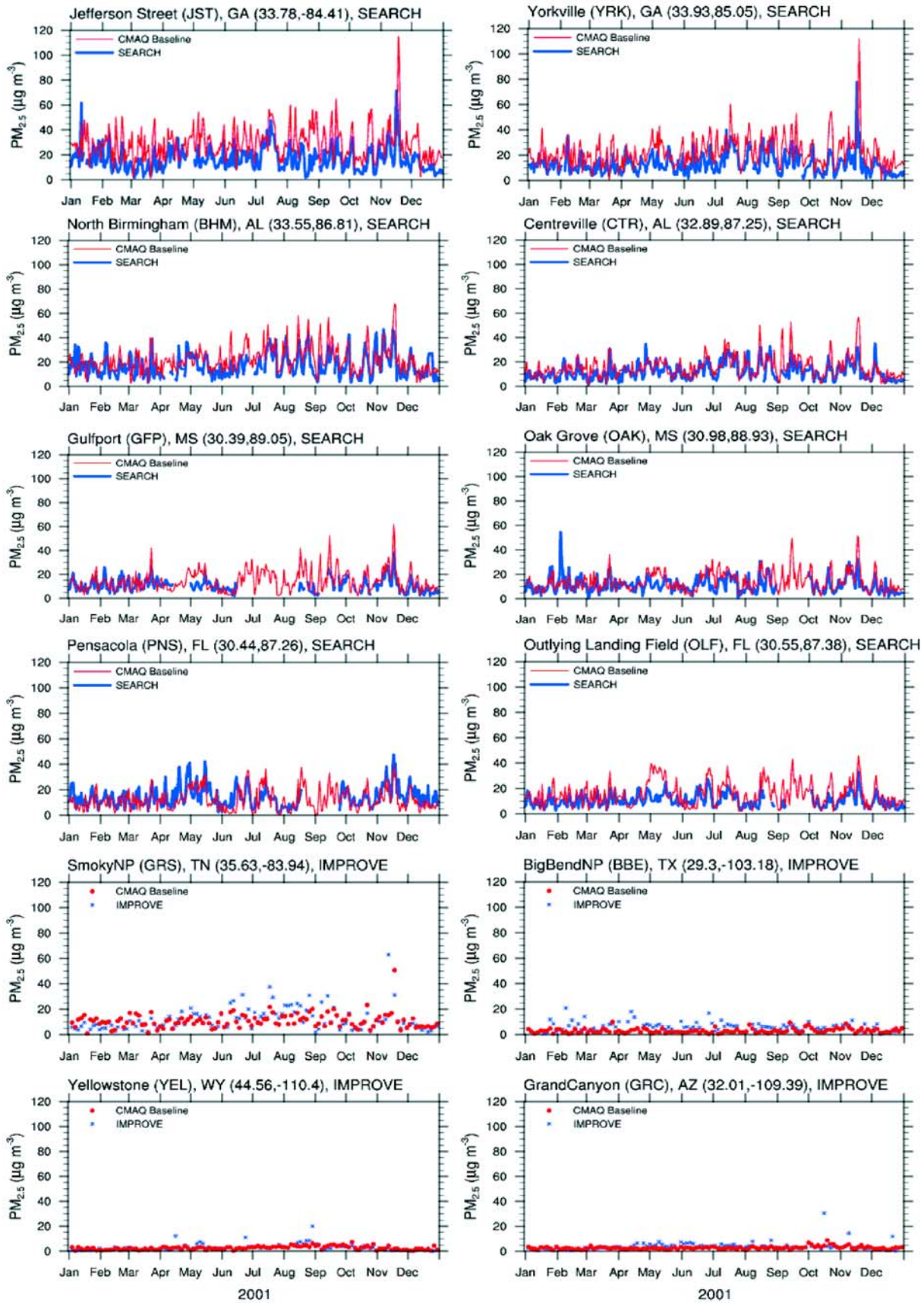

Figure 4. The 24 h average $PM_{2.5}$ at 12 sites including 8 SEARCH sites and 4 IMPROVE sites in 2001. SEARCH data are available every day (indicated by continuous lines), and IMPROVE data are available every 3 days (indicated by dots).

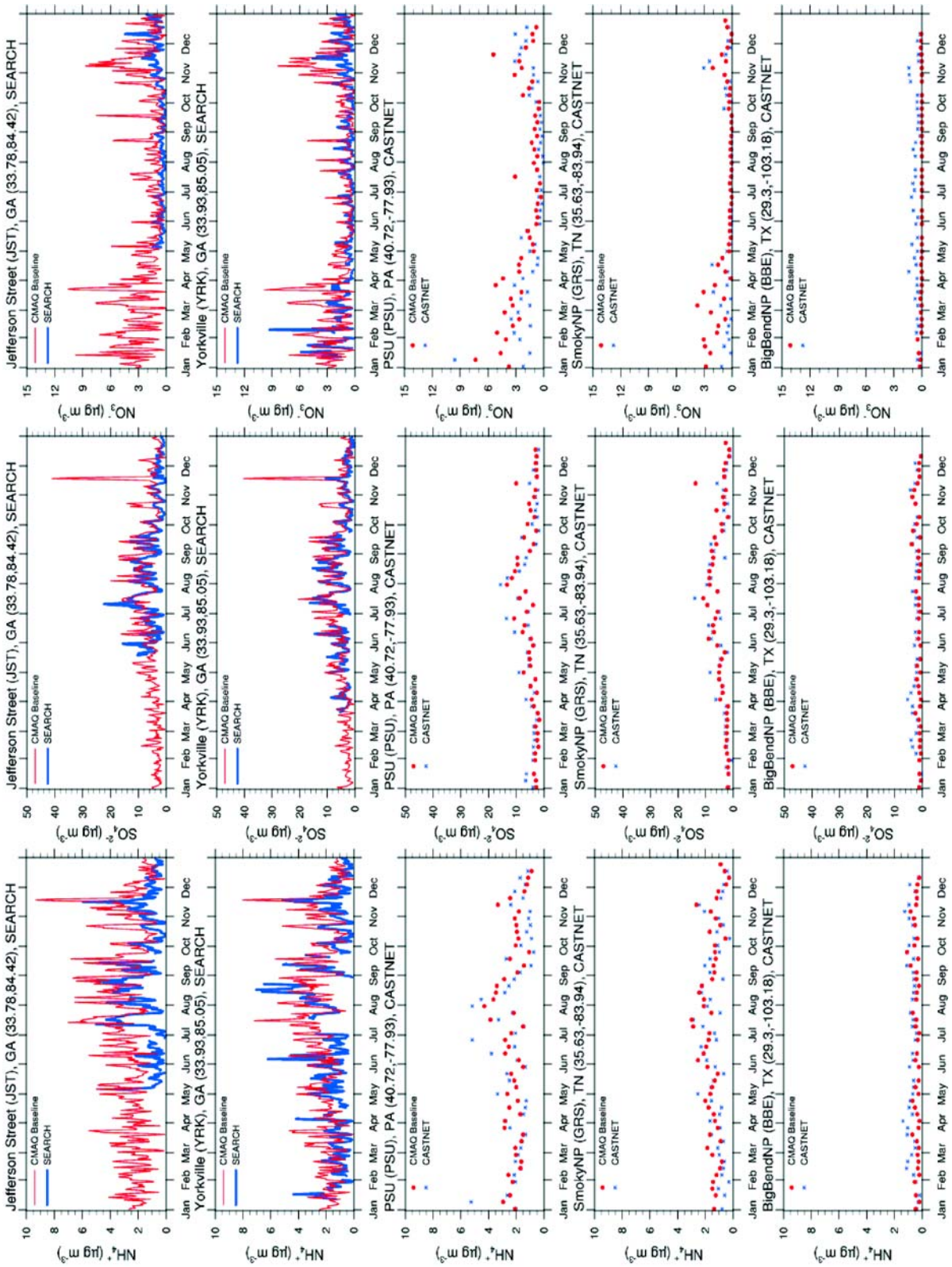

Figure 5. The 24 h average NH_4^+, SO_4^{2-}, and NO_3^- at 5 sites including 2 SEARCH sites (JST and YRK) and 3 CASTNET sites (PSU, GRS, and BBE) in 2001. SEARCH data are available every day (indicated by solid lines), and CASTNET data are available every 3 days (indicated by dots).

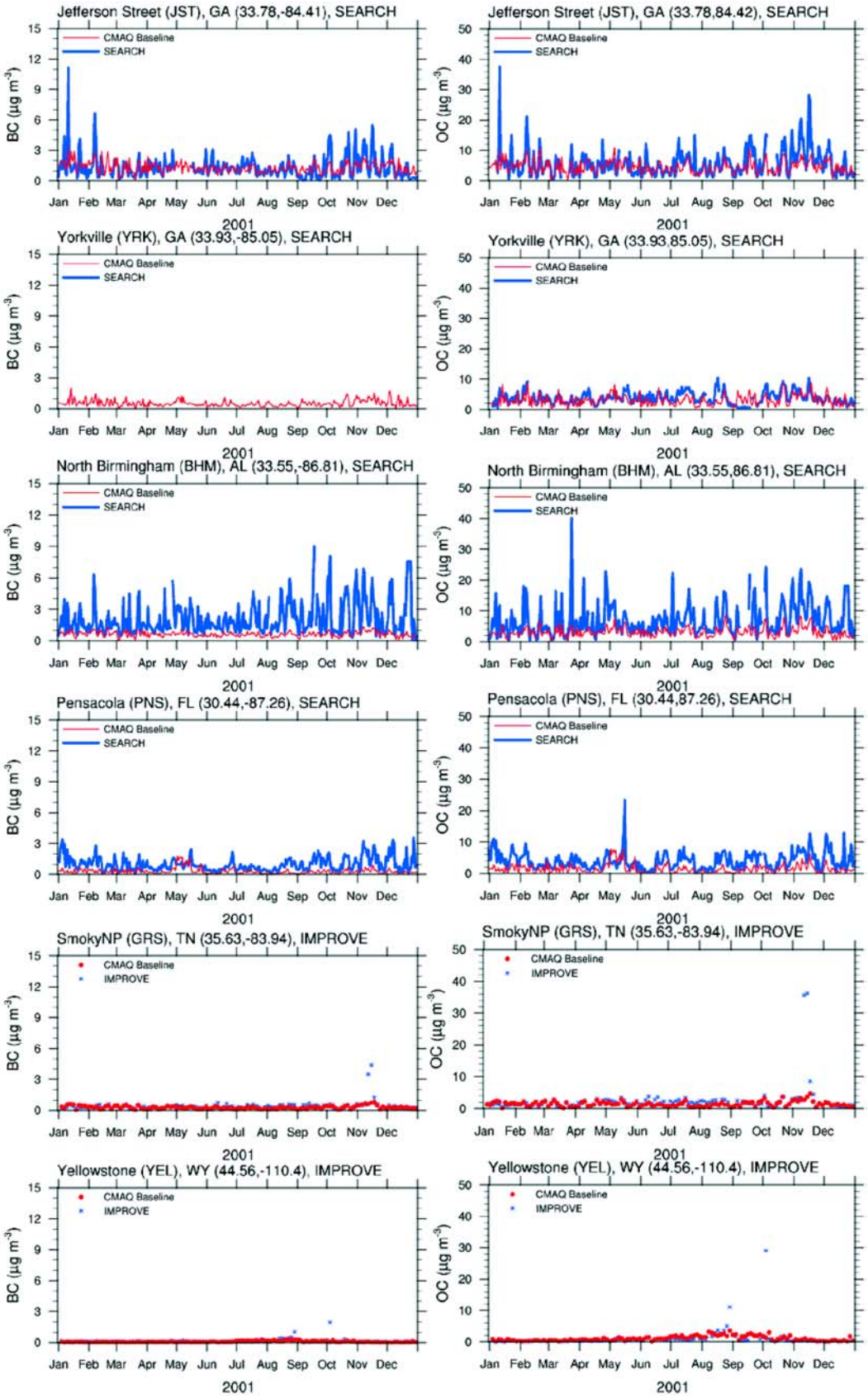

Figure 6. The 24 h average $BC_{2.5}$ and $OC_{2.5}$ at 6 sites including 4 SEARCH sites (JST, YRK, BHM, and PNS) and 2 IMPROVE sites (GRS and YEL) in 2001. SEARCH data are available every day (indicated by solid lines), and IMPROVE data are available every 3 days (indicated by dots). No observed BC data are available at YRK.

YRK, indicating that other unknown inorganic PM may dominate $PM_{2.5}$ on Nov. 16 at this site (this, however, cannot be directly verified, as observed other unknown inorganic PM are not available). While CMAQ reproduces high peaks in $PM_{2.5}$ with 42–60% overpredictions on Nov. 16 at both sites, it fails to reproduce observed high OC at JST and gives much higher SO_4^{2-} than observations at both sites, thus reproducing the $PM_{2.5}$ peaks for a wrong reason. This further indicates the need to evaluate not only $PM_{2.5}$ but also its components, because a good performance for PM components will indicate a good performance for total $PM_{2.5}$, although the reverse does not always hold (as for this case). Comparison of observed vs. simulated BC and OC at JST, BHM, and PNS shows moderate-to-large underpredictions, indicating the possible underestimation of primary BC and OC emissions.

[23] Figures 7 and 8 compare simulated with observed indicator species including NO_y, O_3/NO_x, and O_3/NO_y during afternoon hours (noon to 6 pm) in January and August 2001 at 6 SEARCH sites. The observed values of O_3/NO_z are sparse and thus not included. The afternoon values of $NO_y < 20$ ppb, $O_3/NO_x > 15$ ppb, and $O_3/NO_y > 7$ ppb indicate a NO_x-limited O_3 chemistry, other values above or below these transition values indicate a VOC-limited O_3 chemistry, and transition values represent the ridge line at which O_3 chemistry is equally sensitive to NO_x and VOCs. These transition values are also plotted as horizontal lines in Figures 7 and 8. In January 2001, most observed NO_y values are greater than 20 ppb at JST and BHM but less than 20 ppb at other sites except GFP where NO_y values are above this threshold on some days (e.g., January 1–6, 9–13, and 21–16), indicating a dominancy of VOC-limited O_3 chemistry at JST and BHM throughout January and at GFP on some days and a dominancy of NO_x-limited O_3 chemistry at GFP on some days and at other sites except for JST and BHM throughout January. Other three indicators, on the other hand, all represent a VOC-limited O_3 chemistry at all sites, indicating some self-inconsistencies among these indicators and/or uncertainties associated with the transition values. Given uncertainties in the transition values used in different studies, such inconsistencies are not surprising. *Lu and Chang* [1998] reported a significant discrepancy between their model-derived threshold criteria and those proposed by *Sillman* [1995], e.g., VOC-limited conditions are found to be associated with $NO_y > 5$ ppb by *Lu and Chang* [1998] but with $NO_y > 20$ ppb by *Sillman* [1995]. While *Sillman* [1995] proposed the ratio of the production rates of H_2O_2 and HNO_3 (i.e., P_{H2O2}/P_{HNO3}) < 0.4 for VOC-limited chemistry, *Tonnesen and Dennis* [2000a] found a value of $P_{H2O2}/P_{HNO3} < 0.06$ to 0.07 is more appropriate for their application. Some indicators may be more robust than others under certain conditions. For example, *Tonnesen and Dennis* [2000b] found that $H_2O_2/(O_3+NO_2)$, H_2O_2/HNO_3, O_3/NO_x, and $HCHO/NO_2$ perform better than $HCHO/NO_y$ over the NYC area in summer. *Lu and Chang* [1998] found that O_3/NO_z, $HCHO/NO_y$, and H_2O_2/HNO_3 are effective indicators over San Joaquin Valley, CA in summer. *Hammer et al.* [2002] have shown that H_2O_2/HNO_3 is a robust indicator over Germany in summer. A large set of indicators should therefore be used to collectively distinguish NO_x and VOC-limited conditions. Simulated indicator values are consistent with observations at most sites during most time periods in January. Exceptions are at BHM and GPF where simulated NO_y values are mostly below 20 ppb and values of other indicators are all below the transition values but with magnitudes higher than observations. In August, most observed NO_y values at JST and BHM are still above 20 ppb and observed O_3/NO_x and O_3/NO_y values are mostly below the threshold, indicating VOC-limited O_3 chemistry. Observed NO_y values are mostly below 20 ppb and those of other indicators are mostly above the threshold at other sites, indicating NO_x-limited O_3 chemistry. Simulated indicator values are overall consistent with observations in both months at all sites except for BHM where simulated NO_y values are too low, and simulated values of O_3/NO_x and O_3/NO_y are too high. The inaccuracies in simulated indicator values at BHM are likely caused by too little NO_x, and NO_y in the gas phase but too much nitrate in the particulate phase resulting from a cold bias in 2 m temperatures throughout the year at BHM and too much scavenging of soluble NO_y species in summer resulting from an overprediction in precipitation.

[24] Although the aforementioned indicator ratios are developed for O_3 chemistry, they can provide insight into the sensitivity of secondary PM to changes in VOCs, or NO_x emissions, given the commonality of the gaseous precursors of O_3 and secondary PM. Additional indicator species/ratios have been developed for secondary PM formation. For example, *Ansari and Pandis* [1998] developed a gas ratio, GR:

$$GR = \frac{[TNH4] - 2[TSO4]}{[TNO3]} = \frac{[NH_3] + [NH_4^+] - 2[SO_4^{2-}]}{[NO_3^-] + [HNO_3]} \quad (7)$$

where [TNH4] is the total molar concentration of NH_3, = $[NH_3] + [NH_4^+]$, [TSO4] is the total molar concentration of SO_4^{2-}, = $[SO_4^{2-}]$, [TNO3] is the total molar concentration of NO_3^-, = $[NO_3^-] + [HNO_3]$, and [TNH4]-2[TSO4] indicates free NH_3. *Ansari and Pandis* [1998] applied the GR ratio to predict how PM mass responds to reductions in SO_4^{2-}, NO_3^-, and NH_3 under various ambient temperature, RH, and chemical conditions (e.g., the concentrations of TSO4, TNO3, and TNH4). For example, under conditions with T = 298 K, RH = 90%, values of GR < 0.5, 0.5–1.2, or >1.2 at TNO3 concentration = 10 ppb, and TNH4 $\leq$ 40 ppb indicate that PM responses to an increase in TSO4 are approximately constant (e.g., $d[PM_{2.5}]/d[TSO4] = -0.28$), nonlinear enhanced increase (e.g., $d[PM_{2.5}]/d[TSO4]$ is in the range of −0.28 to 1.34), or linear constant increase (e.g., $d[PM_{2.5}]/d[TSO4] = 1.34$), respectively, a value of GR < 1, 1–3, or >3 at TSO4 concentration = 2.5 ppb, and TNO3 concentration = 5 ppb indicates that PM responses to an increase in NH_3 is nonlinearly near-constant increase, nonlinearly reduced increase, or insensitive, respectively, a value of GR < 0.4, 0.4–2, or >2 at TSO4 concentration = 2.5 ppb, and TNH4 concentration = 5 ppb indicates that PM responses to an increase in NO_3^- is insensitive, nonlinearly enhanced increase, and linearly near-constant increase, respectively. A value of GR < 1 or >1 also indicates that NO_3^- concentration is most sensitive to changes in NH_3 because of abundance of TNO4 or most sensitive to changes in TNO4 because of abundance of free NH_3 [*Pinder et al.*, 2008]. Equation (7) assumes that SO_4^{2-} is fully neutralized by

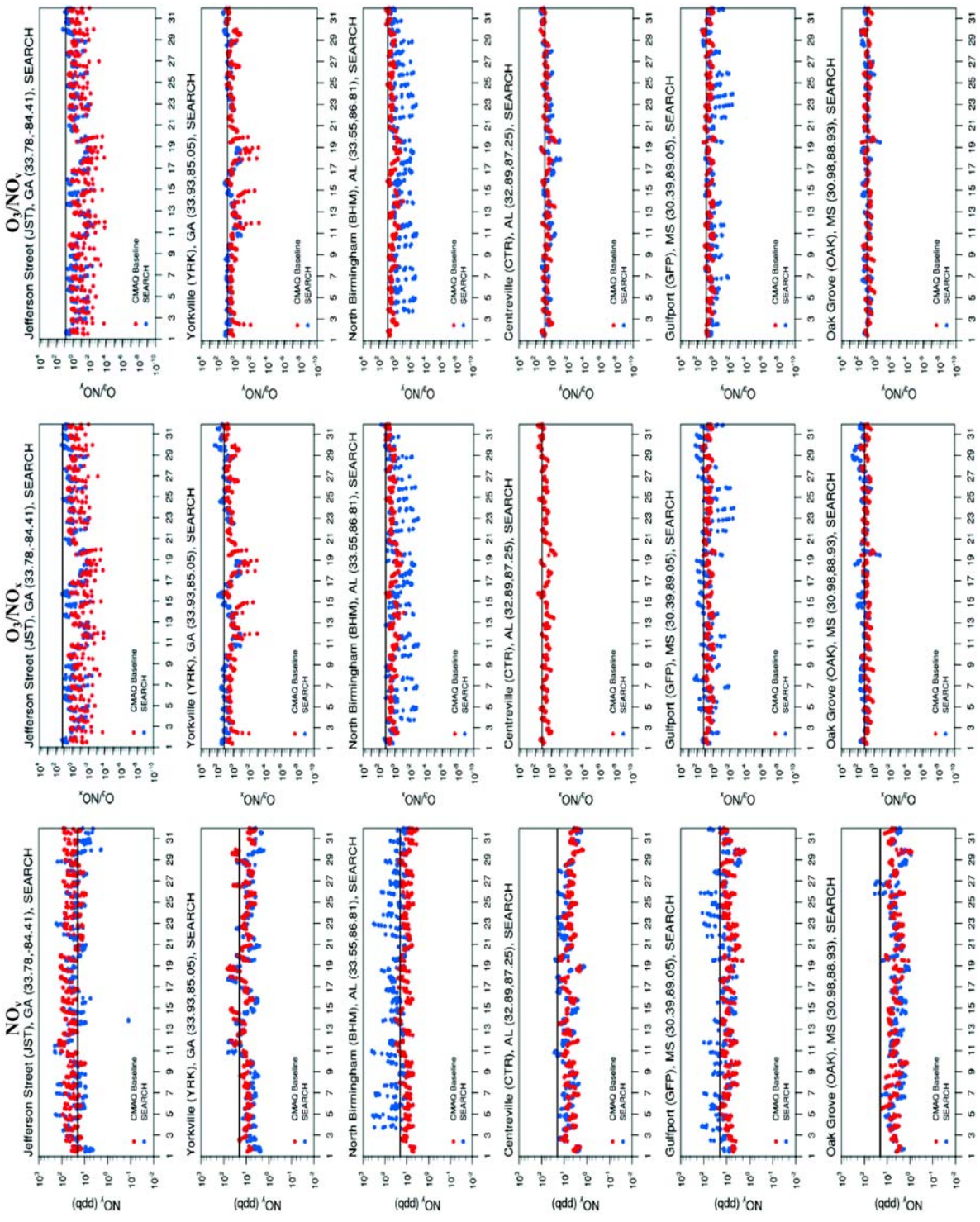

Figure 7. Observed versus simulated indicators at 6 SEARCH sites in January 2001. Observed NO_x mixing ratios are not available at CTR. The horizontal line indicates the threshold; for example, the values of $NO_y < 20$ ppb, $O_3/NO_x > 15$ ppb, and $O_3/NO_y > 7$ ppb indicate a NO_x-limited O_3 chemistry, and other values above or below these transition values indicate a VOC-limited O_3 chemistry.

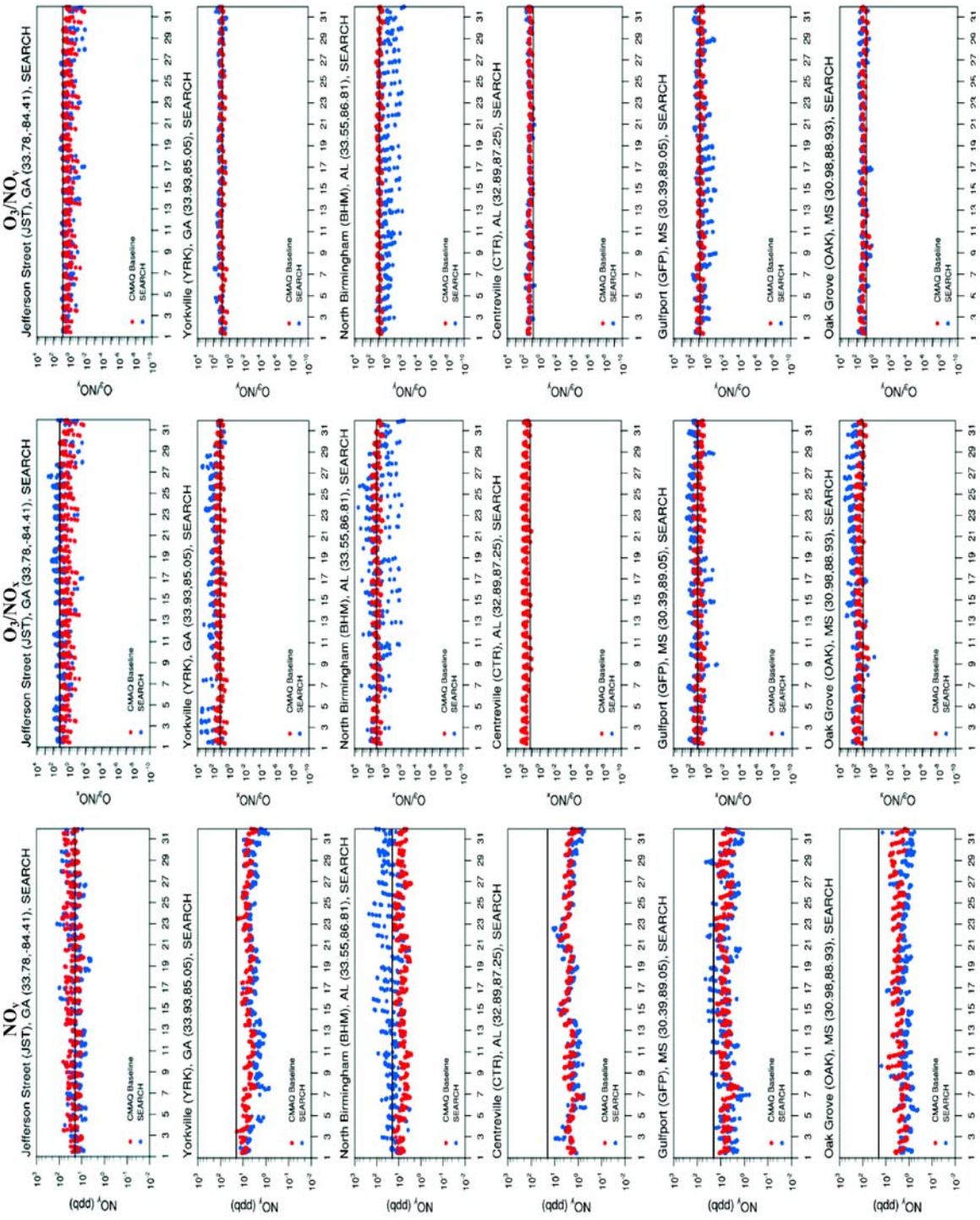

Figure 8. Observed versus simulated indicators at 6 SEARCH sites in August 2001. Observed NO_x mixing ratios are not available at CTR. The horizontal line indicates the threshold, for example, the values of $NO_y < 20$ ppb, $O_3/NO_x > 15$ ppb, and $O_3/NO_y > 7$ ppb indicate a NO_x-limited O_3 chemistry, other values above or below these transition values indicate a VOC-limited O_3 chemistry.

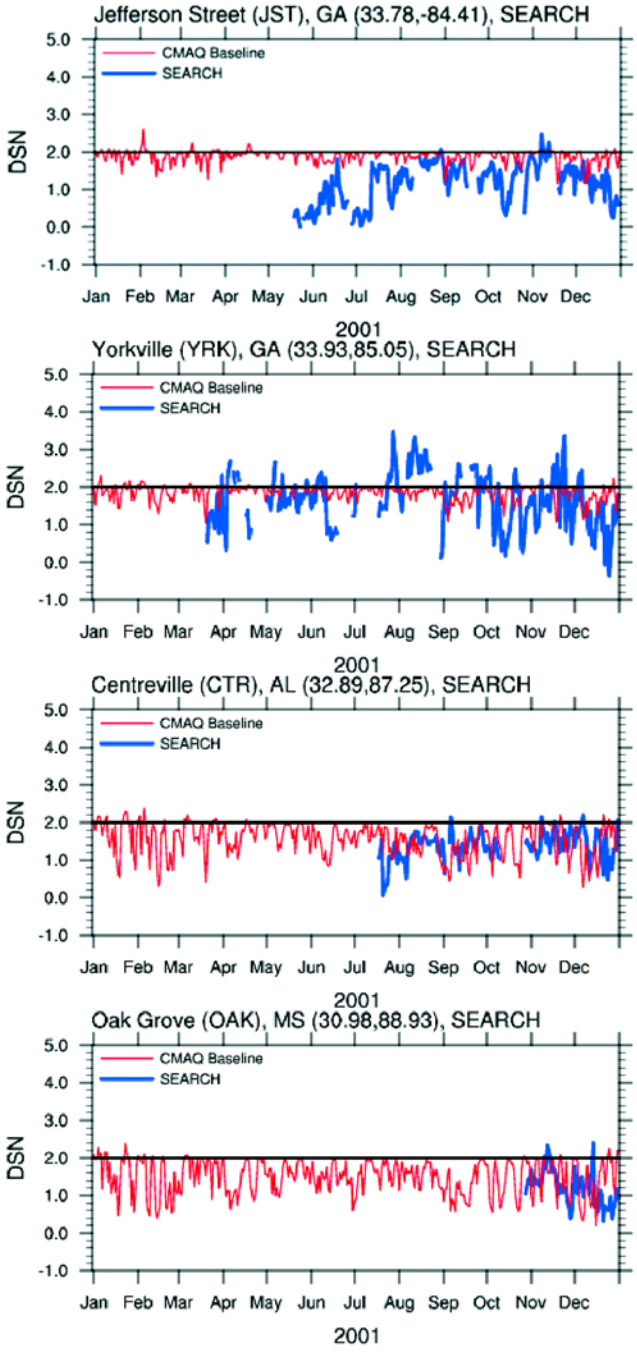

Figure 9. Observed versus simulated degree of sulfate neutralization (DSN) at 4 SEARCH sites: JST, YRK, CTR, and OAK in 2001.

NH_3, which may not be valid under all ambient conditions. *Pinder et al.* [2008] used a ratio to indicate the degree of sulfate neutralization (DSN):

$$DSN = \frac{\left[NH_4^+\right] - \left[NO_3^-\right]}{\left[SO_4^-\right]} \quad (8)$$

A value of DSN of <2 or ≥2 indicates insufficient or full neutralization. The use of DSN allows the GR ratio in equation (7) to be extended to an adjusted GR ratio (AdjGR) that is applicable to conditions with insufficient neutralization of sulfate:

$$AdjGR = \frac{[\mathrm{TNH4}] - \mathrm{DSN} \times [\mathrm{TSO4}]}{[\mathrm{TNO3}]} = \frac{[\mathrm{NH_3}] + [\mathrm{NO_3^-}]}{[\mathrm{NO_3^-}] + [\mathrm{HNO_3}]} \quad (9)$$

AdjGR shows an increased accuracy as an indicator of PM nitrate response to changes in SO_2 or NH_3 emissions, particularly in summer conditions in the eastern U.S. [*Pinder et al.*, 2008]. *Blanchard* [2000] has also developed the excess NH_3 (EA) indicator that is similar to the GR. While the GR or the AdjGR are useful indicators for the sensitivity of second PM formation to their precursor emissions, they require measurements of TNO3, TNH4, TSO4, or $[NH_3]$g, $[HNO_3]$g, $[NO_3^-]$, none of the existing surface network contain a complete set of those values in 2001, therefore they cannot be evaluated. However, DSN can be evaluated given availability of PM composition measurements. Figure 9 compares simulated vs. observed DSN at four sites where observations are available: JST, YRK, CTR, and OAK. While simulated DSN values are in relatively good agreement with observations at CTR and OAK, they are much higher than observations at JST, either higher or lower at YRK. Both simulation and observation have DSN< 2 during most months, indicating that SO_4^{2-} at those sites is mostly insufficiently neutralized and NH_4NO_3 may be formed to neutralize free NH_3. Under such conditions, the response of PM concentrations to changes in TSO4 is nonlinear [*Ansari and Pandis*, 1998].

[25] Overall, CMAQ is able to reproduce observed ratios of indicator species within the same range of ratios, providing some confidence that the predicted change in O_3 and $PM_{2.5}$ may be accurate. Considering the relatively coarse horizontal grid resolution used in the model simulation, CMAQ's performance in reproducing observed temporal variations for O_3, indicator species, $PM_{2.5}$ and $PM_{2.5}$ species appears reasonable.

5. Evaluation With Satellite Measurements

5.1. Tropospheric Column Mass Abundance

[26] Figure 10 shows the observed and simulated seasonal mean tropospheric CO column abundance. The overall statistical performance of CMAQ for seasonal and annual mean CO and other column variables is summarized in Table 2. The seasonal mean MOPITT CO column abundance is the highest in spring, followed by winter, fall, and summer. Elevated CO column masses were observed over Pacific and west coast of the U.S. in spring and winter, which can be partially attributed to trans-Pacific transport of CO during the 2001 spring Asian dust storm. A number of field and modeling studies have reported elevated CO, O_3, peroxyacetylnitrate (PAN), SO_4^{2-}, $PM_{2.5}$, and PM_{10} at surface and in the free troposphere in the western U.S. during springtime [e.g., *Jacob et al.*, 1999; *Jaffe et al.*, 2003; *Heald et al.*, 2006; *Wang et al.*, 2009]. The model reproduces well the observed seasonal variability, with higher CO column mass in winter and spring and lower one in fall and summer. Underpredictions occur over the entire domain in spring, overpredictions occur over the eastern U.S. in summer, and both occur in some regions in winter. The NMBs and NMEs for simulated CO column masses are −6.7% to 26.0% and 7.9−29.1%, respectively. Possible

sources of errors for CMAQ simulation may include uncertainties in the CO emissions and boundary conditions used. Figure 11 shows annual mean CO vertical profiles from MOPITT and CMAQ as well as the standard deviations (with respect to a spatial and/or temporal scale) of the retrievals at four sites (i.e., CHI, HOU, LAX, and GRC) and those based on the domain-wide average for four seasons. The CMAQ profiles compare well with those retrieved from MOPITT radiances at all sites and domain-wide. The largest discrepancies occur at CHI with overpredictions in spring and summer dominating the annual overprediction. Differences between simulated and satellite values are most apparent in spring and summer at all sites and in the domain-wide seasonal mean profiles, consistent with the column CO comparison (see Figure 10). While domain-wide underpredictions occur from surface to 250 mb in spring, domain-wide overpredictions occur throughout the vertical domain in summer.

[27] Figure 12 shows the GOME-observed and CMAQ-simulated seasonal mean tropospheric NO_2 column abundance. CMAQ generally reproduces well the seasonal variations in the NO_2 column abundance, with higher values in winter than other seasons and in the eastern U.S. than in the western U.S. The model underpredicts the NO_2 column abundance with an NMB of −7.5% for summer, and overpredicts that in winter, spring, and fall with NMBs of 18.4%, 5.5%, and 5.7%, respectively. NMEs range from 39.8–53.6%, and correlation coefficients range from 0.74 to 0.87. Despite overall good performance, some discrepancies exist in the magnitudes and spatial distributions, particularly over the eastern U.S. For example, CMAQ gives higher column NO_2 over source regions in the eastern U.S. in winter, but lower values in other seasons over most of the areas in the eastern U.S. (except for a few major cities where simulated column NO_2 values are higher than observations), consistent with some recent studies [e.g., *Stockwell et al.*, 2007]. These discrepancies can be attributed to several factors including uncertainties in the NO_2 emissions, the NO_2 oxidation rates simulated by the CB-IV gas-phase mechanism, and the boundary conditions used. Also, the use of a plume-in-grid treatment for large U.S. power plants has been shown earlier to result in improved column NO_2 performance (vs. GOME data) in summer [*Vijayaraghavan et al.*, 2009]; such a treatment was not employed in this study. Moreover, several studies indicate that power plant and industry NO_x emissions over the eastern U.S. decreased by 50% between 1999 and 2003/2004 [e.g., *Frost et al.*, 2006; *Hudman et al.*, 2007]. It is not clear how accurate the NEI v3.0 used in this study represents the actual emissions over the U.S. in 2001.

[28] Figure 13 shows the observed TOR/TCO and simulated seasonal mean total tropospheric O_3 abundance in DU for 2001. The observed TOR/TCO values are consistent in terms of magnitude (ranging from 28.71–42.8 DU and 29.1–41.4 DU, respectively), although differences exist in their spatial distributions, particularly in winter and spring. Such differences can be attributed to several factors, including different instruments and retrieval algorithms. The simulated TOR and TCO over the domain range from 35.31 to 47.04 DU and from 35.1 to 46.5 DU, respectively. Moderate overpredictions occur in winter and spring and underpredictions occur in summer, due to the use of inaccurate upper boundary layer conditions for O_3 and the uncertainties in the gas-phase chemistry. The NMBs and NMEs range from −13.0% to 34.9% and 18.1% to 35.7% for TOMS/SBUV, and −7.0% to 25.0% and 20.0% to 29.0% for GOME, respectively. Figure 14 shows monthly mean time series plots of GOME and CMAQ TCO at three sites that usually experience elevated O_3: CHI, HOU, and LAX. CMAQ TCO is typically comparable to GOME TCO (±10 DU) except in summer when the CMAQ TCO is consistently lower than GOME TCO. In contrast, at HOU in spring, GOME TCO is lower than the modeled TCO. This is consistent with the results of *Liu et al.* [2006a], in which some of the bias in simulated TCO were attributed to the upper tropospheric biases that were partly caused by a large midlatitude spatio-temporal variability under the stratospheric influence in winter and spring.

5.2. AODs

[29] Figure 15 shows the observed and simulated seasonal mean total column AODs. The highest monthly mean MODIS AODs are up to 0.4 in summer in the eastern U.S. and up to 0.6 in spring over the western U.S. As compared with other seasons, the Pacific coast off the Pacific west has elevated values of 0.2–0.3 in spring, as a result of a major Asian dust storm in April 2001. *Chu et al.* [2005] reported that the highest monthly mean MODIS AODs of 0.4–0.7 occurred in the latitude band associated with Asian dust outbreaks in spring 2001. *Heald et al.* [2006] reported more than 50% of elevated surface SO_4^{2-} during the same period in the northwest U.S. from the IMPROVE network. A persistent Asian fine dust concentration of 1.2 $\mu g\ m^{-3}$ was simulated at the surface in the western U.S. in spring 2001 [*Fairlie et al.*, 2007]. The dust and elevated SO_4^{2-} from long range transport of Asian dust should therefore have contributed to the elevated MODIS-derived AODs along the west coast of U.S. Among all seasons, MODIS-derived AODs over Atlantic Ocean off the east coast of U.S. and around Cuba are the highest in summer, reflecting the highest impact from long range transport of African Soudano-Sahel and Saharan dust that typically peaks in summer [*Prospero et al.*, 1981; *Cakmur et al.*, 2001; *Kaufman et al.*, 2005]. The monthly aerosol composites for 2001 derived based on MODIS in Figure 5 of *Kaufman et al.* [2005] support the fact that these high AODs are primarily due to smoke particles over Atlantic off the east coast of U.S. in February–September and primarily due to dust particles around Cuba in April–September, which are consistent with high MODIS-derived AODs over these regions in spring and summer months shown in Figure 15. As expected, the CMAQ estimated total AODs are the highest in summer because of the highest PM, consistent with observations. However, CMAQ fails to reproduce the observed spatial distributions and magnitudes of AODs over CONUS, Cuba, and the Pacific and Atlantic oceans. Moderate underpredictions occur for all seasons with NMBs of −37.1 to −17.6% and NMEs of 48.1–88.2%, and with much larger underpredictions in the western U.S. than in the eastern U.S. These findings are consistent with those of *Roy et al.* [2007a] who derived AODs based on CMAQ PM predictions for summer 2001 using a different method (i.e., a semiempirical mass extinction method of *Malm et al.* [1994]) for comparison with AODs from MODIS and AERONET and reported a systematic underprediction of derived AODs. Several possible

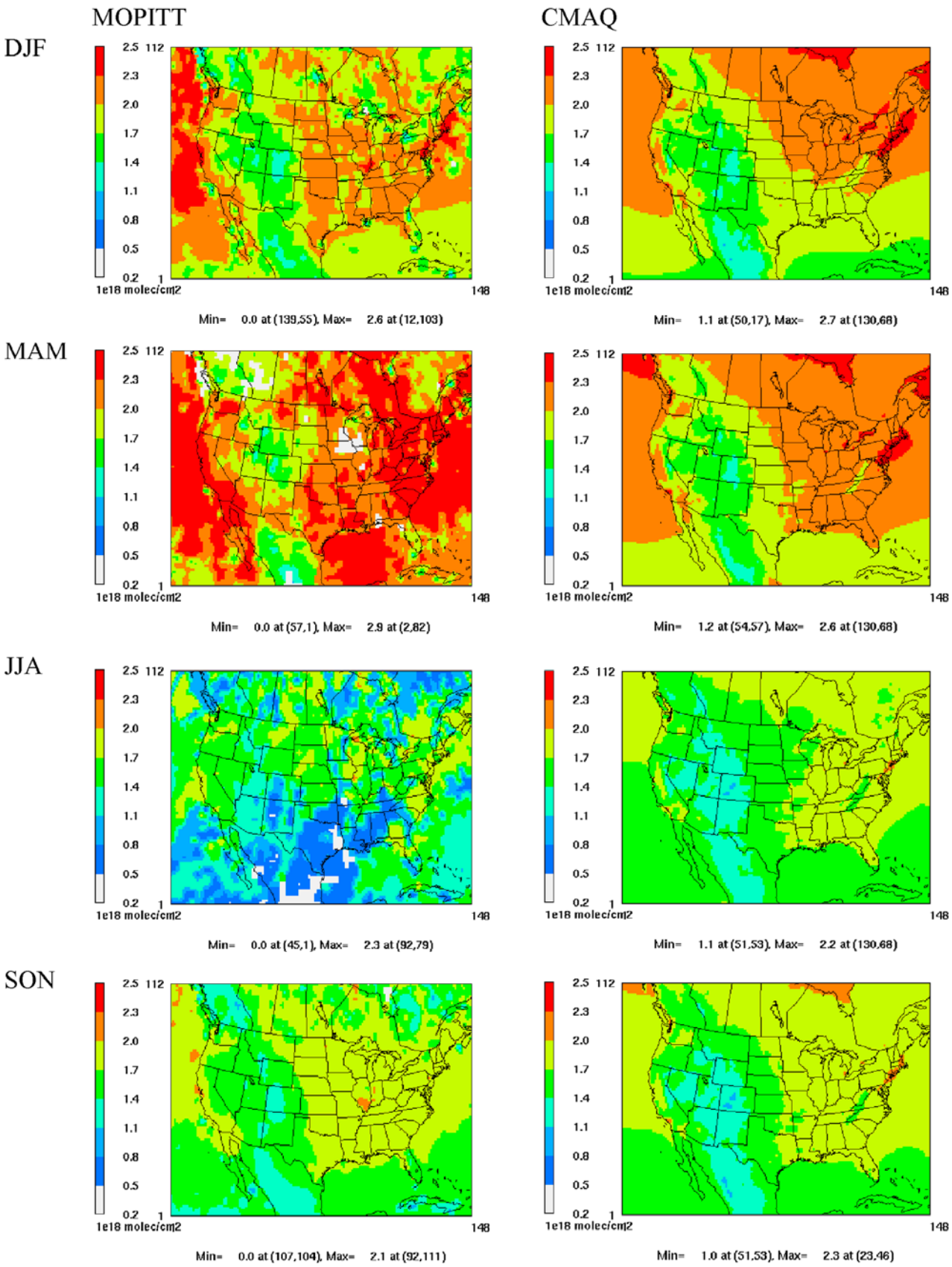

Figure 10. The (left) observed versus (right) simulated seasonal mean tropospheric CO column abundance ($\times 10^{18}$ molecules cm^{-2}) in 2001. The observed CO columns are obtained from MOPITT. MOPITT CO column values during JJA 2001 are plotted based on data in August (no data are available during June–July).

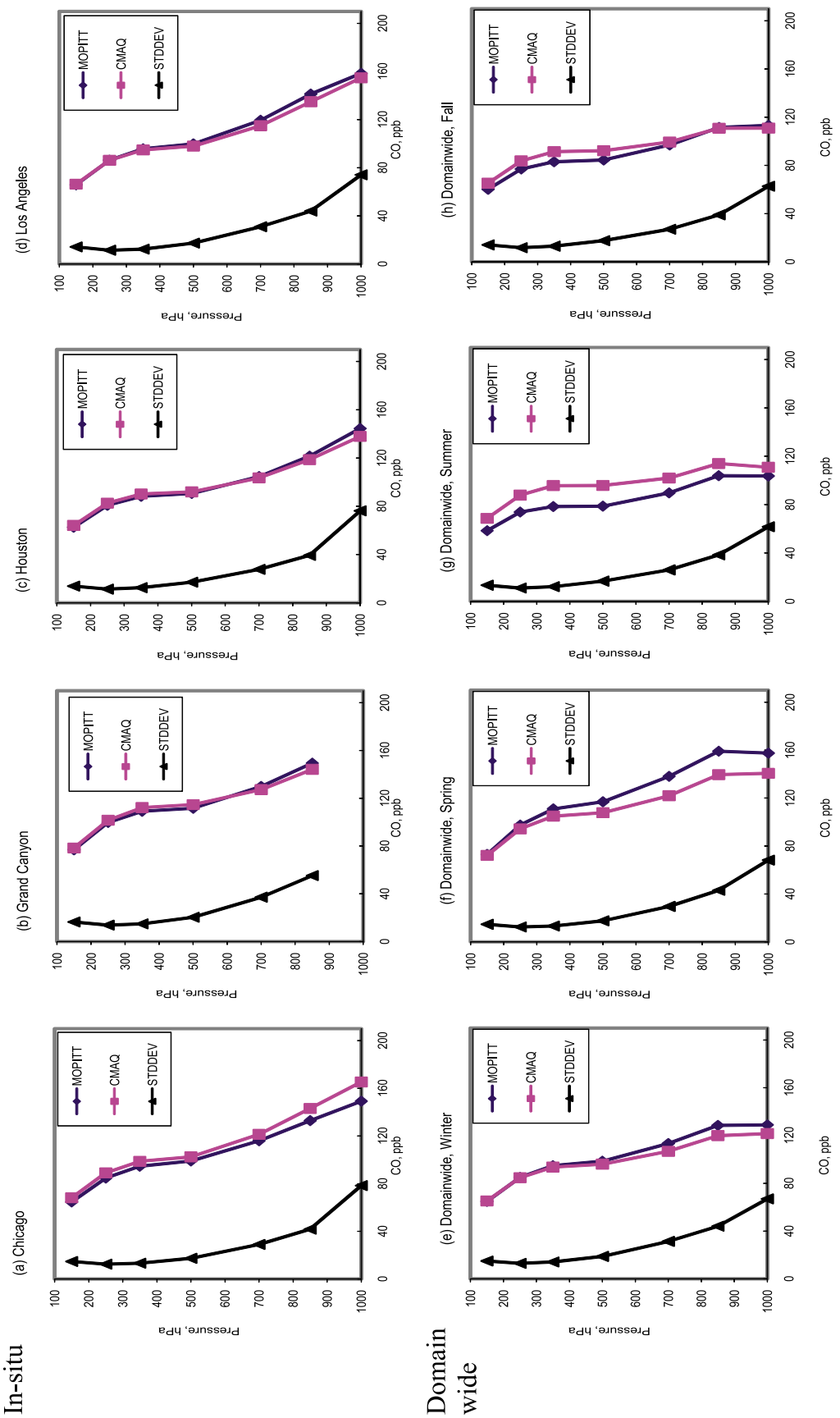

Figure 11. Annual average CO vertical profiles from MOPITT and CMAQ at individual sites (a) Chicago, (b) Grand Canyon, (c) Houston, and (d) Los Angeles, and domain-wide seasonal mean CO vertical profiles from (e–h) MOPITT and CMAQ in 2001.

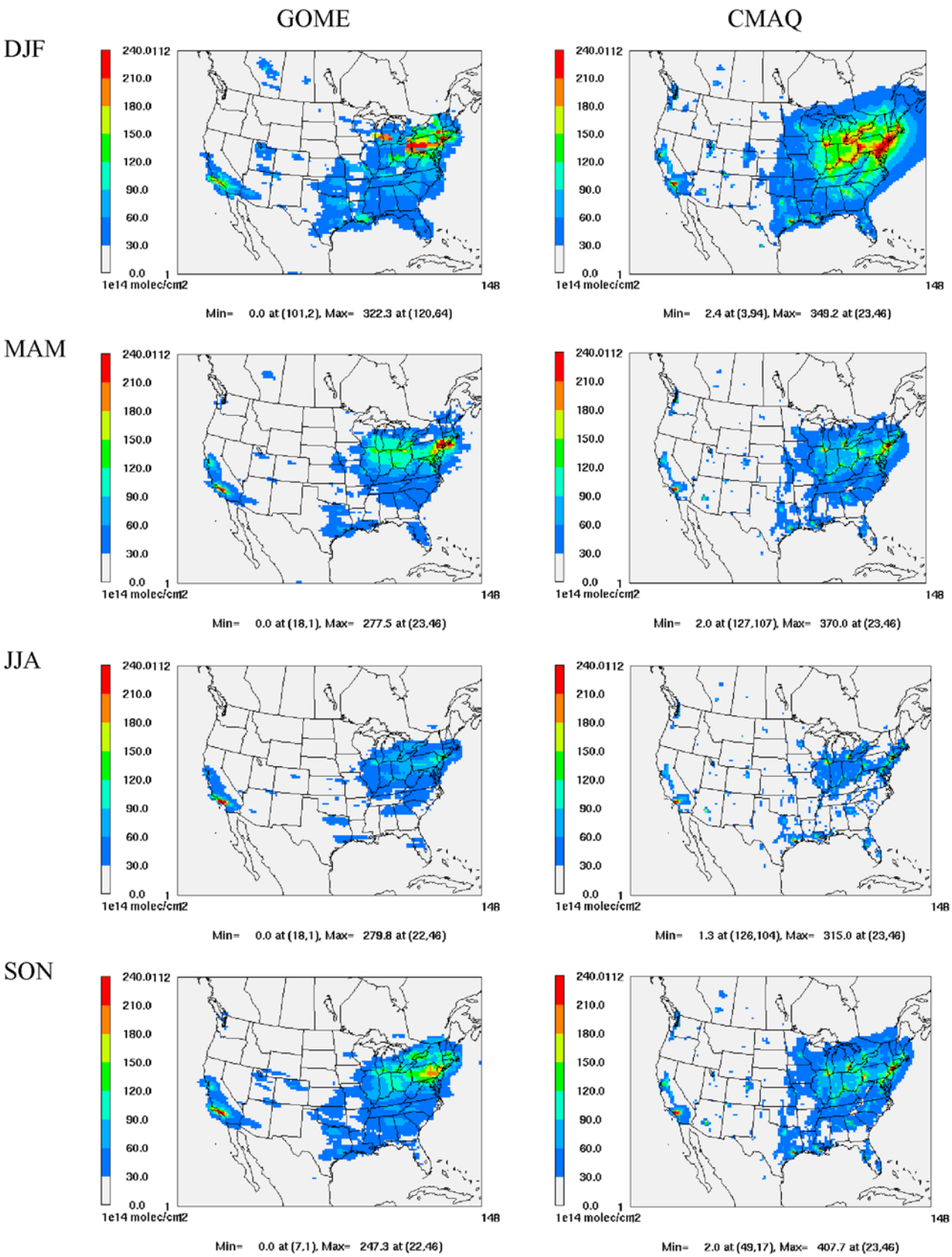

Figure 12. The (left) observed versus (right) simulated seasonal mean tropospheric NO_2 column abundance ($\times 10^{14}$ molecules cm^{-2}) in 2001. The observed NO_2 columns are obtained from GOME.

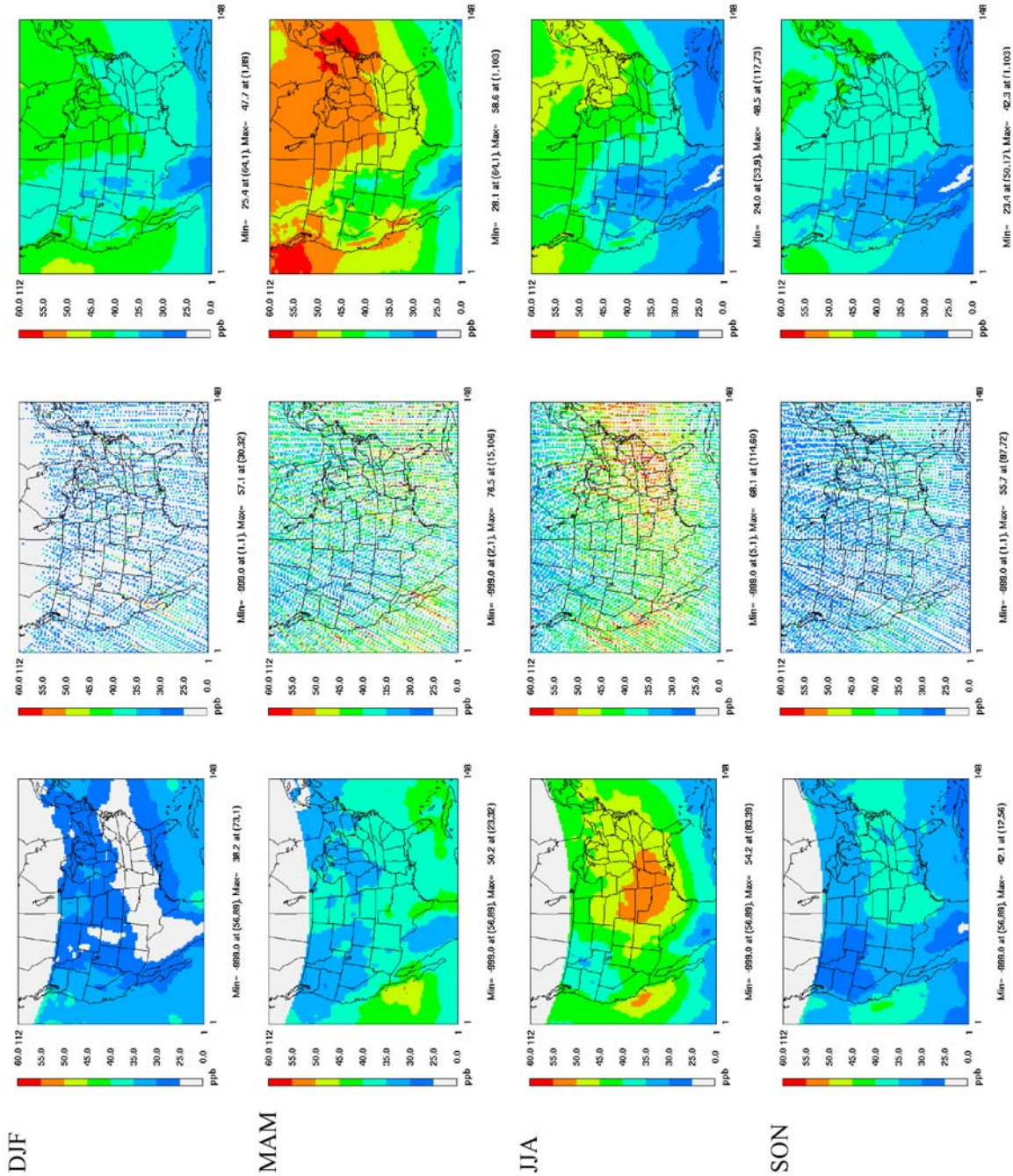

Figure 13. The observed (TOMS/SBUV and GOME) versus simulated (CMAQ) seasonal mean tropospheric O_3 residual (TOR) in DU in 2001. The observed TOR and simulated total column ozone (TCO) are obtained from TOMS/SBUV and GOME, respectively.

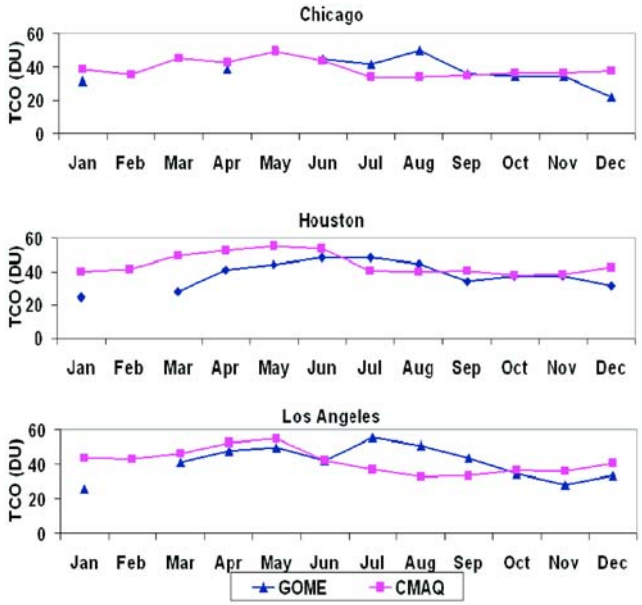

Figure 14. Trends in the monthly mean of total column ozone (TCO) of GOME and CMAQ TCO in 2001 at (top) Chicago, (middle) Houston, and (bottom) Los Angeles.

factors contribute to underpredictions of $PM_{2.5}$ and associated AODs. First, the underestimate of biomass fire emissions in the summer in the western U.S. and the emissions of primary OC and BC over the entire domain can lead to underpredictions of $PM_{2.5}$. For example, in the 2001 NEI, biomass burning related emission estimates are spatially resolved from state to county to grid in CONUS and temporally resolved from annually to hourly based on allocation factors, all of these factors lead to inaccuracies in fire emission estimation [*Roy et al.*, 2007b; *Pouliot et al.*, 2008]. *Roy et al.* [2007b] used the MODIS-derived fire pixel count data to reallocate wildfire emissions over Florida in May 2001 and found reallocated emissions improve CMAQ model predictions of BC and OC. *Mathur* [2008] also demonstrated an improved CMAQ performance in terms of CO and $PM_{2.5}$ predictions by accounting for emissions from wildfires in Alaska using data assimilation of MODIS AODs in the model during 19–23 July 2004 CMAQ simulation that uses 2004 projected emissions from 2001 NEI. *Pouliot et al.* [2008] improved the method of *Roy et al.* [2007b] to derive a 2005 fire inventory for CONUS and reported a 76% higher primary $PM_{2.5}$ emissions from wildfires in 2005 as compared with those in 2002 NEI, due to interannual variability as well as differences in acres burned and the methodology, i.e., the National Fire Danger Rating System (NFDRS) fuel models and fuel loading used for wildfire emission estimation. Differences also exist in prescribed fires and agricultural fires, despite to a lesser extent (e.g., emissions in 2005 are higher than those in 2002 by 22% and 3%, respectively). Second, the inaccurate predictions in the $PM_{2.5}$ concentrations due to inaccurate aerosol chemistry and dynamic treatments (e.g., inaccurate treatments in secondary organic aerosols (SOA) lead to underpredictions in SOA (thus OC and $PM_{2.5}$)). Third, the lack of model treatments of sea salt and dust in CMAQ version 4.4 that can contribute to the observed AODs of $PM_{2.5}$ [*Kaufman et al.*, 2005]. Fourth, uncertainties exist in boundary conditions of $PM_{2.5}$ and its major components used in CMAQ. Using multiple-year measurements from the Aerosol Robotic network (AERONET), *Kaufman et al.* [2001] derived baseline maritime AODs (excluding impact of anthropogenic emissions) to be 0.052 and 0.071 at 500 nm over the Pacific and Atlantic Oceans, respectively. CMAQ-derived AODs are zero or near zero in some areas over the Pacific and Atlantic Oceans. AODs lower than or close to the background values even occur in some areas over CONUS, Cuba, Southern Canada, Northern Mexico where $PM_{2.5}$ and AODs are largely affected by anthropogenic sources, reflecting too low values for the boundary conditions for $PM_{2.5}$ in these regions. In addition, the absolute NMBs for AOD predictions are higher than those for surface $PM_{2.5}$ predictions (see Table 2), indicating possible low boundary $PM_{2.5}$ values and underpredictions of $PM_{2.5}$ above the surface. As shown by *Zhang et al.* [2007], $PM_{2.5}$ concentrations decrease exponentially with the height, but remain > 50% of surface concentrations up to an altitude of 3.2 km in summer 2001 and 2002. Underpredictions of $PM_{2.5}$ between surface and 3.2 km can therefore cause underestimations in AODs.

[30] While CMAQ simulations are affected by uncertainties in the model inputs and formulation, satellite retrievals also have limitations and uncertainties. First, although various error correction techniques (e.g., cloud mask) and retrieval algorithms are used to modify satellite products, uncertainties exist in individual satellite measurements. Compared with AOD (denoted as τ) measurements from AERONET Sun photometers, the uncertainty of MODIS-derived monthly AODs is reported to be $\Delta\tau = \pm 0.05 \pm 0.15\tau$ over land and $\Delta\tau = \pm 0.03 \pm 0.05\tau$ over the ocean because of retrieval errors such as subpixel clouds, snow/ice, and water contamination [*Chu et al.*, 2003, 2005; *Remer et al.*, 2005]. MODIS aerosol measurements over the ocean have been retrieved with special cloud and sediment masks; some residual contamination from very thin cirrus contribute to errors in coarse-mode AODs derived from MODIS but should not affect retrieval of fine-mode AODs [*Remer et al.*, 2005; *Kaufman et al.*, 2005]. For the MOPITT data, the total column CO uncertainty [*Edwards et al.*, 2004] is $-0.5 \pm 12.1\%$, due to retrieval errors such as random errors corresponding to random radiance errors, biases arising from inadequate a priori, and biases associated with the calibrated radiances or the forward model [*Deeter et al.*, 2003]. The uncertainty in the GOME NO_2 tropospheric column data is caused by errors due to the short tropospheric path and uncertainties in clouds, both of which can lead to an underestimation of the amount of NO_2. In general, *Richter and Burrows* [2002] predict that GOME-derived tropospheric NO_2 is low by up to a factor of 2. Second, the effect of clouds (obscuration of the clear-sky path from space to the ground) and the effect of ground albedo (which can result in a lack of contrast between the atmosphere and the surface) make it difficult to obtain the atmospheric quantity from the measurements. For example, the AODs over land may be overestimated by MODIS [*Kinne et al.*, 2003; *Chin et al.*, 2004; *Matsui et al.*, 2004]. The nonsphericity of mineral dusts may likely cause an overestimation in AODs by MODIS over the Pacific [*Remer et al.*, 2005; *Chu et al.*, 2005] and uncertainties in both optical thickness and size parameter

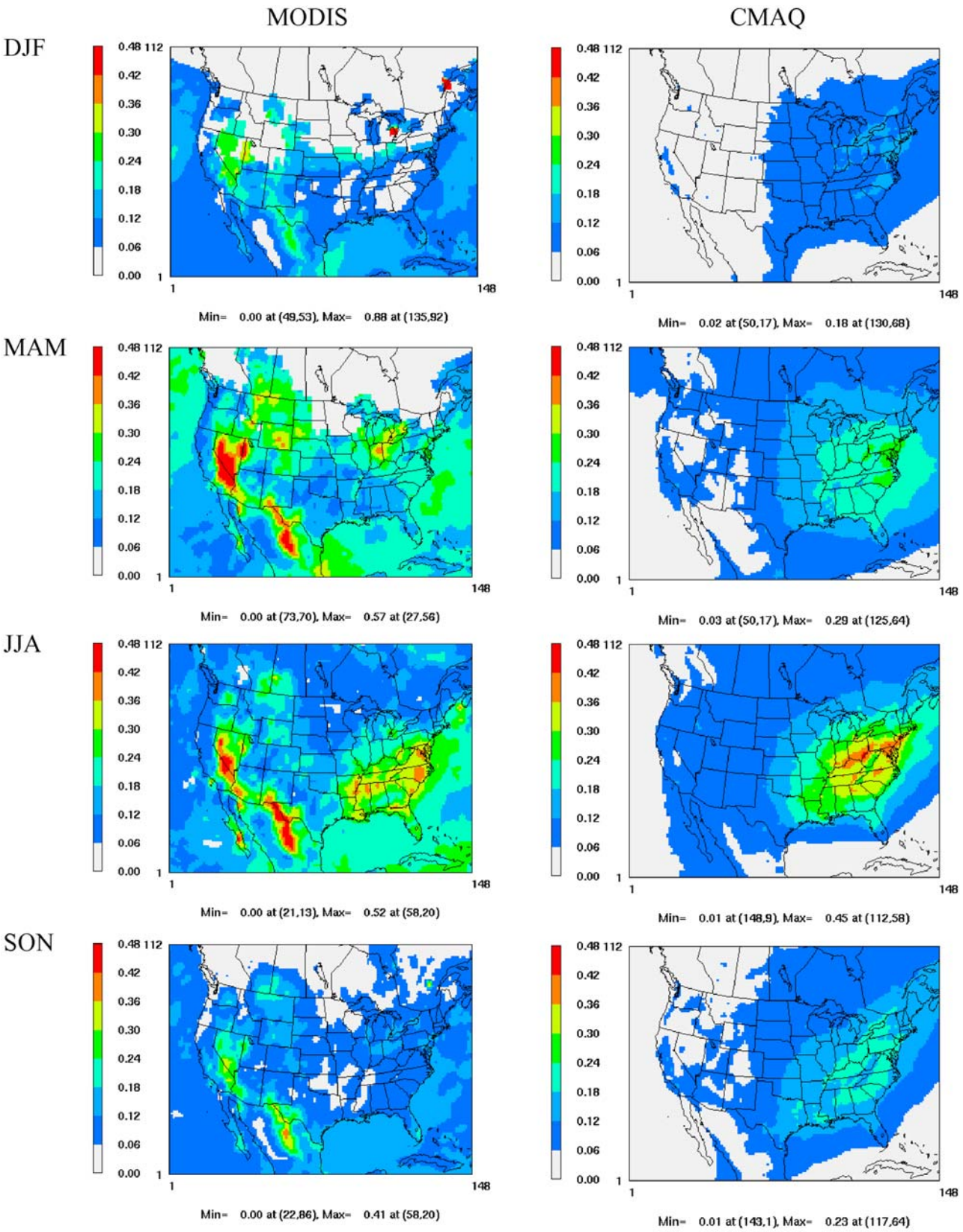

Figure 15. The (left) observed and (right) simulated seasonal mean total column aerosol optical depths (AODs) in 2001. The observed AODs are obtained from MODIS.

retrievals [*Remer et al.*, 2005]. Third, uncertainties exist in the PM size distribution and composition assumed in MODIS retrieval algorithm, which are different from CMAQ results. For example, while CMAQ use three lognormally distributed modes to represent PM size distribution, the aerosol models in the land algorithm (e.g., continental, urban/industrial, developing world) used in MODIS AODs assume different modes with different mean radii and standard deviations [*Remer et al.*, 2005]. Fourth, CMAQ gives higher AODs over regions with frequent cloud coverage and high SO_2 emissions because of the dominancy of sulfate contributions in AODs in these regions, whereas MODIS missed such high AODs because of cloud masks (e.g., during some periods on July 18 and August 12 2001 [see *Roy et al.*, 2007a, Figure 8]. This may help explain the high values of CMAQ-derived AODs in summer over a corridor crossing southern Indiana, Ohio, and Pennsylvania and in fall over some regions in the eastern U.S. with high SO_4^{2-} levels that were not captured by MODIS, despite a domain-wide underprediction by CMAQ during those seasons.

6. Conclusions

[31] A comprehensive evaluation of the CMAQ simulation for the full year of 2001 over the continental U.S. has been performed using both satellite and ground-based measurements to assess current CMAQ's capability in reproducing concentrations and long-term variation trends of major criteria pollutants such as O_3 and $PM_{2.5}$ and related variables such as indicator species and wet deposition fluxes at surface and column mass concentrations of CO, NO_2, TOR, and AODs. An evaluation is conducted for near-surface variables over the whole domain and at 22 individual representative sites. Results indicate an overall satisfactory performance for annual mean maximum 1 h and 8 h average O_3 mixing ratios, with NMBs of −11.6% to 0.1% and −4.6% to 3%, respectively. Model performance for annual 24 h average concentrations of $PM_{2.5}$ and its components is satisfactory or marginally satisfactory, with annual mean NMBs of 4.2−35.3% for $PM_{2.5}$, −22.3 to 26.0% for NH_4^+, −13.0 to 43.5% for SO_4^{2-}, −33.9% to 45.6% for NO_3^-, −52.6% to −5.4% for BC, and −37.6% to 24.8% for OC at various network sites. The larger NMBs for simulated $PM_{2.5}$ and its components occur at urban or suburban sites under the STN and SEARCH networks where MM5/CMAQ simulation at a horizontal grid resolution of 36 km has difficulties in capturing local-scale meteorology, emissions, and concentrations. The NMBs of annual wet deposition fluxes are 19.6% for SO_4^{2-}, −13.0% for NO_3^-, and 31.6% for NH_4^+, consistent with values reported in the literature. The evaluation at 22 individual sites shows that CMAQ reproduces well the daily variations and magnitudes of maximum 8 h O_3 at most locations except for areas over complex terrain such as LAX, RIV, BBE, and GRC and areas with complex interplay of meteorology and emissions such as LAX and NYC. CMAQ can also reproduce higher $PM_{2.5}$ over the SEARCH sites such as JST and YRK and lower values at the IMPROVE sites such as GRS and BBE that are within a factor of 2 of the observed values. Simulated indicator values are overall consistent with observations during most time periods in January and August at most sites except for BHM and GPF where CMAQ underpredicts NO_y but overpredicts O_3/NO_x and O_3/NO_y and JST and YRK where CMAQ overpredicts the degree of sulfate neutralization during most time periods. Such an agreement provides some confidence in accuracy of the predicted change in O_3 and $PM_{2.5}$. While the levels of NO_y suggest VOC-limited O_3 chemistry at JST and BHM in both January and August, other indicators indicate NO_x-limited O_3 chemistry in both months at all sites. Given uncertainties in threshold values for each indicator and inconsistencies among indicators used, caution should be taken when using indicators to determine NO_x or VOC-limited O_3 chemistry and a set of indicators (instead of a single one) are highly recommended for a complete assessment. The reasons for discrepancies between observed and simulated surface O_3, $PM_{2.5}$, indicator species, and wet deposition include uncertainties in the emissions of gases precursors and primary PM, in particular, the inaccurate strengths and temporal profiles of wildfire emissions domain-wide and the overestimation of precursor emissions at urban/suburban sites in the southeastern U.S., the uncertainties in model treatments such as heterogeneous N_2O_5 hydrolysis and partitioning of total nitrate in gas and particulate phases via equilibrium and condensation, and biases in meteorological variables such as temperature and precipitation, due to incapabilities of MM5 in capturing major meteorological and topography characteristics such as complex terrain, land-sea air exchange, impact of sea breezes, extreme wet or dry weathers, and strong diurnal variations in temperatures and mixing heights over some areas.

[32] CMAQ reproduces well the seasonal variations and magnitudes of satellite-derived column CO and NO_2, with NMBs of annual mean values of −0.6% and 6.9%, respectively. For CO vertical profiles, underpredictions occur domain-wide below 250 mb in spring, and overpredictions occur domain-wide throughout the vertical domain in summer. Possible sources of errors may include uncertainties in the CO emissions and boundary conditions used by CMAQ. The total tropospheric O_3 column abundances predicted by CMAQ give a relatively good agreement with the TOR from TOMS/SBUV and TCO from GOME in summer and fall with NMBs of −13.0% to 6.3% against TOMS/SBUV and −7.0% to 11.0% against GOME, but relatively poor performance for winter and spring, with NMBs of 28.7−34.9% against TOMS/SBUV and 20−25% against GOME. As such, CMAQ fails to reproduce observed seasonal variation for TOR/TCO. Given an overall good agreement between simulated and observed surface O_3 mixing ratios with small-to-moderate underpredictions and the fact that upper layer O_3 contributes the most to the O_3 columns, such a poor performance is attributed to the use of inaccurate upper boundary layer conditions for O_3. Among all variables evaluated using satellite data, CMAQ performs the worst for AODs, with the annual mean NMBs of −31.9%. It also fails to reproduce the spatial distributions in all seasons and the seasonal variability over the western U.S. Possible reasons for underpredictions of PM_{10} thus total AODs include the underestimations of biomass fire emissions in summer in the western U.S. and emissions of primary OC and BC over the entire domain, lack of model treatments of sea salt and dust that can contribute to the observed total AODs, uncer-

tainties in boundary conditions of $PM_{2.5}$ and its major components used in CMAQ simulations, as well as limitations and uncertainties in satellite data such as the effects of clouds, the effect of ground albedo, and uncertainties and assumptions associated with retrieval algorithms.

[33] **Acknowledgments.** This work was performed under the National Aeronautics and Space Administration award NNG04GJ90G. The authors thank Ken Schere, George Pouliot, and Warren Peters, U.S. EPA, for providing CMAQ model inputs, Alice Gilliland and Steve Howard, U.S. NOAA/EPA, for providing the scripts for extracting data from observational databases and CMAQ, Shaocai Yu, U.S. EPA/NOAA, for providing FORTRAN script for statistical calculations, Jack Fishman and John K. Creilson, NASA Langley Research Center, for providing Tropospheric Ozone Residual data, Xiong Liu and Kelly Chance, Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, for providing GOME Total Column Ozone data, Shiang-Yuh Wu, Department of Environmental Quality, the state of Virginia, for assisting in TOR data processing, Kai Wang, a graduate student at NCSU, for postprocessing some satellite data, and J.-P. Huang, a former postdoctoral researcher at NCSU, for postprocessing some CMAQ results during the early stage of this work.

References

Ansari, A. S., and S. N. Pandis (1998), Response of inorganic PM to precursor concentrations, *Environ. Sci. Technol.*, *32*, 2706–2714, doi:10.1021/es971130j.

Bey, I., D. J. Jacob, R. M. Yantosca, J. A. Logan, B. D. Field, A. M. Fiore, Q. Li, H. Y. Liu, L. J. Mickley, and M. G. Schultz (2001), Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, *J. Geophys. Res.*, *106*(D19), 23,073–23,096, doi:10.1029/2001JD000807.

Bhave, P., G. Sarwar, W. Appel, and R. Dennis (2006), Revised treatment of N_2O_5 hydrolysis in CMAQ, paper presented at 5th Annual Conference, Commun. Model. and Anal. Syst. Cent., Chapel Hill, N. C.

Binkowski, F. S., and S. J. Roselle (2003), Models-3 community multiscale air quality (CMAQ) model aerosol component: 1. Model description, *J. Geophys. Res.*, *108*(D6), 4183, doi:10.1029/2001JD001409.

Blanchard, C. L. (2000), Ozone process insights from field experiments: Part III. Extent of reaction and ozone formation, *Atmos. Environ.*, *34*, 2035–2043, doi:10.1016/S1352-2310(99)00458-6.

Boylan, J. W., and A. G. Russell (2006), PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models, *Atmos. Environ.*, *40*, 4946–4959, doi:10.1016/j.atmosenv.2005.09.087.

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, doi:10.1115/1.2128636.

Cakmur, R. V., R. L. Miller, and I. Tegen (2001), A comparison of seasonal and interannual variability of soils dust aerosols over the Atlantic Ocean as inferred by the TOMS AI and AVHRR AOT retrievals, *J. Geophys. Res.*, *106*(D16), 18,287–18,303, doi:10.1029/2000JD900525.

Carmichael, G. R., L. K. Peters, and R. D. Saylor (1991), The STEM-II Regional-Scale Acid Deposition and Photochemical Oxidant Model: I. An overview of model development and applications, *Atmos. Environ., Part A*, *25*, 2077–2090.

Chameides, W. L., C. Luo, R. Saylor, D. Streets, Y. Huang, M. Bergin, and F. Giorgi (2002), Correlation between model-calculated anthropogenic aerosols and satellite-derived cloud optical depths: Indication of indirect effect?, *J. Geophys. Res.*, *107*(D10), 4085, doi:10.1029/2000JD000208.

Chin, M., A. Chu, R. Levy, L. Remer, Y. Kaufman, B. Holben, T. Eck, P. Ginoux, and O. Gao (2004), Aerosol distribution in the Northern Hemisphere during ACE-Asia: Results from global model, satellite observations, and Sun photometer measurements, *J. Geophys. Res.*, *109*, D23S90, doi:10.1029/2004JD004829.

Chu, D. A., Y. J. Kaufman, G. Zibordi, J. D. Chern, J. Mao, C. Li, and B. N. Holben (2003), Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS), *J. Geophys. Res.*, *108*(D21), 4661, doi:10.1029/2002JD003179.

Chu, D. A., et al. (2005), Evaluation of aerosol properties over ocean from Moderate Resolution Imaging Spectroradiometer (MODIS) during ACE-Asia, *J. Geophys. Res.*, *110*, D07308, doi:10.1029/2004JD005208.

Dean, K. G., J. Dehn, K. R. Papp, S. Smith, P. Izbekov, R. Peterson, C. Kearney, and A. Steffke (2004), Integrated satellite observations of the 2001 eruption of Mt. Cleveland, Alaska, *J. Volcanol. Geotherm. Res.*, *135*, 51–73, doi:10.1016/j.jvolgeores.2003.12.013.

Deeter, M. N., et al. (2003), Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument, *J. Geophys. Res.*, *108*(D14), 4399, doi:10.1029/2002JD003186.

Easter, R. C., S. J. Ghan, Y. Zhang, R. D. Saylor, E. G. Chapman, N. S. Laulainen, H. Abdul-Razzak, L. R. Leung, X. Bian, and R. A. Zaveri (2004), MIRAGE: Model description and evaluation of aerosols and trace gases, *J. Geophys. Res.*, *109*, D20210, doi:10.1029/2004JD004571.

Eder, B., and S.-C. Yu (2006), An evaluation of model performance of EPA models-3/CMAQ, *Atmos. Environ.*, *40*, 4811–4824, doi:10.1016/j.atmosenv.2005.08.045.

Edwards, D. P., et al. (2004), Observations of carbon monoxide and aerosols from the Terra satellite: Northern Hemisphere variability, *J. Geophys. Res.*, *109*, D24202, doi:10.1029/2004JD004727.

Fairlie, T. D., D. J. Jacob, and R. J. Park (2007), The impact of transpacific transport of mineral dust in the United States, *Atmos. Environ.*, *41*, 1251–1266, doi:10.1016/j.atmosenv.2006.09.048.

Fishman, J., A. E. Wozniak, and J. K. Creilson (2003), Global distribution of tropospheric ozone from satellite measurements using the empirically corrected tropospheric O_3 residual techniques: identification of the regional aspects of air pollution, *Atmos. Chem. Phys.*, *3*, 893–907.

Fishman, J., J. K. Creilson, A. E. Wozniak, and P. J. Crutzen (2005), The interannual variability of stratospheric and tropospheric ozone determined from satellite measurements, *J. Geophys. Res.*, *110*, D20306, doi:10.1029/2005JD005868.

Frost, G. J., et al. (2006), Effects of changing power plant NO_x emissions on ozone in the eastern United States: Proof of concept, *J. Geophys. Res.*, *111*, D12306, doi:10.1029/2005JD006354.

Ghan, S. J., N. S. Laulainen, R. C. Easter, R. Wagener, S. Nemesure, E. G. Chapman, Y. Zhang, and L. R. Leung (2001), Evaluation of Aerosol Direct Radiative Forcing in MIRAGE, *J. Geophys. Res.*, *106*(D6), 5295–5316, doi:10.1029/2000JD900502.

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, doi:10.1016/j.atmosenv.2006.01.023.

Gilliland, A. B., R. L. Dennis, S. J. Roselle, and T. E. Pierce (2003), Seasonal NH3 emission estimates for the eastern United States based on ammonium wet concentrations and an inverse modeling method, *J. Geophys. Res.*, *108*(D15), 4477, doi:10.1029/2002JD003063.

Gilliland, A. B., K. W. Appel, R. W. Pinder, and R. L. Dennis (2006), Seasonal NH_3 emissions for the continental united states: Inverse model estimation and evaluation, *Atmos. Environ.*, *40*, 4986–4998, doi:10.1016/j.atmosenv.2005.12.066.

Gong, S. L., et al. (2003), Canadian Aerosol Module: A size-segregated simulation of atmospheric aerosol processes for climate and air quality models: 1. Module development, *J. Geophys. Res.*, *108*(D1), 4007, doi:10.1029/2001JD002002.

Grell, G. A., 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, doi:10.1016/j.atmosenv.2005.04.027.

Griffin, R. J., K. Nguyen, D. Dabdub, and J. H. Seinfeld (2003), A coupled hydrophobichydrophilic model for predicting secondary organic aerosol formation, *J. Atmos. Chem.*, *44*, 171–190, doi:10.1023/A:1022436813699.

Hammer, M.-U., B. Vogel, and H. Vogel (2002), Findings on H_2O_2/HNO_3 as an indicator of ozone sensitivity in Baden-Württemberg, Berlin-Brandenburg, and the Po valley based on numerical simulations, *J. Geophys. Res.*, *107*(D22), 8190, doi:10.1029/2000JD000211.

Heald, C. L., D. J. Jacob, R. J. Park, B. Alexander, T. D. Fairlie, and R. M. Yantosca (2006), Transpacific transport of Asian anthropogenic aerosols and its impact on surface air quality in the United States, *J. Geophys. Res.*, *111*, D14310, doi:10.1029/2005JD006847.

Heland, J., et al. (2002), First comparison of tropospheric NO_2 column densities retrieved from GOME measurements and in situ aircraft profile measurements, *Geophys. Res. Lett.*, *29*(20), 1983, doi:10.1029/2002GL015528.

Hogrefe, C., B. Lynn, K. Civerolo, J.-Y. Ku, J. Rosenthal, C. Rosenzweig, R. Goldberg, S. Gaffin, K. Knowlton, and P. L. Kinney (2004), Simulating changes in regional air pollution over the eastern United States due to changes in global and regional climate and emissions, *J. Geophys. Res.*, *109*, D22301, doi:10.1029/2004JD004690.

Horowitz, L. W. (2006), Past, present, and future concentrations of tropospheric ozone and aerosols: methodology, ozone evaluation, and sensitivity to aerosol wet removal, *J. Geophys. Res.*, *111*, D22211, doi:10.1029/2005JD006937.

Hudman, R. C., et al. (2007), Surface and lightning sources of nitrogen oxides over the United States: Magnitudes, chemical evolution, and outflow, *J. Geophys. Res.*, *112*, D12S05, doi:10.1029/2006JD007912.

Intergovernmental Panel on Climate Change (IPCC) (2007), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., 996 pp., Cambridge Univ. Press, Cambridge, U. K.

Jacob, D. J., J. A. Logan, and P. P. Murti (1999), Effect of rising Asian emissions on surface ozone in the United States, *Geophys. Res. Lett.*, *26*, 2175–2178, doi:10.1029/1999GL900450.

Jacobson, M. Z. (1997), Development and application of a new air pollution modeling system, part II, Aerosol module structure and design, *Atmos. Environ.*, *31*, 131–144, doi:10.1016/1352-2310(96)00202-6.

Jacobson, M. Z. (2001), Strong radiative heating due to the mixing state of black carbon in atmospheric aerosols, *Nature*, *409*, 695–697, doi:10.1038/35055518.

Jacobson, M. Z. (2002), Control of fossil-fuel particulate black carbon plus organic matter, possibly the most effective method of slowing global warming, *J. Geophys. Res.*, *107*(D19), 4410, doi:10.1029/2001JD001376.

Jacobson, M. Z. (2005), *Fundamentals of Atmospheric Modeling*, 2nd ed., 813 pp., Cambridge Univ. Press, New York.

Jaffe, D. A., J. Snow, and O. Cooper (2003), The 2001 Asian dust events: Transport and impact on surface aerosol concentrations in the U.S., *Eos Trans. AGU*, *84*(46), 501, doi:10.1029/2003EO460001.

Kaufman, Y. J., A. Smirnov, B. N. Holben, and O. Dubovik (2001), Baseline maritime aerosol: Methodology to derive the optical thickness and scattering properties, *Geophys. Res. Lett.*, *28*, 3251–3254, doi:10.1029/2001GL013312.

Kaufman, Y. J., I. Koren, L. A. Remer, D. Tanre, P. Ginoux, and S. Fan (2005), Dust transport and deposition observed from the Terra-Moderate Resolution Imaging Spectroradiometer (MODIS) spacecraft over the Atlantic Ocean, *J. Geophys. Res.*, *110*, D10S12, doi:10.1029/2003JD004436.

Kinne, S., et al. (2003), Monthly averages of aerosol properties, A global comparison among models, satellite data, and AERONET ground data, *J. Geophys. Res.*, *108*(D20), 4634, doi:10.1029/2001JD001253.

Kotchenruther, R. A., and P. V. Hobbs (1998), Humidification factors for aerosols from biomass burning in Brazil, *J. Geophys. Res.*, *103*, 32,081–32,089, doi:10.1029/98JD00340.

Kotchenruther, R. A., P. V. Hobbs, and D. A. Hegg (1999), Humidification factors for atmospheric aerosols off the mid-Atlantic coast of the United States, *J. Geophys. Res.*, *104*, 2239–2251, doi:10.1029/98JD01751.

Liu, X., K. Chance, C. E. Sioris, R. J. D. Spurr, T. P. Kurosu, R. V. Martin, and M. J. Newchurch (2005), Ozone profile and tropospheric ozone retrievals from the Global Ozone Monitoring Experiment: Algorithm description and validation, *J. Geophys. Res.*, *110*, D20307, doi:10.1029/2005JD006240.

Liu, X., et al. (2006a), First directly retrieved global distribution of tropospheric column ozone from GOME: Comparison with the GEOS-CHEM model, *J. Geophys. Res.*, *111*, D02308, doi:10.1029/2005JD006564.

Liu, X., K. Chance, and T. P. Kurosu (2006b), Improved ozone profile retrievals from GOME data with degradation correction in reflectance, *Atmos. Phys. Chem. Discuss.*, *6*, 8285–8300.

Lohmann, U., P. Stier, C. Hoose, S. Ferrachat, S. Kloster, E. Roeckner, and J. Zhang (2007), Cloud microphysics and aerosol indirect effects in the global climate model ECHAM5-HAM, *Atmos. Chem. Phys.*, *7*, 3425–3446.

Lu, C.-H., and J. S. Chang (1998), On the indicator-based approach to assess ozone sensitivities and emissions features, *J. Geophys. Res.*, *103*, 3453–3462, doi:10.1029/97JD03128.

Malm, W. C., J. F. Sisler, D. Huffman, R. A. Eldred, and T. A. Cahill (1994), Spatial and seasonal trends in particle concentration and optical extinction in the United States, *J. Geophys. Res.*, *99*, 1347–1370.

Mathur, R. (2008), Estimating the impact of the 2004 Alaskan forest fires on episodic particulate matter pollution over the eastern United States through assimilation of satellite-derived aerosol optical depths in a regional air quality model, *J. Geophys. Res.*, *113*, D17302, doi:10.1029/2007JD009767.

Matsui, T., S. M. Kreidenweis, R. A. Pielke, B. Schichtel, H. B. Yu, M. Chin, D. A. Chu, and D. Niyogi (2004), Regional comparison and assimilation of GOCART and MODIS aerosol optical depth across the eastern US, *Geophys. Res. Lett.*, *31*, L21101, doi:10.1029/2004GL021017.

Morris, R., B. Koo, D. McNally, T. W. Tesche, and G. Tonnesen (2004), Application of multiple models to simulation fine particulate in the southeastern US, paper presented at National RPO Modeling Meeting, U.S. Environ. Prot. Agency, Denver, Colo.

Penner, J. E., et al. (2001), Aerosols, their direct and indirect effects, in *Climate Change 2001, The Scientific Basis*, edited by J. T. Houghton et al., pp. 291–348, Cambridge Univ. Press, Cambridge, U. K.

Pinder, R. W., R. Strader, C. I. Davidson, and P. J. Adams (2004), A temporally and spatially resolved ammonia emission inventory for dairy cows in the United States, *Atmos. Environ.*, *38*, 3747–3756, doi:10.1016/j.atmosenv.2004.04.008.

Pinder, R. W., R. L. Dennis, and P. V. Bhave (2008), Observable indicators of the sensitivity of $PM_{2.5}$ nitrate to emission reductions: Part I: Derivation of the adjusted gas ratio and applicability at regulatory-relevant time scales, *Atmos. Environ.*, *42*(6), 1275–1286, doi:10.1016/j.atmosenv.2007.10.039.

Pleim, J., G. Gipson, S. Roselle, and J. Young (2003), New features of the 2003 release of the CMAQ model, paper presented at the Models3 Users' Workshop, U.S. Environ. Prot. Agency, Research Triangle Park, N. C.

Pósfai, M., J. R. Anderson, P. R. Buseck, and H. Sievering (1999), Soot and sulfate aerosol particles in the remote marine troposphere, *J. Geophys. Res.*, *104*, 21,685–21,693, doi:10.1029/1999JD900208.

Pouliot, G., T. Pace, B. Roy, T. Pierce, and D. Mobley (2008), Development of a biomass burning emissions inventory by combining satellite and ground-based information, *J. Appl. Remote Sens.*, *2*, 021501, doi:10.1117/1.2939551.

Prospero, J. M., R. A. Glaccum, and R. T. Nees (1981), Atmospheric transport of soil dust from Africa to South America, *Nature*, *289*, 570–572, doi:10.1038/289570a0.

Pun, B. K., and C. Seigneur (2001), Sensitivity of particulate matter nitrate formation to precursor emissions in the California San Joaquin Valley, *Environ. Sci. Technol.*, *35*(14), 2979–2987, doi:10.1021/es0018973.

Queen, A., and Y. Zhang (2008), Examining the sensitivity of MM5-CMAQ predictions to explicit microphysics schemes and horizontal grid resolutions, Part II—PM Concentrations and wet deposition predictions, *Atmos. Environ.*, *42*, 3856–3868, doi:10.1016/j.atmosenv.2007.12.066.

Queen, A., Y. Zhang, R. Gilliam, and J. Pleim (2008), Examining the sensitivity of MM5-CMAQ predictions to explicit microphysics schemes and horizontal grid resolutions, Part I—Database, evaluation protocol, and precipitation predictions, *Atmos. Environ.*, *42*, 3842–3855, doi:10.1016/j.atmosenv.2007.12.067.

Remer, L. A., et al. (2005), The MODIS aerosol algorithm, products and validation, *J. Atmos. Sci.*, *62*, 947–973, doi:10.1175/JAS3385.1.

Richter, A., and J. P. Burrows (2002), Retrieval of tropospheric NO_2 from GOME measurements, *Adv. Space Res.*, *29*(11), 1673–1683, doi:10.1016/S0273-1177(02)00100-X.

Roy, B., R. Mathur, A. B. Gilliland, and S. C. Howard (2007a), A comparison of CMAQ-based aerosol properties with IMPROVE, MODIS, and AERONET data, *J. Geophys. Res.*, *112*, D14301, doi:10.1029/2006JD008085.

Roy, B., G. A. Pouliot, A. Gilliland, T. Pierce, S. Howard, P. V. Bhave, and W. Benjey (2007b), Refining fire emissions for air quality modeling with remotely sensed fire counts: A wildfire case study, *Atmos. Environ.*, *41*, 655–665, doi:10.1016/j.atmosenv.2006.08.037.

Russell, A., and R. Dennis (2000), NARSTO critical review of photochemical models and modeling, *Atmos. Environ.*, *34*, 2283–2324, doi:10.1016/S1352-2310(99)00468-9.

Seigneur, C. (2001), Current status of air quality models for particulate matter, *J. Air Waste Manage. Assoc.*, *51*, 1508–1521.

Seigneur, C. (2005), Air pollution: current challenges and future opportunities, *AIChE J.*, *51*, 356–364, doi:10.1002/aic.10458.

Seinfeld, J. H. (2004), Air pollution: a half century of progress, *AIChE J.*, *50*, 1096–1108, doi:10.1002/aic.10102.

Sillman, S. (1995), The use of NO_y, H_2O_2, and HNO_3 as indicators for ozone-NO_x-hydrocarbon sensitivity in urban locations, *J. Geophys. Res.*, *100*, 4175–4188.

Sillman, S., and D. He (2002), Some theoretical results concerning O_3-NO_x-VOC chemistry and NO_x-VOC indicators, *J. Geophys. Res.*, *107*(D22), 4659, doi:10.1029/2001JD001123.

Sillman, S., D. He, C. Cardelino, and R. E. Imhoff (1997), The use of photochemical indicators to evaluate ozone-NO_x-hydrocarbon sensitivity: Case studies from Atlanta, New York, and Los Angeles, *J. Air Waste Manage. Assoc.*, *47*, 642–652.

Stockwell, W. R., T. Gonzalez, R. M. Fitzgerald, D. Lu, Y. Zhang, and J. D. Fuentes (2007), Thoughts on modeling aerosol formation using examples from CMAQ and WRF-Chem, paper presented at International Aerosol Modeling Algorithms (IAMA) Conference, Calif. Air Resour. Board, Davis.

Tonnesen, G. S., and R. L. Dennis (2000a), Analysis of radical propagation efficiency to assess ozone sensitivity to hydrocarbons and NO_x: 1. Local indicators of instantaneous odd oxygen production sensitivity, *J. Geophys. Res.*, *105*, 9213–9225, doi:10.1029/1999JD900371.

Tonnesen, G. S., and R. L. Dennis (2000b), Analysis of radical propagation efficiency to assess ozone sensitivity to hydrocarbons and NO_x: 2. Long-lived species as indicators of ozone concentration sensitivity, *J. Geophys. Res.*, *105*, 9227–9241, doi:10.1029/1999JD900372.

U.S. Climate Change Science Program (2008), *Climate Projections Based on Emissions Scenarios for Long-Lived and Short-Lived Radiatively Active Gases and Aerosols*, edited by H. Levy II et al., 100 pp., Dep. of Commer., Washington, D. C.

U.S. Environmental Protection Agency (U.S. EPA) (2001), Draft guidance for demonstrating attainment of air quality goals for $PM_{2.5}$ and regional haze, report, Research Triangle Park, N. C.

U.S. Environmental Protection Agency (U.S. EPA) (2005), Evaluating ozone control programs in the eastern United States: NO_x budget trading program progress and compliance, *EPA-454-K-05-001*, Washington, D. C.

U.S. Environmental Protection Agency (U.S. EPA) (2007), Guidance on the use of models and other analyses for demonstrating attainment of air quality goals for ozone, $PM_{2.5}$, and regional haze, *Final Rep. EPA-454/B-07-002*, 27,711 pp., Research Triangle Park, N. C.

Vijayaraghavan, K., H. E. Snell, and C. Seigneur (2008), Practical aspects of using satellite data in air quality modeling, *Environ. Sci. Technol.*, *42*(22), 8187–8192, doi:10.1021/es7031339.

Vijayaraghavan, K., Y. Zhang, C. Seigneur, P. Karamchandani, and H. E. Snell (2009), Export of reactive nitrogen from coal-fired power plants in the USA: Estimates from a plume-in-grid modeling study, *J. Geophys. Res.*, *114*, D04308, doi:10.1029/2008JD010432.

Wang, K., Y. Zhang, C. Jang, S. Phillips, and B. Wang (2009), Modeling intercontinental air pollution transport over the trans-Pacific region in 2001 using the Community Multiscale Air Quality modeling system, *J. Geophys. Res.*, *114*, D04307, doi:10.1029/2008JD010807.

Wu, S., L. J. Mickley, E. M. Leibensperger, D. J. Jacob, D. Rind, and D. G. Streets (2008), Effects of 2000–2050 global change on ozone air quality in the United States, *J. Geophys. Res.*, *113*, D06302, doi:10.1029/2007JD008917.

Wu, S.-Y., S. Krishnan, Y. Zhang, and V. Aneja (2008), Modeling atmospheric transport and fate of ammonia in North Carolina, Part I. Evaluation of meteorological and chemical predictions, *Atmos. Environ.*, *42*, 3419–3436, doi:10.1016/j.atmosenv.2007.04.031.

Yu, S.-C., R. Dennis, S. Roselle, A. Nenes, J. Walker, B. Eder, K. Schere, J. Swall, and W. Robarge (2005), An assessment of the ability of 3-D air quality models with current thermodynamic equilibrium models to predict aerosol NO_3^-, *J. Geophys. Res.*, *110*, D07S13, doi:10.1029/2004JD004718.

Zhang, Y. (2008), Online coupled meteorology and chemistry models: History, current status, and outlook, *Atmos. Chem. Phys.*, *8*, 2895–2932.

Zhang, Y., C. Seigneur, J. H. Seinfeld, M. Jacobson, and F. S. Binkowski (1999), Simulation of aerosol dynamics: A comparative review of algorithms used in air quality models, *Aerosol Sci. Technol.*, *31*(6), 487–514, doi:10.1080/027868299304039.

Zhang, Y., K. Vijayaraghavan, C. Seigneur, and G. Tonnesen (2002), Evaluation of probing tools implemented in CAMx, *Final Rep. CP099-1-2*, Coord. Res. Counc., Inc., Alpharetta, Ga.

Zhang, Y., B. Pun, K. Vijayaraghavan, S.-Y. Wu, C. Seigneur, S. Pandis, M. Jacobson, A. Nenes, and J. Seinfeld (2004), Development and application of the model for aerosol dynamics, reaction, ionization and dissolution (MADRID), *J. Geophys. Res.*, *109*, D01202, doi:10.1029/2003JD003501.

Zhang, Y., K. Vijayaraghavan, and C. Seigneur (2005), Evaluation of three probing techniques in a three-dimensional air quality model, *J. Geophys. Res.*, *110*, D02305, doi:10.1029/2004JD005248.

Zhang, Y., P. Liu, B. Pun, and C. Seigneur (2006a), A comprehensive performance evaluation of MM5-CMAQ for the Summer 1999 Southern Oxidants Study episode, Part I. Evaluation protocols, databases and meteorological predictions, *Atmos. Environ.*, *40*, 4825–4838, doi:10.1016/j.atmosenv.2005.12.043.

Zhang, Y., P. Liu, A. Queen, C. Misenis, B. Pun, C. Seigneur, and S.-Y. Wu (2006b), A Comprehensive performance evaluation of MM5-CMAQ for the Summer 1999 Southern Oxidants Study Episode, Part II. Gas and aerosol predictions, *Atmos. Environ.*, *40*, 4839–4855, doi:10.1016/j.atmosenv.2005.12.048.

Zhang, Y., P. Liu, B. Pun, and C. Seigneur (2006c), A comprehensive performance evaluation of MM5-CMAQ for the Summer 1999 Southern Oxidants Study episode, Part III. Diagnostic and mechanistic evaluations, *Atmos. Environ.*, *40*, 4856–4873, doi:10.1016/j.atmosenv.2005.12.046.

Zhang, Y., J.-P. Huang, D. K. Henze, and J. H. Seinfeld (2007), The role of isoprene in secondary organic aerosol formation on a regional scale, *J. Geophys. Res.*, *112*, D20207, doi:10.1029/2007JD008675.

Zhang, Y., X.-M. Hu, L. R. Leung, and W. I. Gustafson Jr. (2008), Impacts of regional climate changes on biogenic emissions and air quality, *J. Geophys. Res.*, *113*, D18310, doi:10.1029/2008JD009965.

Zhang, Y., X.-Y. Wen, K. Wang, K. Vijayaraghavan, and M. Z. Jacobson (2009), Probing into regional O_3 and particulate matter pollution in the United States: 2. An examination of formation mechanisms through a process analysis technique and sensitivity study, *J. Geophys. Res.*, *114*, D22305, doi:10.1029/2009JD011900.

M. Z. Jacobson, Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, USA.

H. E. Snell, Atmospheric and Environmental Research, Inc., Lexington, MA 02421, USA.

K. Vijayaraghavan, Atmospheric and Environmental Research, Inc., San Francisco, CA 94111, USA.

X.-Y. Wen and Y. Zhang, Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, 1125 Jordan Hall, Campus Box 8208, 2800 Faucette Dr., Raleigh, NC 27695, USA.