ELSEVIER

Temporal features in observed and simulated meteorology and air quality over the Eastern United States

C. Hogrefe[a],[*], P.S. Porter[b], E. Gego[c], A. Gilliland[d], R. Gilliam[d], J. Swall[d], J. Irwin[e], S.T. Rao[d]

[a]Atmospheric Sciences Research Center, University at Albany, 251 Fuller Road, Albany, NY 12203, USA
[b]University of Idaho, 1776 Science Center Dr, Idaho Falls, ID 83402, USA
[c]GEGO, 308 Evergreen Drive, Idaho Falls, ID 83401, USA
[d]NOAA Atmospheric Sciences Modeling Division, Mail Drop E243-02, Research Triangle Park, NC 27711, USA[1]
[e]John S. Irwin and Associates, 1900 Pony Run Road, Raleigh, NC 27615-7415, USA

Abstract

Over the next several years, grid-based photochemical models such as the community multiscale air quality (CMAQ) model, the regional modeling system for aerosols and deposition (REMSAD), the comprehensive air quality model with extensions (CAMx), and other regional models will be used by regulatory agencies in the United States for designing emission control strategies to meet and maintain the National Ambient Air Quality Standards (NAAQS) for O_3, $PM_{2.5}$, and regional haze. In this study, temporal scale analysis is applied as a technique to evaluate an annual simulation of meteorology, O_3, and $PM_{2.5}$ and its chemical components over the continental US utilizing two modeling systems. The spectral decomposition of total $PM_{2.5}$ mass from hourly observations and CMAQ and REMSAD model predictions revealed that days of high $PM_{2.5}$ concentrations are generally characterized by positive forcing from fluctuations having periods equal to or greater than a day (i.e., the diurnal, synoptic, and longer-term components) while the magnitude of intra-day fluctuations showed only small differences between average and episodic conditions. Both modeling systems did not capture most of the variability of the high-frequency, intra-day component for all variables for which hourly measurements were available. Furthermore, it is illustrated that correlations were insignificant on the intra-day time scale for all variables, suggesting that these models in the setup used for this study were not skillful in simulating the higher-frequency variations in meteorological variables and the levels of all pollutants. The models exhibited greatest skills at capturing longer-term (seasonal) fluctuations for temperature, wind speed, O_3, sulfate and nitrate. Correlations for total $PM_{2.5}$, ammonium, elemental carbon (EC), organic carbon (OC) and crustal $PM_{2.5}$ correlations were highest for the synoptic time scale implying problems with factors other than meteorology, such as emissions or lateral chemical boundary conditions, in capturing the baseline fluctuations.
© 2006 Elsevier Ltd. All rights reserved.

Keywords: Model evaluation; Air quality modeling; Particulate matter; Spectral analysis; CMAQ

[*]Corresponding author. Bureau of Air Quality Analysis and Research, New York State Department of Environmental Conservation, 625 Broadway, Albany, NY 12233-3259, USA. Tel.: +1 518 402 8402; fax: +1 518 402 9035.

E-mail address: chogrefe@dec.state.ny.us (C. Hogrefe).

[1]In partnership with the US Environmental Protection Agency.

1. Introduction

Motivated by growing concerns about the detrimental effects of fine particulate matter ($PM_{2.5}$) on human health, the US Environmental Protection Agency (EPA) recently promulgated a National Ambient Air Quality Standard (NAAQS) for $PM_{2.5}$. The $PM_{2.5}$ standard includes a 24-h average limit ($65\,\mu g\,m^{-3}$ for the 98th percentile of the 24-h average for each year averaged over three consecutive years) and annual average limit ($15\,\mu g\,m^{-3}$ averaged over three consecutive years). Except for a few urban areas in the US, the annual standard is the primary concern. Over the next several years, grid-based photochemical models such as the community multiscale air quality (CMAQ) model (Byun and Ching, 1999), the regional modeling system for aerosols and deposition (REMSAD) (ICF Consulting, 2002), the comprehensive air quality model with extensions (CAMx) (Environ, 2004), and other regional models are expected to be used by regulatory agencies in the design of emission control strategies aimed at meeting and maintaining the NAAQS for O_3 and $PM_{2.5}$. The evaluation of these models for a simulation of current conditions is a necessary prerequisite for using them to simulate future conditions (US EPA, 2005a). Compared to the evaluation of summertime episodic O_3 simulations, the issue is more complex for regulatory simulations of $PM_{2.5}$ because simulations need to be performed for an entire annual cycle while taking into account the chemical speciation of $PM_{2.5}$ in order to identify its origin and develop meaningful strategies for meeting $PM_{2.5}$ attainment levels (Stein and Lamb, 2002; Blanchard and Hidy, 2003). From a measurement perspective, Malm et al. (2004) have presented an analysis of temporal and spatial patterns in speciated $PM_{2.5}$ measurements from the interagency monitoring to protect visual environments (IMPROVE) network. To carry out similar analysis for model predictions, extended annual simulations are necessary, and such extended simulations have been performed only recently. The evaluation of such annual simulations has been the topic of several recent studies (e.g. Gilliam et al., 2006; Gego et al., 2006; Swall and Davis, 2006; Porter et al., 2004; Eder and Yu, 2006). In this study, we add to this emerging body of knowledge by applying temporal scale analysis as an additional technique to evaluate an annual simulation of $PM_{2.5}$ and its chemical components over the eastern US. The concept of scale analysis is widely used for research in physical sciences, including meteorology, climatology, and air pollution (Goody et al. 1998; Salcedo et al. 1999) and also has been applied for several air quality model evaluation studies during the past several years (Hogrefe et al., 2001a, b; Biswas et al., 2001, Porter et al., 2004). Here, we first apply the technique to identify the temporal scales that are the largest contributors to the temporal variability in general and to periods of elevated $PM_{2.5}$ concentrations in particular. Next, the ability of two air quality models (CMAQ and REMSAD) to reproduce the variability and temporal evolution of total and speciated $PM_{2.5}$ fluctuations on different time scales extracted from observations is evaluated. Finally, the findings of these analyses are synthesized and discussed.

2. Models and database

2.1. Model setup

Meteorological fields for the air quality simulations were prepared by the PSU/NCAR MM5 model (Grell et al., 1994) version 3.6.1 over the continental United States at a horizontal cell size of 36 km for the time period from 1 January to 31 December 2001 (McNally, 2003; Gilliam et al., 2006). The Eta data assimilation system (EDAS) data were used for the boundary conditions in this simulation. The MM5 fields were then processed by the Meteorology-Chemistry Interface Program (MCIP) version 2.2. The emission inventory was based on the USEPA National Emissions Inventory for 2001. Emissions were processed by the SMOKE processor (Carolina Environmental Program, 2005) which incorporated the MOBILE6 module (US EPA, 2003) for mobile source emissions and the BEIS3.12 model for biogenic emissions (US EPA, 2005b). The seasonality of the ammonia emissions, an important consideration for prediction of $PM_{2.5}$, was estimated based on seasonal information from Gilliland et al. (2003), Pinder et al. (2004), and Goebes et al. (2003). These meteorological and emission fields were then provided as input to two photochemical models, namely CMAQ (February 2004 version) and REMSAD version 7.061 (ICF Consulting, 2002), both run at a horizontal resolution of 36 km over the continental United States. Lateral chemical boundary conditions for both models were prepared from a global simulation with the GEOS-CHEM model (Bey et al., 2001).

Additional details of the setup of these model simulations can be found online at US EPA (2004).

2.2. Observations

This study utilizes a variety of observations from different networks. Hourly observations of surface temperature and wind speed were obtained from the Techniques Data Laboratory (TDL) data set maintained by the Data Support Section at the National Center for Atmospheric Research (NCAR-DSS). Hourly surface O_3 observations, hourly $PM_{2.5}$ concentrations measured by tapered element oscillation microbalance (TEOM) monitors (Rupprecht and Pataschnik Co., Inc., 2000), and 24-h average $PM_{2.5}$ concentrations measured at monitors following the federal reference method (FRM) protocol were retrieved from EPA's air quality system (AQS) database. 24-h average $PM_{2.5}$ concentrations measured on Teflon filters were also obtained from the speciated trends network (STN). Speciated $PM_{2.5}$ measurements were obtained from the IMPROVE network, the clean air status and trends network (CASTNet), and STN. Because of differences in measurement techniques and instrumentation, sampling frequencies, and site location criteria, model performance had to be examined on a species-by-species and network-by-network basis. The analysis presented in this paper focuses on the Eastern United States. Monitoring sites were only included in the analysis if at least 70% of the data were available. All monitors were assigned to the model grid cells in which the monitor was located. If more than one monitor was located in the same grid cell, the same model predicted value was assigned to all monitors. All analyses presented in this paper were performed over the entire annual cycle from 1 January to 31 December 2001 with the exception of O_3 for which only the time period from 1 May to 30 September 2001, the so-called O_3-season, was considered.

2.3. Methods of analysis

Time series of atmospheric pollutant concentrations contain fluctuations occurring on many different timescales. Following the approach outlined in Rao et al. (1997) and Hogrefe et al. (2000), a spectral decomposition technique was applied to estimate temporal variations in observed and predicted time series. To this end, time series of meteorological variables and pollutant concentra-

tions were spectrally decomposed into fluctuations occurring on the intra-day (time period less than 12 h), diurnal (12–48 h), synoptic (2–21 days), and longer-term or baseline (greater than 21 days) time scales using the Kolmogorov–Zurbenko (KZ) filter as described in Hogrefe et al. (2000). Note that the intra-day and diurnal components could be estimated only for the variables that were measured on an hourly basis, while the synoptic and longer-term (baseline) components could be estimated for variables measured hourly, daily or weekly.

3. Results and discussion

3.1. Spectral features of observed and predicted $PM_{2.5}$ time series

Hogrefe et al. (2000) have shown that days of high daily maximum O_3 concentrations coincided with days of increased forcing on the diurnal, synoptic and baseline time scales in both observations and model predictions for a 3-month summer period. In this study, we expand upon this analysis by comparing the component forcing between all days during 2001 and only those days characterized by high $PM_{2.5}$ concentration. In other words, while the daily average $PM_{2.5}$ concentration is affected by each of the four time-scale components, our goal is to identify the timescales that show the strongest relationship between the strength of the forcing and the daily average concentration. For this purpose, high $PM_{2.5}$ days were defined at each monitor as those days on which the daily average $PM_{2.5}$ concentration exceeded the 90th percentile of all concentration values. To estimate the components' strength on each day, the daily mean was used as a surrogate for the strength of the baseline and synoptic components on that day, while for the diurnal and intra-day component, the standard deviation over the 24-h period under consideration was used. The relative strength of a component on a high $PM_{2.5}$ day is defined as the difference between the strength of a particular forcing on that day and the average strength of that forcing over all days.

This analysis was performed for total $PM_{2.5}$ measured by TEOM monitors from the AQS database and the CMAQ and REMSAD predictions at the corresponding grid cells. Fig. 1 illustrates that days of high $PM_{2.5}$ concentrations are characterized by an increase of the component forcing on the diurnal, synoptic, and longer-term (baseline) time scales, while there is little difference
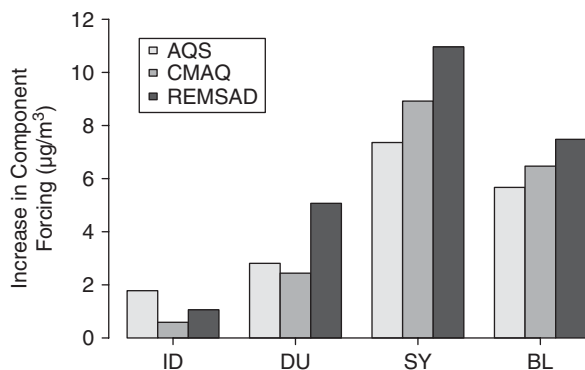
Fig. 1. The difference of component forcing as defined in the text (ID is the intra-day forcing, DU is the diurnal forcing, SY is the synoptic forcing, and BL is the baseline forcing) between high $PM_{2.5}$ days and the average over all days.

between the strength of the intra-day component on all days vs. episodic days. These results are evidence that high $PM_{2.5}$ concentrations tend to occur when the synoptic forcing is positive and the baseline forcing is stronger than average. While both models capture the overall difference in observed component increases (largest increases for the synoptic component, followed by the baseline, diurnal and intra-day component), they tend to overestimate the increase of the synoptic and baseline forcing.

The analysis presented in Fig. 1 was performed for total $PM_{2.5}$ mass measured by TEOM instruments and predicted by CMAQ and REMSAD at the corresponding grid cells. However, it is likely that the relative importance of different time scales to total $PM_{2.5}$ mass varies by chemical component and season. Unfortunately, hourly speciated $PM_{2.5}$ measurements were not available at a sufficient number of sites during 2001 to perform this analysis for observations, but the analysis could be carried out for model predictions. The goal of the following analysis is to identify the most important chemical components for each time scale and season in the model predictions. This analysis can then provide a context for the model evaluation performed in Sections 3.2 and 3.3. To this end, Fig. 2 displays spatial patterns of the chemical species that contribute the most to elevated $PM_{2.5}$ concentrations in the CMAQ simulations for each season (columns) and each temporal component (rows). In wintertime, most regions show nitrate to be the strongest component forcing on high $PM_{2.5}$ days for the intra-day, diurnal, and synoptic components. Exceptions are some urban regions such as New York City and Atlanta where CMAQ predicts the

strongest $PM_{2.5}$ forcing stemming from primary emissions of either carbon or crustal material. On the baseline timescale, these two chemical components are the dominant contributors to increased component forcing over the northern portion of the analysis domain, presumably at least in part due to the increased amount of $PM_{2.5}$ emissions from wood burning during winter months in the temporal allocation of annual total $PM_{2.5}$ emission inventories. During spring (March–May), nitrate again remains the strongest contributor to increased component forcing on high $PM_{2.5}$ days for all temporal components except for the synoptic component for which sulfate becomes more dominant over the eastern part of the analysis domain. Primary organic carbon (OC) is the most important contributor to increased component forcing on all time scales for northern Florida during spring, likely due to the temporal allocation of wildfire emissions during this time period. During summertime, sulfate is the strongest contributor to synoptic and baseline scale forcing on high $PM_{2.5}$ days at virtually all modeling grid cells in the analysis domain, while secondary OC seems to play an important role on the diurnal and intra-day time scale in the southern part of the analysis domain. While there is little change in the dominance of sulfate on the synoptic scale component in fall, secondary OC from the oxidation of anthropogenic VOC sources becomes dominant on the baseline time scale and, over smaller regions, on the diurnal and intra-day time scales as well.

In summary, the spectral decomposition of speciated CMAQ $PM_{2.5}$ predictions reveals the complex structure of seasonal and spatial variations in those model predictions. Depending on the region and season, either $PM_{2.5}$ species related to primary emissions such as primary OC or secondary $PM_{2.5}$ species such as sulfate are the dominant contributors to high $PM_{2.5}$ concentrations. In the following Section 3.2, we will evaluate model performance for the different spectral components and chemical species, and discuss the implications of that analysis to the likelihood of the 'model-predicted' spatial, temporal and chemical variations presented here reflecting reality.

3.2. Comparison of observed and predicted component variances

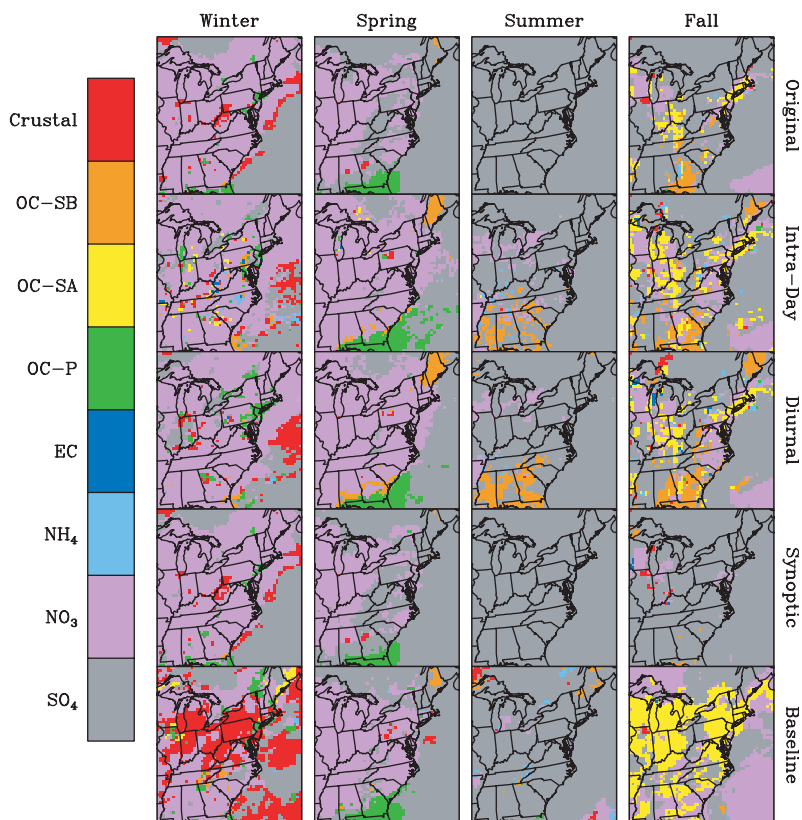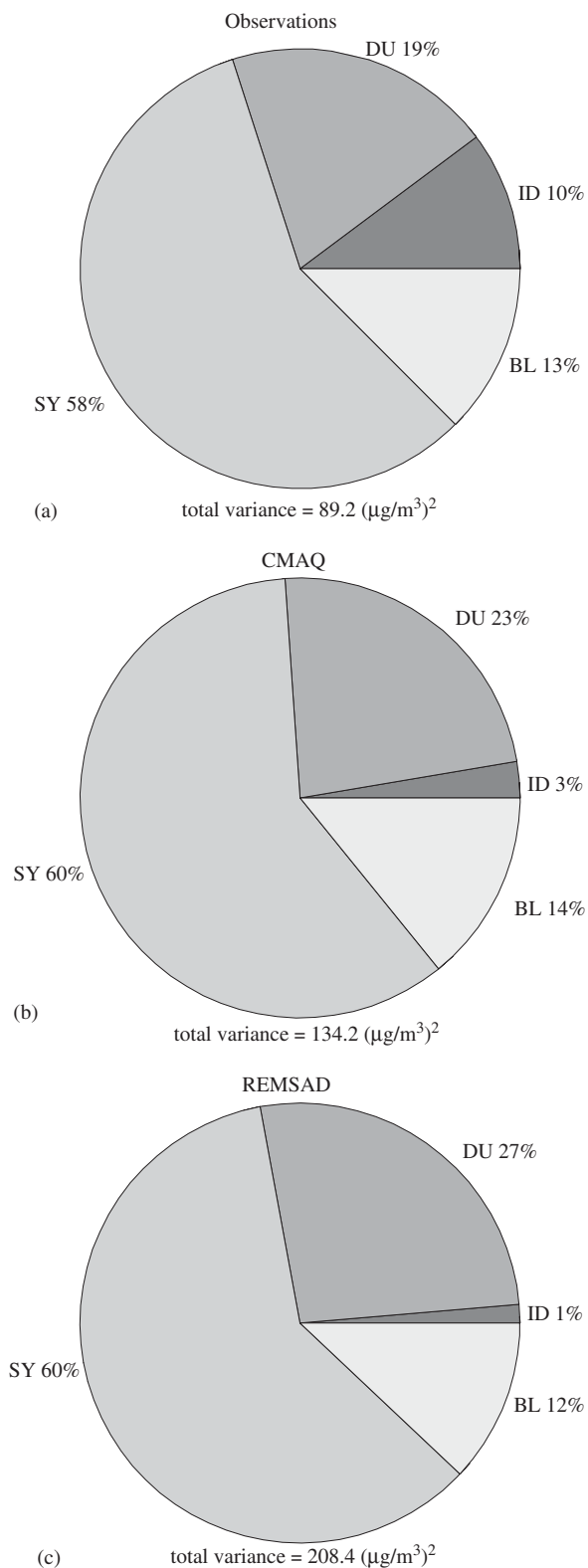The first measure of model performance applied in this study was the comparison of observed and

Fig. 2. Maps indicating the chemical species that show the largest differences between days of elevated $PM_{2.5}$ concentrations and average conditions in the CMAQ simulations for each season (vertical columns) and each temporal component (horizontal rows). "EC" represents elemental carbon aerosols, "OC-P" represents primary organic carbon aerosols, "OC-SA" represents secondary anthropogenic carbon aerosols, and "OC-SB" represents secondary biogenic carbon aerosols.

predicted variability on the different spectral components. Figs. 3a–c present pie charts of the relative contributions of the different spectral components to total model variance for hourly $PM_{2.5}$ observations from TEOM monitors in the AQS database and the corresponding CMAQ and REMSAD predictions. In addition to these relative contributions, the variance of the original time series is shown below each pie chart. The numbers presented in these figures were averaged over all available TEOM monitors. In terms of the relative contributions, both CMAQ and REMSAD capture the contributions on the synoptic and baseline time scales, while they underestimate the relative importance of the intra-day time scale and overestimate the relative importance of the diurnal time scale. Both models also overestimate the total variance, REMSAD to a larger extent than CMAQ. This pie chart highlights the importance of simulating processes on the synoptic and baseline scales to

account for a large fraction of observed $PM_{2.5}$ variability.

Figs. 4–6 present scatter plots of predicted vs. observed standard deviations for different variables and networks. The different spectral components are distinguished by different colors and symbols in each scatter plot. Each point in the scatter plots represents the species standard deviation at a monitor and corresponding model grid cell. REMSAD O_3 predictions were not compared against observations because REMSAD is not considered to be an adequate tool for simulating oxidant photochemistry. The comparison of the standard deviations of hourly O_3 observations and CMAQ predictions in Fig. 4a shows that the magnitude of intra-day fluctuations is underestimated at virtually every monitor. Similarly, the standard deviation of the diurnal component is underestimated by CMAQ at most AQS monitors as evidenced by the location of most points below

Observations

DU 19%

ID 10%

BL 13%

SY 58%

(a)

total variance = 89.2 $(\mu g/m^3)^2$

CMAQ

DU 23%

ID 3%

BL 14%

SY 60%

(b)

total variance = 134.2 $(\mu g/m^3)^2$

REMSAD

DU 27%

ID 1%

BL 12%

SY 60%

(c)

total variance = 208.4 $(\mu g/m^3)^2$

the 1:1 line, indicating lower daytime/nighttime O_3 differences in the model compared to observations. This may be partially due to less frequent and intense nighttime titration events in CMAQ, a feature often observed in modeling studies (Dennis et al., 2004). For the synoptic O_3 components, most points are near the 1:1 line, while the magnitude of O_3 baseline fluctuations appears to be overestimated at most locations. Since baseline fluctuations may be partially affected by fluctuations in boundary conditions, this finding implies that further analysis of the GEOS-CHEM-derived boundary conditions utilized in this simulation may be warranted. Figs. 4b,c show the corresponding results for observed hourly $PM_{2.5}$ at AQS sites against CMAQ and REMSAD predictions, respectively. Consistent with Fig. 3, the largest standard deviation of both observations and model predictions is in the synoptic component, followed by the diurnal, baseline and intra-day components. Both CMAQ and REMSAD underestimate the standard deviation of the intra-day component at virtually all locations. The standard deviations of the CMAQ predictions (Fig. 4b) show a relatively even scatter around the 1:1 line for the diurnal, synoptic, and baseline components, while REMSAD overpredicts these standard deviations at a majority of monitoring locations.

In addition to the hourly O_3 and $PM_{2.5}$ concentrations observed and predicted at the AQS monitors depicted in Fig. 4, this analysis also was performed for daily average sulfate and nitrate concentrations from the IMPROVE network (Fig. 5) and sulfate, nitrate, EC and OC concentrations from the STN network (Fig. 6). Only the synoptic and baseline components could be estimated from these daily measurements. The results for sulfate at IMPROVE sites show that CMAQ captures the standard deviation of both the synoptic and baseline components, while REMSAD tends to underestimate the magnitude of these sulfate fluctuations. The difference between CMAQ and REMSAD is especially pronounced on the baseline time scale, indicating a strong underestimation of the seasonal sulfate fluctuations by REMSAD (Figs. 5a, b). For nitrate, both models tend to

Fig. 3. Pie chart showing the relative contributions of the different temporal components to the total variance of hourly $PM_{2.5}$ concentrations, averaged over all AQS monitors in the analysis domain: (a) Observations from TEOM monitors, (b) CMAQ predictions, and (c) REMSAD predictions.
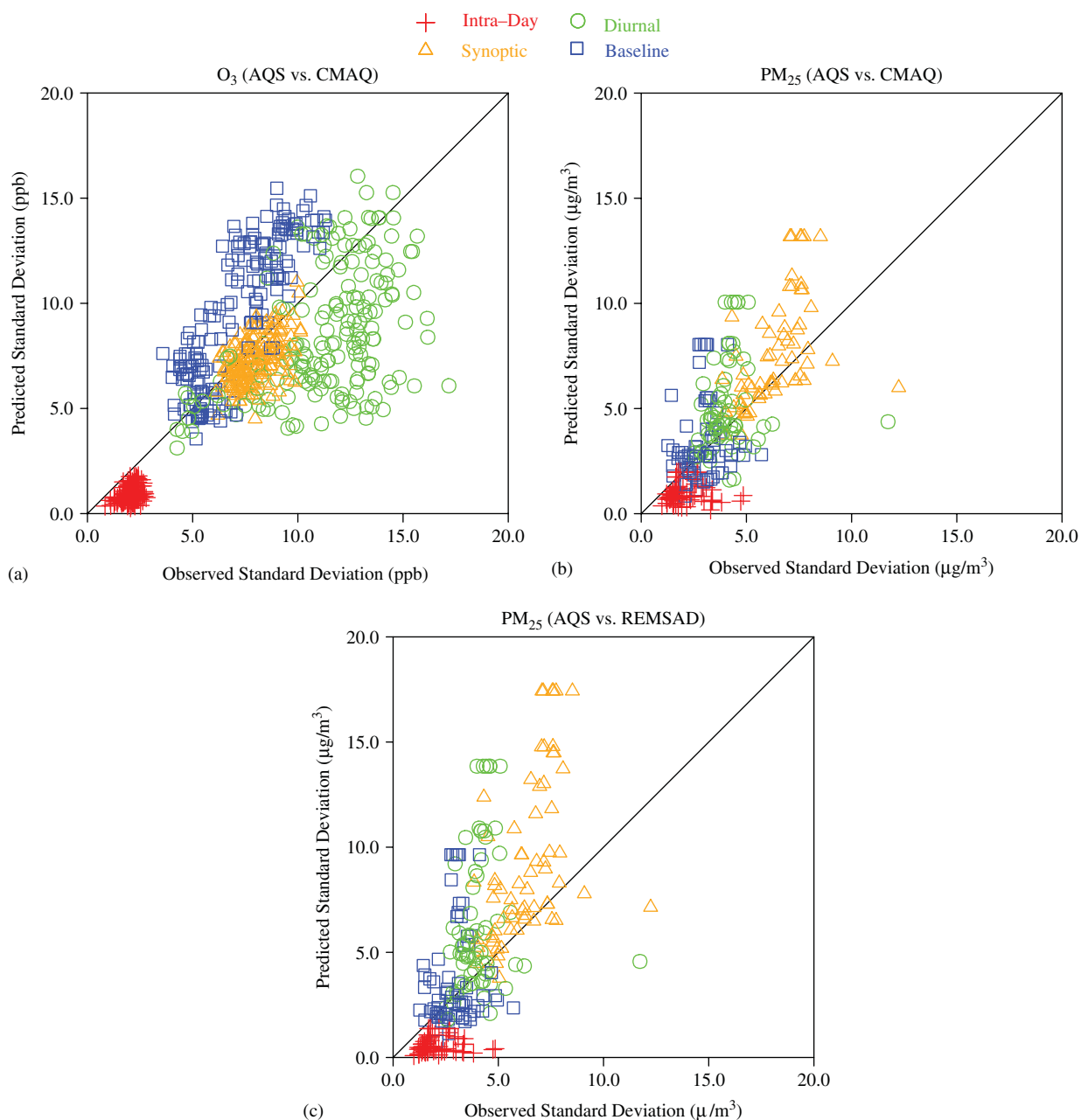
Fig. 4. Scatter plots of the predicted vs. observed standard deviations for different spectral components at all monitors included in the analysis. The different temporal components are distinguished by their colors: (a) CMAQ vs. observations for hourly O_3 at AQS monitors, (b) CMAQ vs. observations for hourly $PM_{2.5}$ at AQS monitors, and (c) REMSAD vs. observations for hourly $PM_{2.5}$ at AQS monitors.

overestimate the magnitude of both synoptic and baseline fluctuations (Figs. 5c, d).

For sulfate, the results are similar to Figs. 5a, b when considering observations from the STN network and the corresponding CMAQ and REMSAD predictions (Figs. 6a, b). This is an indication of the regional-scale nature of sulfate variations that

exhibit similar temporal features at both the rural IMPROVE and the urban STN monitors on the synoptic and baseline scales. The observed and predicted standard deviations of the nitrate synoptic and baseline components, on the other hand, show a greater magnitude and more variability at the STN monitors (Figs. 6c, d) compared to the IMPROVE
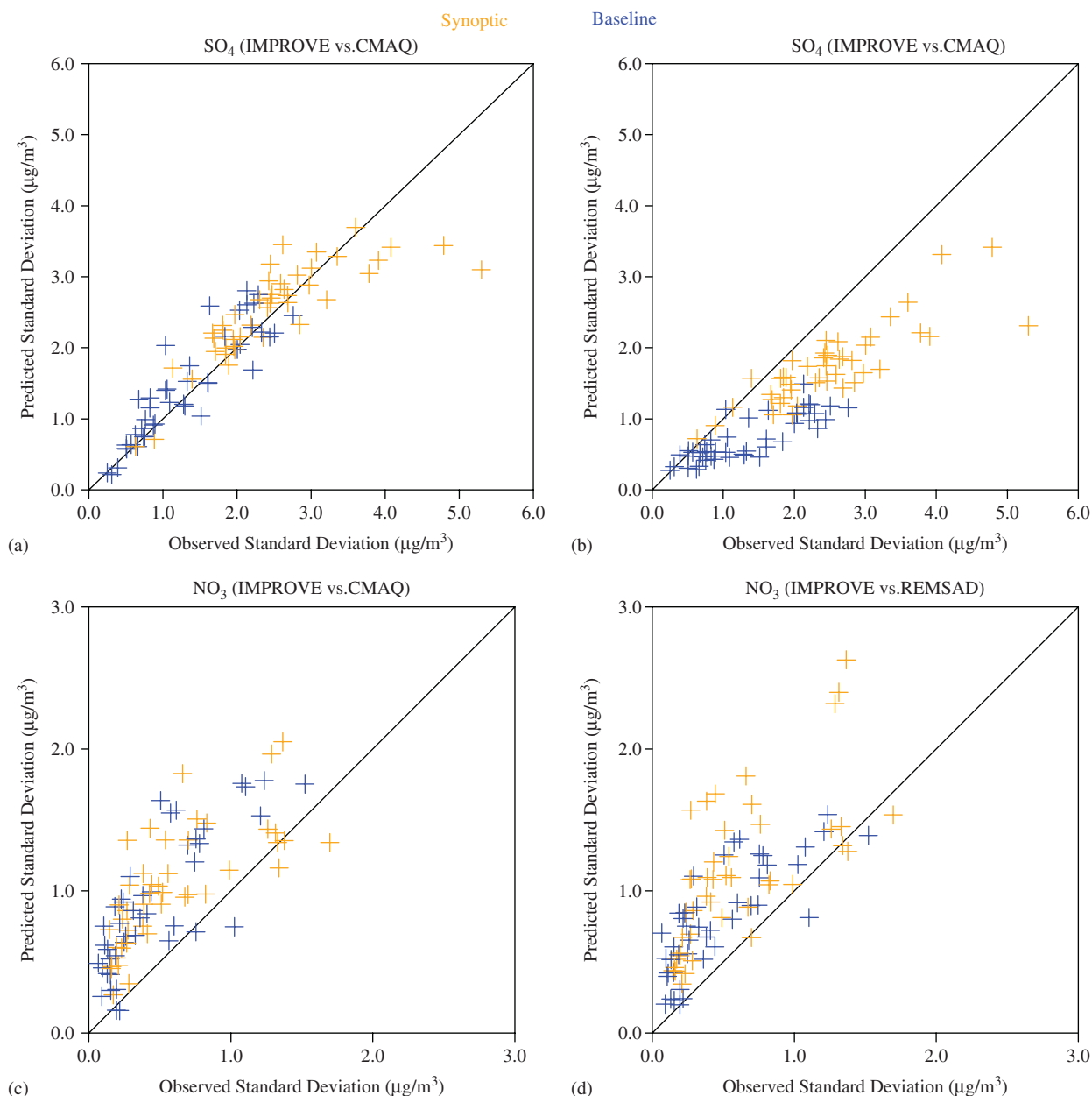
Fig. 5. As in Fig. 4, but for daily average SO_4 and NO_3 concentrations at IMPROVE sites.

monitors (Figs. 5c, d). Finally, results for EC and OC at the STN monitors indicate that neither CMAQ nor REMSAD systematically over- or underpredict the magnitude of synoptic and baseline scale variations, but differences between observed and predicted standard deviations can be as large as a factor of two at a number of stations, especially for OC (Figs. 6e–h).

In summary, neither of the two modeling systems considered in this study capture the variability of

the high-frequency, intra-day component for any of the variables for which hourly measurements are available. As pointed out by Hogrefe et al. (2001a, b) and Biswas et al. (2001), the failure of the air quality model to capture variability on this scale for pollutant concentrations is not surprising given that this variability was also not captured for several key meteorological fields such as temperature and wind speed that are used as input to the air quality model. As discussed by Gilliam et al. (2006),
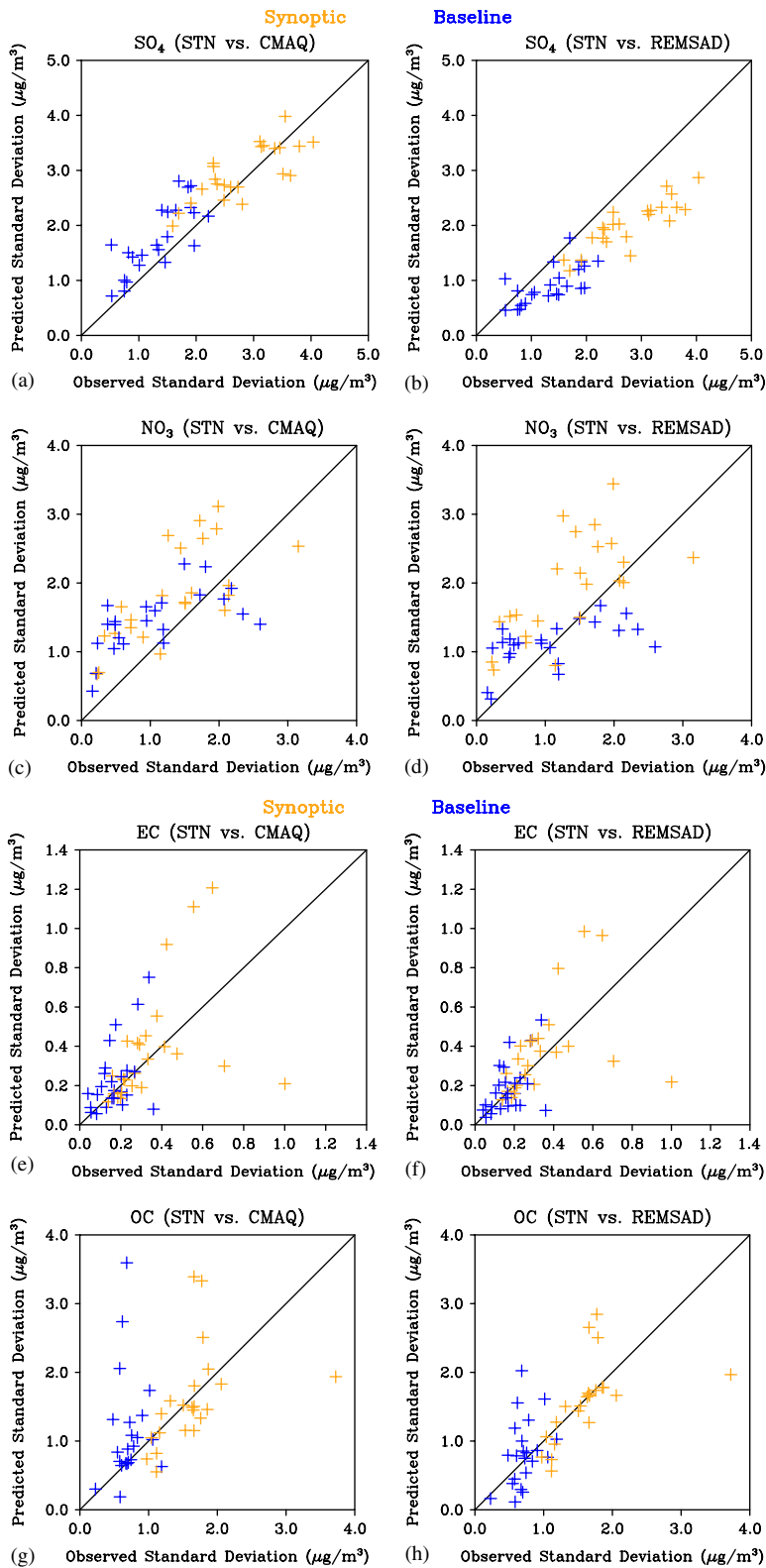
Fig. 6. (a–d) As in Fig. 4, but for daily average SO$_4$ and NO$_3$ concentrations at STN sites. (e–h) As in Fig. 4, but for daily average EC and OC concentrations at STN sites.

this points to possible problems with the simulation of boundary layer mixing processes, boundary layer growth and collapse, and convective clouds in the MM5 simulation. Presumably, higher grid resolution along with better parameterizations of boundary layer processes would be needed to capture more of this variability in MM5, a necessary prerequisite for capturing it in CMAQ as well.

3.3. Correlations on different time scales

As another measure of model performance, correlations between different temporal components embedded in time series of the observed and predicted variables were computed for temperature, wind speed and O_3 as well as total and speciated $PM_{2.5}$ from the different measurement networks (Table 1a, b). The correlations were computed at each site for a given variable/network/model combination, and Table 1a, b list the median value of the correlation across all sites for a given variable/ network/model combination. As noted in Section 3.2, correlations were computed for the time period from 1 January to 31 December for all variables except O_3, for which only the period from 1 May to 30 September was considered. For the meteorological variables (temperature and wind speed), correlations increase with increasing time scale, i.e. correlations are lowest for the intra-day component ($r < 0.2$) and highest for the baseline component ($r > 0.9$). Correlations are relatively high for the diurnal component ($r > 0.6$), in part due to the inherent cyclical nature of this component. Furthermore, correlations are lower when the time series of the diurnal amplitudes are considered (not shown). It is not surprising that the correlation is highest on

Table 1

(a) Correlations between different temporal components embedded in hourly time series of observed and predicted temperature, wind speed, ozone and total $PM_{2.5}$. Median values are shown for each network/variable

	#Sites	Intra-day		Diurnal		Synoptic		Baseline	
Temperature TDL/MM5	738		0.18		0.90		0.95		0.99
Wind speed TDL/MM5	735		0.02		0.60		0.84		0.90
O_3 AQS/CMAQ	688		0.07		0.74		0.67		0.89
$PM_{2.5}$ TEOM	67								
CMAQ REMSAD		0.01	0.03	0.25	0.25	0.70	0.63	0.04	0.10

(b) Correlations between different synoptic and baseline components embedded in time series of observed and predicted $PM_{2.5}$ from different networks. Median values are shown for each network/variable

	#Sites	Synoptic		Baseline	
		CMAQ	REMSAD	CMAQ	REMSAD
$PM_{2.5}$ FRM (24-h average every 3rd day)	938	0.68	0.65	0.60	0.51
$PM_{2.5}$ STN (24-h average every 3rd day)	25	0.60	0.63	0.38	0.35
SO_4 improve (24-h average every 3rd day)	44	0.77	0.70	0.89	0.77
SO_4 CASTnet (weekly average every week)	48	0.85	0.72	0.94	0.88
SO_4 STN (24-h average every 3rd day)	23	0.72	0.70	0.85	0.74
NO_3 improve STN (24-h average every 3rd day)	44	0.46	0.54	0.88	0.78
NO_3 CASTnet (weekly average every week)	48	0.51	0.46	0.89	0.83
NO_3 STN (24-h average every 3rd day)	23	0.39	0.42	0.83	0.66
NH_4 CASTnet (weekly average every week)	48	0.71	0.72	0.55	0.45
NH_4 STN (24-h average every 3rd day)	23	0.63	0.66	0.52	0.37
EC STN (24-h average every 3rd day)	23	0.41	0.39	0.15	0.32
OC STN (24-h average every 3rd day)	22	0.48	0.55	0.24	0.28
Crustal STN (24-h average every 3rd day)	23	0.34	0.29	−0.35	−0.39

the synoptic and baseline time scales since MM5 model predictions were nudged towards analysis fields using four-dimensional data assimilation techniques. These results are consistent with the findings of Hogrefe et al. (2001a) who showed that two mesoscale meteorological models did not capture the temporal evolution of fluctuations on the intra-day scale for temperature, wind speed and water vapor mixing ratio. They also pointed out that this would put a limit on the ability of any air quality model utilizing these input fields to capture pollutant fluctuations on this scale. Indeed, the O_3 correlations calculated in this study and presented in Table 1 follow a similar pattern as temperature and wind speed, with correlations on the intra-day time scale being less than 0.1 and correlations on the baseline time scale being 0.87, strongly supporting the notion that it is crucial to capture all relevant scales in the meteorological fields if they are to be captured in the air quality model.

Among all observations of total $PM_{2.5}$ mass considered in this study, only measurements with TEOM instruments obtained from the AQS database were available on an hourly basis and could be used to compare observed and predicted intra-day and diurnal components of this parameter. For this comparison, it is striking that the correlations between TEOM observations and model predictions are poor on the diurnal and baseline components for both CMAQ and REMSAD despite the high correlations of the meteorological variables on these time scales. Plausible reasons for the discrepancy on the diurnal scale may be misrepresentations of the strength of vertical mixing in the model or the magnitude of primary $PM_{2.5}$ emissions from area and mobile sources. For the baseline time scale, Hogrefe et al. (2004) illustrated that the low correlations stem from the higher $PM_{2.5}$ predictions by both models during wintertime while TEOM measurements show a decrease. Part of this decrease in TEOM measurements during wintertime is likely caused by the high operating temperatures of most of the currently deployed TEOM instruments (30 and 50 °C). Volatilization losses can occur when the sample is heated from ambient temperature to the operating temperature, and such losses tend to be higher during colder ambient temperatures (Allen et al., 1997; Schwab et al., 2004). On the other hand, CMAQ and REMSAD utilize MM5-simulated temperatures to calculate the partitioning between gas and particle phase. In other words, there is an inherent difference between the measurement tech-

nique utilizing a constant operating temperature and the modeling approach utilizing time-varying predicted ambient temperature, and this difference exhibits seasonality, thereby affecting the baseline comparisons. Support for this explanation comes from the higher baseline correlations when CMAQ and REMSAD are compared against $PM_{2.5}$ filter observations from FRM monitors and the STN network (Table 1b). This highlights the importance of conducting $PM_{2.5}$ model evaluation on a network-by-network basis. Combining measurements from different networks for model evaluation can only be performed in circumstances where it has been demonstrated that the variable and metric of interest is not affected by differences in sampling protocols, monitor siting criteria, and monitoring techniques. For example, Gego et al. (2005) demonstrated that monthly average sulfate concentrations measured by IMPROVE and CASTNet in the eastern US show similar temporal and spatial features and one might consider combining them when evaluating longer-term sulfate predictions from air quality models.

Correlations between the synoptic and baseline components of sulfate measured by the IMPROVE, CASTNet, and STN networks, and predicted by CMAQ and REMSAD are consistently greater than 0.7, with baseline correlations exceeding 0.85 for CMAQ and 0.74 for REMSAD. It is noteworthy that there is relatively little difference in model performance across the different networks, a result that is consistent both with the findings of Gego et al. (2005), who compared speciated $PM_{2.5}$ measurements from different networks and also with the regional-scale nature of sulfate concentrations in the eastern United States. Furthermore, correlations for REMSAD are consistently lower than those for CMAQ for $PM_{2.5}$ across all networks. For nitrate, correlations on the baseline time scale are similar to those for sulfate, but correlations on the synoptic time scale are lower. For the baseline, CMAQ correlations are consistently higher than those for REMSAD for sulfate and nitrate. In contrast to baseline correlations for sulfate and nitrate, correlations are relatively low for ammonium. A likely contributor to these lower correlations is the seasonal characterization of NH_3 emissions. The seasonality for NH_3 emissions is a well-known uncertainty that is currently being investigated from both bottom-up inventory development and from top-down estimation methods (Gilliland et al., 2003, 2006). Model predicted

concentrations of elemental carbon (EC), OC, and crustal material are strongly influenced by emissions of primary $PM_{2.5}$ since there is no secondary formation mechanism for EC and crustal material in CMAQ and REMSAD. Consequently, the relatively weak correlations between the observed and predicted baseline components for these species point to potential problems in the temporal allocation of $PM_{2.5}$ emissions during emission processing.

To investigate this issue, we constructed the baseline component of EC observations, CMAQ and REMSAD predictions, and total $PM_{2.5}$ emissions at several STN monitoring locations. This analysis is shown in Figs. 7a and b. The strong correlation between $PM_{2.5}$ emissions and model-predicted EC concentrations is clearly visible at the Decatur, GA monitor and, to a slightly lesser extent, at the Bronx, NY monitor. In both cases, the relatively poor correlation between observations and model predictions seems to be largely driven by the temporal signature of the $PM_{2.5}$ emissions. Therefore, in order to improve model performance on longer time scales for primary species such as EC and crustal material, it is necessary to improve the temporal characterization of primary $PM_{2.5}$ emissions.

In summary, the results of the correlation analysis presented in Table 1a, b illustrate that the models exhibit greatest skills at capturing longer-term (seasonal) fluctuations for temperature, wind speed, O_3, sulfate and nitrate. For the variables for which hourly measurements were available, correlations were insignificant on the intra-day time scale, suggesting that these simulations are not skillful in simulating the higher-frequency variations in meteorological variables and pollutant levels for all species presented. This finding also confirms that problems in the meteorological model to capture high-frequency fluctuations related to boundary layer processes and convective cloud formation (Gilliam et al., 2006) adversely impact the ability of air quality models to capture high-frequency fluctuations of pollutant levels. The fact that for total $PM_{2.5}$, ammonium, EC, OC and crustal $PM_{2.5}$ correlations are highest for the synoptic time scale implies problems with factors other than meteorology, such as emissions or boundary conditions, in capturing the baseline fluctuations. This suggests that capturing meteorological fluctuations on all scales is a necessary but not sufficient prerequisite for capturing pollutant fluctuations on all time scales, especially for the simulation of $PM_{2.5}$.
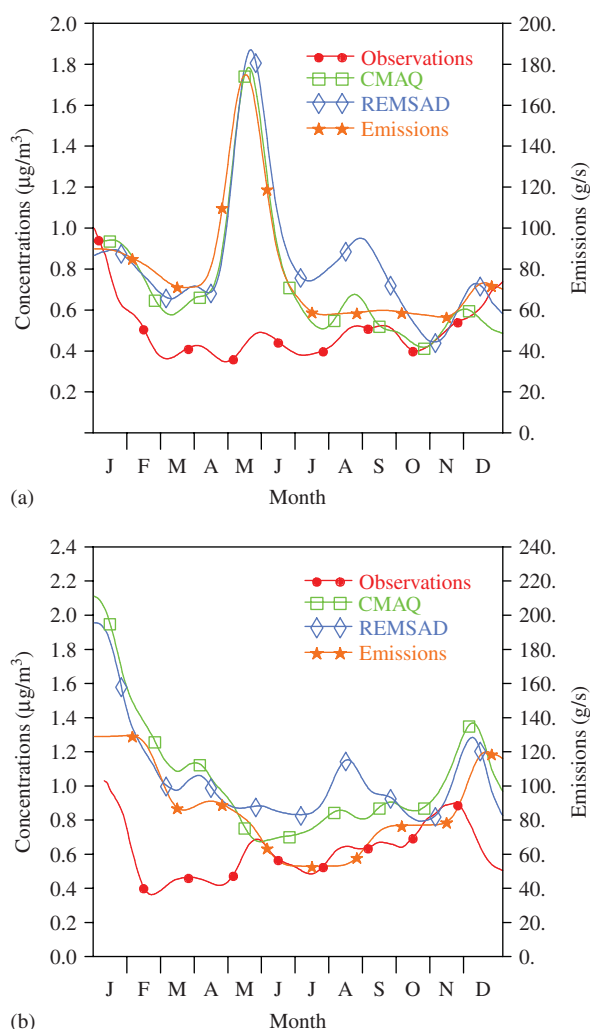
3.4. Discussion

The results presented in Sections 3.2 and 3.3 provide a context for interpreting the CMAQ-predicted contributions of different species and time scales during different seasons to high $PM_{2.5}$ concentrations presented in Section 3.1. First, the low correlations on the intra-day time scale for all hourly variables (meteorology, O_3 and total $PM_{2.5}$), along with the underestimation of variance on this time scale, lessen the importance of the results in the 'intra-day' row of results in Fig. 2. Furthermore, the model-predicted dominance of other primary $PM_{2.5}$ in winter and secondary anthropogenic organic

Fig. 7. Baseline of EC observations, CMAQ and REMSAD predictions, and $PM_{2.5}$ emissions (a) Decatur, Georgia (b) Bronx, New York.

aerosols in fall over large portions of the modeling domain requires further corroborative studies such as source apportionment, given the fairly poor correlations even on the baseline time scale between model predictions and observations from the STN network. For the same reasons, the model-predicted dominance of primary anthropogenic carbon aerosol on high $PM_{2.5}$ days on all time scales during spring in Florida should be viewed with caution. On the other hand, the CMAQ-predicted dominance of nitrate and sulfate on synoptic-scale forcing on high-$PM_{2.5}$ days in winter and summer, respectively, is supported by higher correlations and a closer match with observed variability across several networks. In summary, these results derived from both observations and model predictions support the notion set forth in previous observational studies (e.g. Malm et al., 2004) that inorganic secondary aerosols remain a dominant component of total $PM_{2.5}$ mass over much of the eastern US over all seasons, but that the role of primary $PM_{2.5}$ emissions and organic aerosol formation from anthropogenic and biogenic sources cannot be neglected.

4. Summary

The spectral decomposition of total $PM_{2.5}$ mass from TEOM observations and CMAQ and REMSAD model predictions revealed that days of high $PM_{2.5}$ concentrations are in general characterized by positive forcing from fluctuations having periods equal to or greater than a day (i.e., the diurnal, synoptic, and longer-term components) while the magnitude of intra-day fluctuations showed only small differences between average and episodic conditions. Both modeling systems captured this feature. However, both modeling systems did not capture most of the variability of the high-frequency, intra-day component for all variables for which this component could be estimated from observations. Presumably, higher grid resolution and improved characterization of boundary layer and convective processes in the meteorological model would be needed to capture the magnitude of these fluctuations. Furthermore, the results of the correlation analysis presented illustrate that correlations were insignificant on the intra-day time scale for all variables, suggesting that these models in the setup used for this study are not skillful in simulating the higher-frequency variations in meteorological variables and the levels of all

pollutants. The models exhibit greatest skills at capturing longer-term (seasonal) fluctuations for temperature, wind speed, O_3, sulfate and nitrate. The fact that for total $PM_{2.5}$, ammonium, EC, OC and crustal $PM_{2.5}$ correlations are highest for the synoptic time scale implies problems with factors other than meteorology, such as emissions or boundary conditions, in capturing the baseline fluctuations. This suggests that capturing meteorological fluctuations on all scales is a necessary but not sufficient prerequisite for capturing pollutant fluctuations on all time scales, especially for the simulation of $PM_{2.5}$.

References

Allen, G., Sioutas, C., Koutrakis, P., Reiss, R., Lurmann, F.W., Roberts, P.T., 1997. Evaluation of the TEOM method for measurement of ambient particulate mass in urban areas. Journal of the Air & Waste Management Association 47, 682–689.

Bey, I., Jacob, D.J., Yantosca, R.M., Logan, J.A., Field, B.D., Fiore, A.M., Li, Q., Liu, H.Y., Mickley, L.J., Schultz, M.G., 2001. Global modeling of tropospheric chemistry with assimilated meteorology: model description and evaluation. Journal of Geophysical Research 106 (D19), 23073–23096.

Biswas, J., Hogrefe, C., Rao, S.T., Hao, W., Sistla, G., 2001. Evaluating the performance of regional-scale photochemical modeling systems: part III—precursor predictions. Atmospheric Environment 35, 6129–6149.

Blanchard, C.L., Hidy, G.M., 2003. Effects of changes in sulfate, ammonia, and nitric acid on particulate nitrate concentrations in the Southeastern United States. Journal of the Air & Waste Management Association 53, 283–290.

Byun, D.W., Ching, J.K.S. (Eds.), 1999. Science Algorithms of the EPA Models-3 Community Multiscale Air Quality Model (CMAQ) modeling system. EPA/600/R-99/030. US Environmental Protection Agency, Office of Research and Development, Washington, DC, p. 20460.

Carolina Environmental Program, 2005. Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System, available online at http://cf.unc.edu/cep/empd/products/smoke/index.cfm

Dennis, R., Roselle, S., Gilliam, R., Arnold, J., 2004. High time-resolved comparisons for in-depth probing of CMAQ fine particle and gas predictions. Presented at the 27th NATO/CCMS International Technical Meeting on Air Pollution Modeling and Its Applications, Banff, Canada, 25–29 October.

Eder, B., Yu, S., 2006. A performance evaluation of the 2004 release of Models-3 CMAQ. Atmospheric Environment, this issue, doi:10.1016/j.atmosenv.2005.08.045.

ENVIRON Corp., 2004. User's guide to the Comprehensive Air Quality Model with Extensions (CAMx) Version 4.10S, available online at http://www.camx.com/pdf/CAMx.User.Guide.v4.10s.August2004.pdf

Gego, E., Porter, P.S., Hogrefe, C., Swall, J., Gilliland, A., Irwin, J.S., Rao, S.T., 2006. Objective comparison of CMAQ and REMSAD performances. Atmospheric Environment, this issue, doi:10.1016/j.atmosenv.2005.12.045.

Gego, E., Porter, P.S., Irwin, J.S., Hogrefe, C., Rao, S.T., 2005. Assessing the comparability of ammonium, nitrate and sulfate concentrations measured by three air quality monitoring networks. Pure and Applied Geophysics 162, 1–21.

Gilliam, R., Hogrefe, C., Rao, S.T., 2006. Alternative methods for evaluating meteorological models used in air quality applications. Atmospheric Environment, this issue, doi:10.1016/j.atmosenv.2005.01.023.

Goody, R., Anderson, J., North, G., 1998. Testing climate models: an approach. Bulletin of the American Meteorological Society 79, 2541–2549.

Gilliland, A.B., Dennis, R.L., Roselle, S.J., Pierce, T.E., 2003. Seasonal NH_3 emission estimates for the eastern United States based on ammonium wet concentrations and an inverse modeling method. Journal of Geophysical Research 108 (D15), 4477.

Gilliland, A.B., Appel, K.W., Pinder, R., Dennis, R.L., 2006. Seasonal NH_3 emissions for annual 2001 CMAQ simulation: inverse model estimation and evaluation. Atmospheric Environment, this issue, doi:10.1016/j.atmosenv.2005.12.066.

Goebes, M.D., Strader, R., Davidson, C., 2003. An ammonia emission inventory for fertilizer application in the United States. Atmospheric Environment, 2539–2550.

Grell, G.A., Dudhia, J., Stauffer, D., 1994. A description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5). NCAR Technical Note, TN-398 + STR. National Center for Atmospheric Research, Boulder, CO (138pp).

Hogrefe, C., Rao, S.T., Zurbenko, I.G., Porter, P.S., 2000. Interpreting the information in ozone observations and model predictions relevant to regulatory policies in the eastern United States. Bulletin of the American Meteorological Society 81, 2083–2106.

Hogrefe, C., Rao, S.T., Kasibhatla, P., Kallos, G., Tremback, C., Hao, W., Olerud, D., Xiu, A., McHenry, J., Alapaty, K., 2001a. Evaluating the performance of regional-scale photochemical modeling systems: part I—meteorological predictions. Atmospheric Environment 35, 4159–4174.

Hogrefe, C., Rao, S.T., Kasibhatla, P., Hao, W., Sistla, G., Mathur, R., McHenry, J., 2001b. Evaluating the performance of regional-scale photochemical modeling systems: part II—ozone predictions. Atmospheric Environment 35, 4175–4188.

Hogrefe, C., Jones, J.M., Gilliland, A., 2004. Evaluation of an annual simulation of particulate matter over the continental US—which spatio-temporal features can be captured? Preprints of the 27th NATO/CCMS International Technical Meeting on Air Pollution Modeling and Its Applications, Banff, Canada, 25–29 October, pp. 166–173.

ICF Consulting, 2002. User's Guide to the Regional Modeling System for Aerosols and Deposition (REMSAD) Version 7. Systems Applications International/ICF Consulting, San Rafael, CA 94903, 153pp.

Malm, W.C., Schichtel, B.A., Pitchford, M., Ashbaugh, L.L., Eldred, R.A., 2004. Spatial and monthly trends in speciated fine particle concentration in the United States. Journal of Geophysical Research 109.

McNally, D., 2003. Annual application of MM5 for calendar year 2001. Prepared for US EPA by D. McNally, Alpine Geophysics, Arvada, CO, March, 179pp.

Pinder, R.W., Strader, R., Davidson, C.I., Adams, P.J., 2004. A temporally and spatially resolved ammonia emission inventory for dairy cows in the United States. Atmospheric Environment 38, 3747–3756.

Porter, P.S., Swall, J., Gilliam, R., Gego, E.L., Hogrefe, C., Gilliland, A., Irwin, J.S., Rao, S.T., 2004. Temporal signatures of air quality observations and model outputs: how well do air quality models capture relevant time scales? 27th NATO/CCMS International Technical Meeting on Air Pollution Modeling and Its Applications, Banff, Canada, 25–29 October.

Rao, S.T., Zurbenko, I.G., Neagu, R., Porter, P.S., Ku, J.-Y., Henry, R.F., 1997. Space and time scales in ambient ozone data. Bulletin of the American Meteorological Society 78, 2153–2166.

Rupprecht and Pataschnik Co., Inc., 2000. TEOM Series 1400a ambient particulate monitor, available online at http://www.rpco.com/products/ambprod/amb1400/

Salcedo, R.L.R., Alvim Ferraz, M.C.M., Alves, C.A., Martins, F.G., 1999. Time series analysis of air pollution data. Atmospheric Environment 33, 2361–2372.

Schwab, J.J., Spicer, J., Demerjian, K.L., Ambs, J.L., Felton, H.D., 2004. Long-term field characterization of tapered element oscillating microbalance and modified tapered element oscillating microbalance samplers in urban and rural New York State llocations. Journal of the Air & Waste Management Association 54, 1264–1280.

Stein, A.F., Lamb, D., 2002. Chemical indicators of sulfate sensitivity to nitrogen oxides and volatile organic compounds. Journla of Geophysical Research 107, 4449.

Swall, J., Davis, J., 2006. A Bayesian statistical approach for the evaluation of CMAQ. Atmospheric Environment, this issue, doi:10.1016/j.atmosenv.2005.12.058.

US Environmental Protection Agency (US EPA), 2005a. Guidance on the Use of Models and Other Analyses in

Attainment Demonstrations for the 8-h Ozone NAAQS. 2005, EPA-454/R-05-002, Environmental Protection Agency, Research Triangle Park, NC.

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

US Environmental Protection Agency (US EPA), 2004. Interstate Air Quality Rule technical information, available online at http://www.epa.gov/interstateairquality/technical.html

US Environmental Protection Agency (US EPA), 2005b. Biogenic Emissions Inventory System (BEIS) Modeling, available online at http://www.epa.gov/asmdnerl/biogen.html