Atmospheric Environment 43 (2009) 2561–2570

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

A combined model–observation approach to estimate historic gridded fields of $PM_{2.5}$ mass and species concentrations

Christian Hogrefe [a,d,*], Barry Lynn [b], Richard Goldberg [c], Cynthia Rosenzweig [c], Eric Zalewsky [d], Winston Hao [d], Prakash Doraiswamy [a], Kevin Civerolo [d], Jia-Yeong Ku [d], Gopal Sistla [d], P.L. Kinney [e]

[a] Atmospheric Sciences Research Center, University at Albany, 251 Fuller Road, Albany, NY 12203, USA
[b] Weather It Is, LTD, Efrat, Israel
[c] NASA-Goddard Institute for Space Studies, 2880 Broadway # 1, New York, NY 10025, USA
[d] New York State Department of Environmental Conservation, 625 Broadway, Albany, NY 12233-3259, USA
[e] Mailman School of Public Health, Columbia University, 60 Haven Ave., New York, NY 10032, USA

ARTICLE INFO

Article history:
Received 7 July 2008
Received in revised form
13 February 2009
Accepted 13 February 2009

Keywords:
Long-term air quality modeling
Aerosol model evaluation
Air quality characterization
Data integration
Community Multiscale
Air Quality (CMAQ) model

ABSTRACT

This paper introduces a methodology for estimating gridded fields of total and speciated fine particulate matter ($PM_{2.5}$) concentrations for time periods and regions not covered by observational data. The methodology is based on performing long-term regional scale meteorological and air quality simulations and then integrating these simulations with available observational data. To illustrate this methodology, we present an application in which year-round simulations with a meteorological model (the National Center for Atmospheric Research/Penn State Mesoscale Model, hereafter referred to as MM5) and a photochemical air quality model (the Community Multiscale Air Quality Model, hereafter referred to as CMAQ) have been performed over the northeastern United States for 1988–2005. Model evaluation results for total $PM_{2.5}$ mass and individual species for the time period from 2000 to 2005 show that model performance varies by species, season, and location. Therefore, an approach is developed to adjust CMAQ output with factors based on these three variables. The adjusted model values for total $PM_{2.5}$ mass for 2000–2005 are compared against independent measurements not utilized for the adjustment approach. This comparison reveals that the adjusted model values have a lower root mean square error (RMSE) and higher correlation coefficients than the original model values. Furthermore, the $PM_{2.5}$ estimates from these adjusted model values are compared against an alternate method for estimating historic $PM_{2.5}$ values that is based on $PM_{2.5}/PM_{10}$ ratios calculated at co-located monitors. Results reveal that both methods yield estimates of historic $PM_{2.5}$ mass that are broadly consistent; however, the adjusted CMAQ values provide greater spatial coverage and information for $PM_{2.5}$ species in addition to total $PM_{2.5}$ mass. Finally, strengths and limitations of the proposed approach are discussed in the context of potential uses of this method.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Spatio-temporal information about ambient levels of air pollution is used in a variety of applications such as assessing the impact of emission control programs (Gégo et al., 2007) and investigating the link between air pollution and human health (e.g. Bell et al., 2004; Samoli et al., 2005, and references therein). Most often, this information is derived from ambient monitors from routine or special purpose monitoring networks. For example, most studies investigating the link between air pollution and human health have relied on monitoring data to characterize ambient pollutant concentrations (e.g. Bell et al., 2004; Samoli et al., 2005, and references therein). While monitored values can be considered as "ground truth" if one neglects the effect of measurement errors, assumptions about their spatial homogeneity or lack thereof need to be made in such studies. Frequently, in the absence of dense urban PM monitoring networks, values observed at a "central monitor" were considered to be representative for ambient pollutant levels within a metropolitan area (e.g. Bell et al., 2004; Samoli et al., 2005, and references therein). In addition, analyses of the long-term association between ambient PM levels and health outcomes are complicated by changes in monitoring locations,

* Corresponding author at: 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).

1352-2310/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2009.02.031

instrumentations, or monitoring priorities. For example, with the introduction of National Ambient Air Quality Standards (NAAQS) for annual average and 24-h average $PM_{2.5}$ concentrations in 1997, a widespread $PM_{2.5}$ monitoring network was established in the late 1990s but at the same time the number of PM_{10} monitors began to decline (Lall et al., 2004).

To potentially overcome some of the spatial and temporal data gaps of ambient monitoring networks, recent studies have begun to explore the potential benefits of incorporating concentration fields simulated by grid-based air quality modeling systems into health impact assessments (Marmur et al., 2004; Bell, 2006). Such an approach would be akin to the practice of creating meteorological reanalysis fields in which observations and dynamical models are combined in a framework that is aimed at providing a consistent long-term record of meteorological fields with good temporal and spatial coverage.

In this paper, we present results from a study aimed at performing long-term air quality simulation over the northeastern portion of the U.S. for the purpose of characterizing $PM_{2.5}$ concentrations for the time period from 1988 to 2005. We describe the design of the model simulations in Section 2, present evaluation results for these simulations in Section 3, introduce a method for combining observations and model predictions for improved $PM_{2.5}$ characterization in Section 4, and discuss strengths and weaknesses of the proposed approach in Section 5.

2. Database and model setup

2.1. Modeling system

2.1.1. Meteorological simulations

Meteorological conditions for the time period from January 1, 1988 to December 31, 2005 were simulated with the MM5 meteorological model (Grell et al., 1994). The simulations were performed on two nested grids with 36 km and 12 km horizontal grid spacing, respectively. Among the physics options chosen for the simulation are the ETA scheme for representing the planetary boundary layer, the Kain–Fritsch cloud scheme, and the Rapid Radiation Transfer Model (RRTM) radiation scheme. Throughout the model simulation, MM5 was nudged towards NCEP reanalysis fields using four-dimensional data assimilation.

2.1.2. Emissions processing

All emissions processing including mobile sources and biogenic sources was performed within the Sparse Matrix Operator Kernel Emissions (SMOKE) processing system (Houyoux et al., 2000). To date, there is no single source of anthropogenic emissions covering the entire period of this study from 1988 to 2005 at a level of detail sufficient to support photochemical modeling. Consequently, the emission inventories used in this study were compiled primarily from three different sources. First, for simulating the time period from 1988 to 2000, annual anthropogenic emission inventories for area and nonroad sources were obtained from the U.S. Environmental Protection Agency (U.S. EPA) National Emission Trends database for 1990 and 1996–2000 (U.S. EPA, 2005). The 1990 emissions were used for the simulation of 1988–1990, while emissions from 1991 to 1995 were estimated by interpolation between 1990 and 1996 for these source categories. Onroad mobile source emissions were estimated with the MOBILE6 model using annual county-level vehicle miles traveled (VMT) and MM5 temperatures from 1988 to 2000. For modeling 1988–2000 point sources, stack-level inventories for 1990, 1996, 1999 and 2001 were available. The 1990 inventory was used for 1988–1994, the 1996 inventory was used for 1995–1998, the 1999 inventory was used for 1999, and the 2001 inventory was used for 2000. Second, for simulating 2001, the anthropogenic emission inventories prepared for the EPA Clean Air Interstate Rule were utilized (http://www.epa.gov/CAIR/technical.html). Third, for the years from 2002 to 2005, we utilized the 2002 base case and 2009 future year emission inventories prepared for the Ozone Transport Commission (OTC) modeling platform (OTC, 2007). For the 2002 modeling, the 2002 base case emissions were used directly, and for the 2003–2005 modeling, annual emission and VMT inventories were estimated through linear interpolation between 2002 and 2009. Finally, for 2000–2005 modeling, available hourly point source emission data from Continuous Emission Monitors (CEM) were utilized instead of annual total emissions for all cases when CEM data could be matched against inventory sources within SMOKE. The same temporalization, spatial allocation, and speciation profiles were used to process the inventories for all years. Biogenic sources for the entire modeling period from 1988 to 2005 were estimated with the BEIS3.12 model taking into account MM5 temperature, radiation, and precipitation. Table 1 provides a summary of domain-total annual anthropogenic emissions for 1990 and 2000. It can be seen that NO_x emissions are dominated by mobile and point sources while VOC emissions are dominated by mobile and area sources and SO_2 emissions are dominated by point sources. It is also evident that significant emission reductions for all pollutants have occurred between 1990 and 2000.

2.1.3. Air quality modeling

Using the meteorological and emission fields described above, hourly gridded fields of concentrations, wet deposition, and dry deposition for about 80 gas-phase and aerosol species were simulated with the Community Multiscale Air Quality (CMAQ) model (Byun and Schere, 2006), version 4.5.1. The simulations were performed with two nested grids of 36 km and 12 km, corresponding to the MM5 grids except for a ring of buffer cells. The boundary conditions for the 36 km grid correspond to climatological values, while the boundary conditions for the 12 km grid were derived from the 36 km simulation. The height of the first model layer was approximately 40 m. Gas phase chemistry was represented by the CB-IV mechanism (Gery et al., 1989) while aerosol chemistry was simulated with the aero3 module. For all subsequent analyses, only simulations from the 12 km CMAQ grid were utilized.

2.2. Observations

For $PM_{2.5}$, there are a variety of distinct observational networks and measurement techniques. In this study, CMAQ predictions of $PM_{2.5}$ were compared against three types of observations. Because of differences in measurement techniques, instrumentation, and site location criteria, model performance was examined on a species-by-species and network-by-network basis. In addition, even though the focus of this model application is on estimating historic $PM_{2.5}$ concentrations, as a model performance screening analysis for time periods prior to the availability of $PM_{2.5}$ measurements, CMAQ-simulated PM_{10} concentrations were compared against filter-based PM_{10} measurements. All monitors

Table 1
Annual total anthropogenic emissions over the 12 km modeling domain processed by SMOKE for 1990 and 2000. All emissions are shown in kilotons.

		NO_x	VOC	CO	SO_2	NH_3	$PM_{2.5}$
Area + nonroad	1990	1933	3580	10,379	1018	502	790
	2000	1757	3353	11,627	691	513	661
Mobile	1990	2895	2877	31,924	67	42	77
	2000	2180	1461	18,190	65	68	41
Point	1990	3443	1209	7743	8725	15	172
	2000	1884	334	1190	5088	11	328

were assigned to the model grid cells in which the monitor was located. All model evaluation analyses for $PM_{2.5}$ were performed from 2000 to 2005 since earlier years did not have sufficient observational data coverage, while the comparison against PM_{10} observations was performed for 1988–2005.

First, CMAQ predictions of total $PM_{2.5}$ were compared against filter-based 24-h average $PM_{2.5}$ concentrations based on the Federal Reference Method (FRM) which were obtained from the EPA Air Quality System (AQS) for a total of 262 monitors located in the 12 km CMAQ modeling domain. Measurements at these sites typically are taken once every three days, however, some sites follow a once per day or once every six days sampling schedule. CMAQ predicted total $PM_{2.5}$ was calculated as the sum of all modeled Aitken- and accumulation-mode species except water following equation (2) in Appel et al. (2008).

Second, to analyze CMAQ predictions of the various chemical species contributing to the total $PM_{2.5}$ mass, filter-based 24-h average concentrations were obtained for 27 Chemical Speciation Network (CSN) monitors located in the 12 km CMAQ domain that had sufficient data coverage from 2000 to 2005. At most of these monitors, measurements are taken once every third or sixth day. The following species measured at the CSN monitors were compared to the corresponding speciated CMAQ aerosol components: sulfate (SO_4), nitrate (NO_3), ammonium (NH_4), elemental carbon (EC), and organic carbon (OC). For the calculation of CMAQ OC concentrations from the simulated concentrations of primary and secondary anthropogenic organic aerosol and secondary biogenic organic aerosol mass, we followed the approach described in equation (1) by Appel et al. (2008). Prior to analysis, OC measurements from the CSN network were blank corrected using monthly average site- and monitor-specific values calculated from blank values archived in AQS. In addition, for the evaluation of the CMAQ aerosol component A25 which represents all primary $PM_{2.5}$ mass not attributed to sulfate, nitrate, EC and OC emissions during inventory processing, we followed the Interagency Monitoring of Protected Visual Environments (IMPROVE) network definition of crustal material to calculate an observed "other $PM_{2.5}$" quantity as follows:

$$\text{Observed other } PM_{2.5} = 2.2 \times [Al] + 2.49 \times [Si] + 1.63 \times [Ca] + 2.42 \times [Fe] + 1.94 \times [Ti] \quad (1)$$

Third, speciated $PM_{2.5}$ measurements were also obtained from AQS for 13 IMPROVE monitors located in the 12 km CMAQ domain that had sufficient data coverage from 2000 to 2005. Similar to the CSN monitors, sampling typically occurs every third day. With the exception of ammonium, all of the species listed above for the CSN monitors are also available for the IMPROVE monitors, and we followed the same approach to match observed species to CMAQ-simulated aerosol concentrations. It should be noted, however, that the operational definition of OC and EC differs significantly between the two networks as described by Chow et al. (2001) and DeBell et al. (2006), resulting in different splits between OC and EC at co-located monitors.

In addition, CMAQ-simulated PM_{10} concentrations, defined as the sum of CMAQ $PM_{2.5}$, coarse soil PM, and coarse other PM, were compared against filter-based 24-h average PM_{10} measurements at 135 monitors extracted from the AQS database that were inside the 12 km CMAQ domain and had sufficient data coverage from 1988 to 2005. Finally, for the computation of $PM_{2.5}/PM_{10}$ ratios discussed in Section 4.3, measurements of 24-h average PM_{10} mass from filter-based monitors were obtained from the AQS database at a total of 38 sites that had both $PM_{2.5}$ and PM_{10} measurements for the five-year period from 2000 to 2004 and a sufficient long-term record of PM_{10} from 1988 to 2005.

3. Model evaluation

Fig. 1 shows the CMAQ-simulated composition of $PM_{2.5}$ mass over the entire simulation time period as stacked time series charts. For these time series, concentrations were averaged over all non-water grid cells. Fig. 1a shows daily concentrations while baseline concentrations estimated with the Kolmogorov–Zurbenko (KZ) iterated moving average filter (Rao et al., 1997) are shown in Fig. 1b. It can be seen that SO_4 is the major component of $PM_{2.5}$ in the modeling domain but that its importance has decreased over time as a result of the SO_2 emission reductions since the mid-1990s. There also is a clear seasonal variation in the composition of $PM_{2.5}$, with sulfate higher in summertime and nitrate and crustal/other primary $PM_{2.5}$ higher in wintertime. In addition, the time series exhibit significant interannual variations. While the accuracy of the simulated species time series shown in these figures cannot be evaluated due to the lack of detailed speciated $PM_{2.5}$ for most of the

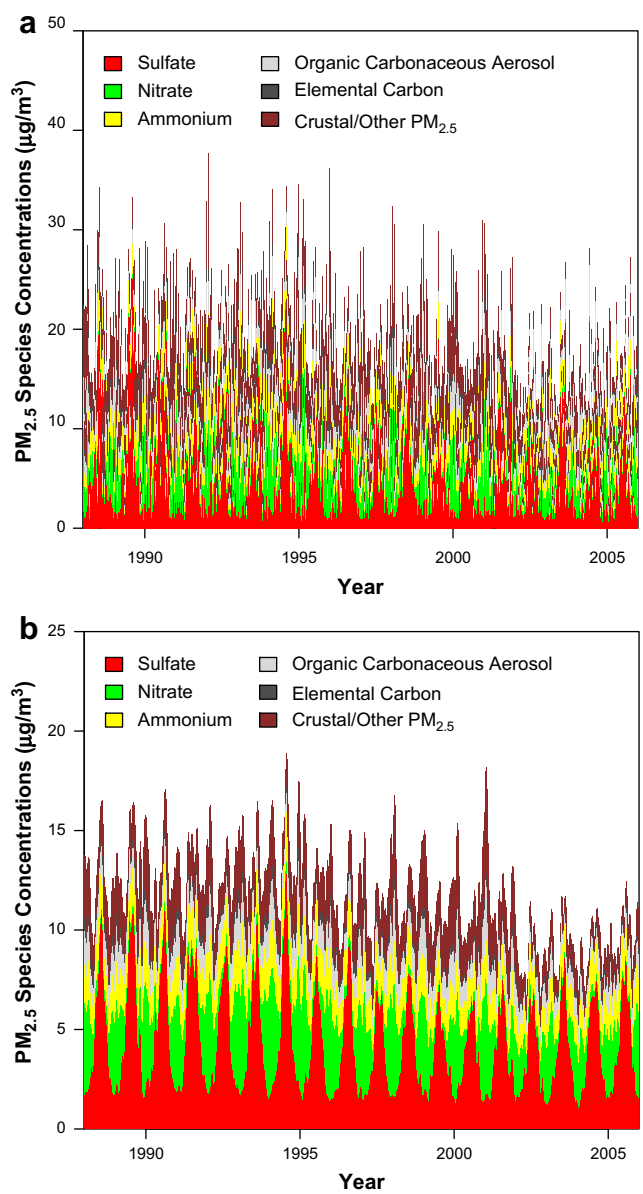

Fig. 1. CMAQ-simulated time series of total and speciated $PM_{2.5}$ concentrations averaged over all non-water grid cells in the modeling domain. a) Daily 24-h average concentrations, b) baseline of 24-h average concentrations estimated by applying a KZ(15,5) filter to the daily values.

simulation time period, as a screening analysis, CMAQ-simulated PM_{10} mass concentrations can be compared against filter-based measurements at AQS sites. Furthermore, for the last six years of the simulation period, measurements of total $PM_{2.5}$ mass are available at FRM monitoring sites. Finally, speciated observations from the IMPROVE and CSN network are also available for this time period.

First, as a screening analysis, CMAQ-simulated PM_{10} mass for 1988–2005 was compared against filter-based measurements of daily PM_{10} mass collected at 135 AQS sites every third or sixth day. The results of this analysis indicate that CMAQ PM_{10} concentrations tend to be lower than observations with a median bias of $-6.5\ \mu g\ m^{-3}$ across all sites. The median station-specific correlation coefficient between the observed and the simulated time series is 0.53 with 63% of all sites showing correlation coefficients greater than 0.5. While there are a number of factors that complicate the comparison of observed and simulated PM_{10} concentrations such as the siting of monitors in locations not representative of the grid-cell average simulated by the model or the loss of volatile components from the PM_{10} filters, and while it is difficult to relate model performance for PM_{10} to model performance for $PM_{2.5}$ due to variations in PM_{10} speciation, size distribution, and measurement techniques, this screening analysis nevertheless suggests that the modeling system exhibits skill in capturing fluctuations in PM concentrations over the entire simulation time period.

Second, we compared CMAQ predictions of total $PM_{2.5}$ to filter-based 24-h $PM_{2.5}$ mass measurements at 262 FRM monitors in the modeling domain for 2000–2005. At these sites, the station-specific bias values (calculated as predictions–observations) ranged from -5.4 to $+12.6\ \mu g\ m^{-3}$ with a median value of $-0.3\ \mu g\ m^{-3}$ while 82% of all stations exhibited bias values between -3 and $+3\ \mu g\ m^{-3}$. The station-specific RMSE values ranged from 4.6 to $18.6\ \mu g\ m^{-3}$ with a median value of $7.5\ \mu g\ m^{-3}$ while 92% of all stations exhibited RMSE values less than $10\ \mu g\ m^{-3}$. Station-specific correlation coefficients between the observed and the simulated 6-year time series ranged from 0.32 to 0.75 with a median value of 0.62 while 94% of all stations exhibited a correlation coefficient greater than 0.5. As discussed in Hogrefe et al. (2006), correlation coefficients between observed and simulated $PM_{2.5}$ concentrations vary by time scale. Since no spectral decomposition was performed in the present analysis, the correlation coefficients for the six-year time series analyzed here reflect the model's overall ability to capture fluctuations on the synoptic, seasonal, and longer-term time scales. Overall, these model evaluation statistics are consistent with other studies performed for shorter time periods (Boylan et al., 2006a,b; Eder and Yu, 2006; Hogrefe et al., 2007; Mathur et al., 2008; McKeen et al., 2007; Morris et al., 2005; Tesche et al., 2006; Yu et al., 2008). These statistics also appear to indicate better model performance for $PM_{2.5}$ compared to the PM_{10} results shown above. However, we caution that the analysis of $PM_{2.5}$ and PM_{10} described in this section was not carried out at co-located sites, therefore, differences in model performance could be at least partially attributable to differences in monitor locations.

Finally, to assess the contribution of the various $PM_{2.5}$ components to total $PM_{2.5}$ mass and to determine whether model performance varies by species as reported in other studies (Boylan et al., 2006a,b; Eder and Yu, 2006; Hogrefe et al., 2007; Morris et al., 2005; Tesche et al., 2006), Fig. 2 shows stacked bar charts of observed and CMAQ-simulated $PM_{2.5}$ species concentrations for each season, averaged over a) 13 IMPROVE sites and b) 27 CSN sites. For each season, all available observations and matching CMAQ predictions for the 2000–2005 time period were utilized, and seasons were defined as December–February for winter, March–May for spring, June–August for summer, and September–November for fall. Note that at the IMPROVE sites, no ammonium concentrations are measured. At both the IMPROVE and the CSN sites, CMAQ

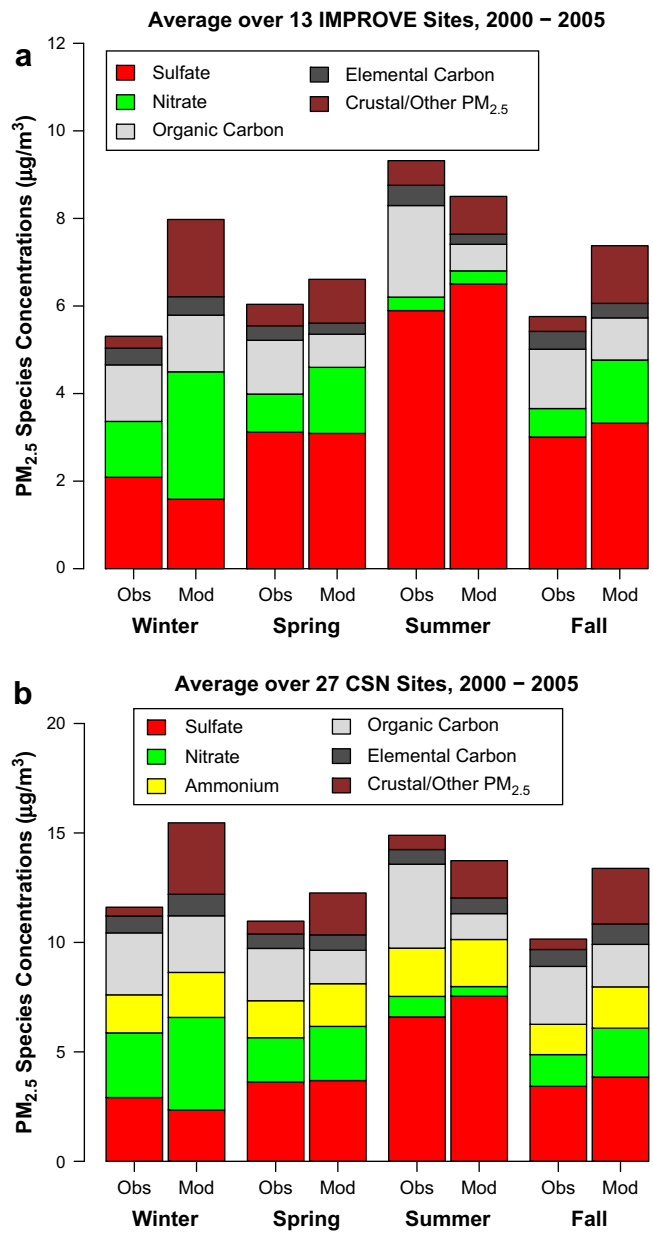

Fig. 2. Comparison of average observed (Obs) and CMAQ-simulated (Mod) $PM_{2.5}$ species concentrations for the time period January 1, 2000–December 31, 2005, calculated separately for each season. a) Averaged over 13 IMPROVE monitors located in the modeling domain, b) averaged over 27 CSN monitors located in the modeling domain.

underestimates sulfate during winter, is close to observed concentrations during spring, and tends to overestimate sulfate during summer and fall. Conversely, nitrate concentrations tend to be overestimated during winter, spring and fall at both the IMPROVE and the CSN sites and underestimated during summer at CSN sites. Concentrations of OC are higher at the CSN sites than the IMPROVE sites in both observations and CMAQ predictions. While the higher observed OC values could in part be due to differences in measurement methodologies (Chow et al., 2001; DeBell et al., 2006), the higher CMAQ values indicate additional contributions from primary sources or secondary formation in urban areas compared to rural locations. CMAQ is underestimating OC in all seasons at both IMPROVE and CSN monitors. The underestimation is most pronounced in summertime, consistent with previous studies indicating missing pathways for the formation of secondary organic

aerosol in CMAQ (Morris et al., 2006; Yu et al., 2007). Concentrations of EC are lower than OC during all seasons at all sites for both observations and CMAQ. For crustal matter, modeled as unspeciated primary $PM_{2.5}$ in CMAQ, predicted concentrations are consistently higher than observations at all sites and seasons, this overprediction is most pronounced during wintertime. It should be noted that, by definition, the comparison of the CMAQ unspeciated primary $PM_{2.5}$ component to speciated observations is challenging. In our analysis, we follow the framework described in U.S. EPA (2007) and relate this quantity to observed soil oxide concentrations. To the extent that primary unspeciated $PM_{2.5}$ concentrations originate from emission sources other than dust, the overestimation shown in Fig. 2 is partially expected but cannot be resolved in the absence of updated $PM_{2.5}$ emission speciation profiles.

4. Adjusting CMAQ-predicted $PM_{2.5}$ concentrations

4.1. Methodology

The results shown in Section 3 for total $PM_{2.5}$ mass reveal that CMAQ has some skill in simulating this parameter. Therefore, it appears justified to view these simulations as potentially valuable estimates of ambient concentrations when observations are not available. However, these results also reveal that total $PM_{2.5}$ predictions are subject to errors with respect to available observations, necessitating an approach for reducing these errors before using these fields for health impact studies. Moreover, good model performance for total $PM_{2.5}$ mass can be caused by compensating errors in $PM_{2.5}$ species predictions such as the tendency to overestimate sulfate and underestimate nitrate and OC during summertime. More generally, the model evaluation results derived from six years of model simulations and corresponding observations in this study confirmed findings from previous studies (e.g. Eder and Yu, 2006; Hogrefe et al., 2006; Mathur et al., 2008; Yu et al., 2008) that model performance for $PM_{2.5}$ varies by species, season, and land-use. Therefore, any approach to correct CMAQ errors when using gridded CMAQ fields for ambient air quality characterization should at least take these three factors into account. Consequently, we developed the following approach for adjusting CMAQ-simulated $PM_{2.5}$ concentrations:

- At each CSN and IMPROVE monitoring location, the observed time series of 24-h average sulfate, nitrate, OC, EC, and crustal material for 2000–2005 were matched against the corresponding CMAQ values.
- The baseline of these observed and simulated species concentrations was calculated by applying a KZ(15,5) moving average filter. As discussed in Rao et al. (1997), the repeated application of this filter results in daily estimates of a seasonally-varying time series despite measurements being available only every third or sixth day.
- The observed and simulated 6-year baseline time series were then averaged to compute an average-year seasonal cycle by averaging the six values for each day (January 1, January 2, etc.).
- Next, the ratio of observed to modeled average seasonal baseline values was calculated for each day of the year for each species for each monitor.
- These ratios (or "adjustment factors") were then interpolated in space with an inverse-distance-weighting interpolation scheme to generate a gridded map of adjustment factors for each day-of-year and each species. To account for the asymmetry between ratios corresponding to model over- vs. underpredictions (e.g. observed/modeled ratios of 0.5 and 2 both correspond to a model error of a factor of 2, but their arithmetic average is different from unity), the ratios were log-transformed prior to spatial interpolation and then re-transformed after interpolation.
- Finally, the raw CMAQ $PM_{2.5}$ species simulations for 1988–2005 were adjusted by multiplying each concentration value with the appropriate adjustment factor for the corresponding day-of-year, grid cell, and species. Since no adjustment factors could be calculated for ammonium because it is not measured at IMPROVE monitors, adjusted ammonium concentrations were estimated by assuming a constant degree of neutralization between the raw and the adjusted fields, calculating this degree of neutralization for the unadjusted fields following the method described in U.S. EPA (2007), and then calculating the adjusted ammonium concentrations based on this degree of neutralization and the adjusted sulfate and nitrate values.

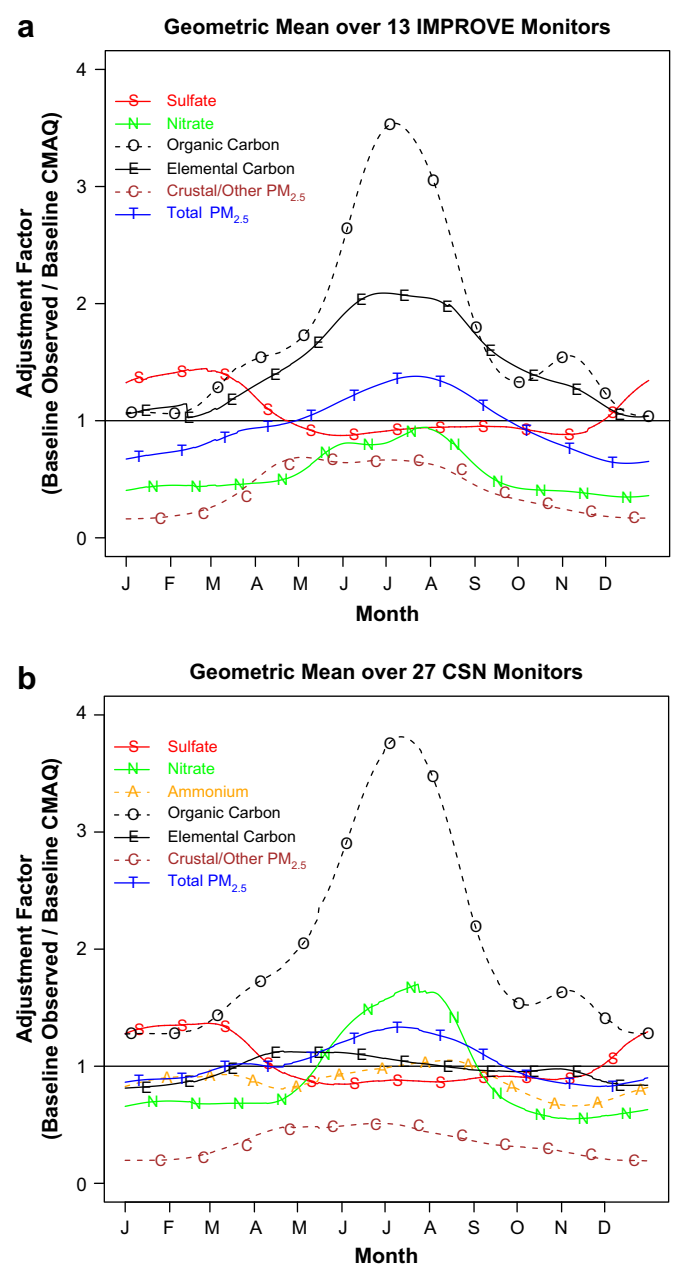

Fig. 3. Species-specific model adjustment factors as function of day-of-year. a) Averaged over 13 IMPROVE monitors located in the modeling domain, b) averaged over 27 CSN monitors located in the modeling domain.

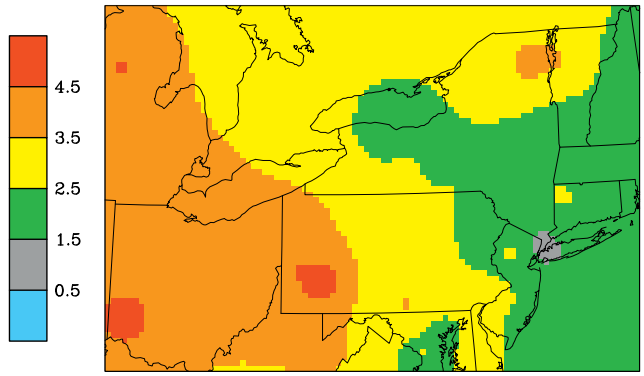

Fig. 4. Illustration of spatially interpolated adjustment factors. The figure displays a map of the gridded OC adjustment factor estimated for June 1 based on interpolating the adjustment factors calculated at the IMPROVE and CSN monitoring sites to the 12 km modeling grid through an inverse-distance-weighting scheme.

To illustrate the variation of these adjustment factors by day-of-year and species, Fig. 3a and b show seasonal time series of adjustment factors at IMPROVE and CSN monitors. For this display, the geometric mean of the adjustment factors for a given day and species over all monitors within each network was calculated. Overall, these adjustment factor curves are consistent with the discussion of the bar charts displayed in Fig. 2 but provide a more gradual seasonal variation compared to the grouping in fixed three-month seasonal time blocks. In particular, it is evident that sulfate is underestimated in the winter months and overestimated in the summer months, while OC is underestimated during all seasons, with the geometric mean underestimation ranging from a factor of 1.2 during wintertime at IMPROVE sites to a factor of 3.8 during summertime at CSN sites. To illustrate how these factors can vary in space, Fig. 4 displays a map of gridded interpolated adjustment factors for OC for June 1. It can be seen that these adjustment factors range from 1.5 to 2.5 for much of the New England region to more than 3.5 for parts of the western portion of the modeling domain. It should be noted that the interpolated gridded maps of adjustment factors incorporate adjustment factors derived at both CSN and IMPROVE monitors. Therefore, differences in monitoring protocols, especially for OC, between the two networks introduce ambiguity to these interpolated maps, and the resulting adjusted CMAQ OC values should be viewed only as a first-order adjustment that overcome some of the deficiencies in the current CMAQ version in terms of SOA formation. In other words, the differences in CMAQ OC underestimation for summertime between the IMPROVE and the

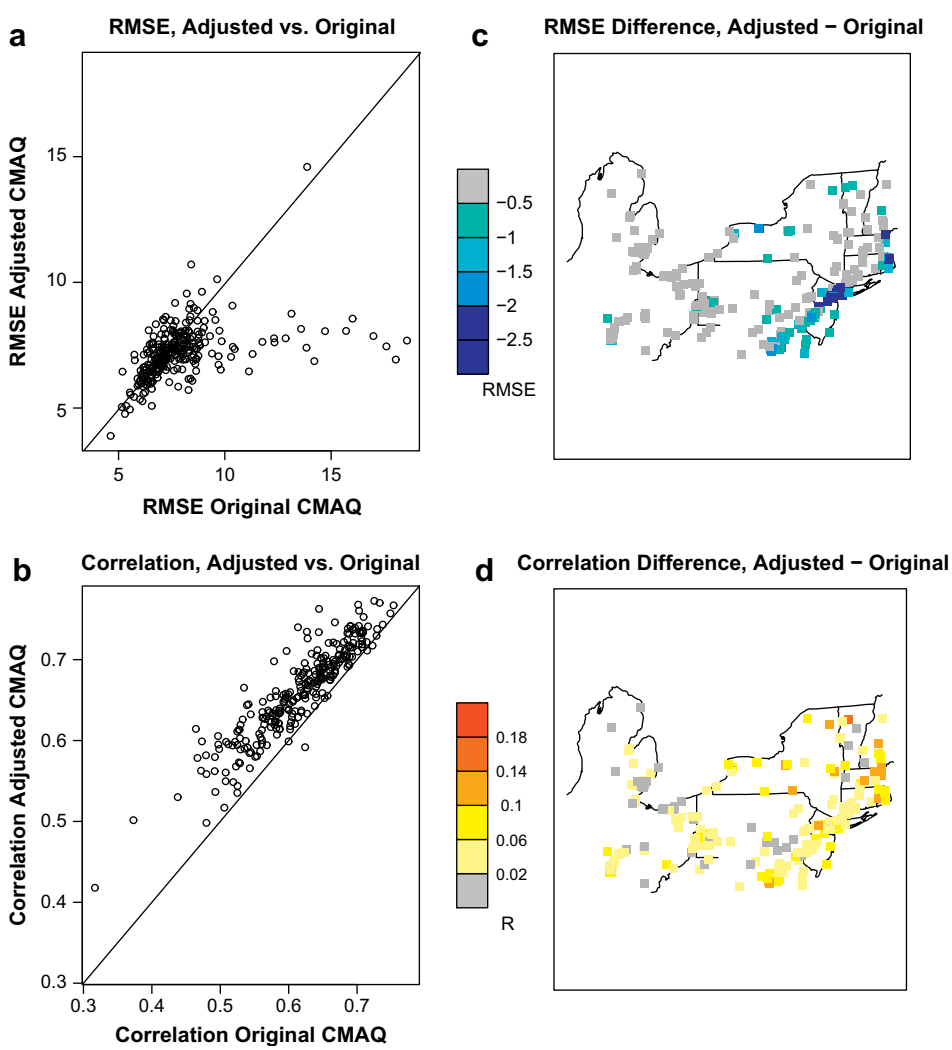

Fig. 5. a) Scatter plot of root mean square error (RMSE) of the adjusted CMAQ $PM_{2.5}$ mass values at 262 FRM sites for 2000–2005 vs. the RMSE values calculated for the original CMAQ simulations at the same sites. b) As in a) but for correlation coefficients. c) Spatial map of differences in RMSE of $PM_{2.5}$ mass at 262 FRM sites for 2000–2005 calculated as RMSE of adjusted CMAQ values minus RMSE of original CMAQ values. d) As in c) but for correlation coefficients.

CSN monitors (see Fig. 3a and b) may be in part due to differences in monitoring protocols, but both curves point to missing sources of OC in CMAQ and interpolating between all of these monitors to create a gridded map of adjustment factors should help to provide a better estimate of summertime OC in the adjusted CMAQ fields compared to the original fields.

4.2. Method evaluation

After the original model-predicted $PM_{2.5}$ species concentrations were adjusted using the methodology described above, total $PM_{2.5}$ mass concentrations were calculated from both the original and the adjusted species concentrations and compared against 24-h $PM_{2.5}$ mass measurements at FRM monitors taken between 2000 and 2005. The $PM_{2.5}$ mass measurements at these FRM sites are completely independent from the species measurements at the IMPROVE and CSN sites that were used to calculate the adjustment factors. Results of the comparison of original and adjusted CMAQ $PM_{2.5}$ mass against observations are shown in Fig. 5. The left two panels display scatter plots of RMSE and correlation coefficients computed for original and adjusted CMAQ $PM_{2.5}$ mass. For these plots, the metrics were computed separately at 262 FRM sites and each point in the scatter plot corresponds to a specific site. In the case of RMSE, almost all points are below the 1:1 line indicating that at most of the sites, the adjusted CMAQ $PM_{2.5}$ concentrations have a lower RMSE than the original $PM_{2.5}$ concentrations. Likewise, almost all of the points in the correlation scatter plot are above the 1:1 line, indicating that correlations between time series of adjusted CMAQ $PM_{2.5}$ concentrations and observations were higher than those between time series of the original CMAQ $PM_{2.5}$ concentrations and observations. In order to examine whether there are regions in which the adjustment was more beneficial than others, the right two panels in Fig. 5 provide maps of differences in RMSE and correlation coefficients. More negative values in Fig. 5c indicate a larger reduction in RMSE resulting from the adjustment, while more positive values in Fig. 5d indicate a larger increase in correlation coefficients resulting from the adjustment. It can be seen that the adjustment was most effective in reducing RMSE in the urban corridor from Washington, D.C. to New York City.

4.3. Comparison against an alternate method to estimate historic $PM_{2.5}$ concentrations

The results shown above indicate that for the 2000–2005 time period, the adjustment approach based on observed CSN and IMPROVE speciation data improved CMAQ predictions for total $PM_{2.5}$ mass at FRM locations, thereby also enhancing the confidence in the use of adjusted CMAQ $PM_{2.5}$ concentrations in unmonitored areas and on days without measurements. Unfortunately, because of the lack of widespread $PM_{2.5}$ monitoring before ca. 2000, the quality of the adjusted CMAQ $PM_{2.5}$ concentrations for the 1988–1999 time period cannot be evaluated. However, it is possible to compare these estimated historical $PM_{2.5}$ values against values estimated with an alternate methodology proposed by Lall et al. (2004). This method was based on establishing $PM_{2.5}/PM_{10}$ ratios at co-located monitors for time periods when both pollutants were measured and then applying these ratios to historic measurements of PM_{10} for time periods when no $PM_{2.5}$ observations were available. In our study, we applied this methodology as follows: First, we determined all stations that had filter-based monitors for both $PM_{2.5}$ and PM_{10} for 2000–2004 and PM_{10} measurements for 1988–2005. This resulted in the selection of 38 monitors in the modeling domain. Next, at each of these sites the average PM_{10} and $PM_{2.5}$ concentrations were calculated over all days for which both measurements were available during 2000–2004, and the ratio of these averages was determined. A histogram of these ratios across the 38 sites is displayed in Fig. 6 and reveals that the ratio typically ranges from 0.5 to 0.8, consistent with the values reported in Lall et al. (2004). Finally, these station-specific ratios were applied to the 1988–2005 PM_{10} measurements at the corresponding sites to estimate historic $PM_{2.5}$ values.

Fig. 7a–d shows a comparison of the estimated 1988–2005 $PM_{2.5}$ values from the adjusted CMAQ simulations and from the $PM_{2.5}/PM_{10}$ ratio method. The scatter plot in panel a) is constructed from daily estimated $PM_{2.5}$ concentrations for all 38 sites at which $PM_{2.5}/PM_{10}$ ratios were available and for all days during 1988–2005 for which PM_{10} measurements were available, resulting in a total of 58,722 data pairs. Out of these pairs, 52% agree with each other within a factor of 1.5, and 75% agree with each within a factor of 2. The overall correlation coefficient is 0.50 and is influenced by both temporal fluctuations and spatial gradients between monitoring locations in these two datasets. Panels b)–d) show histograms of station-specific comparison metrics between the two datasets. First, differences in station-specific average $PM_{2.5}$ estimates derived from the two approaches are displayed in panel b). This histogram indicates that at most sites, long-term mean $PM_{2.5}$ concentrations derived from the two methods agree to within a few $\mu g\ m^{-3}$, with a tendency for the adjusted CMAQ $PM_{2.5}$ concentrations to be slightly lower than the $PM_{2.5}$ concentrations derived from the ratio method. Panel c), which displays root mean square differences (RMSD) between the two estimates, shows that these differences are less than 10 $\mu g\ m^{-3}$ at most sites. In conjunction with Panel b), this indicates that while on average both methods give similar estimates, the values on individual days could be significantly different. This is also evident from the scatter plots in Panel a). Finally, the histogram of correlation coefficients in Panel d) confirms that at most sites, the correlation coefficients between the time series of estimated $PM_{2.5}$ concentrations from both methods range between 0.5 and 0.7, again indicating that on a day-to-day basis, differences between estimates from the two methods can be significant but also indicating that there is an overall agreement in the broader temporal fluctuations of the estimates derived from both methods.

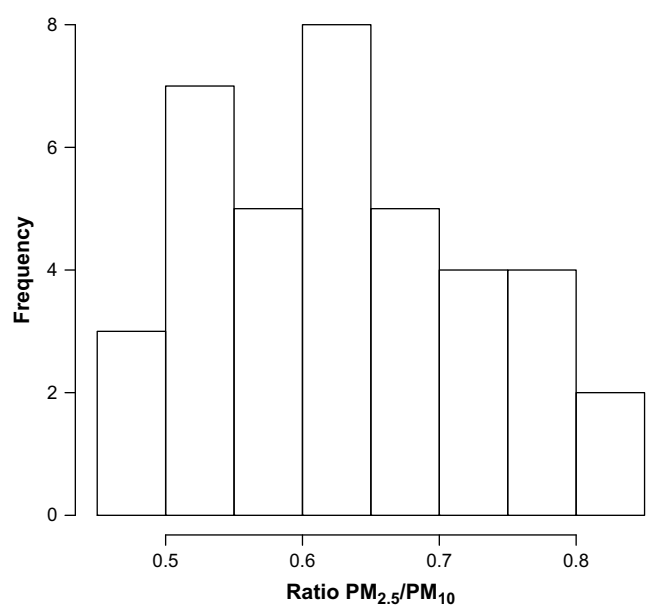

Fig. 6. Histogram of ratios of average PM_{10} and $PM_{2.5}$ concentrations across 38 co-located sites calculated over all days for which both PM_{10} and $PM_{2.5}$ measurements were available.

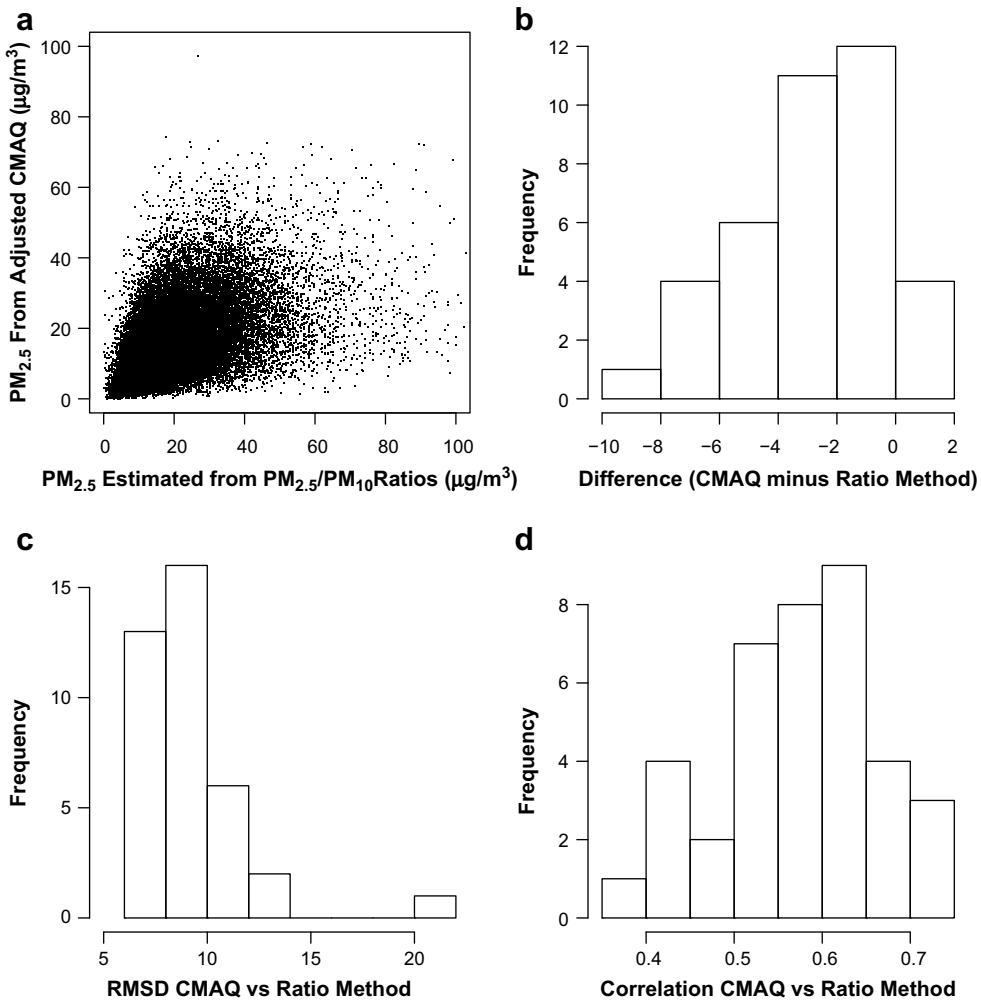

Fig. 7. a) Scatter plot of 24-h average $PM_{2.5}$ concentrations derived from the adjusted CMAQ $PM_{2.5}$ mass values vs. 24-h average $PM_{2.5}$ concentrations estimated from the $PM_{2.5}/PM_{10}$ ratio method. The comparison was performed for 1988–2005 at the location of the 38 sites where co-located $PM_{2.5}$ and PM_{10} measurements were available for 2000–2004 and each point in the scatter plot represents a single daily value at a single site. b) Histogram of the absolute difference of 1988–2005 average estimated $PM_{2.5}$ concentrations between the adjusted CMAQ method and the $PM_{2.5}/PM_{10}$ ratio method across the 38 monitors. c) As in b) but for root mean square differences (RMSD) between the two methods. d) As in b) but for correlation coefficients between the site-specific daily $PM_{2.5}$ time series estimated by the two methods.

Overall, these results reveal that the $PM_{2.5}$ concentrations from the adjusted CMAQ simulations are broadly consistent with an alternate approach to estimating historic $PM_{2.5}$ concentrations but overcome some of the deficiencies associated with this alternate approach, i.e. its limitation in terms of spatial coverage (co-located $PM_{2.5}$ and PM_{10} monitors need to be available for a period sufficiently long to establish reliable $PM_{2.5}/PM_{10}$ ratios) and temporal coverage (historical filter-based PM_{10} measurements are typically available only every 3rd or 6th day). In addition, the ratio method is subject to uncertainties introduced by possible sampling artifacts (the loss of semi-volatile components from the filters), seasonal and interannual variability in the $PM_{2.5}/PM_{10}$ ratios, and trends in species composition, while the CMAQ estimates are affected by uncertainties in historic emission inventories and aerosol model formulation. Therefore, both approaches should be viewed as complementary.

5. Discussion

The methodology introduced and evaluated in the previous sections provides a means to estimate daily concentrations of $PM_{2.5}$ total mass and species for time periods and locations when no monitoring data is available. This data record can be potentially useful for applications such as health impact assessments when long-term records of health outcomes such as mortality or morbidity are matched against ambient pollution concentrations, and for studying interactions between climate variability and air quality. However, there are several limitations of the methodology that have to be kept in mind when conducting such analyses:

- As shown in Section 3, CMAQ model performance varies by species and tends to be best for sulfate and ammonium and worst for nitrate and OC. Therefore, determining accurate adjustment factors is most critical for these species. Unfortunately, these are also the species for which the spatial interpolation of adjustment factors determined at individual CSN and IMPROVE monitors is subject to the highest uncertainty because of the spatial representativeness of such observations and because of differences in measurement methods for OC between these networks.
- The species-by-species model performance evaluation presented in Section 3 was performed for the time period from 2000 to 2005. Subsequently, the adjustment factors were also calculated based on this time period and were then applied to

the 1988–2005 CMAQ simulations. To the extent that there were either trends or interannual variations in species composition over this time period, the use of adjustment factors that only vary by day-of-year introduces additional uncertainty to the estimated adjusted CMAQ $PM_{2.5}$ values.

- While the adjusted CMAQ fields overcome some of the errors inherent in the original model simulations, they are limited by the spatial resolution of the model simulations, in this case 12 km. Therefore, they cannot provide improved information on local-scale phenomena such as elevated EC concentrations near roadways that may be of interest in exposure studies. To this extent, it might be useful to develop approaches that combine the adjusted CMAQ fields with other methods such as land-use-based regression approaches that attempt to resolve sub-grid-scale phenomena. Alternatively, raw or adjusted CMAQ fields could be used to provide boundary conditions to smaller-scale dispersion models such as AERMOD to create a hybrid modeling system as described in Isakov and Özkaynak (2007).

6. Summary

In this study, a methodology is introduced to estimate gridded fields of total and speciated $PM_{2.5}$ concentrations for time periods and regions not covered by observational data. To illustrate this methodology, we present an application in which year-round simulations with the MM5/CMAQ modeling system have been performed over the northeastern United States for 1988–2005. Model evaluation results for total $PM_{2.5}$ mass and individual species for the time period from 2000 to 2005 show that model performance varies by species, season, and location. Broadly consistent with other studies, model performance was generally found to be best for sulfate and ammonium and worst for nitrate and OC. Sulfate concentrations tend to be overestimated during summertime and underestimated during wintertime while the opposite is true for nitrate. OC concentrations are substantially underestimated during summertime at almost all sites. Based on these evaluation results, an approach is developed to adjust CMAQ $PM_{2.5}$ simulations with factors that vary by day-of-year, species, and location. The adjusted model values for total $PM_{2.5}$ mass are compared against independent measurements not utilized for the adjustment approach. This comparison reveals that the adjusted model values have a lower root mean square error (RMSE) and higher correlation coefficients than the original model values. Furthermore, the $PM_{2.5}$ estimates from these adjusted model values are compared against an alternate method for estimating historic $PM_{2.5}$ values that is based on $PM_{2.5}/PM_{10}$ ratios calculated at co-located monitors. Results reveal that both methods yield estimates of historic $PM_{2.5}$ mass that are broadly consistent; however, the adjusted CMAQ values provide greater spatial coverage and information for $PM_{2.5}$ species in addition to total $PM_{2.5}$ mass. Limitations of the proposed approach include assumptions about the temporal stability of adjustment factors for the historic time period, questions about the validity of the spatial interpolation of adjustment factors estimated at individual monitors, and the inability of the method to resolve sub-grid-scale variations in $PM_{2.5}$ concentrations that might be of concern especially for primary $PM_{2.5}$ species in urban areas. In the future, the daily gridded fields of ambient $PM_{2.5}$ concentrations constructed from observations and CMAQ simulations described in this study will be used to study the link between air quality and health. These results then will be compared to health impact studies that relied solely on ambient measurements of air pollution to assess the utility of long-term CMAQ simulations in enhancing the characterization of ambient air quality for health impact studies. Furthermore, future work should also be directed at developing methodologies that take into account uncertainties in the estimated air pollution surfaces during the health analysis. Finally, the eventual goal of such research would be to create a framework in which the impact of emission control programs is quantified both in terms of changes in ambient pollutant concentrations and in terms of changes in health outcomes through the integrated use of observations and photochemical models.

Acknowledgments

The work presented here was supported by the National Oceanic and Atmospheric Administration under award NAO40AR4310185185, but it has not been subjected to its required peer and policy review. Therefore, the statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the sponsoring agency and no official endorsement should be inferred. Additionally, part of the work presented in this paper was performed by the New York State Department of Environmental Conservation, but the views expressed in this paper do not necessarily reflect the views or policies of the New York State Department of Environmental Conservation. Finally, Christian Hogrefe also gratefully acknowledges partial support for this work through a research fellowship from the Oak Ridge Institute for Science and Education (ORISE).

References

Appel, K.W., Bhave, P.V., Gilliland, A.B., Sarwar, G., Roselle, S.J., 2008. Evaluation of the community multiscale air quality (CMAQ) model version 4.5: sensitivities impacting model performance; part II—particulate matter. Atmos. Environ. 42, 6057–6066.

Bell, M.L., McDermott, A., Zeger, S.L., Samet, J.M., Dominici, F., 2004. Ozone and short-term mortality in 95 US urban communities, 1987–2000. J. Am. Med. Assoc. 292, 2372–2378.

Bell, M.L., 2006. The use of ambient air quality modeling to estimate individual and population exposure for human health research: a case study of ozone in the Northern Georgia Region of the United States. Environ. Int. 32, 586–593. doi:10.1016/j.envint.2006.01.005.

Boylan, J.W., Odman, M.T., Wilkinson, J.G., Russell, A.G., 2006a. Integrated assessment modeling of atmospheric pollutants in the Southern Appalachian Mountains: part II. Fine particulate matter and visibility. J. Air Waste Manag. Assoc. 56, 12–22.

Boylan, J.W., Odman, M.T., Wilkinson, J.G., Russell, A.G., 2006b. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. Atmos. Environ. 40, 4946–4959.

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

Chow, J.C., Watson, J.G., Crow, D., Lowenthal, D.H., Merrifield, T., 2001. Comparison of IMPROVE and NIOSH carbon measurements. Aerosol Sci. Technol. 34, 23–34.

DeBell, L., Gebhart, K., Hand, J., Malm, W.C., Pitchford, M.L., Schichtel, B.A., White, W.H., 2006. Spatial and seasonal patterns and temporal variability of haze and its constituents in the United States, report IV. ISSN: 0737-5352-74. Available online at: http://vista.cira.colostate.edu/Improve/Publications/Reports/2006/2006.htm.

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

Gégo, E., Porter, P.S., Gilliland, A., Rao, S.T., 2007. Observation-based assessment of the impact of nitrogen oxides emissions reductions on ozone air quality over the Eastern United States. J. Appl. Meteorol. Climatol. 46 (7), 994–1008.

Gery, M.W., Whitten, G.Z., Killus, J.P., Dodge, M.C., 1989. A photochemical kinetics mechanism for urban and regional scale computer modeling. J. Geophys. Res. 94, 12925–12956.

Grell, G.A., Dudhia, J., Stauffer, D., 1994. A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Technical Note, NCAR/TN-398+STR.

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

Hogrefe, C., Hao, W., Civerolo, K., Ku, J.-Y., Gaza, R.S., Sedefian, L., Sistla, G., Schere, K., Gilliland, A., Mathur, R., 2007. Daily photochemical simulations of ozone and fine particulates over New York State: findings and challenges. J. Appl. Meteorol. 46, 961–979.

Houyoux, M.R., Vukovich, J.M., Coats Jr., C.J., Wheeler, N.J.M., Kasibhatla, P., 2000. Emission inventory development and processing for the seasonal model for regional air quality. J. Geophys. Res. 105, 9079–9090.

Isakov, V., Özkaynak, H., 2007. A modeling methodology to support evaluation of public health impacts of air pollution reduction programs. In: Proceedings of the 29th NATO/SPS International Technical Meeting on Air Pollution Modeling and its Application, September 24–28, 2007, Aveiro, Portugal, pp. 598–605.

Lall, R., Kendall, M., Ito, K., Thurston, G.D., 2004. Estimation of historical annual $PM_{2.5}$ exposures for health effects assessment. Atmos. Environ. 38, 5217–5226. doi:10.1016/j.atmosenv.2004.01.053.

Marmur, A., Russell, A.G., Mulholland, J.A., 2004. Air quality modeling of $PM_{2.5}$ mass and composition in Atlanta: results from a two-year simulation and implications for use in health studies. In: Proceedings of the 3rd Annual CMAS Models-3 Users' Conference, 18–20 October, 2004. Chapel Hill, NC. Available online at: http://www.cmascenter.org/conference/2004/abstracts/Air%20Quality%20Management/marmur_abstract.pdf.

Mathur, R., Yu, S., Kang, D., Schere, K.L., 2008. Assessment of the wintertime performance of developmental particulate matter forecasts with the Eta-Community Multiscale Air Quality modeling system. J. Geophys. Res. 113. doi:10.1029/2007JD008580.

McKeen, S., Chung, S.H., Wilczak, J., Grell, G., Djalalova, I., Peckham, S., Gong, W., Bouchet, V., Moffet, R., Tang, Y., Carmichael, G.R., Mathur, R., Yu, S., 2007. Evaluation of several $PM_{2.5}$ forecast models using data collected during the ICARTT/NEAQS 2004 field study. J. Geophys. Res. 112 (D10S20).

Morris, R.E., McNally, D.E., Tesche, T.W., Tonnesen, G., Boylan, J.W., Brewer, P., 2005. Preliminary evaluation of the Community Multiscale Air Quality Model for 2002 over the Southeastern United States. J. Air Waste Manag. Assoc. 55, 1694–1708.

Morris, R.E., Koo, B., Guenther, A., Yarwood, G., McNally, D., Tesche, T.W., Tonnesen, G., Boylan, J., Brewer, P., 2006. Model sensitivity evaluation for organic carbon using two multi-pollutant air quality models that simulate regional haze in the Southeastern United States. Atmos. Environ. 40 (26), 4960–4972.

Ozone Transport Commission, 2007. Draft Modeling Technical Support Document. Available online at: http://www.otcair.org/projects_details.asp?FID=101&fview=modeling#.

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. Bull. Am. Meteorol. Soc. 78, 2153–2166.

Samoli, E., Analitis, A., Touloumi, G., Schwartz, J., Anderson, H.R., Sunyer, J., 2005. Estimating the exposure–response relationships between particulate matter and mortality within the APHEA multicity project. Environ. Health Perspect. 113, 88–95.

Tesche, T.W., Morris, R., Tonnesen, G., McNally, D., Boylan, J., Brewer, P., 2006. CMAQ/CAMx annual 2002 performance evaluation over the Eastern US. Atmos. Environ. 40, 4906–4919.

U.S. EPA, 2005. Criteria Pollutant Emissions Summary Files Extracted from the National Emissions Inventory (NEI) Database. Available online at: http://www.epa.gov/ttn/chief/net/critsummary.html.

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. Environmental Protection Agency, Research Triangle Park, NC, EPA-454/B-07-002, 262 pp.

Yu, S., Bhave, P.V., Dennis, R.L., Mathur, R., 2007. Seasonal and regional variations of primary and secondary organic aerosols over the continental United States: observation-based estimates and model evaluation. Environ. Sci. Technol. 41, 4690–4697.

Yu, S., Mathur, R., Schere, K., Kang, D., Pleim, J., Young, J., Tong, D., Pouliot, G., McKeen, S.A., Rao, S.T., 2008. Evaluation of real-time $PM_{2.5}$ forecasts and process analysis for $PM_{2.5}$ formation over the Eastern United States using the Eta-CMAQ forecast model during the 2004 ICARTT study. J. Geophys. Res. 113, D06204. doi:10.1029/2007JD009226.