

Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component

2. Model evaluation

Michelle R. Mebust

National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, USA

Brian K. Eder,¹ Francis S. Binkowski,¹ and Shawn J. Roselle¹

Air Resources Laboratory, National Oceanic and Atmospheric Administration, Research Triangle Park, North Carolina, USA

Received 18 October 2001; revised 14 March 2002; accepted 22 May 2002; published 26 March 2003.

[1] An initial evaluation of the Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component reveals CMAQ's varying ability to simulate observed visibility indices and aerosol species concentrations. The visibility evaluation, using National Weather Service observations from 139 airports for 11–15 July 1995, shows that CMAQ reasonably captured the general spatial and temporal patterns of visibility degradation, including major gradients, maxima and minima. However, CMAQ's two visibility prediction methods, Mie theory approximation and mass reconstruction, both underpredict visibility degradation (i.e., overpredict visibility). The mean bias, normalized mean bias (NMB), mean error and normalized mean error (NME) for the Mie calculations are -5.9 dv, -21.7% , 7.0 dv and 25.4% , respectively. For the reconstruction simulations, these statistics are -9.8 dv, -35.5% , 10.0 dv and 36.2% , respectively. Most simulated values ($\sim 90\%$ Mie and $\sim 85\%$ reconstruction) fall within a factor of two of the observations, although $r^2 = 0.25$ (Mie) and $r^2 = 0.24$ (reconstruction). The speciated aerosol evaluation uses observations of sulfate, nitrate, $PM_{2.5}$, PM_{10} and organic carbon obtained from 18 stations of the Interagency Monitoring of Protected Visual Environments (IMPROVE) network in June 1995. This evaluation reveals that, with the exception of sulfate (mean bias: $0.15 \mu\text{g}/\text{m}^3$, NMB: 3.1%), the model consistently underpredicts aerosol concentrations of nitrate ($-0.10 \mu\text{g}/\text{m}^3$, -33.1%), $PM_{2.5}$ ($-3.9 \mu\text{g}/\text{m}^3$, -30.1%), PM_{10} ($-5.66 \mu\text{g}/\text{m}^3$, -29.2%) and organic carbon ($-0.78 \mu\text{g}/\text{m}^3$, -33.7%). Sulfate was simulated best by the model ($r^2 = 0.63$, mean error = $1.75 \mu\text{g}/\text{m}^3$, NME = 36.2%), followed by $PM_{2.5}$ (0.55 , $5.00 \mu\text{g}/\text{m}^3$, 38.5%), organic carbon (0.25 , $0.94 \mu\text{g}/\text{m}^3$, 40.6%), PM_{10} (0.13 , $9.85 \mu\text{g}/\text{m}^3$, 50.8%) and nitrate (0.01 , $0.33 \mu\text{g}/\text{m}^3$, 104.3%). Except for nitrate, 75–80% of simulated concentrations fall within a factor of two of the IMPROVE observations.

INDEX TERMS: 0305 Atmospheric Composition and Structure: Aerosols and particles (0345, 4801); 0345 Atmospheric Composition and Structure: Pollution—urban and regional (0305); 0365 Atmospheric Composition and Structure: Troposphere—composition and chemistry; 0368 Atmospheric Composition and Structure: Troposphere—constituent transport and chemistry;

KEYWORDS: Models-3/CMAQ, PM, model evaluation, air quality modeling, visibility, aerosol species

Citation: Mebust, M. R., B. K. Eder, F. S. Binkowski, and S. J. Roselle, Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component, 2, Model evaluation, *J. Geophys. Res.*, 108(D6), 4184, doi:10.1029/2001JD001410, 2003.

1. Introduction

[2] Ambient air concentrations of particulate matter (atmospheric suspensions of solid or liquid materials, i.e., aerosols) continue to be a major concern for the U.S.

¹On assignment to National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, USA.

Environmental Protection Agency (EPA). High particulate matter (PM) concentrations are associated not only with adverse human health effects, including increased morbidity and mortality arising from altered respiratory and cardiovascular function, but they also contribute to acid precipitation, regional climate change and visibility degradation [U.S. Environmental Protection Agency, 1995, 2000b].

[3] The Clean Air Act and its Amendments require the EPA to establish National Ambient Air Quality Standards (NAAQS) for PM and to assess current and future air quality regulations designed to protect human health and

welfare. In addition, the Regional Haze Rule (Regional Haze Regulation: Final Rule, 64 Federal Register 126, pp. 35,714–35,774, 1 July 1999) requires States to “establish goals and emission reduction strategies for improving visibility” in designated national parks and wilderness areas. Reliable tools for performing these tasks are air quality models, such as the EPA’s Models-3 Community Multiscale Air Quality model (CMAQ) [Byun and Ching, 1999]. CMAQ simulates air concentrations and deposition of various pollutants (including PM), and also calculates visibility indices. These model simulations, which can be conducted on numerous spatial and temporal scales, support both regulatory assessment by EPA Program Offices, as well as scientific studies by research institutions.

[4] Within CMAQ is an aerosol component, or module, designed to simulate the complex processes involving PM. PM is commonly separated into PM_{2.5} (particles with aerodynamic diameters $\leq 2.5 \mu\text{m}$), also known as fine particles, and PM₁₀ (aerodynamic diameters $\leq 10 \mu\text{m}$). Note that PM₁₀ includes PM_{2.5}; particles with aerodynamic diameters between 2.5 and 10 μm are called coarse particles. However, the distinction between fine and coarse particles depends not just on diameter size, but also on the mode of particle generation. Fine particles result from combustion processes and gas-to-particle conversion; coarse particles are generated from wind-driven mechanical disruption.

[5] PM_{2.5} and PM₁₀ are not single entities, but consist of varying mixtures of chemical species, each having its own emission, transport and deposition characteristics. PM mixtures contain primary emissions from industry, power plants, transportation and biogenic sources, as well as secondary particles formed by gas-to-particle conversion. Aerosol species considered within the CMAQ aerosol module include sulfate, nitrate, ammonium, water, primary organic aerosols, secondary organic aerosols of both anthropogenic and biogenic origin, elemental carbon, primary material not otherwise specified and wind-blown dust [Binkowski and Roselle, 2003]. Using calculated concentrations of the above species, the aerosol module determines visibility degradation by two different methods: (1) an approximation of Mie extinction of visible light and (2) an empirical, mass reconstruction approach.

[6] CMAQ, like all models, requires evaluation against observational data to determine its value to the air quality and regulatory communities. One such evaluation [Arnold and Dennis, 2001] assessed CMAQ’s capability for predicting July 1995 ozone concentrations, as well as ozone and NO_x mixing ratios, for the eastern United States. The authors judged model performance using several endpoints over a variety of domain scales and using different chemical mechanisms. Broadly speaking, the model performed well on days with high observed ozone concentrations, but performed less well when ozone observations were lower and when large differences existed between mean and maximum ozone concentrations. Arnold and Dennis [2001] propose that this disparity in model performance is likely due to the model’s inability to correctly predict NO_x concentrations on the low ozone days, a theory borne out by their preliminary investigations. The reader is referred to Arnold and Dennis [2001] for further details.

[7] Another evaluation is currently being conducted for June 1990 and June 1995 air concentrations of sulfate, nitrate

and SO₂, as well as wet deposition of sulfate (R. Dennis, personal communication, 2001). Preliminary analyses show that CMAQ performed reasonably well; model performance was very good for sulfate, nitrate predictions were biased high by about 50%, SO₂ predictions were biased high by a factor of 2 and wet sulfate predictions exhibited no bias.

[8] Unfortunately, evaluating model performance with respect to PM is currently difficult, due to the lack of adequate data for comparison. However, two data sets exist that permit an initial evaluation of the CMAQ aerosol module. The first consists of National Weather Service (NWS) hourly visibility observations. In a preliminary model evaluation [Eder *et al.*, 2000], NWS visibility observations, covering the eastern half of the United States during a widespread regional haze episode from 11–15 July 1995, were compared against CMAQ visibility parameters calculated using Mie theory. This comparison revealed general agreement between modeled and observed values. The second PM data set contains speciated aerosol measurements gathered by the Interagency Monitoring of Protected Visual Environments (IMPROVE) network.

[9] The present evaluation extends that of Eder *et al.* [2000] by comparing IMPROVE observations collected for the eastern United States during June 1995 against CMAQ simulation results, and also by evaluating the mass reconstruction method for calculating July 1995 visibility parameters.

2. CMAQ Model

2.1. General Description

[10] The CMAQ Eulerian grid model simulates concurrently the atmospheric and land processes affecting the transport, transformation, and deposition of air pollutants and their precursors, on both regional and urban scales. CMAQ performs these simulations following first principles and employs a “one atmosphere” philosophy that tackles the complex interactions not only among multiple atmospheric pollutants, but also between regional and urban scales. Pollutants considered within CMAQ include tropospheric ozone, particulate matter, airborne toxics and acidic and nutrient species. The model also calculates visibility parameters. A detailed discussion of the CMAQ modeling system appears in the work of Byun and Ching [1999].

2.2. CMAQ Aerosol Component Overview

[11] The CMAQ aerosol component, described by Binkowski and Roselle [2003] and Binkowski [1999], was derived from the Regional Particulate Model (RPM) [Binkowski and Shankar, 1995], which is an extension of the Regional Acid Deposition Model (RADM) [Chang *et al.*, 1990]. Particle size distributions within CMAQ are represented as the superposition of three lognormal subdistributions, or modes.

[12] Two different modes, the Aitken and accumulation modes, each having variable standard deviations, represent PM_{2.5} particles in the CMAQ aerosol component. Aitken mode particles are those in the mass distribution with diameters smaller than about 0.1 μm . Accumulation mode particle diameters range between 0.1 and 2.5 μm . Conceptually, the Aitken mode represents fresh particles resulting from nucleation and/or direct emission. The accumulation

mode denotes aged particles. Each mode receives primary emitted material, is subject to wet and dry deposition and may grow from condensation of gaseous precursors. The two modes interact through coagulation and the Aitken mode may grow into the accumulation mode and partially merge with it. Fine mode species considered within the CMAQ aerosol component include sulfate, nitrate, ammonium, water, primary organic aerosols, secondary organic aerosols from both anthropogenic and biogenic origin, elemental carbon, and primary aerosol material not otherwise specified.

[13] The coarse particle mode within CMAQ, representing particles having aerodynamic diameters between 2.5 and 10 μm , consists of wind-blown dust and other large particles of unspecified origin. The anthropogenic portion of these particles results mainly from industrial processes. As mentioned above for fine particles, coarse mode particles in the model also undergo wet and dry deposition. Coarse particles do not interact with fine particles in the current implementation of the model. A discussion of possible interaction between the two modes appears in the work of *Binkowski and Roselle* [2003]. In addition to calculating concentrations of the individual species listed previously, the CMAQ aerosol component also calculates $\text{PM}_{2.5}$, PM_{10} and coarse mode concentrations.

[14] Aerosol processes treated within the CMAQ aerosol component include (1) new particle formation, (2) intermodal and intramodal coagulation, and (3) particle growth by addition of mass. The aerosol module uses differential equations to represent conservation of particle number, surface area, and species mass for each mode and then solves these equations analytically. Further details on the aerosol module are given by *Binkowski* [1999] and *Binkowski and Roselle* [2003].

2.3. Visibility in the CMAQ Aerosol Component

[15] Visibility has no precise definition, but according to *Malm et al.* [1994] “visibility, in the most general sense, reduces to understanding the effect that various types of aerosol and lighting conditions have on the appearance of landscape features.” Fine particles scatter light more efficiently than do larger particles; those particles with diameters between approximately 0.1 and 1.0 μm (included in $\text{PM}_{2.5}$) are the most effective [*Malm*, 1999]. In the eastern United States, the $\text{PM}_{2.5}$ constituents leading to visibility degradation consist primarily of sulfates, nitrates, carbonaceous particles and crustal material [*NEECAUM*, 2001]. Carbonaceous particles mainly absorb light; the other species scatter it. Sulfates and nitrates, being hygroscopic, change from solid particles to solution droplets when the relative humidity exceeds 70% [*Malm*, 1999]. As a result, these species contribute more to visibility degradation than do nonhygroscopic particles. For a thorough discussion of visibility, see *Malm* [1999].

[16] Visibility observations may take three different forms, the first two of which are visual range (V_r) and extinction coefficient (β_{ext}). V_r is the farthest distance at which an observer can discern the outline of a black object on the horizon. As aerosol loading increases, V_r decreases. V_r is an intuitive, although subjective, measurement used primarily for landing aircraft and maneuvering vehicles safely. The β_{ext} measures the attenuation, per unit distance,

of image-forming light between an observer and an object. This attenuation results from light scattering and/or absorption by atmospheric particles and gas molecules. These atmospheric particles may have either anthropogenic or biogenic origin. As visibility decreases from aerosol loading, β_{ext} increases. V_r and β_{ext} inversely relate to each other through the standard form of the Koschmieder equation [*Pitchford and Malm*, 1994]:

$$V_r = \frac{3.91}{\beta_{ext}} \quad (1)$$

with V_r in km and β_{ext} in km^{-1} .

[17] Unfortunately, neither V_r nor β_{ext} is linear with perceived visual changes resulting from uniform haze. Therefore, *Pitchford and Malm* [1994] developed the linear deciview (dv) scale, defined using either V_r or β_{ext} , to measure visibility. The deciview scale is analogous to the decibel scale used in acoustics. A measured physical quantity (here the β_{ext}) is compared to a reference value (the β_{ext} in a pristine or Rayleigh atmosphere, equal to 0.01 km^{-1}); the logarithm of this ratio is then multiplied by 10. The resulting deciview scale changes linearly with respect to perceived visual changes in the atmosphere [*Pitchford and Malm*, 1994]. A visibility measurement of zero dv indicates a pristine atmosphere; as visibility worsens, deciviews increase in magnitude. The equation to calculate haziness on the deciview scale appears below:

$$\text{haziness}(dv) = 10 \ln \frac{\beta_{ext}}{0.01 \text{ km}^{-1}}, \quad (2)$$

$$\text{where } \beta_{ext} [\text{km}^{-1}] = \beta_{sp} + 0.01 [\text{km}^{-1}]. \quad (3)$$

β_{sp} is the aerosol extinction coefficient and must be adjusted by $0.01 [\text{km}^{-1}]$, representing molecular scattering, to arrive at β_{ext} .

[18] The CMAQ aerosol component calculates visibility indices (β_{ext} and deciview) using two different methods: an approximation to Mie theory and a mass reconstruction technique. These methods are discussed below. The present evaluation used deciviews for comparison against observational data.

2.3.1. Mie Theory Approximation

[19] The aerosol β_{ext} (km^{-1}) can be calculated from such ambient aerosol characteristics as index of refraction, volume concentration and size distribution and is usually obtained from a convolution of the size distribution with the Mie extinction efficiency [*Willeke and Brockmann*, 1977]. CMAQ calculates the β_{ext} using a very efficient and reasonably accurate approximation to the Mie efficiency [*Binkowski*, 1999]. In this evaluation, the calculated β_{ext} values are then converted within CMAQ to deciviews, using equation (2).

2.3.2. Mass Reconstruction Method

[20] CMAQ also contains an empirical approach, based upon mass reconstruction, for calculating the extinction coefficient β_{ext} . This method, used by the IMPROVE network and documented by *Malm et al.* [1994] and *Malm* [2000], is well-suited to cases in which routine measure-

ments of aerosol species mass concentrations are available, but particle size distribution information is not.

[21] The mass reconstructed β_{ext} equation used within CMAQ, a slightly modified version (J. Sisler, personal communication, 1998) of that from *Malm et al.* [1994], appears below:

$$\begin{aligned} \beta_{sp}(\text{km}^{-1}) = & 0.003f(RH) [\text{ammonium} + \text{sulfate} + \text{nitrate}] \\ & + 0.004 [\text{all organic species}] \\ & + 0.01 [\text{elemental carbon}] \\ & + 0.001 [\text{unspeciated PM}_{2.5}]. \end{aligned} \quad (4)$$

The brackets denote species mass concentration in $\mu\text{g m}^{-3}$. The coefficients in equation (4) represent scattering efficiencies ($\text{m}^2 \text{mg}^{-1}$), except for light absorbing carbon, which instead has an absorption efficiency ($\text{m}^2 \text{mg}^{-1}$). RH values, averaged for the lowest layer in the vertical profile, are calculated by the meteorology model referenced in section 2.4.

[22] Equation (4) normally contains a term associated with coarse mass concentrations; however, the coarse mass term is not currently implemented within CMAQ because uncertainties in wind-blown dust emissions are too large. Upon determination of better emissions estimates, the coarse mass term will be reinserted into the visibility calculations. Again, in the present evaluation, the calculated β_{ext} values were converted within CMAQ to deciviews, for comparison against observations.

2.4. Model Simulation Characteristics

[23] The model simulations were performed with the 2000 public release version of the Models-3/CMAQ code. For this evaluation, the modeling domain covers the eastern United States (see, for example, Figure 1). Each domain grid cell is 36 km by 36 km. The domain's vertical profile contains 21 layers of varying thickness. Layer 1 is 38 m deep and subsequent layer depths increase with height. The uppermost layer is 3 km thick and extends to about 16 km. The meteorological fields were derived from MM5, the Fifth-Generation Pennsylvania State University/National Center for Atmospheric Research (NCAR) Mesoscale Model [Grell *et al.*, 1994]. MM5 is a three-dimensional, limited area, primitive-equation model that has undergone continuous development, improvement and testing for more than 20 years. This state-of-the-science meteorological model is widely accepted and used by the international scientific community to simulate a wide variety of mesoscale processes [Tanrikulu *et al.*, 2000; Seaman, 2000]. Anthropogenic and biogenic emissions were processed with the Models-3 Emission Processing and Projection System (MEPPS) [Benjey *et al.*, 1999], which relies on the National Emissions Inventory [U.S. Environmental Protection Agency, 2000a].

[24] The simulation period for the visibility data comparison was 6–15 July 1995 and that for the speciated data comparison was 1–30 June 1995. When the visibility evaluation was performed, only the July 1995 simulation results were available for study. By the time the speciated aerosol evaluation was conducted, the June 1995 simulation results were accessible. The first few days of each simulation were regarded as “spin-up” time and data from these days were not used in subsequent analyses. Initial conditions and boundary conditions corresponded to estimates of

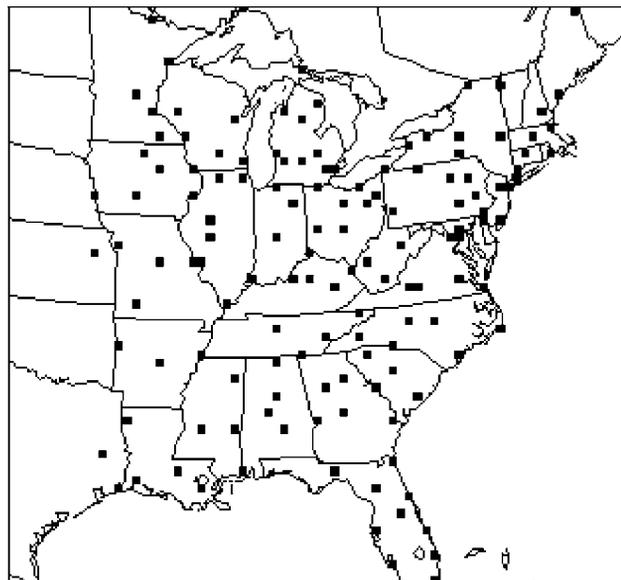


Figure 1. National Weather Service sites (visibility data).

clean air concentrations. Effects from the selection of initial conditions are mitigated by discarding results from the “spin-up” days and the modeling domain is large enough so that boundary condition effects do not dominate far into the domain (G. Gipson, personal communication, 2001).

[25] At the time the present evaluation was conducted, output data were available from only two sets of CMAQ simulations. The July 1995 model runs, originally conducted to demonstrate CMAQ's functionality, employed the Carbon Bond IV (CB-IV) mechanism [Gery *et al.*, 1989]. The basic CB-IV mechanism in CMAQ contains 36 species and 93 reactions, 11 of which are photolytic [Gipson and Young, 1999]. Output data from the July 1995 simulations were used for the visibility evaluation. The June 1995 simulations, performed specifically for CMAQ evaluation, used version 2 of the Regional Acid Deposition Model chemical mechanism (RADM2) [Stockwell *et al.*, 1990]. The basic RADM2 mechanism includes 57 species and 158 reactions, 21 of which are photolytic [Gipson and Young, 1999]. Output data from the June 1995 simulations were used for the speciated aerosol evaluation. Running CMAQ with CB-IV undoubtedly produces different results than running the model under the same conditions using RADM2. Investigating and understanding these differences would require side-by-side model simulations and additional analyses that are beyond the scope of this evaluation.

3. Data Used in Evaluation

3.1. NWS Visibility Data

[26] The July 1995 visibility data used to evaluate CMAQ originated as visual range (V_r) estimates gathered at 174 NWS monitoring stations (mainly airports and military bases) in the eastern United States (see Figure 1). These V_r measurements were converted to extinction coefficient (β_{ext}) format at the Center for Air Pollution and Trends Analysis (CAPITA), an aerosol data repository maintained at Washington University in St. Louis, Missouri, from



Figure 2. IMPROVE sites (speciated aerosol data).

which we obtained the data (R. Husar, personal communication, 2000).

[27] Only noon (local time) observations were used, to assure overhead sunlight and a well-mixed planetary boundary layer. To prevent hydrometeors from biasing the evaluation, observations gathered during periods of precipitation or relative humidities exceeding 90% were discarded. The β_{ext} values were then converted to deciviews according to equation (2) and compared to model output for noontime calculations on the corresponding days. CMAQ output was also filtered to remove predictions biased by rain or by relative humidities above 90%, thus assuring a more valid correspondence between model output and observations.

[28] To remove potential complications associated with boundary effects, data from any NWS monitoring station lying within 10 grid cells (360 km) of the western boundary

was discarded prior to analysis. Because west-to-east flow was assumed, data from monitoring sites lying near the other three boundaries were not excluded in the analysis.

[29] The visibility data used in the present evaluation originated from human observations of V_r and thus have limitations. These observations depend upon such factors as observer visual acuity, number of targets, target configuration, distances to targets, and the physical and optical properties of the targets. Stations having few targets introduce artificial stratification into the data. For example, assuming that targets exist at 15 and 20 km from the observation station, an actual V_r of 18 km would be under-reported as 15 km. Artificial stratifications of this type affect the conclusions drawn from comparisons between observations and model output. Despite these limitations, this data set has spatial and temporal (i.e., daily measurements) coverage unmatched in other data sets.

3.2. IMPROVE Speciated Data

[30] For the second phase of the model evaluation, which focused on the month of June 1995, speciated aerosol data from the IMPROVE network were used. IMPROVE is a collaborative monitoring effort governed by a steering committee composed of representatives from federal, regional, and state organizations [Pitchford and Scruggs, 2000]. The network was designed to (1) establish current visibility and aerosol conditions; (2) identify the chemical species and emission sources responsible for visibility degradation; and (3) document long-term visibility trends [Malm, 2000] at over 100 locations nationwide. However, the majority of IMPROVE sites are located in western states; as a result, only 18 sites fell within the model evaluation domain (Figure 2). Additional information concerning these sites can be found in Table 1. All the sites are rural, except the urban WASH (Washington, D. C.) site.

[31] The IMPROVE speciated data for June 1995 were downloaded from the CAPITA web site (<http://capita.wustl.edu/CAPITA/DataSets/IMPROVE/impnes.html>, 2000). These data, as well as additional measurement information, may also be obtained from the IMPROVE web site ([vista](http://vista.cira.colostate.edu/improve/Data/GraphicViewer/metadata.asp)).

Table 1. IMPROVE Site Information^a

Station Code	Station Location	Location State	Long, °W	Lat, °N	Elev, m
ACAD	Acadia NP	ME	68.308	44.415	129
BOWA	Boundary Waters Canoe Area	MN	91.950	47.950	524
BRIG	Brigantine NWR	NJ	74.472	39.475	9
CHAS	Chassahowitzka NWR	FL	82.567	28.750	2
DOSO	Dolly Sods/Otter Creek Wilderness	WV	79.205	39.143	1158
GRGU	Great Gulf Wilderness	NH	71.217	44.300	439
GRSM	Great Smoky Mountains NP	TN	83.987	35.710	815
JEFF	Jefferson/James River Face Wilderness	VA	79.433	37.667	299
LYBR	Lye Brook Wilderness	VT	73.123	43.243	1010
MACA	Mammoth Cave NP	KY	86.075	37.277	248
MOOS	Moosehorn NWR	ME	67.283	45.117	76
OKEF	Okefenokee NWR	GA	82.117	30.765	49
ROMA	Cape Romain NWR	SC	79.583	33.033	3
SHEN	Shenandoah NP	VA	78.450	38.543	1098
SHRO	Shining Rock Wilderness	NC	83.283	35.650	1621
SIPS	Sipsey Wilderness	AL	87.382	34.358	279
UPBU	Upper Buffalo Wilderness	AR	93.245	35.880	723
WASH	Washington	D.C.	77.063	38.932	16

^aThe information in Table 1 is from CAPITA (<http://capita.wustl.edu/CAPITA/DataSets/IMPROVE/impnes.html>, 2000), except for the elevation data (Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Fort Collins) (<http://vista.cira.colostate.edu/improve/Data/GraphicViewer/metadata.asp>, 2001). NP, National Park; NWR, National Wildlife Refuge.

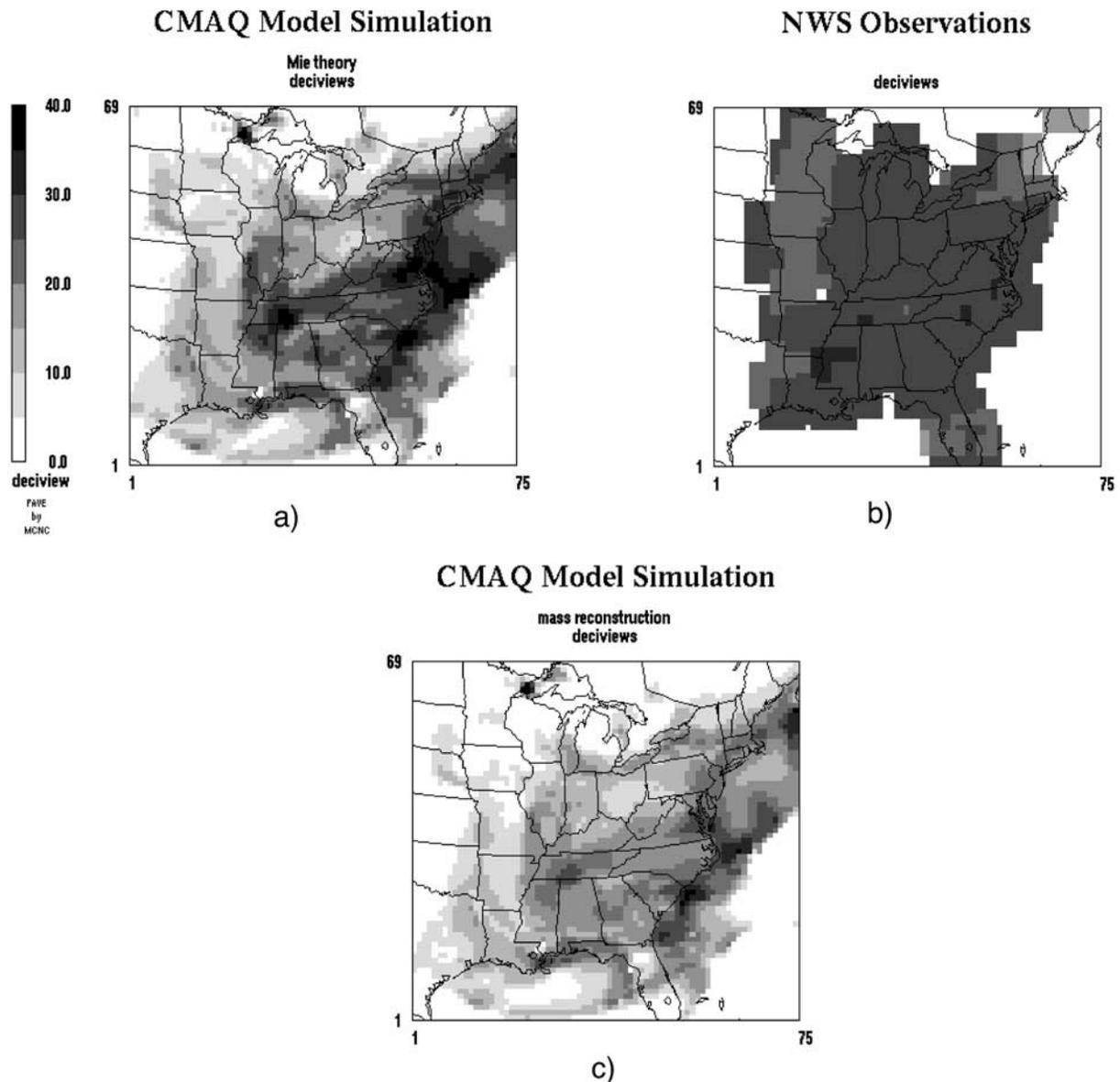


Figure 3. Visibility simulations and observations for 1600 UTC, 11 July 1995. (a) CMAQ Model Simulation, Mie theory, (b) NWS observations, and (c) CMAQ Model Simulation, mass reconstruction.

cira.colostate.edu/improve) maintained by Colorado State University.

[32] In 1995, IMPROVE monitors collected 24-hour integrated samples on Wednesdays and Saturdays (midnight to midnight, local time (R. Eldred, personal communication, 2000)). Given CMAQ's one-month simulation period and IMPROVE's twice-per-week sampling schedule, a total of eight days are available for comparison in June: 3, 7, 10, 14, 17, 21, 24 and 28. According to *UC-Davis* [1995] and *Malm et al.* [1994], each sampler consists of four separate modules (A, B, C and D). Modules A, B, and C collect $PM_{2.5}$ particles, while module D collects PM_{10} particles. Module A contains a Teflon filter that provides most of the $PM_{2.5}$ data. Analysis methods used to measure $PM_{2.5}$ particles include gravimetric measurements, particle-induced X-ray emission (PIXE), proton elastic scattering

analysis (PESA) and X-ray fluorescence (XRF). Module B has a denuder to remove acidic gases, so that filter analysis will result only in particle concentration measurements. The nylon filter from Module B is analyzed by ion chromatography for nitrate, sulfate, chloride, and nitrite. Module C collects carbon on tandem quartz filters that are subsequently analyzed in eight temperature fractions using the Thermal Optical Reflectance (TOR) combustion technique. The analysis results in concentrations of organic carbon and elemental carbon. Module D contains a Teflon filter analyzed by gravimetric analysis for PM_{10} mass concentrations.

[33] Total organic carbon mass concentrations are determined by multiplying measured organic carbon concentrations from Module C by the molar correction factor of 1.4, since the IMPROVE technique assumes that the average organic molecule is 70% carbon [Sisler, 1996]. However,

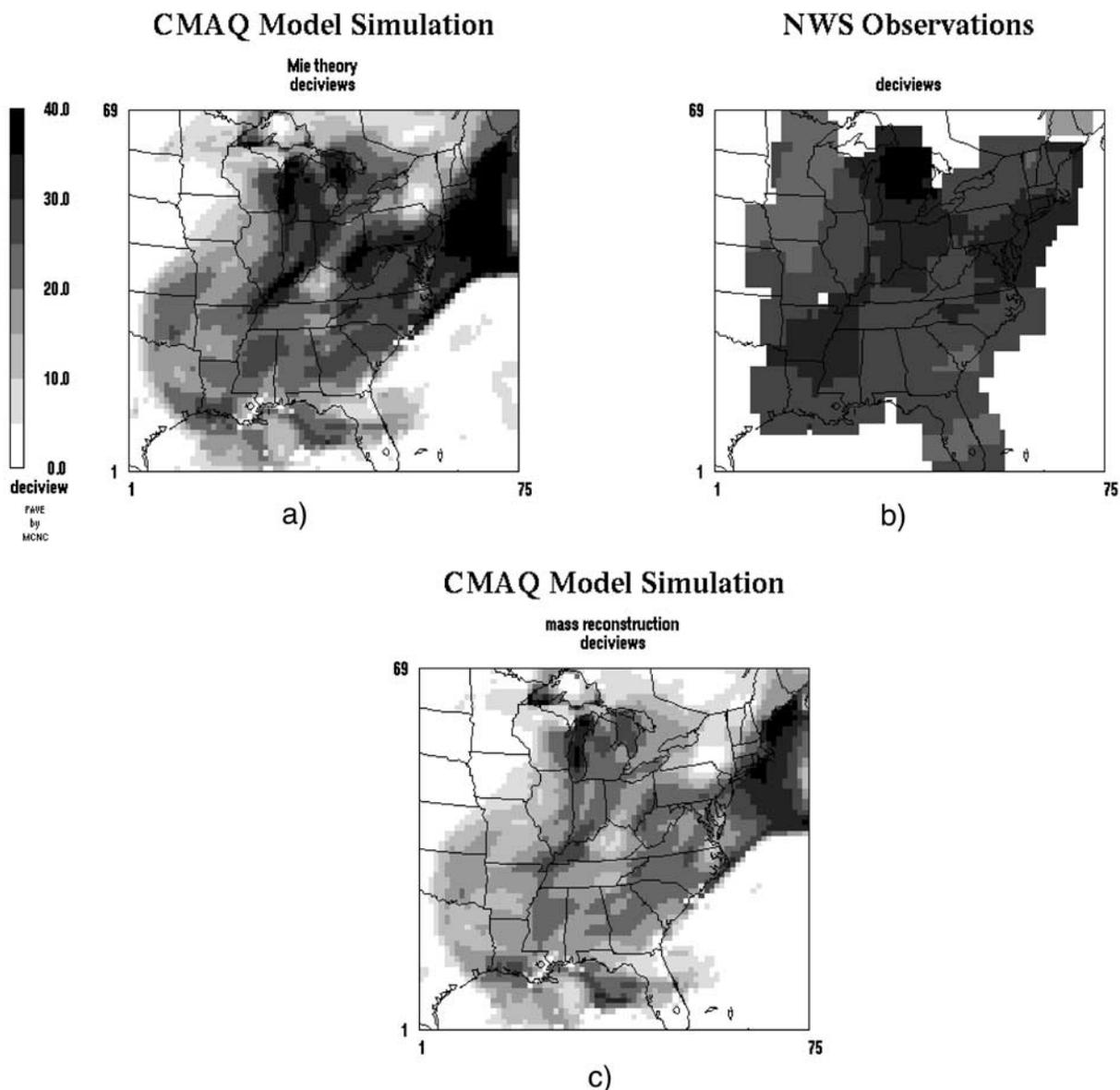


Figure 4. Visibility simulations and observations for 1600 UTC, 13 July 1995. (a) CMAQ Model Simulation, Mie theory, (b) NWS observations, and (c) CMAQ Model Simulation, mass reconstruction.

recent research indicates that a correction factor of 1.4 may be too low. *Turpin and Lim* [2001] suggest that 1.6 ± 0.2 is the appropriate ratio of organic molecular weight per carbon weight for urban aerosols, while 2.1 ± 0.2 more accurately reflects the ratio for nonurban aerosols.

[34] The five species considered in the present evaluation include sulfate, nitrate, organic carbon, and gravimetric $PM_{2.5}$ and PM_{10} . (See Appendix A for caveats concerning sulfate, nitrate, and organic carbon measurements).

4. Evaluation Results

4.1. Visibility

[35] The evaluation of CMAQ for visibility targeted a widespread regional haze and ozone episode that developed over the eastern half of the United States during 11–15 July

1995. This episode was associated with the westward retrogradation of the Bermuda High, a semipermanent anticyclone normally centered over the island of Bermuda. The meteorological conditions during this 5-day period were characterized by anomalously hot temperatures, light winds, and general subsidence, which resulted in minimal cloud cover and precipitation. These conditions, detailed by *Seaman and Michelson* [1995], were conducive to the formation of visibility impairing particles.

4.1.1. NWS Data

[36] Plots of the observed visibility data at 1600 UTC for three days of this 5-day episode appear in Figures 3 through 5. In order to produce a spatial field comparable to CMAQ output, data from the 139 observations were gridded on the same 36-km grid as the model simulations. The eighty-one (9×9) grid cells surrounding each observation station were

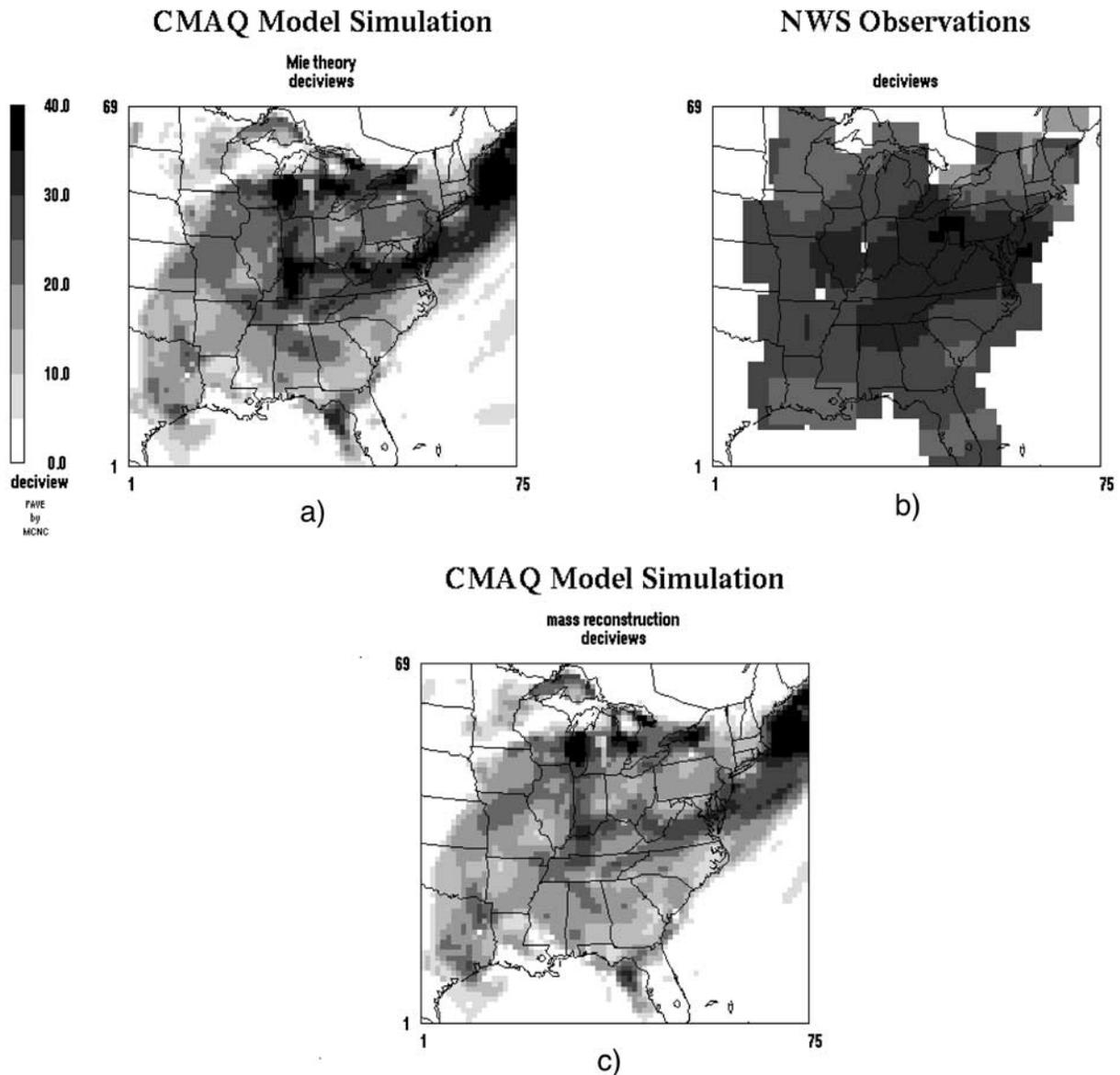


Figure 5. Visibility simulations and observations for 1600 UTC, 15 July 1995. (a) CMAQ Model Simulation, Mie theory, (b) NWS observations, and (c) CMAQ Model Simulation, mass reconstruction.

assigned the deciview value of that station. If a particular cell contained information from more than one station, an average deciview value was used. This simple method of interpolation produced fairly smooth patterns, allowing for regional-scale pattern comparison between observations and CMAQ simulations.

[37] Panels b of Figures 3 through 5 depict the formation and evolution of the large regional haze episode discussed above. At the beginning of the episode (11 July), poorest visibilities (>25 dv) are found from the southeastern United States stretching northward into the Great Lake States. Visibility degradation increases as the episode progresses through 13 July, as poorest visibilities (>30 dv) are now concentrated in three distinct areas: the Mid-Atlantic states, the central Great Lakes, and an area encompassing Arkansas and Mississippi. By 15 July, these three areas merge into

one large area stretching from the Mid-Atlantic to the Ohio River Valley states.

4.1.2. CMAQ Mie Theory Simulation Results

[38] Comparison of CMAQ's Mie theory simulation with the observed spatial patterns reveals a reasonable amount of agreement, as the model generally captures the main gradients and areas of maximum/minimum visibility. Of special note is the area of poor observed visibility (>25 dv) on 11 July (Figure 3) that is well-matched by model results as an area of poor simulated visibility (>20 dv). One exception to this agreement is found along the Atlantic Coast, stretching from New England to Georgia. Here, CMAQ simulates the poorest visibility (40 dv) in an area where observed visibility is considerably better (20–30 dv). Some of this discrepancy can be attributed to the fact that much of this area is offshore and therefore void of observations.

Table 2. Summary Statistics for Visibility Evaluation at 1600 UTC, 11–15 July 1995

Time Period		Visibility, dv					r^2		
		Mean	Mean Error	Normalized Mean Error ^a	Mean Bias	Normalized Mean Bias ^a	Obs	Mie	Recon
11 July (n = 122)	obs	26.0	–	–	–	–	1.00	0.10	0.11
	Mie	19.3	8.0	30.9%	–6.7	–25.8%	0.10	1.00	0.96
	recon	15.7	10.3	39.7%	–10.2	–39.4%	0.11	0.96	1.00
12 July (n = 128)	obs	27.2	–	–	–	–	1.00	0.22	0.19
	Mie	21.5	6.5	23.7%	–5.7	–21.0%	0.22	1.00	0.94
	recon	17.5	9.8	36.0%	–9.8	–35.9%	0.19	0.94	1.00
13 July (n = 128)	obs	28.3	–	–	–	–	1.00	0.48	0.45
	Mie	22.9	6.2	22.8%	–5.3	–19.6%	0.48	1.00	0.96
	recon	18.5	9.9	36.2%	–9.8	–34.7%	0.45	0.96	1.00
14 July (n = 123)	obs	28.2	–	–	–	–	1.00	0.43	0.43
	Mie	23.1	6.4	22.8%	–5.0	–17.7%	0.43	1.00	0.98
	recon	19.2	9.2	32.6%	–9.0	–32.0%	0.43	0.98	1.00
15 July (n = 114)	obs	27.6	–	–	–	–	1.00	0.21	0.18
	Mie	20.9	8.0	26.1%	–6.6	–21.7%	0.21	1.00	0.96
	recon	17.6	10.5	34.5%	–10.0	–36.1%	0.18	0.96	1.00
11–15 July (n = 615)	obs	27.4	–	–	–	–	1.00	0.25	0.24
	Mie	21.6	7.0	25.4%	–5.9	–21.7%	0.25	1.00	0.96
	recon	17.7	10.0	36.2%	–9.8	–35.5%	0.24	0.96	1.00

^a $NME = \frac{\sum_{i=1}^N |m_i - o_i|}{\sum_{i=1}^N o_i} \times 100\%$, $NMB = \frac{\sum_{i=1}^N (m_i - o_i)}{\sum_{i=1}^N o_i} \times 100\%$, where m_i is the model estimate at station i , o_i is the observed value at station i , and N is the number of model-observed pairs for all valid monitoring data.

[39] CMAQ simulations for 13 July (Figure 4) capture the general evolution of the episode. Two of the three areas of greatest visibility degradation (the Mid-Atlantic states and the central Great Lakes) identified by the observations are indeed depicted by the model. The simulation, however, missed the third area encompassing Arkansas and Mississippi. The merger of the three observed areas of visibility degradation by 15 July (Figure 5) is only partially captured by the model. While the general outline and positioning of the large haze “blob” is depicted in the model output, its center is much cleaner than indicated by the observations.

[40] While spatial pattern comparison is important, it needs to be augmented by calculation of summary statistics of the observed data (recorded at the 139 sites) and CMAQ simulation results in the grid cells associated with those sites. Both daily and composite statistics for 1600 UTC, 11–15 July 1995 are given in Table 2, as are the equations for the normalized mean error (NME) and normalized mean bias (NMB).

[41] One common feature seen in Table 2 is the underprediction of the Mie theory simulations. The mean bias ranges from -5.0 dv (NMB = -17.7%) on 14 July to -6.7 dv (NMB = -25.8%) on 11 July, with an average of -5.9 dv (NMB = -21.7%) for the 5-day period. Some of this negative bias is attributable to the previously discussed limitations in the NWS data set; a station having too few targets for determining large visual ranges (V_r) leads to deciview overestimation (i.e., underestimated visibilities, especially in good visibility conditions). The mean error ranges from 6.2 dv (NME = 22.8%) on 13 July to 8.0 dv on 11 July (NME = 30.9%) and 15 July (NME = 26.1%). The mean error averages 7.0 dv (NME = 25.4%) for the 5-day simulation period.

[42] Correlation coefficients were also calculated between the Mie theory simulations and the observations, with r^2 s ranging from 0.10 on 11 July to 0.48 on 13 July. Figure 6a is a scatterplot of Mie theory output versus observations

(n = 615) for the entire simulation, which had an r^2 of 0.25. For reference, 1:1 and 2:1 ratio lines are presented along with the best-fit regression line. The artificial stratification of the NWS observations is obvious in this figure, as is CMAQ’s tendency to underestimate the deciviews (i.e., overestimate visibility). Despite these limitations, over 90% of the simulated values fall within a factor of two of the observations.

4.1.3. CMAQ Mass Reconstruction Simulation

[43] The spatial patterns of the mass reconstruction simulation are very similar to those produced by Mie theory. There is, however, a consistent discrepancy in magnitude between the two simulation techniques (see Table 2), with the mass reconstruction method averaging 3.9 dv lower than Mie theory. The mean bias ranges from -9.0 dv (NMB = -32.0%) on 14 July to -10.2 dv (NMB = -39.4%) on 11 July, averaging -9.8 dv (NMB = -35.5%) for the 5-day period. The mean error varies from 9.2 dv (NME = 32.6%) on 14 July to 10.5 dv (NME = 34.5%) on 15 July, with an average of 10.0 dv (NME = 36.2%) for the five days.

[44] Correlation calculations between the mass reconstruction simulations and observations produced r^2 s ranging from 0.11 on 11 July to 0.45 on 13 July. Figure 6b is a scatterplot of reconstruction method output versus observations (n = 615) for the entire simulation, which had $r^2 = 0.24$. Figure 6b is very similar to Figure 6a. Approximately 85% of the mass reconstruction simulated values fall within a factor of two of the observations, down slightly from Mie theory calculations.

[45] A separate comparison involving the two visibility modeling techniques, Mie theory versus mass reconstruction, revealed excellent agreement, as seen in Figure 6c. The r^2 s for the two simulations vary between 0.94 (12 July) to 0.98 (14 July). This consistency is encouraging, given that the model calculates the results from each method independently. Although both methods use the same species concentrations calculated by CMAQ, the Mie calculations are based on theory and use aerosol size distribution

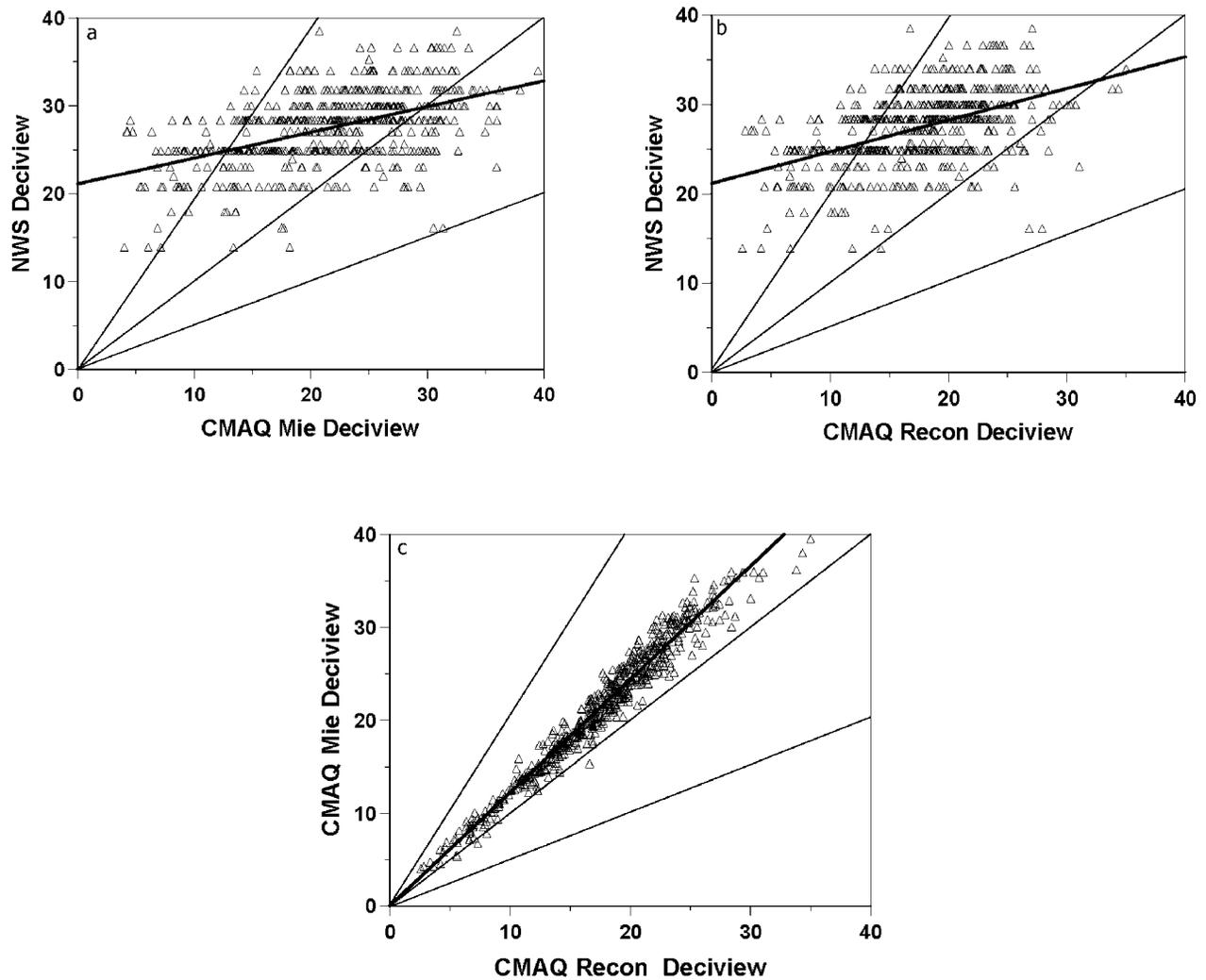


Figure 6. Scatterplots for visibility evaluation at 1600 UTC, 11–15 July 1995. (a) Mie theory, (b) Mass reconstruction, and (c) Mie theory versus mass reconstruction.

parameters determined by CMAQ, whereas the mass reconstruction method originated from empirical measurements that make a priori assumptions regarding aerosol size distribution.

4.2. Speciated Data Evaluation

[46] The evaluation of CMAQ for speciated data targeted the month of June 1995. A cursory examination of meteorological conditions for the month revealed warmer than

Table 3. Summary Statistics for Speciated Aerosol Data

Species	Source	Mean, $\mu\text{g m}^{-3}$	Mean Error, $\mu\text{g m}^{-3}$	Normalized Mean Error ^a	Mean Bias, $\mu\text{g m}^{-3}$	Normalized Mean Bias ^a	r^2
sulfate (n = 129)	CMAQ	4.98	1.75	36.2%	0.15	3.1%	0.63
	IMPROVE	4.83					
nitrate (n = 129)	CMAQ	0.21	0.33	104.3%	-0.10	-33.1%	0.01
	IMPROVE	0.31					
PM _{2.5} (n = 129)	CMAQ	9.06	5.00	38.5%	-3.90	-30.1%	0.55
	IMPROVE	12.96					
PM ₁₀ (n = 129)	CMAQ	13.74	9.85	50.8%	-5.66	-29.2%	0.13
	IMPROVE	19.40					
OC (n = 112)	CMAQ	1.53	0.94	40.6%	-0.78	-33.7%	0.25
	IMPROVE	2.32					

^a $NME = \frac{\sum_{i=1}^N |m_i - o_i|}{\sum_{i=1}^N o_i} \times 100\%$, $NMB = \frac{\sum_{i=1}^N (m_i - o_i)}{\sum_{i=1}^N o_i} \times 100\%$, where m_i is the model estimate at station i , o_i is the observed value at station i , and N is the number of model-observed pairs for all valid monitoring data.

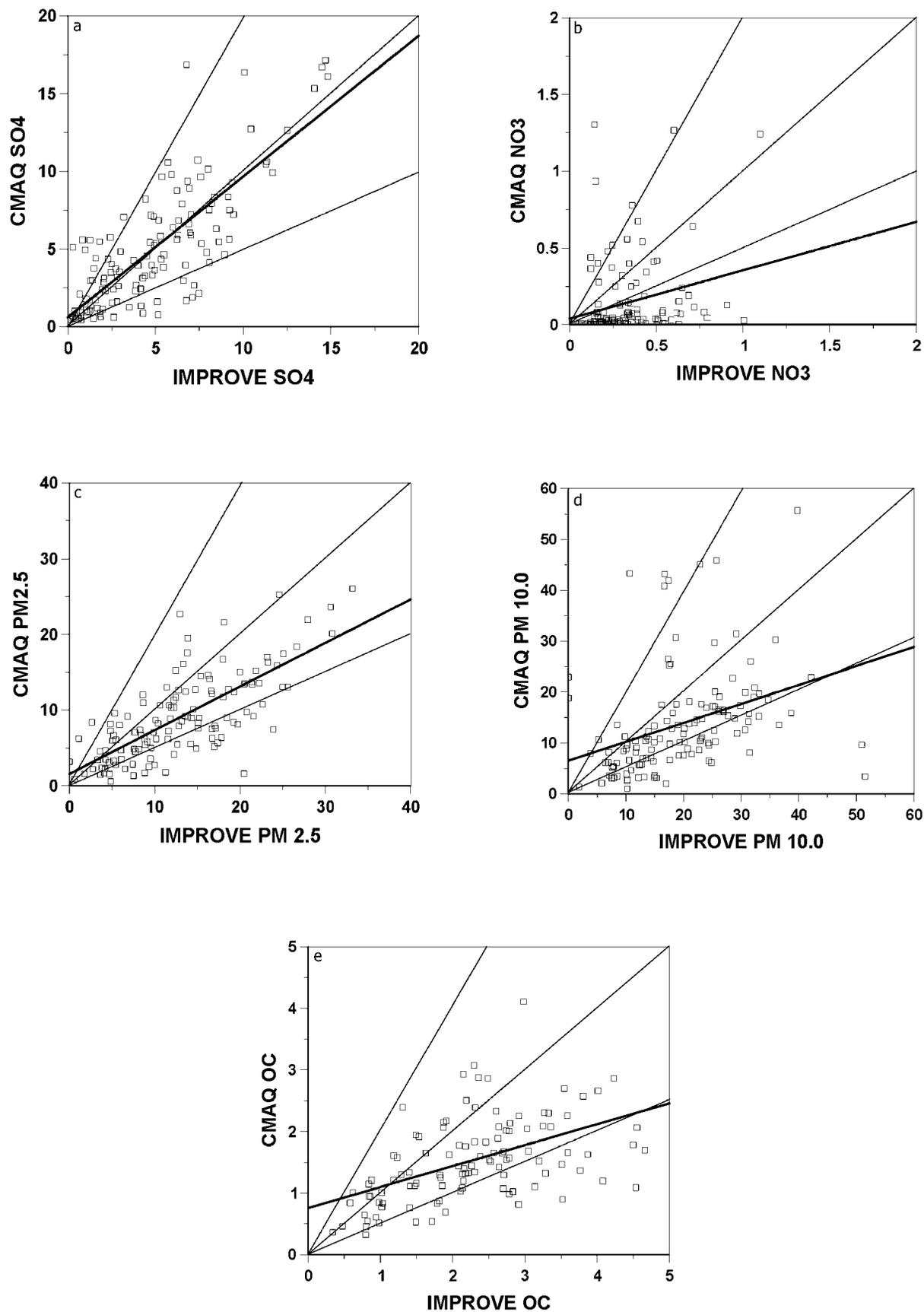


Figure 7. Scatterplots for speciated concentrations ($\mu\text{g}/\text{m}^3$), June 1995. (a) sulfate, (b) nitrate, (c) PM_{2.5}, (d) PM_{10.0}, and (e) organic carbon.

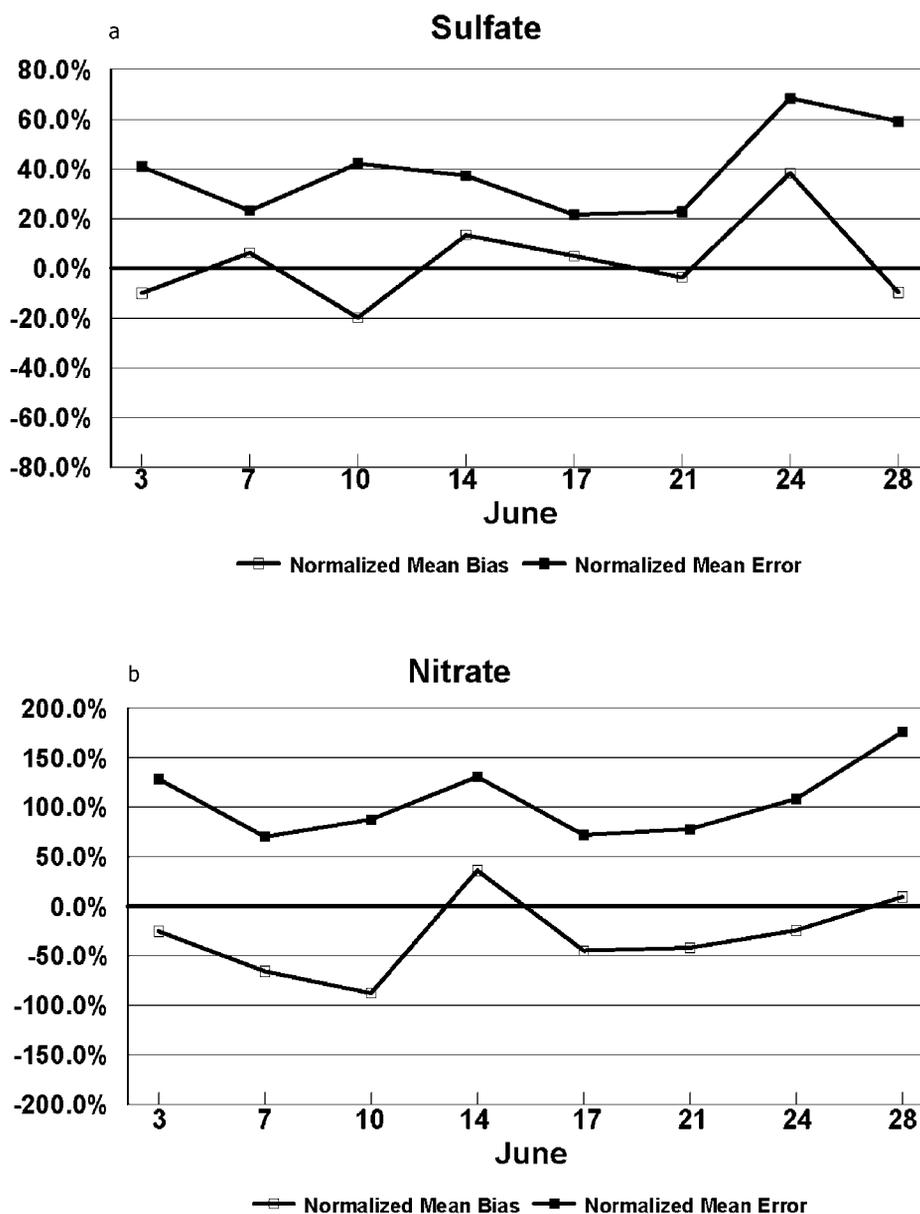


Figure 8. Temporal plots of normalized mean biases (%) and normalized mean errors (%), June 1995. (a) sulfate, (b) nitrate, (c) $PM_{2.5}$, (d) PM_{10} , and (e) organic carbon.

normal conditions (+1 to 2°C) in the upper half of the model domain and slightly cooler than normal conditions (0 to -1°C) in the southern half. Precipitation was above normal from Florida northward into the Virginias and westward into Kansas and Oklahoma. Below normal precipitation was observed in New England, the Great Lakes states, and the southwest portion of the domain.

[47] Given the IMPROVE network's twice-a-week sampling schedule in June 1995, speciated aerosol data existed for the following eight days in June: 3, 7, 10, 14, 17, 21, 24, and 28, providing a maximum of 144 observations (18 stations, 8 days). As discussed earlier, the IMPROVE data represent 24-hour averages. Since CMAQ output is hourly, 24-hour model averages were calculated for comparison against the IMPROVE observations. Because so few observations are available, spatial plots of the observations are

not practical. Therefore, this segment of the evaluation will focus on scatterplots and the examination, across space and time, of the normalized mean error (NME) and normalized mean bias (NMB) associated with each evaluated species.

4.2.1. Sulfate

[48] Examination of Table 3 and Figures 7 through 9 reveals that CMAQ simulates sulfate (SO_4) considerably better than the other species, which are consistently under-predicted. The mean bias is only $0.15 \mu\text{g}/\text{m}^3$ (NMB = 3.1%). Across time, the NMB is within $\pm 25\%$ on all but one day (24 June), while across space it is within $\pm 25\%$ at 15 of the 18 sites. The mean error is considerably larger than the mean bias, however, at $1.75 \mu\text{g}/\text{m}^3$ (NME = 36.2%). The NME is within 50% on all but two days (24 and 28 June) and at all but two locations (ACAD and MOOS). A regression analysis of simulated versus

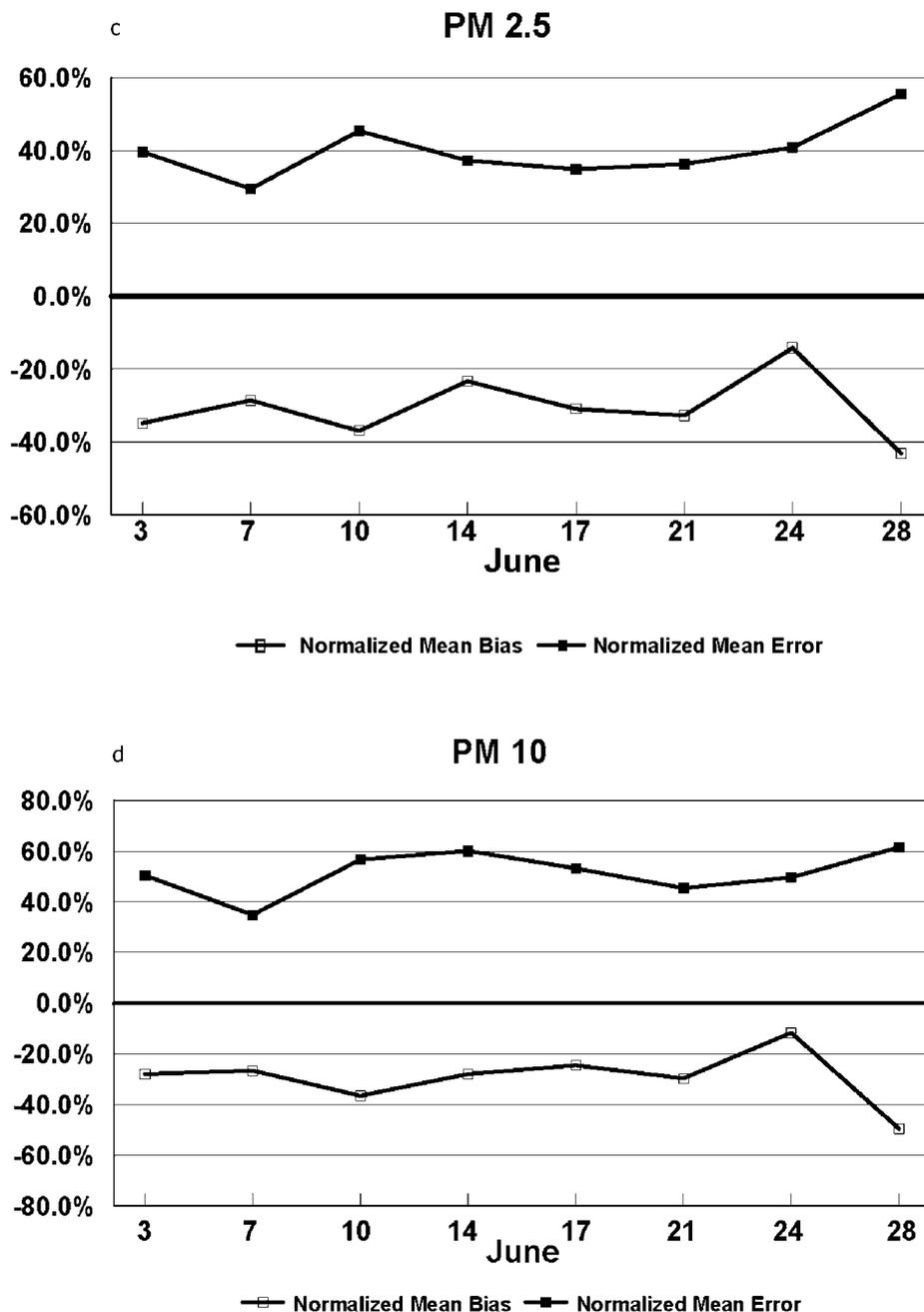


Figure 8. (continued)

observed sulfate concentrations, which produces $r^2 = 0.63$, appears in Figure 7. The best fit regression line is very close to the 1:1 line, with over 80% of the simulated sulfate concentrations falling within a factor of two of the observations.

[49] The relative success of CMAQ in simulating sulfate is not surprising, given its derivation from RPM and RADM (see section 2.2). These models, designed to address the acid rain problem, have provided an excellent foundation for CMAQ.

4.2.2. Nitrate

[50] In contrast to sulfate comparisons, examination of Table 3 and Figures 7 through 9 reveals a very poor level of agreement between simulated and observed concentrations

of nitrate (NO_3), as CMAQ greatly underpredicts nitrate concentrations. The mean bias is $-0.10 \mu\text{g}/\text{m}^3$ ($\text{NMB} = -33.1\%$). Across time, the NMB is within $\pm 25\%$ on only three days and at only one location (SHRO). The mean error, at $0.33 \mu\text{g}/\text{m}^3$, is larger in magnitude than the mean bias and the $\text{NME} = 104.3\%$. The NME is greater than 50% on all days and at all but one location; in fact, the NME exceeds 100% on four days and at two locations. This overall lack of agreement is also reflected in the regression analysis, which produces $r^2 = 0.01$, and the fact that less than 20% of the simulated nitrate concentrations fall within a factor of two of the observations.

[51] A preliminary investigation into CMAQ's poor performance with respect to nitrate suggests that ammonia

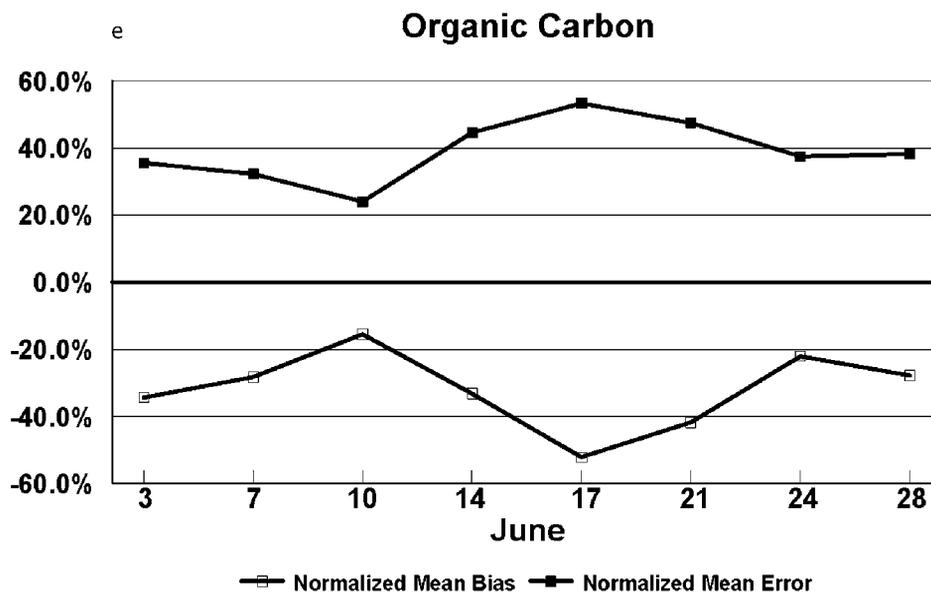


Figure 8. (continued)

emissions based on the National Emissions Inventory (NEI) [U.S. Environmental Protection Agency, 2000a] are probably too low for summer conditions (A. Gilliland et al., Seasonal NH_3 emissions estimates for the eastern United States based on NH_4^+ concentrations, submitted to *Journal of Geophysical Research*, 2003). Emissions from the NEI only provide annually averaged ammonia emission values, but seasonal variability is expected, due to the nature of ammonia emissions. If the ammonia emissions are indeed too low, they would no doubt contribute to the nitrate underprediction seen in the model output for June and July 1995. Efforts are currently underway to determine more realistic ammonia emission estimates [Gilliland et al., 2001] that will eventually allow a more accurate model evaluation with respect to nitrate.

4.2.3. $\text{PM}_{2.5}$

[52] Due in part to the fact that a large component of $\text{PM}_{2.5}$ is sulfate, the level of agreement seen in the $\text{PM}_{2.5}$ evaluation is fairly comparable to the sulfate evaluation, as shown in Table 3 and Figures 7 through 9. The mean bias is $-3.90 \mu\text{g}/\text{m}^3$ (NMB = -30.1%). Across time, the NMB is within $\pm 25\%$ on only two days, while across space the NMB is within $\pm 25\%$ at only 3 of the 18 sites. The mean error, at $5.00 \mu\text{g}/\text{m}^3$, is larger in magnitude than the mean bias and the NME = 38.5% . The NME is within 50% on all but one day (28 June) and at all but four locations (ACAD, BOWA, DOSO, and LYBR). The regression analysis of simulated versus observed $\text{PM}_{2.5}$, which produces an r^2 of 0.55, shows that over 75% of the simulated $\text{PM}_{2.5}$ concentrations fall within a factor of two of the observations.

4.2.4. PM_{10}

[53] Comparisons between simulated and observed PM_{10} concentrations (Table 3 and Figures 7 through 9) reveal that CMAQ underpredicts concentrations, resulting in a mean bias of $-5.66 \mu\text{g}/\text{m}^3$ (NMB = -29.2%). Across time, the NMB is within $\pm 25\%$ on only two days (17 and 24 June) and is negative on all but two days (14 and 28 June). Across space, the NMB is within $\pm 25\%$ at only three locations

(GRSM, SHEN, and SIPS) and is negative at all except two locations (SHEN and WASH). The mean error of $9.85 \mu\text{g}/\text{m}^3$ is larger in magnitude than the mean bias and the NME = 50.8% . The NME is greater than 50% on all except three days and at half of the sites. This overall underprediction and lack of agreement is also reflected in the regression analysis, which produces an r^2 of 0.13. Note, however, that nearly 75% of the simulated PM_{10} concentrations do at least fall within a factor of two of the observations.

[54] It is obvious from this evaluation that processes in the CMAQ aerosol module involving PM_{10} need better representation. Efforts are underway to more accurately model wind-blown dust episodes and also to include the generation and transport of sea salt aerosols, which can affect coastline air quality during onshore flow. Emissions inventories are also undergoing improvement to correct discrepancies. The results of these enhancements should allow for more accurate PM_{10} modeling.

4.2.5. Organic Carbon

[55] As with most of the other species evaluated here, Table 3 and Figures 7 through 9 reveal that CMAQ systematically underpredicts concentrations of organic carbon (OC), resulting in a mean bias of $-0.78 \mu\text{g}/\text{m}^3$ (NMB = -33.7%). Across time, the NMB is within $\pm 25\%$ on only two days (10 and 24 June) and is negative on all days. Across space, the NMB is within $\pm 25\%$ at only three locations (GRSM, SHEN, and SIPS) and it is negative at all but two locations (SHEN and WASH). The mean error of $0.94 \mu\text{g}/\text{m}^3$ is comparable in magnitude to the mean bias and the NME = 40.6% . The NME is within 50% on all days except one (17 June) and at 13 of the 18 locations. The underprediction and lack of agreement seen for organic carbon is also apparent in the regression analysis, which produces an r^2 of 0.25. However, nearly 80% of the simulated organic carbon concentrations fall within a factor of two of observations.

[56] The lack of agreement between simulated and observed organic carbon concentrations results, in part,

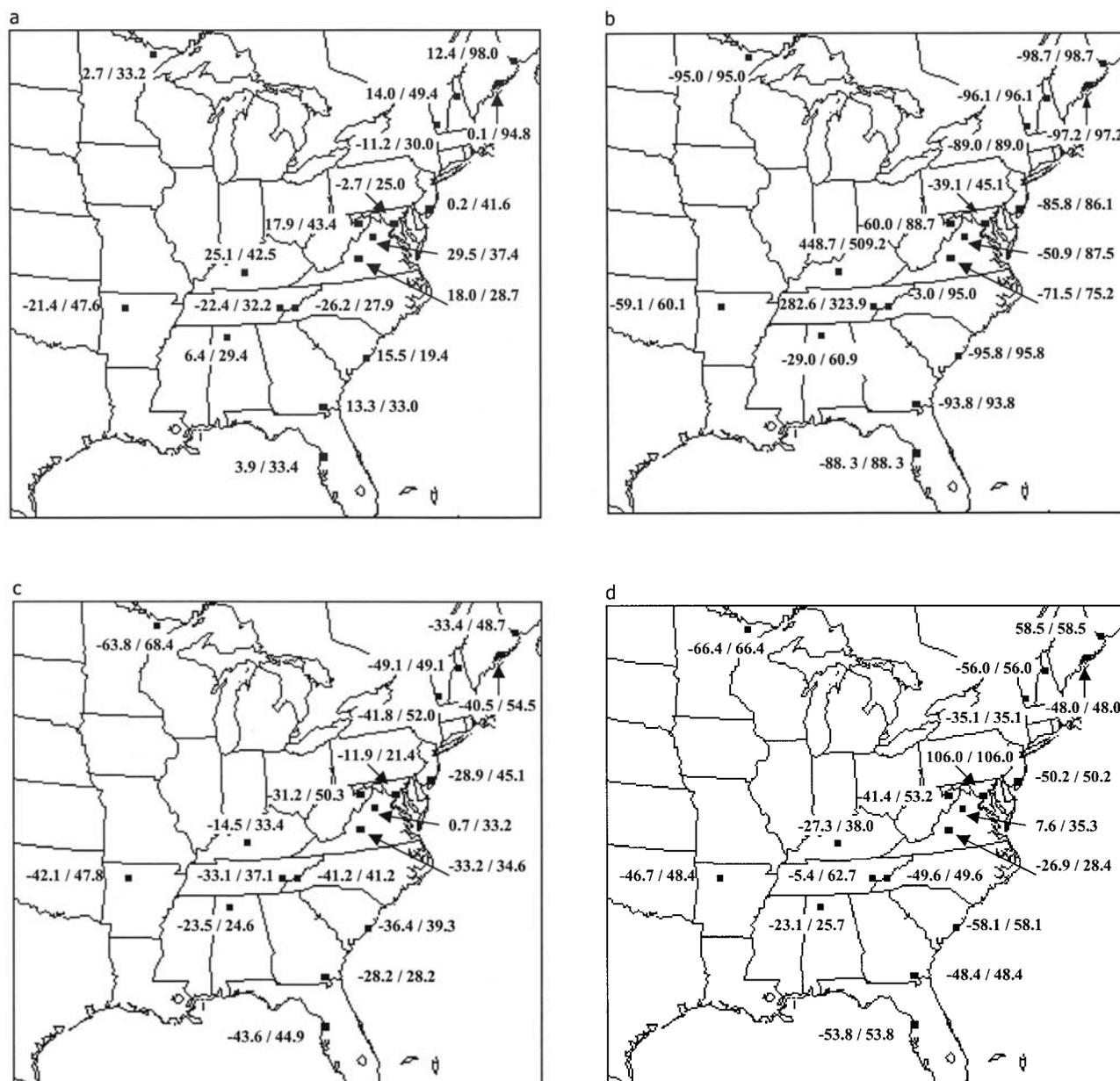


Figure 9. Spatial plots of normalized mean biases (%) and normalized mean errors (%), June 1995. (a) Sulfate, (b) nitrate, (c) PM_{2.5}, (d) PM₁₀, and (e) organic carbon.

from the crude physical representation of organics within the CMAQ aerosol component. Improvements to this representation are underway, following Schell *et al.* [2001]. Further difficulties arise from incomplete knowledge regarding organic aerosol constituents, making it difficult not only to adequately model organic species, but to compare model results with observations.

5. Summary

[57] An initial evaluation of the Models-3 Community Multiscale Air Quality model (CMAQ) aerosol component reveals that CMAQ varies in its ability to simulate observed visibility indices and aerosol species concentrations. The

visibility evaluation, using NWS observations from 139 airports, coincided with an air stagnation event that occurred during 11–15 July 1995, while the aerosol species evaluation used observations of sulfate, nitrate, PM_{2.5}, PM₁₀, and organic carbon, gathered from 18 stations of the IMPROVE network during June 1995. This initial evaluation has helped reveal those processes within CMAQ's aerosol module that require improvement.

5.1. Visibility Evaluation

[58] The visibility evaluation compares the NWS observations against CMAQ simulations, using both Mie efficiency theory approximation and the IMPROVE mass reconstruction technique. Comparison of model results with

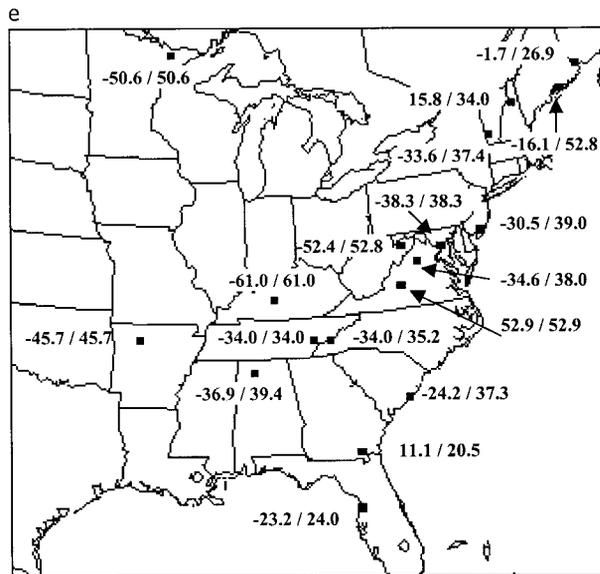


Figure 9. (continued)

observed spatial and temporal patterns of visibility reveals reasonable agreement, as both techniques capture main visibility patterns, including gradients, maxima and minima. Specifically, CMAQ captures the poor visibility experienced by the Ohio Valley and Mid-Atlantic regions throughout the stagnation period, as well as the good visibility seen in the western portion of the model domain. Also captured by the model was the sudden decline in visibility observed in the Great Lakes states, especially Michigan, on 13 July.

[59] Although basic spatial patterns were captured by the techniques, both techniques systematically underpredict visibility degradation (i.e., overpredict visibility). The mean bias for the 5-day Mie theory simulation is -5.9 dv (NMB = -21.7%), while the mean error is 7.0 dv (NME = 25.4%). For the mass reconstruction technique the mean bias is -9.8 dv (NMB = -35.5%) and the mean error is 0.0 dv (NME = 36.2%). Regression analyses confirm these findings, while revealing modest levels of agreement between the observations and the Mie theory ($r^2 = 0.25$) and the mass reconstruction ($r^2 = 0.24$) simulations. Despite the modest agreement, over 90% of the Mie theory simulated values and 85% of the mass reconstruction simulated values fall within a factor of two of the NWS observations.

[60] A number of error sources contribute to the systematic low bias seen in CMAQ visibility calculations. The previously discussed artificial stratification apparent in the NWS observations contributes to some of the discrepancy. The evaluation of CMAQ visibility calculations is also somewhat hampered by the fact that, due to uncertainty in coarse particle emissions, the aerosol module does not account for the contribution of coarse mode particles to extinction.

5.2. Speciated Data Evaluation

[61] The speciated aerosol evaluation compares IMPROVE observations against CMAQ concentrations for five species (sulfate, nitrate, $PM_{2.5}$, PM_{10} , and organic carbon) and reveals that, with the exception of sulfate, the model

systematically underpredicts aerosol concentrations by 30%.

[62] Specifically, the mean bias for sulfate is only 0.15 $\mu\text{g}/\text{m}^3$ (NMB = 3.1%), while the mean error is 1.75 $\mu\text{g}/\text{m}^3$ (NME = 36.2%). Regression analysis reveals an r^2 of 0.63 , with over 80% of the simulated concentrations falling within a factor of two of the observations. As noted previously, CMAQ's relative success in simulating sulfate is due in part to its derivation from earlier computer models designed to simulate acid rain.

[63] In contrast to sulfate, a very poor level of agreement is found between simulated and observed concentrations of nitrate, as revealed by the mean bias (-0.10 $\mu\text{g}/\text{m}^3$ (NMB = -33.1%)) and mean error (0.33 $\mu\text{g}/\text{m}^3$ (NME = 104.3%)). This lack of agreement is also reflected in the regression analysis ($r^2 = 0.01$) and the fact that less than 20% of the simulated nitrate concentrations fall within a factor of two of the observations. As discussed in section 4.2, recent research suggests that the ammonia emissions used for these model simulations were most likely too low for summer conditions. Ongoing efforts to correct the emissions should eventually lead to more realistic simulations of nitrate concentrations. In addition, a heterogeneous N_2O_5 reaction has been added to the most recent version of CMAQ's aerosol module, leading to a more realistic treatment of nitrate concentrations.

[64] Due in part to the fact that a large component of $PM_{2.5}$ is sulfate, the level of agreement seen in the $PM_{2.5}$ evaluation is fairly comparable to the sulfate evaluation, although the mean bias at -3.90 $\mu\text{g}/\text{m}^3$ (NMB = -30.1%) and mean error at 5.00 $\mu\text{g}/\text{m}^3$ (NME = 38.5%) are larger. The regression analysis (r^2 of 0.55) reveals that over 75% of the simulated $PM_{2.5}$ concentrations fall within a factor of two of the observations.

[65] CMAQ's propensity for underprediction continues with PM_{10} concentrations, as the mean bias is -5.66 $\mu\text{g}/\text{m}^3$ (NMB = -29.2%). The mean error, at 9.85 $\mu\text{g}/\text{m}^3$ (NME = 50.8%), is almost twice as large as the mean bias – a characteristic also revealed in the regression analysis (r^2 of 0.13). Note, however, that nearly 75% of the simulated PM_{10} concentrations do fall within a factor of two of the observations. Ongoing efforts to improve CMAQ's ability to simulate PM_{10} concentrations include the addition of both wind-blown dust episodes and the generation and transport of sea salt aerosols to the aerosol module. Emissions inventories are also undergoing improvement and the issue of wildfire emissions is being examined.

[66] CMAQ simulations of organic carbon are likewise too low, resulting in a mean bias of -0.78 $\mu\text{g}/\text{m}^3$ (NMB = -33.7%). The mean error of 0.94 $\mu\text{g}/\text{m}^3$ is comparable in magnitude to the mean bias and the NME = 40.6% . The underprediction and lack of agreement seen for organic carbon is also seen in the regression analysis, which produces an r^2 of 0.25 . However, nearly 80% of the simulated organic carbon concentrations fall within a factor of two of observations. The CMAQ aerosol component has recently been updated to include an improved representation of organic aerosols [Schell *et al.*, 2001] and testing of the new aerosol module is underway.

[67] Inadequacies in evaluation data sets have also been identified, including underestimation of visibility in the NWS data and measurement uncertainties in the IMPROVE

data. Fortunately, the EPA has recently implemented the National PM_{2.5} Monitoring Network, consisting of mass monitoring (1100 sites), routine chemical speciation (300 sites) and supersite characterization. These network measurements will eventually produce a valuable database, allowing a more thorough evaluation of the CMAQ model.

Appendix A

[68] The IMPROVE network samples sulfur-containing aerosols in two different ways [Malm *et al.*, 1994]. The first method, PIXE analysis on the module A Teflon filter, yields sulfur concentration measurements, [S]. Because the molecular weight of sulfate is three times that of sulfur, calculating 3[S] yields sulfate measurements. The second method, ion chromatography of the module B nylon filter, produces sulfate ion measurements directly. However, there is some concern that, although the module B denuder presumably removes SO₂ gas [Malm *et al.*, 1994], unremoved SO₂ may adsorb onto the nylon filter and introduce error into the sulfate ion measurements [UC-Davis, 1995]. To address this concern, Malm *et al.* [1994] compared IMPROVE sulfate measurements against sulfate predictions using 3[S] calculations. Their resulting analysis indicated excellent agreement between sulfate and 3[S] concentrations, with a regression line slope of 1.085 and $r^2 = 0.98$. Still, Malm *et al.* [1994] chose to use the 3[S] method in their subsequent analyses because sulfur measurements are more precise than those for sulfate. However, information from the IMPROVE web site (see http://vista.cira.colostate.edu/improve/Data/DataQuery/SO4_NO3_replacement.htm) reveals the following: "The sulfur measurements from the Module A Teflon were underestimated at some eastern U.S. sites, primarily during the summers of 1992–1995 on days with the highest sulfur loadings coinciding with high humidity. This problem was rectified by increasing the size of the Teflon filters at affected sites (Table 1). For derived values in the IMPROVE database the sulfate ion measurement from the Module B filter is used instead of the Sulfur *3 estimate of sulfate ion mass from module A for time periods before the Teflon filter size was increased at the sites listed below (Table 1)."

[69] Information from Table 1 on the IMPROVE web page listed above revealed that, of the 18 IMPROVE sites used in the current analysis, 11 had their filter size increased prior to June 1995 and data from a 12th site exhibited no evidence of S underestimation, leaving 6 sites with monitors potentially having the sulfur measurement problem. Regression analysis of the June 1995 data used here revealed excellent agreement between sulfate and 3[S] measurements, with $r^2 = 0.98$. Given not only this level of agreement, but also the fact that a majority of the data came from sites unaffected by the sulfur measurement problem, the 3[S] data were analyzed for this evaluation.

[70] The IMPROVE sampling method for particle nitrate uses an acidic vapor denuder followed by a nylon substrate for measuring particle nitrate [Malm *et al.*, 1994]. If nitric acid is not properly removed from the sample stream, it may be misinterpreted as particle nitrate, resulting in inflated particle nitrate mass concentrations [Ames and Malm, 2001]. Another complicating factor is that nitrates volatilize during sampling [UC-Davis, 1995].

[71] According to Chow *et al.* [2001], the definition of organic carbon depends on the analysis methods used to measure it. Different methods produce different results. Like nitrate, organic species also volatilize during sampling [UC-Davis, 1995].

[72] **Acknowledgments.** The authors thank Jerry Gipson, NERL-EPA, for helpful discussions regarding gas-phase chemistry and for technical support in extracting and mapping model output. Disclaimer: This document has been reviewed and approved by the U.S. Environmental Protection Agency for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

- Ames, R. B., and W. C. Malm, Comparison of sulfate and nitrate particle mass concentrations measured by IMPROVE and the CDN, *Atmos. Environ.*, 35, 905–916, 2001.
- Arnold, J. R., and R. L. Dennis, First results from operational testing of the U.S. EPA Models-3/Community Multiscale Model for Air Quality (CMAQ), in *Air Pollution Modeling and Its Application XIV*, edited by S.-E. Gryning and F. A. Schiermeier, pp. 651–658, Kluwer Acad., Norwell, Mass., 2001.
- Benjey, W. G., J. M. Godowitch, and G. L. Gibson, Emission subsystem, in *Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System*, edited by D. W. Byun and J. K. S. Ching, Rep. EPA-600/R-99/030, pp. 4–1–4–107, U.S. Environ. Prot. Agency, U.S. Govt. Print. Off., Washington, D. C., 1999.
- Binkowski, F. S., The aerosol portion of Models-3 CMAQ, in *Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System*, edited by D. W. Byun and J. K. S. Ching, Rep. EPA-600/R-99/030, pp. 10-1–10-23, U.S. Environ. Prot. Agency, U.S. Govt. Print. Off., Washington, D. C., 1999.
- Binkowski, F. S., and S. J. Roselle, Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component, 1, Model description, *J. Geophys. Res.*, 108, doi:10.1029/2001JD001409, in press, 2003.
- Binkowski, F. S., and U. Shankar, The regional particulate matter model, 1, Model description and preliminary results, *J. Geophys. Res.*, 100, 26,191–26,209, 1995.
- Byun, D. W., and J. K. S. Ching, Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) modeling system, Rep. EPA-600/R-99/030, U.S. Environ. Prot. Agency, U.S. Govt. Print. Off., Washington, D.C., 1999.
- Chang, J. S., et al., The regional acid deposition model and engineering model, in *National Acid Precipitation Assessment Program, Acidic Deposition: State of Science and Technology*, vol. 1, NAPAP SOS/T Rep. 4, Natl. Acid Precip. Assess. Program, Washington, D.C., 1990.
- Chow, J. C., J. G. Watson, D. Crow, D. H. Lowenthal, and T. Merrifield, Comparison of IMPROVE and NIOSH carbon measurements, *Aerosol Sci. Technol.*, 34, 23–34, 2001.
- Eder, B. K., M. R. Mebust, F. S. Binkowski, and S. J. Roselle, A preliminary evaluation of Models-3 CMAQ using visibility parameters, paper presented at the International Symposium on the Measurement of Toxic and Related Air Pollutants, Research Triangle Park, N. C., 12–14 Sept. 2000.
- Gery, M. W., G. Z. Whiteen, J. P. Killus, and M. C. Dodge, A photochemical kinetics mechanism for urban and regional scale computer modeling, *J. Geophys. Res.*, 94, 12,925–12,956, 1989.
- Gilliland, A., R. Dennis, S. Roselle, T. Pierce, and L. Bender, Estimating the seasonality of airborne NH₃ emissions with an inverse modeling technique, paper presented at the N2001 Second International Nitrogen Conference, Potomac, Md., 14–18 Oct., 2001.
- Gipson, G. L., and J. O. Young, Gas-phase chemistry, in *Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System*, edited by D. W. Byun and J. K. S. Ching, Rep. EPA-600/R-99/030, pp. 8–1–8–98, U.S. Environ. Prot. Agency, U.S. Govt. Print. Off., Washington, D. C., 1999.
- Grell, G. A., J. Dudhia, and D. R. Stauffer, A description of the fifth-generation Penn State/NCAR mesoscale model (MM5), *NCAR Tech. Note NCAR/TN-398+STR*, Natl. Cent. for Atmos. Res., Boulder, Colo., 1994.
- Malm, W. C., Introduction to visibility, Coop. Inst. for Res. in the Atmos., Colo. State Univ., Fort Collins, 1999. (Available at http://yampa.cira.colostate.edu/IMPROVE/education/intro_to_visibility.pdf.)
- Malm, W. C., Overview and summary, in *Spatial and Seasonal Patterns and Temporal Variability of Haze and Its Constituents in the United States*, pp. S-1–S-19, Coop. Inst. for Res. in the Atmos., Colo. State Univ., Fort Collins, 2000.

- Malm, W. C., J. F. Sisler, D. Huffman, R. A. Eldred, and T. A. Cahill, Spatial and seasonal trends in particle concentration and optical extinction in the United States, *J. Geophys. Res.*, 99, 1347–1370, 1994.
- Northeast States for Coordinated Air Use Management (NESCAUM), Regional haze and visibility in the northeast and mid-Atlantic states, edited by G. Kleiman and A. Marin, Boston, Mass., 2001.
- Pitchford, M. L., and W. C. Malm, Development and applications of a standard visual index, *Atmos. Environ.*, 28, 1049–1054, 1994.
- Pitchford, M. L., and M. Scruggs, IMPROVE network-Current and future configurations, in *Spatial and Seasonal Patterns and Temporal Variability of Haze and Its Constituents in the United States*, pp. 1–1–18, Coop. Inst. for Res. in the Atmos., Colo. State Univ., Fort Collins, 2000.
- Schell, B., I. J. Ackermann, H. Hass, F. S. Binkowski, and A. Ebel, Modeling the formation of secondary organic aerosol within a comprehensive air quality model system, *J. Geophys. Res.*, 106, 28,275–28,293, 2001.
- Seaman, N. L., Meteorological modeling for air-quality assessments, *Atmos. Environ.*, 34, 2231–2259, 2000.
- Seaman, N. L., and S. A. Michelson, Mesoscale meteorological structure of a high-ozone episode during the 1995 NARSTO-Northeast study, *J. Appl. Meteorol.*, 39(3), 384–398, 1995.
- Sisler, J. F., Optical and aerosol data, in *Spatial and Seasonal Patterns and Long Term Variability of the Composition of the Haze in the United States: An Analysis of Data From the IMPROVE Network*, chap. 2, pp. 2-1–2-8, Coop. Inst. for Res. in the Atmos., Colo. State Univ., Fort Collins, 1996.
- Stockwell, W. R., P. Middleton, J. S. Chang, and X. Tang, The second generation regional acid deposition model chemical mechanism for regional air quality modeling, *J. Geophys. Res.*, 95, 16,343–16,367, 1990.
- Tanrikulu, S., D. R. Stauffer, N. L. Seaman, and A. J. Ranzieri, A field-coherence technique for meteorological field-program design for air quality studies, part II, Evaluation in the San Joaquin Valley, *J. Appl. Meteorol.*, 20, 317–334, 2000.
- Turpin, B. J., and H.-J. Lim, Species contributions to PM_{2.5} mass concentrations: Revisiting common assumptions for estimating organic mass, *Aerosol Sci. Technol.*, 35, 602–610, 2001.
- UC-Davis, IMPROVE data guide: A guide to interpret data, Univ. of Calif., Davis, 1995. (Available at <http://vista.cira.colostate.edu/improve/Publications/otherDocs/IMPROVEDataGuide/IMPROVEDataGuide.htm>.)
- U.S. Environmental Protection Agency, Air quality criteria for particulate matter, *Rep. EPA/600/P-95/001aF*, Research Triangle Park, N. C., 1995.
- U.S. Environmental Protection Agency, National air pollutant emission trends: 1990–1998, *Rep. EPA/454/R-00-002*, Off. of Air Qual. Plann. and Stand., Research Triangle Park, N. C., 2000a. (Available at <http://www.epa.gov/ttn/chieff/trends/trends98>.)
- U.S. Environmental Protection Agency, National air quality and emissions trends report, 1998, *Rep. EPA/454/R-00-003*, Off. of Air Qual. Plann. and Stand., Research Triangle Park, N. C., 2000b.
- Willeke, K., and J. E. Brockmann, Extinction coefficients for multimodal atmospheric particle size distributions, *Atmos. Environ.*, 11(10), 995–999, 1977.

F. S. Binkowski, 717 Tinkerbell Rd., Chapel Hill, NC 27517, USA.

B. K. Eder, M. R. Mebust, and S. J. Roselle, National Exposure Research Laboratory, U.S. EPA, E243-03, Research Triangle Park, NC 27711, USA. (mmebust@hpcc.epa.gov)