SUPPORTING MATERIAL

Source apportionment of fine particulate matter over the Eastern U.S. Part I. Source sensitivity simulations using CMAQ with the Brute Force method

Michael Burr and Yang Zhang

Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC, 27695

CONTENT

1. Observational dataset used for model evaluation

2. Model evaluation results

Table S1. The observational networks and satellites used in model evaluation, as well as the variables evaluated the sampling frequency, and the number of sites within the 12-km domain

Table S2. Performance statistics for surface concentrations of PM components simulated by CMAQ in January 2002

Table S3. Performance statistics for surface concentrations of PM components simulated by CMAQ in July 2002

Figure S1. Spatial distribution monthly-mean 8-h O3 (top) and 24-h average PM2.5 concentrations simulated by CMAQ overlaid with observations from AIRS-AQS, CASTNET, and SEARCH in January and July.

1. Observational dataset used for model evaluation

Table S1 summarizes information on the networks and satellites used in this study. Simulation results are compared with surface observations from five monitoring networks, including the Clean Air Status and Trends Network (CASTNET) (http://www.epa.gov/castnet), the Speciation Trends Network (STN), the Interagency Monitoring of Protected Visual Environments (IMPROVE) network (http://vista.cira.colostate.edu/improve), the Southeastern Aerosol Research and Characterization (SEARCH) network (http://www.atmospheroc-research.com/studies/SEARCH/index.html), and the Aerometric Information Retrieval System – Air Quality Subsystem (AIRS-AQS) (http://www.epa.gov/ttn/airs/airsaqs). Different networks use different protocols to distinguish OC from EC. STN and AIRS-AQS use the thermal optical transmittance (TOT) method, which is different from the thermal optical reflectance (TOR) protocol used at IMPROVE and SEARCH (the latter is consistent with the protocol used to estimate OC and EC emissions in the NEI used in CMAQ). The observed OC data from STN and AIRS–AQS are thus not comparable with those of IMPROVE and SEARCH as well as the model predictions. TC, instead of separate EC and OC, at the STN and AIRS–AQS sites is therefore used to evaluate simulated TC. The wet deposition fluxes of SO_4^{2-}, NO_3^-, and NH_4^+ are evaluated against observations from the National Acid Deposition Program (NADP).

2. Model evaluation results

In January, max 1-h and 8-h average O_3 mixing ratios are slightly overpredicted at the AIRS-AQS and CASTNET sites and slightly underpredicted at the SEARCH sites. In July, max 1-h O_3 mixing ratios are slightly underpredicted at all sites, and max 8-h average O_3 mixing ratios are slightly underpredicted at the AIRS-AQS and CASTNET sites but slightly overpredicted the SEARCH sites. Zhang et al. (2006b) suggested using NMBs < 15% and NMEs < 30% as indicators of satisfactory O_3 performance. With NMBs ranging from- 8.2% to 7.2% and NMEs from 12.3 – 22.7%, the model performance for max 1-h and 8-h O_3 mixing ratios is considered to be satisfactory.

Figure S1 shows the spatial distribution of simulated monthly-mean max 8-h O_3 mixing ratios overlaid with observations from AIRS-AQS, CASTNET, and SEARCH in January and July. CMAQ does well in capturing the general spatial distribution in January, with higher values across the southern half of the domain. It also does well in capturing higher values in the mountains of western N.C. CMAQ underpredicts 8-h O_3 mixing ratios across the majority of the domain in July, particularly throughout the upper Midwest and Northeast U.S. Possible reasons for biases in simulated 8-h O_3 include uncertainties in emissions of gaseous precursors such as NO_2 and VOCs, biases in meteorological predictions such as temperature and precipitation, as well as uncertainties in boundary conditions of O_3.

24-h average $PM_{2.5}$ concentrations at all networks are overpredicted in January with NMBs of 18.8 – 52.1% but underpredicted in July with NMBs of -39.2% to -26.3%. As shown in Figure S1, CMAQ captures the general spatial distribution throughout the Mid-Atlantic States in January, though it overpredicts in the Atlanta area, possibly due to overestimation of emissions. The majority of the upper-Midwest shows an overprediction, with the exception of Chicago and St. Louis, which show underpredictions. Similarly, $PM_{2.5}$ is overpredicted throughout most of New England in January, indicating overestimation of emissions in the large urban areas of the Northeast (e.g., New York City (NYC), Washington, D.C.). In July, $PM_{2.5}$ is consistently underpredicted across the entire domain.

It is also important to evaluate $PM_{2.5}$ components in order to assess whether good model performance is a result of compensating errors (Zhang et al., 2009). As shown in Tables S3 and S4, SO_4^{2-} and NH_4^+ are underpredicted at the CASTNET, IMPROVE, and STN sites in January with NMBs of -23.3% to -13.9%. However, SO_4^{2-} and NH_4^+ are overpredicted at the SEARCH sites in January with NMBs of 14.3% and 68.4%, respectively. NO_3^- is overpredicted at all sites in January with NMBs of 3.7 – 56.6%. With the exception of a 25.4% positive bias for BC at the SEARCH sites, all carbon species are underpredicted at all sites in January, with NMBs of -30.5 to -3.8%. Large values for MNB and MNB (> 100%) occur for SO_4^{2-}, NO_3^-, NH_4^+, and BC at SEARCH sites and NH_4^+ at STN sites in January. These large values are likely inflated by very small hourly observations at these sites. In July, SO_4^{2-} is underpredicted at the CASTNET, IMPROVE, and STN sites with NMBs of

11.2 – 19.8%. SO_4^{2-} is, however, overpredicted at the SEARCH sites, with an NMB of 31.2%. NO_3^- is underpredicted at all sites in July, with NMBs of -87.9 to -65.3%. NH_4^+ is also underpredicted at all sites in July, with NMBs of -52.5 to -13.7%. All carbon species are underpredicted at all networks in July, with NMBs of -63.2 to -10.7%. Large values for MNB and MNE also occur for SO_4^{2-} and BC at SEARCH sites in July, likely due to small hourly observations at these sites. Additionally, values of infinity for MNB and MNE for NH_4^+ at AIRS-AQS in January and July and BC at IMPROVE in July indicate very small hourly observations approaching 0. While performance at the SEARCH sites is relatively poor for all PM species (particularly in January), the overall model performance for PM and its composition is consistent with current PM model performance reported in the literature (e.g., Eder and Yu, 2006; Zhang et al., 2006b, 2009) and is considered to be satisfactory based on recommended criteria in these papers.

Several reasons may exist for the discrepancies between observed and simulated $PM_{2.5}$ concentrations. Uncertainties in meteorological predictions can heavily influence $PM_{2.5}$ predictions. For example, the underprediction of $PM_{2.5}$ and its components in July may be caused by a large positive bias for precipitation during the summer months and the overprediction of $PM_{2.5}$ in January may be related to a large negative temperature bias (~0.8 ˚C) in MM5 predictions (Olsen, 2009). The large negative bias for temperature may also help explain the underprediction of SO_4^{2-} and overprediction of NO_3^- in January. There also exist uncertainties in the emissions of primary $PM_{2.5}$ and precursors for secondary $PM_{2.5}$. For example, large underpredictions of all carbon species in both months may be due to underestimation of wildfire emissions (Tian et al., 2009). The anomalous large model biases in the concentrations of NH_4^+, SO_4^{2-}, and NO_3^- in January and July at the SEARCH sites may be partly due to the inaccuracies of emissions precursor species (e.g., NH_3, SO_2, and NO_x) in these areas (Zhang et al, 2006a).

Tables S2 and S3 also show the performance of CMAQ for wet deposition fluxes of SO_4^{2-}, NO_3, and NH_4^+. In January, CMAQ slightly overpredicts wet deposition of SO_4^{2-+} with an NMB of 1.3%. Conversely, the wet deposition fluxes of NH_4^+ and NO_3^- is underpredicted in January with NMBs of -20.3% and -5.3%, respectively. In July, the trend in the predictions of the wet deposition fluxes is similar to that in January but with a much greater underprediction in the wet deposition fluxes of NH_4^+ and NO_3^- (NMBs of -38.9% and -55.6%, respectively). The large underpredictions in wet deposition fluxes of NH_4^+ and NO_3^- in July are likely related to large surface underpredictions of concentrations of these species (NMBs of -13.7% to -52.5% for NH_4^+, NMBs of -65.3% to -87.9% for NO_3^-) while the underprediction in January may be more related to the underprediction of precipitation by MM5. The MNB and MNE statistical measures are not applicable for wet deposition fluxes due to several hourly observations with no precipitation.

These evaluation results are overall consistent with those of Morris et al. (2009), differences in model performance exist at SEARCH and STN sites. For example, for Jan. $PM_{2.5}$ concentrations, the fractional biases (FBs) from our work and Morris et al. (2009) are 36% and -2% for SEARCH and 20% and -10% for STN, respectively. For Jul. $PM_{2.5}$ concentrations, FBs from our work and Morris et al. (2009) are -26% and -35 for SEARCH and -32% and -50% for STN, respectively. These differences can be attributed to some differences in emissions inventories for non-VISTAS States, observational data, and model evaluation scripts used in both works.

Table S1. The observational networks and satellites used in model evaluation, as well as the variables evaluated the sampling frequency, and the number of sites within the 12-km domain

Network	Variables		Sampling Frequency	Number of Sites (PM)	Number of Sites (Gas)
	Gas	PM			
AIRS-AQS	O_3	$PM_{2.5}$, NH_4^+, SO_4^{2-}, NO_3^-,	Hourly	~650	~192 (January) ~683 (July)
IMPROVE		$PM_{2.5}$, SO_4^{2-}, NO_3^-, BC, OC, TC	24-hour average (1-in-3 day)	~32	
CASTNET	O_3	SO_4^{2-}, NO_3^-, NH_4^+	7-day average (hourly O_3)	~45	~45
STN		$PM_{2.5}$, SO_4^{2-}, NO_3^-, NH_4^+, TC	24-hour average (daily, 1-in-3, or 1-in-6 day)	~97	
SEARCH	O_3	$PM_{2.5}$, SO_4^{2-}, NO_3^-, NH_4^+, BC, OC, TC	24-hour average and continuous	~100	~3
NADP		Wet Deposition of SO_4^{2-}, NO_3^-, NH_4^+	Weekly Totals	~100	

Table S2. Performance statistics for surface concentrations of PM components simulated by CMAQ in January 2002

Variable	Network	January								
		MeanObs	MeanSim	Number	NMB	NME	MNB	MNE	RMSE	COR
24-h avg. SO_4^{2-}	CASTNET	2.5	1.9	173	-23.9	25.7	-22.0	24.5	0.8	54.8
	IMPROVE	2.4	2.0	292	-13.9	32.3	-8.7	31.6	1.1	71.5
	SEARCH	2.2	2.5	1 334	14.3	64.2	105.5	142.0	2.0	15.6
	STN	3.4	2.6	645	-23.3	45.0	47.4	92.5	2.8	29.7
	AIRS-AQS	2.9	2.5	625	-12.4	33.8	46.6	85.0	1.3	61.4
24-h avg. NO_3^-	CASTNET	2.3	2.3	173	3.7	42.4	53.5	87.8	1.3	73.3
	IMPROVE	1.5	1.8	292	20.8	56.1	32.2	86.3	1.3	76.2
	SEARCH	1.7	2.6	1 319	56.6	114.5	294.3	345.1	2.7	25.2
	STN	3.7	3.8	565	3.8	56.5	34.0	75.4	3.0	47.7
	AIRS-AQS	3.1	3.7	545	17.4	48.4	30.3	63.7	2.3	64.7
24-h avg. NH_4^+	CASTNET	1.3	1.3	173	-0.1	21.1	6.3	25.4	0.4	79.3
	SEARCH	0.9	1.5	822	68.4	97.2	340.8	361.3	1.2	31.4
	STN	2.2	1.9	173	-11.6	55.6	118.7	146.7	2.6	31.0
	AIRS-AQS	1.6	1.8	626	16.3	38.7	Inf	Inf	0.9	67.5
24-h avg. BC	IMPROVE	0.4	0.4	294	-3.8	44.1	1.6	43.7	0.3	50.3
	SEARCH	1.5	1.9	743	25.4	96.7	130.3	162.2	2.1	16.2
24-h avg. OC	IMPROVE	2.3	1.5	294	-7.3	43.6	4.9	48.3	1.1	53.6
	SEARCH	3.4	3.0	62	-14.6	35.4	0.1	34.6	1.3	71.6
24-h avg. TC	SEARCH	5.4	4.4	1 132	-19.3	59.8	65.8	118.3	4.6	40.9
	STN	5.3	3.7	633	-30.2	52.6	19.0	82.7	4.1	31.6
	IMPROVE	2.1	1.9	294	-7.4	42.3	2.9	45.2	1.4	55.5
	AIRS-AQS	4.9	3.7	620	-26.3	45.7	-0.9	61.4	3.1	38.9
SO_4^{2-}_WD	NADP	1.3	1.3	335	1.3	72.0	NaN	NaN	1.5	30.6
NO_3^-_WD	NADP	1.6	1.5	335	-5.3	90.0	NaN	NaN	2.2	11.7
NH_4^+_WD	NADP	0.3	0.2	335	-20.3	85.5	NaN	NaN	0.5	20.2

Table S3. Performance statistics for surface concentrations of PM components simulated by CMAQ in July 2002

		July								
Variable	**Network**	**MeanObs**	**MeanSim**	**Number**	**NMB**	**NME**	**MNB**	**MNE**	**RMSE**	**COR**
24-h avg. SO_4^{2-}	CASTNET	7.3	6.5	183	-11.2	18.3	-11.3	19.1	1.9	86.2
	IMPROVE	6.7	5.7	349	-15.0	33.2	-3.2	45.1	3.2	81.4
	SEARCH	5.8	7.8	986	31.2	76.7	194.2	218.2	6.2	26.8
	STN	7.7	6.1	1 016	-19.8	46.5	10.3	57.9	6.3	50.5
	AIRS-AQS	6.9	6.5	942	-6.5	37.4	19.8	58.3	3.7	76.4
24-h avg. NO_3^-	CASTNET	0.4	0.1	183	-71.2	85.4	-28.3	91.2	57.8	12.4
	IMPROVE	0.4	0.1	349	-65.3	90.8	-65.4	92.3	0.5	19.1
	SEARCH	0.6	0.1	1 444	-78.1	97.3	-69.3	100.3	91.9	7.9
	STN	2.2	0.3	938	-87.9	92.4	-68.2	87.1	6.1	4.1
	AIRS-AQS	0.9	0.3	964	-69.4	84.7	-67.3	85.7	1.0	27.3
24-h avg. NH_4^+	CASTNET	2.1	1.3	183	-39.5	42.2	-35.6	39.8	1.1	66.2
	SEARCH	2.1	1.8	1 412	-13.7	44.9	43.8	84.2	1.3	41.6
	STN	3.3	1.5	1 016	-52.5	64.4	33.8	96.3	5.8	20.4
	AIRS-AQS	2.2	1.6	942	-26.5	43.1	Inf	Inf	1.4	73.6
24-h avg. BC	IMPROVE	0.4	0.4	317	-13.4	68.9	Inf	Inf	0.5	48.3
	SEARCH	0.8	0.7	1 233	-10.7	62.4	56.5	106.0	0.8	46.6
24-h avg. OC	IMPROVE	2.7	1.3	317	-51.3	70.4	-42.3	78.6	3.5	53.5
	SEARCH	3.4	1.7	62	-51.6	51.3	-50.4	49.7	1.4	71.2
24-h avg. TC	SEARCH	5.0	2.1	1 256	-57.8	61.7	-50.6	60.9	4.0	36.8
	STN	9.1	3.4	98	-63.2	73.6	-40.8	80.3	11.7	48.6
	IMPROVE	3.1	1.7	317	-46.3	68.7	-37.4	76.5	3.8	53.2
	AIRS-AQS	7.5	3.2	931	-58.7	66.7	-44.4	70.7	8.0	58.4
SO_4^{2-}_WD	NADP	2.1	2.2	324	4.6	67.0	NaN	NaN	2.2	35.2
NO_3^-_WD	NADP	1.7	0.7	324	-55.6	72.8	NaN	NaN	1.6	21.9
NH_4^+_WD	NADP	0.4	0.2	324	-38.9	69.5	NaN	NaN	0.4	27.6

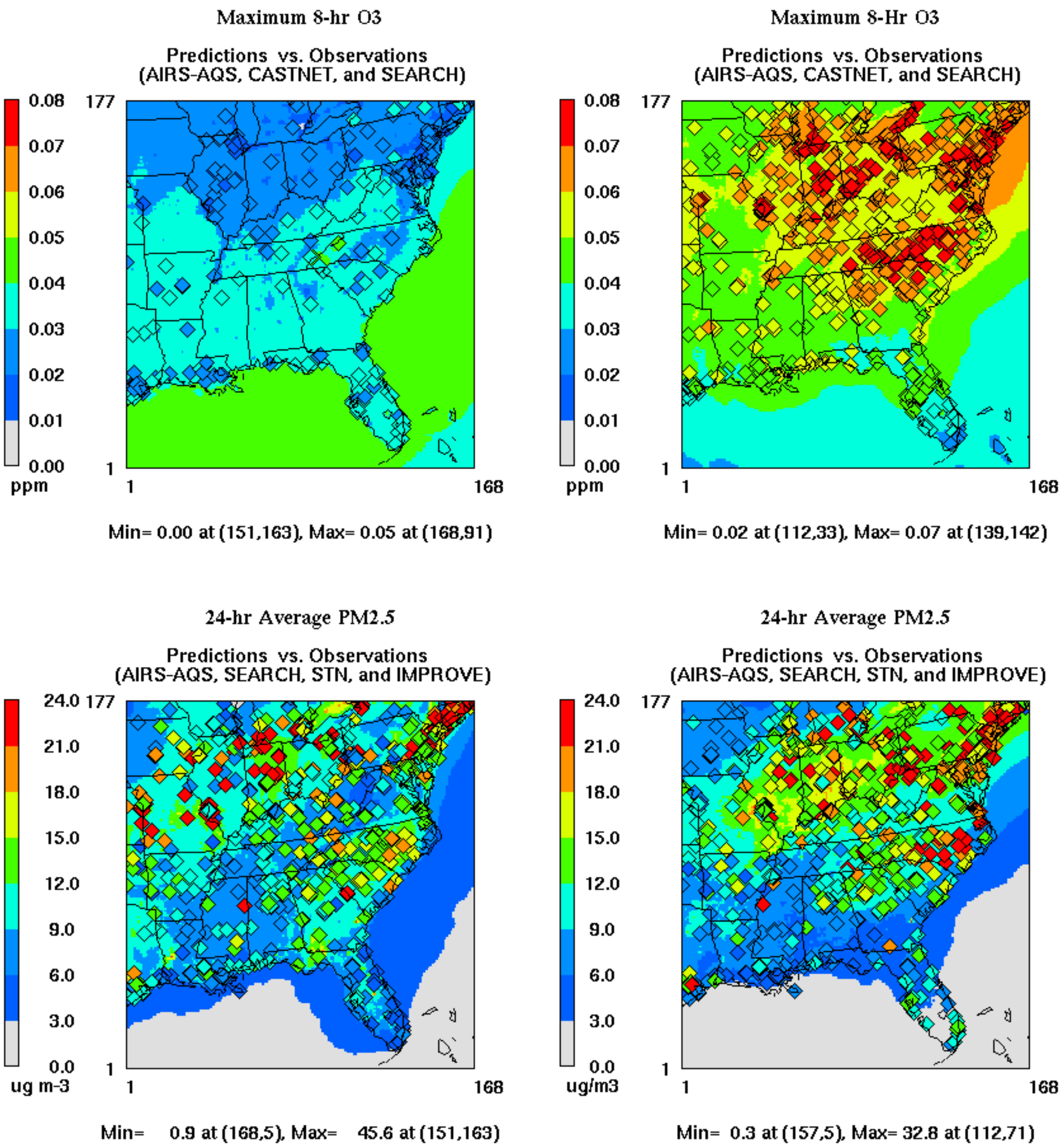

Figure S1. Spatial distribution monthly-mean 8-h O_3 (top) and 24-h average $PM_{2.5}$ (bottom) concentrations simulated by CMAQ overlaid with observations from AIRS-AQS, CASTNET, and SEARCH in January (left) and July (right).