MEMORANDUM

To: Jim DeMocker, EPA Office of Policy Analysis and Review (OPAR)

From: Sharon Douglas & Tom Myers, ICF International

Date: 24 November 2009

Re: Evaluation of CMAQ Model Performance for the 812 Prospective II Study

Introduction

Section 812 of the Clean Air Act Amendments (CAAA) of 1990 requires EPA to periodically assess the effects of the Clean Air Act (CAA) on air quality, the environment, public health, and the economy. EPA is currently developing the second prospective analyses of the benefits and costs of the CAA. As part of this study, the Community Multiscale Air Quality (CMAQ) model was applied to simulate national and regional-scale pollutant concentrations and deposition. The air quality modeling relied on tools and databases that had recently been developed and evaluated by EPA for other national- and regional-scale air quality modeling studies and was used to estimate the differences in air quality for 2000, 2010 and 2020 with and without the emissions reductions expected from the 1990 CAAA.

The methods and results of the emissions processing and air quality modeling that were conducted to support the development of the second prospective CAA Section 812 benefit-cost study are summarized by Douglas et al. (2008). For fine particulate matter ($PM_{2.5}$) and related species, the CMAQ model was applied for an annual simulation period (January through December). A 36-km resolution modeling domain that encompasses the contiguous 48 states was used for the annual modeling. For ozone and related species, the CMAQ model was applied for a five-month simulation period that captures the key ozone-season months of May through September. Two 12-km resolution modeling domains (that when combined cover the contiguous 48 U.S. states) were used for the ozone-season modeling. Altogether, the CMAQ model was applied for a total of 21 simulations (comprising seven core scenarios and three modeling domains). Model-ready meteorological input files for 2002 were provided by EPA.

The outputs from the CMAQ model provide the basis for the calculation of health and ecological benefits of the CAA. Ozone, $PM_{2.5}$, speciated $PM_{2.5}$, and PM_{10} concentrations and nitrogen and sulfur deposition amounts were extracted from the model outputs. Visibility was calculated using a variety of the CMAQ output species.

An integral component of all modeling studies is the evaluation of model performance for the base-case simulation. For this study, the Atmospheric Model Evaluation Tool (AMET) (UNC, 2008) was used to evaluate the CMAQ modeling results. The modeling results for ozone, $PM_{2.5}$ and other pollutant species were compared with observed data. The evaluation was done for the 2000 with-CAAA90 scenario, which represents the base year for the modeling analysis. The modeling results were compared with data for calendar year 2002, since this is the year represented by the meteorological inputs. The results of this evaluation are summarized in this memorandum.

Definitions of Key Statistical Measures

The AMET tool generates a wide variety of statistical measures and graphical analysis products to facilitate the evaluation of CMAQ model performance. Table 1 summarizes the statistical measures discussed in this memorandum. Additional statistical summaries/other measures are available in the AMET output files (upon request).

Table 1. Definition and Description of Measures/Metrics Used for Model Performance Evaluation

Metric	Definition
# of data pairs	*The number of observation/simulation data pairs*
Mean observation value	*The average observed concentration*
Mean simulation value	*The average simulated concentration*
Mean bias	$\left(\frac{1}{N}\right)\sum_{l=1}^{N}(S_l - O_l)$ *where N is the number of data pairs, and S_l and O_l are the simulated and observed values at site l, respectively, over a given time interval.*
Normalized bias	$\left(\frac{1}{N}\right)\sum_{l=1}^{N}(S_l - O_l)/O_l \cdot 100\%$
Normalized mean bias	$\sum_{l=1}^{N}(S_l - O_l)/\sum_{l=1}^{N}O_l \cdot 100\%$
Fractional bias	$\left(\frac{1}{N}\right)\sum_{l=1}^{N}(S_l - O_l)/0.5(S_l + O_l)\cdot 100\%$
Mean error	$\left(\frac{1}{N}\right)\sum_{l=1}^{N}\lvert S_l - O_l\rvert$
Normalized error	$\left(\frac{1}{N}\right)\sum_{l=1}^{N}\lvert S_l - O_l\rvert/O_l \cdot 100\%$
Normalized mean error	$\sum_{l=1}^{N}\lvert S_l - O_l\rvert\Big/\sum_{l=1}^{N}O_l \cdot 100\%$
Fractional error	$\left(\frac{1}{N}\right)\sum_{l=1}^{N}\lvert S_l - O_l\rvert/0.5(S_l + O_l)\cdot 100\%$
Correlation	$\left(N\left(\sum SO\right)-\left(\sum S\right)\left(\sum O\right)\right)\Big/\sqrt{\left(N\sum S^2-\left(\sum S\right)^2\right)\left(N\sum O^2-\left(\sum O\right)^2\right)}$
Index of agreement	*A measure of how well the model represents the pattern of perturbation about the mean value; ranges from 0 to 1.*

In calculating the statistical measures, AMET pairs the CMAQ model output with the observed data for the appropriate locations and time intervals.

For this analysis, ozone data were extracted from the EPA Air Quality System (AQS) dataset. Statistics were calculated using hourly concentrations, daily maximum 1-hour concentrations, and daily maximum 8-hour average concentrations. This evaluation focuses on model performance for daily maximum 8-hour average ozone, since this is the primary ozone metric used in the cost-benefit analysis.

The $PM_{2.5}$ dataset used for this analysis consists of data from the Speciation Trends Network (STN), Interagency Monitoring of PROtected Visual Environments (IMPROVE), and Clean Air Status and Trends Network (CASTNet) monitoring networks. Statistical measures were calculated using paired daily average values of $PM_{2.5}$ and selected species for the STN and IMPROVE data, and weekly average values of selected species for the CASTNet data. This evaluation focuses on model performance for total $PM_{2.5}$, since this is the primary $PM_{2.5}$ metric used in the cost-benefit analysis.

Finally, deposition measurements from the National Acid Deposition Program (NADP) were used in the evaluation of deposition for selected species. In this case, the weekly measurements were matched with the appropriate time intervals from the model output.

For extraction of the model output and matching with the station values, concentration and deposition information were taken from the grid cell in which the monitoring site is located.

Summary of Model Performance for Ozone

CMAQ model performance for ozone was examined for the 12-km Eastern U.S. (EUS) and Western U.S. (WUS) modeling domains for each month and for the entire ozone-season simulation period.

Eastern U.S. (EUS)

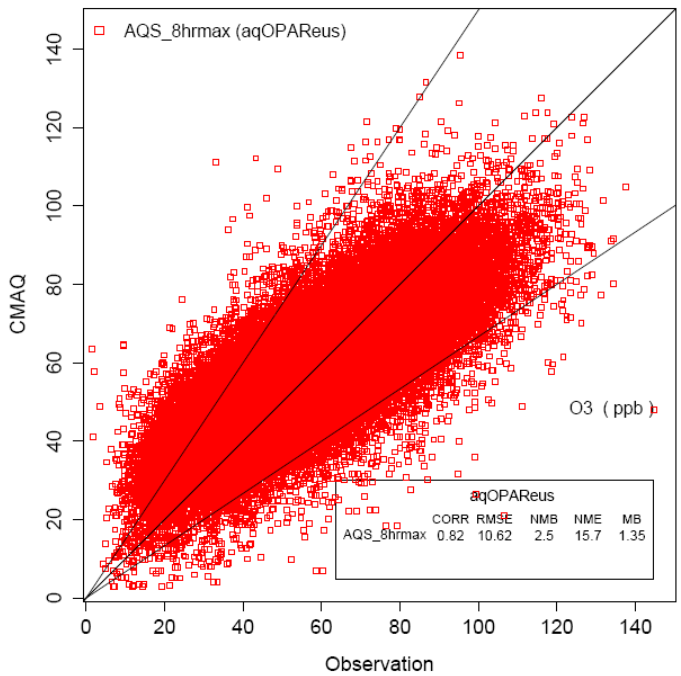
Summary metrics and statistical measures for 8-hour ozone for the EUS domain ozone are presented in Table 2. For certain of the measures, model performance goals and criteria used for prior studies are provided for comparison. For ozone, the recommended ranges for the normalized bias and normalized error are from prior EPA guidance but are still widely used for urban- and regional-scale model performance evaluation (EPA, 2007).

Table 2. Summary Model Performance Statistics for Ozone for the Eastern U.S. (EUS) Modeling Domain: Daily Maximum 8-Hour Average Ozone.

	May	Jun	Jul	Aug	Sep	O3 Season	Goal
No. of Observations	22431	21724	22626	22816	21312	110909	
Mean Observed (ppb)	48.4	56.2	56.6	54.4	48.2	52.8	
Mean Simulated (ppb)	52.9	56.3	57.8	54.2	49.2	54.1	
Normalized Bias (%)	13.4	6.0	7.5	5.7	10.2	8.6	± 15
Normalized Error (%)	19.2	17.2	18.1	17.8	21.6	18.8	≤ 35
Fractional Bias (%)	10.4	3.3	4.7	2.9	5.9	5.4	
Fractional Error (%)	16.8	15.6	16.4	16.4	18.7	16.8	
Correlation (unitless)	0.71	0.85	0.79	0.85	0.83	0.82	
Index of Agreement (unitless)	0.79	0.9	0.87	0.9	0.89	0.88	

Graphical summaries can also provide information about model performance. Simulated daily maximum 8-hour ozone concentrations for each site in the EUS domain are compared in the scatter plot in Figure 1. The scatter plot displays the values for the entire ozone season and visually depicts the range and frequency of agreement represented by the individual observation-simulation pairs included in the calculation of the cumulative statistical measures. Also included on the scatter plot is some statistical information further summarizing model performance.

Figure 1. Comparison of Simulated and Observed Daily Maximum 8-Hour Average Ozone (ppb) for the Eastern U.S. (EUS) Modeling Domain for the Ozone-Season Simulation Period.



The bar charts in Figure 2 summarize the month-to-month variations in the model performance for ozone for the EUS. The mean observed and simulated values for each month and the entire ozone season simulation period are graphically compared in Figure 2a. The normalized bias and error are graphically displayed in Figure 2b.

Figure 2a. Comparison of Mean Observed and Simulated Daily Maximum 8-Hour Average Ozone (ppb) for Each Month and the Entire Ozone Season: EUS Domain.

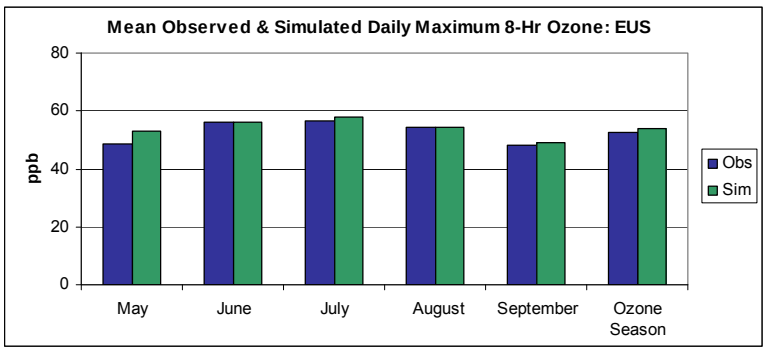


Figure 2b. Normalized Bias (%) and Error (%) for Simulated Daily Maximum 8-Hour Average Ozone for Each Month and the Entire Ozone Season: EUS Domain.

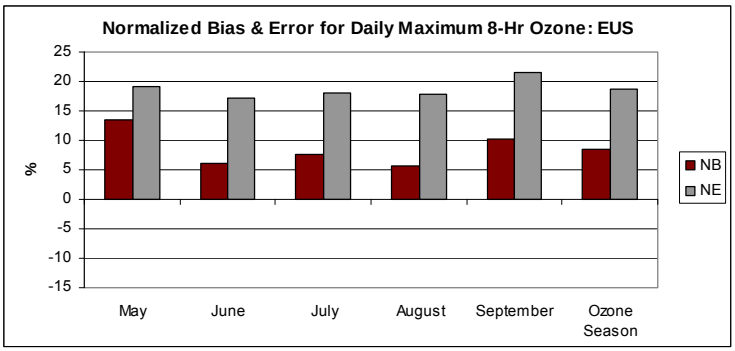
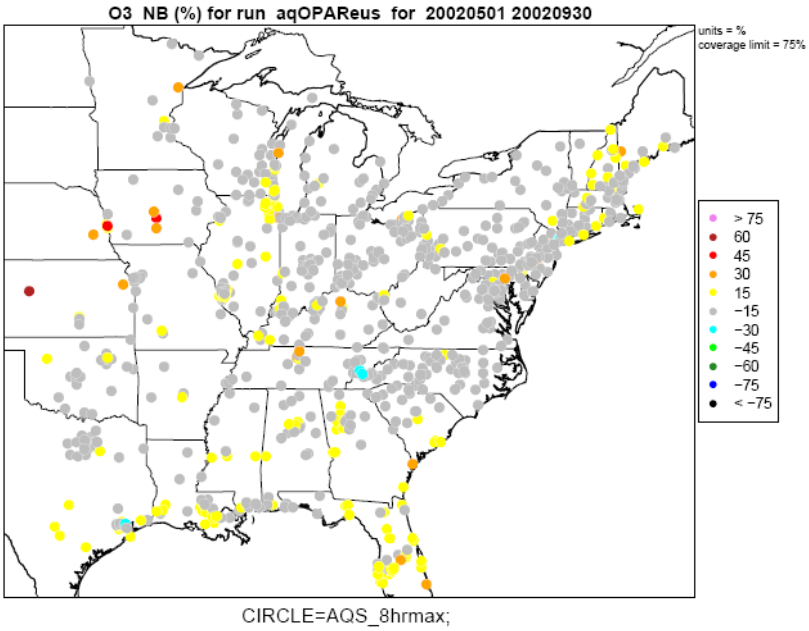


Figure 3 graphically displays spatial variations in model performance, based on the sign and magnitude of the normalized bias. Each dot represents a monitoring site location, and the color of the dot indicates the value of the normalized bias for that site for the entire ozone-season simulation period. Gray dots correspond to a normalized bias within ± 15 percent. Yellow dots indicate an overestimation of daily maximum 8-hour ozone by 15 to 30 percent, on average.

Figure 3. Normalized Bias (%) for Simulated Daily Maximum 8-Hour Average Ozone for the Entire Ozone Season: EUS Domain.



Western U.S. (WUS)

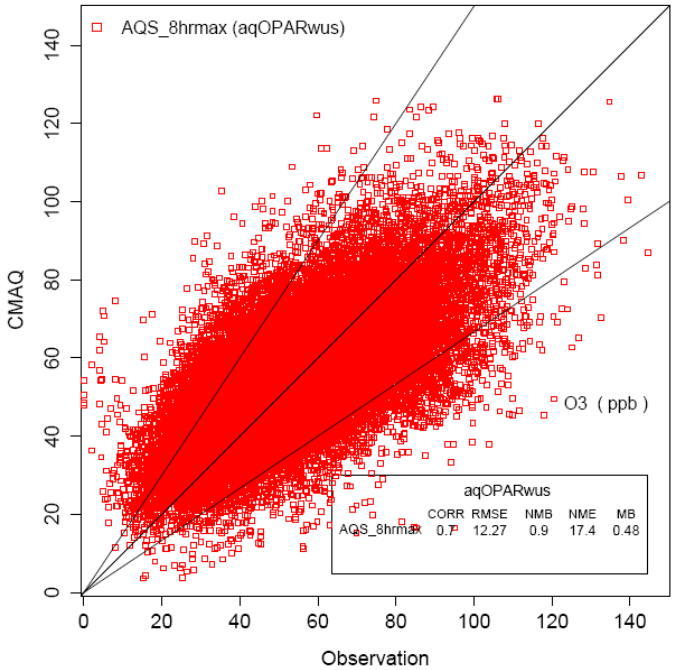
Summary metrics and statistical measures for the WUS domain ozone are presented in Table 3.

Table 3. Summary Model Performance Statistics for Ozone for the Western U.S. (WUS) Modeling Domain: Daily Maximum 8-Hour Average Ozone.

	May	Jun	Jul	Aug	Sep	O3 Season	Goal
No. of Observations	9490	9252	9564	9668	8987	46961	
Mean Observed (ppb)	52.1	56.0	55.5	56.8	50.7	54.3	
Mean Simulated (ppb)	54.2	55.5	53.7	57.2	50.2	54.2	
Normalized Bias (%)	7.7	4.4	2.5	7.6	5.0	5.5	± 15
Normalized Error (%)	17.0	18.1	21.0	23.1	21.4	20.1	≤ 35
Fractional Bias (%)	5.0	1.7	-1.4	2.4	0.9	1.7	
Fractional Error (%)	15.2	17.1	20.0	19.9	19.8	18.4	
Correlation (unitless)	0.69	0.75	0.69	0.67	0.66	0.69	
Index of Agreement (unitless)	0.81	0.84	0.82	0.81	0.8	0.82	

Simulated daily maximum 8-hour ozone concentrations for each site in the WUS domain are compared in the scatter plots in Figure 4.

Figure 4. Comparison of Simulated and Observed Daily Maximum 8-Hour Average Ozone (ppb) for the Western U.S. (WUS) Modeling Domain for the Ozone-Season Simulation Period.



The bar charts in Figure 5 summarize the month-to-month variations in the model performance for ozone for the WUS. The mean observed and simulated values for each month and the entire ozone season simulation period are graphically compared in Figure 5a. The normalized bias and error are displayed in Figure 5b.

Figure 5a. Comparison of Mean Observed and Simulated Daily Maximum 8-Hour Average Ozone (ppb) for Each Month and the Entire Ozone Season: WUS Domain.

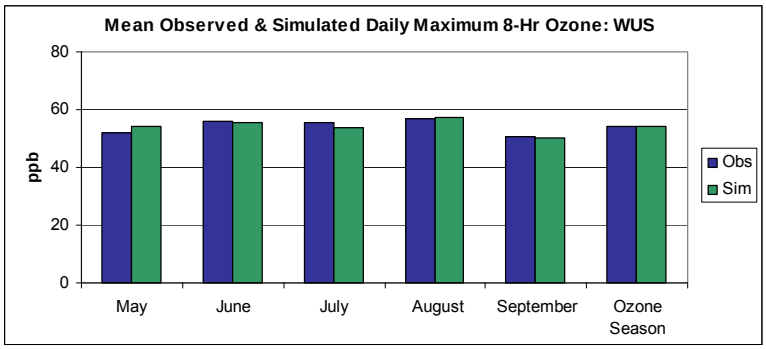


Figure 5b. Normalized Bias (%) and Error (%) for Simulated Daily Maximum 8-Hour Average Ozone for Each Month and the Entire Ozone Season: WUS Domain.

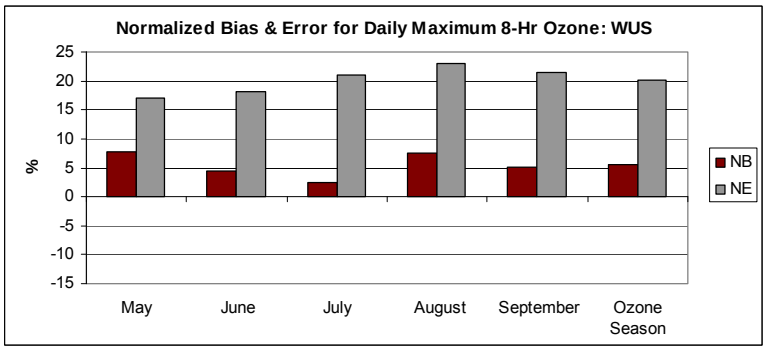
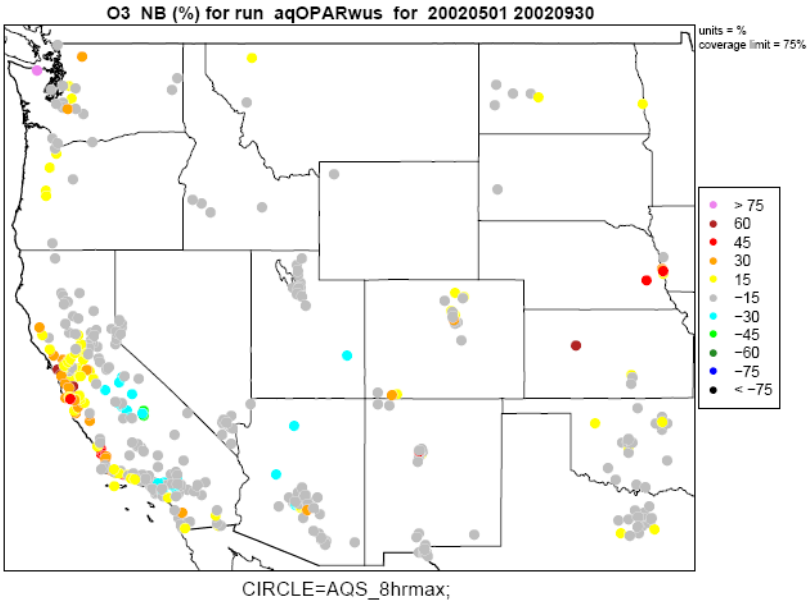


Figure 6 displays the regional distribution of the normalized bias for the WUS domain. Again gray dots correspond to a normalized bias within ± 15 percent and yellow dots indicate an overestimation of daily maximum 8-hour ozone by 15 to 30 percent, on average.

Figure 6. Normalized Bias (%) for Simulated Daily Maximum 8-Hour Average Ozone for the Entire Ozone Season: WUS Domain.



Summary of Model Performance for PM2.5

CMAQ model performance for $PM_{2.5}$ and selected species was examined for the full continental U.S. (CONUS) modeling domain and for two subregions representing the eastern and western U.S. Performance measures were calculated for each calendar quarter and for the entire annual simulation period. To accommodate differences in the measured species, measurement techniques and measurement frequency among the STN, IMPROVE and CASTNet datasets, statistics were calculated separately for each dataset.

STN

Summary metrics and statistical measures calculated using STN data for $PM_{2.5}$ and selected species for the full CONUS domain and the eastern and western portions of the domain are presented in Tables 4 and 5. Table 4 provides annual model performance results for $PM_{2.5}$ and selected species including sulfate (SO4), nitrate (NO3), ammonium (NH4), organic carbon (OC), and elemental carbon (EC). Table 5 provides quarterly and annual model performance results for $PM_{2.5}$ (the calendar quarters are defined as: January – March, April – June, July – September, and October – December). For certain of the measures, model performance goals and criteria used for prior studies are provided for comparison. The results of a number of different studies are listed in the EPA modeling guidance document (EPA, 2007). Few of the studies also list goals and/or criteria. For this analysis, we selected the goals presented by Boylan (2005) for comparison.

Table 4a. Summary Model Performance Statistics for $PM_{2.5}$ and Selected Species for the Continental U.S. (CONUS) Modeling Domain: 24-Hour Average $PM_{2.5}$ (STN).

	SO4	NO3	NH4	OC	EC	PM2.5	Goal
No. of Observations	12069	10682	12069	11653	11972	12257	
Mean Observed (μgm^{-3})	3.5	2.0	1.5	3.3	0.7	12.9	
Mean Simulated (μgm^{-3})	3.5	2.2	1.7	1.7	0.8	14.4	
Normalized Mean Bias (%)	-1.3	11.4	14.9	-47.5	22.2	11.3	
Normalized Mean Error (%)	36.9	78.3	51.4	61.0	66.2	48.5	
Fractional Bias (%)	-5.8	-21.6	19.8	-51.0	12.2	7.2	± 30
Fractional Error (%)	40.7	87.6	53.0	76.0	57.3	46.3	≤ 50
Correlation (unitless)	0.81	0.49	0.63	0.33	0.41	0.51	
Index of Agreement (unitless)	0.9	0.68	0.78	0.49	0.61	0.7	

Table 4b. Summary Model Performance Statistics for $PM_{2.5}$ and Selected Species for the Eastern Portion of the CONUS Modeling Domain: 24-Hour Average $PM_{2.5}$ (STN).

	SO4	NO3	NH4	OC	EC	PM2.5	Goal
No. of Observations	9900	8572	9900	9471	9775	10047	
Mean Observed (μgm^{-3})	3.9	1.7	1.6	2.9	0.6	12.8	
Mean Simulated (μgm^{-3})	4.0	2.4	1.9	1.6	0.8	15.2	
Normalized Mean Bias (%)	1.1	41.1	24.3	-46.3	28.2	18.9	
Normalized Mean Error (%)	35.9	81.3	48.8	59.7	66.7	47.8	
Fractional Bias (%)	-3.4	-10.1	23.5	-50.0	15.9	12.1	± 30
Fractional Error (%)	39.3	84.7	50.0	75.8	55.8	45.3	≤ 50
Correlation (unitless)	0.8	0.7	0.73	0.32	0.44	0.58	
Index of Agreement (unitless)	0.89	0.77	0.83	0.5	0.62	0.73	

Table 4c. Summary Model Performance Statistics for $PM_{2.5}$ and Selected Species for the Western Portion of the CONUS Modeling Domain: 24-Hour Average $PM_{2.5}$ (STN).

	SO4	NO3	NH4	OC	EC	PM2.5	Goal
No. of Observations	2169	2110	2169	2182	2197	2210	
Mean Observed (μgm^{-3})	1.5	3.2	1.3	4.7	0.9	13.6	
Mean Simulated ((μgm^{-3}))	1.1	1.5	0.8	2.3	0.9	10.7	
Normalized Mean Bias (%)	-29.4	-52.5	-36.0	-51.0	3.2	-21.0	
Normalized Mean Error (%)	48.0	71.8	65.8	64.6	64.8	51.6	
Fractional Bias (%)	-17.0	-68.3	3.2	-55.5	-4.4	-15.3	± 30
Fractional Error (%)	47.1	99.3	66.4	76.6	63.9	50.8	≤ 50
Correlation (unitless)	0.5	0.52	0.52	0.26	0.32	0.42	
Index of Agreement (unitless)	0.56	0.56	0.54	0.47	0.56	0.6	

Table 5a. Summary Model Performance Statistics for $PM_{2.5}$ by Quarter for the Continental U.S. (CONUS) Modeling Domain: 24-Hour Average $PM_{2.5}$ (STN).

	Q1	Q2	Q3	Q4	Annual	Goal
No. of Observations	2562	2988	3444	3263	12257	
Mean Observed (μgm^{-3})	12.1	11.8	14.9	12.6	12.9	
Mean Simulated (μgm^{-3})	17.4	12.4	13.0	15.3	14.4	
Normalized Mean Bias (%)	44.2	4.9	-12.5	21.9	11.3	
Normalized Mean Error (%)	69.0	42.5	34.5	55.8	48.5	
Fractional Bias (%)	32.8	0.0	-14.3	16.3	7.2	± 30
Fractional Error (%)	54.9	43.1	39.7	49.5	46.3	≤ 50
Correlation (unitless)	0.44	0.58	0.66	0.49	0.51	
Index of Agreement (unitless)	0.58	0.76	0.79	0.67	0.7	

Table 5b. Summary Model Performance Statistics for $PM_{2.5}$ by Quarter for the Eastern Portion of the CONUS Modeling Domain: 24-Hour Average $_{PM2.5}$ (STN).

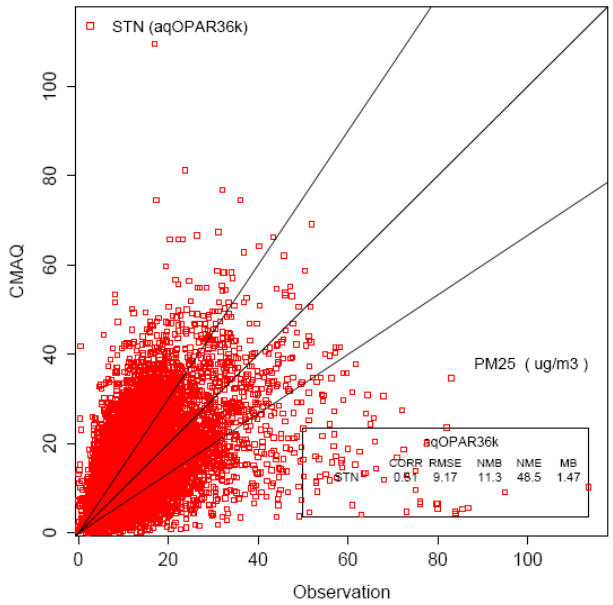
	Q1	Q2	Q3	Q4	Annual	Goal
No. of Observations	2089	2462	2815	2681	10047	
Mean Observed (μgm^{-3})	11.4	12.2	15.7	11.3	12.8	
Mean Simulated (μgm^{-3})	18.4	13.1	14.0	15.9	15.2	
Normalized Mean Bias (%)	62.5	7.3	-11.1	40.2	18.9	
Normalized Mean Error (%)	71.9	41.1	33.3	56.9	47.8	
Fractional Bias (%)	41.7	1.3	-12.7	25.2	12.1	± 30
Fractional Error (%)	54.2	42.1	38.7	48.3	45.3	≤ 50
Correlation (unitless)	0.67	0.62	0.66	0.7	0.58	
Index of Agreement (unitless)	0.62	0.77	0.8	0.74	0.73	

Table 5c. Summary Model Performance Statistics for $PM_{2.5}$ by Quarter for the Western Portion of the CONUS Modeling Domain: 24-Hour Average $PM_{2.5}$ (STN).

	Q1	Q2	Q3	Q4	Annual	Goal
No. of Observations	473	526	629	582	2210	
Mean Observed (μgm^{-3})	15.4	9.6	11.2	18.5	13.6	
Mean Simulated (μgm^{-3})	13.0	8.6	8.8	13.0	10.7	
Normalized Mean Bias (%)	-15.7	-9.9	-21.6	-29.6	-21.0	
Normalized Mean Error (%)	59.6	50.5	42.4	52.6	51.6	
Fractional Bias (%)	-6.5	-5.9	-21.2	-24.6	-15.3	± 30
Fractional Error (%)	58.2	47.6	44.2	54.9	50.8	≤ 50
Correlation (unitless)	0.34	0.35	0.42	0.39	0.42	
Index of Agreement (unitless)	0.55	0.57	0.61	0.58	0.6	

Graphical summaries for $PM_{2.5}$ using the STN data follow. Simulated annual average $PM_{2.5}$ concentrations for each site in the CONUS domain are compared in the scatter plot in Figure 7.

Figure 7. Comparison of Simulated and Observed 24-Hour Average $PM_{2.5}$ (μgm^{-3}) for the Continental U.S. (CONUS) Modeling Domain for the Annual Simulation Period (STN).



The bar charts in Figure 8 and 9 summarize the variations in the model performance by species. Observed and simulated annual average values for selected species and total $PM_{2.5}$ are graphically compared in Figure 8. Fractional bias and error and are graphically displayed in Figure 9.

Figure 8a. Comparison of Observed and Simulated Annual Average $PM_{2.5}$ and Selected Species (μgm^{-3}) for the CONUS Domain (STN).

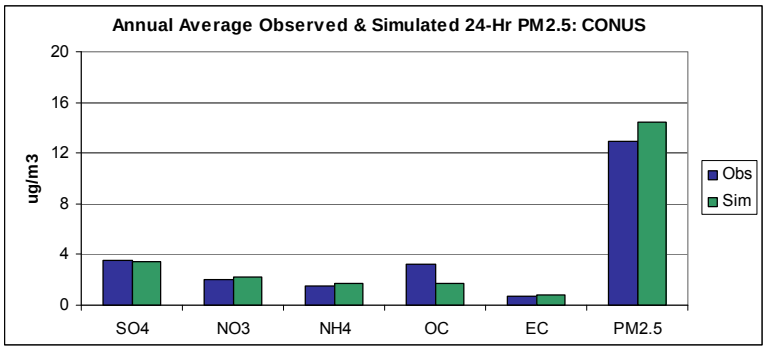


Figure 8b. Comparison of Observed and Simulated Annual Average $PM_{2.5}$ and Selected Species (μgm^{-3}) for the Eastern Portion of the CONUS Domain (STN).

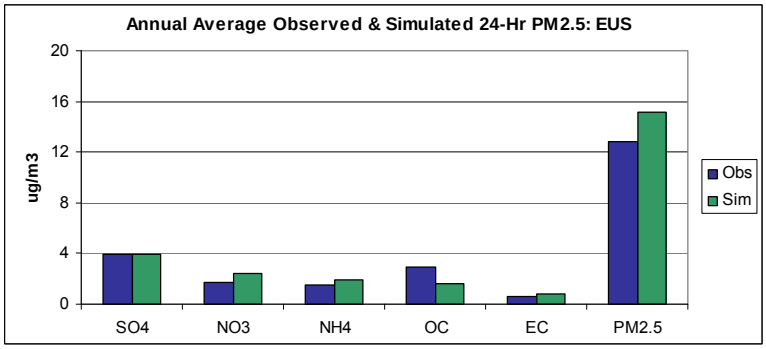


Figure 8c. Comparison of Observed and Simulated Annual Average $PM_{2.5}$ and Selected Species (μgm^{-3}) for the Western Portion of the CONUS Domain (STN).

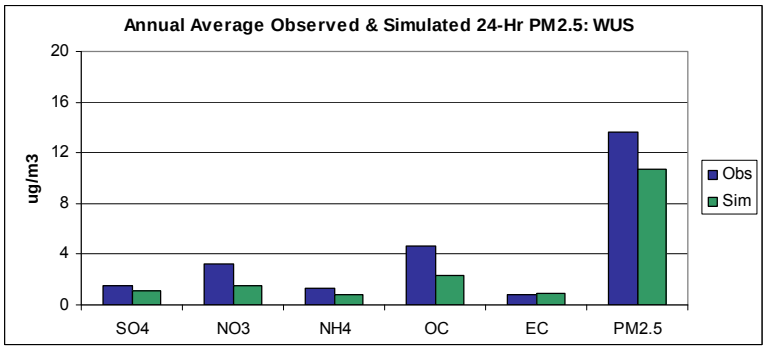


Figure 9a. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ and Selected Species for the CONUS Domain for the Annual Simulation Period (STN).

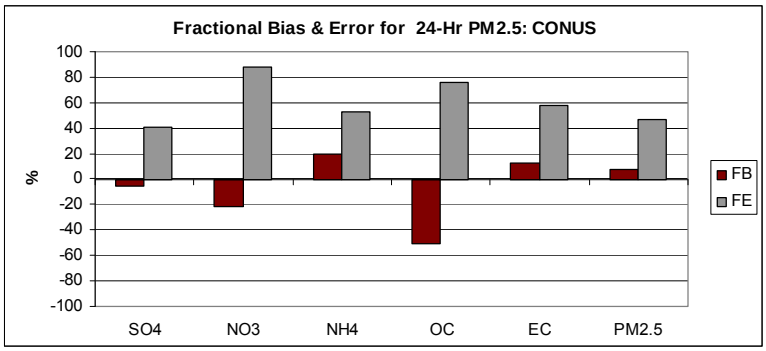


Figure 9b. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ and Selected Species for the Eastern Portion of CONUS Domain for the Annual Simulation Period (STN).

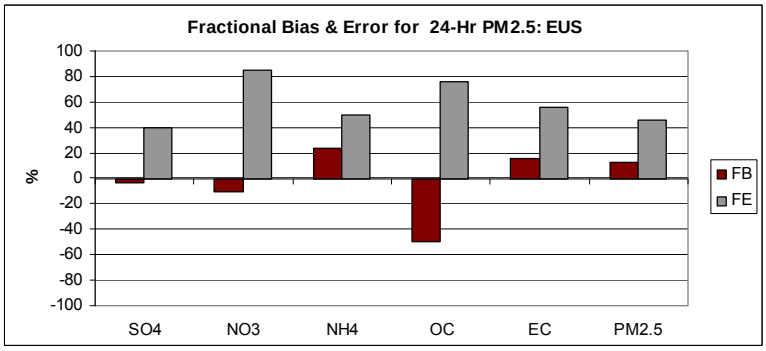
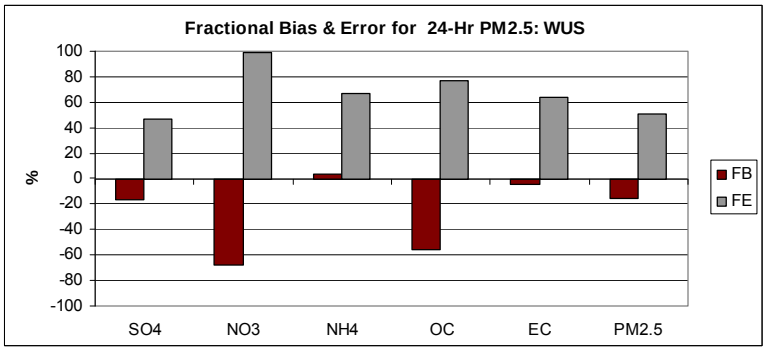


Figure 9c. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ and Selected Species for the Western Portion of CONUS Domain for the Annual Simulation Period (STN).



The bar charts in Figures 10 and 11 summarize the variations in $PM_{2.5}$ model performance by quarter. The mean observed and simulated values for each quarter and for the annual simulation period are graphically compared in Figure 10. Corresponding fractional bias and error values are graphically displayed in Figure 11.

Figure10a. Comparison of Observed and Simulated Average $PM_{2.5}$ (μgm^{-3}) for Each Quarter and the Annual Simulation Period for the CONUS Domain (STN).

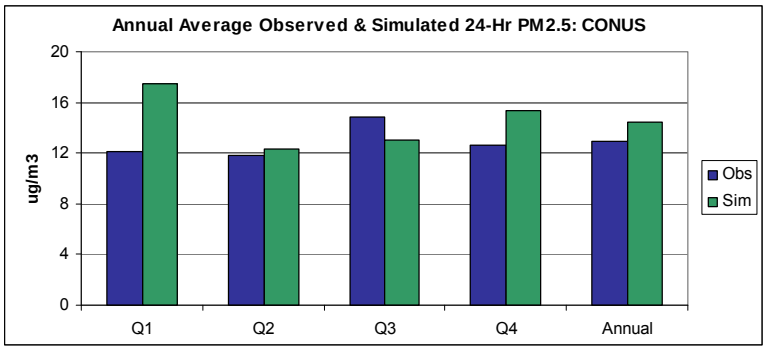


Figure 10b. Comparison of Observed and Simulated Average $PM_{2.5}$ (μgm^{-3}) for Each Quarter and the Annual Simulation Period for the Eastern Portion of the CONUS Domain (STN).

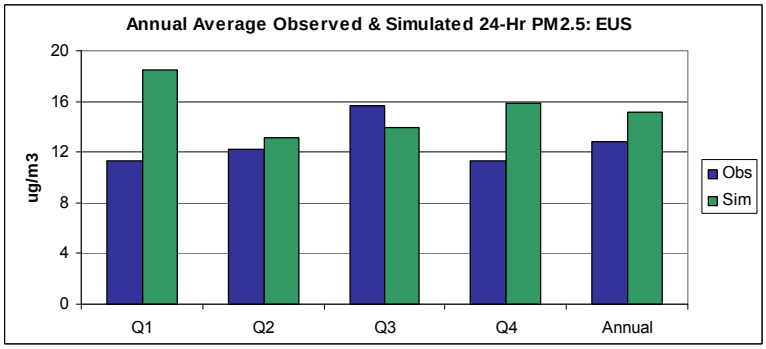


Figure 10c. Comparison of Observed and Simulated Average $PM_{2.5}$ (μgm^{-3}) for Each Quarter and the Annual Simulation Period for the Western Portion of the CONUS Domain (STN).

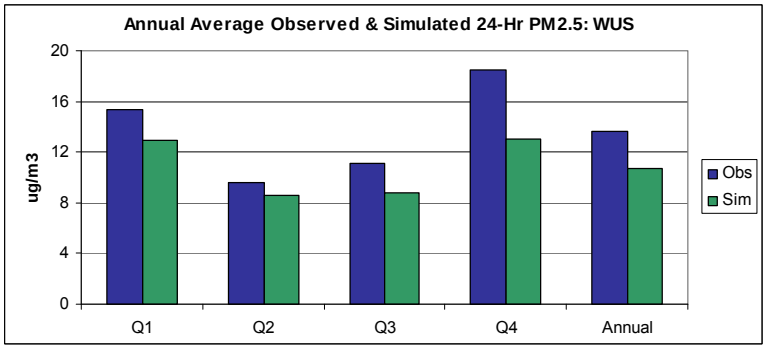


Figure 11a. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ (μgm^{-3}) for Each Quarter and the Annual Simulation Period for the CONUS Domain (STN).

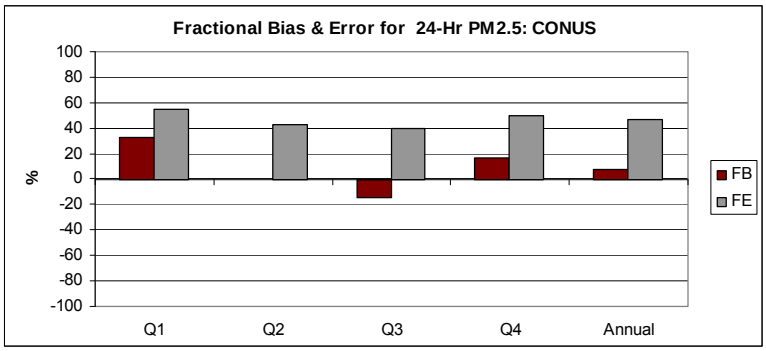


Figure 11b. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ (μgm^{-3}) for Each Quarter and the Annual Simulation Period for the Eastern Portion of the CONUS Domain (STN).

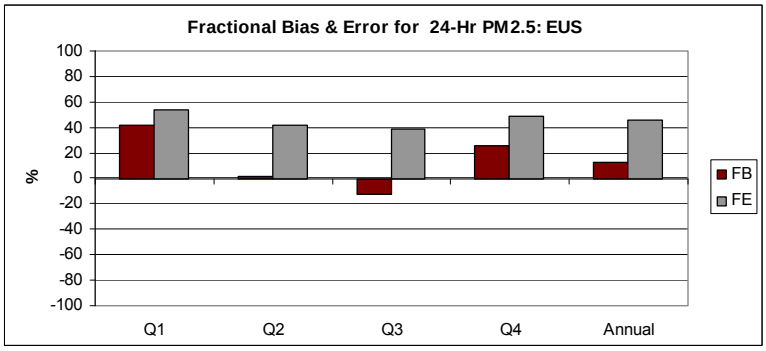


Figure 11c. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ (μgm^{-3}) for Each Quarter and the Annual Simulation Period for the Western Portion of the CONUS Domain (STN).

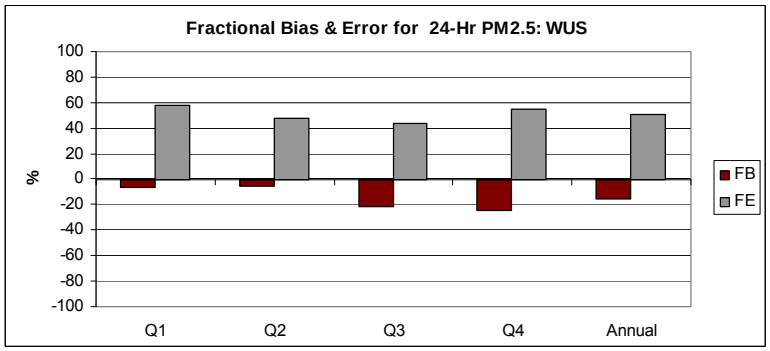
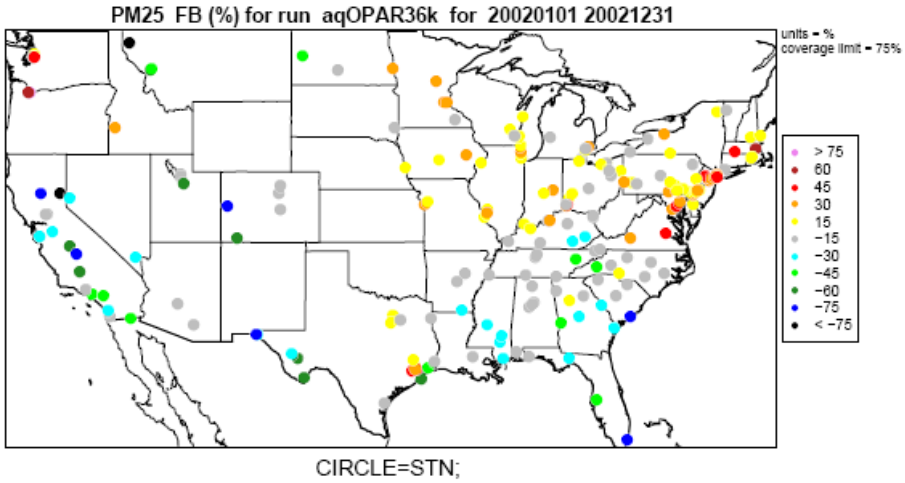


Figure 12 displays the regional distribution of the fractional bias for $PM_{2.5}$ for the CONUS domain. In this plot, gray dots correspond to a fractional bias within ± 15 percent, on average. Light blue dots correspond to a negative fractional bias within 15 to 30 percent. Yellow dots correspond to a positive fractional bias within 15 to 30 percent.

Figure 12. Fractional Bias (%) for Simulated $PM_{2.5}$ for the Annual Simulation Period for the CONUS Domain (STN).



IMPROVE

Summary metrics and statistical measures calculated using IMPROVE data for $PM_{2.5}$ and selected species for the full CONUS domain and the eastern and western portions of the domain are presented in Tables 6 and 7. Table 6 provides annual model performance results for SO4, NO3, NH4, OC, EC, and $PM_{2.5}$. Table 7 provides quarterly and annual model performance results for $PM_{2.5}$. Again, goals presented by Boylan (2005) are also listed for comparison.

Table 6a. Summary Model Performance Statistics for $PM_{2.5}$ and Selected Species for the Continental U.S. (CONUS) Modeling Domain: 24-Hour Average $PM_{2.5}$ (IMPROVE).

	SO4	NO3	NH4	OC	EC	PM2.5	Goal
No. of Observations	15788	15663	718	15494	15475	15622	
Mean Observed (μgm^{-3})	1.6	0.5	1.1	1.3	0.3	5.9	
Mean Simulated (μgm^{-3})	1.7	0.8	1.4	1.3	0.3	6.9	
Normalized Mean Bias (%)	1.5	51.4	27.9	-5.4	16.5	16.6	
Normalized Mean Error (%)	39.4	119.6	54.2	70.9	71.2	61.8	
Fractional Bias (%)	13.3	-43.1	21.6	1.8	6.9	12.7	± 30
Fractional Error (%)	47.2	114.9	52.0	65.6	58.9	55.6	≤ 50
Correlation (unitless)	0.87	0.59	0.67	0.42	0.44	0.54	
Index of Agreement (unitless)	0.93	0.68	0.78	0.58	0.62	0.7	

Table 6b. Summary Model Performance Statistics for $PM_{2.5}$ and Selected Species for the Eastern Portion of the CONUS Modeling Domain: 24-Hour Average $PM_{2.5}$ (IMPROVE).

	SO4	NO3	NH4	OC	EC	PM2.5	Goal
No. of Observations	6831	6826	718	6667	6663	6761	
Mean Observed (μgm^{-3})	2.9	0.7	1.1	1.5	0.3	8.1	
Mean Simulated (μgm^{-3})	2.9	1.3	1.4	1.1	0.3	9.8	
Normalized Mean Bias (%)	-0.9	88.1	27.9	-21.5	5.8	21.7	
Normalized Mean Error (%)	36.0	135.0	54.2	63.0	53.0	61.1	
Fractional Bias (%)	-0.7	-22.7	21.6	-26.3	-0.9	13.1	± 30
Fractional Error (%)	41.2	112.6	52.0	67.8	50.6	55.6	≤ 50
Correlation (unitless)	0.84	0.66	0.67	0.22	0.53	0.52	
Index of Agreement (unitless)	0.91	0.66	0.78	0.4	0.7	0.69	

Table 6c. Summary Model Performance Statistics for $PM_{2.5}$ and Selected Species for the Western Portion of the CONUS Modeling Domain: 24-Hour Average $PM_{2.5}$ (IMPROVE).

	SO4	NO3	NH4	OC	EC	PM2.5	Goal
No. of Observations	8957	8837	0	8827	8812	8861	
Mean Observed (μgm^{-3})	0.6	0.4	NA	1.2	0.2	4.2	
Mean Simulated (μgm^{-3})	0.7	0.4	NA	1.3	0.3	4.6	
Normalized Mean Bias (%)	9.7	2.5	NA	9.0	28.8	9.3	
Normalized Mean Error (%)	50.9	99.1	NA	78.0	91.9	62.8	
Fractional Bias (%)	24.0	-58.8	NA	23.0	12.8	12.3	± 30
Fractional Error (%)	51.7	116.6	NA	63.9	65.1	55.6	≤ 50
Correlation (unitless)	0.5	0.51	NA	0.47	0.41	0.36	
Index of Agreement (unitless)	0.68	0.68	NA	0.62	0.59	0.56	

Table 7a. Summary Model Performance Statistics for $PM_{2.5}$ by Quarter for the Continental U.S. (CONUS) Modeling Domain: 24-Hour Average $PM_{2.5}$ (IMPROVE).

	Q1	Q2	Q3	Q4	Annual	Goal
No. of Observations	3681	3746	4047	4148	15622	
Mean Observed (μgm^{-3})	4.2	6.4	8.2	4.7	5.9	
Mean Simulated (μgm^{-3})	7.4	6.1	6.7	7.3	6.9	
Normalized Mean Bias (%)	77.5	-5.3	-18.2	54.1	16.6	
Normalized Mean Error (%)	95.6	46.8	46.5	79.3	61.8	
Fractional Bias (%)	47.0	-8.4	-23.1	36.2	12.7	± 30
Fractional Error (%)	64.5	47.6	50.3	60.2	55.6	≤ 50
Correlation (unitless)	0.69	0.6	0.54	0.63	0.54	
Index of Agreement (unitless)	0.67	0.77	0.72	0.69	0.7	

Table 7b. Summary Model Performance Statistics for $PM_{2.5}$ by Quarter for the Eastern Portion of the CONUS Modeling Domain: 24-Hour Average $PM_{2.5}$ (IMPROVE).

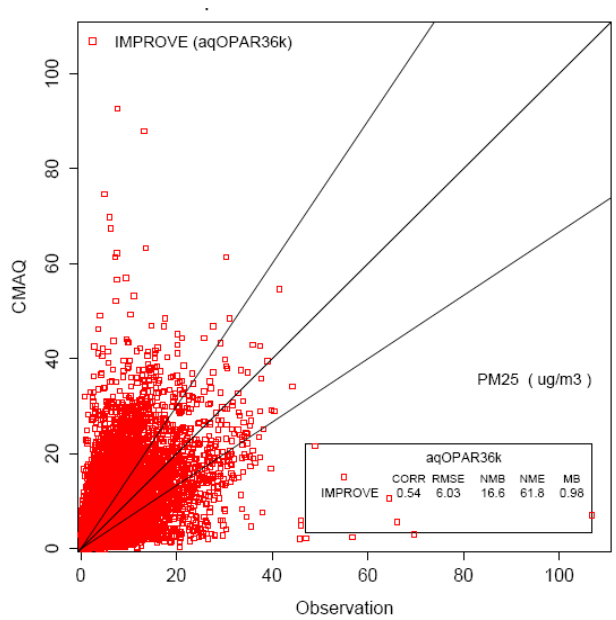
	Q1	Q2	Q3	Q4	Annual	Goal
No. of Observations	1573	1563	1767	1858	6761	
Mean Observed (μgm^{-3})	6.5	8.5	11.2	6.0	8.1	
Mean Simulated (μgm^{-3})	12.1	8.1	8.6	10.4	9.8	
Normalized Mean Bias (%)	87.1	-4.8	-23.1	72.5	21.7	
Normalized Mean Error (%)	100.2	43.4	38.7	86.1	61.1	
Fractional Bias (%)	49.3	-12.1	-30.3	45.0	13.1	± 30
Fractional Error (%)	64.3	47.1	48.9	61.6	55.6	≤ 50
Correlation (unitless)	0.6	0.64	0.7	0.65	0.52	
Index of Agreement (unitless)	0.55	0.79	0.8	0.65	0.69	

Table 7c. Summary Model Performance Statistics for $PM_{2.5}$ by Quarter for the Western Portion of the CONUS Modeling Domain: 24-Hour Average $PM_{2.5}$ (IMPROVE).

	Q1	Q2	Q3	Q4	Annual	Goal
No. of Observations	2108	2183	2280	2290	8861	
Mean Observed (μgm^{-3})	2.5	4.9	5.8	3.7	4.2	
Mean Simulated (μgm^{-3})	3.9	4.6	5.2	4.8	4.6	
Normalized Mean Bias (%)	58.8	-5.8	-11.0	29.6	9.3	
Normalized Mean Error (%)	86.7	51.0	58.0	70.3	62.8	
Fractional Bias (%)	45.3	-5.8	-17.5	29.0	12.3	± 30
Fractional Error (%)	64.5	47.9	51.4	59.0	55.6	≤ 50
Correlation (unitless)	0.53	0.35	0.18	0.53	0.36	
Index of Agreement (unitless)	0.64	0.58	0.38	0.68	0.56	

Graphical summaries for $PM_{2.5}$ using the IMPROVE data follow. Simulated annual average $PM_{2.5}$ concentrations for each site in the CONUS domain are compared in the scatter plots in Figure 13.

Figure 13. Comparison of Simulated and Observed 24-Hour Average $PM_{2.5}$ (µgm-3) for the Continental U.S. (CONUS) Modeling Domain for the Annual Simulation Period (IMPROVE).



The bar charts in Figures 14 and 15 summarize the variations in model performance by species. The mean observed and simulated values for selected species and total $PM_{2.5}$ are graphically compared in Figure 14. Fractional bias and error and are graphically displayed in Figure 15. In a few cases, values are slightly outside the scale used for plotting but can be found in the prior tables.

Figure 14a. Comparison of Mean Observed and Simulated Annual Average $PM_{2.5}$ and Selected Species (μgm^{-3}) for the CONUS Domain (IMPROVE).

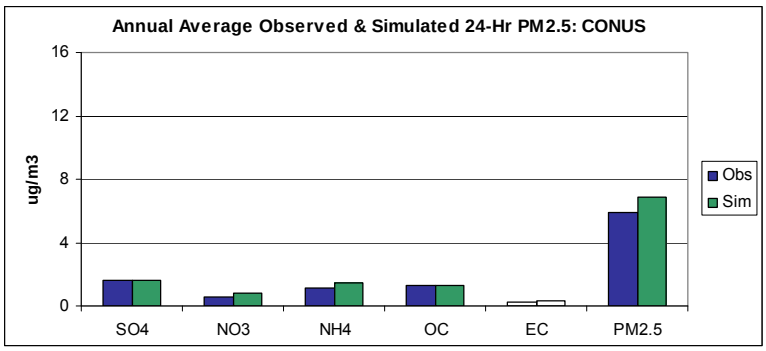


Figure 14b. Comparison of Mean Observed and Simulated Annual Average $PM_{2.5}$ and Selected Species (μgm^{-3}) for the Eastern Portion of the CONUS Domain (IMPROVE).

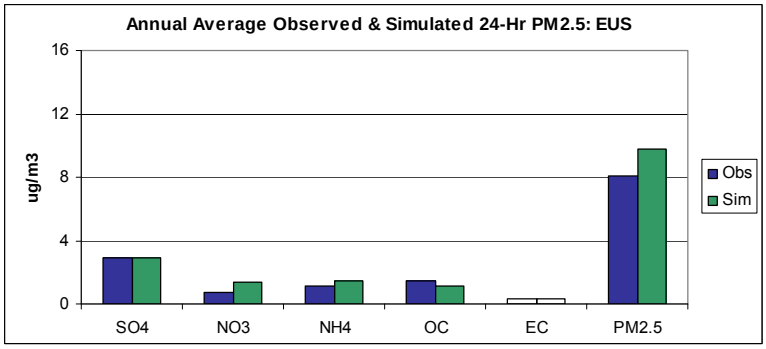


Figure 14c. Comparison of Mean Observed and Simulated Annual Average $PM_{2.5}$ and Selected Species (μgm^{-3}) for the Western Portion of the CONUS Domain (IMPROVE).

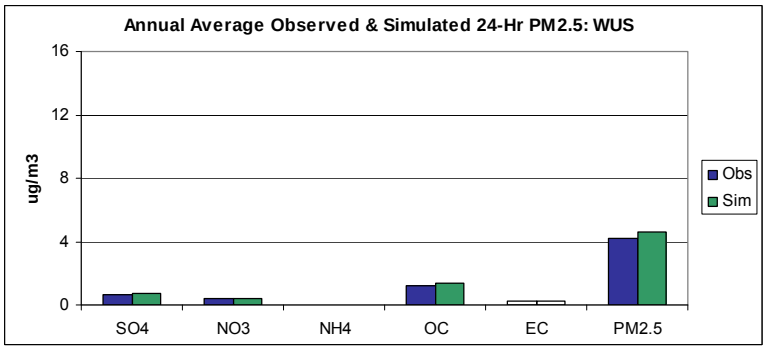


Figure 15a. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ and Selected Species for the CONUS Domain for the Annual Simulation Period (IMPROVE).

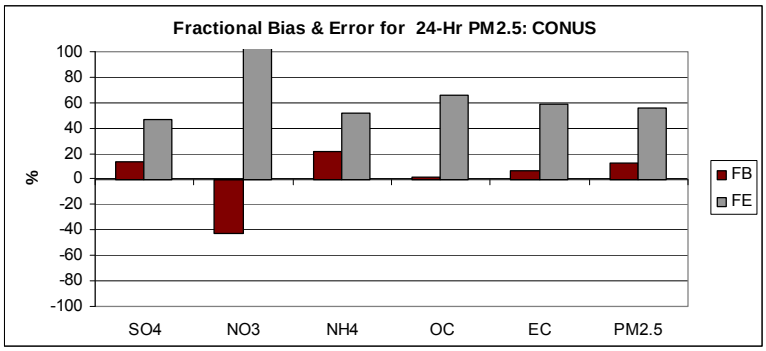


Figure 15b. Fractional Bias (%) and Error (%)Simulated $PM_{2.5}$ and Selected Species for the Eastern Portion of CONUS Domain for the Annual Simulation Period (MPROVE).

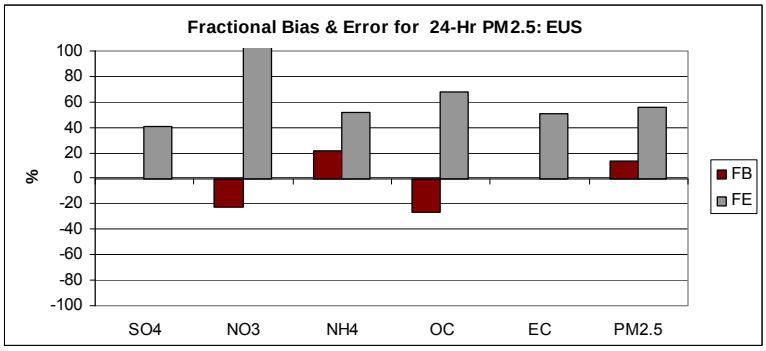
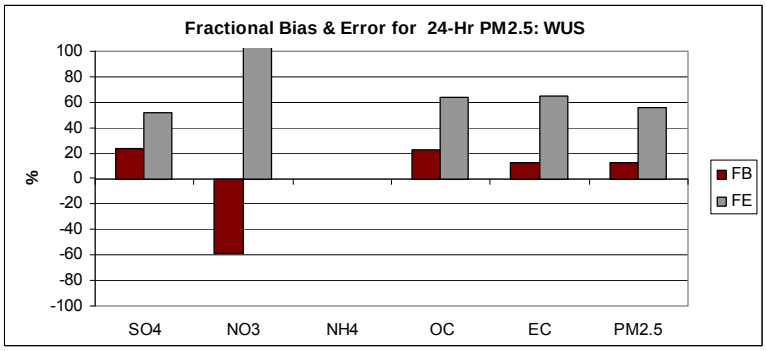


Figure 15c. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ and Selected Species for the Western Portion of CONUS Domain for the Annual Simulation Period (IMPROVE).



The bar charts in Figures 16 and 17 summarize the variations in the model performance by quarter. The mean observed and simulated values for each quarter and for the annual simulation period are graphically compared in Figure 16. Corresponding fractional bias and error values are graphically displayed in Figure 17.

Figure 16a. Comparison of Observed and Simulated Average $PM_{2.5}$ (μgm^{-3}) for Each Quarter and the Annual Simulation Period for the CONUS Domain (IMPROVE).

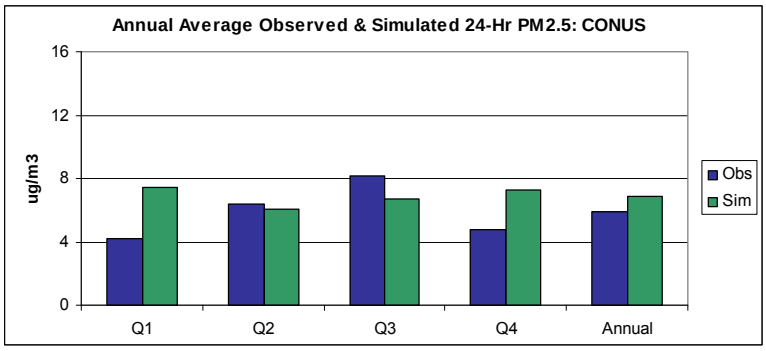


Figure 16b. Comparison of Observed and Simulated Average $PM_{2.5}$ (μgm^{-3}) for Each Quarter and the Annual Simulation Period for the Eastern Portion of the CONUS Domain (IMPROVE).

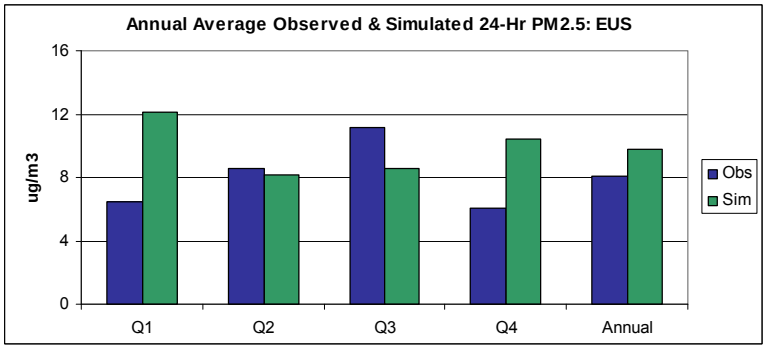


Figure 16c. Comparison of Observed and Simulated Average $PM_{2.5}$ (μgm^{-3}) for Each Quarter and the Annual Simulation Period for the Western Portion of the CONUS Domain (IMPROVE).

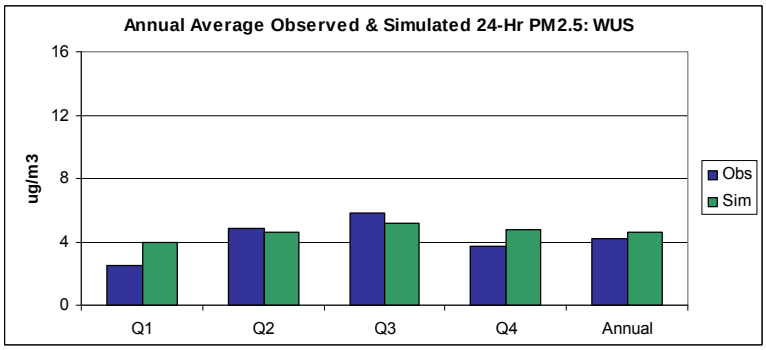


Figure 17a. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ for Each Quarter and the Annual Simulation Period for the CONUS Domain (IMPROVE).

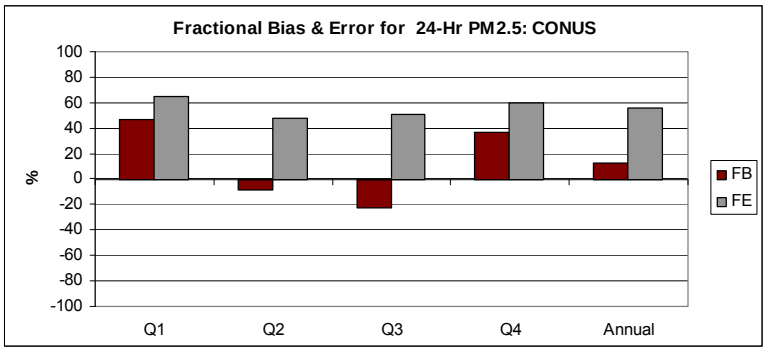


Figure 17b. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ for Each Quarter and the Annual Simulation Period for the Eastern Portion of the CONUS Domain (MPROVE).

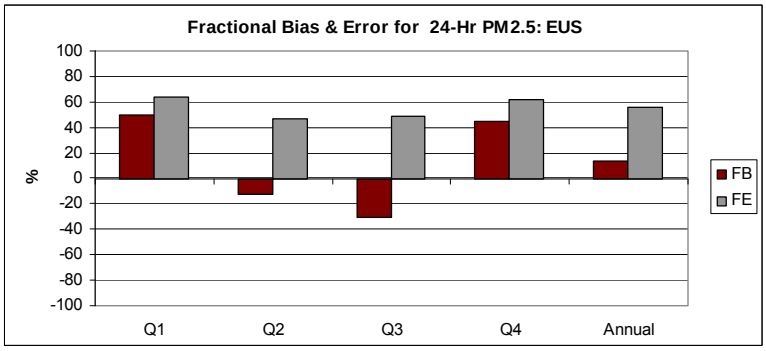


Figure 17c. Fractional Bias (%) and Error (%) for Simulated $PM_{2.5}$ for Each Quarter and the Annual Simulation Period for the Western Portion of the CONUS Domain (IMPROVE).

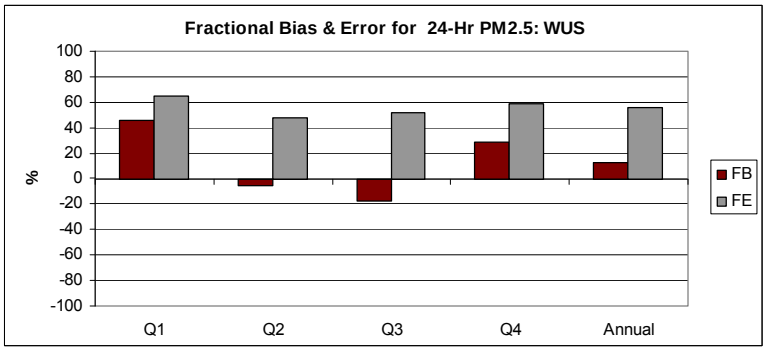
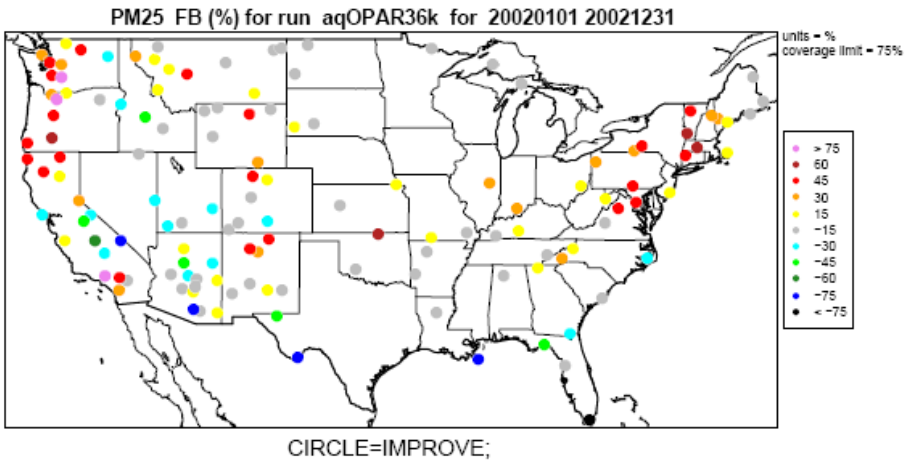


Figure 18 displays the regional distribution of the fractional bias for $PM_{2.5}$ from the IMPROVE dataset for the CONUS domain. In this plot, gray dots correspond to a fractional bias within ± 15 percent, on average. Light blue dots correspond to a negative fractional bias within 15 to 30 percent. Yellow dots correspond to a positive fractional bias within 15 to 30 percent.

Figure 18. Fractional Bias (%) for Simulated $PM_{2.5}$ for the Annual Simulation Period for the CONUS Domain (IMPROVE).



CASTNet

Summary metrics and statistical measures calculated using CASTNet data for $PM_{2.5}$ species for the full CONUS domain and the eastern and western portions of the domain are presented in Table 8. Table 8 provides annual model performance results for the CASTNet species (SO4, NO3, and NH4). Again, goals presented by Boylan (2005) are listed for comparison.

Table 8a. Summary Model Performance Statistics for $PM_{2.5}$ Species for the Continental U.S. (CONUS) Modeling Domain: Weekly Average Measurements (CASTNet).

	SO4	NO3	NH4	Goal
No. of Observations	4088	4088	4080	
Mean Observed (μgm^{-3})	3.0	1.0	1.1	
Mean Simulated (μgm^{-3})	2.9	1.5	1.3	
Normalized Mean Bias (%)	-2.8	52.9	18.7	
Normalized Mean Error (%)	24.6	101.2	41.5	
Fractional Bias (%)	-1.5	-1.9	17.6	± 30
Fractional Error (%)	29.9	96.0	40.0	≤ 50
Correlation (unitless)	0.66	0.35	0.4	
Index of Agreement (unitless)	0.77	0.52	0.51	

Table 8b. Summary Model Performance Statistics for $PM_{2.5}$ Species for the Eastern Portion of the CONUS Modeling Domain: Weekly Average Measurements (CASTNet).

	SO4	NO3	NH4	Goal
No. of Observations	2999	2999	2993	
Mean Observed (μgm^{-3})	3.8	1.1	1.4	
Mean Simulated (μgm^{-3})	3.7	1.9	1.7	
Normalized Mean Bias (%)	-2.6	64.7	19.6	
Normalized Mean Error (%)	23.4	104.0	40.4	
Fractional Bias (%)	-4.8	18.0	18.9	± 30
Fractional Error (%)	25.1	93.0	36.9	≤ 50
Correlation (unitless)	0.58	0.33	0.3	
Index of Agreement (unitless)	0.7	0.49	0.39	

Table 8c. Summary Model Performance Statistics for $PM_{2.5}$ Species for the Western Portion of the CONUS Modeling Domain: Weekly Average Measurements (CASTNet).

	SO4	NO3	NH4	Goal
No. of Observations	1089	1089	1087	
Mean Observed (μgm^{-3})	0.7	0.4	0.3	
Mean Simulated (μgm^{-3})	0.7	0.3	0.3	
Normalized Mean Bias (%)	-5.3	-31.3	6.9	
Normalized Mean Error (%)	43.2	81.4	55.9	
Fractional Bias (%)	7.6	-56.6	13.9	± 30
Fractional Error (%)	43.3	104.1	48.4	≤ 50
Correlation (unitless)	0.35	0.41	0.32	
Index of Agreement (unitless)	0.47	0.59	0.5	

The bar charts in Figures 19 and 20 summarize the variations in annual model performance by species. The mean observed and simulated values for the available CASTNet species are graphically compared in Figure 19. Fractional bias and error and are graphically displayed in Figure 20. In a few cases, values are slightly outside the scale used for plotting but can be found in the prior tables.

Figure 19a. Comparison of Mean Observed and Simulated Annual Average $PM_{2.5}$ Species (μgm^{-3}) for the CONUS Domain (CASTNet).

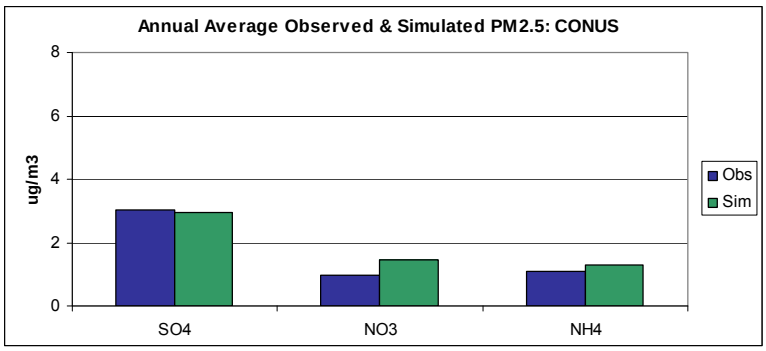


Figure 19b. Comparison of Mean Observed and Simulated Annual Average $PM_{2.5}$ Species (μgm^{-3}) for the Eastern Portion of the CONUS Domain (CASTNet).

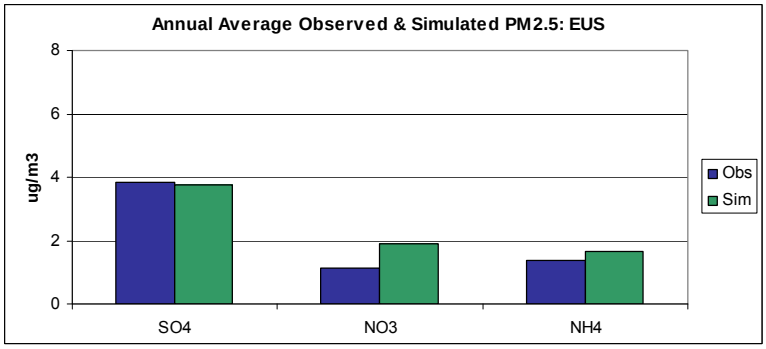


Figure 19c. Comparison of Mean Observed and Simulated Annual Average $PM_{2.5}$ Species (μgm^{-3}) for the Western Portion of the CONUS Domain (CASTNet).

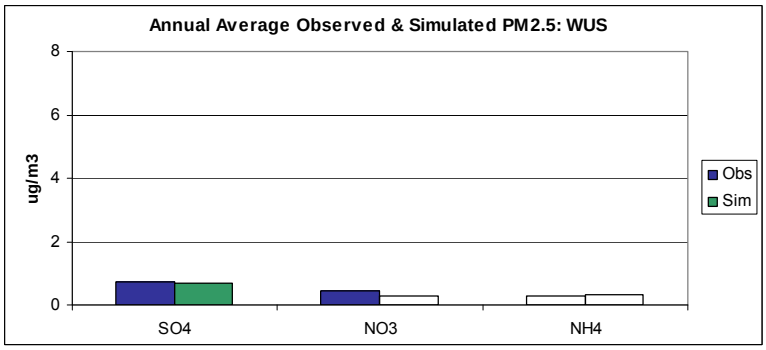


Figure 20a. Fractional Bias (%) and Error (%) for Annual Average $PM_{2.5}$ Species for the CONUS Domain (CASTNet).

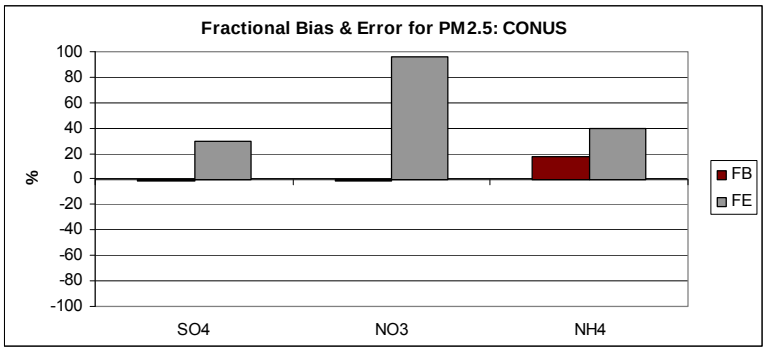


Figure 20b. Fractional Bias (%) and Error (%) for Annual Average $PM_{2.5}$ Species for the Eastern Portion of the CONUS Domain (CASTNet).

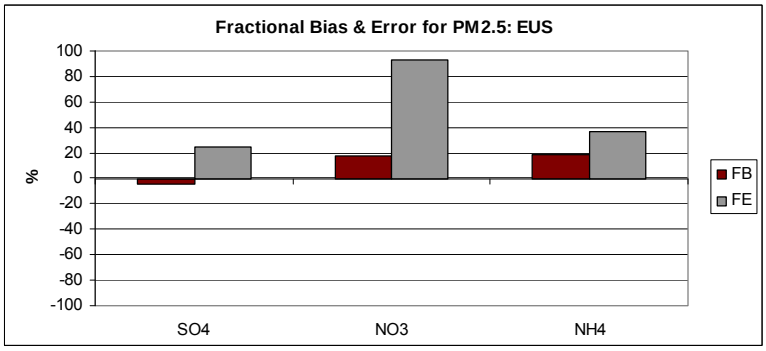
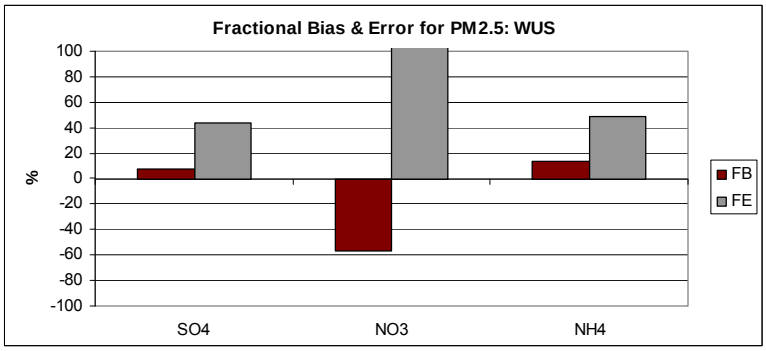


Figure 20c. Fractional Bias (%) and Error (%) for Annual Average $PM_{2.5}$ Species for the Western Portion of the CONUS Domain (CASTNet).



Summary of Model Performance for Selected Deposition Species

CMAQ model performance for selected deposition species was examined for the full CONUS domain and the eastern and western subregions. Performance measures were calculated for each calendar quarter and for the entire annual simulation period using NADP data. Table 9 provides annual model performance results for the NADP species (SO4, NO3, and NH4). No model performance goals for deposition were identified in the literature.

Table 9a. Summary Model Performance Statistics for Deposition Species for the Continental U.S. (CONUS) Modeling Domain: Weekly Average Measurements (NADP).

	SO4	NO3	NH4
No. of Observations	7978	7978	7978
Mean Observed (kg ha^{-1})	0.3	0.2	0.1
Mean Simulated (kg ha^{-1})	0.2	0.2	0.0
Normalized Mean Bias (%)	-2.2	-28.9	-20.2
Normalized Mean Error (%)	65.5	59.8	66.3
Fractional Bias (%)	-17.9	-43.2	-30.8
Fractional Error (%)	80.8	82.4	84.8
Correlation (unitless)	0.57	0.53	0.53
Index of Agreement (unitless)	0.74	0.69	0.69

Table 9b. Summary Model Performance Statistics for Deposition Species for the Eastern Portion of the Continental U.S. (CONUS) Modeling Domain: Weekly Average Measurements (NADP).

	SO4	NO3	NH4
No. of Observations	6172	6172	6172
Mean Observed (kg ha^{-1})	0.3	0.3	0.1
Mean Simulated (kg ha^{-1})	0.3	0.2	0.1
Normalized Mean Bias (%)	-0.8	-25.3	-17.0
Normalized Mean Error (%)	65.1	58.3	65.6
Fractional Bias (%)	-16.6	-35.4	-22.9
Fractional Error (%)	77.6	76.7	80.4
Correlation (unitless)	0.52	0.49	0.5
Index of Agreement (unitless)	0.71	0.67	0.67

Table 9c. Summary Model Performance Statistics for Deposition Species for the Western Portion of the Continental U.S. (CONUS) Modeling Domain: Weekly Average Measurements (NADP).

	SO4	NO3	NH4
No. of Observations	1806	1806	1806
Mean Observed (kg ha^{-1})	0.1	0.1	0.0
Mean Simulated (kg ha^{-1})	0.0	0.0	0.0
Normalized Mean Bias (%)	-25.3	-59.9	-48.8
Normalized Mean Error (%)	72.4	73.1	72.4
Fractional Bias (%)	-22.2	-69.9	-57.9
Fractional Error (%)	91.7	101.9	99.8
Correlation (unitless)	0.4	0.29	0.47
Index of Agreement (unitless)	0.57	0.44	0.57

The bar charts in Figures 21 and 22 summarize the variations in annual model performance by species. The mean observed and simulated values for the available NADP species are graphically compared in Figure 21. Fractional bias and error are graphically displayed in Figure 22. In a few cases, values are slightly outside the scale used for plotting but can be found in the prior tables.

Figure 21a. Comparison of Mean Observed and Simulated Deposition by Species (kg ha^{-1}) for the CONUS Domain (NADP).

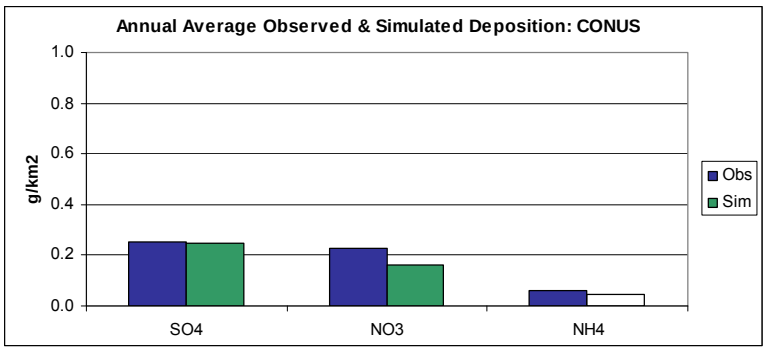


Figure 21b. Comparison of Mean Observed and Simulated Deposition by Species (kg ha^{-1}) for the Eastern Portion of the CONUS Domain (NADP).

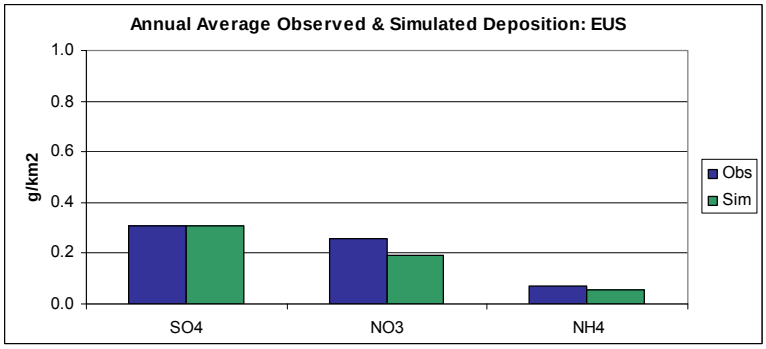


Figure 21c. Comparison of Mean Observed and Simulated Deposition by Species (kg ha^{-1}) for the Western Portion of the CONUS Domain (NADP).

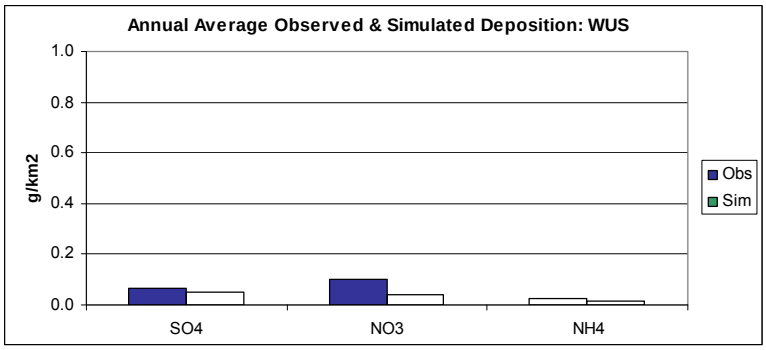


Figure 22a. Fractional Bias (%) and Error (%) for Annual Deposition by Species for the CONUS Domain (NADP).

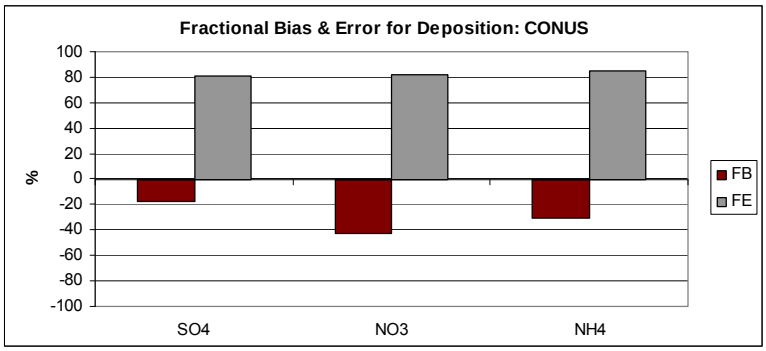


Figure 22b. Fractional Bias (%) and Error (%) for Annual Deposition by Species for the Eastern Portion of the CONUS Domain (NADP).

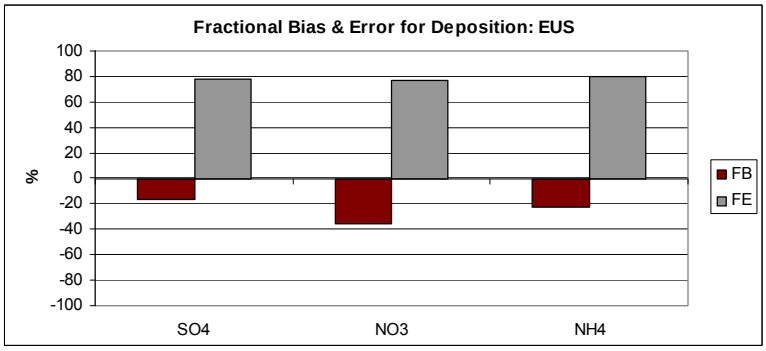
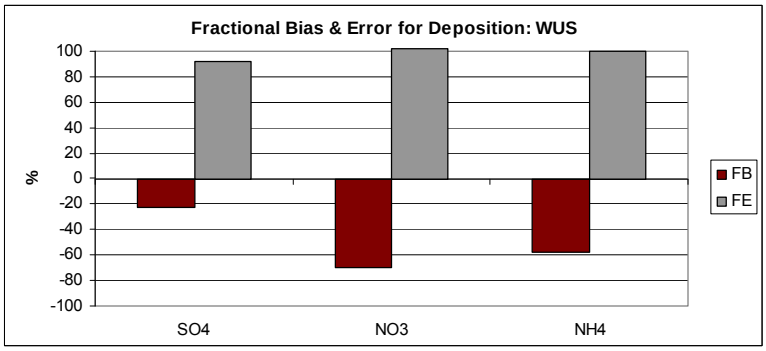


Figure 22c. Fractional Bias (%) and Error (%) for Annual Deposition by Species for the Western Portion of the CONUS Domain (NADP).



Discussion

By most measures, CMAQ model performance for 8-hour ozone for the five-month simulation period is reasonable. Bias and error statistics are well within the recommended ranges for acceptable model performance for both domains, all months, and the entire ozone season. For the EUS, the simulated values are highly correlated with the observations (the correlation ranges from 0.71 to 0.85 for the five months, with an overall value of 0.82). Comparison of the simulated and observed mean values shows that the simulated values are close to observed and the month-to-month variations are consistent. An overall index of agreement of 0.88 indicates that the variations about the mean are well represented. Model performance appears to be relatively consistent throughout the domain, with some overestimation of ozone at some sites scattered throughout the domain, especially in the Southeast, some portions of the Northeast, and the Chicago area.

Although the statistical measures are also good, model performance for ozone is slightly worse for the WUS domain, compared to the EUS domain. Here the correlation ranges from 0.66 to 0.75, with an overall value of 0.69. The corresponding overall index of agreement is 0.82. Spatially, model performance is characterized by a low bias (within ± 15 percent) for most sites throughout the domain, with some exceptions. Ozone is overestimated by more than 15 percent for several sites located in coastal California and underestimated by more than this amount for several sites in California's Central Valley.

CMAQ model performance is reasonable for total PM2.5, but the results for the individual $PM_{2.5}$ species are somewhat mixed. There are variations in performance when considering the different monitoring networks. For the comparison with the STN data, annual average $PM_{2.5}$ is overestimated in the eastern U.S. and overall, but underestimated in the western U.S. The fractional bias and error statistics are within (or in one case nearly within) the goals used for this analysis. Results for the full region and both subregions are characterized by underestimation of nitrate (NO3) and organic carbon (OC). The underestimation of nitrate is most pronounced in the western portion of the domain. The quarterly statistics for the full domain and the eastern portion of the domain show that $PM_{2.5}$ tends to be overestimated during Q1 and Q4 (October – December and January – March, respectively; colder months) and underestimated during Q3 (July – September; warmer months). For the western portion of the domain, there is a tendency for underestimation of $PM_{2.5}$ for all four quarterly periods. The correlation is 0.58 for the eastern portion of the domain, 0.42 for the western portion of the domain, and 0.51 for the full domain. The index of agreement is 0.73, 0.6 and 0.7 for these same three regions.

While the STN monitoring network is comprised mostly of urban and suburban sites, the IMPROVE network covers primarily Class I (i.e., national parks and wilderness) areas. On average, the observed concentrations for $PM_{2.5}$ and its constituent species are lower for the IMPROVE sites. Corresponding model performance errors are larger for the IMPROVE data. Annual average $PM_{2.5}$ is overestimated for the full domain and both subregions. For all three regions, the fractional bias is well within the goal of 30 percent but the fractional error statistics are slightly higher than the goal of 50 percent. Results for the full region based on the IMPROVE data are characterized by underestimation of nitrate. Organic carbon, on the other hand is underestimated in the east and overestimated in the western part of the domain. The quarterly statistics indicate that $PM_{2.5}$ is overestimated during Q1 and Q4 (colder months) and underestimated during Q3 (warmer months). The correlation for total $PM_{2.5}$ is 0.52 for the eastern portion of the domain, 0.36 for the western portion of the domain, and 0.54 for the full domain. The index of agreement is 0.69, 0.56 and 0.7 for these same three regions.

Comparison with the CASTNet data show that of the three PM2.5 species measured by CASTNet good performance is achieved for sulfate (SO4) and ammonium (NH4). NO3 is characterized by larger errors and is overestimated in the east and underestimated in the west.

Comparison of simulated deposition amounts with NADP data indicates that deposition for all three measured species (SO4, NO3, and NH4) is underestimated. For the full domain, the fractional bias ranges from -18 percent for SO4, to -31 percent for NH4, to -43 percent for NO3. The bias and error values are larger for the western, compared to the eastern, portion of the domain.

In summary, model performance is consistent with that for other national-scale and regional-scale CMAQ model applications, and the results can be used (with some uncertainty) to evaluate the effects of the CAAA, especially at the national scale. Model performance is reasonable to good for ozone, even with the relatively coarse 12-km grid resolution. Model performance is also reasonable for total $PM_{2.5}$. The larger errors associate with the simulation of NO3 and OC are typical for most national- and regional-scale PM2.5 applications (EPA, 2007).

References

Boylan, J. 2005. “PM Model Performance Goals and Criteria.” Presented at the 2005 National RPO Modeling Meeting, Denver, Colorado.

Douglas, S. G., J. L. Haney, A. B. Hudischewskyj, T. C. Myers, and Y. Wei. 2008. “Second Prospective Analysis of Air Quality in the U.S.: Air Quality Modeling.” Prepared for the EPA Office of Policy Analysis and Review, Washington, D.C. Prepared by ICF International, San Rafael, California (08-099).

UNC. 2008. “Atmospheric Model Evaluation Tool (AMET) User’s Guide.” Prepared for the U.S. EPA, Office of Research and Development, Research Triangle Park, North Carolina. Prepared by the Institute for the Environment, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina.

U.S. EPA. 2007. “Guidance on the Use of Model and Other Analyses for Demonstration Attainment of Air Quality Goals for Ozone, PM2.5, and Regional Haze.” Prepared by the U.S. EPA Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina (EPA-454/B-07-002).