APPENDIX H

MODELING DEMONSTRATION

Bureau of Air Quality
Department of Environmental Protection

This page blank for copying purposes.

APPENDIX H-1

CMAQ MODEL PERFORMANCE AND ASSESSMENT FOR BASE YEAR 2002 $PM_{2.5}$ MASS AND SPECIATION

Bureau of Air Quality
Department of Environmental Protection

This page blank for copying purposes.

TSD-3

CMAQ Model Performance and Assessment for Base Year 2002: $PM_{2.5}$ Mass and Speciation

Bureau of Air Quality Analysis and Research
Division of Air Resources
New York State Department of Environmental Conservation
Albany, NY 12233

December 14, 2007

Introduction

One of the required tasks for attainment demonstration for the $PM_{2.5}$ National Ambient Air Quality Standard (NAAQS) is the evaluation and assessment of the air quality modeling system that has been utilized to predict future air quality over the region of interest. As part of the attainment demonstration, the MM5/SMOKE/CMAQ modeling system was applied to simulate concentration fields of $PM_{2.5}$ mass and major species for the base year 2002 emissions with the corresponding meteorological information. The modeling databases for meteorology using MM5 (TSD-1, 2007), the emissions using SMOKE (TSD-2a, TSD-2b, 2006), and application of the SMOKE/CMAQ system (TSD-2c, 2007) provide the simulated pollutant fields which are compared to measurements to establish the credibility of the simulation. In the following sections a comparison between the measured and predicted concentrations is performed and results are presented, demonstrating on an overall basis the utility of the modeling system in this application.

The results presented here serve as an illustration of some of the model evaluation and assessment, as outlined in the EPA modeling guidance document (US EPA, 2007), performed on the Base Year 2002 CMAQ simulation. Additional information can be made available by request from the New York State Department of Environmental Conservation.

Summary of measured data

The ambient air quality data for annual 2002 simulation were obtained from the following sources:

- EPA Air Quality System (AQS)
- EPA fine particulate Speciation Trends Network (STN)
- Interagency Monitoring of PROtected Visual Environments (IMPROVE)
- Pinnacle State Park in Addison, NY operated by Atmospheric Science Research Center, University at Albany, Albany, NY
- Harvard Forest, Petersham, MA operated by Harvard University, Boston, MA
- Atmospheric Investigation, Regional Modeling, Analysis and Prediction (AIRMAP) operated by University of New Hampshire, Durham, NH
- NorthEast Ozone & Fine Particle Study (NE-OPS), led by Penn State University and other research groups in Philadelphia, PA

Measured data from sites within the Ozone Transport Region (OTR) plus the rest of Virginia were included here, which hereafter will be referred to as the "OTR+" region. The model-based data were obtained at the grid-cell corresponding to the monitor location and no interpolation was performed. The analysis examines the model response to the OTR+ region with emphasis on the New York City non-attainment area (NYC NAA), which consists of two counties in Connecticut (Fairfield and New Haven) and 10 counties each in New Jersey (Bergen, Essex, Hudson, Mercer, Middlesex, Monmouth, Morris, Passaic, Somerset, and Union) and New York (Bronx, Kings, Nassau, New York, Orange, Queens, Richmond, Rockland, Suffolk, and Westchester).

Fine particulate ($PM_{2.5}$) mass

The 24-hour average Federal Reference Method (FRM) $PM_{2.5}$ mass data collected routinely at SLAMS/NAMS sites across the OTR+ region were extracted from AQS for 225 sites (see http://www.epa.gov/ttn/airs/airsweb/aqsweb/aqswebhome.html). Although 39 of these locations had duplicate monitors in 2002, only data from the primary monitors are presented here. Fifty six of these FRM monitors – 45 unique locations and 11 collocated monitors – are located within the NYC NAA. Hourly $PM_{2.5}$ mass data from 62 sites across the OTR+ region were also included in this analysis. Twenty-eight of these hourly sites are located within the NYC NAA. Hourly $PM_{2.5}$ mass were also obtained from the AIRMAP Program (http://airmap.unh.edu) at Thompson Farm, NH; Pinnacle State Park (http://www.asrc.cestm.albany.edu); and the NE-OPS site in Philadelphia, PA (http://lidar1.ee.psu.edu). In addition to examining the daily averages based on the $PM_{2.5}$ data, composite and site-specific diurnal variations are also presented here. Figures 1 and 2 show the locations of the FRM and hourly $PM_{2.5}$ monitors across the OTR+ region.

Fine particulate speciation

The 24-hour average $PM_{2.5}$ and fine particulate speciation data, consisting of sulfate (SO_4), nitrate (NO_3), elemental carbon (EC), organic carbon/organic mass (OC/OM), and soil/crustal matter from Class I areas across the OTR + region, collected every third day, were obtained from the IMPROVE web site (http://vista.cira.colostate.edu/IMPROVE/Default.htm) for 21 IMPROVE sites (which includes Dolly Sods, WV, even though this site is not in the OTR+ region. In addition to these parameters, the EPA STN (http://www.epa.gov/ttn/amtic/speciepg.html) also reports ammonium (NH_4) to AQS; data from this network are collected every third or sixth day. The STN data from 50 sites in the OTR+ region were used in this analysis. Organic mass is assumed to equal 1.8×OC, and soil/crustal matter is assumed to consist of oxides of Al, Ca, Fe, Si, and Ti. The STN OC data were first corrected by subtracting a monitor-specific constant blank concentration, and these STN blank data are available from http://www.epa.gov/airtrends/aqtrnd03/pdfs/2_chemspecofpm25.pdf. The IMPROVE OC blanks are assumed to equal zero. Nine of the STN sites are located within the NYC NAA. Figure 3 shows the locations of the STN and IMPROVE locations across the OTR+ region. In addition to these daily average speciation data, a preliminary examination of diurnal variations using continuous hourly SO_4 and NO_3 data from the New York State Department of Environmental Conservation (http://www.dec.ny.gov) site IS52 (Bronx) is also presented here.

Evaluation of CMAQ

The following sections provide information on model performance for the above referenced pollutants over the OTR+ portion of the 12-km modeling domain. The statistical formulations that have been computed for each species are as follows: P_i and O_i are the individual daily average predicted and observed concentrations, respectively; $\overline{P}$

and $\overline{O}$ are the average concentrations, respectively, and N is the sample size. These statistical measures are listed in the EPA modeling guidance document (USEPA, 2007).

Each statistic is computed two ways. First, each statistic is computed at each site over the entire year and each quarter; second, each statistic is computed on each day using all available sites. Note that the July 6-9, 2002 period was excluded from this analysis since the observed $PM_{2.5}$ and OM data at many sites were greatly affected by Canadian forest fires. Additionally, those days on which the observed and predicted averages differed by more than a factor of 25 were not included in this analysis, since they can greatly affect some of the normalized metrics.

Observed average, in µg m^{-3}:

$$\overline{O} = \frac{1}{N}\sum O_i$$

Predicted average, in µg m^{-3} (only use P_i when O_i is valid):

$$\overline{P} = \frac{1}{N}\sum P_i$$

Correlation coefficient, R^2:

$$R^2 = \frac{\left[\sum (P_i - \overline{P})(O_i - \overline{O})\right]^2}{\sum (P_i - \overline{P})^2 \sum (O_i - \overline{O})^2}$$

Normalized mean error (NME), in %:

$$NME = \frac{\sum |P_i - O_i|}{\sum O_i} \times 100\%$$

Root mean square error (RMSE), in µg m^{-3}:

$$RMSE = \left[\frac{1}{N}\sum (P_i - O_i)^2\right]^{1/2}$$

Mean fractional error (MFE), in %:

$$MFE = \frac{2}{N}\sum \left|\frac{P_i - O_i}{P_i + O_i}\right| \times 100\%$$

Mean absolute gross error (MAGE), in µg m^{-3}:

$$MAGE = \frac{1}{N}\sum \left|P_i - O_i\right|$$

Mean normalized gross error (MNGE), in %:

$$MNGE = \frac{1}{N}\sum \frac{|P_i - O_i|}{O_i} \times 100\%$$

Mean bias (MB), in μg m^{-3}:

$$MB = \frac{1}{N}\sum (P_i - O_i)$$

Mean normalized bias (MNB), in %:

$$MNB = \frac{1}{N}\sum \frac{(P_i - O_i)}{O_i} \times 100\%$$

Mean fractionalized bias (MFB), in %:

$$MFB = \frac{2}{N}\sum \left[\frac{P_i - O_i}{P_i + O_i}\right] \times 100\%$$

Normalized mean bias (NMB), in %:

$$NMB = \frac{\sum (P_i - O_i)}{\sum O_i} \times 100\%$$

<u>Daily $PM_{2.5}$ mass</u>

Model performance statistics, based on daily average $PM_{2.5}$ levels, are presented in this section across the OTR+ region. Figure 4 displays the time series of observed and predicted $PM_{2.5}$ mass averaged across all FRM monitors. As stated earlier the July 6-9 period was not included in this analysis since many monitors across the region were substantially impacted by Canadian wildfires. Overall, the mean bias (predicted minus observed) was about 3.8 μg m^{-3}, ranging from -17.7 μg m^{-3} to +24.0 μg m^{-3}. The highest overprediction tended to occur during the colder months, quarters 1 and 4, whereas the days on which the model tended to underpredict $PM_{2.5}$ were more likely to occur during the summer months.

Figures 5 and 6 display the mean fractional error (MFE) and mean fractionalized bias (MFB), respectively, at each FRM monitor across the OTR+ region over the entire year. Boylan and Russell (2006) propose that these two metrics are the most useful for model evaluation since they set upper bounds on error/bias, are not likely dominated by a few "outlier" points, and do not assume that observations are absolute truth. Boylan and Russell (2006) further point out that MFE<75% and MFB<60% define a level of accuracy close to the best a model can be expected to achieve in terms of acceptable model performance. The analysis shows that at 221 of the 225 FRM sites the MFE was below 75%, while at 215 sites the MFB was below 60% in the OTR + region. However, such is not the case for the NYC NAA, a region with high emissions density and complex meteorology. While the MFE and MFB criteria are met at each of the 10 primary FRM monitors in CT and 14 primary FRM monitors in NJ, the MFE criteria was met at 17 of the 21 primary FRM monitors in NY and the MFB criteria was only met at 12 of the primary NY FRM monitors. Hence, on an overall basis the model predicts $PM_{2.5}$ mass reasonably well, although not consistently at all locations and all days in the NYC NAA.

Tables 1 and 2 list the median and range in MFE and MFB for the OTR+ region for the entire year excluding July 6-9 that are associated with Canadian fires. The statistical estimates are based on comparing predicted with measurements from FRM, continuous hourly, STN, and IMPROVE locations. In Table 1, the MFE and MFB values are computed at each site, while in Table 2 the MFE and MFB are computed on each day using the data from all available sites. The median MFE values are about 40-50%, ranging from <20% to about 100%. The median MFB values are about 30%, ranging within ±100% (substantial underprediction to large overprediction).

Figure 7 displays the quarterly average $PM_{2.5}$ mass at all OTR+ monitors and also at those monitors within the NYC NAA. Note that, not surprisingly, the observed average $PM_{2.5}$ mass tends to be higher in the NYC NAA than across the entire OTR+. CMAQ tends to overpredict $PM_{2.5}$ at these locations, especially during the colder months. On the other hand, CMAQ performs well during the July-September period on a composite average basis.

$PM_{2.5}$ speciation

Composite daily average predicted and observed concentrations of major $PM_{2.5}$ species – SO_4, NO_3, NH_4, EC, OM (defined here operationally as 1.8×blank-corrected organic carbon), and crustal mass (sum of oxides of Al, Ca, Fe, Si, and Ti) – across the OTR+ region were compared in this analysis. The time series of these species at the 50 STN and 21 IMPROVE locations are shown in Figures 8-18. The data displayed are for every third day, corresponding to the nominal sampling schedule. These two speciation networks collect SO_4, NO_3, EC, OM, and crustal mass, while only the STN reports NH_4 at a number of locations. The NH_4 data from the IMPROVE network are not included here, since in the OTR+ region, only the IMPROVE sites of Dolly Sods, WV; James River Face Wilderness, VA; and Shenandoah National Park, VA reported NH_4 data in 2002. As before, the July 6-9 period was excluded.

The model exhibits good agreement for SO_4 at both the urban STN (Figure 8) and rural IMPROVE (Figure 9) sites, both in terms of absolute concentrations and the seasonal variation. During the summer months, photochemical production of SO_4 is higher than at other times of the year, which the model seems to capture. Figures 10 and 11 suggest that the model qualitatively predicts the seasonal variation of NO_3 fairly well, although the observed and modeled concentrations during the cold season, quarters 1 and 4 (when volatilization of NO_3 is lowest), differ by several $\mu g\ m^{-3}$. The seasonal variation in NH_4 (Figure 12) is more complicated, but appears to be consistent with SO_4 and NO_3. During the summer, most NH_4 is tied to SO_4 (which the model tends to predict well), whereas during the winter a substantial fraction of NH_4 is tied to NO_3 (which the model appears to predict poorly). Both EC (Figures 13 and 14) and crustal mass (Figure 17 and 18) are predominantly primary emissions, and the model tends to overpredict both species – particularly at urban locations and during the colder months. Note, too, the effects of emissions from the July 4th fireworks on crustal mass are obvious in the observed data at both the STN and IMPROVE sites, but are not included in the model emissions inventory. The measured crustal mass on this day exceeds 4 μg m^{-3}, while it is generally closer to ~0.5-1.0 μg m^{-3}only.

The model appears to also have difficulty predicting the seasonal variation of OM (Figures 15 and 16). During the summer months a substantial fraction of OM is secondary in nature, while during the colder months much of the OM may be in the form of primary emissions. To illustrate this, Figure 19 displays the quarterly variation in observed OM, as well as the CMAQ predictions of the three components of modeled OM – anthropogenic primary organic aerosol (POA), anthropogenic secondary organic aerosol (SOA), and biogenic SOA. Note that CMAQ predicts that >80% of the modeled OM is anthropogenic POA (with a wintertime maximum), and only predicts slight increases in the SOA fraction during the summer. On the other hand, the observed OM has a summertime maximum when secondary formation is highest.

Table 3 lists the median and range in MFE and MFB calculated at each site over the season used in this analysis. For SO_4, which contributes ~30% to the total $PM_{2.5}$ mass, the MFE at both STN and IMPROVE locations is roughly 40%, and there is a general tendency to underpredict SO_4 (MFB < 0). At all 71 locations (STN and IMPROVE), the MFE and MFB criteria suggested by Boylan and Russell (2006) are met. This is not surprising, since SO_2 emissions are generally well-characterized and ambient SO_4 levels tend to be somewhat spatially uniform compared to other $PM_{2.5}$ species. The range in MFE and MFB tends to be lower for SO_4 compared to the other species as well. The urban-rural differences for the other species are much more evident than for SO_4. Note that the degree of overprediction of NO_3, NH_4, and EC is considerably higher at the STN locations compared to IMPROVE. CMAQ tends to underpredict OM as well, although some of this discrepancy may be attributed to the fact that OM is operationally defined and is highly dependent on the blank correction and multiplier to account for other components of OM not directly measured. Similarly, the predicted crustal mass tends to be much higher than that observed. This, too, is due in part to relating the CMAQ predictions of unspeciated, primary particulates to the measured crustal mass which is operationally defined as the sum of major metal oxides. For these and other reasons, Boylan and Russell (2006) state that for the non-SO_4 or less abundant $PM_{2.5}$ species, the MFE and MFB guidelines should be relaxed.

Diurnal variations $PM_{2.5}$, SO_4, and NO_3

The hourly $PM_{2.5}$ mass and speciation data allow for examination of the model predictions with high temporal resolution. Figure 20 displays the composite diurnal variation of $PM_{2.5}$ mass averaged across the 28 locations in the NYC NAA, over the entire year (with the exception of July 6-9). For comparison, the average diurnal profiles during quarters 1 and 3 are shown in Figures 21 and 22, respectively. Quarter 1 exhibits the largest discrepancy between observed and predicted $PM_{2.5}$, while quarter 3 exhibits the smallest discrepancy between observed and predicted $PM_{2.5}$ mass. Regardless of time of year, CMAQ predicts a pronounced double peak, with maxima during the morning and late afternoon/early evening time periods, likely corresponding to average commuting patterns. During quarter 1, there is a double peak in the observed diurnal profile, albeit smaller in amplitude. The double peak is not nearly as evident during quarter 3, when during the middle of the day the observed $PM_{2.5}$ levels are actually slightly higher than the predicted levels.

There were only a handful of locations in 2002 where hourly SO_4 and/or NO_3 were measured. One such location was IS52 in Bronx, NY. To illustrate how the model performed over the course of a day at this single location, Figures 23 and 24 display the average diurnal variation of SO_4 and NO_3, respectively. On an average basis, there is little variation in SO_4 at this site over the day (see Figure 23), and the observed and predicted SO_4 levels generally are within 10-20% of each other. The observed SO_4 levels on average do exhibit a small peak during the summer months, but the lack of a strong diurnal indicates the regional nature of this pollutant. In contrast to SO_4, there is a more pronounced diurnal variation in NO_3 at this location (see Figure 24), with a morning peak related to NO_x emissions and a still shallow boundary layer, and an afternoon minimum related to a deep boundary layer, deposition, and the thermal instability of ammonium nitrate (NH_4NO_3). Similar to $PM_{2.5}$ mass, the model overprediction tends to be highest during the nighttime and early morning hours.

Summary

Various model evaluation statistics are presented here for $PM_{2.5}$ mass and major species over the entire 2002 CMAQ simulation. In general, the CMAQ results were best for daily average $PM_{2.5}$ and SO_4 mass. By comparison, CMAQ does not appear to be able to reproduce the day-to-day or seasonal variation in other species like NO_3, OM, EC, or crustal mass. The model appears to meet model performance criteria for $PM_{2.5}$ reasonably well over the OTR+ region as a whole, but not as well if one focuses on an emissions-dense, complex urban setting such as the NYC NAA.

References

TSD-1 (2006) Meteorological Modeling of 2002 using Penn State/NCAR 5th Generation Mesoscale Model (MM5)

TSD-2a (2006) Processing of 2002 Biogenic Emissions for OTC / MANE-VU Regional and Urban Modeling.

TSD-2b (2006) Processing of 2002 Anthropogenic Emissions: OTC Regional and Urban 12km Base Year Simulation.

TSD-2c (2006) $PM_{2.5}$ modeling using the SMOKE/CMAQ system over the Ozone Transport Region (OTR)

Boylan, J. W., and Russell, A. G., 2006. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. Atmospheric Environment 40, 4946-4959.

United States Environmental Protection Agency (USEPA), 2007. Guidance on the use of models and other analyses for demonstrating attainment of air quality goals for ozone, $PM_{2.5}$, and regional haze. EPA-454/B-07-002, Research Triangle Park, NC.

Table 1. The median and range in mean fractional error (MFE) and mean fractionalized bias (MFB) values for daily $PM_{2.5}$ mass at FRM, hourly, and speciation sites across the OTR+ region. These statistics were computed at each site over the entire year.

	Median MFE (minimum to maximum MFE)	**Median MFB (minimum to maximum MFB)**
FRM (225 sites)	42.9% (28.1 to 109%)	22.6% (-72 to 109%)
Hourly (62 sites)	49.3% (17.4 to 88%)	35.3% (-41.9 to 134%)
STN (50 sites)	40.1% (28.1 to 83%)	12.3% (-27.3 to 81%)
IMPROVE (21 sites)	53.1% (35.3 to 73%)	30.1% (-12.0 to 64%)

Table 2. Same as Table 1, except that the statistics were computed on each day using all available sites.

	Median MFE (minimum to maximum MFE)	**Median MFB (minimum to maximum MFB)**
FRM (225 sites)	42.5% (17.6 to 87%)	26.4% (-83 to 88%)
Hourly (62 sites)	50.0% (18.6 to 99%)	35.9% (-68 to 99%)
STN (50 sites)	42.1% (17.9 to 84%)	27.0% (-80 to 75%)
IMPROVE (21 sites)	48.4% (22.9 to 102%)	30.3% (-101 to 101%)

Table 3. Same as Table 1, except for daily average species concentrations at the 50 STN and 21 IMPROVE sites.

	Median MFE (minimum to maximum MFE)	**Median MFB (minimum to maximum MFB)**
STN SO_4	39.8% (27.2 to 60%)	-11.1% (-40.9 to 52%)
IMPROVE SO_4	42.4% (34.5 to 60%)	-22.6% (-54 to 2.2%)
STN NO_3	80% (57 to 123%)	25.3% (-76 to 118%)
IMPROVE NO_3	104% (86 to 125%)	58% (-2.4 to 122%)
STN NH_4	43.0% (27.8 to 83%)	27.8% (-13.2 to 81%)
STN EC	48.8% (29.4 to 106%)	29.5% (-43.2 to 102%)
IMPROVE EC	57% (41.3 to 87%)	7.2% (-64 to 85%)
STN OM	62% (33.7 to 106%)	-34.9% (-94 to 64%)
IMPROVE OM	54% (40.1 to 83%)	-4.0% (-82 to 43.8%)
STN crustal	120% (72 to 159%)	116% (56 to 159%)
IMPROVE crustal	101% (65 to 126%)	97% (36.0 to 124%)

Figure 1. Locations of the FRM monitors across the OTR+ (open squares) and NYC NAA (filled squares). Forty five of the 225 monitors are within the NYC NAA.

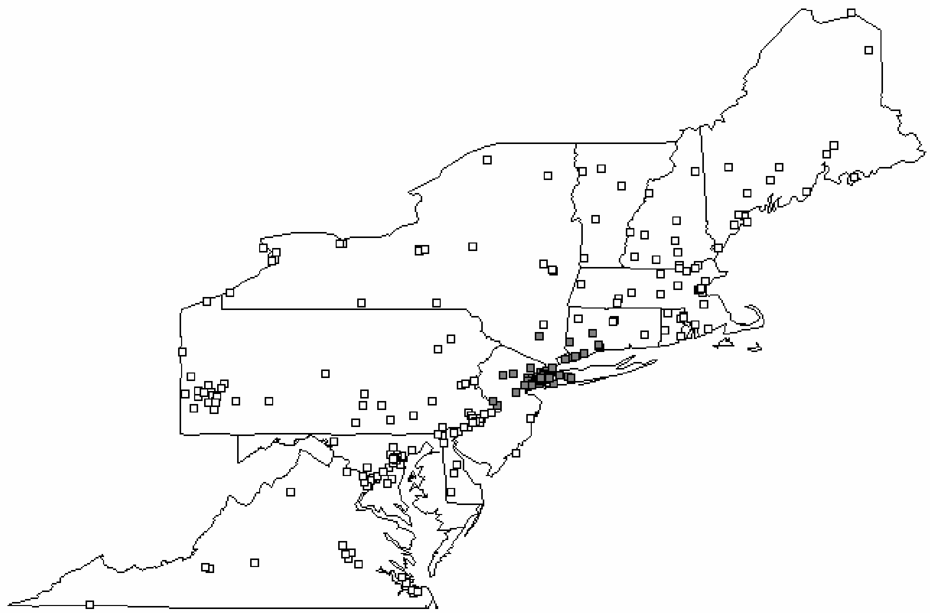


Figure 2. Locations of the hourly TEOM monitors across the OTR+ (open squares) and NYC NAA (filled squares). Twenty-eight of the 62 monitors are within the NYC NAA.

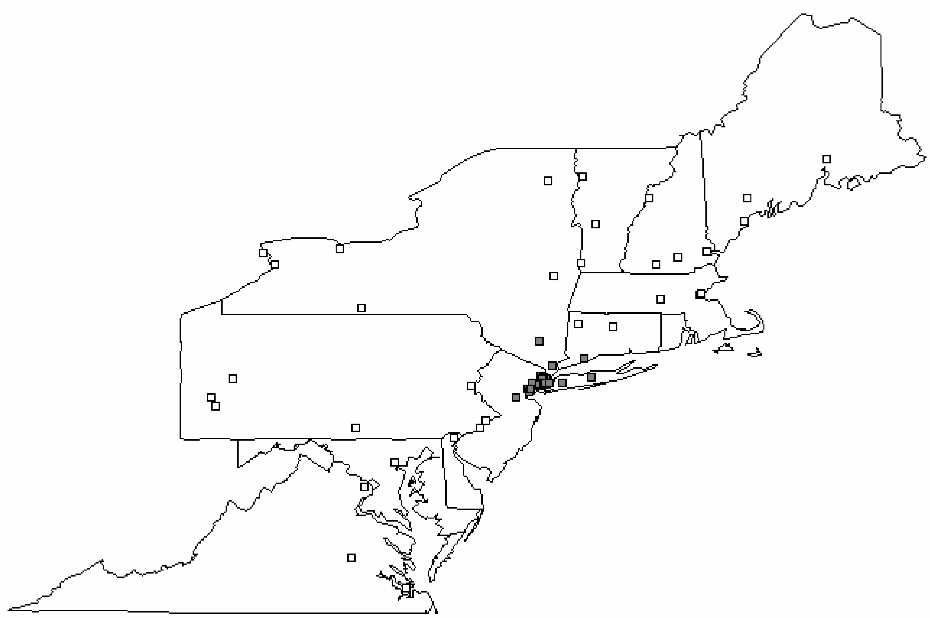


Figure 3. Locations of the speciation monitors across the OTR+ (open squares denote STN, open triangles denote IMPROVE) and the NYC NAA (filled squares denote STN). Nine of the 50 STN monitors are within the NYC NAA. There are 21 IMPROVE monitors across the OTR+ region.

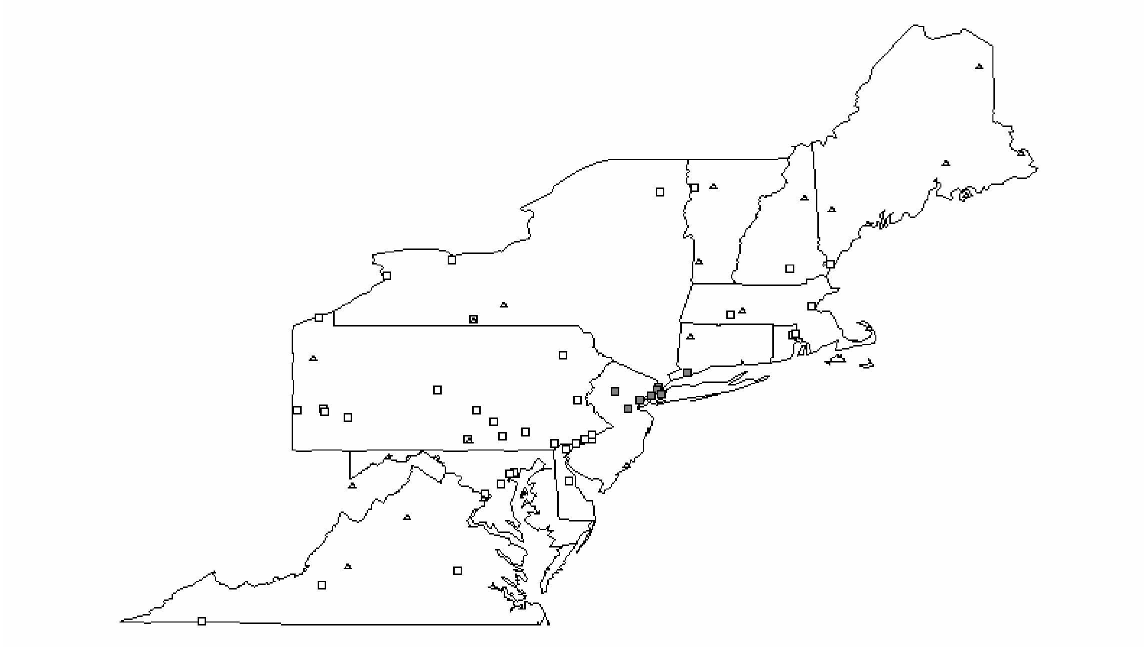


Figure 4. Time series of $PM_{2.5}$ mass based on the composite average of all 225 FRM monitors across the OTR+ region. The observed values are denoted with the black line, the model predictions are denoted with a thick gray line and asterixes.

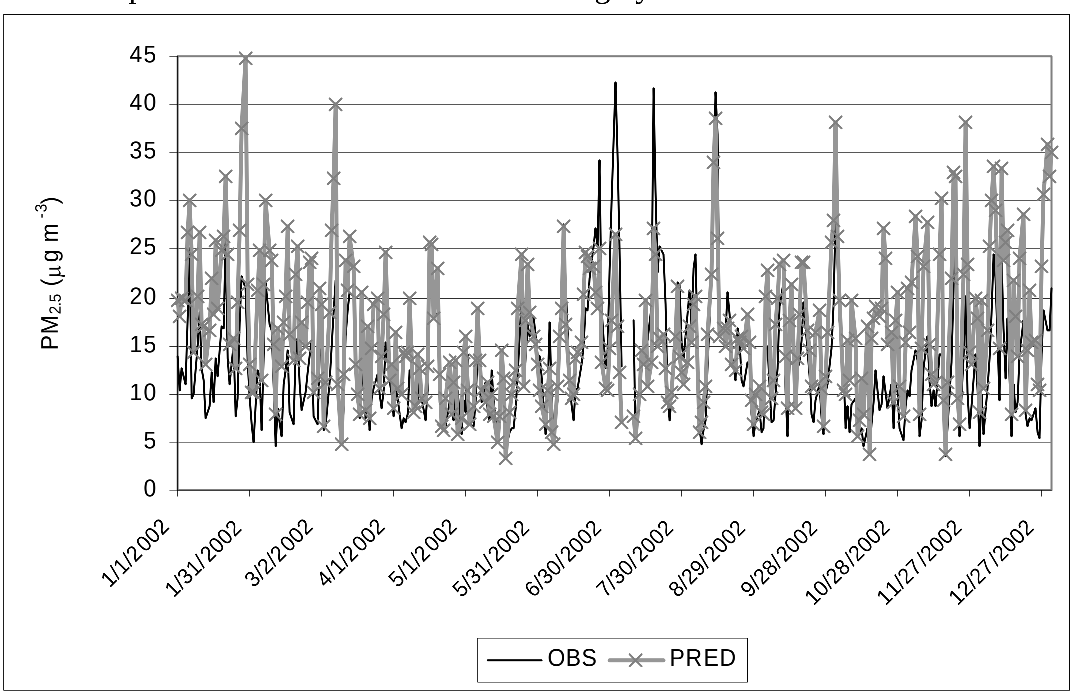


Figure 5. Mean fractional error (MFE, %) at each FRM location over the entire year: blue, <30%; green, 30-45%; orange, 45-60%; red, 60-75%; pink, >75%.

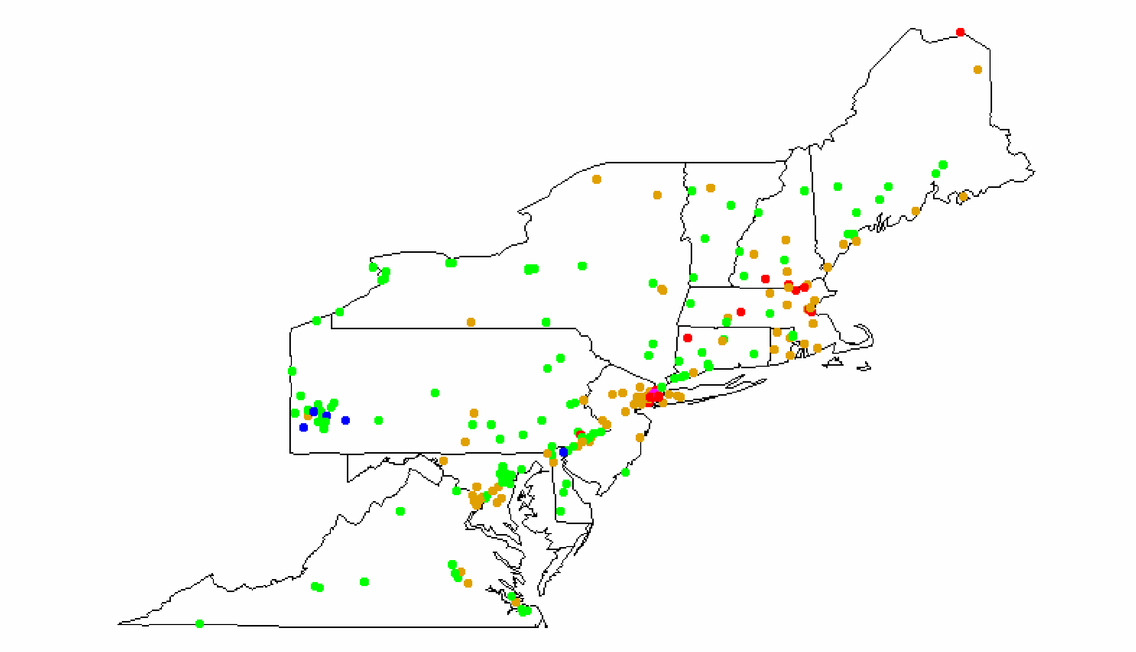


Figure 6. Mean fractionalized bias (MFB, %) at each FRM location over the entire year: gray, <-45%; blue, -45 to -15%; green, -15 to 15%; orange, 15-45%; red, 45-75%; pink, >75%.

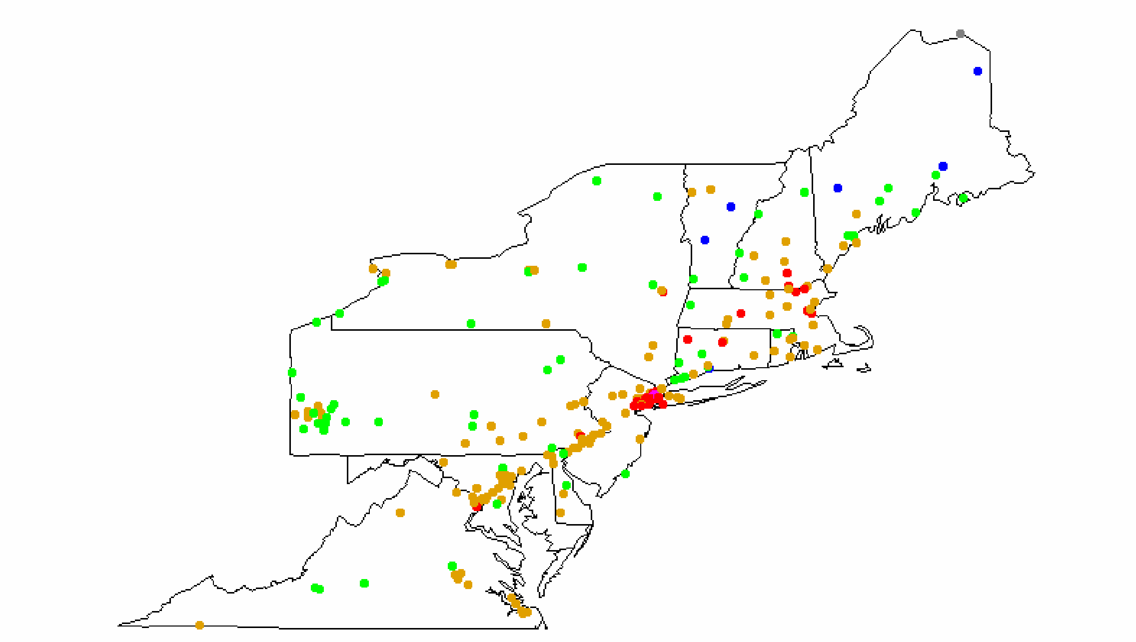


Figure 7. Observed and predicted quarterly average $PM_{2.5}$ mass at FRM locations, for the entire OTR+ region and for the NYC NAA.

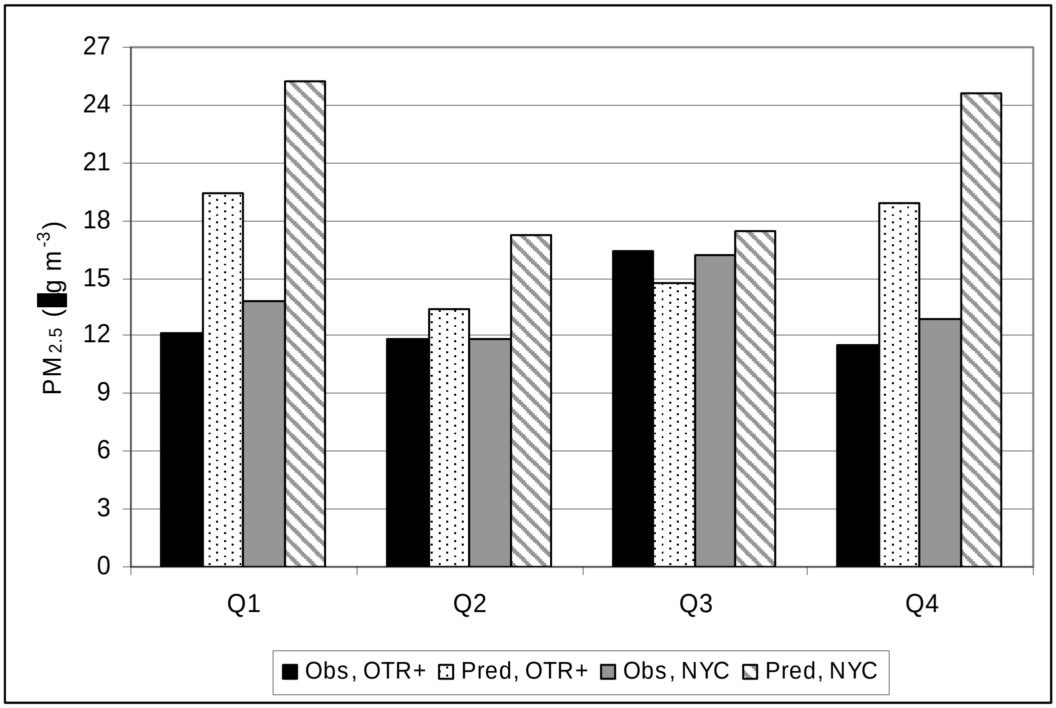


Figure 8. Time series of SO_4 mass based on the composite average of all 50 STN monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

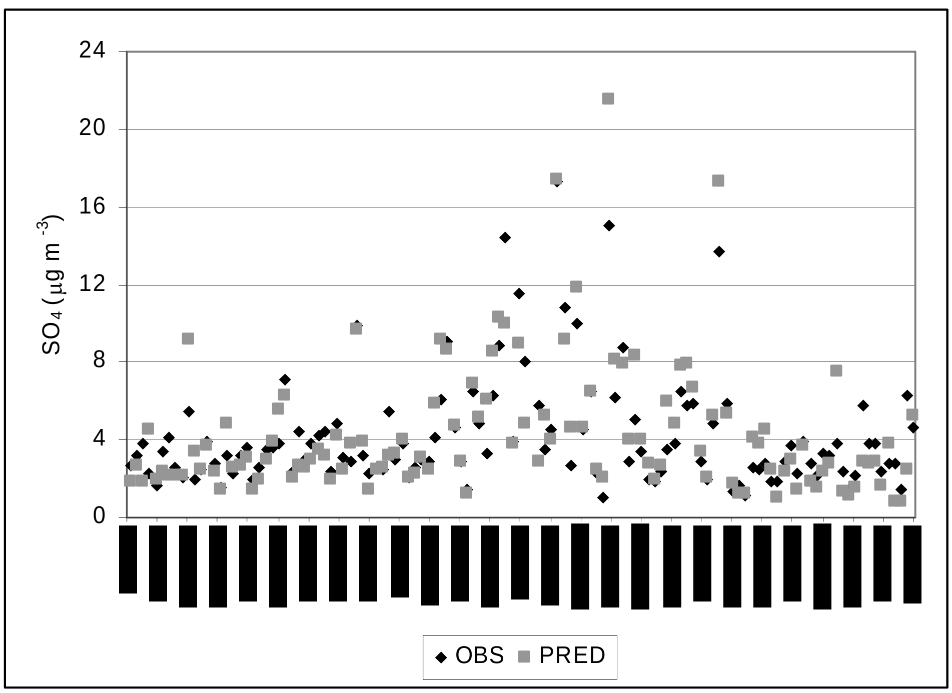


Figure 9. Time series of SO_4 mass based on the composite average of all 21 IMPROVE monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

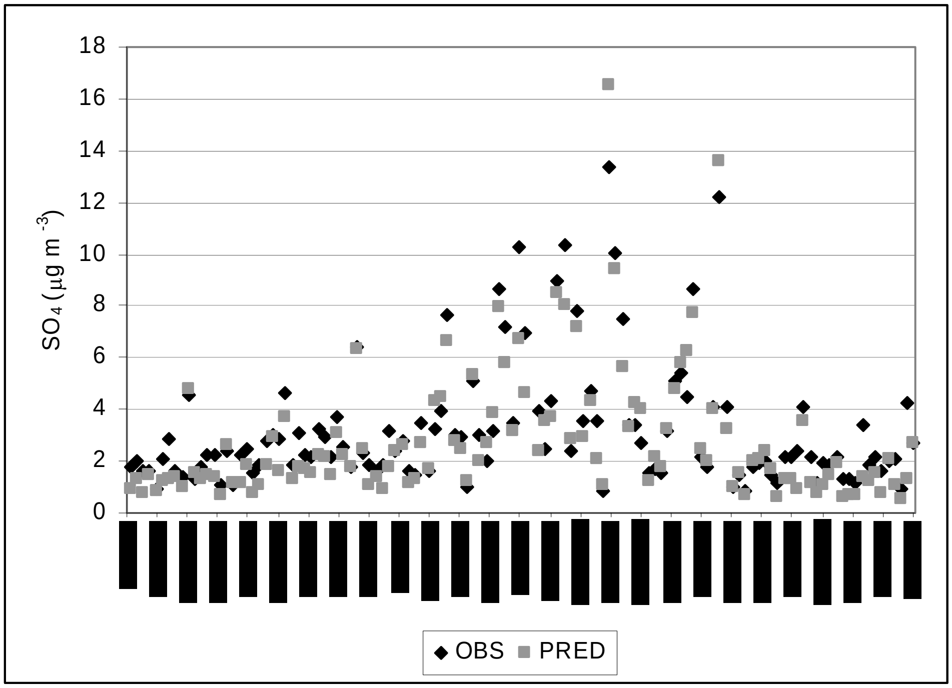


Figure 10. Time series of NO_3 mass based on the composite average of all 50 STN monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

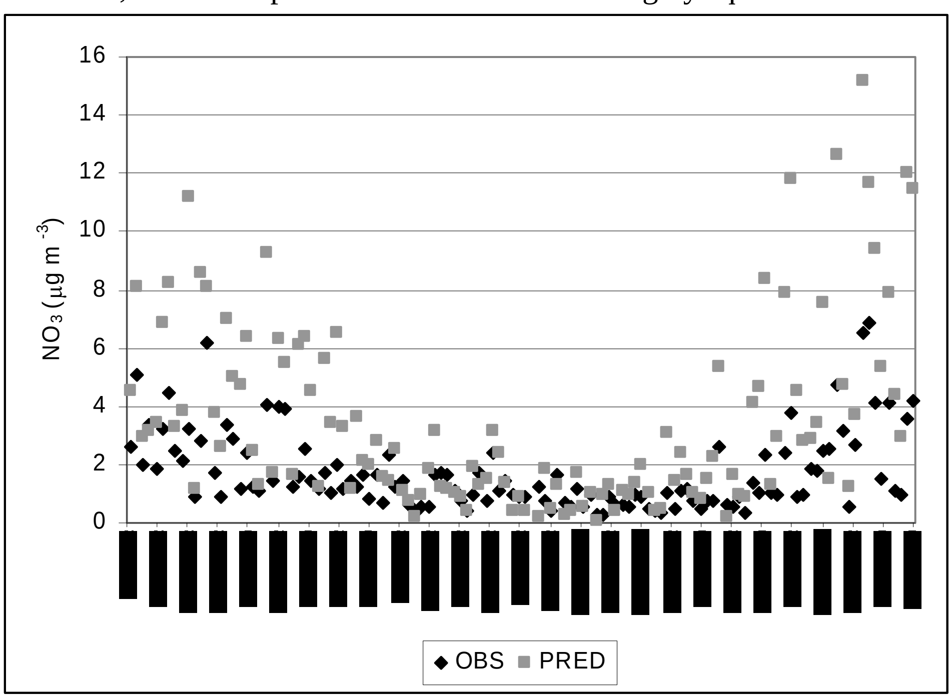


Figure 11. Time series of NO_3 mass based on the composite average of all 21 IMPROVE monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

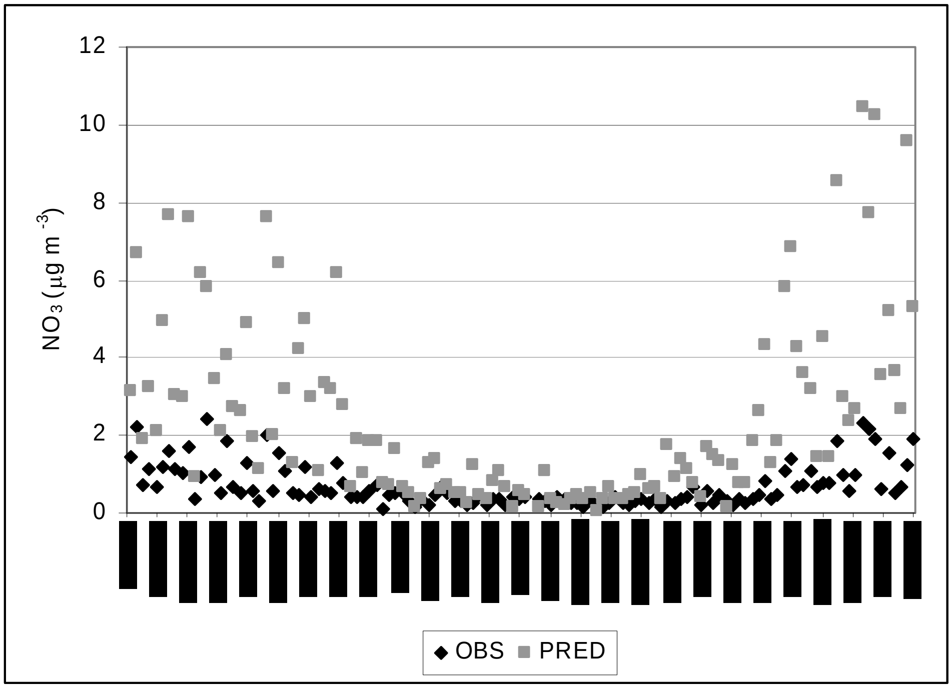


Figure 12. Time series of NH_4 mass based on the composite average of all 50 STN monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

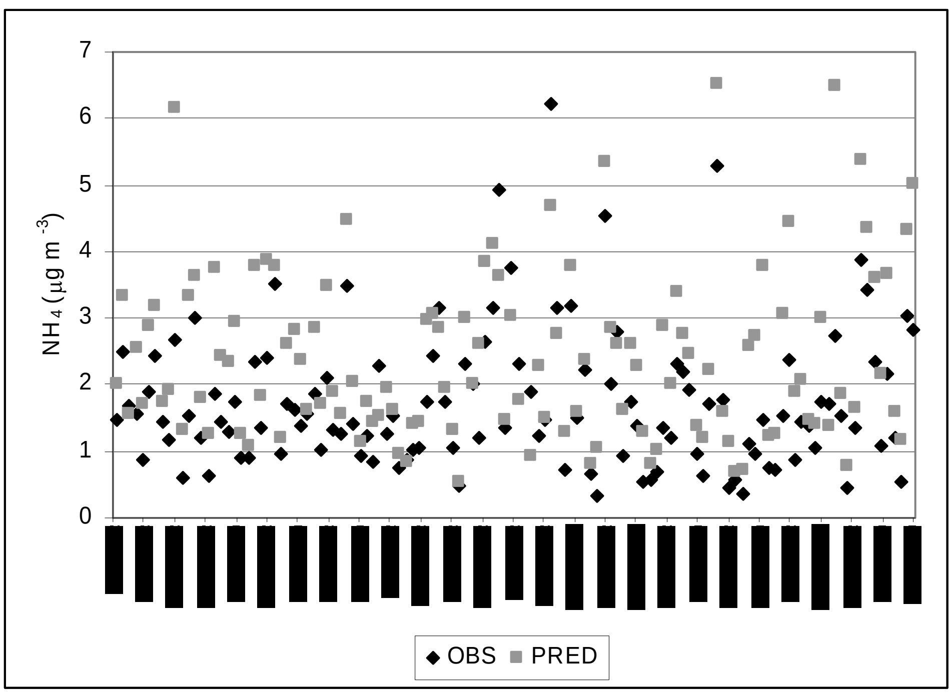


Figure 13. Time series of EC mass based on the composite average of all 50 STN monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

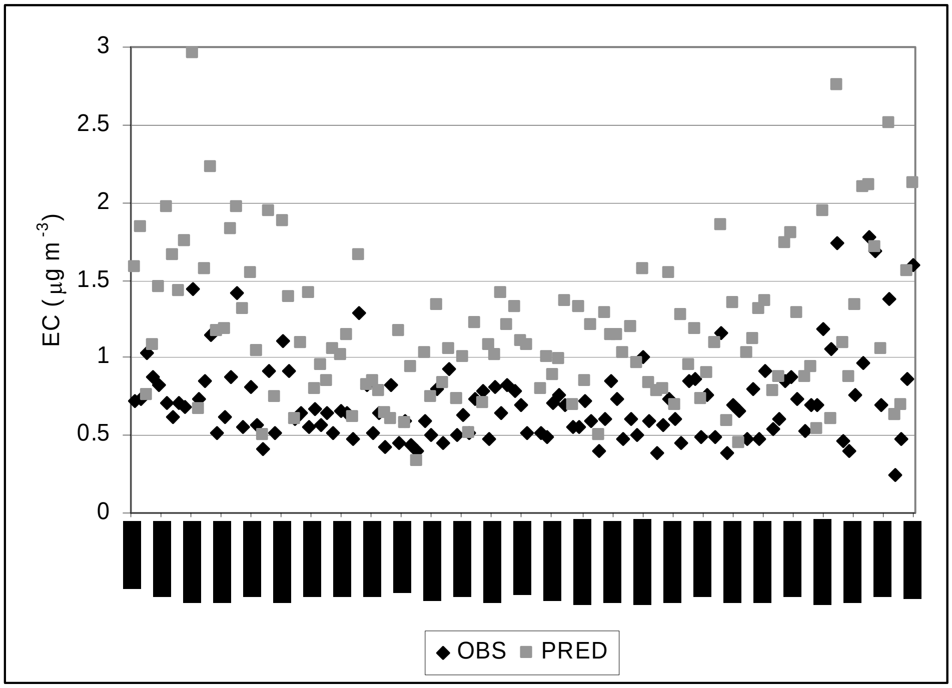


Figure 14. Time series of EC mass based on the composite average of all 21 IMPROVE monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

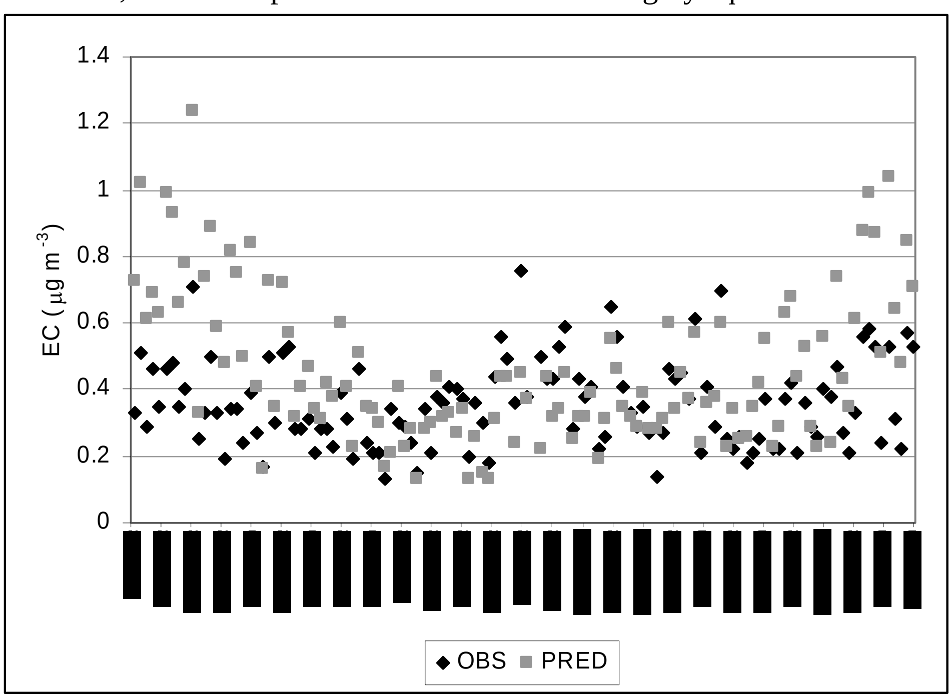


Figure 15. Time series of OM mass based on the composite average of all 50 STN monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

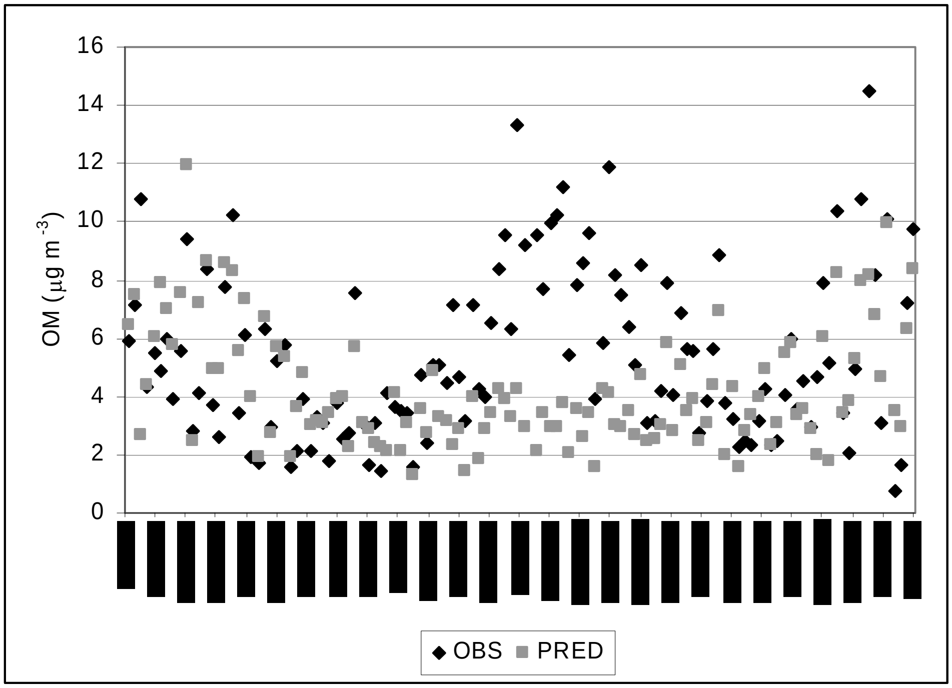


Figure 16. Time series of OM mass based on the composite average of all 21 IMPROVE monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

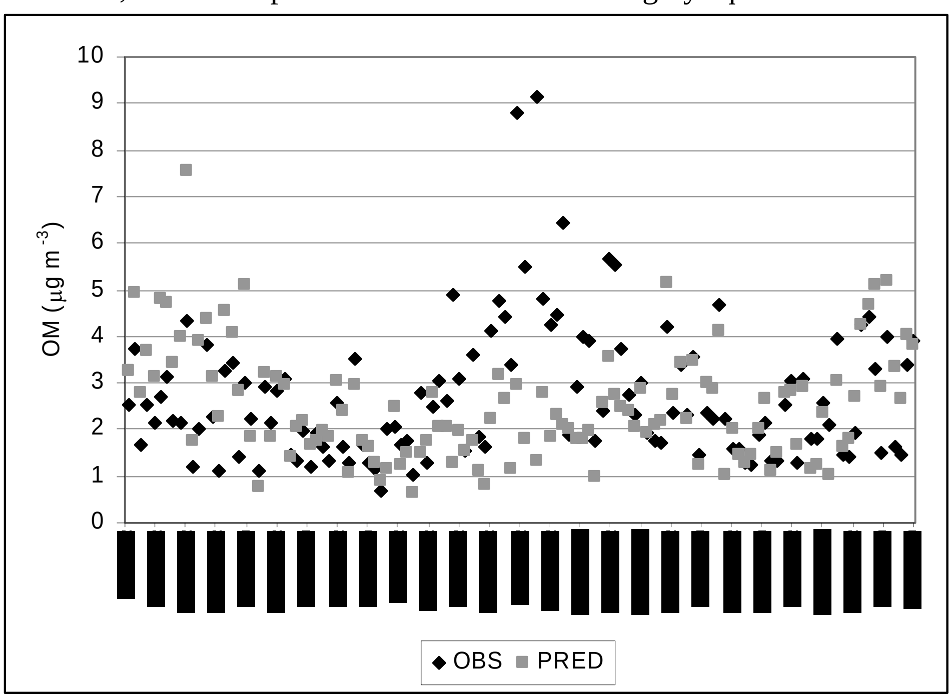


Figure 17. Time series of crustal mass based on the composite average of all 50 STN monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

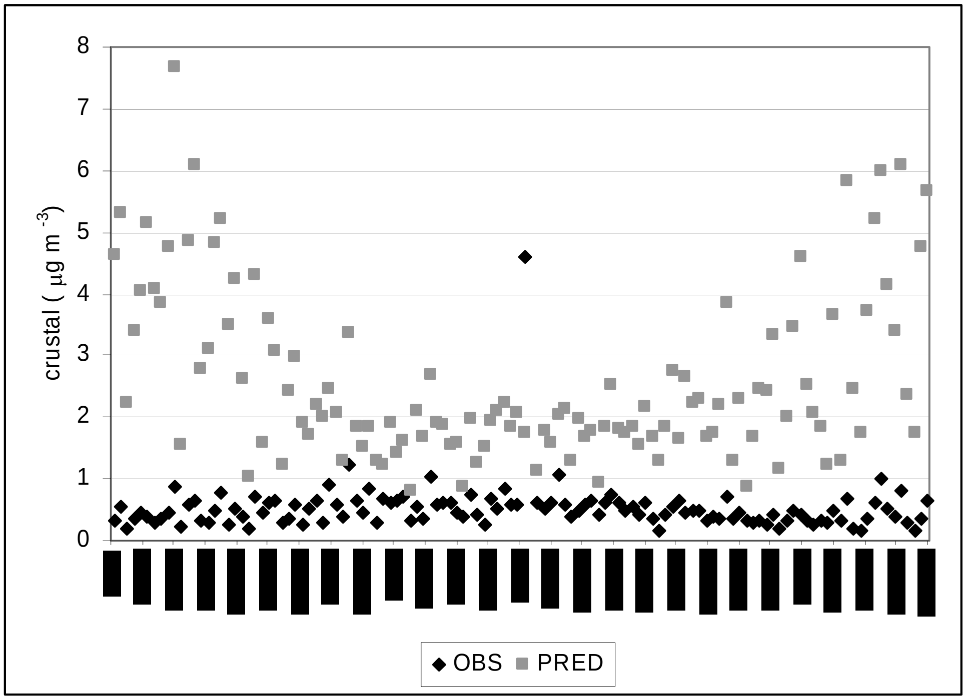


Figure 18. Time series of crustal mass based on the composite average of all 21 IMPROVE monitors across the OTR+ region. The observed values are denoted with the black diamonds, the model predictions are denoted with gray squares.

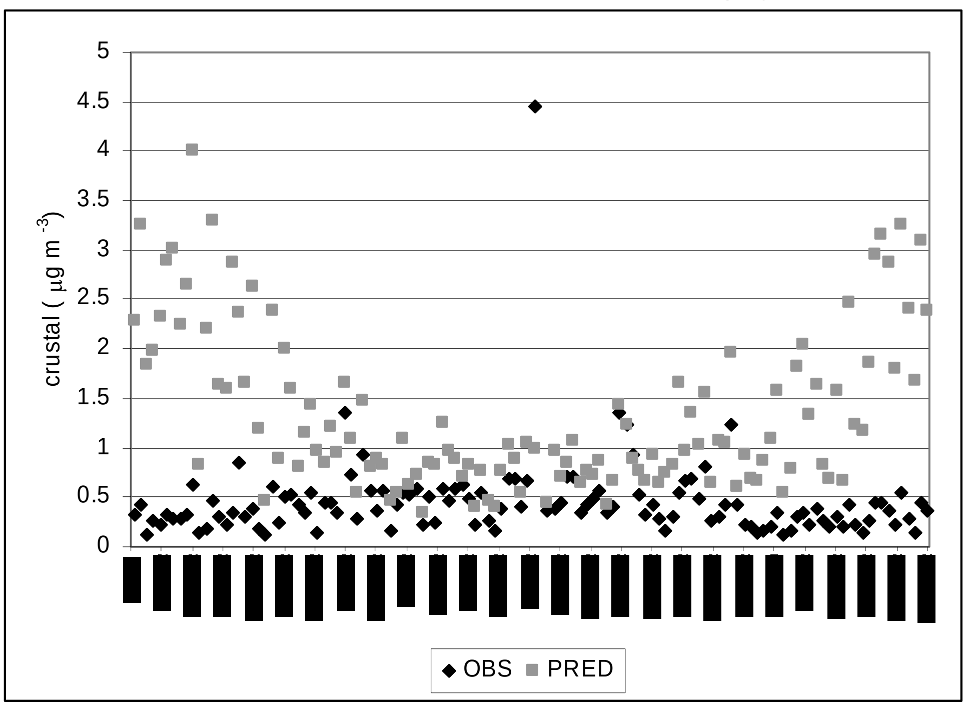


Figure 19. Quarterly variation in observed OM and predicted components of OM – anthropogenic primary organic aerosol (POA), anthropogenic secondary organic aerosol (SOA), and biogenic SOA – averaged over sites in the NYC NAA.

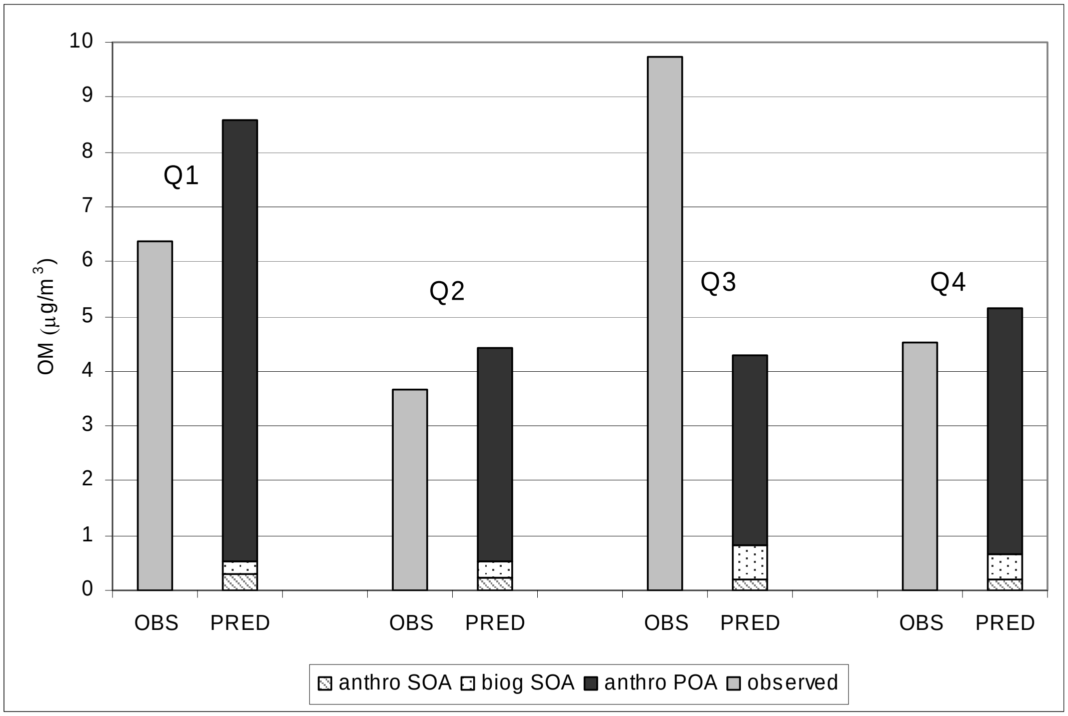


Figure 20. Composite average diurnal variation in observed (black line) and predicted (gray line) $PM_{2.5}$ mass over the 28 hourly monitors in the NYC NAA, entire year.

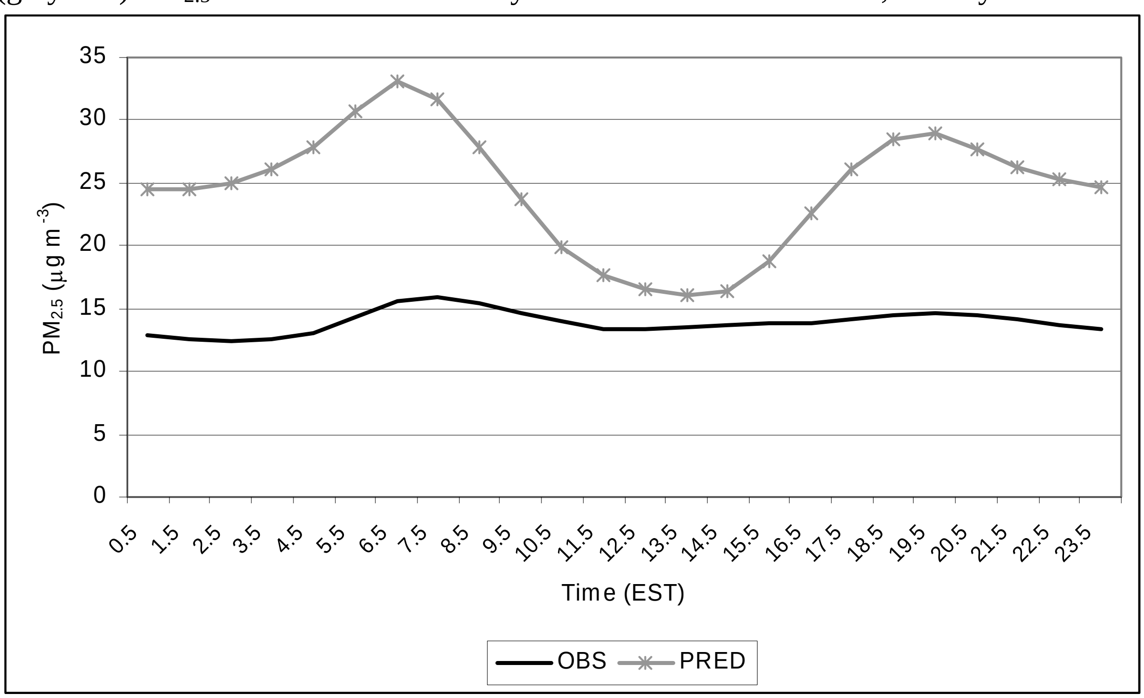


Figure 21. Composite average diurnal variation in observed (black line) and predicted (gray line) $PM_{2.5}$ mass over the 28 hourly monitors in the NYC NAA, quarter 1.

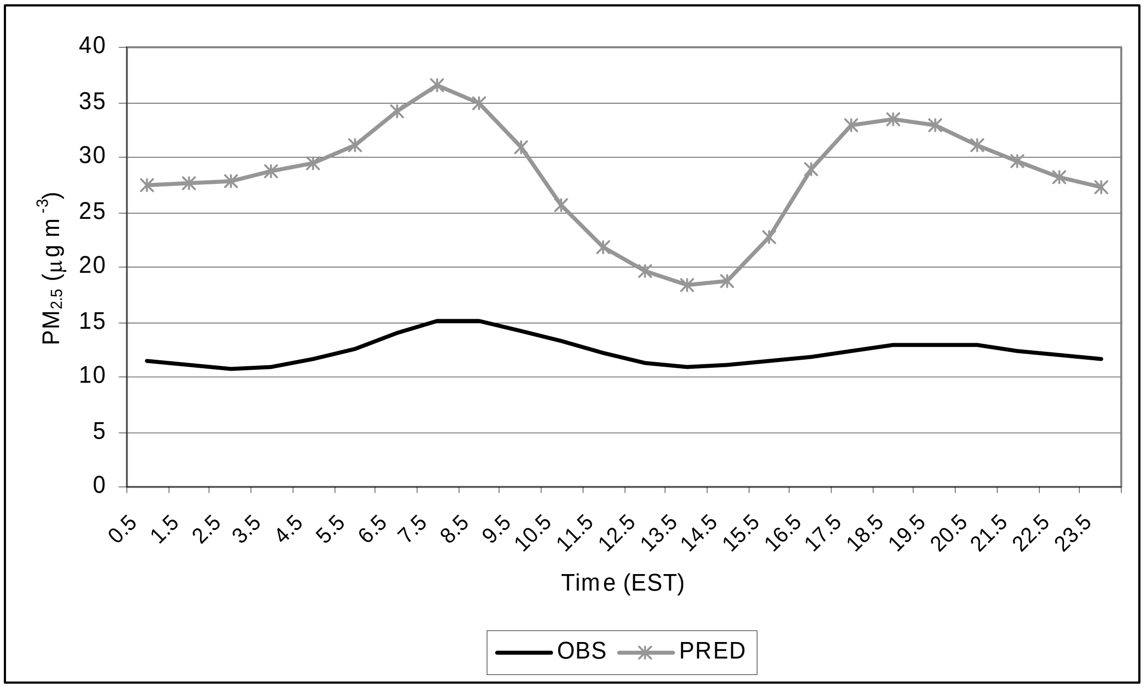


Figure 22. Composite average diurnal variation in observed (black line) and predicted (gray line) $PM_{2.5}$ mass over the 28 hourly monitors in the NYC NAA, quarter 3.

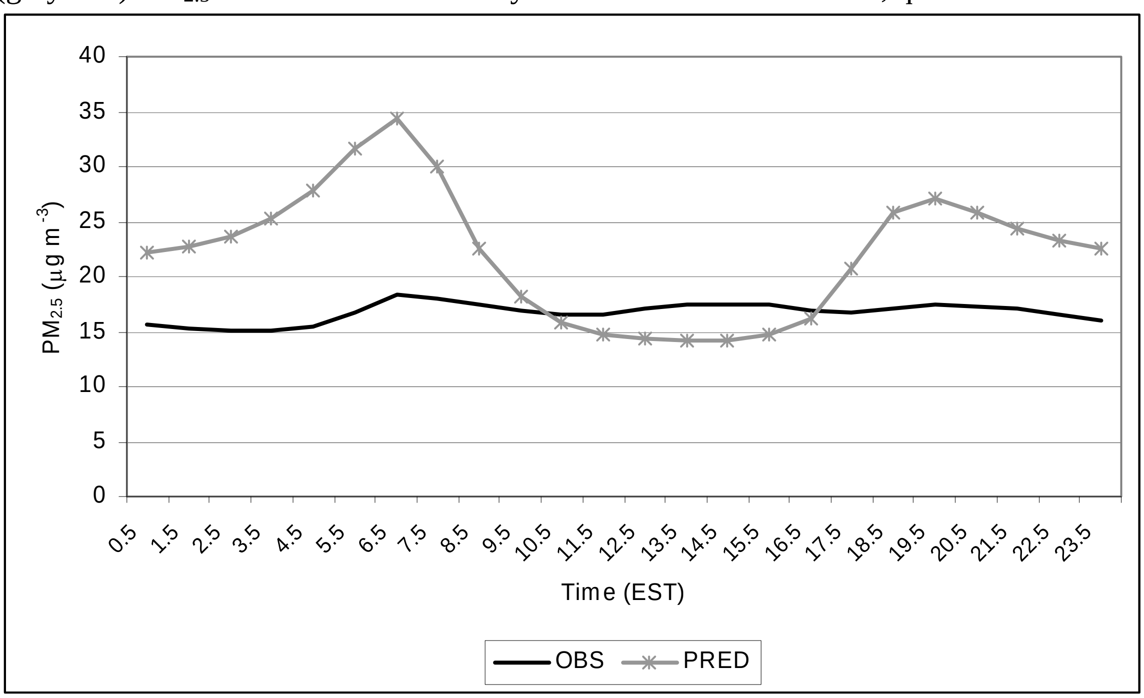


Figure 23. Average diurnal variation at IS52 (Bronx, NY) of SO_4, entire year.

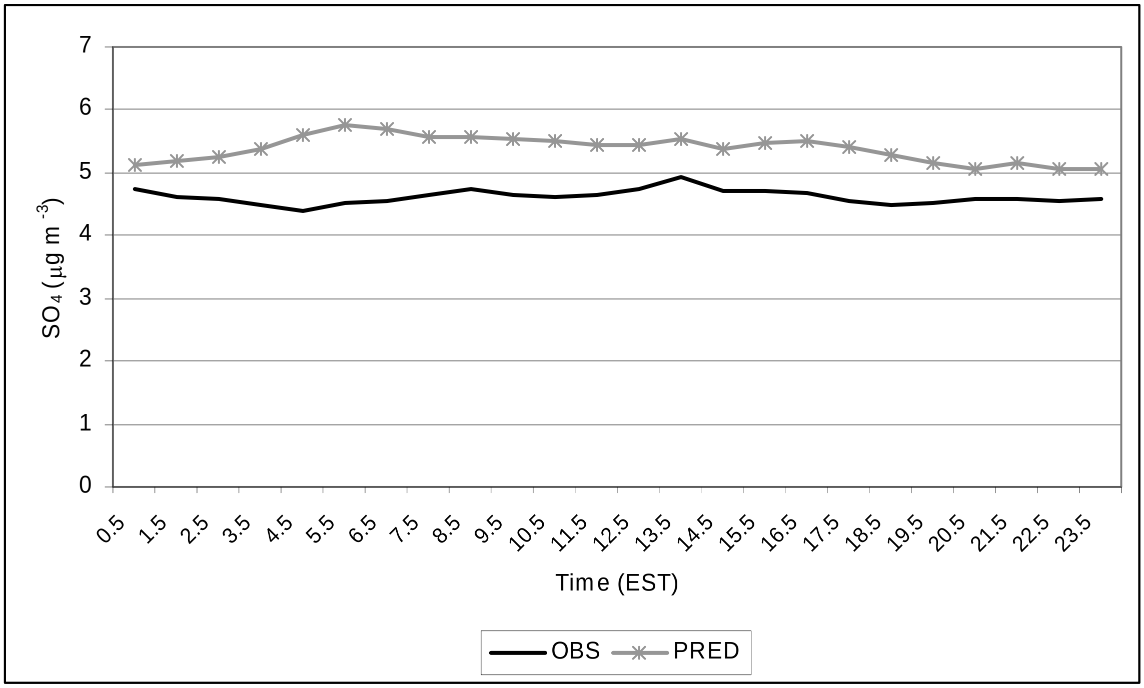
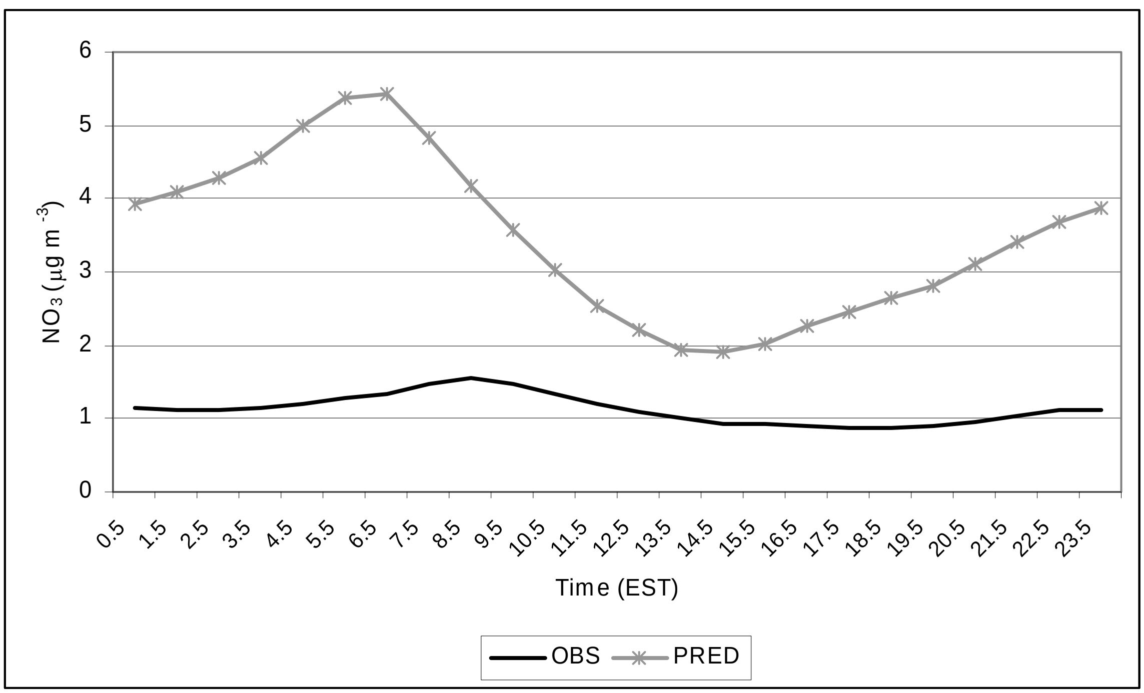


Figure 24. Same as Figure 23, except for NO_3.



APPENDIX H-2

METEOROLOGICAL MODELING OF 2002 USING PENN STATE/NCAR 5TH GENERATION MESOSCALE MODEL (MM5)

Bureau of Air Quality
Department of Environmental Protection

This page blank for copying purposes.

TSD-1a

Meteorological Modeling of 2002 using Penn State/NCAR 5th Generation Mesoscale Model (MM5)

Bureau of Air Quality Analysis and Research

Division of Air Resources

New York State Department of Environmental Conservation

Albany, NY 12233

March 19, 2006

Meteorological Modeling using Penn State/NCAR 5th Generation Mesoscale Model (MM5)

Version 3.6 of MM5 was used to generate annual 2002 meteorology for the OTC modeling work. Prof. Dalin Zhang of the University of Maryland performed the MM5 simulations in consultation with NYSDEC staff. The model was applied in Lambert conformal map projection and utilized MPP Version developed for clusters. The two-way nested domain consisted of coarse (36km) and fine (12km) mesh corresponding to 149x129 and 175x175 grids, respectively, in this application (see Figure 1).

The Lambert projection used in this work followed the Regional Planning Organization (RPO) national domain setup with the center at (40°N, 97°W) and parallels at 33°N and 45°N. Map projection parameters in reference to the projection center point are as follows: Southwest corner for the 36 km grid is at (-2664km, -2304km) and the northeast corner at (2664km, 2304km). In the case of the 12km grid, the southwest corner is at (252km, -900km) and the northeast corner at (2340km, 1188km). In the vertical direction, the terrain following σ-coordinate system was used with the pressure at each σ-level determined from a reference state that is estimated using the hydrostatic equation from a given sea-level pressure and temperature with a standard lapse rate. There are 30 unevenly spaced σ levels, giving 29 vertical layers, with higher resolution within the planetary boundary layer (PBL). The σ levels are:

1.0000, 0.9974, 0.9940, 0.8980, 0.9820, 0.9720, 0.9590, 0.9430, 0.9230, 0.8990,

0.8710, 0.8390, 0.8030, 0.7630, 0.7180, 0.6680, 0.6180, 0.5680, 0.5180, 0.4680,

0.3680, 0.3180, 0.2680, 0.2180, 0.1680, 0.1230, 0.0800, 0.0400, 0.0000

The surface layer was set at about 10m, the level at which surface winds were typically observed, and the model top was set at 50hPa with a radiative top boundary condition. The time steps for the 36km and 12km domains were 75 and 25 seconds, respectively.

The important model physics options used for this MM5 simulation include:

- Kain-Fritsch (1993) convective scheme for both 36- and 12-km domains
- Explicit moisture scheme (without the mixed phase) containing prognostic equations for cloud water (ice) and rainwater (snow) (Dudhia 1989; Zhang 1989)
- Modified version of the Blackadar planetary boundary layer (PBL) scheme (Zhang and Anthes 1982; Zhang and Zheng 2004)
- Simple radiative cooling scheme (Grell et al. 1994)
- Multi-layer soil model to predict land surface temperatures using the surface energy budget equation (Dudhia 1996)

Note that the Blackadar PBL scheme has been modified in order to correct the phase shift of surface wind speed and temperature diurnal cycle, following a study that compared five different PBL schemes: the Gayno-Seaman TKE scheme (Shafran et al. 2000), Burk-

Thompson (1989), Blackadar (Zhang and Anthes 1982), MRF (Hong and Pan 1996), and Mellor-Yamada-Jajic (Mellor and Yamada 1974; Jajic 1990, 1994). The details of the study can be found at Zhang and Zheng (2004).

Nudging Processes

The MM5 provides options for nudging observations for each domain during the model integration process (Stauffer and Seaman, 1990; Stauffer et al. 1991). The Eta analyses of upper-air winds, temperature and water-vapor mixing ratio as well as their associated surface fields were used for nudging every 6 hours, and the Eta surface wind fields blended with surface wind observations were used to nudge every 3 hours. While only the surface winds were nudged, their influences could extend into the PBL as well (see Stauffer et al. 1991). Based on UMD's prior experience in numerical experiments, the following nudging coefficients have been used:

- Upper-air wind fields: 5. 0E-4s^{-1} for Domain 1 (36km), and 2. 5E-4s^{-1} for Domain 2 (12km);
- Upper-air temperature fields: 1.0E-5s^{-1} for both Domains;
- Surface winds: 5. 0s^{-1}E-4s^{-1} for Domain 1, and 2.5E-4s^{-1} for Domain 2; and
- Surface temperature and moisture: not nudged due to instability consideration.

ASSESSMENT

This assessment covers the period of May through September 2002.

National Weather Service (NWS) and CASTNet data – Surface temperature, Wind Speed, and Humidity

NWS (TDL) and CASTNet (www.epa.gov/castnet/) surface measurements of temperature, wind speed, and humidity (note there were no humidity measurements for CASTNet) were used to compare with the MM5 outputs. The evaluation was performed with METSTAT program developed by Environ Corporation (www.camx.com/files/metstat.15feb05.tar.gz) When comparing to NWS data, the METSTAT interpolates the first layer MM5 (at 10m height) temperature and humidity data to a height of 2m, the level that corresponds to the NWS measurement of these parameters. However, no such interpolation was made for wind speed and direction. In the case of CASTNet surface measurements, no such changes were needed as CASTNet data were reported at a height of 10m. In this analysis, no exclusion was made for calm conditions. The reported calm winds (zero wind speed measured) were treated as is in this evaluation effort. The METSTAT calculated standard statistical measures – average, bias, error and index of agreement between the measured and predicted parameters.

Figure 2 displays the temperature and wind speed comparison of MM5 and measured data from NWS and CASTNet networks for August 2002. MM5 performance for both in magnitude and diurnal timing, temperature can be considered to be quite good for both NWS and CASTNet data, while MM5 underpredicted NWS and overpredicted CASTNet daytime wind speed, respectively. It should be pointed out that there are differences in how the meteorological information is collected and reported by the two networks as well as in MM5. The CASTNet measurements are based on hourly averaged wind speed while NWS reports 2min average at 10min before the hour, whereas MM5 predictions are reflective of the last time-step of the hour of computation. Interestingly, MM5 appears to track quite well the nighttime minimum wind speed for both networks. In the case of humidity (not shown), MM5 tracks the NWS observed humidity trend well, but MM5 missed the observed semi-diurnal cycles. Comparisons for the five months including bias and root mean square error from both NWS and CASTNet are available on request from NYSDEC.

The above assessment is based on domain-wide averages to provide an overall response of the model over the five months. Another way of assessing the model is to examine the degree of correlation between the measured and predicted parameters. Figures 3a and 3b displays such a comparison for wind speed and temperature, respectively, for the NWS hourly data covering the period of May through September 2002. For the NWS data, the correlations are in the range from 0.7 to 0.8 for wind speed, above 0.96 for temperature, and in the range of 0.8 to 0.9 for humidity. CASTNet data (not shown) also exhibit similar correlation. These correlations indicate that MM5 simulation has captured both the diurnal and synoptic scale variations. Detailed plots of this comparison are available on request from NYSDEC.

Vertical Profiler – Winds

The Wind-Profiler network measurements along the U. S. East Coast (www.madis-fsl.org/cap) were used to evaluate the vertical profiles from MM5. There are twelve wind-profiler measurement stations from which data were available for comparison. For convenience of comparison, the wind-profiler measurements were interpolated to the MM5 vertical levels. The approach used was simple interpolation between two adjacent wind-profiler layers to the MM5 vertical level, and was limited to that reported by the profiler measurement. The focus of the comparison was to assess if MM5 was able to capture the measured vertical structure, and for this we used the observed Low Level Jet (LLJ) as an indicator. The comparison was performed for June, July and August 2002. In general it is found that MM5 captures the profiler measured vertical wind field structure reasonably well. Figure 4 displays an example of the MM5 and wind profiler comparison for the August 2002 episode at Richmond, VA and Concord, NH. MM5 predicted weaker LLJ winds compared to those based on the wind-profiler measurements. The detailed plots of this comparison are available on request from NYSDEC.

Cloud Cover – Satellite cloud image

Cloud information derived from satellite image data (www.atmos.umd.edu/~srb/gcip/webgcip.htm)were used to assess the MM5 prediction of cloud cover. The 0.5^o by 0.5^o resolution of the satellite data were interpolated into the 12km MM5 grid for comparison. The MM5 total cloud fraction was estimated by MCIP based on the MM5's low cloud, middle cloud and high cloud predictions. In general, MM5 captured the satellite cloud pattern well but underestimates the satellite cloud fraction (see Figure 5 as an example). Part of problem may due to the coarse resolution of the satellite cloud data.

Precipitation comparison

The monthly total observed precipitation data were constructed from 1/8-degree daily precipitation analysis data (http://data.eol.ucar.edu/codiac/dss/id=21.093 produced by Climate Prediction Center, based on 7,000-8,000 hourly/6-hourly gauge reports and radar). The MM5 monthly total precipitation was estimated from the MM5 predicted convective and non-convective rainfall and summed up for each month. In general, MM5 captured the observed spatial patterns in May and September, but no so well for June, July and August (See Figure 6), perhaps reflective of the summertime convective rain activities not captured by MM5. Detailed plots of this comparison are available on request from NYSDEC.

Calm Conditions

Calm conditions are defined as observed wind speed of zero knots and wind direction as 0^{o}. It would be useful to assess how MM5 performs under observed calm conditions, because of potential pollutant buildup that could occur under such conditions. Table 1 lists the summary of the percentage of calm condition at each hour for the August 2002 from the NWS data within the 12km domain. It is apparent from the Table that the calm conditions occur primarily during the night and early morning hours, from 23Z (7 p.m. EDT) to 15Z (11 a.m. EDT) with a peak at 10Z (6 a.m. EDT). To assess MM5 performance, the observed and MM5 predicted wind speeds were divided into calm and non-calm according to observed wind speed. Figure 7 displays such a comparison of the MM5 predicted wind speed to the observed wind speed under the calm and non-calm conditions for the month of August 2002. For the "calm" group, the average wind speed for MM5 varies from 1 m/s during the night and early morning hours and over 1.5 m/s during the day. MM5 is over-predicting during observed calm wind conditions. There are local minima every 3 hours, due to the surface observed wind speed nudging in MM5. In contrast under the non-calm conditions, MM5 underpredicts by about 0.5 m/s for all hours with noticeable local maximum happening at the nudging hours. The MM5 nudging process would pull predictions toward the measured data, while the underprediction of MM5 for the non-calm conditions may due to the adopted PBL scheme in this simulation.

Summary

In this study, we performed an assessment of the MM5 simulation to real-world data, both at the surface level as well as in the vertical. While there are no specific recommended procedures identified for this assessment, similar approaches have been used elsewhere (Dolwick 2005, Baker 2004, and Johnson 2004). Traditionally, the NWS surface measurements are used for such a comparison. Since NWS data had been used through nudging processes in developing the MM5 simulation, the comparisons should not be far removed from each other. In this study, we extended the evaluation by using CASTNet measurements that were not used in the MM5 simulations. Thus comparison with CASTNet data provides for an independent assessment and should complement the comparison with NWS data. We also compared the MM5 results with the wind profiler data and cloud data derived from satellite images to diagnose if the MM5 simulation is yielding the right type of dynamics in the vertical. The analyses shows that in general, the performance of the MM5 is reasonable both at the surface and in the vertical, thereby providing confidence in the use of these data in the CMAQ simulations.

References

Baker, K. 2004: www.ladco.org/tech/photo/photochemical.html

Burk, S. D. and W. T. Thompson, 1989: A vertically nested regional numerical weather prediction model with second-order closure physics. Mon. Wea. Rev., 117, 2305–2324.

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiments using a mesoscale two-dimensional model. J. Atmos. Sci., 46, 3077–3107.

Dudhia, J., 1996: A multi-layer soil temperature model for MM5. Preprints, 6th Annual MM5 Users Workshop, Boulder, CO.

Dolwick, P. 2005: http://cleanairinfo.com/modelingworkshop/presentations/MPE_Dolwick.pdf

Grell, G. A., J. Dudhia , and D. R. Stauffer 1994: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech.Note NCAR/TN-398 1 STR, 122 pp.

Hong, S.-H., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. Mon. Wea. Rev., 124, 2322–2339.

Jajic, Z. I., 1990: The step-mountain coordinate: Physical package. Mon. Wea. Rev., 118, 1429-1443.

Jajic, Z. I., 1994: The step-mountain Eta coordinate model: Further development of the convection, viscous sublayer and turbulent closure schemes. *Mon. Wea. Rev.*, 122, 927-945.

Johnson, M. 2004: www.ladco.org/tech/photo/photochemical.html

Kain, J.S., and J.M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain-Fritsch scheme. Cumulus Parameterization. Meteor. Monogr., 46, Amer. Meteor. Soc., 165-170.

Mellor, G. L., and T. Yamada, 1974: A hierarchy of turbulence closure models for planetary boundary layers. J. Atmos. Sci., 31, 1791–1806.

Shafran, P.C., N.L. Seaman, and G. A. Gayno, 2000: Evaluation of numerical predictions of boundary layer structure during the Lake Michigan ozone study. J. Appl. Meteor., 39, 412-426.

Stauffer, D. R., N. L. Seaman and F. S. Binkowski 1991: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part II: Effects of data assimilation within the planetary boundary layer. Mon. Wea. Rev., 119, 734-754.

Stauffer, D. R. and N. L. Seaman, 1990: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. Mon. Wea. Rev., 118, 1250-1277.

Zhang, D.-L., 1989: The effect of parameterized ice microphysics on the simulation of vortex circulation with a mesoscale hydrostatic model. Tellus, 41A, 132-147.

Zhang, D.-L, and R. A. Anthes, 1982: A high-resolution model of the planetary boundary layer-sensitivity tests and comparisons with SESAME-79 data. J. Appl. Meteor., 21, 1594–1609.

Zhang, D.-L, and W.-Z. Zheng, 2004: Diurnal cycles of surface winds and temperatures as simulated by five boundary-layer parameterizations. J. Appl. Meteor., 43, 157-169.

Table 1 Measured calm and non-calm occurrences over the modeling domain during August 2002 based on NWS data

Hour	#Non-Calm	#Calm	#Total	% Calm
00Z	18209	3924	22133	17.7
01Z	16531	6026	22557	26.7
02Z	15604	6929	22533	30.8
03Z	14983	7245	22228	32.6
04Z	14309	7540	21849	34.5
05z	14073	7735	21808	35.5
06Z	13934	7949	21883	36.3
07Z	13792	8040	21832	36.8
08Z	13542	8273	21815	37.9
09Z	13542	8385	21927	38.2
10Z	13708	8591	22299	**38.5**
11Z	14139	8693	22832	38.1
12Z	15297	7690	22987	33.5
13Z	17336	5192	22528	23
14Z	18522	3439	21961	15.7
15Z	18755	2617	21372	12.2
16Z	19169	2015	21184	9.5
17Z	19555	1617	21172	7.6
18Z	19982	1430	21412	6.7
19Z	20149	1389	21538	6.4
20Z	20565	1288	21853	5.9
21Z	20518	1383	21901	6.3
22Z	20672	1556	22228	7
23Z	20231	2292	22523	10.2

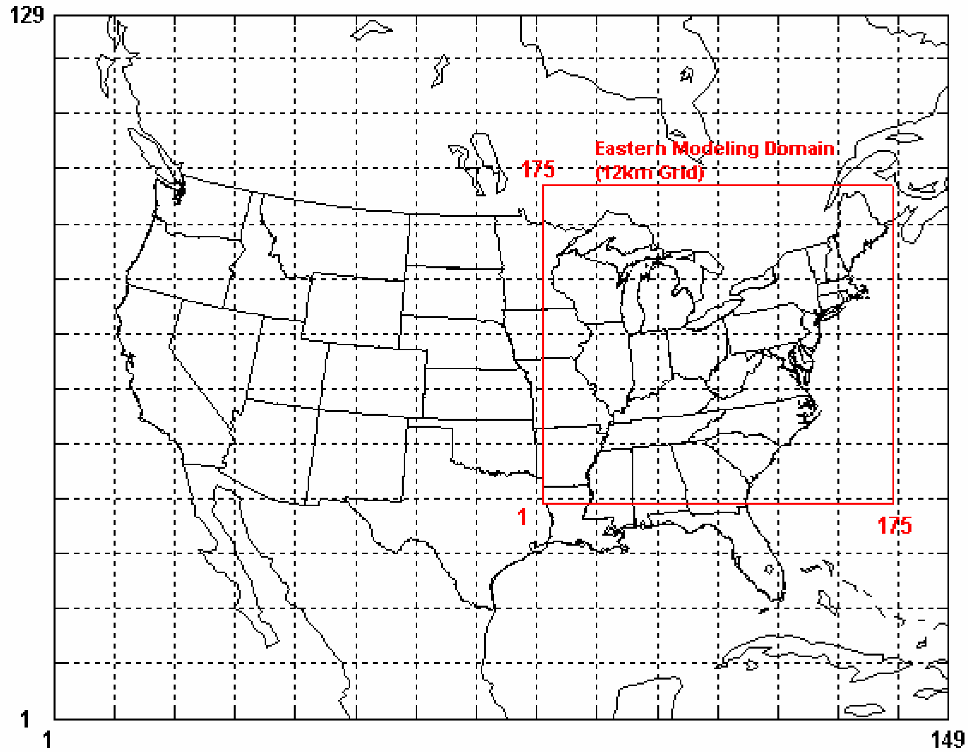


Figure 1: OTC MM5 modeling domain with areal extent of 12km and 36km grids

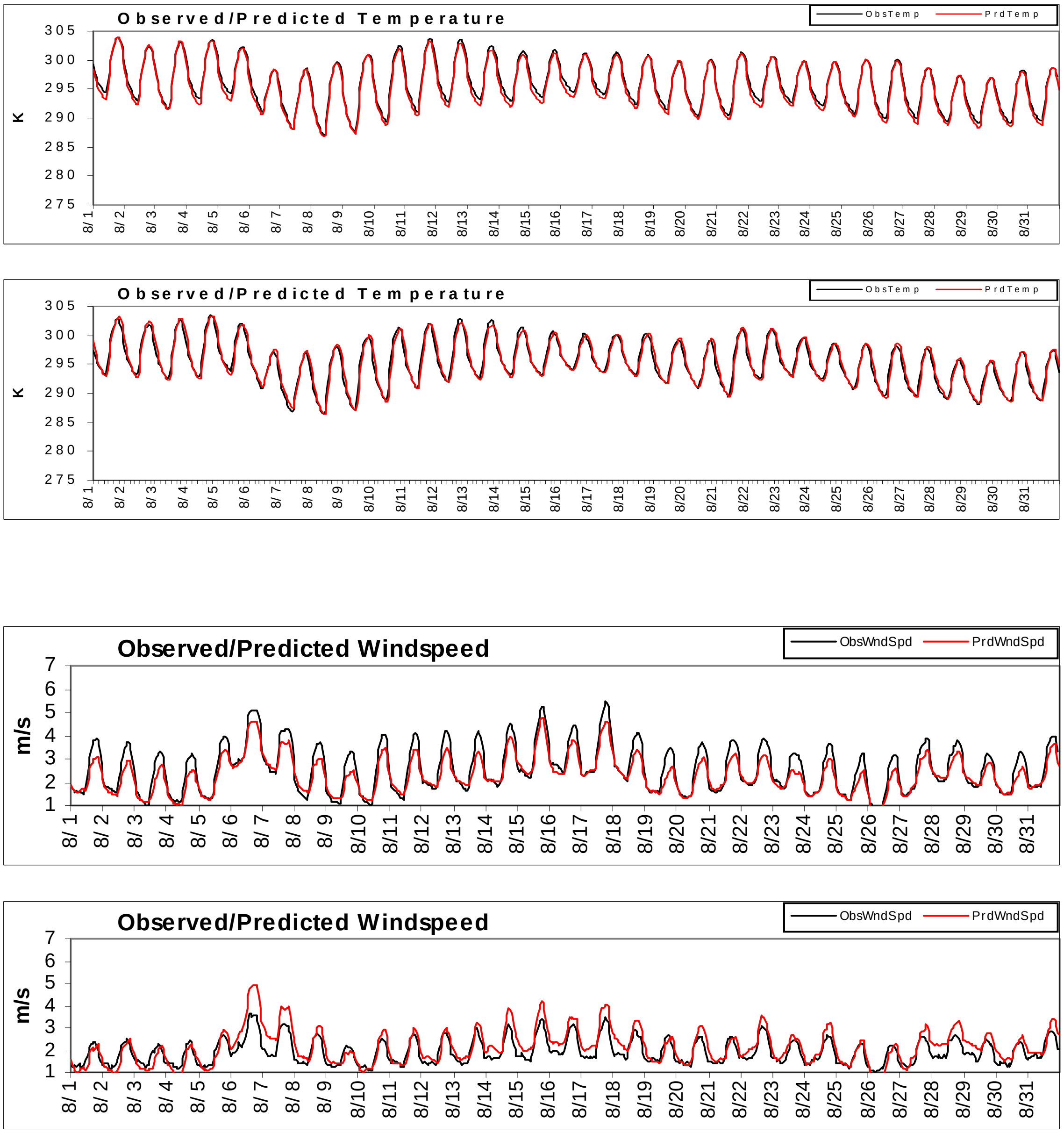


Figure 2: Temperature and Wind speed comparisons for August 2002. In each case the upper panel corresponds to comparison between MM5 and NWS data and the lower panel between MM5 and CASTNet data.

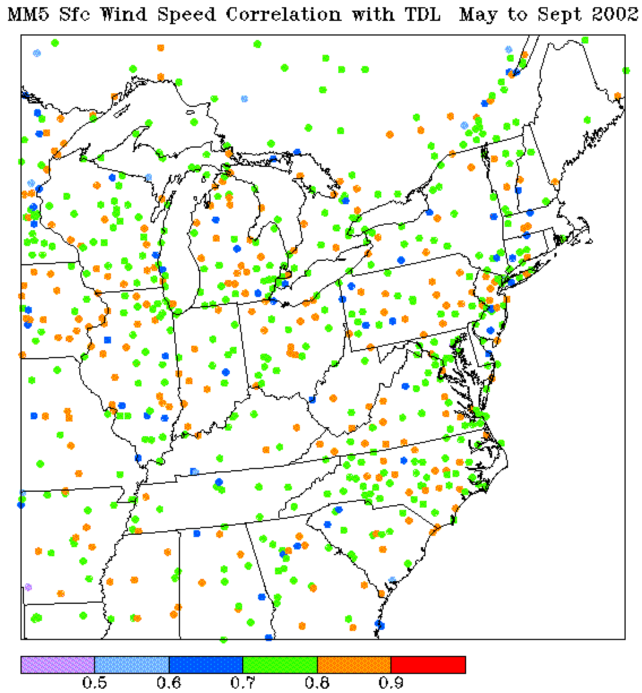


Figure 3a: Spatial correlation estimates between MM5 and NWS data for wind speed from May to September 2002

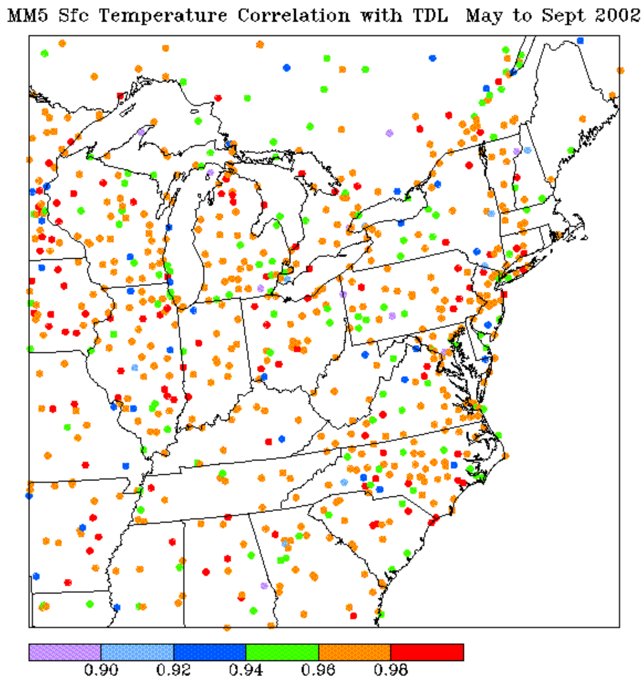
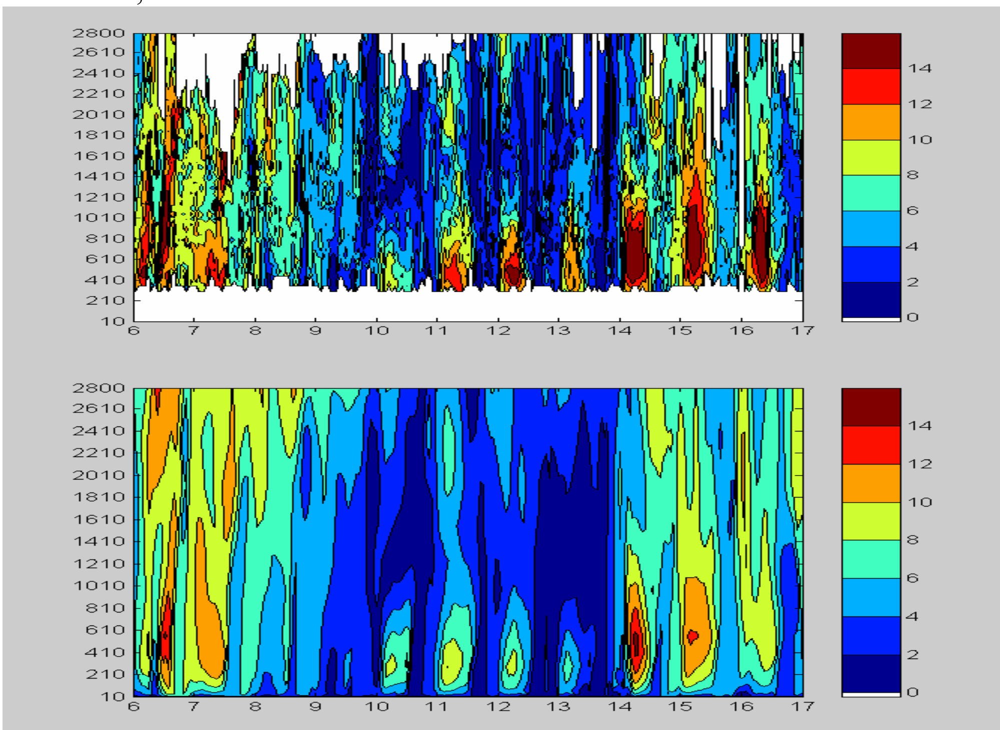


Figure 3b: Spatial distribution of correlation coefficients for Temperature between MM5 and NWS data from May to September 2002.

Richmond, VA



Concord, NH

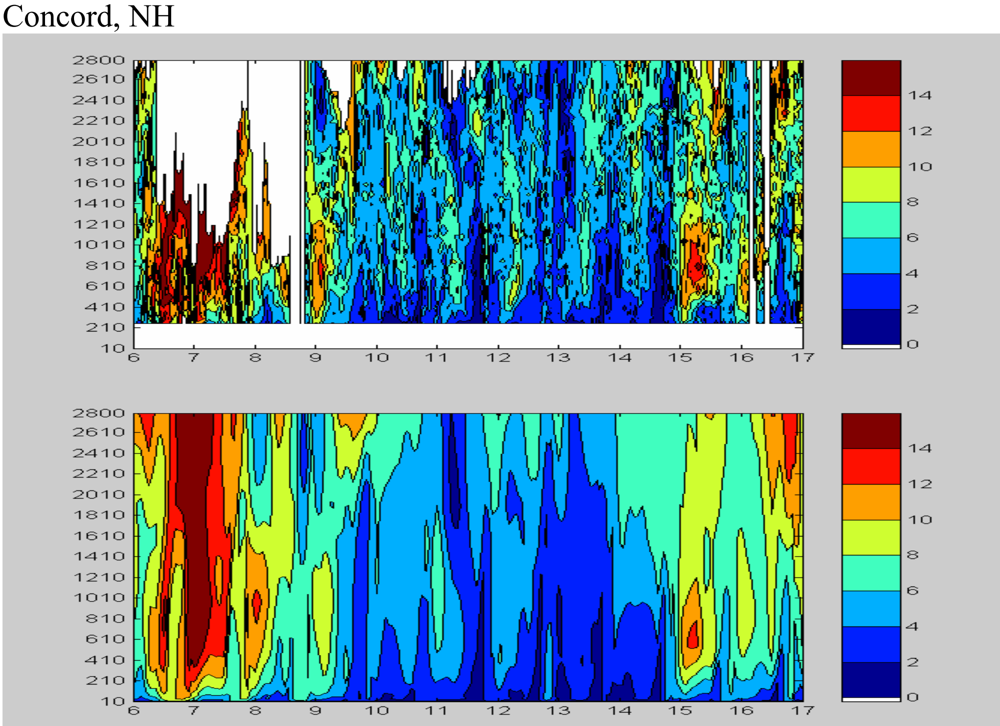


Figure 4: MM5 and Wind profiler comparison for August 6 to 17, 2002 at Richmond, VA and Concord, NH. The upper and lower panes at each station are for MM5 and profiler, respectively. The abcissa represents day and the ordinate the height (m).

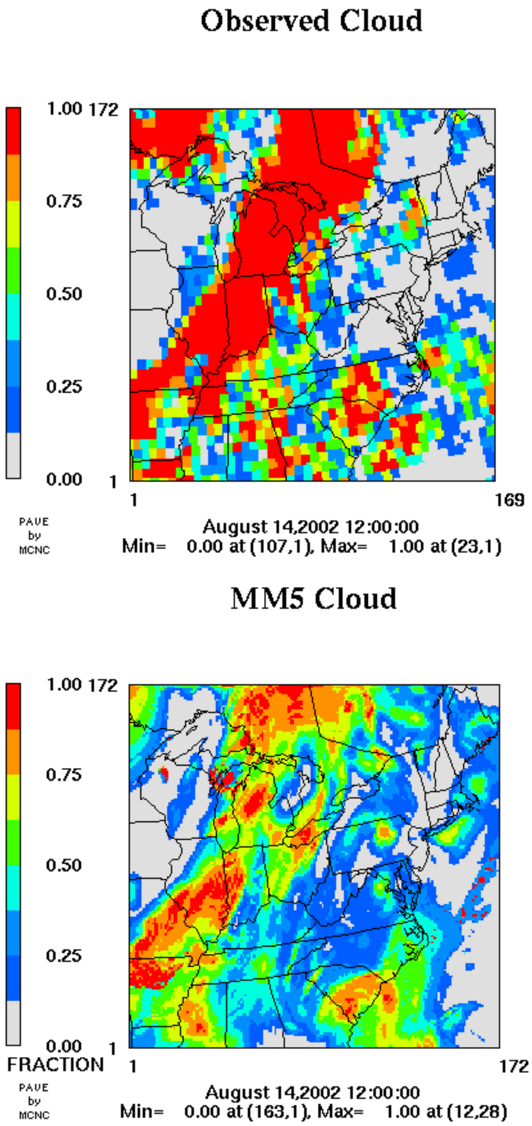


Figure 5: MM5 and Satellite cloud images for August 14, 2002 at 0700 EST

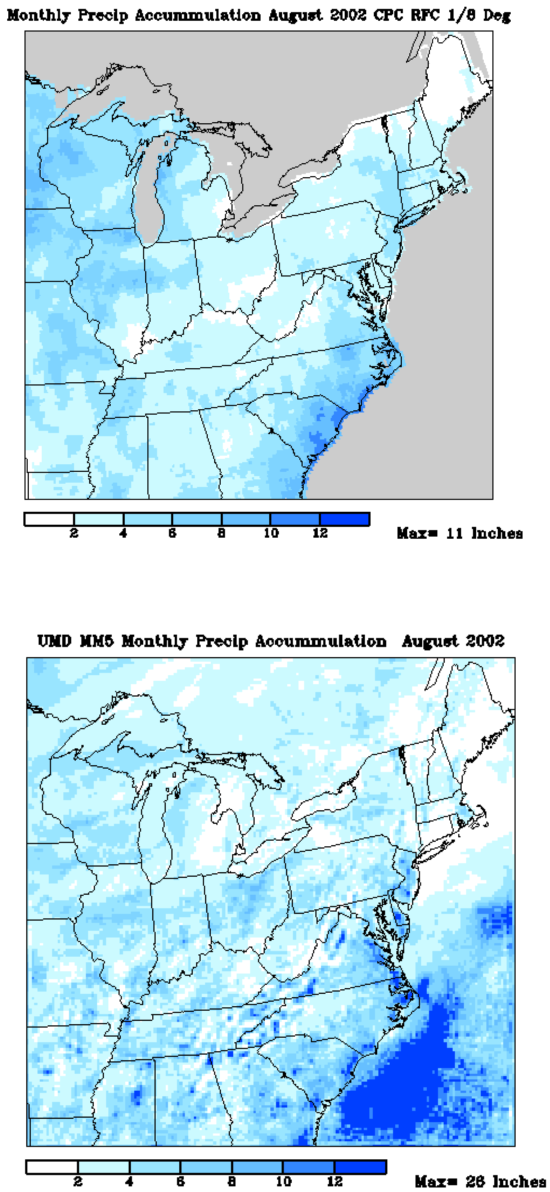


Figure 6: MM5 predicted and measured precipitation over the domain for the month of August 2002

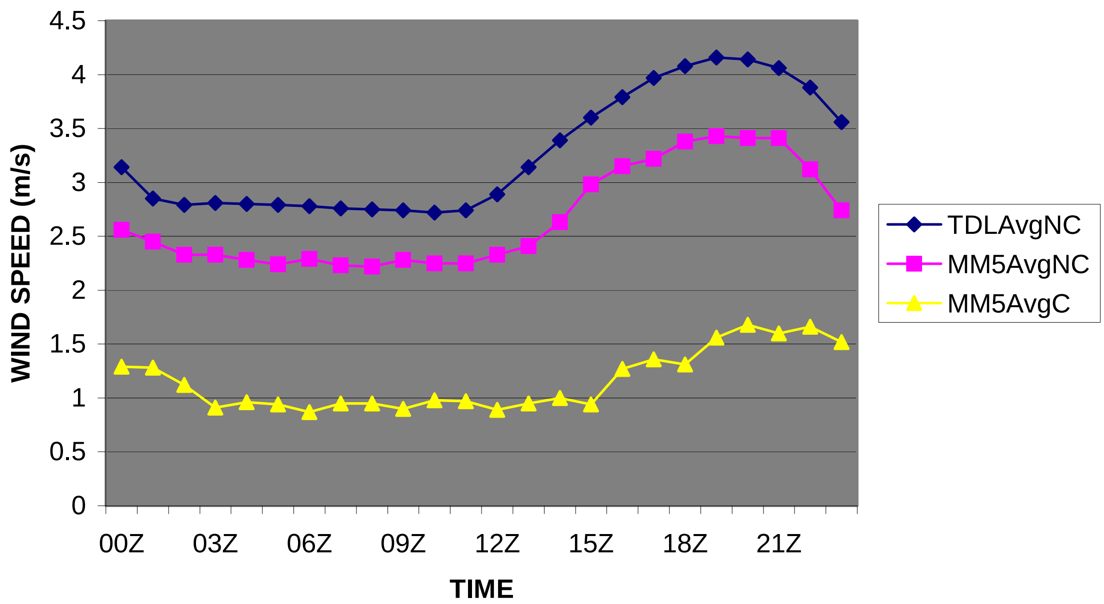


Figure 7: Comparison of averaged wind speed between MM5 and observed under calm (C) and non-calm (NC) conditions.

APPENDIX H-3

PROCESSING OF 2002 ANTHROPOGENIC EMISSIONS: OTC REGIONAL AND URBAN 12KM BASE YEAR SIMULATION

Bureau of Air Quality
Department of Environmental Protection

This page blank for copying purposes.

TSD-2b

Processing of 2002 Anthropogenic Emissions: OTC Regional and Urban 12km Base Year Simulation

Bureau of Air Quality Analysis and Research
Division of Air Resources
New York State Department of Environmental Conservation
Albany, NY 12233

March 19, 2007

Overview

All emissions processing for the revised 2002 OTC regional and urban 12 km base case simulations was performed with SMOKE2.1 compiled on a Red Hat 9.0 Linux operating system with the Portland group fortran compiler version 5.1. The emissions processing was performed on a month-by-month and RPO-by-RPO basis, i.e. SMOKE processing was performed for each month for each of the RPOs (MANE-VU, VISTAS, CENRAP, MRPO) individually as well as for Canada. For each month/RPO combination, a separate SMOKE ASSIGNS file was created, and the length of the episode in each of these ASSIGNS files was set to the entire month. Also, as discussed in Section 3, there was no difference between "episode-average" temperatures and "monthly-average" temperatures for the Mobile6 simulations that used the option of temperature averaging.

This document is structured as follows: A listing of all emission inventories is given in Section 2, organized by RPO and source category. Section 3 discusses the Mobile6 processing approach employed for the different RPOs, while Section 4 describes the processing of biogenic emissions with BEIS3.12. Finally, Sections 5 through7 describe the temporal allocation, speciation, and spatial allocation of the emissions inventories, respectively.

1. Emission Inventories

1.1 MANE-VU

Version 3 of the MANE_VU inventory was utilized to generate CMAQ-ready emissions. This emissions inventory data were obtained from the MANEVU archive in April 2006.

1.1.1 Area Sources

- Files: MANEVU_AREA_SMOKE_INPUT_ANNUAL_SUMMERDAY_040606.txt and MANEVU_AREA_SMOKE_INPUT_ANNUAL_WINTERDAY_040606.txt prepared by PECHAN, downloaded from **ftp.marama.org** (username mane-vu, password exchange)
- Fugitive dust correction: This was applied as county-specific correction factors for SCC's listed at http://www.epa.gov/ttn/chief/emch/invent/index.html#dust; the correction factor file gcntl.xportfrac.txt was obtained from EPA's CAIR NODA ftp site http://www.airmodelingftp.com (password protected).; this adjustment was performed using the SMOKE programs cntlmat and grwinven to generate an adjusted IDA inventory file used for subsequent SMOKE processing

1.1.2 Nonroad Sources

- File: MANEVU_NRD2002_SMOKE_030306 prepared by PECHAN; downloaded from **ftp.marama.org** (username mane-vu, password exchange)

1.1.3 Mobile Sources

- VMT/Speed: MANEVU_2002_mbinv_02022006_addCT.txt prepared by PECHAN and NESCAUM; downloaded from http://bronze.nescaum.org/Private/junghun/MANE-VU/onroad_ver3_update/MANEVU_V3_update.tar

1.1.4 Point Sources

- Files: MANEVU_Point_SMOKE_INPUT_ANNUAL_SUMMERDAY_041006.txt and MANEVU_Point_SMOKE_INPUT_ANNUAL_WINTERDAY_041006.txt prepared by PECHAN were downloaded from **ftp.marama.org** (username mane-vu, password exchange)
- Fugitive dust correction: This was applied as county-specific correction factors for SCC's listed at http://www.epa.gov/ttn/chief/emch/invent/index.html#dust; the correction factor file gcntl.xportfrac.txt was obtained from EPA's CAIR NODA ftp site http://www.airmodelingftp.com (password protected).; this adjustment was performed using the SMOKE programs cntlmat and grwinven to generate an adjusted IDA inventory file used for subsequent SMOKE processing
- Corrected the omission of 2,100 tons/year VOC emissions from several point sources in NJ. NJDEP provided updated IDA files on June 30 that were used for modeling.

1.2 CENRAP

The inventory data were obtained from the CENRAP ftp site in March 2006 and reflect version BaseB of the CENRAP inventory.

1.2.1 Area Sources

- Files:
  - CENRAP_AREA_SMOKE_INPUT_ANN_STATES_081705.txt
  - CENRAP_AREA_MISC_SMOKE_INPUT_ANN_STATE_071905.txt
  - CENRAP_AREA_BURNING_SMOKE_INPUT_ANN_TX_ NELI_071905.txt
  - CENRAP_AREA_MISC_SMOKE_INPUT_NH3_MONTH_{MMM} _072805.txt where {MMM} is JAN, FEB, … DEC
  - CENRAP_AREA_SMOKE_INPUT_NH3_MONTH_{MMM} _071905.txt where {MMM} is JAN, FEB, … DEC
- Fugitive dust correction: This was applied as county-specific correction factors for SCC's listed at http://www.epa.gov/ttn/chief/emch/invent/index.html#dust; the correction factor file gcntl.xportfrac.txt was obtained from EPA's CAIR NODA ftp site http://www.airmodelingftp.com (password protected).; this adjustment was performed using the SMOKE programs cntlmat and grwinven to generate an adjusted IDA inventory file used for subsequent SMOKE processing
- Note about area and nonroad source SMOKE processing for the CENRAP region: All area source inventories (both annual and month-specific) were processed in

one step through SMOKE. SMK_AVEDAY_YN was set to N, so seasonal profiles were used to apportion the annual inventories numbers by month. This setting was also used for the nonroad processing performed in a separate step. This was necessary since the month-specific files had zero in their 'average-day' column and the annual total column reflects the "monthly emissions as annual totals" as per header line. Therefore, seasonal profiles are used to apportion both the annual and month-specific files. As described below, we utilized the temporal profiles and cross-reference files generated by CENRAP. However, we did not verify that this approach indeed leads to the intended monthly allocation of ammonia and nonroad emissions.

1.2.2 Nonroad Sources

- Files:
  - CENRAP_NONROAD_SMOKE_INPUT_ANN_071305.txt
  - CENRAP_NONROAD_SMOKE_INPUT_MONTH_{MMM}_071305.txt where {MMM} is JAN, FEB, … DEC

1.2.3 Mobile Sources

- VMT/Speed files:
  - mbinv02_vmt_cenrap_ce.ida
  - mbinv02_vmt_cenrap_no.ida
  - mbinv02_vmt_cenrap_so.ida
  - mbinv02_vmt_cenrap_we.ida

1.2.4 Point Sources

- File: CENRAP_POINT_SMOKE_INPUT_ANNUAL_DAILY_072505.txt
- Fugitive dust correction: This was applied as county-specific correction factors for SCC's listed at http://www.epa.gov/ttn/chief/emch/invent/index.html#dust; the correction factor file gcntl.xportfrac.txt was obtained from EPA's CAIR NODA ftp site http://www.airmodelingftp.com (password protected).; this adjustment was performed using the SMOKE programs cntlmat and grwinven to generate an adjusted IDA inventory file used for subsequent SMOKE processing.

1.3 VISTAS

All VISTAS emission files were obtained from the Alpine Geophysics ftp site. They reflect version BaseG of the VISTAS inventory with the exception of fire emissions which reflect BaseF and BaseD. These files were downloaded between February and August, 2006.

1.3.1 Area Sources

- Files:

    - arinv_vistas_2002g_2453922_w_pmfac.txt
    - ida_ar_fire_2002_vistaonly_basef.ida
- Note: the header lines of these files indicate that the fugitive dust correction was already applied, so no further correction was performed.

1.3.2 Nonroad Sources

- Files:
    - nrinv_vistas_2002g_2453908.txt
    - marinv_vistas_2002g_2453972.txt

1.3.3 Mobile Sources

- VMT/Speed file: mbinv_vistas_02g_vmt_12jun06.txt

1.3.4 Point Sources

- Files:
    - Annual:
        - egu_ptinv_vistas_2002typ_baseg_2453909.txt
        - negu_ptinv_vistas_2002typ_baseg_2453909.txt
        - ptinv_fires_{MM}_typ.vistas.ida where {MM} is 01, 02, 03, etc. depending on the month; these annual point fire files were generated as part of the VISTAS BaseD inventory and were obtained in January 2005
    - Hour-specific:
        - pthour_2002typ_baseg_{MMM}_28jun2006.ems where {MMM} is jan, feb, mar, etc.
        - pthour_fires_{MM}_typ.vistas.ida where {MM} is 01, 02, 03, etc. depending on the month; these hourly point fire files were generated as part of the VISTAS BaseD inventory and were obtained in January 2005
- Note: No fugitive dust correction was performed for these files.

1.4 MRPO

MRPO emissions for SMOKE modeling were generated by Alpine Geophysics through a contract from MARAMA to convert the MRPO BaseK inventory from NIF to IDA format. The files were downloaded from the MARAMA ftp site **ftp.marama.org** (username mane-vu, password exchange) between April and June 2006.

1.4.1 Area Sources

- Files:
    - Annual:
        - arinv_mar_mrpok_2002_27apr2006.txt
        - arinv_other_mrpok_2002_20jun2006.txt
    - Month-specific:

    - arinv_nh3_2002_mrpok_{mmm}_3may2006.txt where {mmm} is jan, feb, etc.
    - dustinv_2002_mrpok_{mmm}_23may2006.txt where {mmm} is jan, feb, etc.
- Fugitive dust correction: This correction was performed only to the arinv_other_mrpok_2002_20jun2006.txt file using county-specific correction factors for SCC's listed at http://www.epa.gov/ttn/chief/emch/invent/index.html#dust; the correction factor file gcntl.xportfrac.txt was obtained from EPA's CAIR NODA ftp site http://www.airmodelingftp.com (password protected).; this adjustment was performed using the SMOKE programs cntlmat and grwinven to generate an adjusted IDA inventory file used for subsequent SMOKE processing.
- Note about area source SMOKE processing: SMOKE processing was performed separately for the annual and month-specific files. For the annual inventory processing, SMK_AVEDAY_YN was set to N, so seasonal profiles were used to apportion the annual inventories numbers by month. For the month-specific inventory processing, this variable was set to Y so that no seasonal profiles would be applied and the inventory numbers in the 'average day' column would be used. To save a SMOKE processing step, the annual "marine" inventory "arinv_mar_mrpok_2002_27apr2006.txt" was processed together with the annual "other area source" inventory "arinv_other_mrpok_2002_20jun2006.txt" even though it technically is part of the nonroad inventory.

1.4.2 Nonroad Sources

- Files: nrinv_2002_mrpok_{mmm}_3may2006.txt where {mmm} is jan, feb, etc.

1.4.3 Mobile Sources

- VMT/Speed file: mbinv_mrpo_02f_vmt_02may06.txt

1.4.4 Point Sources

- Files: ptinv_egu_negu_2002_mrpok_1may2006.txt
- Fugitive dust correction: This correction was performed only to the arinv_other_mrpok_2002_20jun2006.txt file using county-specific correction factors for SCC's listed at http://www.epa.gov/ttn/chief/emch/invent/index.html#dust; the correction factor file gcntl.xportfrac.txt was obtained from EPA's CAIR NODA ftp site http://www.airmodelingftp.com (password protected).; this adjustment was performed using the SMOKE programs cntlmat and grwinven to generate an adjusted IDA inventory file used for subsequent SMOKE processing.

1.5 Canada

1.5.1 Area Sources

- File: AS2000_SMOKEready.txt obtained from ftp://ftp.epa.gov/EmisInventory/canada_2000inventory
- Fugitive dust correction: We applied "divide-by-four" correction for SCC's listed at http://www.epa.gov/ttn/chief/emch/invent/index.html#dust; this adjustment was performed outside SMOKE with in-house Fortran programs. No county/province-specific correction factors were available for Canada

1.5.2 Nonroad Sources

- File: NONROAD2000_SMOKEready.txt obtained from ftp://ftp.epa.gov/EmisInventory/canada_2000inventory

1.5.3 Mobile Sources

- File: MOBILE2000_SMOKEready.txt obtained from ftp://ftp.epa.gov/EmisInventory/canada_2000inventory
- Fugitive dust correction: applied "divide-by-four" correction for SCC's listed at http://www.epa.gov/ttn/chief/emch/invent/index.html#dust; this adjustment was performed outside of SMOKE with in-house Fortran programs. No county/province-specific correction factors were available for Canada.

1.5.4 Point Sources

There has long been difficulty in obtaining an up-to-date Canadian criteria emissions inventory for point sources. This is due largely to confidentiality rights afforded to Canadian facilities. Thus far, the most recent inventory of Canadian point sources is rooted in the 1985 NAPAP data and is close to two decades old. Because there are a number of high emitting industrial facilities in southern Canada it is of particular importance to have a reasonably accurate inventory of these sources especially when modeling air quality over the Northeast and Midwest United States. Toward this end, an effort was made to obtain more recent Canadian point source data and incorporate it into an inventory database, which could then be used for the 2002 OTC air quality modeling.

Perhaps the most accurate and publicly accessible source of Canadian pollutant data is now available from the National Pollutant Release Inventory (NPRI) database. This database contains 268 substances. Facilities that manufacture, process or otherwise use one of these substances and that meet reporting thresholds are required to report these emissions to Environment Canada on an annual basis. The NPRI data are available at Environment Canada's website and can be found at the link http://www.ec.gc.ca/pdb/npri/npri_home_e.cfm. The page hosts an on-line search engine where one can locate emissions by pollutant or location. In addition, the entire database is available for download as an MS Access or Excel file. The NPRI database contains numerous pages with a rather comprehensive list of information. Detailed information is available about each facility, including location, activity and annual emissions. In addition, facilities having stacks with a height of 50 meters or more are required to report stack parameters.

Unfortunately, one of the limitations of the NPRI database for modeling purposes is that the data are only available at the facility level. Emissions models require process level information, so in order to use this data, a few generalizations had to be made. Each facility has a Standard Industrial Classification (SIC) code associated with it; however, emissions models require Source Classification Codes (SCC's). SCC's are of critical importance as the emissions models use these codes for assignment of temporal and speciation profiles. SIC codes describe the general activity of a facility while SCC codes describe specific processes taking place at each facility. While no direct relationship exists between these two codes, a general albeit subjective association can be made.

For the purposes of creating a model-ready inventory file it was necessary to obtain the whole NPRI database. After merging all the necessary components from the NPRI database required in the SMOKE inventory file, the SIC code from each facility was examined and assigned an SCC code. In most cases, only a SCC3 level code was assigned with confidence. While this is admittedly a less than desirable process, it does allow for the use of the most recent emissions from the NPRI database to be used in modeling. Furthermore, having some level of SCC associated with these emissions will ensure that they will be assigned a temporal and speciation profile by the model, other than the default. Once the model-ready inventory file was developed, it was processed through SMOKE.

2. Mobile6 Processing

2.1 MANE-VU

2.1.1 <u>Mobile6 input files</u>

- Month-specific input files were prepared by PECHAN and NESCAUM and were downloaded from http://bronze.nescaum.org/Private/junghun/MANE-VU/onroad_ver3_update/MANEVU_V3_update.tar
- Added the line "REBUILD EFFECTS   :0.10" to each file before the SCENARIO record to override the Mobile6 default setting of 0.9 (90%) for the "chip reflash" effectiveness

2.1.2 <u>SMOKE/Mobile6 auxiliary files</u>

- SMOKE/Mobile6 auxiliary files were prepared by PECHAN and NESCAUM and were downloaded from http://bronze.nescaum.org/Private/junghun/MANE-VU/onroad_ver3_update/MANEVU_V3_update.tar

2.1.3 Temperature averaging

- Following the setting in the MANEVU_2002_mvref.txt files, the following procedures were used by SMOKE for temporal and spatial temperature averaging in the calculation of emission factors:
  - Spatial averaging: temperatures were averaged over all counties that share a common reference county (i.e. Mobile6 input file)
  - Temporal averaging for May – September emissions processing: no temporal averaging was used, i.e. day-specific temperatures were used to calculate emission factors for each day.
  - Temporal averaging for non-summer-months emissions processing: Temporal averaging over the duration of the episode (i.e. the entire month, see introduction) was used, i.e. monthly average temperatures were used to calculate the emission factors.

2.2 CENRAP

2.2.1 Mobile6 input files

- Mobile6 input files for the CENRAP region for January and July were contained in the files central_M6_{MMM}.zip, north_M6_{MMM}.zip, south_M6_{MMM}.zip, west_M6_{MMM}.zip where {MMM} is either jan or jul. July input files were used for April – September processing, while January input files were used for the remaining months
- All files were downloaded from the CENRAP ftp site in March 2006.

2.2.2 SMOKE/Mobile6 auxiliary files

- SMOKE/Mobile6 auxiliary files were contained in the files central_M6_RD.zip, north_M6_RD.zip, south_M6_RD.zip, and west_M6_RD.zip. The SMOKE MCREF, MVREF, and MCODES files were contained in the file MOBILESMOKE_Inputs.zip. The MCREF and MVREF files were combined for the different regions (“central”, “east”, “west”, “north”)
- All files were downloaded from the CENRAP ftp site in March 2006.

2.2.3 Temperature averaging

- The following procedures were used by SMOKE for temporal and spatial temperature averaging in the calculation of emission factors according to the setting in the mvref files:
  - Spatial averaging: no spatial averaging of temperatures, i.e. the temperatures for the reference county is used to calculate emission factors for all counties that share this reference county (i.e. Mobile6 input file)
  - Temporal averaging: Temporal averaging over the duration of the episode (i.e. the entire month, see introduction) was used, i.e. monthly average temperatures were used to calculate the emission factors.

2.3 VISTAS

2.3.1 Mobile6 input files

- Month-specific Mobile6 input files were obtained from the Alpine Geophysics ftp site in July 2006. They reflect version BaseG of the VISTAS inventory.

2.3.2 SMOKE/Mobile6 auxiliary files

- SMOKE/Mobile6 auxiliary files utilized were obtained from the Alpine Geophysics ftp site in July 2006. They reflect version BaseG of the VISTAS inventory.

2.3.3 Temperature averaging

- The following procedures were used by SMOKE for the temporal and spatial temperature averaging in the calculation of emission factors according to the setting in the mvref_baseg.36k.ag.txt file:
    - Spatial averaging: temperatures averaged over all counties that share a common reference county (i.e. Mobile6 input file)
    - Temporal averaging: Temporal averaging over the duration of the episode (i.e. the entire month, see introduction) was used, i.e. monthly average temperatures were used to calculate the emission factors.

2.4 MRPO

2.4.1 Mobile6 input files

- Month-specific Mobile6 input files for SMOKE modeling were generated by Alpine Geophysics through a contract from MARAMA. They are based on version BaseK of the MRPO inventory. The files were downloaded from the MARAMA ftp site **ftp.marama.org** (username mane-vu, password exchange) in May 2006.

2.4.2 SMOKE/Mobile6 auxiliary files

- SMOKE/Mobile6 auxiliary files for SMOKE modeling were generated by Alpine Geophysics through a contract from MARAMA. They are based on version BaseK of the MRPO inventory. The files were downloaded from the MARAMA ftp site **ftp.marama.org** (username mane-vu, password exchange) in May 2006.

2.4.3 Temperature averaging

- The following procedures were used by SMOKE for the temporal and spatial temperature averaging in the calculation of emission factors according to the setting in the mvreg_mrpo_basek.txt file:
    - Spatial averaging: temperatures averaged over all counties that share a common reference county (i.e. Mobile6 input file)

- Temporal averaging: Temporal averaging over the duration of the episode (i.e. the entire month, see introduction) was used, i.e. monthly average temperatures were used to calculate the emission factors.

3. Biogenic Emission Processing

Hourly gridded biogenic emissions for the 12 km and 36 km modeling domains were calculated by BEIS3.12 through SMOKE, using MCIP-processed MM5 fields for temperature ("TA", layer-1 temperature), solar radiation ("RGRND"), surface pressure ("PRES"), and precipitation ("RN" and "RC"). A 'seasonal switch' file was generated by the SMOKE utility metscan to determine whether winter or summer emission factors should be used for any given grid cell on any given day. Winter emission factors are used from January 1st through the date of the last frost and again from the data of the first frost in fall through December 31st. Summer emission factors are used for the time period in between. This calculation is performed separately for each grid cell.

4. Temporal Allocation

4.1 MANE-VU

4.1.1 Area and nonroad sources

- Generated as part of the MANE-VU version 1 inventory
- amptpro.m3.us+can.manevu.030205.txt
- amptref.m3.manevu.012405.txt
- downloaded from **ftp.marama.org** (username mane-vu, password exchange) in January 2005

4.1.2 Mobile sources

- MANEVU_2002_mtpro_02022006_addCT.txt
- MANEVU_2002_mtref_02022006_addCT.txt
- prepared by PECHAN and NESCAUM and downloaded from http://bronze.nescaum.org/Private/junghun/MANE-VU/onroad_ver3_update/MANEVU_V3_update.tar

4.1.3 Point Sources

- Based on the same files as for the MANE-VU area and nonroad temporal files listed above, but added the CEM-based 2002 state-specific temporal profiles and cross-references for EGU sources for the MANE-VU states that were generated by VISTAS for their BaseD modeling and obtained in February 2005.
- No CEM-based hour-specific EGU emissions were utilized

4.2 CENRAP

The following temporal profiles and cross-reference files were used:

- Area and nonroad sources:
    - amptpro.m3.us+can.cenrap.010605_incl_nrd.txt
    - amptref.m3.cenrap.010605_add_nh3_and_nrd.txt
- Mobile sources:
    - mtpro.cenrap.v3.txt
    - mtref.cenrap.v3.txt
- Point sources:
    - ptpro.{QQ}.cenrap_egus_cem.00-03avg.121205.txt where {QQ} is Q1 for January/February/March, Q2 for April/May/June, etc.
    - ptref.{QQ}.cenrap_egus_cem.00-03avg.121205.txt where {QQ} is Q1 for January/February/March, Q2 for April/May/June, etc.
- All files were downloaded from the CENRAP ftp site in March 2006.

4.3 VISTAS

The following month-specific temporal profiles and cross-reference files were used:

- Area and nonroad sources:
    - atpro_vistas_basef_15jul05.txt
    - atref_vistas_basef_15jul05.txt
- Mobile sources:
    - mtpro_vistas_basef_04jul05.txt
    - mtref_us_can_vistas_basef_04jul05.txt
- Point sources:
    - ptpro_typ_{MMM}_vistasg_28jun2006.txt where {MMM} is jan, feb, mar, etc.
    - ptref_typ_vistas_baseg_28jun2006.txt
- These files were obtained from the Alpine Geophysics ftp site. They reflect version BaseG of the VISTAS inventory for the point source allocation files and version BaseF for the area, nonroad, and mobile source allocation files. These files were downloaded between February and July, 2006.

4.4 MRPO

The following month-specific temporal profiles and cross-reference files were used for all source categories:

- amptpro_typ_us_can_{MMM}_vistas_27nov04.txt where {MMM} is jan, feb, mar, etc.
- amptref_2002_us_can_vistas_17dec04.txt
- These files were obtained from VISTAS in January 2005 and reflect their BaseD modeling. No updated temporal profiles or cross-reference files were developed for use with the MRPO BaseK inventory.

4.5 Canada

For Canada, the SMOKE2.1 default temporal profiles and cross-reference files (amptpro.m3.us+can.txt and amptref.m3.us+can.txt) were utilized.

5. Speciation

The same speciation profiles (gspro.cmaq.cb4p25.txt) and cross-references (gsref.cmaq.cb4p25.txt) were utilized for all regions and all source categories. Different versions of these files were obtained (SMOKE2.1 default, EPA-CAIR modeling, VISTAS, CENRAP and MANE-VU) and compared. After comparing the creation dates and header lines of these files, it was determined that the EPA-CAIR and MANE-VU files had the most recent updates, and consequently the final speciation profile and cross-reference files used for all regions and source categories was based on the EPA-CAIR files with the addition of MANE-VU specific updates.

6. Spatial Allocation

6.1 U.S.

The spatial surrogates for the 12km domain were extracted from the national grid 12km U.S. gridding surrogates posted at EPA's website at
http://www.epa.gov/ttn/chief/emch/spatial/newsurrogate.html
The gridding cross-references were also obtained from this website, but for the processing of MANE-VU area source emissions, MANE-VU specific cross-reference entries posted on the MARAMA ftp site were added.

6.2 Canada

The spatial surrogates for Canadian emissions for the 12km domain were extracted from the national grid 12km Canadian gridding surrogates posted at EPA's website at
http://www.epa.gov/ttn/chief/emch/spatial/newsurrogate.html
The gridding cross-references were also obtained from this website.

Reference:

Pechan: (2006) Technical Support document for 2002 MANE-VU SIP Modeling inventories, version 3. Prepared by E. H. Pechan & Associates, Inc. 3622 Lyckan Parkway, Suite 2005, Durham, NC 27707.

APPENDIX H-4

$PM_{2.5}$ MODELING USING THE SMOKE/CMAQ SYSTEM OVER THE OZONE TRANSPORT REGION (OTR)

Bureau of Air Quality
Department of Environmental Protection

This page blank for copying purposes.

TSD-2c

$PM_{2.5}$ modeling using the SMOKE/CMAQ system over the Ozone Transport Region (OTR)

Bureau of Air Quality Analysis and Research

Division of Air Resources

New York State Department of Environmental Conservation

Albany, NY 12233

February 1, 2006

Air Quality Modeling Domain

The modeling domain utilized in this application represented a sub-set of the inter-RPO's continental modeling domain that covered the entire 48-state region with emphasis on the Ozone Transport Region. The OTC modeling domain at 12km horizontal mesh is displayed in Figure 1 is part of the 36km continental domain that is designed to provide boundary conditions (BCs). The particulars of the two modeling domains are:

The 36km domain covered the continental US by a 149 by 129 mesh in the east-west and north-south directions, respectively. The domain is based on Lambert Conformal Projection with the center at (97ºW 40ºN) and parallels at 33ºN and 45ºN. As evident from Figure 1, the 12km domain utilized in this analysis covers most areas of the eastern US and has 172 by 172 mesh in the horizontal. Both domains utilize 22 layers in the vertical extending to about 16km with 16 layers placed within the lower 3km.

Photochemical Modeling -- CMAQ

The CMAQ (version 4.5.1) with CB4 chemistry, aerosol module for $PM_{2.5}$ and RADM cloud scheme was utilized in this study. Photochemical modeling was performed with the CCTM software that is part of the CMAQ modeling package. Version 4.5.1 of this modeling software was obtained from the CMAS modeling center at http://www.cmascenter.org. The following module options were used in compiling the CCTM executable:

- Horizontal advection: yamo
- Vertical advection: yamo
- Horizontal diffusion: multiscale
- Vertical diffusion: eddy
- Plume-in-Grid: non operational
- Gas phase chemical mechanism: CB-4
- Chemical solver: EBI
- Aerosol module: aero3
- Process analysis: non operational

The following computational choices were made during compilation:

- Compiler version: PGI 6.0
- Fortran compiler flags:-Mfixed -Mextend -Bstatic -O2 -module ${MODLOC} -I.
- C compiler flags: -v -O2 -I${MPICH}/include
- IOAPI library: version 3.0
- NETCDF library: version 3.6.0
- Parallel processing library version: mpich 1.2.6
- Static compilation on 32-bit system

The following choices were made for running the executable:

- Number of processors: 8
- Domain decomposition for parallel processing: 4 columns, 2 rows
- Number of species written to the layer-1 hourly-average concentration output (ACONC) file: 39 (O3, NO, CO, NO2, HNO3, N2O5, HONO, PNA, PAN, NTR, NH3, SO2, FORM, ALD2, PAR, OLE, ETH, TOL, XYL, ISOP, ASO4I, ASO4J, ANO3I, ANO3J, ANH4I, ANH4J, AORGAI, AORGAJ, AORGPAI, AORGPAJ, AORGBI, AORGBJ, AECI, AECJ, A25I, A25J, ACORS, ASEAS, ASOIL)
- Each daily simulation was performed for 24 hours starting at 05:00 GMT (00:00 EST)

The following postprocessing steps were performed using utility tools from the "ioapi" software package obtained from http://www.baronams.com/products/ioapi/AA.html#tools:

- Extract and combine the following species for each hour for the first 16 model layers from the full 3-D instantaneous concentration output file: O3, CO, NO, NO2, NOY_1 (=_NO + NO2 + PAN + HNO3_), NOY_2 (=_NO + NO2 + PAN + HNO3 + HONO + N2O5 + NO3 + PNA + NTR_), HOX (=_OH + HO2_), VOC (=_2*ALD2 + 2*ETH + FORM + 5*ISOP + 2*OLE + PAR + 7*TOL + 8*XYL_), ISOP, PM2.5 (=_ASO4I + ASO4J + ANO3I + ANO3J + ANH4I + ANH4J + AORGAI + AORGAJ + 1.167*AORGPAI + 1.167*AORGPAJ + AORGBI + AORGBJ + AECI + AECJ + A25I + A25J_), PM_SULF (=_ASO4I + ASO4J_), PM_NITR (=_ANO3I + ANO3J_), PM_AMM (=_ANH4I + ANH4J_), PM_ORG_SA (=_AORGAI + AORGAJ_), PM_ORG_PA (=_1.167*AORGPAI + 1.167*AORGPAJ_), PM_ORG_SB(=_AORGBI + AORGBJ_), PM_ORG_TOT (=_AORGAI + AORGAJ + 1.167*AORGPAI + 1.167*AORGPAJ + AORGBI + AORGBJ_), PM_EC (=_AECI + AECJ_), PM_OTH (=_A25I + A25J_), PM_COARS (=_ACORS + ASEAS + ASOIL_), SO2, HNO3, NH3, H2O2
- Extract all species for all model layers for the last hour of each daily instantaneous concentration output file to enable "hot" restarts of modeling simulations
- Create daily files of hourly running-average 8-hr ozone concentrations with time stamps assigned to the first hour of the averaging interval

The following files are archived on LTO2 computer tapes (each tape holds approximately 200 Gb of data) for each day:

- Aerosol/visibility file
- Layer-1 hourly-average concentration output file (contains 39 species)
- Dry deposition file
- Wet deposition file
- Extracted 16-layer species file
- Restart file (last hour of full 3-D instantaneous concentration file)
- Hourly 8-hr concentration file

Photolysis Rates

One of the inputs to CMAQ is the photolysis rates. In this study, photolysis rate lookup tables were generated for each day of 2002 with the JPROC software that is part of the CMAQ modeling package. This software was obtained from the CMAS modeling center at http://www.cmascenter.org. Rather than using climatological ozone column data, daily ozone column measurements from the NASA Earthprobe TOMS instrument were downloaded from ftp://toms.gsfc.nasa.gov/pub/eptoms/data/ozone/Y2002/ and used as input to the JPROC processor. It should be noted that TOMS data were missing for the time period from August 3 – 11, 2002. The missing period was filled as follows-- TOMS data file for August 2 was used as JPROC input for August 3rd through August 7th, and the TOMS data file for August 12th was used as JPROC input for August 8th through August 11th.

Boundary Conditions (BCs)

The boundary conditions for the 12km grid were extracted from the 36km CMAQ simulation. The 36km simulation utilized boundary conditions that were based on a one-way nest approach to GEOS-CHEM global model outputs (Moon and Byun 2004, Baker 2005). As stated above, the intent of the 36km CMAQ simulation was to provide the BCs for the 12km model that would be more reflective of the emissions and meteorology rather than to use either clean or arbitrary pollutant fields. Also, in this study the CMAQ simulations utilized a 15-day ramp-up period, thereby minimizing the propagation of the boundary fields into the areas of concern. A report on the setup and application of the 36km CMAQ and the extraction of the BCs is available from NYSDEC.

Meteorological data

The meteorological data for this study was based on MM5 modeling (see Meteorological Modeling, 2007). The MM5 fields are then processed by MCIP version 3.0, a utility available as part of the CCTM software from CMAS Modeling Center (see http://www.cmascenter.org) to provide CMAQ model-ready inputs.

Emissions

The emissions data for 2002 were generated by individual states within the OTR and were assembled and processed through the Mid Atlantic Northeast Visibility Union (MANE-VU), a Regional Planning Organization (RPO). These emissions were then processed by NYSDEC using SMOKE processor to provide CMAQ compatible inputs (Anthro-Emissions 2006). The 2002 emissions for the non-OTR areas within the modeling domain were obtained from the corresponding RPOs and were processed using SMOKE, in a manner similar to that of the OTR.emissions. Details of this processing are outlined in the report (Pechan 2007), and the hourly biogenic emissions (Bio-Emissions, 2006)

CMAQ simulations

CMAQ simulations were performed using the one-way nesting approach in which we perform the continental CMAQ simulation at 36km grid spacing. For this simulation we utilized clean initial conditions with boundary conditions extracted from the simulation of GEOS-CHEM global chemical model. The interface program used in this application was developed by University of Huston (Moon and Byun 2004), which was applied to obtain hourly 36km boundary concentrations from GEOS-CHEM outputs. The CMAQ 36km simulation was initiated from December 15, 2001 with the first 15 days as spin up period and terminated on December 31, 2002. The simulation utilized the 2002 emissions data available from the RPOs and 2002 MM5 meteorological fields developed by the University of Maryland (TSD-1a). The hourly boundary fields for the 12km CMAQ domain were obtained by application of BCON program to the 3-D concentration fields generated by the 36km CMAQ simulation.

The 12km simulations for both base and future year were assigned the boundary conditions based on the 36km CMAQ simulation and clean initial conditions. The annual simulation was parsed out between different member states or their contractors of the OTR, so as to expedite the process of completing the simulation in a limited time. The approach used is as follows: The annual simulation was parsed out into five parts and each modeling center identified below initiated and completed the simulation, extracted the outputs which were then combined to provide the annual simulation. There was considerable exchange of information in the setup and execution of the modeling system between the centers using benchmark runs to ensure consistency and uniformity between the centers. The process was followed both for the base year 2002 and for the future year 2009. Details on CMAQ setup and run scripts are available from NYSDEC.

Modeling Center	Simulation period	Analysis period
MDE/UMD	Dec 15, 2001 to Feb 28, 2002	Jan 01, 2002 to Feb 28, 2002
NJDEP/Rutgers	Feb 15, 2002 to May 14, 2002	Mar 01, 2002 to May 14, 2002
NYSDEC	May 01, 2002 to Sep 30, 2002	May 15, 2002 to Sep 30, 2002
VA DEQ	Sep 15, 2002 to Oct 30, 2002	Oct 01, 2002 to Oct 30, 2002
NESCAUM	Oct 15, 2002 to Dec 31, 2002	Nov 01, 2002 to Dec 31, 2002

References

Baker, K.: (2005) http://www.ladco.org/tech/photo/present/ozone.pdf

Moon, N. and D. Byun: (2004) A simple user’s guide for “geos2cmaq”code: Linking CMAQ with GEOS-CHEM. Version 1.0. Institute for Multidimensional Air quality Studies (IMAQS), University of Houston, Houston TX.

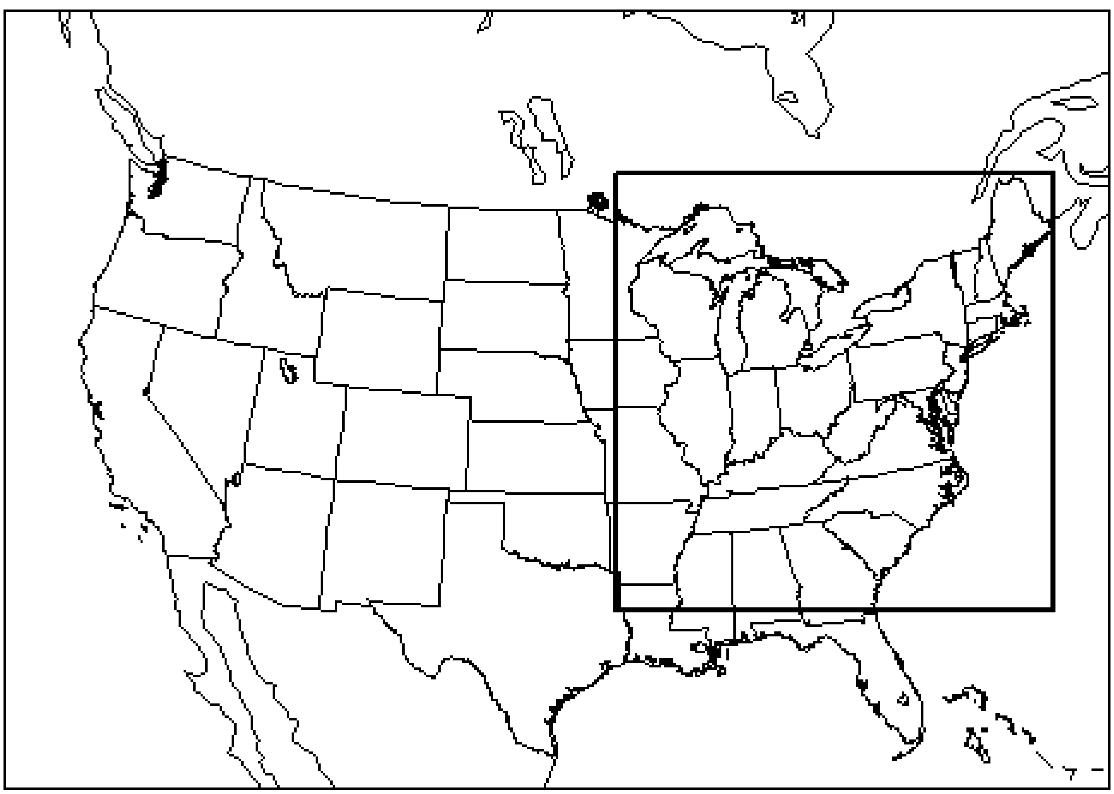
Meteorological Modeling: (2007) Meteorological Modeling of 2002 using Penn State/NCAR 5th Generation Mesoscale Model (MM5). TSD-1a

Pechan: (2006) Technical Support document for 2002 MANE-VU SIP Modeling inventories, version 3. Prepared by E. H. Pechan & Associates, Inc. 3622 Lyckan Parkway, Suite 2005, Durham, NC 27707.

Bio-Emissions: (2006) Processing of Biogenic Emissions for OTC/MANE-VU Modeling. TSD-1b

Anthro-Emissions: (2006) Emission Processing for the Revised 2002 OTC Regional and Urban 12 km Base Case Simulations. TSD-1c

Figure 1 Display of 36- and 12km air quality modeling domains.



APPENDIX H-5

PROCESSING OF 2002 BIOGENIC EMISSIONS FOR OTC / MANE-VU REGIONAL AND URBAN MODELING

Bureau of Air Quality
Department of Environmental Protection

This page blank for copying purposes.

TSD-2a

Processing of 2002 Biogenic Emissions for OTC / MANE-VU Regional and Urban Modeling

Bureau of Air Quality Analysis and Research

Division of Air Resources

New York State Department of Environmental Conservation

Albany, NY 12233

September 19, 2006

Biogenic emissions for the time period from January 1, 2002 – December 31, 2002 were calculated by NYSDEC using the Biogenic Emissions Inventory System (BEIS) version 3.12 integrated within SMOKE2.1. General information about BEIS is available at http://www.epa.gov/AMD/biogen.html while documentation about biogenic emissions processing within SMOKE2.1 is available at http://cf.unc.edu/cep/empd/products/smoke/version2.1/html/ch06s10.html and http://cf.unc.edu/cep/empd/products/smoke/version2.1/html/ch06s17.html . Note that the SMOKE documentation refers to BEIS3.09 and has not yet been updated for BEIS3.12. This affects the number of species modeled as well as the use of different speciation profiles. However, the general processing approach has not changed from BEIS3.09 to BEIS3.12. In short, this processing approach is as follows and was utilized by NYSDEC for its biogenic emission processing for 8-hr ozone and $PM_{2.5}$ modeling:

1. **Normbeis3** reads gridded land use data and emissions factors and produces gridded normalized biogenic emissions for 34 species/compounds. The gridded land use includes 230 different land use types. Both summer and winter emissions factors for each species/compound are provided for each of the 230 land use types. On output, **Normbeis3** generates a file `B3GRD` which contains gridded summer and winter emission fluxes for the modeling domain that are normalized to 30 °C and a photosynthetic active radiation (PAR) of 1000 $\mu mol/m^2s$. In addition, gridded summer and winter leaf area indices (LAI) are also written to `B3GRD`.

2. **Tmpbeis3** reads the gridded, normalized emissions file `B3GRD` and meteorological data from the MCIP-processed MM5 meteorological fields generated by the University of Maryland for MANE-VU/OTC modeling. Specifically, the following MM5/MCIP meteorological variables are used by **Tmpbeis3** to compute hour-specific, gridded biogenic emissions from the normalized emission fluxed contained in `B3GRD:` layer-1 air temperature ("TA"), layer-1 pressure ("PRES"), total incoming solar radiation at the surface ("RGRND"), and convective ("RC") and non-convective ("RN") rainfall. Additionally, the emissions for the 34 species/compounds modeled by BEIS3.12 are converted to CO, NO, and the CB-IV

VOC species utilized in CMAQ via the use of the BEIS3.12-CB-IV speciation profile. In adition, an optional seasonal switch file, `BIOSEASON`, was utilized to decide whether to use summer or winter emissions factors for any given grid cell on any given day. This file was generated by the SMOKE2.1 utility **Metscan** based on MM5 layer-1 air temperatures to determine the date of the last spring frost and first fall frost at each grid cell. Summer emission factors are used by **Tmpbeis3** for the time period between the last spring frost and first fall frost at any given grid cell, and winter emission factors are used for the remaining time period. Documentation for the **Metscan** utility is available at http://cf.unc.edu/cep/empd/products/smoke/version2.1/html/ch05s07.html . An animated GIF file showing the `BIOSEASON` file used by NYSDEC can be found at ftp://ftp.dec.state.ny.us/dar/air_research/chogrefe/biog_reports/b3season_movie.gif

3. For reporting purposes, the hourly, speciated, gridded emissions were aggregated to the county level for each day. For any given grid cell, emissions are distributed among the counties intersecting this grid cell in proportion to the area of each of these counties within the grid cell. The area gridding surrogates needed for this aggregration are based on a file obtained from EPA via http://www.epa.gov/ttn/chief/emch/spatial/new/bgpro.12km_041604.us.gz followed by windowing for the MANE-VU/OTC modeling domain.

Table 1 County and State totals of estimated biogenic emissions (tpy)

State	FIPS	County	NO [TPY]	CO [TPY]	VOC [TPY]
Connecticut	009001	Fairfield	52	894	7150
	009003	Hartford	88	915	8537
	009005	Litchfield	98	1261	12221
	009007	Middlesex	54	615	5587
	009009	New Haven	80	876	7544
	009011	New London	74	906	8960
	009013	Tolland	55	651	5999
	009015	Windham	60	772	8019
Connecticut		**TOTAL**	560	6889	64017
Deleware	010001	Kent	308	1354	15912
	010003	New Castle	143	875	8834
	010005	Sussex	539	2045	21595
Deleware		**TOTAL**	990	4274	46342
DC	011001	Washington	30	150	1726
DC		**TOTAL**	30	150	1726
Maine	023001	Androscoggin	35	885	8204
	023003	Aroostook	741	15531	140877
	023005	Cumberland	49	1298	11528
	023007	Franklin	72	3269	32111
	023009	Hancock	66	2950	27090
	023011	Kennebec	73	1425	12849
	023013	Knox	30	689	6680
	023015	Lincoln	32	849	8072
	023017	Oxford	79	3224	34189
	023019	Penobscot	211	7249	63128
	023021	Piscataquis	146	8638	80748
	023023	Sagadahoc	37	526	4504
	023025	Somerset	173	8413	77850
	023027	Waldo	57	1833	18125
	023029	Washington	144	6459	58678
	023031	York	73	1698	15571
Maine		**TOTAL**	2018	64936	600203
Maryland	024001	Allegany	63	661	8664
	024003	Anne Arundel	79	945	12786
	024005	Baltimore	166	847	8102
	024009	Calvert	59	798	10048
	024011	Caroline	202	648	7907

	024013	Carroll	189	822	7853
	024015	Cecil	86	654	10093
	024017	Charles	78	1079	15042
	024019	Dorchester	134	829	10337
	024021	Frederick	204	1123	10964
	024023	Garrett	102	930	11391
	024025	Harford	141	911	9053
	024027	Howard	75	562	4460
	024029	Kent	177	498	4761
	024031	Montgomery	134	813	6786
	024033	Prince Georges	87	732	10214
	024035	Queen Annes	222	684	7146
	024037	St Marys	99	886	10793
	024039	Somerset	58	498	5796
	024041	Talbot	131	495	5225
	024043	Washington	112	781	7538
	024045	Wicomico	124	796	10304
	024047	Worcester	158	1121	13079
	024510	Baltimore	54	235	1762
Maryland		**TOTAL**	2934	18350	210104
Massachusetts	025001	Barnstable	261	668	5905
	025003	Berkshire	73	1182	11029
	025005	Bristol	107	753	7142
	025007	Dukes	115	252	1728
	025009	Essex	55	794	7128
	025011	Franklin	61	1031	9424
	025013	Hampden	51	904	9201
	025015	Hampshire	61	820	7056
	025017	Middlesex	68	1085	11630
	025019	Nantucket	56	159	1362
	025021	Norfolk	49	615	5513
	025023	Plymouth	170	1197	11876
	025025	Suffolk	26	177	1351
	025027	Worcester	103	1955	23612
Massachusetts		**TOTAL**	1257	11594	113957
New Hampshire	033001	Belknap	25	693	6915
	033003	Carroll	40	1512	14981
	033005	Cheshire	49	1019	10099
	033007	Coos	72	3239	33668
	033009	Grafton	91	2442	23151
	033011	Hillsborough	48	1337	14503
	033013	Merrimack	48	1314	13566
	033015	Rockingham	39	1120	10080
	033017	Strafford	25	686	6617
	033019	Sullivan	45	943	8314
New Hampshire		**TOTAL**	482	14306	141894

New Jersey	034001	Atlantic	135	1225	18890
	034003	Bergen	37	239	2455
	034005	Burlington	151	1827	25255
	034007	Camden	68	491	7751
	034009	Cape May	90	566	7763
	034011	Cumberland	122	773	10699
	034013	Essex	57	199	1831
	034015	Gloucester	119	556	8444
	034017	Hudson	26	125	701
	034019	Hunterdon	81	706	5743
	034021	Mercer	85	475	4889
	034023	Middlesex	98	456	5267
	034025	Monmouth	125	1152	15423
	034027	Morris	63	604	7288
	034029	Ocean	128	1871	27063
	034031	Passaic	41	339	3841
	034033	Salem	123	535	8304
	034035	Somerset	49	518	5548
	034037	Sussex	67	718	7768
	034039	Union	21	168	2191
	034041	Warren	125	517	4505
New Jersey		**TOTAL**	1813	14058	181618
New York	036001	Albany	59	730	6253
	036003	Allegany	129	1218	9526
	036005	Bronx	25	100	657
	036007	Broome	107	879	7861
	036009	Cattaraugus	148	1654	13540
	036011	Cayuga	227	986	7928
	036013	Chautauqua	202	1260	8144
	036015	Chemung	88	521	3911
	036017	Chenango	149	1120	7833
	036019	Clinton	138	1631	13341
	036021	Columbia	96	896	8484
	036023	Cortland	101	616	4280
	036025	Delaware	133	1672	13435
	036027	Dutchess	90	1096	10288
	036029	Erie	165	1127	6898
	036031	Essex	94	2547	20888
	036033	Franklin	228	2337	17197
	036035	Fulton	90	764	5275
	036037	Genesee	201	645	3993
	036039	Greene	47	886	8182
	036041	Hamilton	78	2092	16056
	036043	Herkimer	175	1783	12846
	036045	Jefferson	251	1754	12503
	036047	Kings	15	60	309

	036049	Lewis	154	1693	12116
	036051	Livingston	222	888	6048
	036053	Madison	149	1049	7528
	036055	Monroe	223	990	6237
	036057	Montgomery	106	579	4715
	036059	Nassau	81	408	2859
	036061	New York	16	76	473
	036063	Niagara	335	940	5182
	036065	Oneida	214	1515	10021
	036067	Onondaga	171	929	6259
	036069	Ontario	178	767	6024
	036071	Orange	110	1065	13024
	036073	Orleans	195	635	3314
	036075	Oswego	119	1277	7911
	036077	Otsego	157	1190	7958
	036079	Putnam	32	473	5243
	036081	Queens	20	105	543
	036083	Rensselaer	96	894	7316
	036085	Richmond	47	173	1292
	036087	Rockland	26	300	4006
	036089	St. Lawrence	376	3876	28960
	036091	Saratoga	76	1125	9010
	036093	Schenectady	39	377	3032
	036095	Schoharie	95	737	5496
	036097	Schuyler	87	438	3193
	036099	Seneca	127	438	3305
	036101	Steuben	267	1475	12085
	036103	Suffolk	368	1328	12886
	036105	Sullivan	76	1325	12538
	036107	Tioga	102	730	5400
	036109	Tompkins	96	576	4128
	036111	Ulster	82	1493	15714
	036113	Warren	46	1396	11568
	036115	Washington	183	1109	8355
	036117	Wayne	270	920	5940
	036119	Westchester	35	549	5347
	036121	Wyoming	194	720	3813
	036123	Yates	107	507	4017
New York		**TOTAL**	8313	63436	492483
Pennsylvania	042001	Adams	186	892	8926
	042003	Allegheny	182	948	6727
	042005	Armstrong	108	940	9955
	042007	Beaver	69	600	4895
	042009	Bedford	128	1249	14127
	042011	Berks	280	1377	14146
	042013	Blair	91	729	7579
	042015	Bradford	224	1265	9423

042017	Bucks	144	954	8399
042019	Butler	149	1032	8602
042021	Cambria	128	805	6545
042023	Cameron	25	627	7563
042025	Carbon	53	585	8121
042027	Centre	158	1344	16886
042029	Chester	264	1176	10474
042031	Clarion	85	848	10743
042033	Clearfield	149	1368	13267
042035	Clinton	71	1230	18191
042037	Columbia	106	802	9080
042039	Crawford	204	1297	10839
042041	Cumberland	193	816	9505
042043	Dauphin	116	799	8502
042045	Delaware	35	410	3250
042047	Elk	49	949	8921
042049	Erie	199	1107	8273
042051	Fayette	156	1087	9277
042053	Forest	26	577	7122
042055	Franklin	271	1057	10296
042057	Fulton	93	744	9341
042059	Greene	91	830	6966
042061	Huntingdon	135	1093	12606
042063	Indiana	144	1078	9156
042065	Jefferson	101	865	7362
042067	Juniata	79	588	8263
042069	Lackawanna	58	586	5569
042071	Lancaster	464	1299	9565
042073	Lawrence	114	503	3755
042075	Lebanon	155	623	5827
042077	Lehigh	149	594	6040
042079	Luzerne	75	1013	13215
042081	Lycoming	152	1457	16633
042083	Mc Kean	57	1044	7113
042085	Mercer	175	865	7114
042087	Mifflin	107	620	7508
042089	Monroe	75	773	8856
042091	Montgomery	106	812	6736
042093	Montour	85	321	3306
042095	Northampton	144	506	4416
042097	Northumberland	92	570	6340
042099	Perry	113	804	10216
042101	Philadelphia	29	194	1420
042103	Pike	37	757	9946
042105	Potter	89	1129	9027
042107	Schuylkill	123	1050	15001
042109	Snyder	88	538	6373
042111	Somerset	221	1251	11228

	042113	Sullivan	45	684	5112
	042115	Susquehanna	126	978	6448
	042117	Tioga	176	1313	10942
	042119	Union	71	541	6435
	042121	Venango	72	855	9086
	042123	Warren	76	1031	7352
	042125	Washington	166	1068	7429
	042127	Wayne	89	862	5954
	042129	Westmoreland	199	1297	10589
	042131	Wyoming	60	551	4634
	042133	York	366	1393	12758
Pennsylvania		**TOTAL**	8645	59945	585271
Rhode Island	044001	Bristol	40	90	441
	044003	Kent	41	328	3471
	044005	Newport	37	183	1646
	044007	Providence	39	591	6901
	044009	Washington	54	572	6775
Rhode Island		**TOTAL**	211	1764	19233
Vermont	050001	Addison	186	922	6274
	050003	Bennington	43	896	7349
	050005	Caledonia	58	1149	10239
	050007	Chittenden	74	606	3633
	050009	Essex	61	1315	11795
	050011	Franklin	208	971	5927
	050013	Grand Isle	50	490	3506
	050015	Lamoille	36	727	5627
	050017	Orange	57	1182	10120
	050019	Orleans	120	1570	12842
	050021	Rutland	102	1257	9867
	050023	Washington	47	1099	9502
	050025	Windham	42	1232	10898
	050027	Windsor	57	1330	10796
Vermont		**TOTAL**	1142	14745	118376
Virginia	051001	Accomack	187	959	9472
	051003	Albemarle	140	1246	12533
	051005	Alleghany	35	522	7369
	051007	Amelia	70	915	10717
	051009	Amherst	80	905	10823
	051011	Appomattox	76	830	10447
	051013	Arlington	17	64	531
	051015	Augusta	135	1049	13291
	051017	Bath	46	771	11636
	051019	Bedford	189	1279	13052
	051021	Bland	41	515	7097
	051023	Botetourt	74	780	10211

051025	Brunswick	98	1458	18254
051027	Buchanan	32	722	9557
051029	Buckingham	76	1287	18830
051031	Campbell	112	1078	12933
051033	Caroline	73	1173	16020
051035	Carroll	132	634	6885
051036	Charles City	93	415	4711
051037	Charlotte	84	1219	14277
051041	Chesterfield	69	802	10686
051043	Clarke	56	369	4009
051045	Craig	39	538	7314
051047	Culpeper	105	894	10720
051049	Cumberland	56	814	10677
051051	Dickenson	20	550	6910
051053	Dinwiddie	82	1207	16511
051057	Essex	58	671	7403
051059	Fairfax	111	533	5538
051061	Fauquier	150	1166	14084
051063	Floyd	47	593	6493
051065	Fluvanna	54	775	10756
051067	Franklin	119	1297	15933
051069	Frederick	64	588	8798
051071	Giles	38	508	4918
051073	Gloucester	32	510	5945
051075	Goochland	47	670	10392
051077	Grayson	60	627	8260
051079	Greene	57	434	5727
051081	Greensville	63	735	9009
051083	Halifax	201	1852	22730
051085	Hanover	91	950	12493
051087	Henri	81	427	5468
051089	Henry	59	805	9772
051091	Highland	44	608	8579
051093	Isle Of Wight	178	813	8049
051095	James City	41	314	3989
051097	King And Queen	77	673	7615
051099	King George	62	540	6111
051101	King William	102	712	7846
051103	Lancaster	33	311	3669
051105	Lee	97	680	7221
051107	Loudoun	137	942	8999
051109	Louisa	78	1142	16780
051111	Lunenberg	88	1108	13611
051113	Madison	70	598	7305
051115	Mathews	27	367	4025
051117	Mecklenburg	145	1478	18507
051119	Middlesex	42	480	5561
051121	Montgomery	70	501	5366

051125	Nelson	67	979	12465
051127	New Kent	35	600	8240
051131	Northampton	90	263	2019
051133	Northumberland	88	778	9298
051135	Nottoway	74	894	10670
051137	Orange	98	759	8265
051139	Page	77	540	6705
051141	Patrick	75	884	10255
051143	Pittsylvania	203	1806	22102
051145	Powhatan	47	675	10194
051147	Prince Edward	69	942	12042
051149	Prince George	73	572	6484
051153	Prince William	38	718	10979
051155	Pulaski	61	450	6510
051157	Rappahannock	61	521	7141
051159	Richmond	63	383	4548
051161	Roanoke	63	427	5278
051163	Rockbridge	101	813	9710
051165	Rockingham	189	1020	12959
051167	Russell	56	703	7975
051169	Scott	95	753	9943
051171	Shenandoah	117	757	10570
051173	Smyth	78	603	7159
051175	Southampton	177	1306	15588
051177	Spotsylvania	46	911	12575
051179	Stafford	27	637	8344
051181	Surry	85	784	10024
051183	Sussex	102	1267	16362
051185	Tazewell	77	639	7477
051187	Warren	44	438	6310
051191	Washington	142	632	6822
051193	Westmoreland	101	777	9357
051195	Wise	35	462	5685
051197	Wythe	109	596	7803
051199	York	35	271	3423
051510	Alexandria	38	145	1065
051515	Bedford	22	101	604
051520	Bristol	37	135	1220
051530	Buena Vista	6	43	381
051540	Charlottesville	18	98	528
051550	Chesapeake	71	666	8477
051560	Clifton Forge	27	61	436
051570	Colonial Heights	35	88	662
051580	Covington	24	114	1605
051590	Danville	55	343	3405
051595	Emporia	19	234	3300
051600	Fairfax	18	96	1518
051610	Falls Church	16	98	1120

	051620	Franklin	66	142	1041
	051630	Fredericksburg	14	250	3012
	051640	Galax	45	94	519
	051650	Hampton	24	127	1112
	051660	Harrisonburg	73	143	746
	051670	Hopewell	26	79	711
	051678	Lexington	8	62	620
	051680	Lynchburg	45	250	2135
	051683	Manassas	17	86	743
	051685	Manassas Park	17	50	268
	051690	Martinsville	19	190	1625
	051700	Newport News	63	231	2187
	051710	Norfolk	42	197	2692
	051720	Norton	13	120	1305
	051730	Petersburg	58	171	1419
	051735	Poquoson	17	122	1351
	051740	Portsmouth	34	285	3215
	051750	Radford	27	76	609
	051760	Richmond	29	239	3517
	051770	Roanoke	33	91	770
	051775	Salem	14	61	568
	051790	Staunton	69	205	1550
	051800	Suffolk	118	964	11269
	051810	Virginia Beach	186	924	8724
	051820	Waynesboro	43	120	895
	051830	Williamsburg	3	38	446
	051840	Winchester	42	117	772
Virginia		**TOTAL**	9267	80615	981848