A STUDY OF PROCESS CONTRIBUTIONS TO $PM_{2.5}$ FORMATION DURING THE 2004 ICARTT PERIOD USING THE ETA-CMAQ FORECAST MODEL OVER THE NORTHEASTERN U.S.

Shaocai Yu[+$*], Rohit Mathur[+], Kenneth Schere[+], Daiwen Kang[+$], Jonathan Pleim[+], Jeffrey Young[+], and Daniel Tong[+$]

Atmospheric Sciences Modeling Division, Air Resources Laboratory,
National Oceanic and Atmospheric Administration, RTP, NC 27711
[+]On assignment to National Exposure Research Laboratory,
U.S. Environmental Protection Agency, Research Triangle Park, NC 27711
[$] On assignment from Science and Technology Corporation,
10 Basil Sawyer Drive, Hampton, VA 23666-1393

1. INTRODUCTION

Atmospheric distributions and loadings of fine particulate matter ($PM_{2.5}$; particles with diameters less than 2.5 μm) can adversely affect human and ecosystem health, and are thus of major concern. Hence, it is important for local air quality agencies to accurately forecast $PM_{2.5}$ concentrations to alert the public of the onset, severity and duration of unhealthy air and to encourage people to help reduce emissions-producing activities. In this study, first, we evaluate the Eta-CMAQ forecast model performance for the chemical components (SO_4^{2-}, NO_3^-, and NH_4^+) of $PM_{2.5}$ with the observational data from aircraft flights during the 2004 International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) field experiments. The model spatial performance for $PM_{2.5}$ chemical constituents (SO_4^{2-}, NO_3^-, NH_4^+, OC and EC) is evaluated with the observational data from the IMPROVE, CASTNET, and STN networks. The spatial and temporal performance of the model for surface $PM_{2.5}$ mass over the eastern U.S during this period is examined through comparison with observations from the U.S. EPA Air Quality System (AQS) network. Second, the contributions of various physical and chemical processes governing the distribution of $PM_{2.5}$ are investigated through detailed analysis of model process budgets using the Integrated Process Rate analysis (IPR) along back trajectories from selected locations.

*** *Corresponding author address:*** **Shaocai Yu,** E-243-01, NERL, U.S. EPA, RTP, NC 27711; e-mail: yu.shaocai@epa.gov

2. DESCRIPTION OF THE ETA-CMAQ FORECAST MODEL SUITE AND OBSERVATIONAL DATABASE

The Eta-CMAQ air quality forecasting system (Otte et al., 2005), created by linking the Eta model (Rogers et al., 1996) and the U.S. EPA's CMAQ Modeling System (Byun and Schere, 2006), is deployed over the domain of the eastern U.S. during summer 2004. The domain has horizontal grid cell sizes of 12 km. Twenty-two layers of variable thickness set on a sigma-type coordinate are used to resolve the vertical extent from the surface to 100 hPa. The primary Eta-CMAQ model forecast for next-day's surface-layer $PM_{2.5}$ is based on the current day's 12 UTC Eta cycle. The emissions are projected to 2004 from the 2001 U.S. EPA national emission inventory (Pouliot, 2005). The Carbon Bond chemical mechanism (version 4.2) is used to represent reaction pathways. The developmental CMAQ/Eta model uses the same aerosol module as CMAQ described in Binkowski and Roselle [2003] and its updates are described by Bhave et al. [2004] and Yu et al. [2006]. The aerosol distribution is modeled as a superposition of three lognormal modes that correspond nominally to the ultrafine (diameter (D_p) < 0.1 μm), fine (0.1 < D_p < 2.5 μm), and coarse (D_p > 2.5 μm) particle sizes. The model results for $PM_{2.5}$ concentrations are obtained by summing species concentrations over the first two modes.

Over the eastern U.S., a total of four monitoring networks for $PM_{2.5}$ observations were employed in this evaluation (IMPROVE, STN, CASTNET and AQS). In the IMPROVE network,

two 24-hour samples are collected on quartz filters each week, on Wednesday and Saturday, beginning at midnight local time. The observed $PM_{2.5}$, SO_4^{2-}, NO_3^-, EC and OC data are available at 71 rural sites over the eastern U.S. The STN network (http://www.epa.gov/air/data/aqsdb.html) follows the protocol of the IMPROVE network (i.e. every third day collection) with the exception that most of the sites are located in urban areas. The observed $PM_{2.5}$, SO_4^{2-}, NO_3^-, and NH_4^+ data are available at 178 STN sites over the eastern U.S. The CASTNET (www.epa.gov/castnet/) collected the concentration data at predominately rural sites using filter packs that are exposed for 1-week intervals (i.e., Tuesday to Tuesday). CASTNET species at the 34 sites used in this evaluation include: SO_4^{2-}, NO_3^-, and NH_4^+. In addition, the hourly, near real-time $PM_{2.5}$ data at 309 sites in the eastern U.S. are measured by tapered element oscillating microbalance (TEOM) instruments through the U.S. EPA's Air Quality System (AQS) network. From roughly the July 5 to August 15, 2004, measurements of vertical profiles of $PM_{2.5}$ composition, gas species (CO, NO, NO_2, H_2O_2, HNO_3, SO_2, PAN, toluene etc), and meteorological parameters were carried out by instrumented aircraft (NOAA P-3 and NASA DC-8) deployed as part of the 2004 ICARTT field experiment. The detailed instrumentation and protocols for measurements are described in http://www.al.noaa.gov/ICARTT/FieldOperations/. Following *Mathur et al*. [2005], modeled results were extracted by "flying" the aircraft through the 3-D modeling domain by mapping the locations of the aircraft to the model grid indices (column, row, and layer). Hourly resolved model outputs were linearly interpolated to the corresponding observational times.

3. SUMMARY

For the model performance evaluation, it was found that the domain wide mean values of mean bias and error for $PM_{2.5}$ at the AQS sites are -3.3 and 11.3 $\mu g\ m^{-3}$, respectively. The model underpredicted the observed $PM_{2.5}$ by -21.1%. A comparison of the modeled and observed $PM_{2.5}$ indicates that the model captured a majority (73.3%) of observed $PM_{2.5}$ within a factor of 2. On the basis of the results of a daily time series for the daily 24-hr concentrations, the model had the best performance on 8/16 and 8/13, and significant underpredictions of $PM_{2.5}$ at the beginning (7/16 to 7/26). A close inspection of the modeled and observed data indicates that the majority of the domain was significantly influenced by pollutants from large Alalska forest fires during 7/16 and 7/26 as shown by the aerosol index images from the TOMS satellite observations [http://toms.gsfc.nasa.gov]. The significant under-prediction of $PM_{2.5}$ during this period is mainly attributed to inadequate representation of the transport of pollution associated with biomass burning from outside the domain. Spatially, the model generally slight under predicted observed $PM_{2.5}$ by 21% across space, especially over the northern part of the model domain.

On the basis of results at the IMPROVE, CASTNET and STN sites, the model captured a majority of observed SO_4^{2-}, NH_4^+, $PM_{2.5}$ concentrations within a factor of 2. The model overpredicted the observed mean SO_4^{2-} by 15% at the STN sites, 6% at the CASTNET sites and 11% at the IMPROVE sites. The model also overpredicted the observed mean NH_4^+ by 21% at the STN sites but underpredicted by -6% at the CASTNET sites and 12% at the IMPROVE sites. The model underpredicted the observed mean $PM_{2.5}$ by 15% at the STN sites and 20% at the IMPROVE sites. The model overpredicted the observed SO_2 by 77% at the CASTNET sites. The poor model performance for NO_3^- (correlation <0.4) is related in part to volatility issues associated with NO_3^-, and due to errors in the SO_4^{2-} and total ammonia (NH_3+NH_4^+) simulations (Yu, et al., 2005). The model under predicted most of the observed OC and TC at the IMPROVE sites by more than a factor of 2. This causes an under-prediction of $PM_{2.5}$.

On the basis of results of aircraft observations, the model generally over predicted SO_4^{2-} at high altitudes and unpredicted NH_4^+ compared to P-3 and DC-8 observations although the model results were close to the observations on some days. Mean vertical profiles show overpredictions in SO_4^{2-} above altitudes of ~100 m. These in part could arise from overestimation of H_2O_2 by the CB-IV mechanism, resulting in overprediction of aqueous SO_4^{2-} production. This is also one of the main reasons for the over prediction of SO_4^{2-} at the surface sites of the IMPROVE, STN and CASTNET. The model consistently overpredicted observed H_2O_2 by more than a factor of 2 at the altitude <~1000 m most of time compared to the P-3 observations, leading to the overprediction of SO_4^{2-} because of too much SO_2 cloud oxidation. The modeled SO_2 concentrations are generally higher than the observations at the low altitude (<~300 m) but close to the observations at the high altitude. The model performance for NO_3^- is poor like those at the surface.

To study the contributions of various physical and chemical processes to the distribution of $PM_{2.5}$, we employ a simple approach by using the Hybrid Single Particle Lagrangian Integrated Trajectory model (HYSPLIT) (Draxler, 2003) to determine the back trajectory linking a downwind receptor to upwind source areas and then applying process analysis to the CMAQ grid cell along the trajectory of the air mass transport path. For the back trajectory analysis, the same meteorology applied in the Eta-CMAQ simulation was used to generate input data sets for use in the HYSPLIT back trajectory calculation. Because there is a well-mixed boundary layer during the daytime, the column mean contribution of each process from the layer 1 to layer 14 (~2000 m) was calculated. Figure 1 shows the backward trajectories ending at 11 UTC 17 August, 2004 at the South Allegheny High School (SAHS) and John sites, PA , and ending at 11 UTC 19 August, 2004, at the South Dekalb (SD), McDonough (MD) and Newnan (NN) sites, GA. These sites and times were chosen because their $PM_{2.5}$ concentrations were high (>40 $\mu g\ m^{-3}$) relative to other sites. Table 1 summarizes the total accumulation contributions of each process to $PM_{2.5}$ formation along 24-hr back-trajectories (see Figure 1) on the basis of column means from layer 1 to layer 14 at the five sites. As can be seen, the dominant processes for $PM_{2.5}$ formation and sink vary from the site to site. Horizontal advection process contributes to the loss of $PM_{2.5}$ at most of sites except the SAHS site where it increases $PM_{2.5}$ significantly. In most of cases, vertical diffusion and advection processes make some contributions to the loss of $PM_{2.5}$. The effects of horizontal diffusion on the $PM_{2.5}$ formation are negligible as shown in Table 1. The integrated process budgets along the trajectories at the PA sites indicate large contributions from cloud processing to $PM_{2.5}$, suggesting relatively large contribution of SO_4^{2-} to $PM_{2.5}$ in these air masses. In contrast, the trajectories reaching the sites in GA are characterized by negligible contributions by the cloud process but larger aerosol process, suggesting the large contribution from secondary organic aerosol formation and primary emissions to $PM_{2.5}$ at these locations.

4. REFERENCES

Bhave, P.V., S.J. Roselle, F.S. Binkowski, C.G. Nolte, S.C. Yu, G.L. Gipson, K.L. Schere, CMAQ aerosol module development: Recent enhancements and future plans. Presentation at the 2004 Models-3/CMAQ Conference, October 18-20, 2004, Chapel Hill, NC.

Binkowski, F.S.; Roselle, S.J., Models-3 Community Multi-scale Air Quality (CMAQ) model aerosol component:1. Model description. *J. Geophys. Res.* 2003, 108(D6), 4183, doi:10.1029/2001JD001409.

Byun, D.W., and K.L. Schere (2006), Review of the governing equations, computational algorithms, and other components of the models-3 Community Multi-scale Air Quality (CMAQ) modeling system, *Applied Mechanics Reviews*, 59, 51-77.

Draxler, R. R., (2003), Evaluation of an ensemble dispersion calculation. J. Appl. Meteorol., 42, 308-317.

Mathur, R., et al. (2005), Multiscale air quality simulation platform (MAQSIP): Initial applications and performance for tropospheric ozone and particulate matter, *J. Geophys. Res*., 110, D13308, doi:10.1029/2004JD004918.

Otte, T.L.,et al., 2005. Linking the Eta model with the community multiscale air quality (CMAQ) modeling system to build a national air quality forecasting system. Weather and Forecasting, 20, 367-384, 2005.

Pouliot, G.A., 2005. The emissions processing system for te Eta/CMAQ air quality forecast system, Proc. 7th Conf. on Atmospheric Chemistry, The 85th AMS Annual Meeting, Paper 4.5, Amer. Meteor. Soc., San Diego, CA.

Rogers, E., T. Black, D. Deaven, G. DiMego, Q. Zhao, M. Baldwin, N. Junker, and Y. Lin. 1996: Changes to the operational “early” Eta Analysis/Forecast System at the National Centers for Environmental Prediction. Wea. Forecasting, 11, 391-413.

Yu, S. C., R. Dennis, S. Roselle, A. Nenes, J. Walker, B. Eder, K. Schere, J. Swall, and W. Robarge, An assessment of the ability of three-dimensional air quality models with current thermodynamic equilibrium models to predict aerosol NO3- , J. Geophys. Res., 110, D07S13,doi 10.1029/2004JD004718, 2005.

Yu, S.C., P.V. Bhave, R.L. Dennis, and R. Mathur. Seasonal and regional variations of primary and secondary organic aerosols over the continental United States: Semi-empirical estimates and model evaluation. *Environmental Science & Technology* (Submitted June 2006)

Disclaimer The research presented here was performed under the Memorandum of Understanding between the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Commerce's National Oceanic and Atmospheric Administration (NOAA) and under agreement number DW13921548. This work constitutes a contribution to the NOAA Air Quality Program. Although it has been reviewed by EPA and NOAA and approved for publication, it does not necessarily reflect their policies or views.

Table 1. The total accumulation contributions of each process to $PM_{2.5}$ formation along 24-hr back-trajectories (see Figure 1) on the basis of column means from layer 1 to layer 14 (PBL height) at the five sites (2 in PA and 3 in GA).

Processes for $PM_{2.5}$ ($\mu g\ m^{-3}\ hr^{-1}$)	Site (2004) NN, GA (8/19)	SD, GA (8/19)	MD, GA (8/19)	SAHS, PA (8/17)	JOHN, PA (8/17)
Aerosol process	5.72	6.01	4.51	3.37	2.68
Cloud process	1.4×10^{-4}	1.5×10^{-4}	1.3×10^{-4}	6.60	7.28
Dry deposition	-2.0×10^{-4}	-2.8×10^{-4}	-2.2×10^{-4}	-2.4×10^{-4}	-2.3×10^{-4}
Emission	4.18	1.79	3.85	1.32	1.06
Horizontal advection	-6.61	-21.76	-18.28	5.32	-6.88
Horizontal diffusion	-0.05	0.02	0.00	-0.01	-0.01
Vertical diffusion	-0.97	-0.66	-0.60	-0.54	-0.54
Vertical advection	-1.33	-0.17	-0.28	0.05	-1.25
Total source	9.90	7.82	8.36	16.65	11.02
Total sink	-8.96	-22.59	-19.17	-0.55	-8.67
Net Change	0.93	-14.77	-10.80	16.10	2.35

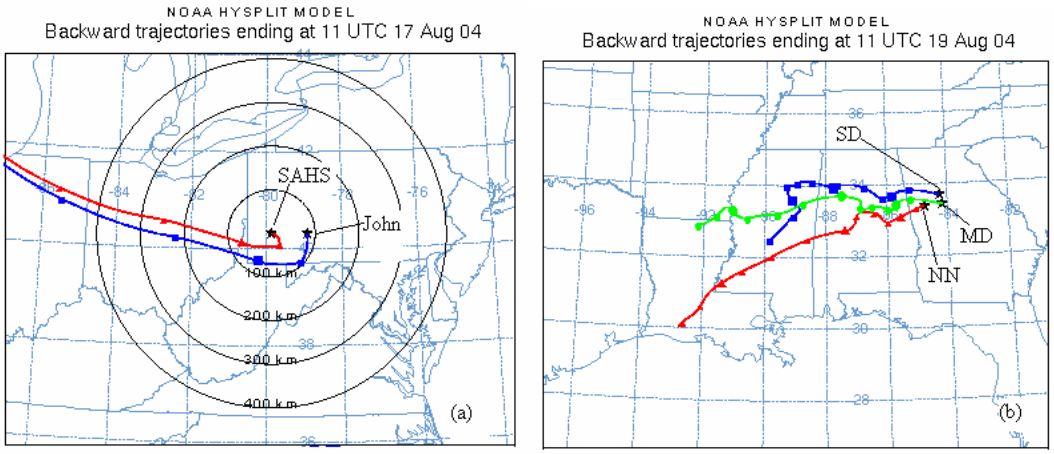

Figure 1. Backward trajectories ending at 11 UTC 17 August, 2004 at the South Allegheny High School (SAHS) and John sites, PA (a), and ending at 11 UTC 19 August, 2004, at the South Dekalb (SD), McDonough (MD) and Newnan (NN) sites, GA (b).