AIR-QUALITY MODELING OF $PM_{2.5}$ MASS AND COMPOSITION IN ATLANTA: RESULTS FROM A TWO-YEAR SIMULATION AND IMPLICATIONS FOR USE IN HEALTH STUDIES

Amit Marmur*, Armistead G. Russell, and James A. Mulholland
School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0512
e-mail: **amit.marmur@ce.gatech.edu**
Voice (404) 385-4565 Fax (404) 894-8266

1. INTRODUCTION

The purpose of this study is to evaluate the potential of using air-quality model simulations of fine particulate matter ($PM_{2.5}$) in epidemiologic studies, and to identify issues involved in the process. The focus here is on using $PM_{2.5}$ simulations generated by the US-EPA's Community Multiscale Air Quality Modeling system (CMAQ) (Byun and Ching, 1999), for a study on the health outcomes of $PM_{2.5}$ in the Atlanta metropolitan area (Tolbert et al., 2001). The use of air-quality models may introduce several benefits when applied in epidemiologic studies, compared to the use of ambient data. First, an average value over a model cell of typical size (e.g., 36km x 36km) may better represent the air quality over an applicable area as compared to a measurement at a single point within that area. Such a value can also assist in evaluating the representativeness and quality of measurements at different locations (stations) all residing within the range of the same model cell. Also, the ability to model episodes for which no measurements were performed may allow expanding the epidemiologic study to geographical areas for which no data was available, to past (historic) episodes, and to complete existing datasets.

2.0 AIR QUALITY MODELING

Fine particulate matter ($PM_{2.5}$) modeling was performed using components of the US-EPA's Models-3 modeling system, including the Penn-State/NCAR Meteorological Model (MM5) (Grell et al., 1999), the Carolina Environmental Program's (CEP) Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System version 1.5 (Houyoux et al., 2003), and CMAQ version 4.3.

* *Corresponding author*

Speciated $PM_{2.5}$ and gas phase pollutants were simulated for a two year period, 2000-2001, using a grid of 36km by 36 km cells. The grid covered the entire eastern and central USA, and was comprised of 78 by 66 cells laterally, and six vertical layers. Meteorological fields were generated by MM5, using the PX LSM. Emissions from each grid cell were generated by SMOKE based on the 1999 National Emission Inventory, projected to the year 2000. Finally, pollutant concentrations, in the form of hourly averages, were calculated by CMAQ, using the saprc99 chemical mechanism.

3.0 $PM_{2.5}$ MEASUREMENT

Total $PM_{2.5}$ mass, major ions (SO_4^{-2}, NO_3^-, NH_4^+) and carbon fractions (elemental carbon, EC; organic carbon, OC) were measured at four locations throughout the Atlanta metropolitan area. The monitoring stations from which data were used are part of two different networks: SEARCH (Southeastern Aerosol Research and Characterization) network (Hansen et al., 2003), which includes the Jefferson Street (JST) and Yorkville (YK) sites, and ASACA (Assessment of Spatial Aerosol Composition in Atlanta) network (Butler et al., 2003), which includes the South-Dekalb (SD), Fort-McPherson (FM) and Tucker (Tu) sites. Figure 1 shows the location of these sites within the Atlanta metro. For comparison with models simulation, the measurements from 2000 and 2001 were used.

4.0 RESULTS AND DISCUSSION

4.1 Spatial analysis of air-quality data

The dataset of measurements was analyzed for spatial trends, based on correlations between sites, for the period of March 1999 – August 2000.

While the average value is quite similar at different sites for many of the species, as shown in Table 1, it is the temporal trend (day-to-day variation) that would drive an epidemiologic study. To address this issue, site inter-correlations were computed, and are presented in the form of correlelograms. These correlelograms show the correlations (R) between each pair of stations, as a function of distance between the two stations.

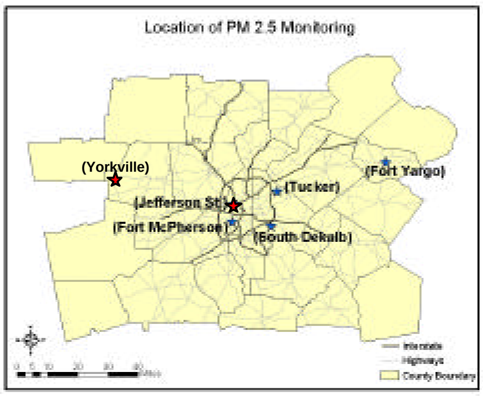

Fig. 1 Location of PM2.5 monitoring stations throughout the Atlanta Metropolitan area.

	JST	FT	TU	SD	YK
$PM_{2.5}$	21.2	19.5	21.2	-	15.3
SO_4^{-2}	5.5	5.2	5.2	5.3	5.6
NO_3^-	1.0	0.8	1.0	0.7	0.8
NH_4^+	2.1	1.9	2.0	1.8	2.6
EC	1.8	1.3	1.2	1.6	0.7
OC	4.4	4.0	4.1	4.3	3.6

Table 1 Average daily values ($\mu g/m^3$) over the period of March 1999- August 2000 for $PM_{2.5}$ species at four sites in the Atlanta metropolitan area

The correlelograms show no downward trend in the R values with increasing distance for total $PM_{2.5}$ and sulfate (Figure 2). This is due to sulfate being a secondary pollutant, and the high fraction of secondary pollutants in total $PM_{2.5}$. The co-located data is highly correlated, indicative of a low measurement error. A slight downward trend is observed for nitrate and ammonium, with a low measurement error. The slight slope is due to these being secondary pollutants. The slope could be indicative of differences in local availability of ammonia gas, controlling the formation of ammonium-nitrate (Russell et al., 1983). Correlations for EC and OC are significantly lower, and a clear downward slope is seen for EC (Figure 3). Higher measurement errors are observed, along with a strong local effect due to primary emissions of EC and OC.

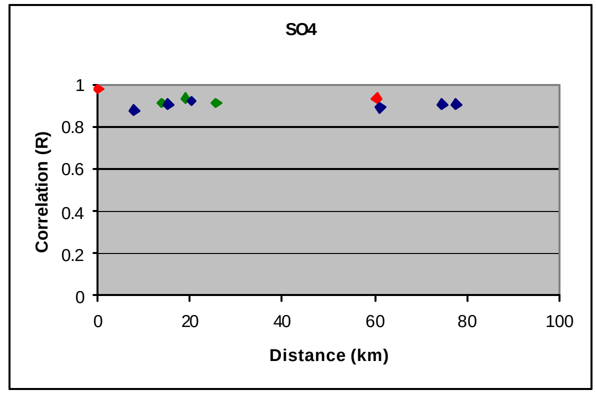

Fig. 2 SO_4^{-2} correlelogram for JST, FT, SD, TU and YK sites shows typical secondary pollutant trend

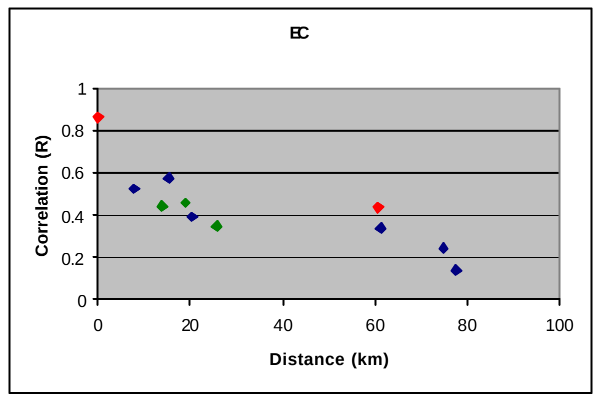

Fig. 3 EC correlelogram for JST, FT, SD, TU and YK sites shows typical primary pollutant trend

From an epidemiologic view point, these correlelograms suggest that the location of measurement is not of great significance for $PM_{2.5}$ and sulfate, as all sites are reasonably well correlated. The same is true for nitrate and ammonium, within the urban range. For the carbon species, on the other hand, correlations were lower, due to both measurement error and local effects. Hence, a local measurement is not as representative of the carbon levels over the entire domain, though the results here do not suggest that any one site has an obvious advantage.

4.2 Air-quality model simulations

CMAQ simulated concentrations of SO_4^{-2}, NO_3^-, NH_4^+, EC, OC and total $PM_{2.5}$ were compared with the observations at four different sites in the Atlanta metropolitan area. All four sites used here are located within the same 36km grid cell for which CMAQ results are reported. Thus, all four daily values are comparable with the same

CMAQ value. Such a comparison allowed us to address two issues. First was to evaluate CMAQ performance based on statistical measures such as the correlation coefficient and the root-mean-square-error (RMSE). Second was to suggest whether a single site exists which is more representative of the health study domain, which is primarily the Atlanta metropolitan area, and hence more suitable for use by the epidemiologic study. This was based on the comparison to the levels generated by CMAQ.

CMAQ performance for sulfate seems to be reasonably good (Table 2, Fig. 3). The annual average value is very well predicted for all sites, and the correlation coefficient is relatively high for JST. Many of the day-to-day variations, as well as the seasonal variations seem to be captured by the model. When comparing the model correlations with correlation between sets of measured data at different sites, it seems that the model output is as correlated as the data. Interpreting the deviation from a perfect correlation as measurement/analysis and "spatial" error, it seems that the error generated by the model is not larger than the total error in the measured data. Given that, use of the air-quality model results holds promise in simulating sulfate levels and trends in accuracies allowing its use in epidemiologic studies.

SO_4^{-2}	JST	FTM	SD	TU	CMAQ
JST	1.00				
FTM	0.73	1.00			
SD	0.59	0.74	1.00		
TU	0.70	0.65	0.67	1.00	
CMAQ	0.73	0.54	0.44	0.49	1.00
Mean (mg/m3)	4.86	4.33	4.27	4.14	4.77
StDev (mg/m3)	3.19	3.06	3.26	3.34	3.07
RMSE (mg/m3)	2.30	3.02	3.41	3.31	-

Table 3 Correlation coefficients (R) between SO_4^{-2} CMAQ simulations and data measured at four sites in the Atlanta metropolitan area, along with mean values, standard deviations and root-mean-square-error (RMSE) between CMAQ simulation and measurement (daily values, January 2000 – December 2001)

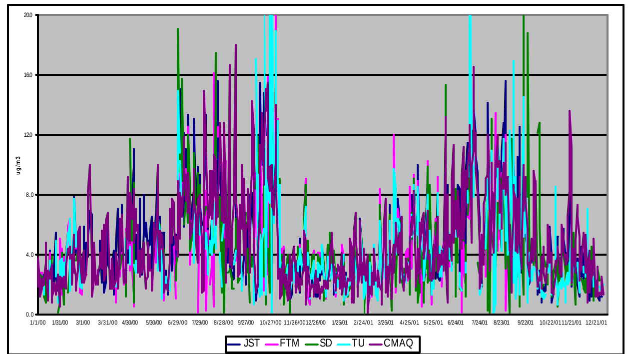

Fig. 4 Daily averages (1/2000-12/2001) of SO_4^{-2} concentrations ($\mu g/m^3$) simulated by CMAQ and measured at four sites in the Atlanta metropolitan area

CMAQ performance for nitrate was less encouraging (Table 3). Even though the correlation based on the entire two-year dataset was quite reasonable (0.59 at JST), the annual average was extremely over-predicted. A detailed look at Figure 5 indicates that the reason for this over prediction is the periodic spikes in the model output. It appears that the model's nitrate mechanism is over sensitive to variations in temperature, or that more ammonia is simulated to be present than is actually the case. The correlation was still relatively high, mainly due to the model capturing the seasonal trends. When compared to the correlations between the different sites, CMAQ did not seem to capture the daily trends in nitrate levels as well as the measurements, and therefore does not currently seem to be as robust for use in epidemiologic studies.

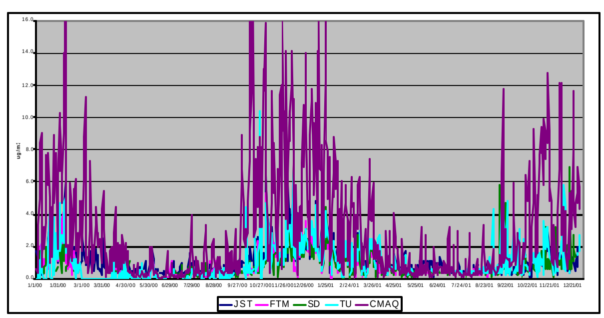

Fig. 5 Daily averages (1/2000-12/2001) of NO_3^- concentrations ($\mu g/m^3$) simulated by CMAQ and measured at four sites in the Atlanta metropolitan area

NO3-	JST	FTM	SD	TU	CMAQ
JST	1.00				
FTM	0.84	1.00			
SD	0.70	0.89	1.00		
TU	0.74	0.81	0.71	1.00	
CMAQ	0.59	0.56	0.49	0.50	1.00
Mean (mg/m3)	1.05	1.03	0.95	1.23	2.99
StDev (mg/m3)	0.93	1.02	1.04	1.24	3.64
RMSE (mg/m3)	3.42	4.26	4.03	4.15	-

Table 4 Correlation coefficients (R) between NO_3^- CMAQ simulations and data measured at four sites in the Atlanta metropolitan area, along with mean values, standard deviations and root-mean-square-error (RMSE) between CMAQ simulation and measurement (daily values, January 2000 – December 2001)

Simulated levels of both EC and OC were lower than observations, but when model correlations were compared to the site-to-site ones, it seems that the simulated levels are as good as the data. There was no significant seasonal trend in the EC and OC data. As previously mentioned, the low site-to-site correlations indicate large measurement/analysis errors and spatial variations. It seems that the air-quality model was able to simulate trends, at some of the sites, better than the multiple measurements at various sites. It is still unclear whether the air-quality model captures correctly all the factors driving EC and OC concentrations, but it is evident that based on the limitations of the available EC/OC measurements with regards to epidemiologic studies, using the simulated results can serve as an alternative.

5.0 CONCLUSIONS

Results from air-quality model simulations for $PM_{2.5}$ were evaluated and compared with data from four monitoring sites in the Atlanta metropolitan area. Model performance for sulfate, EC, OC and total $PM_{2.5}$ was reasonably good, as reflected by the model's ability to simulate the day-to-day variability in pollutant concentrations. For these species, correlations between CMAQ output and measured data were as high as the correlations between the different measured datasets. Model performance for nitrate and ammonium was not as good, as the models ability to simulate the day-to-day variations was limited. The comparison of data from four different sites to CMAQ simulations for sulfate, EC, OC and total PM2.5 indicated JST site as the most correlated with the model. Given the model's ability to simulate the temporal variations, and the fact that its output represents and average value over a relatively large local domain, it appears that using air quality models, such as CMAQ, can serve as a useful tool for epidemiologic studies.

7.0 ACKNOWLEDGEMENTS

This work was supported by subcontractors to Emory University under grants from the USEPA (R82921301-0) and NIEHS (R01ES11199 and R01ES11294). We also thank Georgia Power and Southern Company for continuing support and the individuals at ARA for providing access to data.

8.0 REFERENCES

Butler, A.J., Andrew, M.S., and Russell, A.G., 2003: Daily sampling of $PM_{2.5}$ in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. *Journal of Geophysical Research – Atmospheres*, **108**.

Byun, D.W.; Ching, J.K.S. Science Algorithms of the EPA-Models-3 Community Multiscale Air Quality (CMAQ) Modeling System, Office of Research and Development, Washington, DC 20460, 1999; EPA/600/R-99/030.

Grell, G.A; Dudhia, J.; Stauffer, D.R. A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). 1999, NCAR Tech. Note NCAR/TN-398+STR.

Hansen, D.A., Edgerton, E.S., Hartsell, B.E., Jansen, J.J., Kandasamy, N., Hidy, G.M., and Blanchard, C.L., 2003: The southeastern aerosol research and characterization study: Part 1-overview. *Journal of the Air & Waste Management Association*, **53**: 1460-1471.

Houyoux M.; Vukovich J.; Brandmeyer, J.E.; Seppanen, C.; Holland, A. Sparse Matrix Operator Kernel Emissions Modeling System (SMOKE): User Manual, Version2.0, Carolina Environmental Program, 2003. Russell, A.G., McRae, G.J., Cass, G.R.,: 1983: Mathematical-Modeling of the Formation and Transport of Ammonium-Nitrate Aerosol, *Atmospheric Environment*, **17**:949-964.

Tolbert, P., Klein, M., Metzger, K., Peel, J., Flanders, W.D., Todd, K., Mulholland, J., Ryan, P.B., Frumkin, H., 2001: Particulate pollution and cardiorespiratory emergency department visits in Atlanta, August 1998-August 2000 (ARIES/SOPHIA studies). *Epidemiology*, **12**:265.