

## High Resolution Air Quality Modeling for the Mexico City Metropolitan Zone using a Source-Oriented CMAQ model – Part I: Emission Inventory and Base Case Model Results

Sajjad Ali, Gang Chen, Hongliang Zhang and Qi Ying\*

Department of Civil Engineering, Texas A&M University, College Station, Texas, USA 77845-3136

Iris V. Cureño, Adrián Marín, Humberto A. Bravo\*, Rodolfo Sosa

Centro de Ciencias de la Atmósfera, Sección de Contaminación Ambiental, Universidad Nacional Autónoma de México, Ciudad Universitaria, México D.F. CP 04510

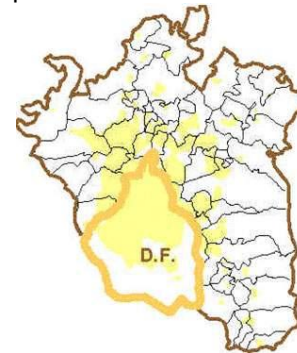
### 1. INTRODUCTION

Degradation of air quality in the urban areas can directly harm the health of a large population, adversely affect the built and natural environment of the surrounding regions, and contribute to global climate change. In this aspect, the rapid growth of megacities (urban areas with a population of over 10 million) in the developing world being a major source of atmospheric pollution is a cause for concern.

The Mexico City Metropolitan Zone (MCMZ) is situated on an elevated basin 2240 m above sea level. MCMZ has a population of around 20 million inhabitants, around 4 million vehicles, and over 40,000 industries that contribute to atmospheric pollution. The MCMZ (Figure 1) covers an area of approximately 7700 km<sup>2</sup> that is constrained by mountain ridges. The surrounding mountains tend to trap pollutants within the MCMZ basin. The high altitude and temperate climate lead to intense sunlight that aids the photochemical processes of ozone and other oxidants.

Several different models have been used in the past to help understand the formation of air pollutants in the MCMZ, including the offline CAMx/MM5 (Lei et al., 2008) and WRF-Chem (Zhang et al., 2009). Horizontal grid resolution up to 3 km have been used in these studies. Although the modeling domains in these studies typically cover MCMZ and surrounding regions, most of the studies did not include anthropogenic emissions of gaseous and PM pollutants from sources outside the MCMZ, the PM emissions due to windblown dust and SO<sub>2</sub> emissions from Popocatepetl, an active volcano 70 km southeast of Mexico City, although several individual studies have shown

that these sources could all contribute significantly to the observed concentrations in the MCMZ. In addition, the MCMZ anthropogenic emissions used in these studies are generally outdated and could not represent the actual emissions during the modeling episodes.



**Figure 1 Mexico City Metropolitan Zone**

The objectives of this study are to 1) develop a modeling emission inventory for the MCMZ and surrounding areas that utilizes the most updated information and includes all major sources that could potentially affect air quality in the MCMZ, and 2) evaluate the emission inventory by comparing 3D air quality model predictions with all available observation data.

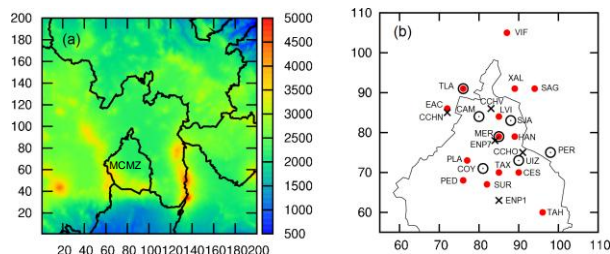
### 2. MODEL APPLICATION

In this study, a source-oriented Community Multiscale Air Quality Model (CMAQ) based on CMAQ v4.7.1 is applied to model air quality in the MCMZ and surrounding regions that cover an area of 200x200 km<sup>2</sup> during a six-day air quality episode from March 2-7, 2006 with 1 km spatial resolution. Figure 1 shows the surface elevation of the domain and the location of the monitoring stations for meteorology and air pollutants. The gas phase mechanism SAPRC-99 and the aerosol module AERO5 are used for this study.

The initial and hourly boundary conditions are generated using the GEOS-Chem model for the nested North America domain with 0.5x0.666 grid

\*Corresponding authors: Qi Ying, Department of Civil Engineering, Texas A&M University, College Station, TX 77843-3136; e-mail: [qying@civil.tamu.edu](mailto:qying@civil.tamu.edu). H.A. Bravo, Centro de Ciencias de la Atmósfera, Sección de Contaminación Ambiental, Universidad Nacional Autónoma de México, Ciudad Universitaria, México D.F. CP 04510; e-mail: [hbravo@servidor.unam.mx](mailto:hbravo@servidor.unam.mx)

resolution (<http://acmg.seas.harvard.edu/geos/>). A separate set of initial and boundary conditions was prepared using the default CMAQ profiles and the differences in the model predictions will be discussed in Section 4.



**Figure 2** (a) Surface elevation (m) of the 200x200 1-km horizontal resolution modeling domain and (b) locations of gas phase pollutants and PM10 (•), PM2.5 (o) and meteorology (x) monitoring sites.

The meteorology conditions during this episode are generated using the Advanced Research Weather Research Forecast (WRF) model v3.1. A 3-level nested domain with Lambert conformal projection was chosen to run the WRF model. The horizontal grid sizes are 9km, 3km and 1km respectively with 29 vertical layers. Initial and boundary conditions for the WRF simulations are prepared using the North American Regional Reanalysis (NARR) data set. The WRF output was processed using the Meteorology and Chemistry Interface Preprocessor (MCIP) program.

The source-oriented CMAQ model is capable of directly determining the source contributions of VOCs to ozone (Ying and Krishnan, 2010) and secondary organic aerosol (SOA) formation (Zhang and Ying, 2010). The focuses of this study are on the development of emission inventory and evaluation of the base case model performance so the source-oriented features of the CMAQ model are not utilized.

### 3. EMISSION INVENTORY

#### 3.1 The 2006 Emission Inventory for MCMZ

Emissions of gaseous and particulate matter for the MCMZ used in this study are provided by Mexico City's Secretary of Environment based on the most recent update of the 2006 emission inventory. The emissions reported in the 2006 emission inventory for point, mobile and area sources are generated by the SPEA program (Processing System of Atmospheric Emissions) version 1.0.0. (Ortiz L.M., 2005), which generates temporally and spatially allocated emissions for all the criteria pollutants.

The information necessary to generate point source emissions is obtained from the COA

(Annual Operation Certificate), which is the official document where the industries report the emissions generated for their operations along with other useful operational information. Hourly emissions from point sources are based on the work-hours of the different industries. Generally industries in Mexico operate 8 hours or more per day and the highest emissions for point sources run from 0600-1800 LST (Figure 3).

The emissions generated by area sources are obtained using emission factors based on population. The population data and the activities information are taken from the AGEBS (Basic Geostatistical Areas) by INEGI (National Institute of Statistics and Geography). The AGEBS are geographic areas made for census purposes in Mexico considering the population and their activities per small areas which are part of the MCMZ. The hourly-temporal distribution for area sources is based in the activity-hours of the population. Therefore, the highest emissions for area sources run from 1000-2000 LST.

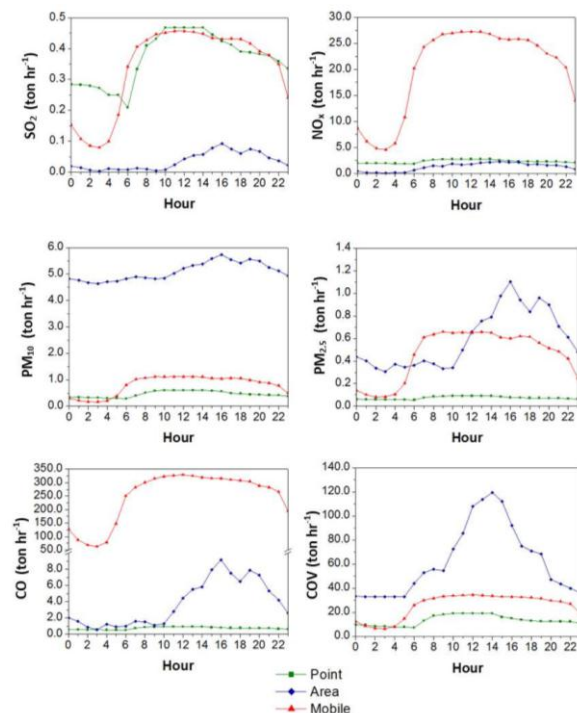
The emissions for mobile sources of major fuel types (i.e. gasoline, diesel, liquefied petroleum gas and natural gas) are calculated using the emission factors generated by the MOBILE5-Mexico model (which considers Mexico specific sulfur content, RVP and altitude) and the kilometers traveled (KMT) in a day for the vehicle. The mobile source emissions are allocated to the main avenues and streets inside of the studied area. The hourly-temporal distribution for mobile sources is based in the higher hours of traffic vehicle observed in the MCMZ (based on a document "Sistema de Información de Condiciones de Tránsito para la Estimación de Emisiones Contaminantes por Fuentes Móviles en la Zona Metropolitana del Valle de México" ETEISA, 2003). Consequently, the highest emissions for mobile sources are presented from 0400-2100 LST (Figure 3).

The speciation profiles used to split the total VOCs emissions into individual SAPRC-99 model species, and PM2.5 emissions into elemental carbon (EC), organic compounds (OC), nitrate, sulfate and other components are from the SPECIATE 3.2 database.

#### 3.2 Other Emission Data Sources

The emissions from all the point sources outside of the MCMZ were included from the 1999 Mexico National Emissions Inventory (MNEI) (<http://www.epa.gov/ttnchie1/net/mexico.html>). The day-specific emission rates of NO<sub>x</sub> and SO<sub>2</sub> from the Francisco Pérez Ríos thermoelectric power plant, located in the Tula industrial complex, Hidalgo State, are based on the actual fuel oil

consumptions records and sulfur content of the fuel oil used by the power plant. Several errors in the original MNEI for the thermoelectric power plant, including stack height, stack diameter and number of stacks, are also corrected using the data acquired directly from the plant. Moreover, the SO<sub>2</sub> emissions from Popocatepetl, an active volcano 70 km southeast of Mexico City are included. Area and mobile sources outside the MCMZ were not included in the current model emission inventory.



**Figure 3** Temporal distribution by source and specie obtained from the 2006 emission inventory for the Mexico City Metropolitan Zone by Mexico City's Secretary of Environment.

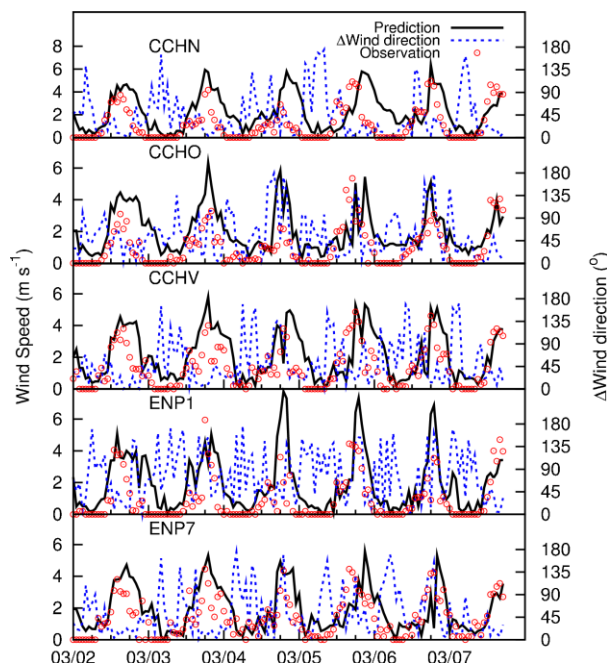
Windblown dust emissions in the entire domain are predicted based on the model described by Fernando (2008) and Shaw et al. (2008), with slight modifications. In summary, the predicted surface friction velocity and soil moisture from WRF as well as the soil type and land use/land cover information is used to predict the vertical flux of PM<sub>10</sub>. Emission rate of PM<sub>2.5</sub> is set to be 6% of PM<sub>10</sub> and windblown dust speciation profiles from SPECIATE 3.2 is used to calculate the emission rates of individual PM<sub>2.5</sub> components. More details of the windblown dust model and the contributions of windblown dust to PM<sub>2.5</sub>/PM<sub>10</sub> will be documented elsewhere.

Biogenic emissions were generated using the Biogenic Emissions Inventory System, Version 3 (BEIS3) included in the SMOKE distribution. The 1-km resolution BELD3 and cover data with 230 different cover types were used to estimate emissions from vegetation and soil.

## 4. RESULTS AND DISCUSSION

### 4.1 Meteorology

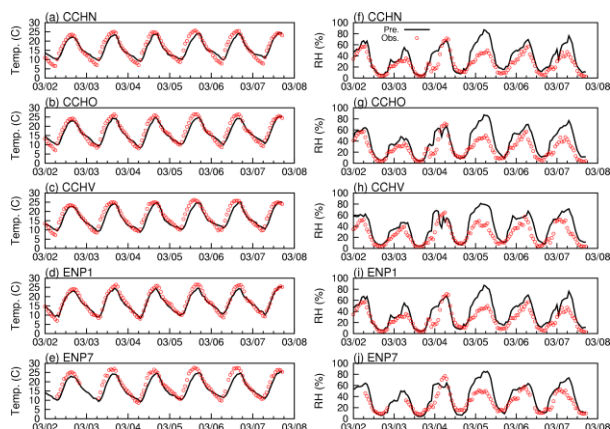
Figure 4 shows that the WRF model generally captures the diurnal variation of the wind speed in the MCMZ. Wind speed is highest in the later afternoon and lower at night and early morning hours. The WRF predictions are slightly higher than the observations. The predicted wind directions agree better with observations when wind speed is higher but there are larger differences when the wind speed is low. The differences in the wind directions are not likely to significantly affect air quality model results because of the slower wind speed.



**Figure 4** Predicted and observed surface wind speed and the difference between the observed and predicted wind direction.

Figure 5 shows that the predicted temperature and relative humidity agrees well with observations. The WRF model seems to over-predict relative humidity at all stations on the morning of March 5<sup>th</sup> and 7<sup>th</sup> by approximately 20% but otherwise the predictions are reasonably well.

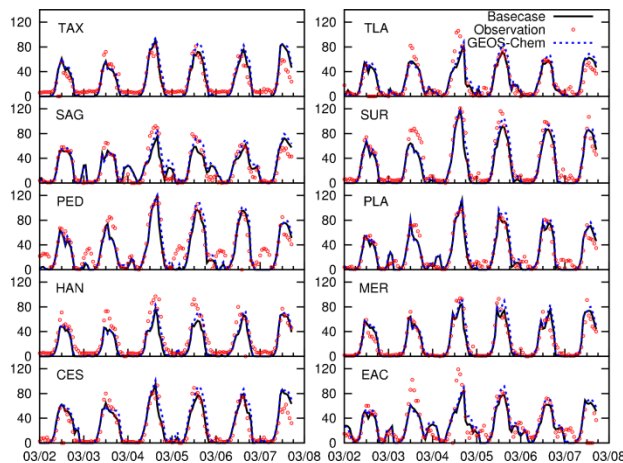




**Figure 5** Predicted and observed surface temperature and relative humidity.

#### 4.2 Gas Phase Pollutants

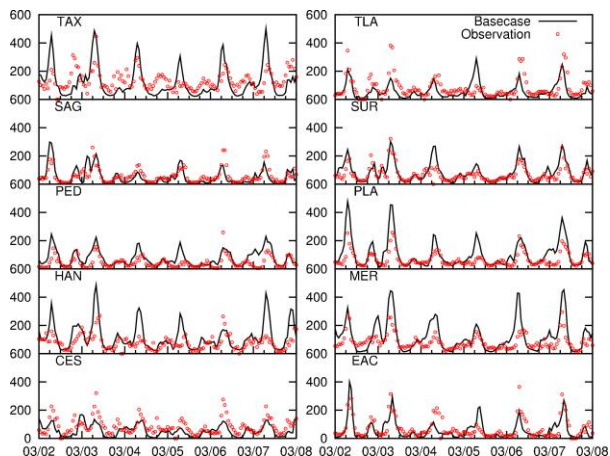
Figure 6 shows the predicted  $O_3$  concentrations at 10 observation stations. The locations of the stations can be seen in Figure 1. The base case results are based on the default CMAQ initial and boundary conditions. The dash lines represent predictions using boundary conditions generated from the GEOS-Chem simulation as discussed before. Both sets of simulations agree well with observations although peak  $O_3$  concentrations are under-predicted on a few days. The simulation using GEOS-Chem boundary conditions predicts slightly higher  $O_3$  than the default boundary conditions. The under-prediction of  $O_3$  peaks are likely caused by the overestimation of  $NO_x$  emissions coupled with the underestimation of VOC emissions (see Figure 7 and Figure 8 below).



**Figure 6** Time series of predicted and observed  $O_3$  concentrations. Units are ppb.

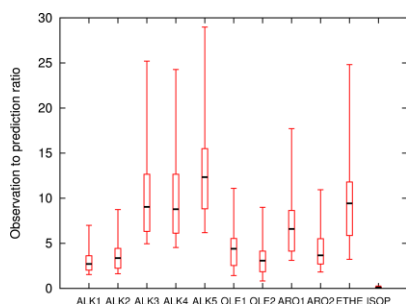
Figure 7 shows that although the model correctly captures the diurnal variation of  $NO_x$  concentrations at all stations, the morning peaks

of  $NO_x$  are over-predicted, especially for stations in the northern part of DF. The over-prediction problem is not likely caused by insufficient vertical mixing as the minimal vertical diffusion coefficients were calculated based on urban fraction and a minimum value of  $2.0 \text{ m}^2 \text{ s}^{-1}$  is used for 100% urban regions, such as most part of DF. It has been reported that  $NO_x$  emissions from MOBILE5 can be over-estimated by a factor of 2 (Singh and Sloan, 2006). Thus, the overestimation is likely caused by uncertainties in the MOBILE5 model. Overestimation of  $NO_x$  emissions will enhance the titration of  $O_3$  during the day, leading to under-predictions of  $O_3$  and over-predictions of  $NO_2$ . The  $NO_2$  concentrations are indeed over-predicted (not shown).  $CO$  concentrations in the urban center stations are also over-predicted. Further experimental and modeling studies to validate/correct the MOBILE5-Mexico emission estimates of mobile source emissions are necessary to provide more confidence in the mobile source emissions.



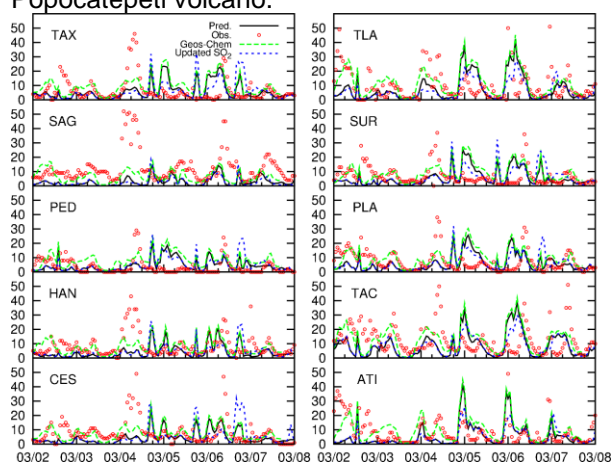
**Figure 7** Time series of predicted and observed  $NO_x$  concentrations. Units are ppb.

Figure 8 shows the box-and-whisker plot of the ranges of the ratio of the observed SAPRC-99 VOC concentrations (based on 54 individual PAMS VOC species) to the predicted SAPRC-99 VOC concentrations. 24-hour average samples were taken at 5 monitoring stations every 6 days. The data shown here is based on an extended modeling exercise for March 2<sup>nd</sup> to March 31<sup>st</sup> with 19 data points for each VOC species. Light alkanes and alkenes are under-predicted by a factor of 2-5 while aromatics and larger alkanes are under-predicted by approximately an order of magnitude. The under-prediction of aromatics and long chain alkanes (ALK5) could have significant impacts on the SOA predictions.



**Figure 8** Observation to prediction ratio of VOC species. Median, minimum, maximum and 25% to 75% quartiles are shown on the plot.

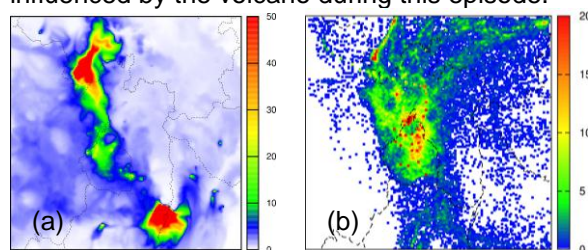
Figure 9 shows the predicted and observed  $\text{SO}_2$  concentrations. The model does not predict the  $\text{SO}_2$  well and under-predicts high  $\text{SO}_2$  concentrations on several instances. A regional  $\text{SO}_2$  peak occurs on the morning of March 4<sup>th</sup> while none of the simulations are able to capture, although the GEOS-Chem simulation appears to yield better results, suggesting the regional increase of  $\text{SO}_2$  might occur in an area larger than the current domain. Back trajectories analysis shows that air masses are coming from the north when the high  $\text{SO}_2$  concentrations occur, thus rule out the possibility of the contributions from the Popocatepetl volcano.



**Figure 9** Time series of predicted and observed  $\text{SO}_2$  concentrations. Units are ppb. Base case results are based on the default MNEI 1999 with  $\text{SO}_x$  emissions reduced by 40%. Updated  $\text{SO}_2$  case uses the  $\text{SO}_2$  emissions based on daily oil-fuel consumption and fuel sulfur content from the Francisco Pérez Ríos thermoelectric power plant in the Tula industrial complex.

Episode-average regional distribution of  $\text{SO}_2$  shown in Figure 10(a) clearly shows the locations of the two major  $\text{SO}_2$  sources, the Tula industrial complex to the northwest and the Popocatepetl volcano to the southeast. Hour 1200-1300 is

selected because the highest  $\text{SO}_2$  concentration in the MCMZ occurs at this hour. The emissions from the Tula industrial complex have greater impacts on the MCMZ than the volcano during this episode. Other sources outside MCMZ do not contribute significantly. Back trajectories calculated using the HYSPLIT model (with the same meteorology files used to drive the CMAQ model) are shown in Figure 10(b). Significant amount of trajectories pass through the downwind regions of the Tula industrial complex. There are fewer trajectories pass through the areas influenced by the volcano during this episode.



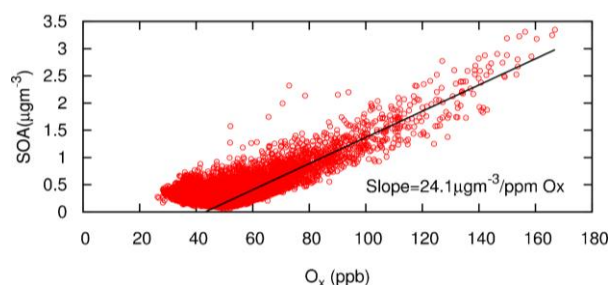
**Figure 10** (a) Episode averaged distribution of predicted surface  $\text{SO}_2$  at 1200-1300 LST. Units are ppb. The maximum concentration is set to 50 ppb to better show the spatial distribution of  $\text{SO}_2$ . (b) Frequency distribution of 24-hour back-trajectories ending at all monitoring stations at all hours throughout the episode. Maximum at trajectory end points should be 144 but is scaled to better show the spatial distributions.

The difficulty in  $\text{SO}_2$  predictions in MCMZ is that most of the  $\text{SO}_2$  is not produced locally but arrives at MCMZ through long range transport. Gas flow rate and exhaust exit temperatures are key parameters that determines the plume rise and the dispersions of  $\text{SO}_2$  but are generally not monitored. The parameters used in the current study are based default US EPA data.

### 4.3 Secondary Organic Aerosol

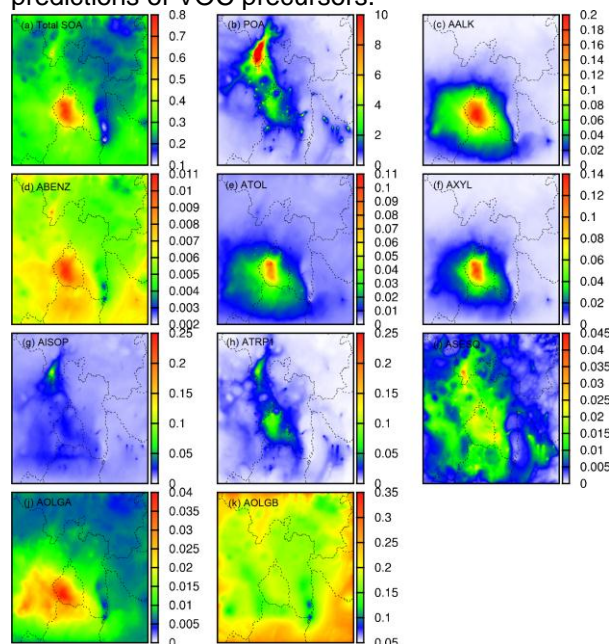
Both SOA and odd oxygen ( $\text{O}_x = \text{NO}_2 + \text{O}_3$ ) are formed as products of VOCs oxidation. Figure 11 shows the correlation between SOA and  $\text{O}_x$  along 12-hour back trajectories ending at all monitoring stations during March 2<sup>nd</sup> to 7<sup>th</sup>.  $\text{O}_x$  concentrations reach as high as 170 ppb. The slope of a linear fit between  $\text{O}_x$  and SOA using data points with  $\text{O}_x$  greater than 70 ppb is  $24.1 \mu\text{g m}^{-3} / \text{ppm O}_x$ . The predicted peak SOA concentrations are a factor of 3-10 lower than AMS measurements (Wood et al., 2010) while the  $\text{O}_x$  concentrations are also under-predicted. While it is possible that the under-prediction of SOA could be caused by under-predictions of VOC precursors as shown in Figure 8, further investigation is necessary to determine

the causes of the gap between predictions and observations.



**Figure 11** Correlation of SOA and O<sub>x</sub>

Figure 12 shows the regional distributions of episode-average SOA concentrations. Alkanes and aromatic compounds are major anthropogenic contributors to SOA while isoprene and monoterpenes are major contributors to biogenic SOA. In addition, oligomerized biogenic SOA contributes to approximately  $0.2 \mu\text{g m}^{-3}$ . It should be noted that these estimations are likely low by an order of magnitude because of under-predictions of VOC precursors.



**Figure 12** Episode average of regional distribution of SOA and SOA components. POA is the PM<sub>2.5</sub> primary organic aerosol.

## ACKNOWLEDGMENT

This research is funded by TAMU-CONACyT collaborative research program, project #2009-0055. IVC is also supported by an exchange student fellowship from CONACyT for her visit at TAMU. The authors also would like to thank Ing. Jorge Sarmiento, Magdalena Armenta, Patricia Camacho, and Francisco Hernández from the

Mexico City Secretary of Environment for providing the 2006 emission inventory and Pablo Sánchez, A. from the Environmental Pollution Section, CCA-UNAM for his help in the project.

## REFERENCES

- Choi, Y.J., Fernando, H.J.S., 2008. Implementation of a windblown dust parameterization into MODELS-3/CMAQ: Application to episodic PM events in the US/Mexico border. *Atmospheric Environment* 42, 6039-6046.
- Lei, W., Zavala, M., de Foy, B., Volkamer, R., Molina, L.T., 2008. Characterizing ozone production and response under different meteorological conditions in Mexico City. *Atmospheric Chemistry and Physics* 8, 7571-7581.
- Molina, L.T., Molina, M.J., 2002. *Air Quality in the Mexico Megacity: An Integrated Assessment*. Kluwer Academic Publishers.
- Molina, M.J., Molina, L.T., 2004. Megacities and atmospheric pollution. *Journal of the Air & Waste Management Association* 54, 644-680.
- Shaw, W.J., Allwine, K.J., Fritz, B.G., Rutz, F.C., Rishel, J.P., Chapman, E.G., 2008. Evaluation of the wind erosion module in DUSTAN. *Atmospheric Environment* 42, 1907-1921.
- Singh, R.B., Sloan, J.J., 2006. A high-resolution NO<sub>x</sub> emission factor model for North American motor vehicles. *Atmospheric Environment* 40, 5214-5223.
- Wood, E.C., Canagaratna, M.R., Herndon, S.C., T.B., O., Kolb, D.R., Worsnop, J.H., Kroll, W.B., Knighton, W.B., Seila, R., Zavala, M., Molina, L.T., DeCarlo, P., Jimenez, J.L., Weinheimer, A.J., Knapp, D.J., Jobson, B.T., Stutz, J., Kuster, W.C., Williams, E.J., 2010. Investigation of the correlation between odd oxygen and secondary organic aerosol in Mexico City and Houston. *Atmospheric Chemistry and Physics Discussions* 10, 3547-3606.
- Ying, Q., Krishnan, A., 2010. Source contributions of volatile organic compounds to ozone formation in southeast Texas. *J. Geophys. Res.* 115, D17306.
- Zhang, H., Ying, Q., 2010. Secondary Organic Aerosol Formation and Source Attribution in Southeast Texas. *Atmospheric Environment* Submitted.
- Zhang, Y., Dubey, M.K., Olsen, S.C., Zheng, J., Zhang, R., 2009. Comparisons of WRF/Chem simulations in Mexico City with ground-based RAMA measurements during the 2006-MILAGRO. *Atmospheric Chemistry and Physics* 9, 3777-3798.