USING CMAQ TO INTERPOLATE AMONG CASTNET MEASUREMENTS

Betty K. Pun* and Christian Seigneur
Atmospheric and Environmental Research, Inc., San Ramon, CA, USA

1. INTRODUCTION

The Clean Air Status and Trends Network (CASTNET) comprises 95 active monitoring stations across the United States (http://www.epa.gov/castnet/). These stations, primarily located in rural areas, provide measurements on weekly atmospheric concentrations of sulfate, total nitrate, ammonium, and sulfur dioxide. Dry deposition fluxes of these species are estimated based on the measured atmospheric concentrations, meteorological data, and information on land use, vegetation, and surface conditions.

Estimating regional concentrations and deposition fluxes from point measurements requires the use of interpolation. Historically, simple inverse weighting algorithms have been applied. An example is shown in Figure 1.

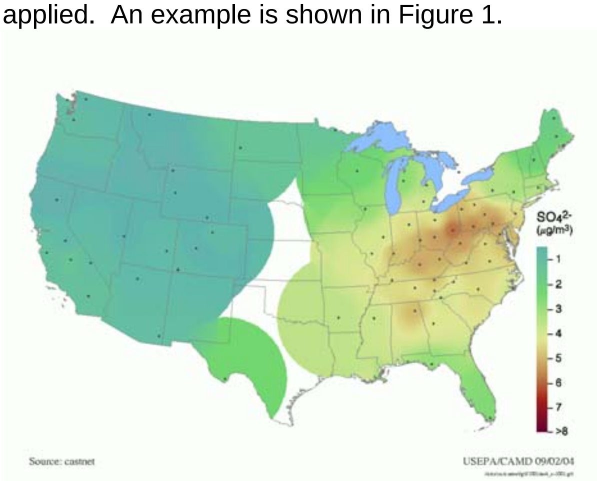

Figure 1. Interpolated concentration fields of particulate sulfate using an inverse distance approach applied to CASTNET measurements in 2001 (source: http://www.epa.gov/castnet/mapconc.html)

Simple algorithms have obvious limitations in areas without good monitor coverage. Where there are few monitors, undue influence form single monitors can result in unrealistic gradients. To generate realistic estimates, it is desirable to incorporate other information into the interpolation process. Such information includes precursor emission density, topography, land use, and other factors that influence the spatial distribution of concentrations and deposition fluxes.

Chemical transport models (CTMs) embody current data and knowledge in emissions, land use, meteorology, and chemistry. Therefore, using the spatial distribution predicted by CTMs such as CMAQ to provide gradient information between point measurements at CASTNET sites provides a strong basis for interpolated estimates of regional concentrations and dry deposition fluxes.

2. METHOD

2.1 Correspondence between CMAQ model species and CASTNET measurements

CASTNET uses open-face, three-stage filter packs to collect atmospheric particles. Particulate matter (PM) sulfate, sulfur dioxide (SO_2), PM nitrate, nitric acid (HNO_3), and PM ammonium are recorded. Because of the volatility of PM nitrate, conversion between HNO_3 and PM nitrate is expected to occur over the sampling period of one week due to changes in ambient temperatures. Total nitrate is expected to be conserved in the samples. PM ammonium is also subject to phase transition; but measurements of ammonia (NH_3) are not available and total ammonium cannot be estimated.

CASTNET filter packs are not configured with inlets to restrict particle size. CMAQ simulates PM concentrations associated with Aitken, accumulation, and coarse modes. At most locations, PM sulfate, nitrate, and ammonium are expected to be associated with fine PM (Aitken and accumulation modes). In areas influenced by sea salt or alkaline dust, nitrate and sulfate may be present in coarse particles.

Dry deposition fluxes are derived from CASTNET measurements of concentrations and dry deposition velocities estimated from local conditions. The same general formulation is used in CTMs. However, algorithms for estimating dry deposition velocities may be different, especially because grid-averaged meteorology and land use are used. Table 1 lists the correspondence between CASTNET data and the model species.

**Corresponding author:* Betty K. Pun, Atmospheric and Environmental Research, Inc., 2682 Bishop Drive, Suite 120, San Ramon, CA 94583; e-mail: pun@aer.com

Table 1. Correspondence between measured/estimated quantities of CASTNET and CMAQ model species used in their interpolation.

	CASTNET quantities	CMAQ variables	Notes
PM sulfate	SO4	ASO4I+ASO4J	1
sulfur dioxide	totalSO2	SO2	
total nitrate	totalNO3	ANO3I+ANO3J+HNO3	1
PM ammonium	NH4	ANH4I+ANH4J	1,2
PM sulfate dry deposition	SO4_FLUX	ASO4Idd+ASO4Jdd	1,3
sulfur dioxide dry deposition	SO2_FLUX	SO2dd	1,3
total nitrate dry deposition	NO3_FLUX+HNO3_FLUX	ANO3Idd+ANO3Jdd+HNO3dd	1,3
PM ammonium dry deposition	NH4_FLUX	ANH4Idd+ANH4Jdd	1,2,3

1. particle size range may not correspond exactly because open-face filters do not have size cut off but model simulates Aitken and accumulation modes.
2. may be subject to volatilization.
3. deposition velocity estimates may differ between CASTNET and CMAQ.

2.2 Interpolation Methodology

The interpolation procedure combines unevenly spaced observation with evenly spaced model predictions using the following guiding principles. First, interpolated concentrations equal measured concentrations at locations with measurements. Second, interpolated concentrations equal modeled concentrations at locations with no nearby observations. Third, interpolation of concentrations near one or more monitors is governed by the magnitude of the concentration(s) at the nearby monitor(s) and the gradients in the modeled concentration fields.

2.2.1 Formulation

To generate an evenly spaced field of interpolated concentrations, an evenly spaced field of corrections is applied to the modeled concentrations. At any monitor (k = ksite), the error (E_{ksite}) is defined in a manner similar to standard model evaluation procedures.

$$E_{ksite} = C_{o,ksite} - C_{s,icell,jcell} \quad (1)$$

where $C_{o,\,ksite}$ is a valid observation at ksite and $C_{s,icell,\,jcell}$ is the simulated grid-average concentration in the grid cell (icell, jcell) where the monitor is located.

To obtain the interpolated concentration at the center of a grid cell, the error at the center of the grid cell is applied to the simulated concentration

$$C_{i,icell,jcell} = E_{icell,jcell} + C_{s,icell,jcell} \quad (2)$$

The error at the center of a grid cell is defined as the weighted sum of the errors at the nearby monitoring sites.

$$E_{icell,jcell} = \sum_{ksite} W_{ksite} \cdot E_{ksite} \quad (3)$$

where the weight applied to each error term is proportional to the inverse squared distance between the center of the grid cell and the monitor location.

$$W_{ksite} = \frac{1/r^2_{icell,jcell,ksite}}{\sum_k 1/r^2_{icell,jcell,k}} \quad (4)$$

A radius of influence is defined (see section 2.2.2) so that if a monitor is located too far from a grid cell, this monitor will not affect the interpolated concentration of the grid cell. This approach works well for an area with multiple monitors located in close proximity. If a monitoring site is located at the center of the grid cell, E_{ksite} = $E_{icell,jcell}$, and $C_{i,icell,jcell}$ = $C_{o,ksite}$. On the other hand, if a grid cell is located far from any monitor, $E_{icell,jcell}$ = 0 and $C_{i,icell,jcell}$ = $C_{s,icell,jcell}$. These are the expected limiting behaviors of the interpolation scheme.

In areas where monitors are sparse, there is a need to temper the influence of individual monitors. Take an example where there is only one monitor in a large area. Within the radius of influence of that monitor, interpolated concentrations are calculated by applying a correction equal to the error at that site. Just outside of the radius of influence, model values are used. An abrupt change in concentration may be introduced in the interpolated field. A progressive weaker influence of the monitor value with distance is desirable.

Within the framework described by Equations (1) through (4), an option is provided to add one or more virtual sites at the radius of influence for each grid cell. These virtual sites are associated with zero errors, and they reduce the weight given to actual sites and the error term used to correct the simulated concentrations (Equation 3). The modified weight is shown below for $n_{virtual}$ virtual sites at the radius of influence ($r_{influence}$).

$$W_{ksite} = \frac{1 / r_{icell,jcell,ksite}^{2}}{n_{virtual} / r_{influence}^{2} + \sum_{k} 1 / r_{icell,jcell,k}^{2}} \qquad (5)$$

The larger the number of virtual sites used in Equation 5, the faster the decrease with distance of the influence of the error at each monitor.

2.2.2 Adjustable parameters

There are two adjustable parameters in the algorithm described in section 2.2.1, (1) the radius of influence of each monitoring site and (2) the number of virtual sites. A radius of influence of 720 km was selected based on the distribution of CASTNET sites to ensure that every location within the continental United States is influenced by at least one CASTNET site. The number of virtual sites is selected based on a visual inspection of the error (Equation 3) and interpolated fields (Equation 2) in a test application with 0, 1, or 4 virtual sites. The use of four virtual sites provides a smooth interpolation field even in areas with few CASTNET monitors (see Figure 2).

3. APPLICATION

Seventy-eight CASTNET sites were operational during 2001, including two pairs of co-located sites. Weekly filter pack measurements and dry deposition fluxes are obtained from http://www.epa.gov/castnet.

CMAQ is applied for the entire year of 2001 (Zhang et al., 2006) over the contiguous United States at a horizontal resolution of 36 km x 36 km. Hourly outputs of surface concentrations and dry deposition fluxes are extracted for the variables listed in Table 1. Average concentrations and deposition fluxes are calculated for periods corresponding to CASTNET measurements.

Figure 2 provides an example of the measured and simulated PM sulfate concentrations during one measurement period. The bottom three panels of Figure 2 show the interpolated concentration fields when zero, one or four virtual sites are used in the interpolation procedure. When 0 or 1 virtual sites are used, the influence of the error of an individual site can dominate over the model gradients in areas with few monitoring sites (e.g., in the southeast part of the modeling domain). When 4 virtual sites are used, the modeled gradients are reproduced in areas far away from monitoring sites, resulting in more realistic interpolation results. These results provide the basis for selecting 4 virtual sites for use with the algorithm described in Section 2.2.

(a)
(b)
(c)
(d)

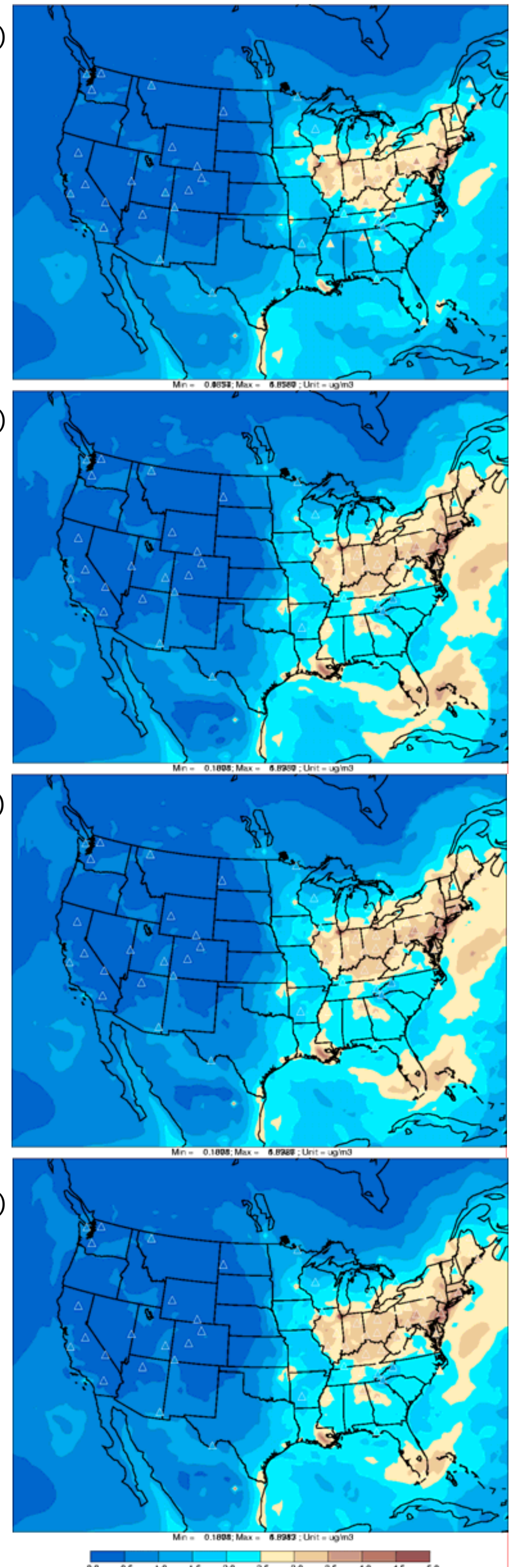

Figure 2. PM sulfate concentration measurements from CASTNET during 2-8 January 2001 and (a) CMAQ simulated concentrations, and interpolated concentrations assuming (b) 0, (c) 1, and (d) 4 virtual sites.

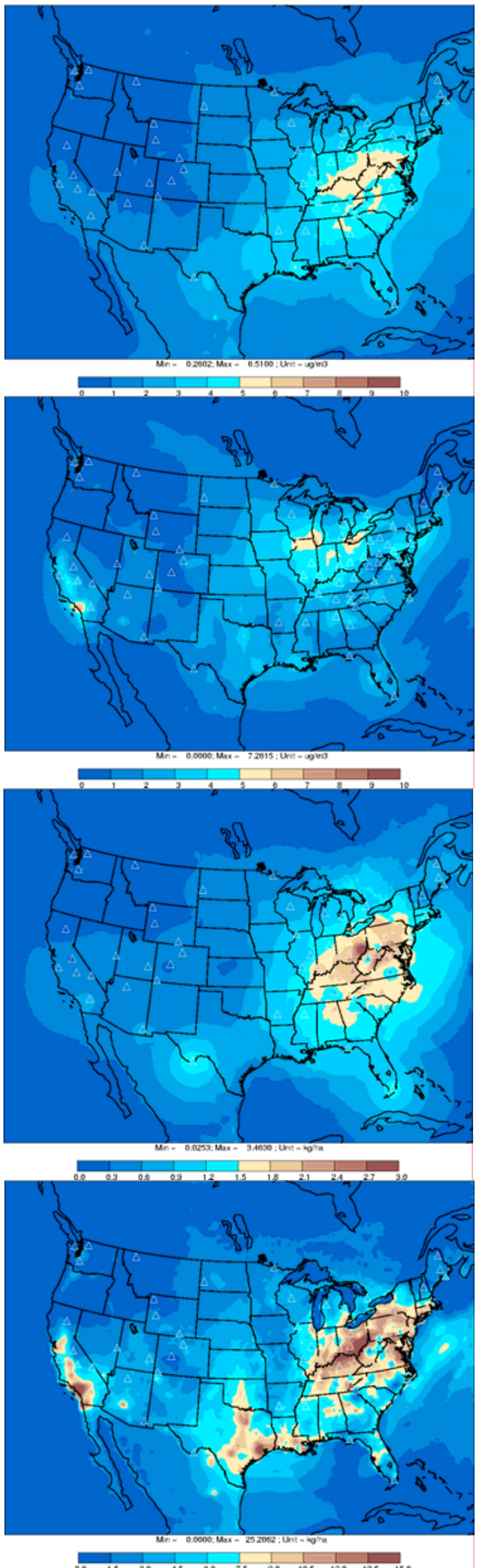

Figure 3. Interpolated 2001 annual concentrations of (a) PM sulfate and (b) total nitrate, and total annual dry deposition fluxes of (c) PM sulfate and (d) total nitrate.

Figure 3 shows the interpolation of annual concentrations of PM sulfate, total nitrate and annual total deposition fluxes of PM sulfate and total nitrate. Only sites with 70% data completeness or better are included in the analysis. The annual concentration is calculated as the average of the available data, and the annual total flux is the sum of the available data, scaled to an annual total when there are missing measurements. These annual concentrations and deposition fluxes are superimposed on the interpolated map using colored symbols.

On an annual average basis, PM sulfate is regionally distributed, with higher concentrations over the eastern United States than the western United States. Over the eastern United States, the highest concentrations over the Ohio River Valley are retained in the interpolated field, and PM sulfate concentrations decrease away from the Ohio River Valley with gradual gradients on the interpolated concentration map.

In Figure 3b, high annual total nitrate concentrations are seen in the lower Midwest and in California. The lower Midwest concentrations are a wintertime phenomenon and are captured by a few sites in Illinois, Indiana, and Ohio. The model filled in the gradients during interpolation. The urban concentrations near Chicago and Los Angeles are features from the model that represent higher urban nitrate during summer. Because of the lack of CASTNET sites in urban areas, these high urban nitrate concentrations will not be apparent if traditional interpolation methods are used. This is a good example of the value that is added by the simulation, which can represent urban emissions, geography, and nitrate photochemistry.

PM sulfate dry deposition fluxes show higher spatial variability than PM sulfate concentrations. Deposition fluxes are subject to variability in deposition velocities and concentrations. Deposition velocity is a function of meteorology and land use, as well as the intrinsic characteristics of particles and gases involved. In the interpolated fields, the influence of individual sites is quite obvious, especially ones that are isolated, e.g., Big Bend National Park in Texas and Everglades National Park in Florida. Model predictions of PM dry deposition fluxes are bias low compared to CASTNET measurements. Errors are large compared to the magnitude of the simulated concentrations, and the interpolation procedure is more strongly influenced by the interpolated error compared to the CMAQ concentration gradients.

PM sulfate dry deposition fluxes show less spatial variability than total nitrate fluxes. The predicted total nitrate dry deposition fluxes are not biased low like PM sulfate fluxes, because dry deposition of HNO_3 dominates the total nitrate fluxes and dry deposition fluxes for gases do not appear to be underpredicted. The interpolated fields show a lot of features in areas without CASTNET measurements, driven by model predictions. In the lower Midwest, very strong spatial gradients are present in the interpolated fields, originating from the variability of the CASTNET estimates of total nitrate dry deposition in the area.

4. EVALUATION

To evaluate the performance of the interpolation scheme, it is desirable to compare the interpolated concentrations against other measurements of regional concentrations or dry deposition fluxes. IMPROVE represents a comprehensive monitoring network for rural concentrations of PM species.

IMPROVE has no or limited measurements of SO_2 and PM ammonium during 2001. For nitrate, IMPROVE monitors use a denuder to remove HNO_3; therefore, they measure PM nitrate, not subject to PM/gas conversion that affects CASTNET measurements. Due to differences between IMPROVE and CASTNET methodologies, PM sulfate is the only species where IMPROVE and CASTNET measurements are comparable.

At IMPROVE monitors, 24-hour samples are obtained every third day, representing a sample completeness of 29 or 43% for a weekly period, and 33% on average. An evaluation of weekly interpolated fields based on individual CASTNET periods using IMPROVE data that span 43% or less of the comparison period is subject to significant errors due to limited sampling. IMPROVE measurements are frequently used to establish seasonal and annual average concentrations. The random error in the estimate of a seasonal or annual average is expected to be smaller than that in the estimate of a weekly average. Therefore, seasonal and annual PM sulfate concentrations are compared between the CASTNET-interpolated fields and the IMPROVE measurements. Standard model performance evaluation metrics are used, including mean bias, mean normalized bias, mean error, mean normalized error, and the coefficient of determination (r^2).

Table 2 shows the performance statistics for annual PM sulfate concentrations for all IMPROVE sites, and for IMPROVE sites in the eastern and western United States.

Table 2. Performance statistics for annual average sulfate concentrations interpolated from CASTNET measurements using CMAQ modeling results compared against IMPROVE annual average data.

Metric	**All**	**Eastern U.S.**	**Western U.S.**
Number of IMPROVE sites (1)	93	34	59
Mean IMPROVE observation (1)	1.63	3.13	0.77
Mean interpolated CASTNET value (1)	1.67	3.16	0.81
Mean bias (2)	0.03	0.03	0.04
Normalized bias (3)	0.07	0.02	0.10
Mean error (2)	0.18	0.31	0.11
Normalized error (3)	0.14	0.10	0.15
r^2	0.96	0.87	0.84

(1) threshold concentration 0.01 $\mu g/m^3$ used.
(2) units = $\mu g/m^3$; (3) unitless

The interpolation scheme performs extremely well for annual PM sulfate for all IMPROVE sites. The small overprediction can be attributed to systematic differences in the particle collection methodology at IMPROVE (size restricted inlets) and CASTNET monitors (open face filter packs). There is no systematic error (low bias) that could indicate weaknesses in the interpolation methodology. Statistics for pairs of observation vs. interpolated values are 7% for normalized bias and 14% for normalized error. This level of performance is noteworthy due to the large number of observed samples with very low concentrations. Generally, errors associated with small concentrations lead to poor performance for the paired statistics. The interpolation scheme can explain 96% of the variability observed at IMPROVE sites.

When comparing eastern and western United States, the mean bias in each region is comparable. Due to lower concentrations in the western United States, the normalized bias is higher than in the eastern United States. Mean error in the eastern United States is higher than in the western United States, but normalized error is lower due to greater concentrations in the eastern United States. The lack of improvement in the r^2 values in the spatially limited data sets compared to the entire United States indicates that there is no underlying difference in the performance of the interpolation procedure for PM sulfate in the subregions.

Figure 4 shows the IMPROVE annual concentrations superimposed on the CMAQ interpolated fields based on CASTNET measurements. The most significant underprediction in the interpolated field occurs at Breton, LA. The annual average observation at the Breton IMPROVE site is 3.4 $\mu g/m^3$, whereas a concentration of 2.5 $\mu g/m^3$ is estimated based on interpolating CASTNET measurements. A plume of high sulfate concentrations is predicted by CMAQ in the New Orleans/Baton Rouge area, due to emissions in the vicinity. Model predictions are tempered by low CASTNET measurements in Caddo Valley, AR. Additional monitors in southeastern Texas and in Louisiana may improve the interpolated concentration estimates in southern Louisiana. However, PM sulfate concentrations affecting Breton may also result from the impact of off-shore sources that are underrepresented in the emissions inputs to models. Coastal gradients in the interpolated PM sulfate field are controlled by CMAQ predictions because there are no observations offshore. Therefore, model weaknesses in the treatment of mixing and deposition near coastal areas will translate into larger errors in the interpolated concentrations at coastal sites.

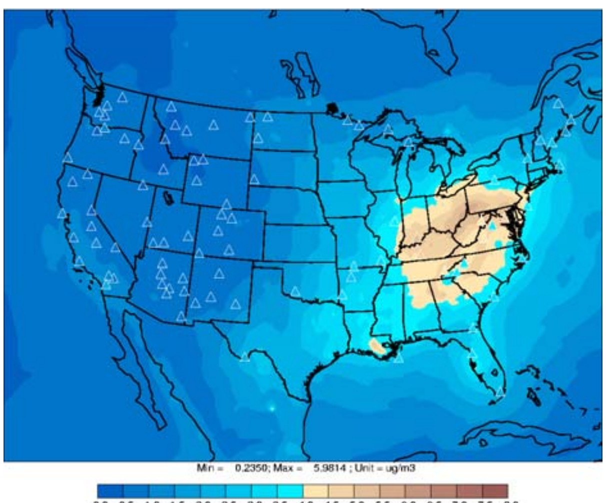

Figure 4. 2001 annual PM sulfate concentrations measured by IMPROVE sites (triangles) superimposed on CMAQ-interpolated CASTNET concentrations.

Performance statistics for seasonal averaged PM sulfate concentrations are also calculated. Despite very different sulfate concentrations, the performance of the interpolation methodology is satisfactory for all seasons.

5. CONCLUSIONS

CMAQ is used to interpolate CASTNET observations of ambient concentrations and estimate dry deposition fluxes. This interpolation scheme produces more realistic gradients where observations are sparse. In addition, where there are sharp gradients due to emissions, the model can add valuable information into the interpolated fields. In isolated areas where pollution may be important but CASTNET coverage may be insufficient (e.g., Los Angeles Basin for total nitrate), the interpolated maps are more realistic than a simple distance-weighting because the model provides essential information based on geography, emissions, meteorology, and chemistry, for the interpolation process.

The interpolation technique is evaluated using sulfate data from IMPROVE. The comparison reveals no bias in the average sulfate concentrations, and no seasonality in the performance. Due to higher density of CASTNET sites in the eastern United States, the interpolation performs better there than in the western United States. Interpolating CASTNET data using the CMAQ model provides reliable estimates (14% error at specific locations) of annual or seasonal concentrations for non-urban areas.

The interpolation scheme is subject to potential weaknesses of the models, e.g., biases in the dry deposition of particles, representation of processes in coastal areas. Based on CASTNET data and a 36-km resolution CMAQ simulation, the interpolation scheme provides accurate estimates in rural areas (e.g., IMPROVE sites), but should be applied with care in urban areas.

6. ACKNOWLEDGEMENTS

This work was funded by the U.S. Environmental Protection Agency (Contract 68-W-03-033, Assignment 2-15, 3-15). Bryan Bloomer, Manager, provided insightful discussion and feedback throughout the project.

7. REFERENCES

Zhang, Y., Vijayaraghavan, K., Huang, J-P., Jacobson, M.Z., Probing into Regional O_3 and PM Pollution: A 1-year CMAQ Simulation and Process Analysis over the United States, *Proc. Eighth Conference on Atmospheric Chemistry*, Amer. Meteor. Soc., Atlanta, GA, February 2006.