Estimating Particle Sulfate Concentrations Using MISR Retrieved Aerosol Properties

Yang Liu, Bret A. Schichtel, and Petros Koutrakis

***Abstract*—Understanding the spatial distribution of fine particle sulfate (SO_4^{2-}) concentrations is important for optimizing emission control strategies and assessing the population health impact due to exposure to SO_4^{2-}. Aerosol remote sensors aboard polar orbit satellites can help expand the sparse ground monitoring networks into regions currently not covered. We developed a generalized additive model (GAM) using MISR fractional aerosol optical depths (AODs) scaled by GEOS-Chem aerosol profiles to predict ground-level SO_4^{2-} concentrations. This advanced spatial statistical model was compared with alternative models to evaluate the effectiveness of including simulated aerosol vertical profiles and adopting an advanced statistical model structure in terms of improving the AOD- SO_4^{2-} association. The GAM is able to explain 70% of the variability in SO_4^{2-} concentrations measured at the surface, and the predicted spatial surface of annual average SO_4^{2-} concentrations are consistent with interpolated contours from ground measurements. Comparisons with alternative models demonstrate significant advantages of using model-scaled lower-air fractional AODs instead of their corresponding column values. The nonlinear association between SO_4^{2-} concentrations and fractional AODs makes the GAM a more suitable model structure than conventional linear regressions.**

***Index Terms*—Fractional AOD, GAM, MISR, particle sulfate, spatial pattern.**

I. Introduction

FINE particulate matter ($PM_{2.5}$) is a complex mixture of airborne particles primarily composed of sulfate, nitrates, ammonium, elemental carbon, organic compounds, and mineral dust. Epidemiological studies worldwide have established a robust association between exposure to $PM_{2.5}$ and various adverse health outcomes such as respiratory diseases, chronic obstructive pulmonary disease, cardiovascular diseases, and premature death [1], [2]. Particulate matter also has important optical properties affecting visibility and the global radiation balance. Measuring the chemical composition of $PM_{2.5}$ is crucial in determining the relative toxicity of different $PM_{2.5}$ constituents as well as their optical properties, hence developing effective emission control policies [3]. Ground-based monitoring networks such as the $PM_{2.5}$ chemical speciation network (CSN) [4] operated by the U.S. Environmental Protection Agency (EPA) and the Interagency Monitoring of Protected Visual Environments (IMPROVE) network [5] have been the basis for studying the spatial and temporal characteristics of major $PM_{2.5}$ components such as sulfate (SO_4^{2-}). Although these ground measurements of particle speciation are considered accurate and used by regulatory agencies such as EPA and National Park Service for compliance monitoring, they are limited both spatially (mostly in urban areas or national parks) and temporally (every three days or every six days) due to high operational cost.

Polar orbiting satellites can cover nearly the entire globe in the matter of a few days. The rapidly advancing satellite aerosol remote sensing technology makes them a potential source of information on the transport and spatial patterns of fine particles. The Multiangle Imaging SpectroRadiometer (MISR) was launched in December 1999 aboard NASA's Earth Observing System (EOS) Terra satellite. MISR aerosol retrieval over land is described in detail elsewhere [6] and given briefly here. In the operational MISR aerosol retrieval, it is assumed that atmospheric aerosols are horizontally homogeneous within a 17.6 km x 17.6 km region. It uses the presence of spatial contrasts within the 17.6-km retrieval region to derive empirical orthogonal functions (EOF) to represent the region-averaged surface-leaving light reflection. MISR defined a set of aerosol mixtures to represent aerosol types globally for computational efficiency. These mixtures are combinations of several aerosol components defined by a size distribution, shape, refraction index, and scale height. MISR observed top-of-atmosphere (TOA) radiances are compared to each of the TOA radiances calculated based on the predefined aerosol mixtures. A set of statistical tests was developed to determine which aerosol mixtures best fit the observations. These tests explicitly include instrument measurement uncertainty in the retrieval results [7]. Previous research that linked various satellite-retrieved aerosol optical depths (AODs) to ground level particle concentrations has shown that particle vertical distribution and change of composition must be considered in order to establish a robust relationship between AOD and ground-level particle abundance [8]–[10].

As the MISR algorithm identifies a set of aerosol components representing the major ambient particle species, these components contain valuable information about particle emission sources and chemistry [11], [12]. We have developed a

Manuscript received February 28, 2009; revised June 11, 2009. First published September 15, 2009; current version published October 28, 2009. This work was supported in part by the Harvard-EPA Center on Particle Health Effects (R-827353 and R-832416) and in part by the NASA Jet Propulsion Laboratory under a subcontract with Emory University (1363692).

Y. Liu is with the Rollins School of Public Health, Emory University, Atlanta, GA 30322 USA.

B. A. Schichtel is with the National Park Service, Colorado State University, Fort Collins, CO 80523 USA.

P. Koutrakis is with the School of Public Health, Harvard University, Boston, MA 02215 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2009.2030153

1939-1404/$26.00 © 2009 IEEE

TABLE I
AEROSOL COMPONENTS ASSUMED IN MISR VERSION 17 RETRIEVALS

No.	Name	Shape	d_{min}[1] (μm)	d_{max}[1] (μm)	d_c[2] (μm)	ω[3]	Refractive index	H[4] (km)
1	spherical_nonabsorbing_0.06	spherical	0.002	0.8	0.06	1	1.45+0*i*	2
2	spherical_nonabsorbing_0.12	spherical	0.002	1.5	0.12	1	1.45+0*i*	2
3	spherical_nonabsorbing_0.26	spherical	0.02	3.0	0.24	1	1.45+0*i*	2
6	spherical_nonabsorbing_2.80	spherical	0.2	100	1.00	1	1.45+0*i*	2
8	spherical_absorbing_0.12_ssa_green_0.9	spherical	0.002	1.5	0.12	0.90	1.45+0.015*i*	2
14	spherical_absorbing_0.12_ssa_green_0.8	spherical	0.002	1.5	0.12	0.80	1.45+0.033*i*	2
19	grains_mode1_h1 (dust)	Grains Model	0.2	2.0	1.00	0.98	1.51+0.004*i*	10
21	spheroidal_mode2_h1 (dust)	spheroids	0.2	12.0	2.00	0.90	1.51+0.004*i*	10

1 –d_{min} and d_{max} are the minimum and maximum particle diameters for a given aerosol component.

2 –d_c is the characteristic diameter of the lognormal size distribution.

3 – ω is single scattering albedo at 558 nm.

4- H is layer scale height.

three-step approach in which MISR AOD data are combined with simulated aerosol vertical profiles to estimate ground-level concentrations of $PM_{2.5}$ constituents such as SO_4^{2-} [13], [14]. The first step is to extract aerosol speciation information contained in the MISR fractional AOD values. The fractional AOD of a MISR aerosol component is the average contribution of this component to total AOD in all successful aerosol three-component mixtures. In practice, this primarily distinguishes spherical from nonspherical and bright from light-absorbing aerosol species in the MISR data. Since AOD is a column aggregate, the second step is to strengthen the correlation between AOD and near-surface SO_4^{2-} concentrations by scaling the column fractional AODs with GEOS-Chem simulated aerosol vertical profiles. Finally, a multivariate linear regression model was developed to estimate SO_4^{2-} concentrations with the scaled fractional AODs as major predictors. The objective of this analysis is to evaluate the effectiveness of each step in terms of their effectiveness in improving the AOD-$PM_{2.5}$ association. We wish to provide guidance to the future development of this technique. We adopt an advanced model structure flexible enough to account for any nonlinear effects of the fractional AODs and spatial and temporal biases. We present the results of estimating the spatial trend of SO_4^{2-} concentrations in the continental United States using the MISR fractional AODs. We first introduce the ground measurements, model simulations, and satellite data in this analysis. We then describe the modeling approach to estimate ground level SO_4^{2-} concentrations. Model fitting results are compared to the results from two alternative models to evaluate the effectiveness of simulated aerosol vertical profiles and the nonparametric model structure. Finally, we use the final GAM to predict annual mean SO_4^{2-} concentrations on a 50-km resolution grid in the continental U.S., and evaluate the spatial patterns of predicted SO_4^{2-} levels with interpolated ground measurements.

II. DATA AND METHODS

A. MISR Level 2 Aerosol Data

This analysis used the Version 17 MISR Level 2 aerosol data. In this dataset, there are 74 different aerosol mixtures that are constructed from up to three of the eight predefined aerosol components (i.e., components 1, 2, 3, 6, 8, 14, 19, and 21) (Table I). MISR aerosol components are designed to be approximately orthogonal to ensure minimal overlapping in the ambient aerosol species they represent. A detailed discussion of MISR data structure, the aerosol components used to construct the aerosol models and the percentage contribution of each component to total AOD is given elsewhere [13]. The MISR aerosol data covering the continental U.S. in 2001 and 2005 were obtained from the NASA Langley Research Center (LARC) Atmospheric Sciences Data Center (http://edg.larc.nasa.gov/~imswww/imswelcome/index.html). We selected these two years because MISR had consistent data versions at the time of this analysis. Previous research has indicated that MISR retrieved aerosol microphysical information is more accurate at higher AOD values. Although setting a lower bound of 0.15 or higher for total AOD would allow only higher quality MISR aerosol microphysics data to be included in the modeling process, it will substantially bias the predicted SO_4^{2-} concentrations. In addition, to ensure a sufficiently large sample size to fit the statistical models, we excluded all AOD values less than 0.05 in the following modeling analysis.

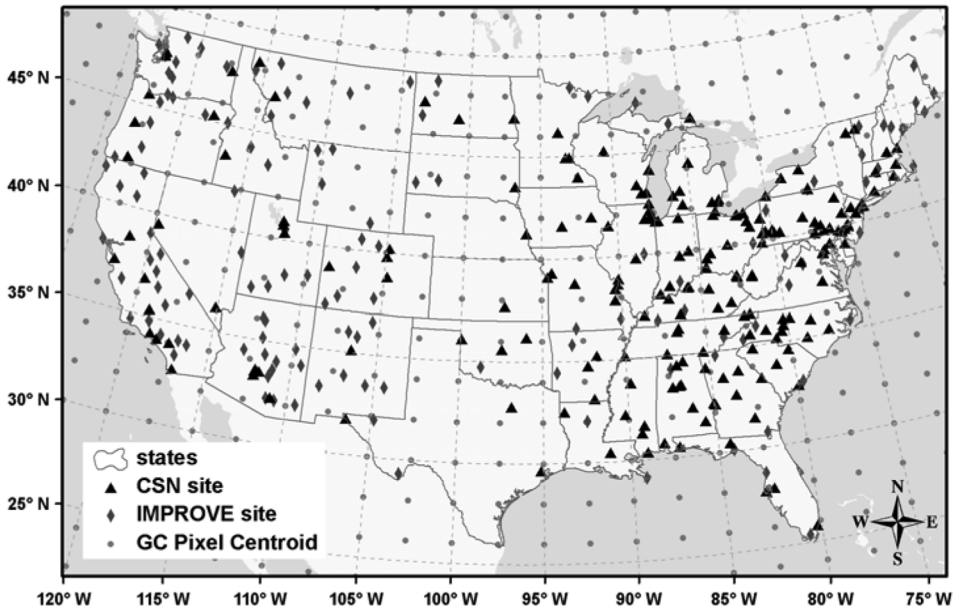


Fig. 1. Geographical locations of EPA CSN sites (triangles) and the IMPROVE sites (diamonds) in the continental U.S. GEOS-Chem $2^\circ \times 2.5^\circ$ grid centroids are shown as grey dots.

B. $\mathrm{PM}_{2.5}$ *Speciation Data*

Daily average particle sulfate concentrations collected by EPA's CSN for 2001 and 2005 were obtained from EPA's Air Quality System (AQS) (http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm). IMPROVE measurements of daily sulfate concentrations during the same periods were obtained from the Visibility Information Exchange Web System (VIEWS) (http://vista.cira.colostate.edu/views/). Fig. 1 shows the spatial distribution of approximately 210 CSN sites mostly in the east and 140 IMPROVE sites. In developing our statistical models, we selected all daily SO_4^{2-} concentration measurements in the AQS except those identified as low quality by the Quality Assurance qualifiers.

C. GEOS-Chem Aerosol Simulation Data

The GEOS-Chem model is a global 3-D chemistry and transport model (CTM) driven by assimilated meteorological data from the Goddard Earth Observing System (GEOS-4) at the NASA Global Modeling and Assimilation Office (GMAO) including winds, convective mass fluxes, boundary layer heights, temperature, clouds, precipitation, and surface properties. The aerosol and gaseous simulations are coupled through formation of sulfate and nitrate, partitioning of total inorganic nitrate, heterogeneous chemistry on aerosols, and aerosol effects on photolysis rates [15]. GEOS-Chem aerosol simulation (version 7-03-06 in the current analysis) produces mass loadings of particle species such as nitrate $\left(\mathrm{NO}_3^-\right)$, sulfate $\left(\mathrm{SO}_4^{2-}\right)$, ammonium $\left(\mathrm{NH}_4^+\right)$, organic (OC) and elemental carbon (EC), and sea salt, and dust at 3-h intervals, $2^\circ \times 2.5^\circ$ spatial resolution, and 20 vertical layers in the troposphere [16]. Although GEOS-Chem aerosol vertical profiles have not been validated over land, Park *et al.* showed using an earlier version that the model can reproduce with no significant bias the observed vertical profiles of sulfate aerosols from the TRACE-P aircraft campaign over the Pacific Ocean [17]. Simulation results for 2001 and 2005 are interpolated to 10–12 a.m. local time values to match the MISR sampling time window. GEOS-Chem simulated aerosol loading profiles are used as scaling factors to calculate lower air MISR AOD, as discussed later. It should be noted that GEOS-Chem (version 8-01-01) is now driven by a newer version of meteorological data (GEOS-5). However, the processes most relevant to the vertical distribution of aerosols, including, boundary layer mixing and aerosol microphysics were not changed (Jintai Lin, personal communication). Therefore, using an older version of GEOS-Chem simulation results is unlikely to have any significant impacts on the major findings of this analysis.

D. Model Development

As mentioned above, a three-step approach has been developed to strengthen and stabilize the relationship between $\mathrm{PM}_{2.5}$ concentrations and MISR AOD, whose original form is discussed in detail elsewhere [13], [14]. Since each of the eight MISR aerosol components can have different contributions to ground-level sulfate concentrations, the first step involves dissembling total MISR column AOD into species-related fractions using the AOD value associated with, and the three aerosol components defined by each of the 74 aerosol mixtures (1). If present, the fractional AOD of a MISR aerosol component is defined as the average contribution of this component to total AOD. For example, if MISR does not observe the presence of any dust particles in a 17.6 km pixel, the fractional AODs for dust components (i.e., 19 and 21) will be zero. By definition, the sum of all the significant fractional AODs is equal to the total column AOD. See (1), shown at the bottom of the next page.

The second step is calculating the lower-air fractional AOD for each component by scaling them with GEOS-Chem simulated aerosol vertical profiles. We assume that the lower-air (i.e., < 1 km above ground) fraction of AOD is more strongly

correlated to surface observations due to active vertical mixing in the boundary layer during daytime. This assumption is tested in one of the alternative models. Because the eight MISR aerosol components are not defined by chemical composition, it is not feasible to match each GEOS-Chem species to a MISR aerosol component. Therefore, we combine GEOS-Chem aerosol mass loadings of SO_4^{2-}, NO_3^-, NH_4^+, and sea salt at each vertical layer to represent the contribution of spherical nonabsorbing particles over land, and use it to scale MISR components 1, 2, 3, and 6. We combine the mass loadings of OC and EC to represent the contribution of spherical absorbing particles, and use it to scale MISR components 8 and 14. We use the mass loadings of dust to represent the contribution of nonspherical particles, and use it to scale MISR components 19 and 21 (2). *GC lower-air aerosol loading* in (2) refers to the sum of GEOS-Chem species mass loadings within one kilometer from the surface, and *GC column aerosol loading* refers to the sum of GEOS-Chem species mass loadings in all 20 layers. We did not correct the particle growth effect at lower troposphere using GEOS-Chem relative humidity fields as the nonlinear spatial and temporal terms included in the GAM ((3) can partially account for it. In addition, applying an individual humidity correction factors to combined species such as EC and OC can introduce uncertainty into the models

$$\begin{aligned}&\text{MISR lower-air fractionalAOD}\\&\quad=\frac{\text{GC lower-air aerosol loading}}{\text{GC column aerosol loading}}\\&\quad\times\text{MISR fractional AOD.}\end{aligned}\tag{2}$$

The last step is to develop a statistical model to link lower-air fractional AODs with sulfate concentrations. Because eight MISR aerosol components are designed to be approximately orthogonal, we can now include them as individual predictors in our statistical model. We develop a generalized additive model (GAM) with lower-air fractional AOD values as the main predictors of SO_4^{2-} concentration (3). A GAM expands the capability of traditional linear regressions by allowing some or all predictor variables to have nonlinear relationships with the dependant variable by using semi-parametric spline smoothers [18]. Like linear regression models, a GAM assumes that the associations of predictor variables with the dependant variable are additive

$$\begin{aligned}\left[SO_4^{2-}\right]&=f_{x,y}(x,y)+f_t(t)\\&\quad+\sum_{i=1}^{8}f_i(\text{MISR lower-air fractional AOD}_\text{i}).\end{aligned}\tag{3}$$

On the left-hand side of (3) is daily SO_4^{2-} concentrations measured at CSN or IMPROVE sites. On the right hand side, $f_{x,y}(x,y)$ is a 2-D smooth spatial bias (analogous to model intercept in a linear regression) reflecting the potential impact of site locations on the associations between SO_4^{2-} and fractional AODs. The x and y are geographic coordinates of CSN and IMPROVE sites under U.S. contiguous Albers projection. $f_t(t)$ is a nonlinear regression term varying smoothly in time (day of year) to account for systematic seasonal variations. Eight MISR lower-air fractional AODs corresponding to the eight MISR aerosol components are included as predictor variables represented by Σf_i (*MISR lower-air fractional* AOD_i). The fitted nonlinear smooth function for each MISR aerosol component is analogous to the regression coefficient in a linear regression model. Our GAM still assumes that the associations between MISR aerosol components and SO_4^{2-} concentrations are constant across the entire modeling domain. Fitting spatially varying smooth functions for each site or clusters of sites is not feasible due to limited data available. Nonetheless, our GAM represents a significant improvement over a traditional linear model as the nonlinear bias terms and smooth functions are difficult to represent in a linear model. The scaled lower-air fractional AODs were tested for collinearity in a linear regression model setting and no significant collinearity was found (results not shown). This is consistent with the fact that the eight MISR aerosol components used to form various aerosol mixtures are designed to be approximately orthogonal to each other in order to optimally represent the aerosol types in ambient atmosphere as well as possible ([9] and Ralph Kahn, personal communication). Therefore, they are all included in the GAM as independent predictor variables. Our preliminary modeling analysis indicated little discrepancy due to different measurement protocols between CSN and IMPROVE so no network indicator was included in the model as a variable. This agrees with a previous comparison of IMPROVE and CASTNet sulfate measurements [5] and a comparison of sulfate concentrations from collocated IMPROVE and CSN samplers [19]. All the lower-air fractional AODs are assumed to have nonlinear associations with SO_4^{2-} concentration although they can take linear forms if linear relationships fit the model well based on standard GAM fitting criteria (generalized cross-validation scores, t statistics, and model adjusted R^2 estimates).

Liu *et al.* [13], [14] already demonstrated in a case study that the fractional AODs approach performs substantially better in estimating ground-level $PM_{2.5}$ concentrations compared to using total AOD as the sole predictor. In this analysis, we develop two alternative models after the initial step of deriving fractional AODs, to evaluate additional impacts of including GEOS-Chem aerosol profiles and using the nonlinear GAM structure. The terminology of the alternative GAM (3a) is identical to (3) except that lower-air fractional AODs are replaced

$$\text{Fractional AOD}_{\text{i(i=1,8)}}=\frac{\sum_{\text{j=1}}^{74}\text{AOD}_{\text{mixture_j}}\times\text{Fraction}_{\text{component i in mixture_j}}}{\text{No. of Successful Mixtures}}\tag{1}$$

TABLE II
COMPARISON OF MODEL FITTING RESULTS FOR FULL GAM AND ALTERNATIVE MODELS

Sample size	Whole continental U.S., n = 1387	
Full model (Eq. 3)		
Significant predictors*	AOD_1, AOD_2, AOD_3, AOD_8, AOD_{19}, AOD_{21}, DOY, (x,y)	
Model adj. R^2	0.70(0.62)#	
Sample size	Eastern U.S, n = 808	Western U.S., n = 579
Full model (Eq. 3)		
Significant predictors*	AOD_1, AOD_2, AOD_3, AOD_8, AOD_{19}, AOD_{21}, DOY, (x,y)	AOD_1, AOD_2, AOD_3, AOD_8, AOD_{19}, AOD_{21}, DOY, (x,y)
Model adj. R^2	0.68 (0.62)	0.63 (0.43)
Alternative GAM (Eq. 3a)		
Significant predictors**	AOD_{1c}, AOD_{2c}, AOD_{3c}, AOD_{8c}, AOD_{19c}, AOD_{21c}, DOY, (x,y)	AOD_{1c}, AOD_{2c}, AOD_{3c}, AOD_{19c}, AOD_{21c}, DOY, (x,y)
Model adj. R^2	0.64 (0.54)	0.58 (0.28)
Alternative GLM (Eq. 3b)		
Significant predictors*	AOD_1, AOD_2, AOD_3, AOD_8, AOD_{19}, AOD_{21}, season	AOD_1, AOD_2, AOD_3, AOD_8, AOD_{19}, season
Model adj. R^2	0.47	0.40

* AOD_1 is lower-air fractional AOD of MISR component 1, and AOD_2 for lower-air MISR component 2, and so on. DOY is day of year. (x, y) is the smooth spatial bias.
# The adjusted model R^2 values in the brackets are calculated with DOY and (x,y) excluded.
** AOD_{1c} is column fractional AOD of MISR component 1, and AOD_{2c} for column MISR component 2, and so on.

with column fractional AODs [i.e., (2) is not implemented]. The alternative general linear model (GLM) uses scaled lower-air fractional AODs as predictors, but takes a more traditional linear regression model structure (3b) as in Liu *et al.* [14]. μ is model intercept, β_i is the regression coefficient for lower-air fractional AOD of aerosol component I, and β_s is the group of regression coefficients for a four-level categorical variable for season

$$[SO_4^{2-}] = f_{x,y}(x,y) + f_t(t) + \sum_{i=1}^{8} f_i(\text{MISR column fractional AOD}_\text{i}) \tag{3a}$$

$$[SO_4^{2-}] = \mu + \sum_{i=1}^{8} \beta_i \times \text{MISR lower-air fractional AOD}_\text{i} + \beta_s \times \text{season}. \tag{3b}$$

III. RESULTS AND DISCUSSION

A. Fitting of the Full Model

After spatial and temporal aligning, the final model fitting dataset has 1,387 site-days. As shown in Table II, the adjusted model R^2 values indicate that after adjusting for the number of variables included in the model, the full GAM explains 70% of the variability in SO_4^{2-} concentrations. All the predictor variables in Table II are significant at $\alpha = 0.05$ level. Both the spatial bias and temporal term are significant, indicating the presence of residual spatial and temporal variability unexplained by fractional AODs. Without these two terms, the adjusted model R^2 is 0.62, approximately a 10% decrease. The unexplained variability in SO_4^{2-} concentrations can be attributed to a few factors. First, as a polar orbiting instrument, MISR measures atmospheric aerosols at a given location around 10:30 a.m. local time. This snapshot may not represent 24-hr average conditions precisely. Second, MISR observes ambient particles whereas CSN and IMPROVE sites measure SO_4^{2-} concentrations as dry mass. Equation (3) does not fully account for spatial and temporal variability in relative humidity and the associated sulfate particle growth [20]. Third, although each MISR aerosol component has a unique combination of size distribution, particle shape and refractive index and the components are designed to be approximately orthogonal, the significant overlap in the size distributions and incomplete orthogonality may limit their capability of fully distinguishing between significant shifts in SO_4^{2-} size distributions in space and time. Fourth, the $2° \times 2.5°$ resolution GEOS-Chem aerosol vertical profiles are only average representations of the actual aerosol vertical profiles for the 17.6 km resolution MISR fractional AODs. Finally, comparison of ground-level point measurements with areal satellite observations may introduce additional model uncertainty because of the spatial averaging of satellite observations. Fig. 2 shows the scatter plot of fitted SO_4^{2-} concentrations versus EPA CSN and IMPROVE observations, and the SO_4^{2-} model underestimates by approximately 19%.

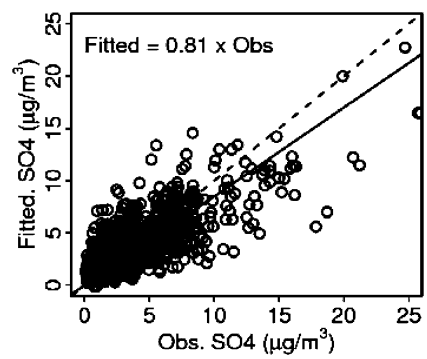


Fig. 2. Scatterplots of predicted daily SO_4^{2-} concentrations versus CSN and IMPROVE observations. The solid line represents simple linear regression results with intercept excluded. The 1:1 line is displayed as a dashed line for reference.

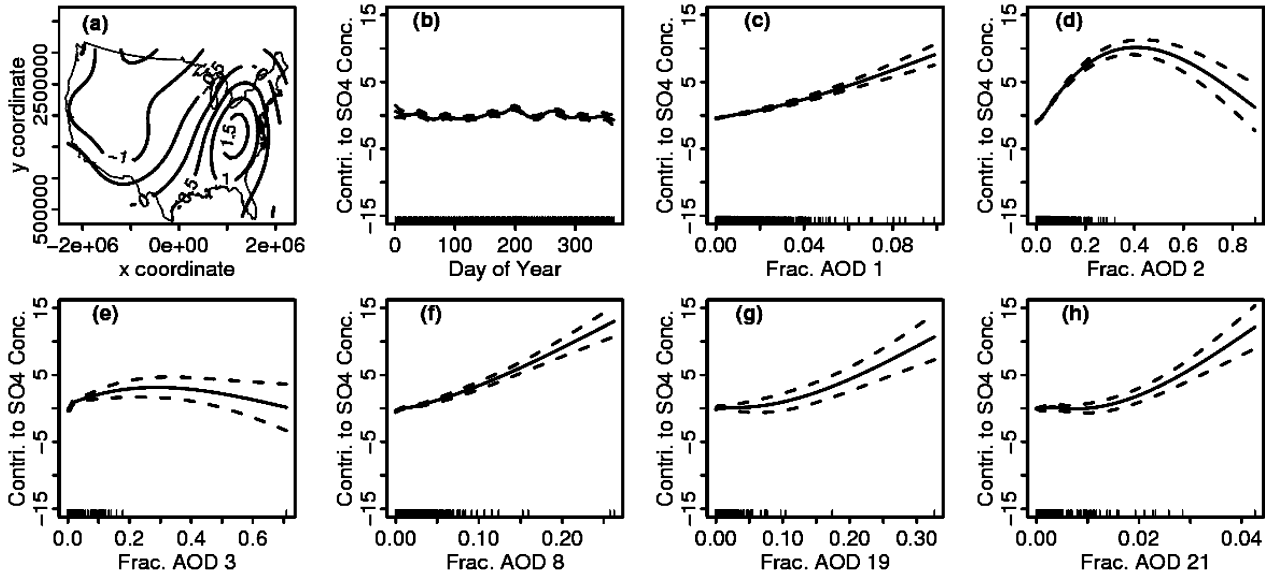


Fig. 3. Fitted smooth regression terms in the full GAM for the spatial bias term (plot a), day of year (plot b), and scaled lower-air fractional AODs for MISR aerosol component 1 (plot c), 2 (plot d), 3 (plot e), 8 (plot f), 19 (plot g), and 21 (plot h). The two dashed lines in plot b through h are confidence intervals of the fitted smooth curves. The smooth terms have units of μg/m^3.

Fig. 3 shows how fitted SO_4^{2-} concentrations vary with each significant variable in the full GAM while holding other variables constant. The contours in plot (a) represent the spatial bias in fitted SO_4^{2-} concentrations as compared to the national average in our modeling domain. Ranging from 1.5 μg/m^3 in DC-Maryland-Virginia region to -1 μg/m^3 in northwest, it is small compared with the fractional AODs [plots (c) through (h)], which can be as high as 15 μg/m^3 for fitted SO_4^{2-} concentrations. The temporal bias term also has a relatively small impact on fitted SO_4^{2-} concentrations [plot (b)]. All scaled MISR lower-air fractional AODs are significant predictors except for components 6 and 14. Component 14 represents highly absorbing particles such as EC and some OC species so it is not surprising that this component is not a significant predictor of bright nonabsorbing sulfate particles. All aerosol components contribute to SO_4^{2-} concentrations positively, i.e., higher fractional AODs predict higher SO_4^{2-} concentrations. Because component 19 is present in only 15% of the data records and is correlated with component 21, its significance is not necessarily robust. Nonetheless, the significance of dust components (19 and 21) could be because dust particles can be coated with SO_4^{2-} during transport passing polluted industrial areas [21]. When included in the model, the positive fractional AOD of component 6 (i.e., the x-axis) yields a decrease in predicted SO_4^{2-} concentrations (i.e., the smooth curve is negative on the y-axis), which is not physically interpretable. As a result, it is excluded from the final model.

It has been reported that particle scattering efficiencies observed at multiple IMPROVE sites vary directly with mass concentration [22], [23]. The authors of these studies speculated that this is caused by the shift of particle size distributions to larger particles due to particle aging and humidification related to in-cloud processing during transport. This means that the AOD—SO_4^{2-} relationship will vary spatially depending on the distance between the monitoring site and SO_2 emission sources as well as the meteorology along the transport pathway. This is reflected in our statistical model as the nonlinear associations between SO_4^{2-} concentrations and scaled MISR fractional AODs. In addition, it might also be related to the fact that these components are not defined exactly to match common chemical species in $PM_{2.5}$.The nonlinear association could also rise from decreased MISR retrieval accuracy and slight overestimation at lower AOD values [24]. Unfortunately, it is not possible to explain the shape of each individual smooth function given the empirical nature of our GAM.

B. Fitting Results of the Alternative Models

The two alternative models have slightly different sets of significant predictors in the east and the west; hence, each model is fitted in the two regions separately (808 site-days in the east, and 579 site-days in the west). For the purpose of comparing model performance, the full GAM is also fitted in each region separately. The significant predictors in the alternative GAM (3a) are identical to the full GAM in the east. The adjusted model R^2 values with and without the spatial bias and temporal term (0.64 and 0.54, respectively) indicate a slightly lower model performance than the full GAM (3) in estimating SO_4^{2-} concentrations. More importantly, in the west the adjusted model R^2 value decreases by 50% without the spatial bias and the temporal term (from 0.58 to 0.28). The greater importance of these terms in (3a) signals column fractional AODs as weaker predictors of SO_4^{2-} concentrations than their scaled lower-air counterparts in the full GAM. This is not surprising, as scaled fractional AODs are less influenced by possible long-range pollution transport events in the upper troposphere, which is often unrelated to ground-level SO_4^{2-} concentrations. In addition, simulated aerosol profiles also address the issue of vertical mixing directly linked to ground-level pollutant concentrations.

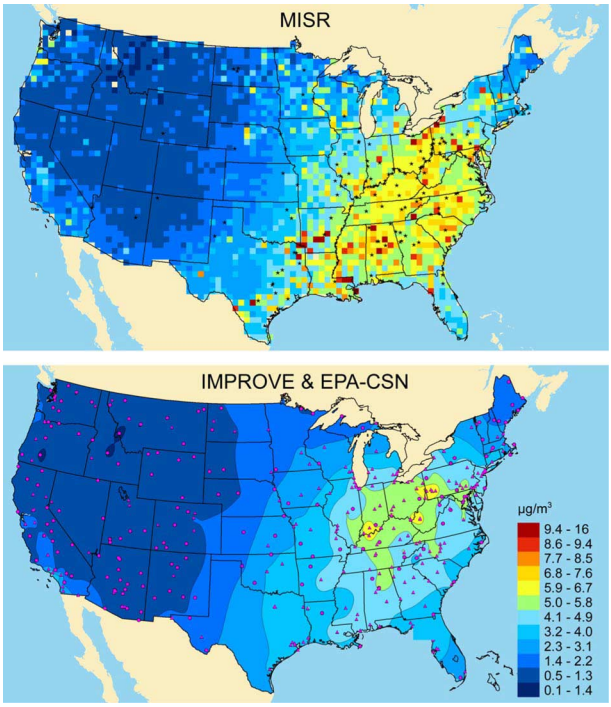


Fig. 4. 2-D surfaces of predicted annual average SO_4^{2-} concentrations (upper plot, large SO_2 point sources ($> 20{,}000$ tons/year) labeled as stars) and annual average SO_4^{2-} concentrations interpolated from IMPROVE and EPA measurements (lower plot, IMPROVE sites labeled as dots, CSN sites labeled as triangles).

The significant predictors in the alternative GLM (3b) are identical to those for the full GAM in the east, and slightly different in the west (component 21 is insignificant). The regression coefficients in the east (not shown here) are similar to those presented in Liu *et al.* [14], with components 1 and 21 having the largest positive coefficients. The regression coefficients in the west (not shown here) are slightly different, probably due to a much larger sample size in the current study. The adjusted model R^2 values (0.47 in the east and 0.40 in the west) indicate lower model predicting power than the full GAM. As shown in Fig. 3, the associations between SO_4^{2-} concentration and some aerosol components deviate far from linear (e.g., fractional AODs of components 2, and 21); therefore, adopting a more stringent linear model form will limit predicting power.

C. Domain Predictions

Due to MISR's relatively low sampling frequency (approximately once a week at mid-latitudes) and data loss due to cloud cover, it is not feasible to study the daily SO_4^{2-} spatial trend. Instead, we first aggregate predicted daily SO_4^{2-} concentrations to monthly averages, then calculate average predicted SO_4^{2-} concentrations over the two-year period using these monthly values; this avoids overweighting the summer months, which have more daily observations. Each 50-km grid cell has an average of 60 MISR observations, with a range from 1 to 117. The upper plot of Fig. 4 shows the spatial pattern of predicted average SO_4^{2-} concentrations with 167 large SO_2 point sources ($> 20{,}000$ tons/year). Predicted SO_4^{2-} concentrations range from below 0.5 μg/m^3 in west Montana to 16 μg/m^3 northeast corner of Texas. Predicted SO_4^{2-} concentrations are on average substantially higher in the east (4.3 μg/m^3) than in the west (1.5 μg/m^3). The lower plot of Fig. 4 shows interpolated contours of average SO_4^{2-} concentrations in 2001 and 2005, based on IMPROVE and EPA CSN measurements. The spatial interpolation was conducted using Kriging with an exponential semivariogram model as implemented in IDL (ITT Visual Information Solutions 2009). The spatial interpolation is to help visualize the spatial patterns in the data and is not meant to estimate the sulfate concentrations between the sites.

The large-scale spatial gradient of MISR-predicted SO_4^{2-} concentrations are consistent with the interpolated contours. The interpolated contours show that the Ohio River Valley has the highest sulfate concentrations where there are a number of large coal-fired power plants. MISR predicts a mean SO_4^{2-} concentration of 6.2 μg/m^3 in this approximately 150 km $\times$ 100 km region, which is in good agreement with interpolated concentrations (5.9–6.7 μg/m^3). Similarly, MISR predicted SO_4^{2-} concentration in Pittsburgh (6.6 μg/m^3) agrees with interpolated results. However, these prominent features are obscured in the prediction map by many fine-scale hot spots with higher MISR predicted SO_4^{2-} concentrations. Fig. 4 shows that many hot spots seen on top of the relatively smooth spatial surface are near major SO_2 emission sources such as large coal-fired power plants. Since the spatial term in the GAM is highly smooth, these hot spots arise from high MISR fractional AODs. The potential causes for the differences between the predicted and spatially interpolated measured concentrations may indicate a more heterogeneous SO_4^{2-} spatial pattern than is captured by the current monitoring networks. After all, the grid cells in our domain are over 10 times the number of ground monitors.

However, some of predicted hot spots may be noises. This is indicated by collocated CSN and IMPROVE monitors that do not observe very high pollution levels. For example, MISR predicted SO_4^{2-} concentrations range from 7.5 to 9.1 μg/m^3 in Washington DC—Baltimore region but interpolated concentrations are between 5.0–5.8 μg/m^3. In addition, interpolated contours do not show any high concentrations along Lake Erie. Other hot spots are not covered by any ground monitors, making them difficult to evaluate. Potential sources of error include uncertainties of the regression models, limitations in the MISR AOD retrieval accuracy, and satellite sampling biases. Predicted SO_4^{2-} concentrations are generally biased high partially because very low AOD values (< 0.05) are excluded to ensure high quality of fractional AODs. In addition, MISR data are only available on relatively cloud-free days but the lowest SO4 concentrations generally coincide with cloudy days during precipitation events. A detailed examination of these pollution hot spots in conjunction with regional air quality model simulations should be conducted to better understand the causes of the discrepancies.

IV. Summary

We previously developed a three-step approach using MISR fractional AODs and GEOS-Chem aerosol simulations to predict ground-level SO_4^{2-} concentrations. In this analysis, we advanced the original approach with a more sophisticated and flexible GAM structure. When compared with EPA CSN and IMPROVE measurements in 2001 and 2005, the full GAM is able to explain 70% of the variability in SO_4^{2-} concentrations. Regional predictions using this model show consistent large spatial patterns with interpolated contours from ground measurements. Our model also identifies areas with elevated SO_4^{2-} levels, which may not be captured by simple interpolation. Comparison with alternative models shows that GEOS-Chem simulated aerosol vertical profiles improve model-predicting power and reduce spatial bias by removing the impact of long-range transport events unrelated to ground-level SO_4^{2-} concentrations. By allowing nonlinear relationships between SO_4^{2-} concentrations and fractional AODs, the GAM has consistent model structure, and proves to be more effective than linear models. The integration of satellite remote sensing, global CTM, and advanced statistical modeling produces valuable information on the spatial and temporal trends of ground-level SO_4^{2-} concentrations. Although daily predictions are limited by satellite sampling frequency, long-term average results provided by this method can be used to examine the effectiveness of emission control polices and validate regional air quality model simulations on broad spatial scales. Ideally, such models can be more effective if additional MISR aerosol components can be designed to match particle speciation definitions in atmospheric chemistry models more precisely. Further research is needed to integrate relative humidity profiles into the model to account for particle growth effects, to better understand the causes of discrepancies in the spatial patterns derived exclusively from the monitoring data and in conjunction with the MISR AOD products, and to understand the impact of co-linearities in the MISR AOD products on the regression model results.

Acknowledgment

The authors would like to thank Dr. R. Kahn of the NASA Goddard Space Flight Center for technical support on MISR data and helpful comments on the manuscript. They would also like to thank the anonymous reviewers for their thoughtful comments on improving the quality of this paper.

References

[1] D. W. Dockery, C. A. Pope, X. P. Xu, J. D. Spengler, J. H. Ware, M. E. Fay, B. G. Ferris, and F. E. Speizer, "An association between air-pollution and mortality in 6 United-States cities," *New England J. Med.*, vol. 329, pp. 1753–1759, Dec. 1993.

[2] C. A. Pope and D. W. Dockery, "Health effects of fine particulate air pollution: Lines that connect," *J. Air Waste Manage. Assoc.*, vol. 56, pp. 709–742, Jun. 2006.

[3] A. Willis, M. Jerrett, R. T. Burnett, and D. Krewski, "The association between sulfate air pollution and mortality at the county scale: An exploration of the impact of scale on a long-term exposure study," *J. Toxicol. Environ. Health A*, vol. 66, pp. 1605–1804, Aug. 2003.

[4] E. L. Gego, P. S. Porter, J. S. Irwin, C. Hogrefe, and S. T. Rao, "Assessing the comparability of ammonium, nitrate and sulfate concentrations measured by three air quality monitoring networks," *Pure Appl. Geophys.*, vol. 162, pp. 1919–1939, Oct. 2005.

[5] W. C. Malm, B. A. Schichtel, R. B. Ames, and K. A. Gebhart, "A 10-year spatial and temporal trend of sulfate across the United States," *J. Geophys. Res. Atmos.*, vol. 107, Nov. 2002.

[6] J. V. Martonchik, R. A. Kahn, and D. J. Diner, "Retrieval of aerosol properties over land using MISR observations," in *Satellite Aerosol Remote Sensing Over Land*, A. A. Kokhanovsky and G. d. Leeuw, Eds. Berlin, Germany: Springer, 2009.

[7] J. V. Martonchik, D. J. Diner, R. A. Kahn, T. P. Ackerman, M. M. Verstraete, B. Pinty, and H. Gordon, "Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 7, pp. 1212–1227, Jul. 1998.

[8] J. A. Engel-Cox and R. M. Hoff, "Science-policy data compact: Use of environmental monitoring data for air quality policy," *Environ. Sci. Policy*, vol. 8, pp. 115–131, 2005.

[9] Y. Liu, J. Sarnat, V. Kilaru, D. Jacob, and P. Koutrakis, "Estimating ground-level $PM_{2.5}$ in the eastern United States using satellite remote sensing," *Environ. Sci. Technol.*, vol. 39, pp. 3269–3278, 2005.

[10] R. B. A. Koelemeijer, C. D. Homan, and J. Matthijsen, "Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe," *Atmos. Environ.*, vol. 40, pp. 5304–5315, Sep. 2006.

[11] R. Kahn, P. Banerjee, and D. McDonald, "Sensitivity of multiangle imaging to natural mixtures of aerosols over ocean," *J. Geophys. Res.-Atmos.*, vol. 106, pp. 18219–18238, Aug. 2001.

[12] R. Kahn, A. Petzold, M. Wendisch, E. Bierwirth, T. Dinter, M. Esselborn, M. Fiebig, B. Heese, P. Knippertz, D. Muller, A. Schladitz, and A. V. Hoyningen-Huene, "Desert dust aerosol air mass mapping in the western sahara, using particle properties derived from space-based multi-angle imaging," *Tellus Ser. B Chem. Phys. Meteorol.*, vol. 61, pp. 239–251, Feb. 2009.

[13] Y. Liu, R. Kahn, and P. Koutrakis, "Estimating $PM_{2.5}$ component concentrations and size distributions using satellite retrieved fractional aerosol optical depth: Part I—Method development," *J. Air Waste Manage. Assoc.*, vol. 57, pp. 1351–1359, Nov. 2007.

[14] Y. Liu, R. Kahn, S. Turquety, R. M. Yantosca, and P. Koutrakis, "Estimating $PM_{2.5}$ component concentrations and size distributions using satellite retrieved fractional aerosol optical depth: Part II—A case study," *J. Air Waste Manage. Assoc.*, vol. 57, pp. 1360–1369, Nov. 2007.

[15] D. J. Jacob, "Heterogeneous chemistry and tropospheric ozone," *Atmos. Environ.*, vol. 34, pp. 2131–2159, 2000.

[16] R. J. Park, D. J. Jacob, B. D. Field, R. M. Yantosca, and M. Chin, "Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy," *J. Geophys. Res. Atmos.*, vol. 109, Aug. 2004.

[17] R. J. Park, D. J. Jacob, P. I. Palmer, A. D. Clarke, R. J. Weber, M. A. Zondlo, F. L. Eisele, A. R. Bandy, D. C. Thornton, G. W. Sachse, and T. C. Bond, "Export efficiency of black carbon aerosol in continental outflow: Global implications," *J. Geophys. Res. Atmos.*, vol. 110, Jun. 2005.

[18] S. N. Wood, "Stable and efficient multiple smoothing parameter estimation for generalized additive models," *J. Amer. Statist. Assoc.*, vol. 99, pp. 673–686, Sep. 2004.

[19] L. J. DeBell, K. A. Gebhart, J. L. Hand, W. C. Malm, M. L. Pitchford, B. A. Schichtel, and W. H. White, Spatial and Seasonal Patterns and Temporal Variability of Haze and Its Constituents in the United States, CIRA IMPROVE Network Report IV to the National Park Service, ISSN: 0737-5352-74, 2006 [Online]. Available: http://vista.cira.colostate.edu/improve/Publications/Reports/2006/2006.htm

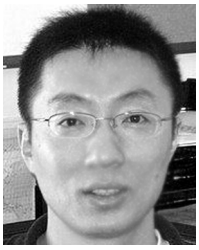
[20] M. Chin, P. Ginoux, S. Kinne, O. Torres, B. N. Holben, B. N. Duncan, R. V. Martin, J. A. Logan, A. Higurashi, and T. Nakajima, "Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and sun photometer measurements," *J. Atmos. Sci.*, vol. 59, pp. 461–483, Feb. 2002.

[21] M. Chin, T. Diehl, P. Ginoux, and W. Malm, "Intercontinental transport of pollution and dust aerosols: Implications for regional air quality," *Atmos. Chem. Phys.*, vol. 7, pp. 5501–5517, Jun. 2007.

[22] D. H. Lowenthal and N. Kumar, "Variation of mass scattering efficiencies in IMPROVE," *J. Air Waste Manage. Assoc.*, vol. 54, pp. 926–934, Aug. 2004.

[23] W. C. Malm and J. L. Hand, "An examination of the physical and optical properties of aerosols collected in the improve program," *Atmos. Environ.*, vol. 41, pp. 3407–3427, May 2007.

[24] Y. Liu, J. A. Sarnat, B. A. Coull, P. Koutrakis, and D. J. Jacob, "Validation of multiangle imaging spectroradiometer (MISR) aerosol optical thickness measurements using aerosol robotic network (AERONET) observations over the contiguous United States," *J. Geophys. Res. Atmos.*, vol. 109, Mar. 2004.



Yang Liu received the B.S. degree in environmental engineering from Tsinghua University, China, in 1997, the M.S. degree in mechanical engineering from the University of California, Davis, CA, in 1999, the Ph.D. degree in environmental science and engineering from the School of Engineering and Applied Sciences, Harvard University, Boston, MA, in 2004.

He joined the Rollins School of Public Health, Emory University, Atlanta, GA, as an Assistant Professor of the Department of Environmental and Occupational Health in 2009. For the past ten years, his primary research interest has been in the applications of satellite remote sensing, atmospheric model simulations, and spatial statistics in air pollution exposure modeling. His current work includes PM exposure modeling in the northeastern U.S. in relation to birth outcomes; an investigation of improving the accuracy satellite derived PM concentrations using space-borne and ground-based lidar profiles; coupling satellite observations with ground monitoring to study the associations between cardiovascular hospital admissions and air pollution in Beijing, China; and study of the heat-related mortality/morbidity in the southwestern and southern U.S. to help develop effective public health indicators.

Bret A. Schichtel received the D.Sc. degree in mechanical engineering from Washington University, St. Louis, MO, in 1996.

He is currently a physical scientist for the National Park Service, Air Resource Division. He has been involved in air quality research for about 20 years conducting both regulatory and research assessments. This work involved the participation in a number of national and regional air quality studies and programs to understand the causes of elevated ozone, particulate matter, visibility, and nitrogen deposition levels at sensitive receptors. He is also actively involved in the Interagency Monitoring of Protected Visual Environments (IMPROVE) program which measures speciated particulate matter in class I areas in support of the Regional Haze Rule.

Petros Koutrakis received the B.S. degree in chemistry from the University of Patras, Greece, in 1980, the M.S. degree in atmospheric chemistry in 1982, and the Ph.D. degree in environmental chemistry in 1984 from the University of Paris, France.

He is a Professor of environmental sciences and the Director of the Exposure, Epidemiology, and Risk Program at the School of Public Health, Harvard University, Boston, MA, and the Director of the EPA/Harvard University Center for Ambient Particle Health Effects. He has conducted numerous air pollution studies in the United States, Canada, Spain, Chile, Kuwait, Cyprus, and Greece that investigate the extent of human exposures to gaseous and particulate air pollutants. Other research interests include the assessment of particulate matter exposures and their effects on the cardiac and pulmonary health.