Adv. Stat. Clim. Meteorol. Oceanogr., 1, 59–74, 2015
www.adv-stat-clim-meteorol-oceanogr.net/1/59/2015/
doi:10.5194/ascmo-1-59-2015

ASCMO
Open Access

Autoregressive spatially varying coefficients model for predicting daily $PM_{2.5}$ using VIIRS satellite AOT

E. M. Schliep[1], A. E. Gelfand[2], and D. M. Holland[3]

[1]Department of Statistics, University of Missouri, Columbia, Missouri, USA
[2]Department of Statistical Science, Duke University, Durham, North Carolina, USA
[3]National Exposure Research Laboratory, US Environmental Protection Agency, Research Triangle Park, North Carolina, USA

Correspondence to: E. M. Schliep (schliepe@missouri.edu)

Received: 16 July 2015 – Revised: 24 November 2015 – Accepted: 1 December 2015 – Published: 16 December 2015

Abstract. There is considerable demand for accurate air quality information in human health analyses. The sparsity of ground monitoring stations across the United States motivates the need for advanced statistical models to predict air quality metrics, such as $PM_{2.5}$, at unobserved sites. Remote sensing technologies have the potential to expand our knowledge of $PM_{2.5}$ spatial patterns beyond what we can predict from current $PM_{2.5}$ monitoring networks. Data from satellites have an additional advantage in not requiring extensive emission inventories necessary for most atmospheric models that have been used in earlier data fusion models for air pollution. Statistical models combining monitoring station data with satellite-obtained aerosol optical thickness (AOT), also referred to as aerosol optical depth (AOD), have been proposed in the literature with varying levels of success in predicting $PM_{2.5}$. The benefit of using AOT is that satellites provide complete gridded spatial coverage. However, the challenges involved with using it in fusion models are (1) the correlation between the two data sources varies both in time and in space, (2) the data sources are temporally and spatially misaligned, and (3) there is extensive missingness in the monitoring data and also in the satellite data due to cloud cover. We propose a hierarchical autoregressive spatially varying coefficients model to jointly model the two data sources, which addresses the foregoing challenges. Additionally, we offer formal model comparison for competing models in terms of model fit and out of sample prediction of $PM_{2.5}$. The models are applied to daily observations of $PM_{2.5}$ and AOT in the summer months of 2013 across the conterminous United States. Most notably, during this time period, we find small in-sample improvement incorporating AOT into our autoregressive model but little out-of-sample predictive improvement.

1 Introduction

Particulate matter (PM) in the atmosphere poses a dangerous public health risk worldwide with effects ranging from reduced vision to respiratory and cardiovascular problems (e.g., US Environmental Protection Agency, 2004; Pope III and Dockery, 2006; Miller et al., 2007; Valavanidis et al., 2008). It has been estimated that a reduction in PM in metropolitan areas in the United States may increase overall life expectancy by as much as 15 % (Pope III et al., 2009). Small PM, such as that with aerodynamic diameters less than $2.5\,\mu g\,m^{-3}$ ($PM_{2.5}$), has been shown to have high toxicity (Valavanidis et al., 2008). New air quality regulations for $PM_{2.5}$ were put into effect by the US EPA in 1997. The EPA designated Federal Reference Methods (FRMs) and Federal Equivalent Methods (FEMs) are used to sample and analyze $PM_{2.5}$. In addition, the EPA worked with federal, state, local, and tribal agencies to establish a network of ambient mass monitoring methods nationwide.

Over the past decade, the demand for spatial models to provide estimates of air quality for inputs in human health analyses and to assess advances in air quality has grown rapidly. Due to the spatial sparsity and varying monitoring schedules of fine particulate monitoring stations, several is-

sues, which rely on surface $PM_{2.5}$ measurements, such as air quality compliance, air quality trends, and daily air quality indices, could benefit from improved spatial prediction of $PM_{2.5}$ in areas where $PM_{2.5}$ is not observed. The use and level of complexity of aerosol optical thickness (AOT), also referred to as aerosol optical depth (AOD), in predicting surface $PM_{2.5}$ has steadily increased in recent years. Early efforts focused solely on simple linear relationships (Wang and Christopher, 2003; Engel-Cox et al., 2004), later ones on multivariate regression (Al-Hamdan et al., 2009; Paciorek and Liu, 2012; Kloog et al., 2011, 2014). Other investigators (Roy et al., 2007; van Donkelaar et al., 2012) use atmospheric model output such as the GEOS-Chem chemical transport model and the Community Multi-Scale Air Quality (CMAQ) model to develop empirical relationships between AOD and $PM_{2.5}$ and use them to calibrate observed AOD in model-based prediction of $PM_{2.5}$.

Of interest here is evaluation of the effectiveness of employing columnar measurements of AOT obtained using passive satellite remote sensing in predicting $PM_{2.5}$ over the conterminous United States. The Geostationary Operational Environmental Satellite (GOES) Aerosols and Smoke Product (GASP), the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on both NASA's Terra and Aqua satellites, and the Multi-angle Imaging Spectroradiometer (MISR) instrument, also on Terra, are customary sources of AOT data found in the literature. The Visible Infrared Imaging Radiometer Suite (VIIRS) instrument, launched in October 2011 on board the Suomi National Polar-orbiting Partnership satellite, a joint National Oceanic and Atmospheric (NOAA)/NASA partnership, was designed to provide continuity between the current MODIS instruments and future VIIRS instruments.

The VIIRS data used for this analysis were obtained from the NOAA Comprehensive Large Array-data Stewardship System (CLASS) (Jackson et al., 2013). The Joint Polar Satellite System Algorithm Engineering Review Board released the VIIRS Aerosol Optical Thickness Intermediate Product (IP) and Environmental Data Record (EDR) to the public with a validated stage 2 level maturity with an effective date of 23 January 2013[1] (Liu et al., 2014). The VIIRS AOT EDR data product has improved spatial resolution over MODIS. The VIIRS AOT is derived from an aggregation of 8 m × 8750 m IP pixels providing a final data product at a 6 km spatial resolution versus the 10 km from MODIS, both at nadir. Additionally, VIIRS, having a larger swath than MODIS (3040 km versus 2330 km), has spatial resolution at the edge of the swath of 12 km (2 times resolution at nadir), whereas the resolution of MODIS is 40 km (4 times resolution at nadir).

AOT satellite measurements have been shown to be a proxy for surface $PM_{2.5}$ with varying degrees of success. For example, Liu et al. (2005) report a strong potential for satellite remote sensing in air quality monitoring. Conversely, Paciorek and Liu (2009) found limited spatial association between the two data sources at daily, monthly average, and annual average temporal resolutions. They conclude that AOT provided little additional value in a model for monthly $PM_{2.5}$ that already accounts for local emissions, meteorology, land use, and regional variability.

One of the main advantages of estimating ground-level $PM_{2.5}$ using satellite AOT data is that satellites often provide spatial coverage over the entire domain of interest in comparison to the relatively sparse point location monitoring data. We would hope to enhance prediction of $PM_{2.5}$ at remote locations by conditioning on observed AOT. However, there are many challenges to this approach. First, the strength of relationship between the two data sources varies greatly in both time and space primarily due to the vertical structure of AOT. When the concentration of AOT is close to the surface, the correlation between AOT and $PM_{2.5}$ will tend to be high; when the concentration of AOT is higher in the column, the correlation between AOT and $PM_{2.5}$ will be low. Specifically, correlation tends to be strongest on days with low cloud cover, low relative humidity, and good vertical mixing within the atmosphere (Al-Hamdan et al., 2009). Hoff and Christopher (2009) provided a detailed discussion of the theoretical relationship between AOT and $PM_{2.5}$, along with a comprehensive summary of studies.

More recently, researchers have used Light Detection And Ranging (lidar) measurements, an active remote sensing technique, to provide additional insights into the relationship between the vertical-resolved aerosol extinction coefficient and $PM_{2.5}$. The aerosol extinction coefficient provides a range-resolved measure at which pulses of light, from a lidar, are attenuated due to presence (loading) of aerosols in the atmosphere. Toth et al. (2014) showed that aerosol particle distributions tend to be more concentrated near the surface in the eastern contiguous United States and more diffuse vertically in the western contiguous. They report that the near-surface-integrated extinction measured by the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) flown aboard the Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) is not representative of total column-integrated extinction (AOT) and that the relationship varies spatially. Additionally, they find that near-surface CALIOP extinction is more highly correlated with ground-level $PM_{2.5}$. This suggests that it is necessary to have knowledge of the ratio of near-surface-integrated extinction to column-integrated AOT in order to effectively capture the utility of satellite AOT as a proxy for surface $PM_{2.5}$. Unfortunately, there are fewer than 100 pairs of collocated CALIOP and $PM_{2.5}$ observations. Very recently, localized measures of the vertical structure of aerosols, using a fixed site lidar, showed that the total height of the aerosols, referred to as the haze layer

[1] see http://www.nsof.class.noaa.gov/notification/pdfs/VIIRSAerosolAOT_APSPEDRProvisionalReleaseReadme_Final.docx for statement of maturity.

height, can be used to normalize AOT (Chu et al., 2015). However, these results are very localized in both time and space.

Extensive missing data in both data sources pose a second challenge and missingness is not necessarily missing at random (Kloog et al., 2014). For satellite data, the presence of clouds within a pixel prevents a valid retrieval of AOT and can also impact retrieval of AOT on adjacent pixels due to the scattering from the clouds. As a result, the use of cloud masks and quality control flags to avoid bias in AOT from nearby clouds can result in a high percentage of missing AOT data. Additionally, the data sources are spatially and temporally misaligned. That is, polar-orbiting satellites obtain AOT data only once per day and the observations are the integral of the atmospheric column. The AOT observations are reported on areal grid cells of varying resolution depending on the satellite and processing. By comparison, $PM_{2.5}$ monitoring stations are given at fixed point locations, and allow for various temporal summaries of ground-level $PM_{2.5}$.

Correlation between data sources has been shown to be better, leading to improved out-of-sample prediction, when modeling daily data versus data aggregated to monthly or yearly scales (Paciorek and Liu, 2012). Using satellite and monitoring station data for five southeastern states in the year 2003, Al-Hamdan et al. (2009) found that including meteorological data with remote sensing data in a multiple regression model increases the amount of explained variation in $PM_{2.5}$. Lee et al. (2012) modeled daily $PM_{2.5}$ data from 69 monitoring sites in the northeast for years 2000–2008 using a linear model that regresses PM on 10 km resolution MODIS satellite AOT with a day-specific coefficient for AOT. Out-of-sample prediction is done using the fitted regression model when AOT is observed. When missing, $PM_{2.5}$ is predicted using the daily average observed value for $PM_{2.5}$ along with a location-specific random effect based on a K-means approach.

Kloog et al. (2014) proposed models to predict daily $PM_{2.5}$ at high resolution (1 km × 1 km) across the northeastern United States using the Multi-Angle Implementation of Atmosphere Correction (MAIAC) algorithm. MAIAC is a MODIS satellite data processing algorithm that performs aerosol retrievals and atmospheric correction over both dark vegetated surfaces and bright deserts based on a time series analysis and image-based processing (Lyapustin et al., 2011). Hu et al. (2014) proposed a two-stage model using both MAIAC-Aqua and MAIAC-Terra AOT 1 km resolution data for the southeastern United States in 2003. At the first stage they fitted a multiple linear regression model with day-specific intercept and coefficient for AOT and include land use and meteorological data as covariates. At the second stage they regress the residuals of $PM_{2.5}$ from the first stage onto location-specific intercept and coefficient for AOT that are obtained via a geographical weighting function. Under these multiple-stage models, it is unclear how or if error is propagated through to the final estimates of $PM_{2.5}$. Therefore, uncertainty may be underestimated.

Van Donkelaar et al. (2012) improved upon the prediction of daily ground-level $PM_{2.5}$ using AOT and daily-generated scale factors from the GEOS-Chem chemical transport model. That is, they apply a linearly interpolated simulation ratio of ground-level $PM_{2.5}$ and AOT from GEOS-Chem to the observed MODIS and MISR AOT to get a satellite-driven $PM_{2.5}$ estimate followed by a climatological ground-based regional bias correction factor derived from a comparison with surface-based $PM_{2.5}$ monitoring data. Their model is fitted to data obtained between the years 2004 and 2009 across North America. Model performance is evaluated using R^2 values by day and by region.

Our contribution is to address some of the foregoing modeling challenges in predicting daily $PM_{2.5}$ and apply formal model comparison across competing candidate models. We propose a daily hierarchical autoregressive spatially varying coefficients model to jointly predict, at point level, daily average $PM_{2.5}$ for consecutive days across the contiguous United States, using AOT obtained from VIIRS. The model encompasses several important features. First, due to the high degree of temporal dependence between daily measurements of $PM_{2.5}$, we include an autoregressive term in the model. This was motivated by exploratory data analysis of daily $PM_{2.5}$ observations from 18 June to 9 July 2013. For these 22 consecutive days, we found correlation between $PM_{2.5}$ on adjacent days ranging between 0.52 and 0.91. Second, since correlation between $PM_{2.5}$ and AOT varies both in time and in space depending on the vertical structure of AOT, our model employs a day-specific spatially varying intercept and coefficient for AOT. Third, due to missingness in the daily AOT surfaces, we define our model hierarchically such that there is a model for AOT at the second level. This allows us to implement simultaneous model-based imputation of AOT at missing grid cells through connection to a model for $PM_{2.5}$ defined conditionally on AOT. The hierarchical modeling framework also ensures that we propagate uncertainty in our estimates of AOT to uncertainty in our estimates of $PM_{2.5}$. Lastly, the model includes meteorology data as covariates since variables such as temperature, wind, absolute humidity, and precipitation have strong effects on the concentration of $PM_{2.5}$ and speciated-$PM_{2.5}$ (Dawson et al., 2007). Moreover, meteorology data have been shown to help explain the vertical structure of AOT, and thus, strengthen the relationship between AOT and $PM_{2.5}$ and improve prediction at unobserved locations (Al-Hamdan et al., 2009; Zeeshan and Oanh, 2014; Wang et al., 2013).

For model comparison, we consider competing submodels nested within our model in terms of fit (in-sample) and prediction (out-of-sample) of $PM_{2.5}$. Specifically, we evaluate the predictive benefit of spatially varying coefficients, autoregression, and, perhaps of most interest, AOT. To our knowledge, there is no work in the literature that includes such extensive model comparison over so many days and over such

a large region. Unfortunately, atmospheric numerical model output from the CMAQ[2] was not available for the subject time period of this analysis. Thus, we are unable to offer any comparison of fusion with AOT to fusion with CMAQ as described in, e.g., Berrocal et al. (2012).

The remainder of this paper is structured as follows. In Sect. 2 we describe the $PM_{2.5}$ and VIIRS satellite-obtained AOT data. In Sect. 3 we define our autoregressive spatially varying coefficients model and methods for obtaining inference in a Bayesian framework. In Sect. 4, the model and three nested submodels are compared based on both model fit and out-of-sample prediction using daily $PM_{2.5}$ and AOT for 18 June–9 July 2013. We end with a summary and directions of future work in Sect. 5.

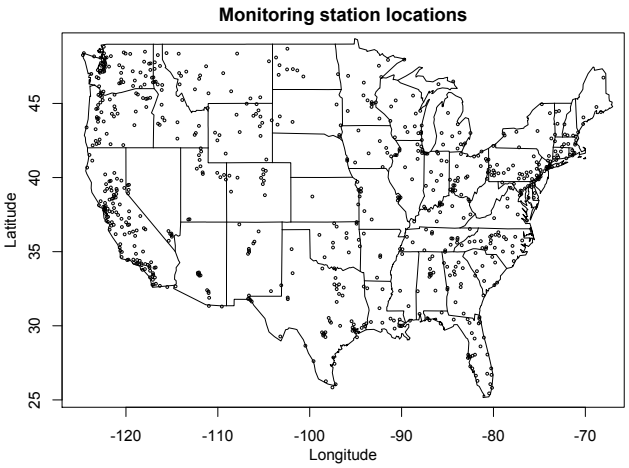

Figure 1. The 764 monitoring station locations of $PM_{2.5}$ within the AQS repository for the summer months of 2013.

2 $PM_{2.5}$ and AOT data

2.1 Monitoring station $PM_{2.5}$ data

We obtain $PM_{2.5}$ data from the Environmental Protection Agency's (EPA) Air Quality System (AQS) Datamart[3], which is a repository of ambient air quality data. In comparison with the FRM $PM_{2.5}$ monitoring network[4], the AQS monitoring data provide a larger number and better spatial distribution of daily monitoring sites to use in this evaluation of daily $PM_{2.5}$ prediction using AOT data. The AQS data consist of continuous, hourly non-FRM $PM_{2.5}$ observations at 764 unique monitoring sites throughout the conterminous United States to which we convert to daily averages. Monitoring locations tend to be concentrated in areas of high population density (see Fig. 1) leaving large spatial areas with few or no observations. Prediction using just AQS data may rely on observed $PM_{2.5}$ hundreds of kilometers away. Additionally, there is missingness in observed $PM_{2.5}$ from these monitoring stations. For the summer months of 2013, between 4 and 8 % of monitoring stations are missing observations on a given day.

2.2 VIIRS satellite-obtained aerosol optical thickness

The VIIRS AOT data product used for this study is the validated stage 2 level maturity EDR AOT available at the National Oceanic and Atmospheric (NOAA) Comprehensive Large Array-data Stewardship System (CLASS)[5]. VIIRS EDR AOT is produced at a 6 km × 6 km resolution at nadir and expands to ≈ 12.8 km × 12.8 km at the edge of the swath. For additional details on the VIIRS data products, including the EDR AOT, see Jackson et al. (2013). AOT is the degree to which particles (dust, smoke pollution) prevent the transmission of light by absorption or scattering of light. AOT is the vertical integral of the aerosol extinction coefficient, a summary of several different measures of the absorption of light in the entire column of the atmosphere. Therefore, AOT tells us how much direct sunlight is prevented from reaching the ground by these aerosol particles and can be used to infer about air quality. AOT is a dimensionless number typically between 0 and 1 where a value of 0.01 corresponds to an extremely clean atmosphere and a value of 0.4 corresponds to a very hazy condition. On a typical day, average AOT across the United States ranges from 0.10 to 0.15.

The VIIRS satellite orbits the earth roughly once per hour, producing swaths of AOT data for each orbit. The orbit occurs at roughly 14:30 LT during daylight savings time and 13:30 LT during non-daylight savings time. On each day, the conterminous United States is observed by the compilation of three or four orbits of the satellite, depending on the orbiting pattern. As seen in Fig. 2, there is some overlap between consecutive swaths to ensure that all locations on the globe are observed. The daily satellite observations across the United States are aggregated to 6km grid boxes from 64 0.75 km grid boxes in an 8×8 square. The latitude and longitude are given for the center of each of the grid boxes. Each swath is made up of panels with dimension 96×400. Unfortunately, these panels and grid boxes change across days. We obtain daily VIIRS satellite data for 22 consecutive days during the summer of 2013 from 18 June to 9 July. For uniformity across days, we aggregate the data to a common 12 km lattice. We average AOT values for grid boxes containing more than one observation.

[2]CMAQ data can be obtained at https://www.cmascenter.org/cmaq/.

[3]The EPA's AQS $PM_{2.5}$ data can be obtained at https://ofmext.epa.gov/AQDMRS/aqdmrs.html.

[4]FRM $PM_{2.5}$ data can be obtained at http://www.epa.gov/ttnamti1/pmfrm.html.

[5]The VIIRS aerosol EDRs are available at http://www.nsof.class.noaa.gov.

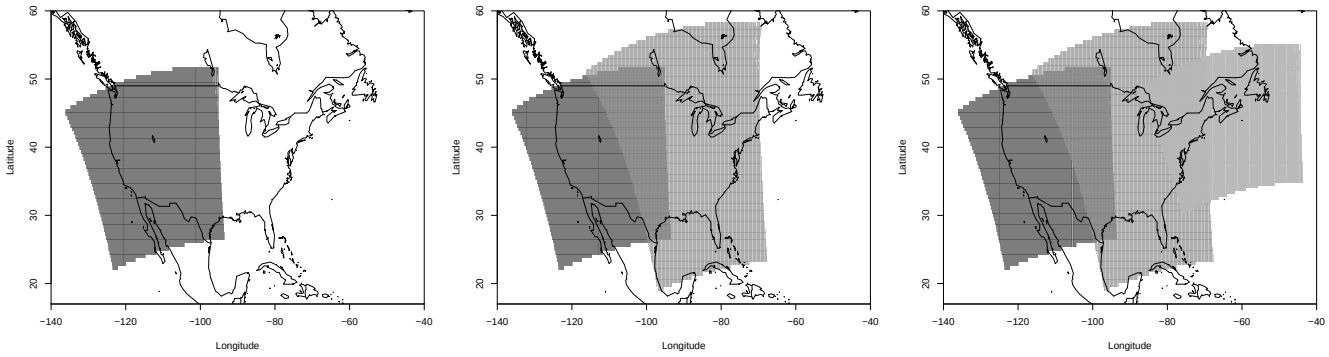

Figure 2. Three swaths covering the continental United States on an illustrative summer day.

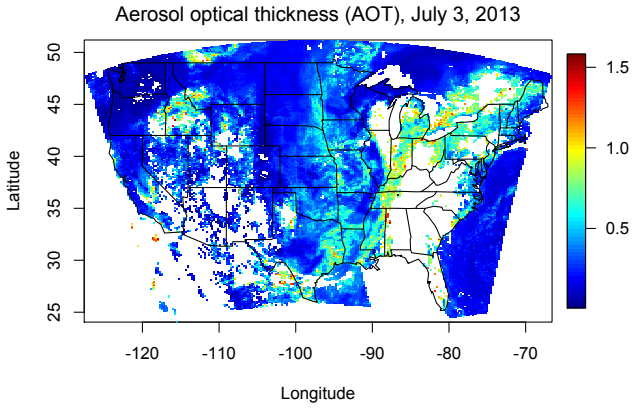

Figure 3. AOT on 3 July 2013.

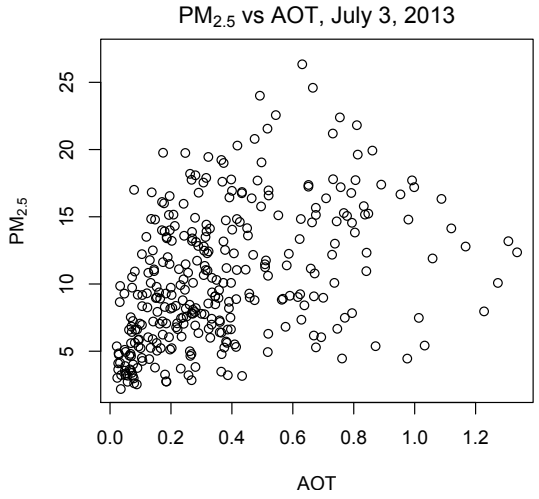

Figure 4. $PM_{2.5}$ vs. AOT on 3 July 2013. Spearman's ρ correlation is 0.47, which was computed using the 522 monitoring station locations that had both $PM_{2.5}$ and AOT observed on that day.

While polar-orbiting satellites produce one measurement of AOT per day across the entire globe, a large number of areal grid cell observations are flagged and removed through a quality assurance protocol (Jackson et al., 2013). Reasons for removal include poor atmospheric conditions, such as dense clouds, or troubling land surface conditions, such as snow, ice, or sand. The highest quality AOT data have the most missing observations and missingness varies both spatially and temporally. Illustratively, observed AOT across the contiguous United States on 3 July 2013 is shown in Fig. 3 where white indicates missing AOT. Here, AOT was observed at 64 % of the grid cells containing one of the 764 monitoring stations. Over the summer months in 2013, the observation rate of AOT at grid cells containing monitoring stations ranges from 45 to 83 % on a given day.

Figure 4 shows a scatter plot of $PM_{2.5}$ and AOT on 3 July 2013 where the Spearman's ρ correlation between the two variables is 0.47. Daily correlation between $PM_{2.5}$ and AOT in the summer months is extremely variable from day to day and ranges between -0.14 and 0.56.

3 An autoregressive spatial varying coefficient model

The model for daily $PM_{2.5}$ is specified on the log-scale, i.e., daily $PM_{2.5}$ is log-transformed. Additionally, AOT is on the square-root scale. Normal Q–Q plots of the log-transformation of $PM_{2.5}$ and square root transformation of AOT are shown for 4 days in the Appendix A1. Both transformations result in slight derivations from normality but the deviations are minimal and the transformations agree with previous work in the literature (e.g., Berrocal et al., 2010 for $PM_{2.5}$, Paciorek and Liu, 2012 for AOD.)

We propose a hierarchical model for $PM_{2.5}$ that incorporates randomness for AOT in order to impute missing satellite data. That is, we do model-based imputation for missing AOT in the context of a conditional model for $PM_{2.5}$ given AOT. Let N denote the fixed number of grid cells in the spatial lattice for AOT. Then, we define the vector of observed AOT on day t as $\boldsymbol{A}_t = \{A_{1,t}, \ldots, A_{n_t,t}\}$, where n_t is the num-

ber grid cells with AOT observations on day t. We model AOT as

$$\boldsymbol{A}_t = \mathbf{M}_t\boldsymbol{\theta}_t + \boldsymbol{\eta}_t, \tag{1}$$

where $\boldsymbol{\theta}_t$ is a vector length N from an intrinsic conditional autoregressive model (IAR) (Besag et al., 1991), $\mathbf{M}_t$ is an $n_t \times N$ indicator matrix for the observed AOT grid cells for day t, and $\boldsymbol{\eta}_t$ is a vector of n_t i.i.d. errors. Therefore, $\boldsymbol{A}_t|\boldsymbol{\theta}_t \sim N(\mathbf{M}_t\boldsymbol{\theta}_t, \sigma_A^2\mathbf{I}_{n_t})$ where $\mathbf{I}_{n_t}$ is a $n_t \times n_t$ identity matrix. The IAR model for $\boldsymbol{\theta}_t$ is defined such that

$$\theta_{t,i}|\boldsymbol{\theta}_{t,-i}, \kappa \sim N\left(\frac{1}{p_i}\sum_{j:j\sim i}\theta_{t,j}, \frac{1}{p_i\kappa}\right),$$

where $j \sim i$ denotes that cells j and i are neighbors, κ is a precision parameter and p_i is the number of neighbors for grid cell i. We employ a neighborhood structure on the regular lattice such that neighbors are the adjacent and diagonal grid cells. The neighborhood structure is constant over time.

Let $\mathrm{P}_t(s)$ denote the observation of $\mathrm{PM}_{2.5}$ at location s on day t. Our most general model for $\mathrm{P}_t(s)$ is defined conditional on $A_{i,t}$ where s is in grid cell i. We employ an autoregressive, spatially varying coefficient model for $\mathrm{P}_t(s)$. The AR(1) model conditional on AOT for $t = 1, \ldots, T$ is defined as

$$\begin{aligned} P_t(s) =& \alpha_{0,t} + \beta_{0,t}(s) + (\alpha_{1,t} + \beta_{1,t}(s))A_{i,t} \\ &+ \boldsymbol{X}_t(s)\boldsymbol{\gamma} + \rho P_{t-1}(s) + \epsilon_t(s). \end{aligned} \tag{2}$$

Here, $\alpha_{0,t}$ and $\alpha_{1,t}$ are the global intercept and AOT coefficients for day t, $\boldsymbol{X}_t(s)$ is a vector of location and day-specific meteorological covariates, $\boldsymbol{\gamma}$ is a vector of coefficients, $\beta_{0,t}(s)$ and $\beta_{1,t}(s)$ are spatially varying intercept and AOT coefficients for day t, and $\epsilon_t(s)$ is i.i.d. (independent and identically distributed) error.

Van Donkelaar et al. (2012) proposed a priori adjustments to AOT that rely on additional sources of computer model output to try to account for the variability in the vertical profile of AOT. Here, the processes $\beta_{0,t}(s)$ and $\beta_{1,t}(s)$ offer a data-driven approach to adjust for the spatiotemporal variability in the relationship between columnar AOT at areal grid cells and ground-level $\mathrm{PM}_{2.5}$ concentrations at point locations.

If AOT is missing at grid cell i on day t, it is imputed within the Markov chain Monte Carlo (MCMC) using Eqs. (1) and (2). To initiate Eq. (2) we must specify a model for $P_0(s)$ in order to predict at unobserved locations. Thus, for $t = 0$, we define

$$\begin{aligned} P_0(s) =& \alpha_{0,0} + \beta_{0,0}(s) + (\alpha_{1,0}(s) + \beta_{1,0}(s))A_{i,0} \\ &+ \boldsymbol{X}_0(s)\boldsymbol{\gamma} + \epsilon_0(s), \end{aligned} \tag{3}$$

where $\alpha_{0,0}$, $\alpha_{1,0}$, $\beta_{1,0}(s)$, $\beta_{0,0}(s)$, $A_{i,0}$, and $\epsilon_0(s)$ are analogous to above.

The spatial processes, $\beta_{0,t}(s)$ and $\beta_{1,t}(s)$, are defined through a linear model of co-regionalization. That is, for $t = 0, \ldots, T$

$$\begin{bmatrix} \beta_{0,t}(s) \\ \beta_{1,t}(s) \end{bmatrix} = \begin{bmatrix} C_{11} & 0 \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} W_{0,t}(s) \\ W_{1,t}(s) \end{bmatrix},$$

where $W_{0,t}(s)$ and $W_{1,t}(s)$ are independent mean 0 Gaussian processes with exponential covariance functions and day-specific spatial decay parameters $\phi_{0,t}$ and $\phi_{1,t}$, respectively. The diagonal elements of the lower-triangular co-regionalization matrix, C_{11} and C_{22}, are non-negative. Lastly, the spatial processes are modeled independently in time as we believe the autoregression will capture the majority of the temporal dependence in the model.

The hierarchical model is fitted within the Bayesian framework, enabling convenient Gibbs sampler loops for imputation. That is, we update the missing AOTs given the parameters and update the parameters given the missing and observed AOTs. We assign prior distributions to all model parameters. When possible, conjugate, non-informative priors are used. For $t = 0, \ldots, T$, $\alpha_{0,t}$ and $\alpha_{1,t}$ are assigned independent normal prior distributions with mean 0 and variance 10^4. The conjugate prior distribution for the meteorology coefficient vector $\boldsymbol{\gamma}$ is $\mathrm{MVN}(0, 1000\mathbf{I}_p)$, where p is the number of covariates in the model.

The diagonal elements of the $\mathbf{C}$ matrix, C_{11} and C_{22}, are assigned independent, diffuse half-normal prior distributions with parameter $\lambda = 0.5$, where the mean is $1/\lambda$ and variance is $(\pi - 2)/2\lambda^2$. The off-diagonal element, C_{21}, is assigned a diffuse mean 0 normal prior with variance 10. These three parameters are updated in a block Metropolis–Hastings step. The day-specific spatial decay parameters, $\phi_{0,t}$ and $\phi_{1,t}$, are assigned uniform prior distributions such that the effective range is less than half the maximum distance. Locations are given in easting and northing coordinates under the Albers conic equal-area projection and distances are computed as Euclidean. The maximum distance between location is 0.8 in Albers standard units. Both error terms, $\eta_{i,t}$ and $\epsilon_t(s)$, are i.i.d. normal random variables with mean 0 and variance parameters σ_A^2 and σ_P^2, respectively. The variance parameters σ_P^2 and σ_A^2 are each assigned IG(2, 2) (inverse-gamma) prior distributions. Lastly, the precision parameter κ is assigned a Gamma(2, 2) prior distribution. These are vague since the priors for σ_P^2, σ_A^2, and $1/\kappa$ have infinite variance.

4 The application

The model is applied to data from 510 monitoring stations located across the conterminous United States that have consecutive daily $\mathrm{PM}_{2.5}$ observations from 18 June to 9 July. This region is spanned by $N = 88\,500$ areal 12 km grid cells of satellite-obtained AOT. The number of grid cells with observed AOT, n_t, ranges from 40 870 to 54 230. Meteorology covariates included in the model are daily average temperature and relative humidity obtained at the monitoring sta-

tions[6]. Additionally, 209 monitoring stations are used for out-of-sample prediction. The number of daily $PM_{2.5}$ observations at these out-of-sample locations ranges from 172 to 203.

We compare our proposed Eq. (2) to three nested submodels, given in detail below. The three submodels are (1) an autoregressive model with daily global intercept and daily spatial-varying coefficients for AOT, (2) a non-autoregressive daily spatially varying coefficients model, and (3) a daily spatial-varying intercept autoregressive model without AOT. We do not consider an autoregressive model with daily spatially varying intercept rather with global daily slope, since we have noted above that there is considerable spatial variability in the association between AOT and $PM_{2.5}$.

Autoregressive model with daily global intercept and daily spatially varying coefficient for AOT

Submodel (S1) is an autoregressive model with only one spatially varying coefficient. Here, the day-specific intercept is global and a spatially varying coefficient for AOT is employed to capture the spatially varying relationship between $PM_{2.5}$ and AOT. The model is written as

$$P_t(s) = \alpha_{0,t} + (\alpha_{1,t} + \beta_{1,t}(s))A_{i,t} + \boldsymbol{X}_t(s)\boldsymbol{\gamma} + \rho P_{t-1}(s) + \epsilon_t(s). \quad (4)$$

The model for AOT is the same as in Eq. (1).

Non-autoregressive daily spatially varying coefficients model

Submodel (S2) is non-autoregressive. That is, both intercept and AOT coefficient are spatially varying and day-specific but there is no temporal dynamics for $PM_{2.5}$. The model is written

$$P_t(s) = \alpha_{0,t} + \beta_{0,t}(s) + (\alpha_{1,t} + \beta_{1,t}(s))A_{i,t} + \boldsymbol{X}_t(s)\boldsymbol{\gamma} + \epsilon_t(s) \quad (5)$$

and is fitted analogously to the autoregressive model but with $\rho = 0$.

Daily spatially varying intercept autoregressive model without AOT

To assess the significance of AOT in our proposed model, we consider a submodel (S3) without AOT. This model still has a spatial random effect to capture the spatial heterogeneity in $PM_{2.5}$ as well as the autoregressive term. That is,

$$P_t(s) = \alpha_{0,t} + \beta_{0,t}(s) + \boldsymbol{X}_t(s)\boldsymbol{\gamma} + \rho P_{t-1}(s) + \epsilon_t(s), \quad (6)$$

where we drop the model for AOT given in Eq. (1).

[6]The meteorological data are reported by the EPA at the AQS monitoring stations and can be obtained at http://aqsdr1.epa.gov/aqsweb/aqstmp/airdata/download_files.html.

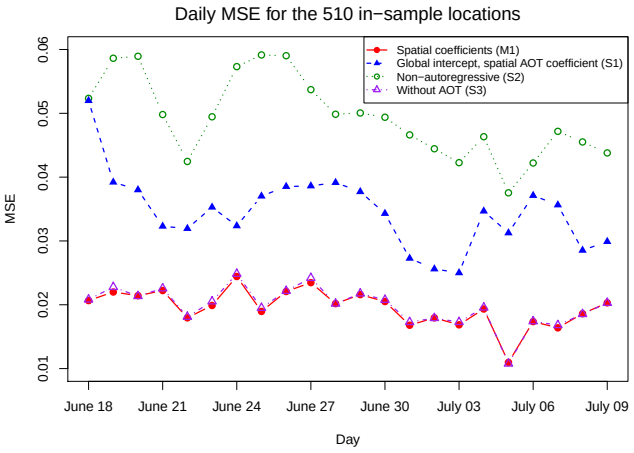

Figure 5. Daily MSE for in-sample locations for 18 June–9 July 2013 under the competing models (Eqs. 2–6).

4.1 Model comparison results

We obtain inference for all model parameters in the Bayesian framework using MCMC and a hybrid Metropolis-within-Gibbs algorithm. Additional details regarding the sampling algorithm are included in Appendix A2. Each of the chains are run for 30 000 iterations and the first 10 000 are discarded as burn-in. We computed Monte Carlo standard errors using the batch means procedure (Flegal et al., 2008) to assess convergence of the chains.

We present the results comparing our proposed model (Eq. 2) and the three submodels, (Eqs. 4–6). In-sample mean square error (MSE) by day is given for the four models in Fig. 5. Additionally, Fig. 6 reports the median absolute deviation (MAD) by day for each of the four models where MAD is computed on the original scale of the data (as opposed to log-transformed). The spatial coefficients model (Eq. 2) reports the lowest MSE and MAD across most days while model (Eq. 6) is very similar. Equation (5) reports the highest MSE and MAD across all days.

We predict $PM_{2.5}$ at the unobserved locations for all 22 days by drawing from the joint posterior predictive distribution. Daily R^2 computed using the out-of-sample locations under Eq. (2) ranges from 0.22 to 0.63. Figure 7 gives daily median absolute prediction deviation (MAPD) for the 209 out-of-sample locations for each of the four models. Note that MAPD is computed using only locations in which $PM_{2.5}$ was observed for that given day. We also compare the models based on the continuous rank probability score (CRPS), a strictly proper scoring rule that evaluates the concentration of a predictive distribution for an observation around that observation (Gneiting and Raftery, 2007). Daily average CRPS are given in Fig. 8. In terms of out-of-sample prediction, Eqs. (2)–(6) are nearly indistinguishable. In terms of large spatial scale dynamic prediction, this suggests that there may

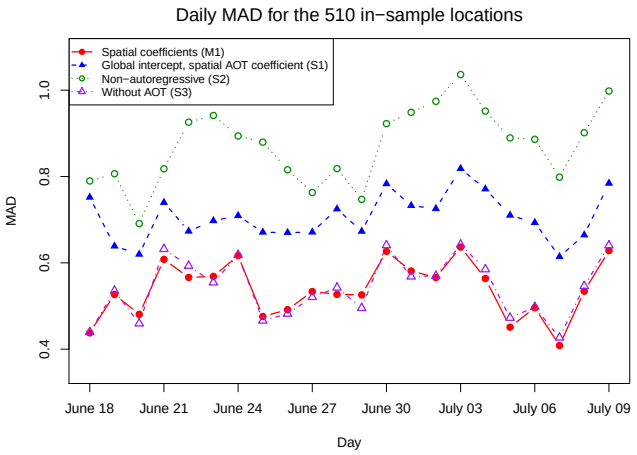

Figure 6. Daily MAD for in-sample locations for 18 June–9 July 2013 under the competing models, (Eqs. 2–6), where MAD is given in units of $\mu g\,m^{-3}$.

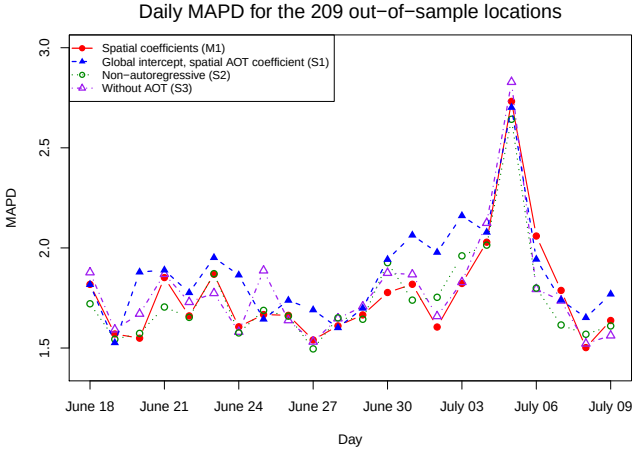

Figure 7. Daily MAPD for the out-of-sample locations under the competing models.

be little benefit in using VIIRS satellite-obtained AOT when predicting $PM_{2.5}$.

4.2 Inference results

Focusing on Eq. (2), we investigate the variability in parameter estimates both temporally and spatially. The day-to-day variability of the global intercept and AOT coefficient terms are shown in Fig. 9. The solid line denotes the posterior mean of the parameter and the dashed lines denote the upper and lower quantities of the 90 % credible interval. The estimates of the daily global coefficients of AOT are generally positive indicating that $PM_{2.5}$ is increasing with AOT, conditional on the effects of the meteorology data and the previous day's $PM_{2.5}$.

Spatial maps of the posterior mean estimates for $\alpha_0+\beta_0(s)$ and $\alpha_1+\beta_1(s)$ for 3 consecutive days are shown in Fig. 10.

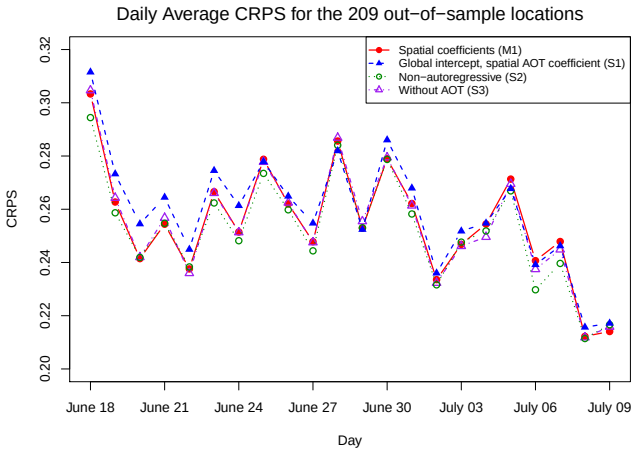

Figure 8. Daily average CRPS for the out-of-sample locations under the competing models.

These processes are capturing the spatial dependence in $PM_{2.5}$ and the spatially varying effect of AOT, respectively, after accounting for the meteorological covariates and autoregression.

There do not appear to be any obvious temporal patterns between consecutive days of either of the spatial processes; this is likely captured by the autoregression. Also, the estimates of $\alpha_1+\boldsymbol{\beta}_1$ appear uniformly larger on 2 July 2013 than on either the previous or subsequent day.

The proportion of AOT coefficient estimates by day (left) and across space (right) that are significantly different than 0 based on the 90 % credible interval are given in Fig. 11. Here, we consider an estimate *significantly different from 0* if the 90 % credible interval does not include 0. We note that here, *significance* is determined marginally; we are not doing simultaneous inference. The daily proportion of locations having coefficient estimates that are significantly different than 0 is heavily skewed and is between 0 and 0.84. There is undoubtedly spatial confounding between the spatial processes $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$. That is to say, the numerous days with few to no locations having significant AOT coefficients does not mean that the linear effect of AOT on $PM_{2.5}$ is insignificant.

Equation (4) only has one spatial process resulting in the proportion of β_1 estimates that are significantly different from 0 to be much larger (Fig. 12). Therefore, to capture the spatially varying significance of AOT, we compute an adjusted estimate of the AOT coefficient under the proposed spatial coefficients Eq. (2). First, notice that at time t, $\alpha_{0,t}+\beta_{0,t}(s)+(\alpha_{1,t}+\beta_{1,t}(s))A_{i,t}$ can be rewritten in vector form as

$$\mathbf{1}\alpha_{0,t}+\boldsymbol{\beta}_{0,t}+\mathbf{D}_A(\mathbf{1}\alpha_{1,t}+\boldsymbol{\beta}_{1,t}), \tag{7}$$

where $\mathbf{D}_A$ is an $n\times n$ diagonal matrix with the $A_{i,t}$'s on the diagonal. Equation (7) can be rewritten as

$$\mathbf{1}\alpha_{0,t}+\mathbf{D}_A(\mathbf{D}_A^{-1}\boldsymbol{\beta}_{0,t}+\mathbf{1}\alpha_{1,t}+\boldsymbol{\beta}_{1,t}), \tag{8}$$

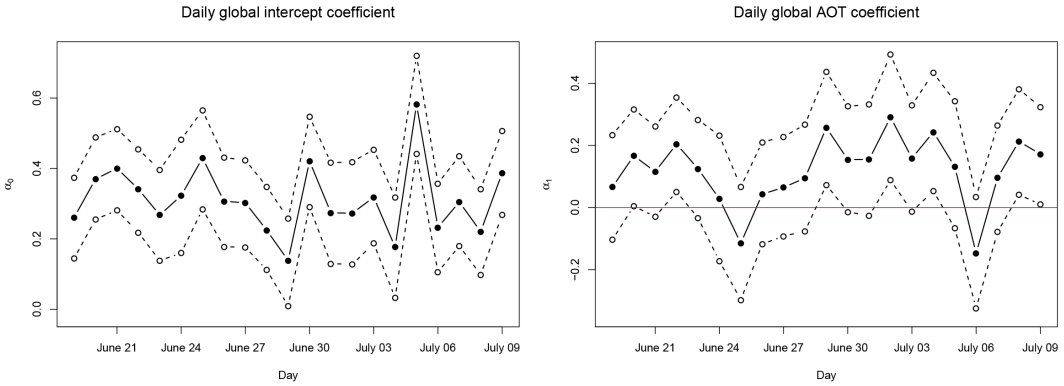

Figure 9. Mean and 90 % credible interval by day for the (left) global intercept, α_0, and (right) coefficient for AOT, α_1, for Eq. (2).

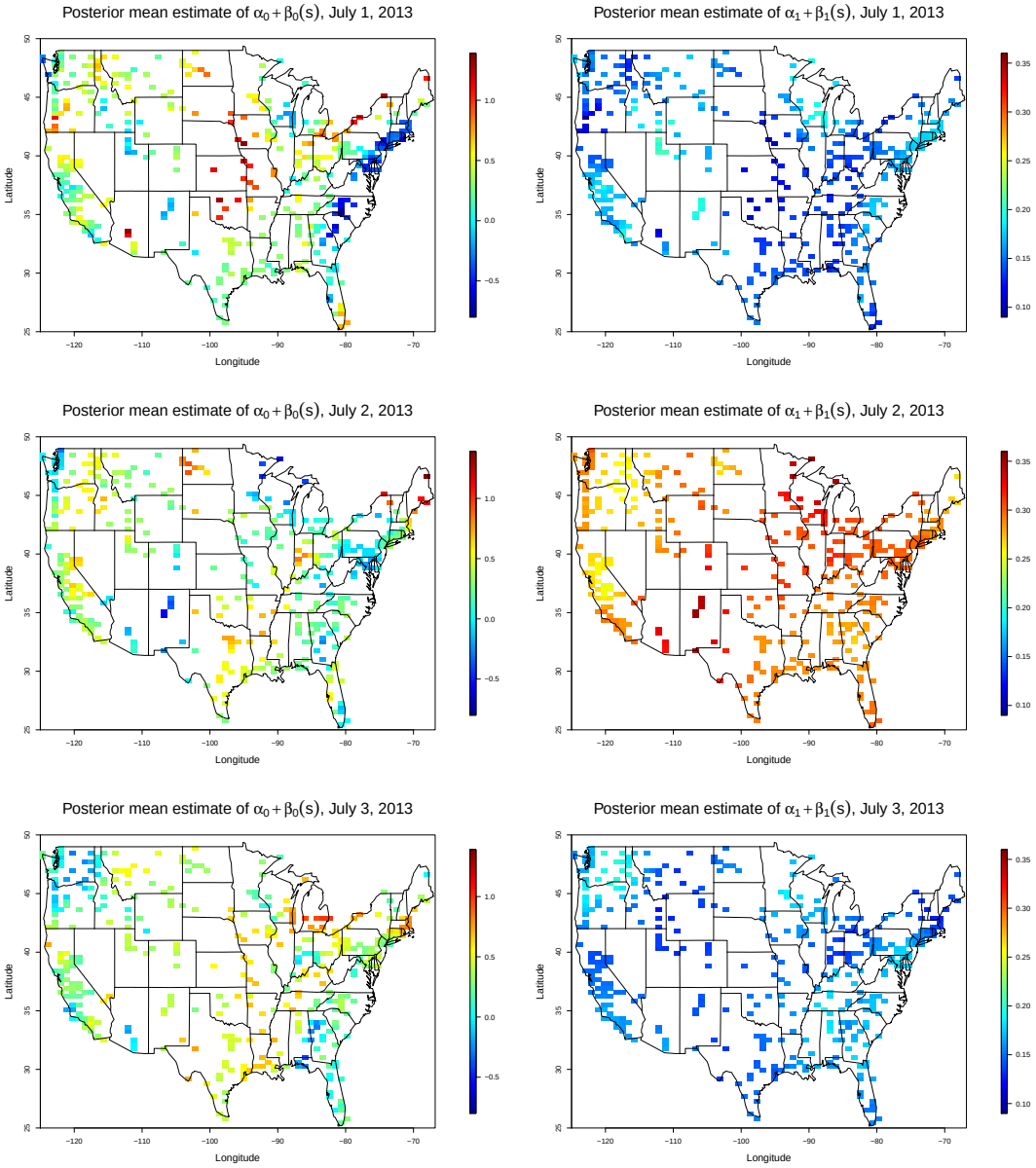

Figure 10. Posterior mean estimates of the coefficients, $\alpha_0 + \beta_0(s)$ and $\alpha_1 + \beta_1(s)$ for 3 consecutive days from Eq. (2).

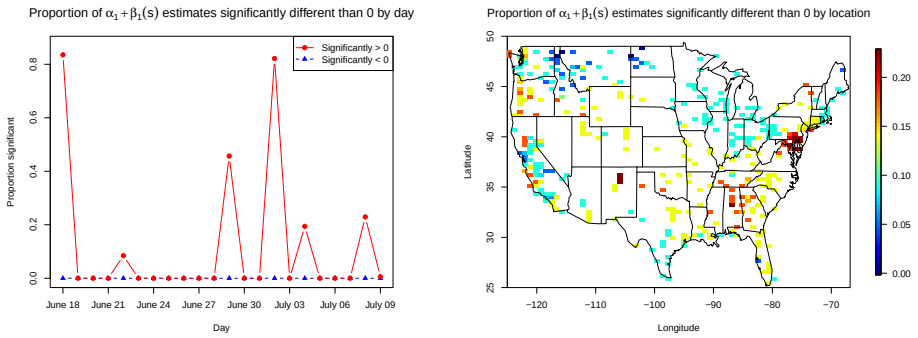

Figure 11. The proportion of $\alpha_1 + \beta_1(s)$ estimates by day (left) and across space (right) that are significantly different than 0, i.e., 90 % credible interval does not include 0, under Eq. (2).

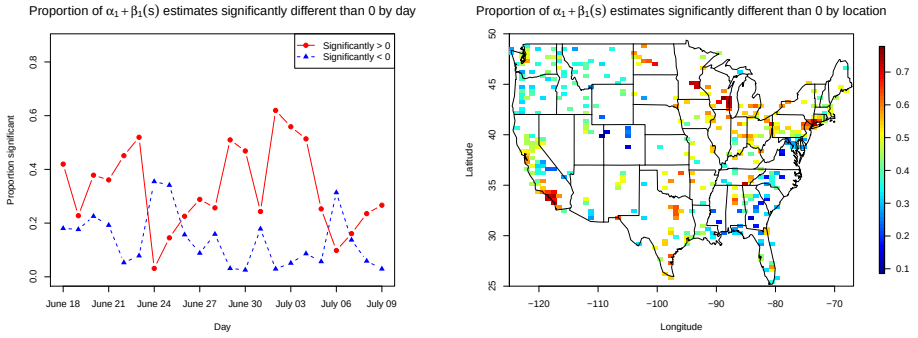

Figure 12. The proportion of $\alpha_1 + \beta_1(s)$ estimates by day (left) and across space (right) that are significantly different than 0, i.e., 90 % credible interval does not include 0, under Eq. (4).

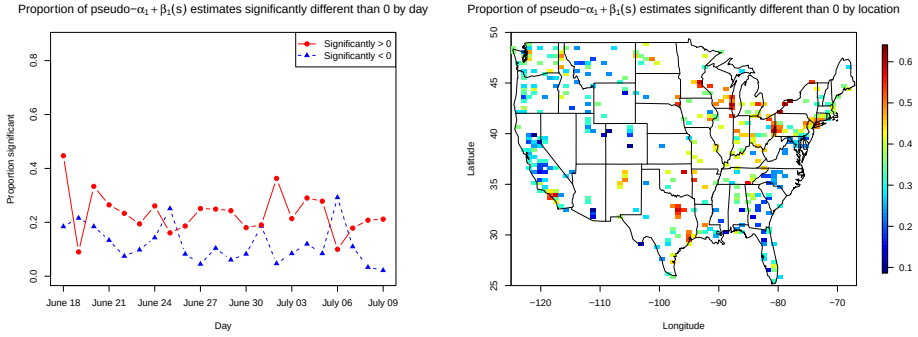

Figure 13. The proportion of pseudo-estimates by day (left) and across space (right) that are significantly different than 0.

where $\mathbf{D}_A^{-1}\boldsymbol{\beta}_{0,t} + \mathbf{1}\alpha_{1,t} + \boldsymbol{\beta}_{1,t}$ is an adjusted estimate of $\mathbf{1}\alpha_{1,t} + \boldsymbol{\beta}_{1,t}$. We computed the posterior mean and 90 % credible interval for this adjusted parameter. The significance of these estimates over time and space are shown in Fig. 13 where now the significance more closely resembles that from Eq. (4). The daily proportion of significant estimates is between 0.23 and 0.63. Spatially, the midwest and central part of the United States tend to have more days with significant spatially varying adjusted estimates.

Posterior estimates and 90 % credible intervals for the remaining model parameters are given in Table 1. In general, $PM_{2.5}$ is increasing with average daily temperature and relative humidity although the effect of relative humidity is not significant. The autocorrelation parameter is significant with mean estimate 0.671 indicating that $PM_{2.5}$ on day $t-1$ is very informative for $PM_{2.5}$ on day t.

Illustratively, Figs. 14 and 15 give spatial maps of the mean and standard deviation of the posterior predictive distribution of $PM_{2.5}$ for 28 June and 3 July 2013, respectively. On 28 June, with the exception of Florida, estimates of $PM_{2.5}$ are higher in the southern half of the country than the northern half. Standard deviations are higher in parts of Nevada, North Dakota, and along the Atlantic coast in North Carolina and Florida. The points on the map indicate the monitoring

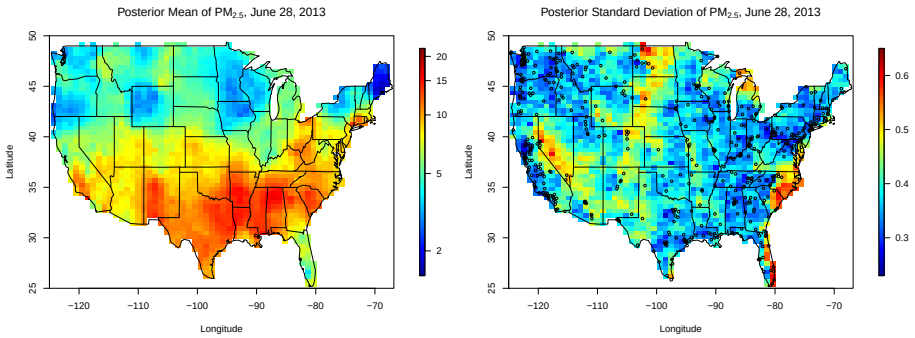

Figure 14. Mean and standard deviation estimates of the posterior predictive distribution of $PM_{2.5}$ for 28 June 2013. The units of the mean are given in $\mu g\,m^{-3}$ whereas the standard deviation units are on the log-scale. The dots on the standard deviation map are the locations of the monitoring stations for the observed data used in fitting the model.

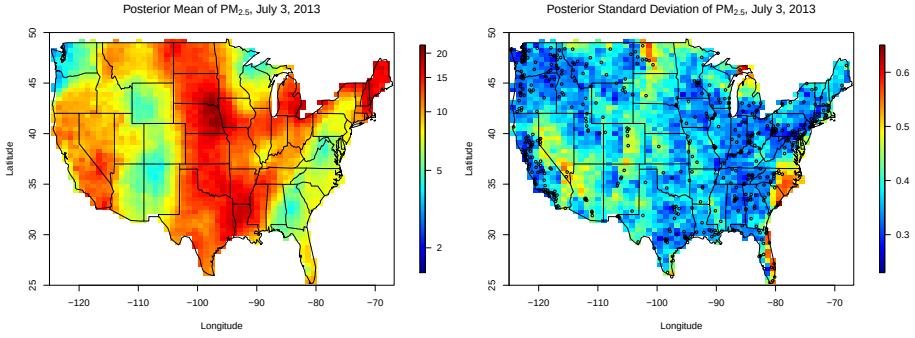

Figure 15. Mean and standard deviation estimates of the posterior predictive distribution of $PM_{2.5}$ for 3 July 2013. The units of the mean are given in $\mu g\,m^{-3}$ whereas the standard deviation units are on the log-scale. The dots on the standard deviation map are the locations of the monitoring stations for the observed data used in fitting the model.

Table 1. The posterior mean and 90 % credible interval for parameters.

Parameter	Mean	90 % credible interval
γ_1 (avg. daily temp)	4.03*	(3.25*, 4.92*)
γ_2 (avg. daily relative humidity)	0.07*	(−0.37*, 0.51*)
ρ	0.671	(0.660, 0.682)
C_{11}	0.405	(0.382, 0.435)
C_{21}	−0.038	(−0.094, −0.003)
C_{22}	0.098	(0.051, 0.184)
σ_P^2	0.035	(0.034, 0.037)
σ_A^2	0.325*	(0.320*, 0.330*)
κ	27.92	(27.85, 27.99)

$*$ denotes $\times 10^{-3}$.

stations used in model fitting, noting that the areas of elevated uncertainty correspond to regions with few or no stations. On 3 July, the posterior mean surface of $PM_{2.5}$ is quite different from that on 28 June while the standard deviation surface is quite similar. Here, estimates of $PM_{2.5}$ are greater than $10\,\mu g\,m^{-3}$ in much of the central United States. A similarity between the 2 days is that $PM_{2.5}$ estimates are lower in Washington than most of the country.

5 Summary and future work

A hierarchical autoregressive model with daily spatially varying coefficients was employed to jointly model consecutive day average $PM_{2.5}$ across the United States using AOT data obtained from VIIRS, previous day $PM_{2.5}$, and meteorological covariates. The model allows for missingness in AOT data and carries out model-based imputation for AOT at missing grid cells using a Gaussian Markov random field (GMRF). Several submodels are considered for quantifying any improvement in daily $PM_{2.5}$ prediction using AOT versus predictions based solely on ground-monitoring data.

Our analyses show small in-sample improvement incorporating AOT into our autoregressive model for daily $PM_{2.5}$ prediction during this time period, but little predictive out-of-sample improvement. Paciorek and Liu (2009) found similar results in that AOT does not improve the monthly and yearly prediction of $PM_{2.5}$. Thus, models without AOT perform almost as well as models that incorporate AOT.

We considered further modeling efforts involving multiple sources of $PM_{2.5}$ data. Daily FRM $PM_{2.5}$ data showed similar correlation levels with AOT as the continuous $PM_{2.5}$ data. Thus, there seemed little to be gained by incorporating another source of $PM_{2.5}$ data. Further, the models proposed here assume the independent error arises from the process model and does not account for measurement error in $PM_{2.5}$. While the model could be modified to account for this additional source of error, this would weaken the explanatory power of the model. We believe that the predictive performance of the models is not penalized by this simplification.

The significance of AOT in out-of-sample prediction may be suppressed by the fact that AOT is a vertically and spatially integrated measure while $PM_{2.5}$ is observed at point locations and, hence, reflects small-scale variability. One might expect more benefit from AOT when predicting $PM_{2.5}$ at the areal scale of the AOT data. Our objective here is point-level prediction using AOT. Areal $PM_{2.5}$ at continental scale with more than 40 000 AOT grid cells presents a computationally challenging upscaling problem. The spatial and temporal misalignment between $PM_{2.5}$ and AOT could also be a cause for low correlation and insignificance of AOT. We are investigating additional covariates, such as distance to nearest road, that might further explain local variability of $PM_{2.5}$ beyond the meteorological variables currently being considered.

Capturing the vertical profile of AOT is a critical component in evaluating the utility of AOT in predicting $PM_{2.5}$. We are exploring spatially and temporally varying adjustments for AOT that utilize both computer model output and covariate information about the vertical profile. It has been demonstrated that physical measures from lidar, or the use of atmospheric models, can be used to account for a portion of the vertical structure. This may aid in the use of AOT to better predict $PM_{2.5}$ in the future (Toth et al., 2014; Zeeshan and Oanh, 2014; Chu et al., 2015; Hutchison et al., 2008).

Due to the quality assurance protocol of AOT from VIIRS, there is a lot of missing AOT data, which may also contribute to the ineffectiveness of AOT in our model in terms of out-of-sample prediction. Improvement in AOT retrieval algorithms, particularly with regard to surface reflectivity, and improved cloud screening may lead to lower levels of missingness in the future. The GOES-R Advanced Baseline Imager (ABI) scheduled to launch in March of 2016 will provide an AOD data product at $2\,\text{km} \times 2\,\text{km}$ spatial resolution every minute of the continental United States (Huff et al., 2012). The GOES-R AOD will allow for hourly composites of AOD, which may help alleviate the missingness of data as seen with the current polar-orbiting instruments.

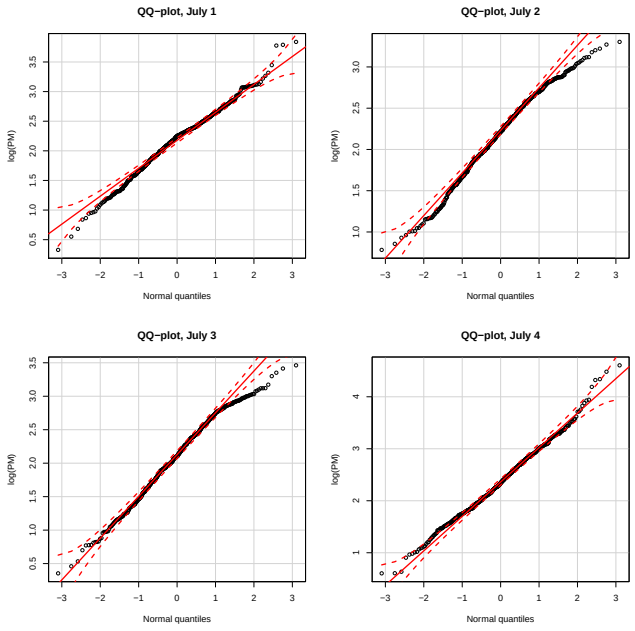

Figure A1. Normal Q–Q plots of the log-transformed $PM_{2.5}$ on 4 consecutive days.

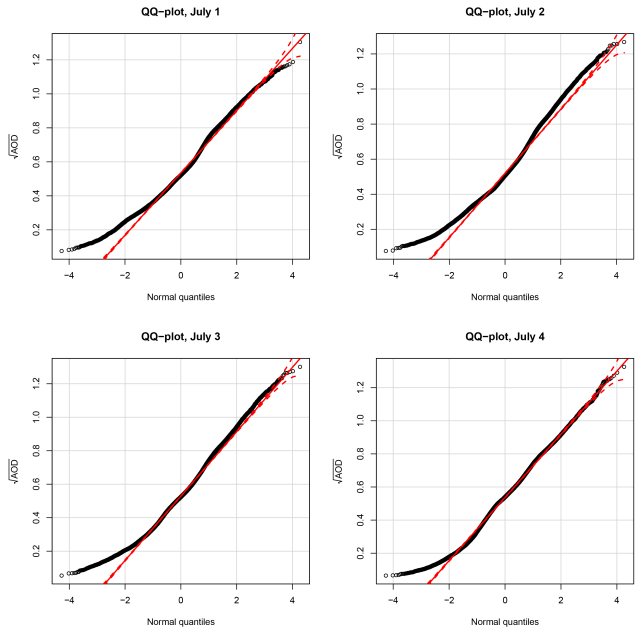

Figure A2. Normal Q–Q plots of the square root-transformed AOT on 4 consecutive days.

Appendix A

A1 Data transformations

Normal Q–Q plots of the log-transformation of $PM_{2.5}$ and square root transformation of AOT are shown for 4 days in Figs. A1 and A2, respectively. Both figures seem to suggest a slight derivation from normality although the deviations are minimal and the transformations agree with previous work in the literature (e.g., Berrocal et al., 2010; Paciorek and Liu, 2012).

A2 Posterior derivations

Full posterior inference is obtained using a hybrid Metropolis-within-Gibbs sampler. Let $\widetilde{A}_{i,t}$ denote AOD at grid cell i on day t. This value is either observed such that $\widetilde{A}_{i,t} = A_{i,t}$ or is predicted within the algorithm. Then, define $\mathbf{Z}_t = [1, \widetilde{\boldsymbol{A}}_t]$ and $\mathbf{K}_t = [\mathbf{I}, \mathrm{diag}(\widetilde{\boldsymbol{A}}_t)]$. Additionally, let $\boldsymbol{\alpha}_t = [\alpha_{t,0}, \alpha_{t,1}]$ and $\boldsymbol{\beta}_t = [\boldsymbol{\beta}_{t,0}, \boldsymbol{\beta}_{t,1}]$.

The full conditional distributions are given in details below.

– $\boldsymbol{\alpha}_0 | \cdot \mathrm{MVN}(\mathbf{U}^{-1}\mathbf{V}, \mathbf{U}^{-1})$,

where

$$\mathbf{U} = \mathbf{Z}_0' \Sigma_P^{-1} \mathbf{Z}_0 + \Sigma_\alpha^{-1},$$
$$\mathbf{V} = \mathbf{Z}_0' \Sigma_P^{-1} (\mathbf{P}_0 - \mathbf{X}_0 \boldsymbol{\gamma} - \mathbf{K}_0 \boldsymbol{\beta}_0),$$

$\Sigma_P = \sigma_P^2 \mathbf{I}$, and Σ_α is the prior covariance matrix of $\boldsymbol{\alpha}$.

– For $t = 1, \ldots, T, \boldsymbol{\alpha}_t | \cdot \mathrm{MVN}(\mathbf{U}^{-1}\mathbf{V}, \mathbf{U}^{-1})$,

where

$$\mathbf{U} = \mathbf{Z}_t' \Sigma_P^{-1} \mathbf{Z}_t + \Sigma_\alpha^{-1},$$
$$\mathbf{V} = \mathbf{Z}_t' \Sigma_P^{-1} (\mathbf{P}_t - \mathbf{X}_t \boldsymbol{\gamma} - \mathbf{K}_t \boldsymbol{\beta}_t - \rho \mathbf{P}_{t-1}).$$

– $\boldsymbol{\gamma} | \cdot \mathrm{MVN}(\mathbf{U}^{-1}\mathbf{V}, \mathbf{U}^{-1})$,

where

$$\mathbf{U} = \sum_{t=1}^{T} \mathbf{X}_t' \Sigma_P^{-1} \mathbf{X}_t + \Sigma_\gamma^{-1},$$
$$\mathbf{V} = \mathbf{X}_0' \Sigma_P^{-1} (\mathbf{P}_0 - \mathbf{Z}_0 \boldsymbol{\alpha}_0 - \mathbf{K}_0 \boldsymbol{\beta}_0) + \sum_{t=1}^{T} \mathbf{X}_t' \Sigma_P^{-1} (\mathbf{P}_t - \mathbf{Z}_t \boldsymbol{\alpha}_t - \mathbf{K}_t \boldsymbol{\beta}_t - \rho \mathbf{P}_{t-1}),$$

and Σ_γ is the prior covariance matrix of $\boldsymbol{\gamma}$.

– $\boldsymbol{\beta}_0 | \cdot \mathrm{MVN}(\mathbf{U}^{-1}\mathbf{V}, \mathbf{U}^{-1})$,

where

$$\mathbf{U} = \mathbf{K}_0' \Sigma_P^{-1} \mathbf{K}_0 + \Sigma_{\beta_0}^{-1},$$
$$\mathbf{V} = \mathbf{K}_0' \Sigma_P^{-1} (\mathbf{P}_0 - \mathbf{Z}_0 \boldsymbol{\alpha}_0 - \mathbf{X}_0 \boldsymbol{\gamma}),$$

and $\Sigma_{\beta_0} = [\mathbf{C} \otimes \mathbf{I}] \begin{bmatrix} \mathbf{R}_{0,0} & 0 \\ 0 & \mathbf{R}_{1,0} \end{bmatrix} [\mathbf{C}' \otimes \mathbf{I}]$.

– For $t = 1, \ldots, T, \boldsymbol{\beta}_t | \cdot \mathrm{MVN}(\mathbf{U}^{-1}\mathbf{V}, \mathbf{U}^{-1})$,

where

$$\mathbf{U} = \mathbf{K}_t' \Sigma_P^{-1} \mathbf{K}_t + \Sigma_{\beta_t}^{-1},$$
$$\mathbf{V} = \mathbf{K}_t' \Sigma_P^{-1} (\mathbf{P}_t - \mathbf{Z}_t \boldsymbol{\alpha}_t - \mathbf{X}_t \boldsymbol{\gamma} - \rho \mathbf{P}_{t-1}),$$
$$\text{and } \Sigma_{\beta_t} = [\mathbf{C} \otimes \mathbf{I}] \begin{bmatrix} \mathbf{R}_{0,t} & 0 \\ 0 & \mathbf{R}_{1,t} \end{bmatrix} [\mathbf{C}' \otimes \mathbf{I}].$$

– $\sigma_P^2 | \cdot \sim \text{I.G.}(a_P + \frac{(T+1)m}{2}$,
$b_P + \frac{1}{2}(\mathbf{P}_0 - \mathbf{Z}_0\boldsymbol{\alpha}_0 - \mathbf{X}_0\boldsymbol{\gamma} - \mathbf{K}_0\boldsymbol{\beta}_0)'$
$(\mathbf{P}_0 - \mathbf{Z}_0\boldsymbol{\alpha}_0 - \mathbf{X}_0\boldsymbol{\gamma} - \mathbf{K}_0\boldsymbol{\beta}_0)$
$+ \frac{1}{2}\sum_{t=1}^{T}(\mathbf{P}_t - \mathbf{Z}_t\boldsymbol{\alpha}_t - \mathbf{X}_t\boldsymbol{\gamma} - \mathbf{K}_t\boldsymbol{\beta}_t - \rho\mathbf{P}_{t-1})'$
$(\mathbf{P}_t - \mathbf{Z}_t\boldsymbol{\alpha}_t - \mathbf{X}_t\boldsymbol{\gamma} - \mathbf{K}_t\boldsymbol{\beta}_t - \rho\mathbf{P}_{t-1}))$,
where $m =$ number of monitoring stations.

– $\sigma_A^2 | \cdot \sim \text{I.G.}(a_A + \frac{\sum_{t=0}^{T} n_t}{2}$,
$b_P + \frac{1}{2}\sum_{t=0}^{T}(\boldsymbol{A}_t - \mathbf{M}_t\boldsymbol{\theta}_t)'(\boldsymbol{A}_t - \mathbf{M}_t\boldsymbol{\theta}_t))$,
where n_t is the number grid cells with AOD observations on day t.

– $\kappa | \cdot \sim \text{Gamma}(a_\kappa + \frac{(T+1)N}{2}$,
$b_\kappa + \frac{1}{2}\sum_{t=0}^{T}\boldsymbol{\theta}_t'\mathbf{Q}\boldsymbol{\theta}_t)$,
where $\mathbf{Q}$ is the second-order neighborhood matrix and N denotes the fixed number of grid cells in the spatial lattice for AOD.

– Sampling $(\phi_{0,t}, \phi_{1,t})$ for $t = 0, \ldots, T$ requires a Metropolis–Hastings algorithm. The candidate values, denoted $\phi_{0,t}^c$ and $\phi_{1,t}^c$ are accepted with probability

$$\min\left(1, \frac{\prod_{t=0}^{T} p(\boldsymbol{\beta}_t | \mathbf{C}, \phi_{0,t}^c, \phi_{1,t}^c)\pi(\phi_{0,t}^c, \phi_{1,t}^c) g(\phi_{0,t}, \phi_{1,t} | \phi_{0,t}^c, \phi_{1,t}^c)}{\prod_{t=0}^{T} p(\boldsymbol{\beta}_t | \mathbf{C}, \phi_{0,t}, \phi_{1,t})\pi(\phi_{0,t}, \phi_{1,t}) g(\phi_{0,t}^c, \phi_{1,t}^c | \phi_{0,t}, \phi_{1,t})}\right),$$

where $\pi(\cdot)$ denotes the prior distribution and $g(\cdot|\cdot)$ denotes the proposal distribution.

– Sampling $\mathbf{C}$ also requires a Metropolis–Hastings algorithm. The three parameters, C_{11}, C_{21}, and C_{22} are updated in a block where the candidate, denoted $\mathbf{C}^c = \begin{bmatrix} C_{11}^c & 0 \\ C_{21}^c & C_{22}^c \end{bmatrix}$ is accepted with probability

$$\min\left(1, \frac{\prod_{t=0}^{T} p(\boldsymbol{\beta}_t | \mathbf{C}^c, \phi_{0,t}, \phi_{1,t})\pi(\mathbf{C}^c) g(\mathbf{C} | \mathbf{C}^c)}{\prod_{t=0}^{T} p(\boldsymbol{\beta}_t | \mathbf{C}, \phi_{0,t}, \phi_{1,t})\pi(\mathbf{C}) g(\mathbf{C}^c | \mathbf{C})}\right).$$

– The autoregressive parameter, ρ, requires a Metropolis–Hastings step where ρ^c is accepted with probability

$$\min\left(1, \frac{\prod_{t=0}^{T} p(\boldsymbol{P}_t | \boldsymbol{\alpha}_t, \boldsymbol{\gamma}, \boldsymbol{\beta}_t, \tau_P^2, \widetilde{\boldsymbol{A}}_t, \rho^c)\pi(\rho^c) g(\rho | \rho^c)}{\prod_{t=0}^{T} p(\boldsymbol{P}_t | \boldsymbol{\alpha}_t, \boldsymbol{\gamma}, \boldsymbol{\beta}_t, \tau_P^2, \widetilde{\boldsymbol{A}}_t, \rho)\pi(\rho) g(\rho^c | \rho)}\right).$$

– Recall that we must infill AOD on the fly for grid cells that contain monitoring stations but are missing AOD. When AOD is observed at grid cell i on day t, $\widetilde{A}_{i,t} = A_{i,t}$. For each grid cell that is missing AOD but contains a monitoring station, we must predict AOD. Let $\{J_t\}$ denote the set of grid cells that contain a monitoring station but are missing AOD on day t. Further, for $j \in J_t$, let s_j denote the location of the monitoring station in the grid. Then, for each j, we sample $\widetilde{A}_{j,t}$ by sampling from the posterior predictive distribution. That is, for $t = 0$ and $j \in \{J_0\}$,

$$\widetilde{A}_{j,t} | \cdot \sim \text{N}(U^{-1}V, U^{-1}),$$

where

$$U = \frac{\alpha_1 + \beta_1(s_j)^2}{\sigma_T^2} + \frac{1}{\sigma_A^2},$$
$$V = \frac{\alpha_1 + \beta_1(s_j)}{\sigma_T^2}(P_0(s_j) - \alpha_0 - \boldsymbol{X}_0(s_j)\gamma - \beta_0(s_j)) + \frac{\theta_{j,0}}{\sigma_A^2}.$$

For $t = 1, \ldots, T$ and $j \in \{J_t\}$,

$$\widetilde{A}_{j,t} | \cdot \sim \text{N}(U^{-1}V, U^{-1}),$$

where

$$U = \frac{\alpha_1 + \beta_1(s_j)^2}{\sigma_T^2} + \frac{1}{\sigma_A^2},$$
$$V = \frac{\alpha_1 + \beta_1(s_j)}{\sigma_T^2}(P_t(s_j) - \alpha_0 - \boldsymbol{X}_t(s_j)\gamma - \beta_0(s_j) - \rho P_{t-1}(s_j)) + \frac{\theta_{j,t}}{\sigma_A^2}.$$

– If grid cell i contains a monitoring station,

$$\theta_{i,t} | \boldsymbol{\theta}_{j \neq i,t}, \cdot \sim N\left(\left(\frac{1}{\sigma_A^2} + \kappa p_i\right)^{-1}\left(\frac{1}{\sigma_A^2}\widetilde{A}_{i,t} + \bar{\theta}_{i,t}\kappa p_i\right), \left(\frac{1}{\sigma_A^2} + \kappa p_i\right)^{-1}\right),$$

where p_i is the number of neighbors of cell i and $\bar{\theta}_{i,t} = \frac{1}{p_i}\sum_{j \sim i}\theta_{j,t}$.
If grid cell i does not contain a monitoring station,

$$\theta_{i,t} | \boldsymbol{\theta}_{j \neq i,t}, \cdot \sim N\left(\bar{\theta}_{i,t}, \frac{1}{\kappa p_i}\right).$$

Acknowledgements. The work of the first author was supported in part by funding through the US EPA's Office of Research and Development using EPA contract EP-13-D-000257. The authors thank James Szykman of the National Exposure Research Laboratory, Environmental Science Division, Landscape Characterization Branch (US EPA) for providing the VIIRS satellite aerosol optical thickness (AOT) data product used in this analysis and valuable discussion that greatly improved this manuscript.

Disclaimer. The US Environmental Protection Agency through its Office of Research and Development partially collaborated in this research. Although it has been reviewed by the Agency and approved for publication, it does not necessarily reflect the Agency's policies or views.

Edited by: C. Paciorek
Reviewed by: two anonymous referees

References

Al-Hamdan, M. Z., Crosson, W. L., Limaye, A. S., Rickman, D. L., Quattrochi, D. A., Estes Jr., M. G., Qualters, J. R., Sinclair, A. H., Tolsma, D. D., Adeniyi, K. A., and Niskar, A. S.: Methods for characterizing fine particulate matter using ground observations and remotely sensed data: potential use for environmental public health surveillance, J. Air Waste Manage., 59, 865–881, 2009.

Berrocal, V. J., Gelfand, A. E., and Holland, D. M.: A bivariate space-time downscaler under space and time misalignment, Ann. Appl. Stat., 4, 1942–1975, 2010.

Berrocal, V. J., Gelfand, A. E., and Holland, D. M.: Space-Time Data Fusion Under Error in Computer Model Output: An Application to Modeling Air Quality, Biometrics, 68, 837–848, 2012.

Besag, J., York, J., and Mollié, A.: Bayesian image restoration, with two applications in spatial statistics, Ann. I. Stat. Math., 43, 1–20, 1991.

Chu, D. A., Ferrare, R., Szykman, J., Lewis, J., Scarino, A., Hains, J., Burton, S., Chen, G., Tsai, T., Hostetler, C., Hair, J., Holben, B., and Crawford, J.: Regional characteristics of the relationship between columnar AOD and surface $PM_{2.5}$: Application of lidar aerosol extinction profiles over Baltimore-Washington Corridor during DISCOVER-AQ, Atmos. Environ., 101, 338–349, 2015.

Dawson, J. P., Adams, P. J., and Pandis, S. N.: Sensitivity of $PM_{2.5}$ to climate in the Eastern US: a modeling case study, Atmos. Chem. Phys., 7, 4295–4309, doi:10.5194/acp-7-4295-2007, 2007.

Engel-Cox, J. A., Holloman, C. H., Coutant, B. W., and Hoff, R. M.: Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality, Atmos. Environ., 38, 2495–2509, 2004.

Flegal, J. M., Haran, M., and Jones, G. L.: Markov chain Monte Carlo: Can we trust the third significant figure?, Statistical Science, 23, 250–260, 2008.

Gneiting, T. and Raftery, A. E.: Strictly proper scoring rules, prediction, and estimation, J. Am. Stat. Assoc., 102, 359–378, 2007.

Hoff, R. M. and Christopher, S. A.: Remote sensing of particulate pollution from space: have we reached the promised land?, J. Air Waste Manage., 59, 645–675, 2009.

Hu, X., Waller, L. A., Lyapustin, A., Wang, Y., Al-Hamdan, M. Z., Crosson, W. L., Estes, M. G., Estes, S. M., Quattrochi, D. A., Puttaswamy, S. J., and Liu, Y.: Estimating ground-level $PM_{2.5}$ concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model, Remote Sens. Environ., 140, 220–232, 2014.

Huff, A. K., Hoff, R. M., Kondragunta, S., Zhang, H., Ciren, P., Xu, C., Christopher, S., Yang, E. S., and Szykman, J.: The NOAA air quality proving ground: Preparing the air quality community for next-generation products from the GOES-R satellite, Air and Waste Management Association, Pittsburgh, PA, Magazine for Environmental Managers, 32–37, 2012.

Hutchison, K. D., Faruqui, S. J., and Smith, S.: Improving correlations between MODIS aerosol optical thickness and ground-based $PM_{2.5}$ observations through 3D spatial analyses, Atmos. Environ., 42, 530–543, 2008.

Jackson, J. M., Liu, H., Laszlo, I., Kondragunta, S., Remer, L. A., Huang, J., and Huang, H.-C.: Suomi-NPP VIIRS aerosol algorithms and data products, J. Geophys. Res.-Atmos., 118, 12673–12689, 2013.

Kloog, I., Koutrakis, P., Coull, B. A., Lee, H. J., and Schwartz, J.: Assessing temporally and spatially resolved $PM_{2.5}$ exposures for epidemiological studies using satellite aerosol optical depth measurements, Atmos. Environ., 45, 6267–6275, 2011.

Kloog, I., Chudnovsky, A. A., Just, A. C., Nordio, F., Koutrakis, P., Coull, B. A., Lyapustin, A., Wang, Y., and Schwartz, J.: A new hybrid spatio-temporal model for estimating daily multi-year $PM_{2.5}$ concentrations across northeastern USA using high resolution aerosol optical depth data, Atmos. Environ., 95, 581–590, 2014.

Lee, H. J., Coull, B. A., Bell, M. L., and Koutrakis, P.: Use of satellite-based aerosol optical depth and spatial clustering to predict ambient $PM_{2.5}$ concentrations, Environ. Res., 118, 8–15, 2012.

Liu, H., Remer, L. A., Huang, J., Huang, H.-C., Kondragunta, S., Laszlo, I., Oo, M., and Jackson, J. M.: Preliminary evaluation of S-NPP VIIRS aerosol optical thickness, J. Geophys. Res.-Atmos., 119, 3942–3962, 2014.

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

Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korkin, S., Remer, L., Levy, R., and Reid, J.: Multiangle implementation of atmospheric correction (MAIAC) – 2. Aerosol algorithm, J. Geophys. Res.-Atmos., 116, D03211, doi:10.1029/2010JD014986, 2011.

Miller, K. A., Siscovick, D. S., Sheppard, L., Shepherd, K., Sullivan, J. H., Anderson, G. L., and Kaufman, J. D.: Long-term exposure to air pollution and incidence of cardiovascular events in women, New Engl. J. Med., 356, 447–458, 2007.

Paciorek, C. J. and Liu, Y.: Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter, Environ. Health Persp., 117, 904–909, 2009.

Paciorek, C. J. and Liu, Y.: Assessment and statistical modeling of the relationship between remotely sensed aerosol optical depth and $PM_{2.5}$ in the eastern United States, Tech. Rep. 167, Health Effects Institute, Boston, MA, 5–83, 2012.

Pope III, C. A. and Dockery, D. W.: Health effects of fine particulate air pollution: lines that connect, J. Air Waste Manage., 56, 709–742, 2006.

Pope III, C. A., Ezzati, M., and Dockery, D. W.: Fine-particulate air pollution and life expectancy in the United States, New Engl. J. Med., 360, 376–386, 2009.

Roy, B., Mathur, R., Gilliland, A. B., and Howard, S. C.: A comparison of CMAQ-based aerosol properties with IMPROVE, MODIS, and AERONET data, J. Geophys. Res.-Atmos., 112, D14301. doi:10.1029/2006JD008085, 2007.

Toth, T. D., Zhang, J., Campbell, J. R., Hyer, E. J., Reid, J. S., Shi, Y., and Westphal, D. L.: Impact of data quality and surface-to-column representativeness on the $PM_{2.5}$/satellite AOD relationship for the contiguous United States, Atmos. Chem. Phys., 14, 6049–6062, doi:10.5194/acp-14-6049-2014, 2014.

US Environmental Protection Agency: Particulate Matter Research Program, Five Years of Progress, Environmental Protection Agency, Office of Research and Development, Washington, DC 20460, EPA 600/R-04/058, 2004.

Valavanidis, A., Fiotakis, K., and Vlachogianni, T.: Airborne particulate matter and human health: toxicological assessment and importance of size and composition of particles for oxidative damage and carcinogenic mechanisms, J Environ. Sci. Heal. C, 26, 339–362, 2008.

van Donkelaar, A., Martin, R. V., Pasch, A. N., Szykman, J. J., Zhang, L., Wang, Y. X., and Chen, D.: Improving the accuracy of daily satellite-derived ground-level fine aerosol concentration estimates for North America, Environ. Sci. Technol., 46, 11971–11978, 2012.

Wang, J. and Christopher, S. A.: Intercomparison between satellite-derived aerosol optical thickness and $PM_{2.5}$ mass: implications for air quality studies, Geophys. Res. Lett., 30, 2095, doi:10.1029/2003GL018174, 2003.

Wang, Z., Liu, Y., Hu, M., Pan, X., Shi, J., Chen, F., He, K., Koutrakis, P., and Christiani, D. C.: Acute health impacts of airborne particles estimated from satellite remote sensing, Environ. Int., 51, 150–159, 2013.

Zeeshan, M. and Oanh, N. K.: Assessment of the relationship between satellite AOD and ground PM_{10} measurement data considering synoptic meteorological patterns and Lidar data, Sci. Total Environ., 473, 609–618, 2014.