Journal of Data Science **12**(2014), 255-275

On the Spatio-Temporal Relationship Between MODIS AOD and $PM_{2.5}$ Particulate Matter Measurements

Aaron T. Porter[1], Jacob J. Oleson[2], Charles O. Stanier[3]

Department of Statistics, The University of Missouri - Columbia.[1]

Department of Biostatistics, The University of Iowa.[2]

Department of Chemical and Biochemical Engineering, The University of Iowa.[3]

Abstract: Particulate matter smaller than 2.5 microns ($PM_{2.5}$) is a commonly measured parameter in ground-based sampling networks designed to assess short and long-term air quality. The measurement techniques for ground based $PM_{2.5}$ are relatively accurate and precise, but monitoring locations are spatially too sparse for many applications. Aerosol Optical Depth (AOD) is a satellite based air quality measurement that can be computed for more spatial locations, but measures light attenuation by particulates throughout in entire air column, not just near the ground. The goal of this paper is to better characterize the spatio-temporal relationship between the two measurements. An informative relationship will aid in imputing $PM_{2.5}$ values for health studies in a way that accounts for the variability in both sets of measurements, something physics based models cannot do. We use a data set of Chicago air quality measurements taken during 2007 and 2008 to construct a weekly hierarchical model. We also demonstrate that AOD measurements and a latent spatio-temporal process aggregated weekly can be used to aid in the prediction of $PM_{2.5}$measurements.

Key words: Air Quality, AOD, Bayesian, Hierarchical, Spatio-Temporal

1. Introduction

Particulate matter with aerodynamic diameter smaller than 2.5 microns ($PM_{2.5}$) has well known associations with many diseases. Accordingly, air quality studies are increasingly prominent in the public health literature. Air quality has been linked to various health outcomes, particularly cardiovascular and respiratory diseases (Dominici et al., 2006, Pope III et al., 2002) but also to other diseases as well (Pope et al., 2009). It is important to understand not only what the diseases are and what causes them, but the spatial and temporal patterns as well. Both spatial and temporal patterns of chronic diseases are informative to begin inferring casual factors of acute diseases.

Historically, studies have been time series studies which followed specific locations over a long period of time. See Dominici et al. (2002, 2003, 2000) for thorough examples. More recent studies have also been focusing on spatial relationships as well as temporal (Jerrett et al., 2005, Paciorek and Liu, 2009). The majority of such studies account for spatial relationships by aggregating across large geographic regions. Typically, the aggregation is necessary because the $PM_{2.5}$ measurements are ground-based and can only be obtained at a small number of monitoring stations. Therefore, it is difficult to produce full spatio-temporal models because the $PM_{2.5}$ stations are so sparse.

Since observed values of $PM_{2.5}$ at finer geographic scales are not readily available, some studies have examined the relationship between $PM_{2.5}$ and aerosol optical depth (AOD). Various sensors aboard many satellites orbiting earth have daily global coverage, and capture the data at various spatial and spectral resolutions. AOD is computed using a radiative transfer model (Remer et al., 2006) and is used by researchers in climate change and weather prediction. Since the Terra satellite has daily global coverage, AOD data recorded by MODerate Resolution Imaging Spectroradiometer (MODIS) aboard the satellite can be utilized to develop daily air quality estimates. Specifically, AOD is the depletion of reflected solar irradiation, caused by the presence of solid and liquid aerosols in the atmosphere, as measured by the satellites. The AOD value denotes the columnar presence of solid and liquid aerosols.

Only a fraction of the AOD value captures the airborne particulate matter (PM) near the earth surface (Kumar et al., 2011). Therefore, we expect AOD and PM values are moderately, but not perfectly, correlated. It is this correlation that is of particular interest in this paper. We know that the $PM_{2.5}$ values can be more finely estimated by using the information available within AOD (Paciorek et al., 2008). Our goal is to investigate the correlation over both space and time between AOD and $PM_{2.5}$ and determine how that correlation varies across both space and time after accounting for important covariates. Paciorek and Liu (2009) discuss the importance of comparing AOD and $PM_{2.5}$ and also the limitations. Given that we can't control for all factors influencing the differences between $PM_{2.5}$ and AOD (e.g., chemical composition, shape, mixing state, relative humidity, and size distribution), quantifying the spatial and temporal correlations between $PM_{2.5}$ values after accounting for AOD and other covariates would allow better predictions of $PM_{2.5}$ values in new locations.

Increasingly, AOD values are being used as a proxy to improve $PM_{2.5}$ predictions since PM values are geographically sparse (Liu et al., 2007, McMillan et al., 2009). In light of this research, a rigorous model of the spatial and temporal of AOD and $PM_{2.5}$ is desired. We describe the data being used in the analyses in Section 2. A Bayesian hierarchical model is built in Section 3 to show how much of the spatio-temporal pattern in PM is actually being described by AOD and other covariates. We close with a discussion in Section 4.

2. Data

$PM_{2.5}$ values are available using the Environmental Protection Agency (EPA) Air Quality System (AQS). We have utilized only the more accurate 24 hour measurements in our analysis, and have discarded the one hour measurements in order to obtain a more consistent outcome measure. The network is available in many urban areas and supplemental sites. In the Chicago area there are thirty-five $PM_{2.5}$ monitoring stations as shown in Figure 1. Data from this study are obtained as part of a larger data assimilation study. The particular data we analyze were collected between April 2007 and December 2008. In the larger study, $PM_{2.5}$ values are being predicted at 12 to 36 km horizontal resolution over the United States and at a 4km resolution for the Chicago area. Given the smaller scale being focused on in Chicago, the work in this paper focuses on the relationship between $PM_{2.5}$ and AOD and their spatio-temporal trends. Although the current study focuses on Chicago, the method could be generalized to any location with both particulate and AOD observations.

AOD data from MODIS onboard Terra satellite were used for these analyses. The Terra and Aqua satellites cover the United States approximately twice per day. The NASA aerosol land team developed an algorithm to extract aerosol over land and ocean (Levy et al., 2007, Remer et al., 2006). The algorithm allows for daily global AOD values at various spatial resolutions and are available through Level 1 and Atmosphere Archive Distribution System (LAADS) (NASA, 2010). MODIS data are useful due to their fine spatial resolutions (here we use a 1000m resolution across 36 channels) and ensure robust retrieval of AOD at 10 km (Kumar et al., 2011, Levy et al., 2007). As shown in Figure 1, measurements are obtained for the entire Chicago area at least once during the study period.

Since one degree latitude is not equivalent in terms of distance to one degree longitude at the latitude of Chicago, all latitude and longitude coordinates were converted to Universal Transverse Mercator (UTM) coordinates, using a standard mathematical approximation by the United States Department of the Army (1973). UTM coordinates are Eastings and Northings, and are measured in meters. One Easting is equal in distance to one Northing anywhere on the globe. In order to prepare the data for use, we selected a 6,000 meter distance window for each location. We then averaged all of the $PM_{2.5}$ and AOD data available for this window within each calender week. This step is necessary because AOD and $PM_{2.5}$ are never measured at the same location at the same time point. The distance of 6,000 meters and time of one week was selected in order to provide an acceptable level of required imputation. After these procedures we proceeded with our data analysis.

Figure 2 contains some descriptive plots of the data. After a square root transformation, $PM_{2.5}$ and AOD the residuals of a linear regression given the covariates of the joint model in Section 3.1 are normally distributed, as confirmed by an Anderson-Darling normality test. We see that ${PM_{2.5}}^{1/2}$ is linearly related to $AOD^{1/2}$. We also see from the graphs of $PM_{2.5}$ and AOD over time by location that a separable model is sufficient to fit these data. This makes sense *a priori*, as the area considered is small enough that there is likely to be little interaction between space and time, especially after we account for covariates.

3. The Hierarchical Model

The goal of this section is to construct a Bayesian hierarchical model to assess the degree of large scale correlation between the two measurements after adjusting for additional predictor variables. To that end, we will model the $PM_{2.5}$ values using AOD values as a predictor variable. Our primary goal is to develop a model that will allow a response surface capable of predicting $PM_{2.5}$ at locations where there are no monitoring stations. We then use the latent spatio-temporal process to increase the accuracy of our measurements.

A joint Bayesian hierarchical model was built to analyze the relationship between AOD and $PM_{2.5}$. We utilize Bayesian methodology, due to the natural way latent spatial and temporal effects can be included in the model. These effects can then be used to generate an explicit understanding of the amount of spatial and temporal dependence through the magnitude of the latent random effects and via the parameters in their priors. Our model is unique in that we model AOD and $PM_{2.5}$ jointly. We are unaware of any research in the literature to perform this joint modeling.

For the sake of comparison, we also construct a vector autoregessive (VAR) model for these data. Our goal in our joint model is to demonstrate that AOD values can be used to predict $PM_{2.5}$ values at locations where there is no $PM_{2.5}$ monitoring station. However, AOD values are noisy, and so it is desirable to compare the performance of the model in terms of predicting AOD and $PM_{2.5}$ based on both sets of data. A VAR model, which we perform using standard frequentist analyses, helps to relieve the issue of predicting the more precise $PM_{2.5}$ values based on noisy AOD values.

3.1 The Hierarchical Data Model

The model we utilized can be written as

$$\begin{aligned} \sqrt{AOD} &= X_{AOD}\beta_{AOD} + \omega_1(s) + \tau_1(t) + \epsilon_1(s,t) \\ \sqrt{PM_{2.5}} &= X_{PM}\beta_{PM} + \omega_2(s) + \tau_2(t) + \epsilon_2(s,t). \end{aligned} \tag{1}$$

In the AOD portion of the joint model, X_{AOD} is a matrix containing an intercept, a factor variable with information on the season, and a factor variable containing data on the land use at the location being analyzed. The season factor variable consists of three indicators for Winter, Spring and Summer, with Autumn serving as the baseline season. The land use factor consists of two indicators for residential and industrial land use, with commercial land use serving as the baseline land use. The remaining terms are all considered random as $\omega_1(s)$ is a latent spatial effect, $\tau_1(t)$ is a latent temporal effect, and $\epsilon_1(s,t)$ is a pure error term. These errors are considered independent and identically distributed. An overview of spatio-temporal methods can be found in Cressie and Wikle (2011). Due to many factors, including exploratory plots of the data and a relatively small scale both geographically and temporally, an additive model was deemed reasonable.

In the $PM_{2.5}$ portion of the joint model, X_{PM} is a matrix containing an intercept and a column corresponding to $AOD^{1/2}$. Once again, $\omega_2(s)$ is a latent spatial effect, $\tau_2(t)$ is a latent temporal effect, and $\epsilon_2(s,t)$ is a pure error term. Exploratory plots again yielded no evidence of spatio-temporal interaction which is further supported by the results of the model.

This model directly allows large scale analysis of the correlation between $PM_{2.5}$ and AOD data by allowing correlations of the full AOD data and full $PM_{2.5}$ data to be computed. It also quantifies how much of the relationship can be modeled with latent spatial and temporal terms. For our particular data set, weekly averages of $PM_{2.5}$ are easily calculated from the daily $PM_{2.5}$ values. The vast majority of air quality measurements are at weekly averaging times or shorter.

3.2 Hierarchical Model Priors

In a Bayesian hierarchical model, care must be taken in selecting appropriate prior distributions for the model parameters. Conjugate priors were selected because of natural interpretations and they make computations more efficient. The beta coefficients were assigned normal priors such that $\beta_{AOD} \sim N(0, \sigma^2_{\beta_{AOD}} I_6)$ and $\beta_{PM} \sim N(0, \sigma^2_{\beta_{PM}} I_2)$. The corresponding variance parameters were assigned conjugate inverse gamma priors, $\sigma^2_{\beta_{AOD}} \sim IG(1,1)$ and $\sigma^2_{\beta_{PM}} \sim IG(1,1)$. While these inverse gamma priors are informative, they are diffuse enough that the mean of the distribution is undefined.

A graph of $PM_{2.5}$ with respect to AOD by location yielded evidence that a separable covariance structure was sufficient for these data. The latent spatial effects, $\omega_1(s)$, should capture any spatial correlation remaining in $AOD^{1/2}$ after accounting for the covariates in X_{AOD}. The correlation between AOD observations should decay with increasing separation so we chose an exponential decay spatial prior of the form $N(0, \sigma^2_{\omega_1}[exp(-\rho_1 d_{ij}]))$. The variance $\sigma^2_{\omega_1}$ controls the amount of spatial variability and ρ_1 is the decay parameter. For the remaining parameters, $\sigma^2_{\omega_1}$ was assigned an $IG(1, 1)$ prior, and ρ_1 a $Unif(0, 1)$ prior. This uniform prior allows for extreme spatial dependence, or near spatial independence, with the latter occurring due to the distances between monitoring stations all being greater than 500 meters, and typically several thousand meters. When ρ is large, these distances overwhelm the correlation structure and are able to produce nearly independent measurements. In general, 95% of the spatial autocorrelation has decayed at a distance of $d^* = -\log(0.05)/\rho$.

Similarly, $\omega_2(s)$ was assigned an exponential decay spatial prior, with variance $\sigma^2_{\omega_2}$ and decay parameter ρ_3 with $\sigma^2_{\omega_1}$ being assigned an $IG(1, 1)$ prior and $\rho_3 \sim Unif(0, 1)$. Thus, these parameters quantify the amount of spatial correlation remaining in ${PM_{2.5}}^{1/2}$ after adjusting for AOD and other important covariates.

In this model, the latent temporal effects, $\tau_1(t), t = 1, \ldots, 93$, capture the temporal correlation in $AOD^{1/2}$ not accounted for in X_{AOD}. We assigned an autoregressive prior of order one (AR(1)). Because there is no previous temporal data, the latent temporal effect for the first week, $\tau_1(1)$, was assigned a $N(0, \sigma^2_{\tau_1})$ prior. Then, the remaining weeks in the AR(1) model, $t = 2, \ldots, 93$, are written as

$$\tau_1(t) = \rho_2 \tau_1(t - 1) + \epsilon(t)$$

for $t = 2, \ldots, 93$. Note that the correlation parameter ρ_3 captures the dependence between successive weeks. Also, $\epsilon(t) \sim N(0, \sigma^2_{\tau_1})$ and the variance parameter was further assigned the inverse gamma prior, $\sigma^2_{\tau_1} \sim IG(1, 1)$. Finally, we assigned ρ_2 a $Unif(-1, 1)$ prior.

Similarly, $\tau_2(t), t = 2, ..., 93$, was assigned an AR(1) prior, with variance $\sigma^2_{\tau_2}$ and correlation parameter ρ_4. Consequently, $\sigma^2_{\tau_2}$ was assigned an IG(1,1) prior, and ρ_4 a Unif(-1,1) prior. Again, the latent temporal effect for the first week was assigned a Normal(0,$\sigma^2_{\tau_2}$) prior.

3.3 Posterior Computation

We note the similarity of our model with that of Oleson and He (2004). Our model uses a bivariate normal outcome, rather than a binomial outcome, but is similar in structure to their model. Additionally, Oleson and He (2004) notes that this prior gives a closed form for the inverse of the covariance matrix. We denote this matrix by A, and note that it is of the form $\sigma^2_{\omega_2}/(1 - \rho_2^2)\Sigma$ where Σ is tridiagonal. The diagonal elements are (1, 1+ρ_2^2,..., 1+ρ_2^2, 1) and the elements immediately above an below the diagonal are $-\rho_2$. This form makes the matrix convenient in a Gibbs step. We assume that, given the fixed effects, all random effects have mean zero. More details on these distributions can be found in Appendix A. For distributions which do not have a closed form, slice sampling was used in order to draw from the full conditional distribution of the parameter.

Imputation of missing $PM_{2.5}$ values and AOD values was performed via Metropolis sampling. Imputation was performed as part of the MCMC chain, and both AOD and $PM_{2.5}$ values were imputed based on the current state of the model at each iteration. Only nine of the thirty-five locations have both AOD and $PM_{2.5}$ values within 6,000 meters of each other during a week within the study period. The data are broken into 93 measurements, each corresponding to a calender week, at each of 9 monitoring stations. We model AOD using weekly averages. This requires imputation of 29.2% of the values.

Because we are modeling these measures jointly, it is convenient to propose new values for AOD and $PM_{2.5}$ measurements using a Gaussian proposal, and then to compute the joint likelihood in order to determine acceptance or rejection. The acceptance rate for AOD and $PM_{2.5}$ imputations were tuned to be near 40%.

3.4 Model Results

Three MCMC chains were run, and 400,000 iterations were collected after a burn-in period of 100,000 iterations for each chain. Convergence was assessed based on the Gelman-Rubin diagnostic (Gelman and Rubin, 1992). The full results for the AOD parameters can be found in Table 1, and the full results from the $PM_{2.5}$ parameters can be found in Table 2.

The model predicting AOD shows some evidence of seasonality, with the majority of the posterior masses of the spring and summer parameters being positive, but there is little evidence for the importance of land use in the AOD model. Additionally, the latent spatio-temporal process places considerably more mass on a positively autocorrelated model than a negatively autocorrelated model. This indicates potential positive temporal association in predicting AOD, but also demonstrates that the temporal latent process is noisy. The model also puts more mass on spatial decay parameters that suggest a latent spatial effect. This would suggest that, at any given time point, prediction of AOD values can be improved by accounting for the spatial structure in the data.

The model predicting $PM_{2.5}$ shows evidence of a relationship between AOD and $PM_{2.5}$ at all temporal aggregation levels. Figure 3 shows four realizations of the model predictions of ${PM_{2.5}}^{1/2}$ graphed with respect to the true values of ${PM_{2.5}}^{1/2}$. We see a strong linear trend in these data. The posterior root mean squared error for these data is 0.197, which is low compared to magnitude of ${PM_{2.5}}^{1/2}$ (the mean being 3.30), especially considering that the model is trying to predict a quantity as variable as $PM_{2.5}$. Plotting four sets of model-based predictions for ${PM_{2.5}}^{1/2}$ by the true ${PM_{2.5}}^{1/2}$ values, as was done in Figure 3, shows that our model accurately recovers the true values. This is promising, because it implies that future air quality studies that collect sparse air quality data can use weekly aggregation and still adequately predict $PM_{2.5}$ values based on AOD.

The model has failed to detect any meaningful latent temporal trend in $PM_{2.5}$ after accounting for AOD, with the autocorrelation varying throughout the entire domain of its prior. This would argue that the temporal process is too noisy at a given location to detect any meaningful structure. However, the model predicts a latent spatial process in the $PM_{2.5}$ values even after accounting for AOD. Spatial correlation matrices can be found in Appendix B. These results yield evidence that a spatial model is useful in kriging $PM_{2.5}$ values, and that the spatial process of AOD is not the only spatial process important in modeling $PM_{2.5}$. A plot of the posterior densities for the spatio-temporal dependency parameters can be found in Figure 4. It is worth noting, for clarity, that values of ρ_1 and ρ_3 nearer to zero are more significant than larger values, based on the parameterization of the model.

4. The VAR Model

We perform a VAR analysis for comparison. The general model we utilize is:

$$\sqrt{AOD_{i,t}} = \beta_{0i} + \sum_{j=1}^{9}\sum_{l=1}^{3}\{\beta_{1ijl}\sqrt{AOD_{j,t-l}}\} + \sum_{j=1}^{9}\sum_{l=1}^{3}\{\beta_{2ijl}\sqrt{PM2.5_{i,j,t-l}}\} + \epsilon_{1it}$$

$$\sqrt{PM2.5_{i,t}} = \beta_{3i} + \sum_{j=1}^{9}\sum_{l=1}^{3}\{\beta_{4ijl}\sqrt{AOD_{j,t-l}}\} + \sum_{j=1}^{9}\sum_{l=1}^{3}\{\beta_{5ijl}\sqrt{PM2.5_{i,j,t-l}}\} + \epsilon_{2it},$$

for $i = 1, ...9$, $t = 2, ..., 93$. All terms in this model are scalars. This model represents an alternative joint model for $PM_{2.5}$ and AOD. Notably, it does not contain any terms inducing concurrent spatial dependence, instead relying on the model to evolve temporally. We have performed the standard imputation of replacing any missing values with the mean value of the outcome. Additionally, we have not included land use in predicting AOD, as it was used only in predicting the imputed values of AOD. We assume that the random errors $\boldsymbol{\epsilon_{1i.}} \sim N(0, \sigma_{1i}^2)$ and that $\boldsymbol{\epsilon_{2i.}} \sim N(0, \sigma_{2i}^2)$. This model was estimated using the fastVAR package in the *R* programming language (Wong, 2012).

Due to the large number of parameters to be estimated, we suppress the full set of estimates and confidence intervals, and instead focus on model fit. The selection of an order three VAR model was performed by comparing all computable order and utilizing the Bayesian Information Criterion (BIC) (Schwarz, 1978) as a model selection criterion. BIC itself was chosen as a suitable model selection criterion because the previous model was performed in the Bayesian framework. The RMSE of this model was estimated to be 0.422, which is much larger than our hierarchical model. Our explanation for this difference is that weekly aggregation is coarse for air quality data, and the simultaneous spatial structure in our hierarchical model is advantageous. The VAR model may be a reasonable model when time is more finely partitioned.

5. Discussion

Our weekly model detected a strong latent spatial process in $PM_{2.5}$ even after accounting for AOD, indicating systematic spatial variability of $PM_{2.5}$ around a mean structure containing only AOD. This latent structure is helpful in kriging $PM_{2.5}$ values when AOD is aggregated weekly. Further refinement of the model utilizing additional covariates in the $PM_{2.5}$ model may allow for improved kriging, but we have shown that AOD is not sufficient to explain the spatial pattern in $PM_{2.5}$ for this particular data set.

The advantage of using statistical models is that they are able to directly incorporate variance components to model the uncertainty in the physical process. This ability is lacking in physics based models. Due to the high variance of these measurements, some researchers may choose statistical methods rather than physics based models. Since daily $PM_{2.5}$ predictions based on AOD are feasible, these predictions would be helpful in future air quality studies.

Our particular model is suitable for small areas over long time frames. The model of Paciorek and Liu (2009) is potentially better suited to larger areas. Its general additive structure allows complex splines to be fit over large areas and this gives the model additional flexibility, but these tend to hide the generating mechanism of the data. Our model utilizes less complex spatial smoothing, but is easily interpretable and informative with regards to the underlying spatio-temporal process of the data. We do admit that our structure may not scale up to large areas. For instance, the model would not be suitable for the entire United States, due to the separability we have utilized. However, our model is general enough that it may be applicable to small geographical scales in many location other than the Chicago area.

6. Acknowledgements

This research was supported, in part, by grant RD-83386501-0 from the Environmental Protection Agency. The authors would like to thank the reviewer for the their helpful comments which greatly improved this paper. The authors are particularly grateful to Naresh Kumar, Environmental Health in Epidemiology and Public Health, University of Miami, Miami, FL for providing the expertise

needed to acquire the $PM_{2.5}$ and AOD data.

7. Appendix A: Full Conditional Distributions

We introduce the following notation: X_s is a matrix indexing spatial locations, and X_t is a matrix indexing temporal periods. The correlation matrix will be denoted by Φ when there is no need to refer to the individual elements. We provide only the posterior distributions of the model component related to AOD due to the similarity with the posteriors of the model component related to $PM_{2.5}$. Each distribution listed is conditioned on the data and other parameters.

i. $\omega_1(s) \sim N_I(\mu, (\frac{1}{\sigma_\epsilon^2} X_s' X_s + \frac{1}{\sigma_{\omega_1}^2} \Phi^{-1}))^{-1})$

where $\mu = (\frac{1}{\sigma_\epsilon^2} X_s' X_s + \frac{1}{\sigma_{\omega_1}^2} \Phi^{-1})^{-1} \frac{1}{\sigma_\epsilon^2} X_s'(AOD - X_{AOD}\beta_{AOD} - X_t \tau_1)$

ii. $\tau_1(t) \sim N_J(\mu, (\frac{1}{\sigma_\epsilon^2} X_t' X_t + \frac{1-\rho_2^2}{\sigma_{\tau_1}^2} A^{-1}))^{-1} \frac{1}{\sigma_\epsilon^2}(AOD - X_{AOD}\beta_{AOD} - X_s \omega_1)$

where $\mu = ((\frac{1}{\sigma_\epsilon^2} X_t' X_t + \frac{1-\rho_2^2}{\sigma_{\tau_1}^2} A^{-1}))^{-1} \sigma_\epsilon^2 X_t'(AOD - X_{AOD}\beta_{AOD} - X_s \omega_1)$

iii. $\sigma^2 \sim IG(\frac{IJ}{2} + a_0, \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J (AOD_{ij} - \omega_i - \tau_j)^2 + b_0)$

iv. $\sigma_{\omega_1}^2 \sim IG(\frac{I}{2} + a_1, \frac{1}{2} \omega' \Phi \omega + b_1)$

v. $\sigma_{\tau_1}^2 \sim IG(\frac{J}{2} + a_1 \frac{1}{2} \tau' A' \Omega^{-1} A \tau + b_2)$

vi. $\rho_1 \propto \frac{|\Phi|^{\frac{1}{2}}}{\sigma_{\omega_1}^I} \exp(-\frac{1}{2\delta_1} \omega' \Phi \omega)$

vii. $\rho_2 \propto \exp\left(\frac{1}{2}\log(1-\rho_2^2) - \frac{1}{2\sigma_{\tau_1}^2}\left\{\sum_{j=1}^J \tau_j^2 - 2\rho_{32} \sum j = 2^J \tau_{j-1} \tau j + \rho_2^2 \sum_{j=2}^{J-1} \theta_j^2\right\}\right)$

viii. $\beta_{AOD} \sim N_6(\mu, (\frac{1}{\sigma_\epsilon^2} X' X + \frac{1}{\sigma_{beta}^2} I_6)^{-1})$

where $\mu = (\frac{1}{\sigma_\epsilon^2} X_{AOD}' X_{AOD} + \frac{1}{\sigma_{beta}^2} I_6)^{-1} \frac{1}{\sigma \epsilon^2} X'(AOD - X_s \omega - X_t \tau)$

ix. $\sigma_\beta^2 \sim IG(\frac{6}{2} + a_0, \frac{1}{2}(AOD - X_s \omega - X_t \tau)'(AOD - X_s \omega - X_t \tau) + b_0)$

8. Appendix B: Correlation Matrices

In this section we present data-based and model based spatial correlation matrices for the square root of AOD and the square root of $PM_{2.5}$ for exploratory purposes and for model comparison. It is worth noting that the data-based spatial correlation matrices only utilize the 37 time points where every location had a measured AOD and $PM_{2.5}$ values. Still, they offer some information as to the general appropriateness of the model and the importance of the model-based estimation. Model-based estimates are based on the posterior median of the spatial correlation parameters for AOD and $PM_{2.5}$ in order to provide a single measure of the correlation, but the model utilizes the entire posterior of these parameters.

The data-based correlation matrix for the square root of AOD is

$$\begin{bmatrix} 1.00 & 0.86 & 0.57 & 0.86 & 0.75 & 0.65 & 0.65 & 0.64 & 0.68 \\ & 1.00 & 0.51 & 0.79 & 0.87 & 0.74 & 0.71 & 0.73 & 0.75 \\ & & 1.00 & 0.47 & 0.67 & 0.33 & 0.46 & 0.53 & 0.54 \\ & & & 1.00 & 0.71 & 0.68 & 0.62 & 0.73 & 0.69 \\ & & & & 1.00 & 0.64 & 0.65 & 0.74 & 0.77 \\ & & & & & 1.00 & 0.65 & 0.65 & 0.64 \\ & & & & & & 1.00 & 0.67 & 0.62 \\ & & & & & & & 1.00 & 0.90 \\ & & & & & & & & 1.00 \end{bmatrix}.$$

The model-based correlation matrix for the square root of AOD is

$$\begin{bmatrix} 1.00 & 0.83 & 0.84 & 0.88 & 0.84 & 0.71 & 0.76 & 0.79 & 0.78 \\ & 1.00 & 0.87 & 0.92 & 0.97 & 0.85 & 0.91 & 0.94 & 0.91 \\ & & 1.00 & 0.85 & 0.91 & 0.78 & 0.84 & 0.89 & 0.89 \\ & & & 1.00 & 0.91 & 0.79 & 0.83 & 0.86 & 0.84 \\ & & & & 1.00 & 0.84 & 0.90 & 0.95 & 0.92 \\ & & & & & 1.00 & 0.93 & 0.88 & 0.89 \\ & & & & & & 1.00 & 0.95 & 0.95 \\ & & & & & & & 1.00 & 0.97 \\ & & & & & & & & 1.00 \end{bmatrix}.$$

We note two aspects regarding the previous set of matrices. First of all, the entire data-based correlation matrix is positive. Therefore, positive spatial autocorrelation matrices, such as the exponential model we have selected, are a reasonable choice in modeling this correlation. Secondly, the data-based correlation does not account for covariates or missing data, which is a strength of the model-based correlation matrix. Even so, both correlation matrices demonstrate a strong spatial process that underlies the AOD values.

The data-based correlation matrix for the square root of $PM_{2.5}$ is

$$\begin{bmatrix} 1.00 & 0.66 & 0.79 & 0.81 & 0.62 & 0.59 & 0.73 & 0.67 & 0.80 \\ & 1.00 & 0.58 & 0.69 & 0.75 & 0.64 & 0.72 & 0.62 & 0.73 \\ & & 1.00 & 0.73 & 0.74 & 0.53 & 0.69 & 0.71 & 0.80 \\ & & & 1.00 & 0.68 & 0.59 & 0.81 & 0.75 & 0.82 \\ & & & & 1.00 & 0.61 & 0.72 & 0.67 & 0.78 \\ & & & & & 1.00 & 0.75 & 0.64 & 0.66 \\ & & & & & & 1.00 & 0.75 & 0.75 \\ & & & & & & & 1.00 & 0.77 \\ & & & & & & & & 1.00 \end{bmatrix}.$$

The model-based correlation matrix for the square root of $PM_{2.5}$ after removing AOD is

$$\begin{bmatrix} 1.00 & 0.03 & 0.04 & 0.10 & 0.04 & 0.01 & 0.01 & 0.02 & 0.01 \\ & 1.00 & 0.10 & 0.20 & 0.55 & 0.05 & 0.17 & 0.33 & 0.19 \\ & & 1.00 & 0.05 & 0.16 & 0.01 & 0.04 & 0.11 & 0.11 \\ & & & 1.00 & 0.16 & 0.01 & 0.04 & 0.07 & 0.04 \\ & & & & 1.00 & 0.04 & 0.14 & 0.37 & 0.23 \\ & & & & & 1.00 & 0.27 & 0.10 & 0.11 \\ & & & & & & 1.00 & 0.37 & 0.37 \\ & & & & & & & 1.00 & 0.57 \\ & & & & & & & & 1.00 \end{bmatrix}.$$

The previous two correlation matrices are not directly comparable, as there is no data-based correlation matrix that can account for $PM_{2.5}$. However, it is clear that there is a strong spatial process in $PM_{2.5}$, and that some of this process can be modeled by a using AOD as a fixed effect. However, the model-based $PM_{2.5}$ correlation matrix still yields certain location pairs with noticeable latent correlation (such as locations 8 and 9). This provides evidence for the importance of including a latent spatial process at this model stage.

References

Cressie, N. and Wikle, C. K. (2011). *Statistics for Spatio-Temporal Data*, 46-54. Wiley, Hoboken, NJ.

Department of the Army (1973). *Universal transverse mercator grid, U. S. Army Technical Manual TM 5-241-8.*

Dominici, F., McDermoot, A., Zeger, S., and Samet, J. (2002). Airborne particulate matter and mortality: Time-scale effects in four u.s. cities. *American Journal of Epidemiology*, **157**:1053-1063.

Dominici, F., McDermott, A., Daniels, M., Zeger, S., and Samet, J. (2003). A special report to the health effects institute on the revised analyses of the nmmaps ii data. *The Health Effects Institute.* Cambridge, MA.

Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L., and Samet, J. M. (2006). Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *Journal of the American Medical Association*, 295:1127-1134.

Dominici, F., Samet, J., and Zeger, S. (2000). Combining evidence on air pollution and daily mortality from the twenty largest us cities: A hierarchical modeling strategy (with discussion). *Royal Statistical Society, Series A*, **163**:263-302.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**:457-472.

Jerrett, M., Burnett, R. T., Ma, R., Pope, C. A. I., Krewski, D., Newbold, K. B., Thurston, G., Shi, Y., Finkelstein, N., Calle, E. E., and Thun, M. J. (2005). Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology*, **16**:727-736.

Kumar, N., Chu, A. D., Foster, A. D., Peters, T., and Willis, R. (2011). Satellite remote sensing for developing time-space resolved estimates of ambient particulate in Cleveland, OH. *Aerosol Science and Technology*, **45**:1090-1108.

Levy, R. C., Remer, L. A., Mattoo, S., Vermote, E. F., and Kaufman, Y. J. (2007). Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of moderate resolution imaging spectroradiometer spectral reflectance. *Journal of Geophysical Research-Atmospheres*, **112**:D13211; doi:10.1029/2006JD007811.

Liu, Y., Franklin, M., Kahn, R., and Koutrakis, P. (2007). Using aerosol optical thickness to predict ground-level $pm_{2.5}$ concentrations in the St. Louis area: a comparison between MISR and MODIS. *Remote Sensing of Environment*, **107**:33-44.

McMillan, N. J., Holland, D. M., Morara1, M., and Feng1, J. (2009). Combining numerical model output and particulate data using Bayesian space-time modeling. *Environmetrics*, **21**:48-65.

NASA (2010). The level 1 and atmosphere archive and distribution system. *http://labsweb.nascom.nasa.gov/*.

Oleson, J. J. and He, C. Z. (2004). Space-time modeling for the missouri turkey hunting survey. *Environmental and Ecological Statistics*, **11**:85-101.

Paciorek, C. J. and Liu, Y. (2009). Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter. *Environmental Health Perspectives*, **117**:904-909.

Paciorek, C. J., Liu, Y., Moreno-Macias, H., and Shobha, K. (2008). Spatiotemporal associations between GOES aerosol optical depth retrievals and ground-level $PM_{2.5}$. *Environmental Science and Technology*, **42**:5800-5806.

Pope, C., Ezzati, M., and Dockery, D. (2009). Fine-particulate air pollution and life expectancy in the united states. *New England Journal of Medicine*, **360**:376-386.

Pope III, C., Burnett, R., Thun, M., Calle, E., Krewski, D., Ito, K., and Thurston, G. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA: *the journal of the American Medical Association*, **287**:1132-1141.

Remer, L. A., Tanre, D., and Kaufman, Y. J. (2006). *Algorithm for remote sensing of tropospheric Aerosol from MODIS*. NASA, Greenbelt, MD: Goddard Space Flight Center.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**:461-464.

Wong, J. (2012). *fast*VAR: *fast*VAR. R package version 1.2.1.19

Table 1: 95% Credible Intervals for the AOD Parameters

Parameters	Posterior Median	95% Credible Interval
Intercept ($\beta_{0_{AOD}}$)	0.171	(-0.339, 0.806)
Winter Season ($\beta_{1_{AOD}}$)	0.161	(-0.125, 0.397)
Spring Season ($\beta_{2_{AOD}}$)	0.209	(0.005, 0.380)
Summer Season ($\beta_{3_{AOD}}$)	0.176	(-0.058, 0.329)
Landuse Industrial ($\beta_{4_{AOD}}$)	0.059	(-0.290, 0.501)
Landuse Residential ($\beta_{5_{AOD}}$)	-0.035	(-0.336, 0.394)
Spatial Variance ($\sigma^2_{\omega_1}$)	0.237	(0.104, 0.738)
Temporal Variance ($\sigma^2_{\tau_1}$)	0.037	(0.027, 0.054)
Pure Error Variance (σ^2_1)	0.012	(0.011, 0.014)
Spatial Decay (ρ_1)	0.003	(0.001, 0.812)
Temporal Autocorrelation (ρ_2)	0.245	(-0.207, 0.833)

Table 2: 95% Credible Intervals for $PM_{2.5}$ Parameters

Parameters	Posterior Median	95% Credible Interval
Intercept ($\beta_{0_{PM}}$)	1.355	(0.973, 1.742)
sqrt(AOD) ($\beta_{1_{PM}}$)	2.381	(0.941, 2.976)
Spatial Variance ($\sigma^2_{\omega_2}$)	0.208	(0.093, 0.924)
Temporal Variance ($\sigma^2_{\tau_2}$)	0.525	(0.329, 0.995)
Pure Error Variance (σ^2_2)	0.173	(0.154, 0.197)
Spatial Decay (ρ_3)	0.094	(0.001, 0.954)
Temporal Autocorrelation (ρ_4)	0.006	(-0.948, 0.951)

FIGURES INCLUDED BELOW

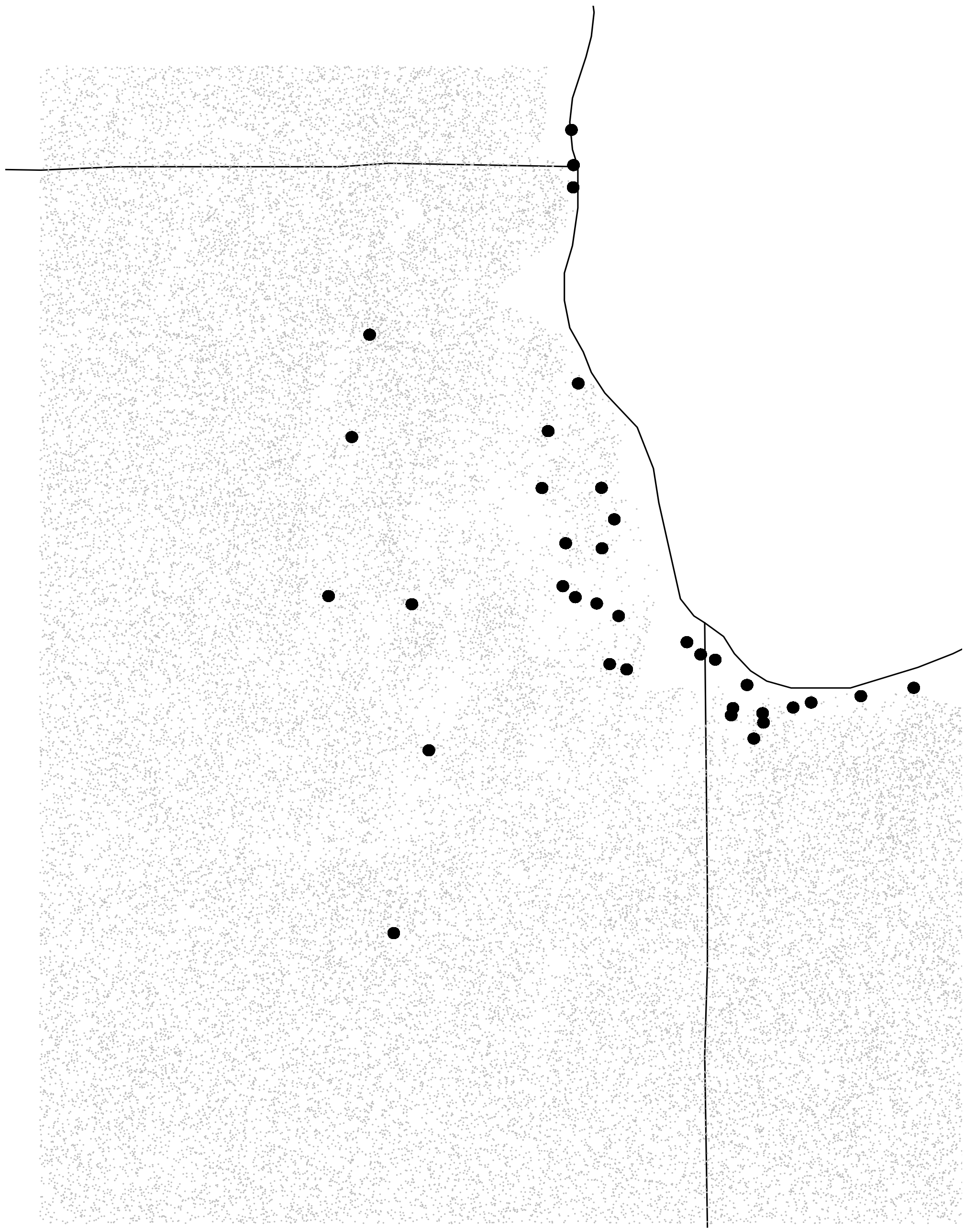

Figure 1: Geographic locations of the 35 $PM_{2.5}$ monitor sites. Black points represent the monitor sites, and the gray points represent a random sample of 10% of the locations where AOD measurements were observed.

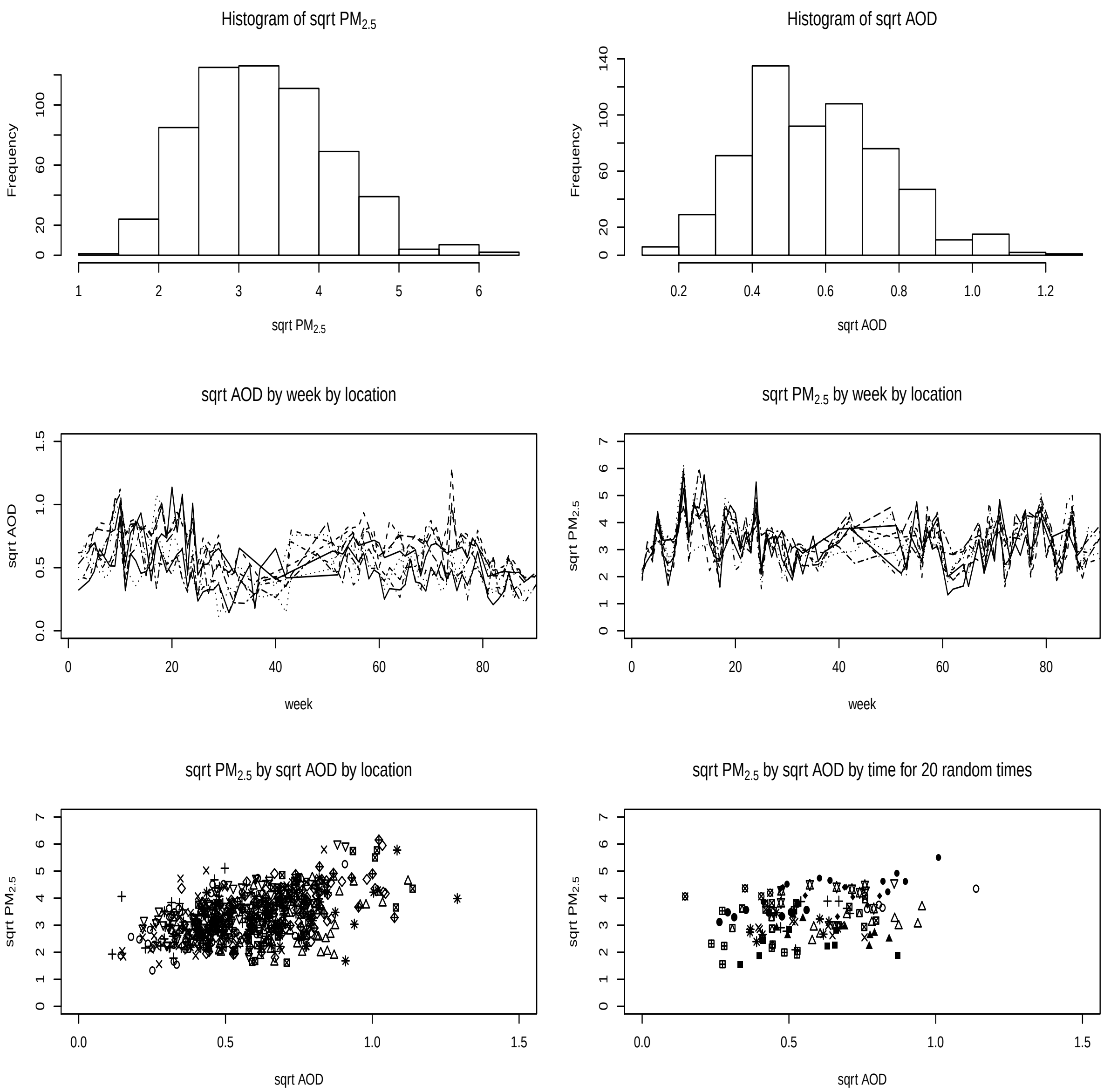

Figure 2: Descriptive Plots for AOD and $PM_{2.5}$. Left to right, top to bottom, the images are: 1 A histogram of the square root of $PM_{2.5}$. 2 A histogram of the square root of AOD. 3 A plot of the square root of AOD by location, demonstrating separability. 4 A plot of the square root of $PM_{2.5}$ by location over time demonstrating separability. 5 Plots of the square root of $PM_{2.5}$ by the square root of AOD stratified over location, with different symbols used for each location. 6 Plots of the square root of $PM_{2.5}$ by the square root of AOD stratified over 20 randomly selected times, with different symbols used for each time point.

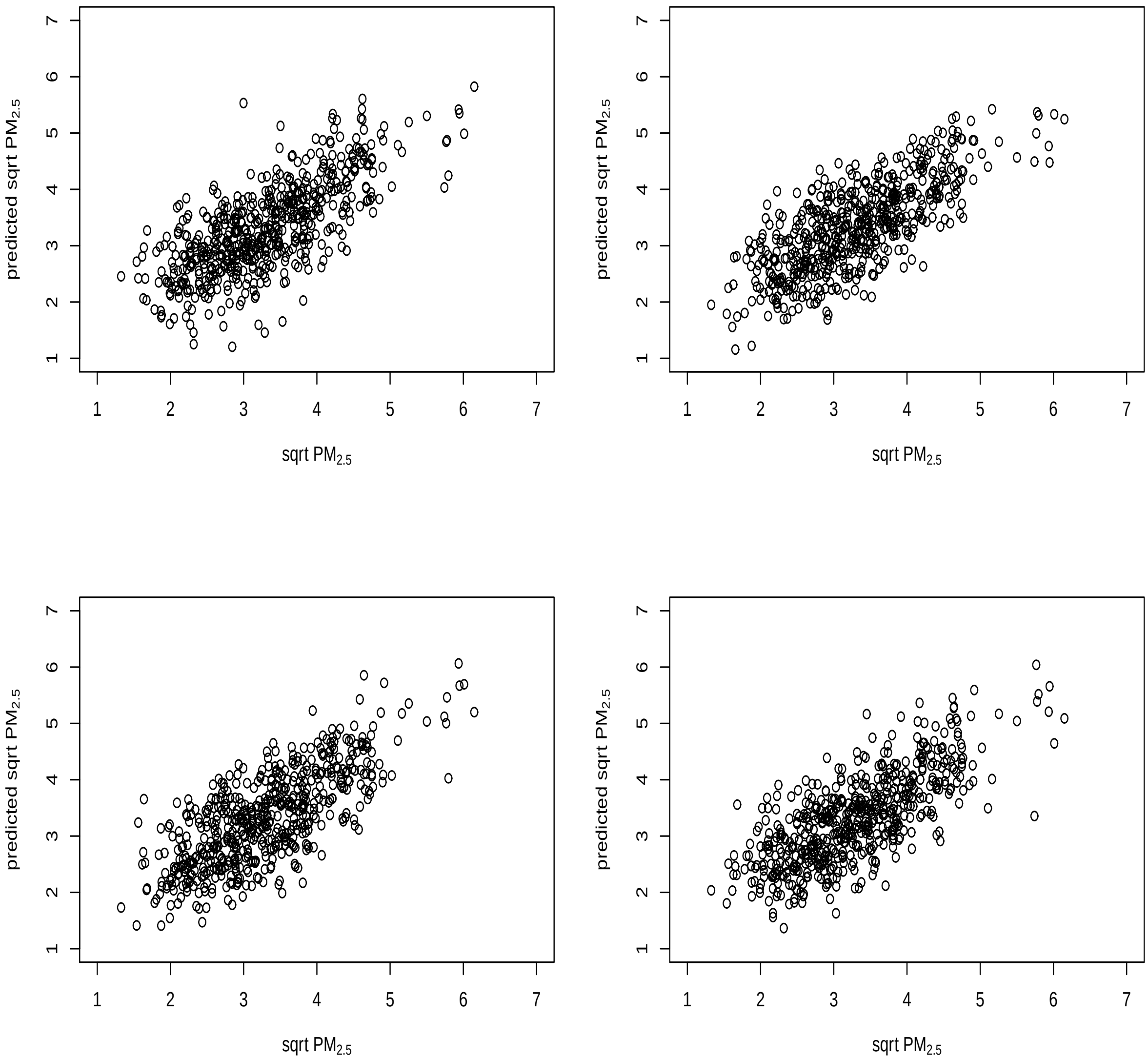

Figure 3: Four realizations of $PM_{2.5}$ predicted by our model plotted by the true $PM_{2.5}$ values.

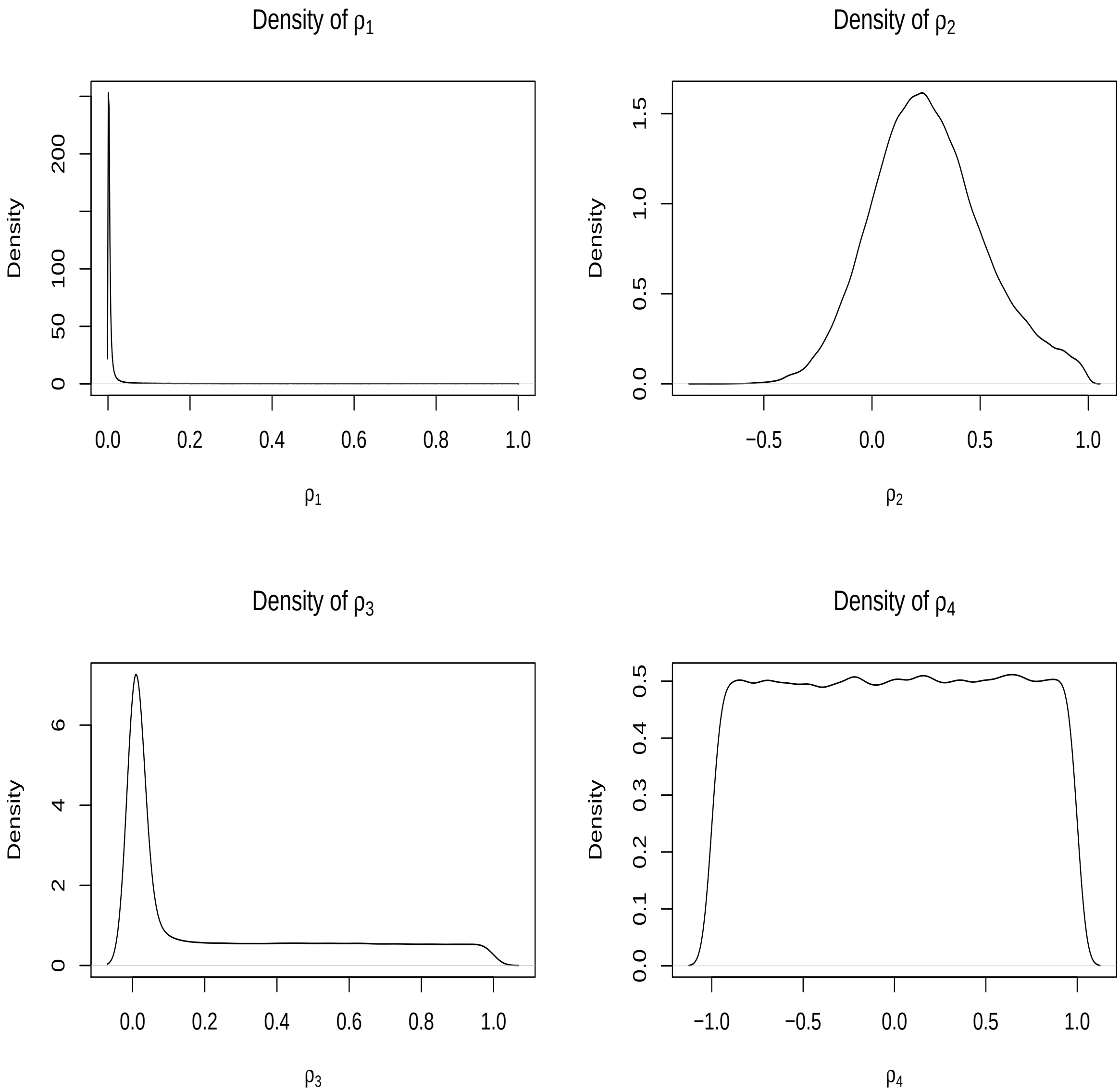

Figure 4: Spatio-temporal Dependency Parameters

Received November 8 2012; accepted October 22, 2013.

Aaron T. Porter
Department of Statistics
Middlebush, The University of Missouri - Columbia
Columbia, Missouri, 65211, U.S.A.

Jacob J. Oleson Department of Biostatistics
105 River Street, The University of Iowa
Iowa City, Iowa, 52242, U.S.A.

Charles O. Stanier
Department of Chemical and Biochemical Engineering
College of Engineering, The University of Iowa
Iowa City, Iowa, 52242, U.S.A.