A Dynamic Process Convolution Approach to Modeling Ambient $PM_{2.5}$ and PM_{10} Concentrations Levels

Catherine A. Calder

November 7, 2003

Department of Statistics, Preprint No. 724

The Ohio State University

Abstract

Elevated levels of particulate matter (PM) in the ambient air have been shown to be associated with certain adverse human health effects. As a result, monitoring networks that track PM levels have been established across the country. Some of the older monitors measure PM less than 10 μm in diameter (PM_{10}) while the newer monitors track PM levels less than 2.5 μm in diameter ($PM_{2.5}$); it is now believed that this fine component of PM is more likely to be related to the negative health effects associated with PM. We propose a bivariate dynamic process convolution model for $PM_{2.5}$ and PM_{10} concentrations. Our aim is to extract information about $PM_{2.5}$ from PM_{10} monitor readings using a latent variable approach and to provide better space-time interpolations of $PM_{2.5}$ concentrations compared to interpolations made using only $PM_{2.5}$ monitoring information. We illustrate the approach using $PM_{2.5}$ and PM_{10} readings taken across the state of Ohio in 2000.

Keywords: spatio-temporal statistics, Bayesian modeling, fine particulate matter, coarse particulate matter

1 Introduction

Elevated levels of particulate matter (PM) in the ambient air have been shown in the epidemiological literature to be linked to increased rates of mortality, morbidity, and respiratory and cardiovascular disease in human populations. Examples of such studies include Goldberg et al. (2001), Katsouyanni et al. (2001), Mar et al. (2000), and Laden et al. (2000). The strength of the association between PM and health effects is usually inversely related to the aerodynamic diameter of the PM. However, until the mid-1980's, most measurements of PM concentrations in the United States were of total suspended PM (TSP) which is dominated by particles too large to penetrate the human thorax and consequently are a poor measure of the inhalation of PM (Lippman, 2000). In the late 1980's, as a result of the establishment of a national depository for air pollution data and of government monitoring mandates, a substantial amount of information about PM_{10} (PM less than 10 μm in aerodynamic diameter) became available. Widespread measurements of $\mathrm{PM}_{2.5}$ (PM less than 2.5 μm in aerodynamic diameter), the component of PM that is most likely responsible for the adverse health effects associated with exposure to PM, are only readily available for about the last four to five years. We propose a model for $\mathrm{PM}_{2.5}$ and PM_{10} that permits more accurate space-time interpolation $\mathrm{PM}_{2.5}$ by including additional information from the more readily available PM_{10} readings.

The basis of our modeling approach is to define two latent spatial temporal processes representing fine and coarse PM and relate these latent processes to our monitor data as follows:

$$\begin{aligned}\mathrm{PM}_{2.5} &= \mathrm{PM}_{fine}\\ \mathrm{PM}_{10} &= \mathrm{PM}_{fine} + \mathrm{PM}_{coarse}\end{aligned}$$

Differences between formation mechanism and composition make this distinction between fine and coarse particles appropriate. According to Lippman (2000), fine particles are formed by chemical reactions, nucleation, condensation, coagulation, and evaporation of fog and cloud droplets in which gases have dissolved and reacted and are composed of sulfate, nitrate, ammonium, hydrogen ions, organic compounds, metal, and particle bound water. Sources include combustion, atmospheric transformation products of $\mathrm{NO_X}$ and $\mathrm{SO_2}$, and high temperature industrial processes. On the other hand, coarse particles (particles greater than 2.5 μm but less than 10 μm in diameter) are formed from mechanical disruption, evaporation of sprays, and suspension of dusts and can include resuspended dusts, coal and fly ash, and metal oxides of crustal elements.

Sources of coarse particles include industrial soil and dust, biological sources, construction and demolition, coal and oil combustion, and ocean spray. Due to these differences in composition and formation processes, we argue that the cross-covariance structure of $PM_{2.5}$ and PM_{10} should be constructed by appropriately specifying the spatial temporal covariance structures of the latent PM_{fine} and PM_{coarse} processes.

We propose a Bayesian dynamic process convolution approach (Higdon, 2002; Calder et al., 2002) for specifying the latent PM_{fine} and PM_{coarse} processes. Spatial anisotropy is permitted in these latent processes by allowing the correlation length of the processes to depend on wind direction and speed. The observed monitoring data are modeled as realizations of these latent process perturbed by measurement error. In addition, covariate effects for the day of the week, temperature, and wind speed are included in the model. Key advantages of our approach include the ability to handle misaligned bivariate space-time data and a latent structure that directly incorporates the additive relationship between PM_{fine} and PM_{coarse}. We illustrate our approach using $PM_{2.5}$ and PM_{10} concentration reading taken across the state of Ohio in 2000.

2 Data

The $PM_{2.5}$ and PM_{10} data used to illustrate our modeling approach were obtained from the Environmental Protection Agency (EPA)'s Aerometric Information Retrieval System/Air Quality Subsystem (AIRS/AQS) database. Measurements taken from monitors located within the state of Ohio in the year 2000 were collected. The sampling schemes of the monitors were somewhat irregular; some $PM_{2.5}$ readings were taken daily and some were taken every third day while PM_{10} readings were either collected daily or every sixth day. Figure 1 shows the locations of the the 44 $PM_{2.5}$ and 67 PM_{10} monitors along with their corresponding sampling schemes. In addition to the irregular sampling scheme, some data are missing probably due to mechanical failure of the monitoring instrument.

Weather data were obtained from the National Climatic Data Center (NCDC) for two purposes: converting the two different types of PM measurements to the same units which is necessary because of the additive nature of our model and for use as covariate information in our model. For simplicity, we assume that the daily weather was fairly consistent over the state of Ohio. Based on this assumption, we only use daily weather data collected at Columbus International Airport (located in the middle of the state) in our analysis. Using daily temperature and pressure readings, the $PM_{2.5}$ measurements, collected in units of

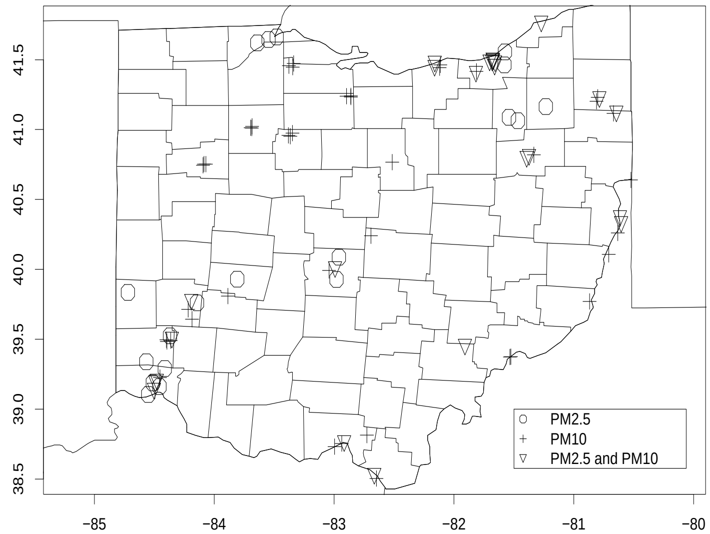

Figure 1: Locations and sampling schemes of the $PM_{2.5}$ and PM_{10} monitors.

$\mu g/m^3$ local conditions (temperature and pressure at the time of the reading), were converted to units of $\mu g/m^3$ standard conditions (25° Celsius, 760 mm Hg), the units in which PM_{10} is observed. The temperature and wind speed readings are used as covariates in our model and the wind speed and wind direction readings are used to specify the anisotropic properties of our latent processes.

3 Modeling Approach

The basic structure for our underlying space-time process is the discrete process convolution framework proposed in Higdon (1998). Discrete process convolutions approximate a Gaussian process using a constructive approach. A spatial field field $\psi(s)$ is defined for every point s in the spatial domain $\mathcal{D}$ as follow:

$$\psi(s) = \sum_{m=1}^{M} \kappa(s - \omega_m) x(\omega_m)$$

where $\kappa(\cdot)$ is a smoothing kernel and $\mathbf{x} = (x(\omega_1), x(\omega_2), \ldots, x(\omega_M))'$ is a discrete white noise process defined on the lattice $\{\omega_1, \omega_2, \ldots, \omega_M\}$ that covers $\mathcal{D}$.

A dynamic process convolution model, as in Higdon (2002) and Calder et al. (2002), extends the discrete process convolution approach by defining the underlying process $\mathbf{x}$ to be a temporally dependent process that is spatially smoothed at each time-step by a two-dimensional smoothing kernel κ. We note that this dynamic process convolution model is different than a three-dimensional discrete process convolution model that would have independent white noise processes defined on a three-dimensional lattice smoothed with three-dimensional kernels. The dynamic process convolution approach has several computational and interpretive advantages which are discussed in Calder (2003).

We propose a bivariate dynamic process convolution model for jointly modeling $\mathrm{PM}_{2.5}$ and PM_{10}. Our approach is similar to the constructive approach for specifying multivariate spatial models proposed Ver Hoef and Barry (1998) of convolving dependent underlying processes and to the dynamic factor process convolution models for multivariate space-time data discussed in Calder (2003). In our model, latent space-time PM_{fine} and PM_{coarse} fields are constructed using the dynamic process convolution framework by convolving independent underlying random walks defined on two lattices. The model for the $\mathrm{PM}_{2.5}$ and PM_{10} concentration levels is specified by adjusting the levels of the PM_{fine} and PM_{coarse} fields with a mean level and a measurement error term:

$$\begin{aligned} \begin{pmatrix} \mathbf{y}_t^{2.5} \\ \mathbf{y}_t^{10} \end{pmatrix} &= \begin{pmatrix} \mu_t^{2.5} \\ \mu_t^{10} \end{pmatrix} + \begin{pmatrix} \mathbf{K}_t^{fine.2.5} & \mathbf{0} \\ \mathbf{K}_t^{fine.10} & \mathbf{K}_t^{coarse.10} \end{pmatrix} \begin{pmatrix} \mathbf{x}_t^{fine} \\ \mathbf{x}_t^{coarse} \end{pmatrix} + \epsilon_t \\ \begin{pmatrix} \mathbf{x}_t^{fine} \\ \mathbf{x}_t^{coarse} \end{pmatrix} &= \begin{pmatrix} \mathbf{x}_{t-1}^{fine} \\ \mathbf{x}_{t-1}^{coarse} \end{pmatrix} + \nu_t \end{aligned} \tag{1}$$

where $\mathbf{y}_t^{2.5}$ and $\mathbf{y}_t^{10}$ are $(N^{2.5} \times 1)$ and $(N^{10} \times 1)$ vectors of the $\mathrm{PM}_{2.5}$ and PM_{10} data at time t, respectively. The vectors $\mu_t^{2.5}$ and μ_t^{10} are the mean levels of the $\mathrm{PM}_{2.5}$ and PM_{10} processes at time t, respectively, and include weekday/weekend, temperature, and wind speed effects:

$$\begin{aligned} \mu_t^{2.5}(n) &= \beta_0^{fine} + \beta_1^{fine}\mathrm{weekday}_t + \beta_2^{fine}\mathrm{temp}_t + \beta_3^{fine}\mathrm{wind}_t \\ \mu_t^{10}(n) &= (\beta_0^{fine} + \beta_0^{coarse}) + (\beta_1^{fine} + \beta_1^{coarse})\mathrm{weekday}_t \\ &\quad +(\beta_2^{fine} + \beta_2^{coarse})\mathrm{temp}_t + (\beta_3^{fine} + \beta_3^{coarse})\mathrm{wind}_t \end{aligned} \tag{2}$$

for $n = 1, 2, \ldots, N_t^{2.5}$ for $\mu_t^{2.5}(n)$ and $n = 1, 2, \ldots, N_t^{10}$ for $\mu_t^{10}(n)$ and where $\mathrm{weekday}_t$ is an indicator function that returns the value 1 if day t is a weekday and 0 if day t is a weekend. The β parameters are partitioned into fine and coarse components to preserved the additive structure of the components of the model. The underlying processes $\mathbf{x}^{fine}$ and $\mathbf{x}^{coarse}$ are defined on lattices of size M^{fine} and M^{coarse}

respectively. We defer discussion of the choice of lattice locations to later in this section. The measurement error and underlying process error terms are normally distributed with diagonal covariance matrices:

$$\epsilon_t \sim N\left(\mathbf{0}, \begin{pmatrix} \lambda_{\epsilon_t}^{2.5} & \mathbf{0} \\ \mathbf{0} & \lambda_{\epsilon_t}^{10} \end{pmatrix}\right) \qquad \nu_t \sim N\left(\mathbf{0}, \begin{pmatrix} \lambda_{\nu_t}^{fine} & \mathbf{0} \\ \mathbf{0} & \lambda_{\nu_t}^{coarse} \end{pmatrix}\right)$$

where $\lambda_{\epsilon_t}^{2.5}$, $\lambda_{\epsilon_t}^{10}$, $\lambda_{\nu_t}^{fine}$, and $\lambda_{\nu_t}^{coarse}$ are diagonal matrices with identical diagonal elements. In our analysis, the diagonal elements of $\lambda_{\epsilon_t}^{2.5}$ and $\lambda_{\epsilon_t}^{10}$ are given diffuse inverse gamma (IG) prior distributions. For simplicity and to help with potential identifiability issues (see the discussion below), the diagonal elements of $\lambda_{\nu_t}^{fine}$ and $\lambda_{\nu_t}^{coarse}$ are set equal to 0.01.

The $\mathbf{K}_t$ matrices in Equation 1 are composed of the appropriate smoothing kernels evaluated at the locations of the data at time t. Since the dimension of the $\mathbf{K}_t$ matrices depends on time, our model can handle different numbers of PM measurements at different times without the need for imputation. $\mathbf{K}_t^{fine.2.5}$ relates the $PM_{2.5}$ data to the fine underlying process using the kernel κ^{fine}, $\mathbf{K}_t^{fine.10}$ relates the PM_{10} data to the fine underlying process using κ^{fine}, and $\mathbf{K}_t^{coarse.10}$ relates the PM_{10} data to the coarse underlying process using the coarse kernel κ^{coarse}. Both κ^{fine} and κ^{coarse} are Gaussian kernels with covariance structures chosen as a result of a preliminary analysis based on examining various directional empirical variograms. (See Cressie (1993) for a discussion of variogram analysis and Kern (2000) for information on choosing the variance of Gaussian smoothing kernels and assigning the appropriate spacing of the underlying grid in discrete process convolution models based on features of empirical variograms.) In our preliminary analysis, we binned both the $PM_{2.5}$ and PM_{10} data according to the wind speed (less than eight miles per hour (mph) and greater than eight mph) and according to wind direction (north-south, northeast-southwest, east-west, and northwest-southeast). Empirical variograms were constructed for the data that fell into each of the eight categories for $PM_{2.5}$ and PM_{10} ignoring the temporal dependence in the data. Based on these variograms, it appeared as though the spatial correlation length for $PM_{2.5}$ depended both on wind speed (correlation length was longer the stronger the wind) and wind direction (correlation length was longer in the direction of the wind). As a result we chose the covariance matrix of the Gaussian kernel for the fine underlying process on day t to be

$$\Sigma_t^{fine} = \begin{pmatrix} a_t^2\cos^2(\theta_t - \pi/2) + b_t^2\sin^2(\theta_t - \pi/2) & \sin(\theta_t - \pi/2)\cos(\theta_t - \pi/2)(b_t^2 - a_t^2) \\ \sin(\theta_t - \pi/2)\cos(\theta_t - \pi/2)(b_t^2 - a_t^2) & a_t^2\sin^2\theta_t - \pi/2 + b_t^2\cos^2(\theta_t - \pi/2) \end{pmatrix}$$

where θ_t is the direction of the wind on day t (measured clockwise from the North and expressed in radians) and set $a_t = 1/\sqrt{2}$ and $b_t = 0.75/\sqrt{2}$ when the wind speed on day t was low (less than eight mph) and $a_t = 2.5/\sqrt{2}$ and $b_t = 1/\sqrt{2}$ when the wind speed was high (greater than eight mph). This elliptical covariance form allows the spatial dependence to be stretched in the direction of the wind (Higdon, 1998). The covariance matrix of the smoothing kernel for the coarse underlying process was chosen to be a circular normal with standard deviation $0.5/\sqrt{2}$. Based on these choices of kernels, the locations of the fine lattice points were chosen to be $1/\sqrt{2}$ degrees apart and the locations of the coarse lattice points were chosen to be $0.5/\sqrt{2}$ degrees apart. Lattice locations more than $\sqrt{1.25}$ times the distance between the lattice points from PM monitors were deleted and not used in our analysis. This guarantees that all observations fall within the smallest square delineated by latitude and longitude lines containing the remaining lattice locations while allowing us to remove lattice locations that are not close monitoring stations. Thus, the value of the spatial components of the model far from data points is zero and the PM levels at these locations would be predicted to be the overall daily mean levels for the state. Figure 2 shows the locations of the underlying fine and coarse lattices along with examples of one standard deviation ellipses of the smoothing kernels (shrunk by a factor of 2).

Convergence problems can arise in the Markov Chain Monte Carlo (MCMC) algorithm used to fit our model due to the weak identifiability (the prior distributions of all model parameters are proper, so technically the model is identifiable) of parameters in the mean component and in the dynamic process convolution component of the model. Unless the parameters in these components are constrained to a certain extent, the values of the parameters can 'trade-off' in the MCMC algorithm. See Gelfand and Sahu (1999) for a discussion of the implications of Bayesian identifiability in MCMC based model fitting algorithms. In our model, we choose to address the identifiability/MCMC convergence problem by placing somewhat informative priors on several model parameters. Each of the β parameters are given a $N(0, 100)$ prior distribution. The initial values of the underlying processes, $x_0^{fine}(m)$ for $m = 1, 2, \ldots, M^{fine}$ and $x_0^{coarse}(m)$ for $m = 1, 2, \ldots, M^{coarse}$, were constrained to be nearly zero by giving them $N(0, 0.001)$ prior distributions. In addition, the diagonal elements of $\lambda_{\nu_t}^{fine}$ and $\lambda_{\nu_t}^{coarse}$ were set equal to 0.01 as mentioned before. This tight prior is strong enough to keep the spatial temporal component of the model from fitting the mean level but does not completely constrain the spatial flexibility of the model as a result of the influence of nearly

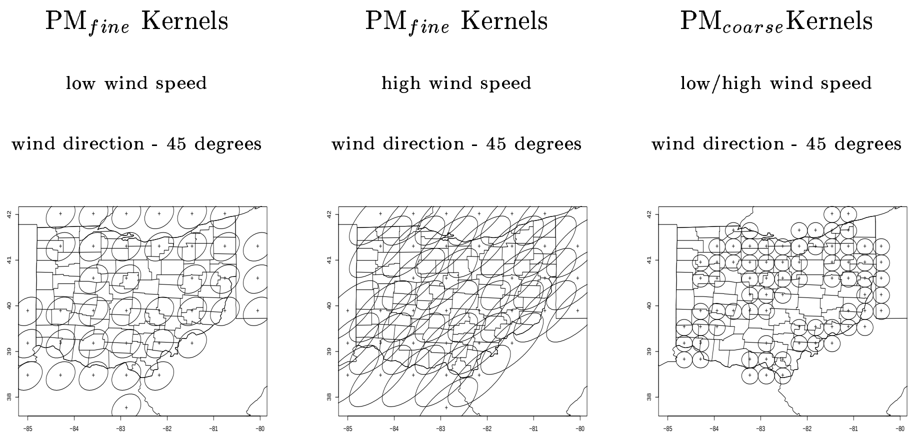

Figure 2: Underlying lattice locations for the PM$_{fine}$ and PM$_{coarse}$ processes along with examples of one standard deviation ellipses of the PM$_{fine}$ smoothing kernel (κ^{fine}) and PM$_{coarse}$ smoothing kernel (κ^{coarse}) smoothing kernel (shrunk by a factor of two).

50-100 PM data points on each day. It appears from a sensitivity analysis that these prior assumptions are flexible enough to allow the model to fit the data reasonably well but are strong enough to permit the MCMC algorithm to converge.

Another limitation of our modeling approach is that PM data cannot be modeled on the log scale while maintaining the computational (discussed in the next section) and interpretive features of the modeling approach. As a result, we chose to eliminate a few large observations ($PM_{2.5}$ concentration readings greater than $50\mu g/m^3$ and PM_{10} concentration readings greater than $100\mu g/m^3$) for which our normal error structure was not appropriate. We do not believe that eliminating these values from the analysis strongly influenced our results due to the fact that these large readings are most likely outliers and are not useful for predicting average PM levels across the study region.

4 Results

The model was fit to the AIRS/AQS PM data using an MCMC algorithm. Since all full conditional distributions of model parameters can be found in closed form, the MCMC algorithm is simply a Gibbs sampler (Gelfand and Smith, 1990) that iteratively samples from the full conditional distributions of the parameters. The underlying $\mathbf{x}$ full conditionals can be derived recursively using the Forward Filtering Backward Sampling

	β_0 (mean level)	β_1 (weekday effect)	β_2 (temperature)	β_3 (wind speed)
fine	15.8 (15.3,16.3)	1.607 (0.780,1.35)	0.181 (0.174,0.190)	-0.451 (-0.489,-0.413)
coarse	10.9 (9.77,11.3)	2.11 (1.53,2.67)	0.201 (0.182,0.218)	-0.174 (-0.243,-0.106)

Table 1: Posterior means and 95 percent credible intervals for the mean level and covariate effect parameters.

Algorithm (see West and Harrison (1997)). The MCMC algorithm was run for 3000 iterations in addition to 500 samples taken to 'burnin' the chain. Convergence of the Markov chain was determined by examining trace plots of model parameters.

Table 1 summarizes the posterior distributions of the model parameters in the mean component of the model (Equation 2). These values are as would be expected. The mean concentration level of $PM_{2.5}$ is around 16 μg/m^3 and PM_{10} levels are on average about 11 μg/m^3 higher. The effects of temperature and weekday are positive as would be expected: temperature is perhaps a catalyst for chemical reactions that result in the formation of PM and weekdays have greater pollution from car traffic and factories. The negative effect of wind speed is also logical; particles are able to disperse further away on days with high wind speed making the concentration of particles in the ambient air smaller. The posterior means and 95 percent credible intervals for the error variance parameters, $\lambda^{2.5}_{\epsilon_n}$ and $\lambda^{10}_{\epsilon_n}$, are 29.2 (28.5,29.8) and 63.4 (61.7,65.1), respectively.

Figures 3 and 4 provide graphical summaries of our results. Figure 3 contains the time-series of the posterior mean $PM_{2.5}$ and PM_{10} concentration levels at a location where neither $PM_{2.5}$ nor PM_{10} were measured. Interpolated $PM_{2.5}$ and PM_{10} concentration levels across the entire state for three consecutive days are shown in Figure 4. These plots were constructed by defining a fine-scale grid over the region and corresponding $\mathbf{K}^*_t$. Using these $\mathbf{K}^*_t$ matrices, the posterior means of the underlying $\mathbf{x}$ processes are linked to the PM levels at the locations on the fine scale grids adjusting the mean level of the process using the posterior means as estimates of the β parameters in the mean component of the model. On these three days there is not a great deal of temporal variability, but there are slight changes to the PM levels in certain locations. A noticeable feature of these figures is the higher levels of PM near the large cities in the state. The figures showing the $PM_{2.5}$ levels pick up the anisotropy incorporated into the model for the underlying PM_{fine} process while the anisotropy appears to be much less apparent in the PM_{10} figures due to the isotropy of the PM_{coarse} component of the model.

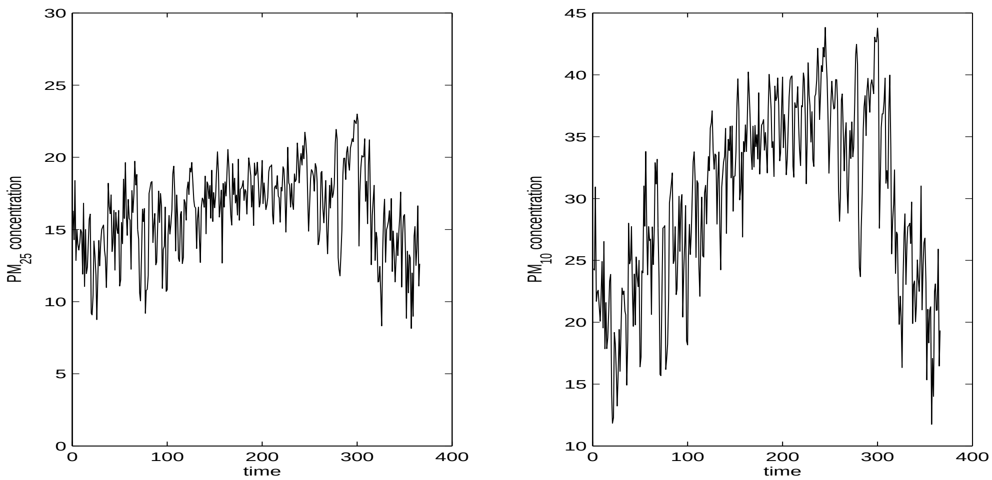

Figure 3: Posterior mean of the predicted PM concentration levels at latitude 40°N and longitude -83°W

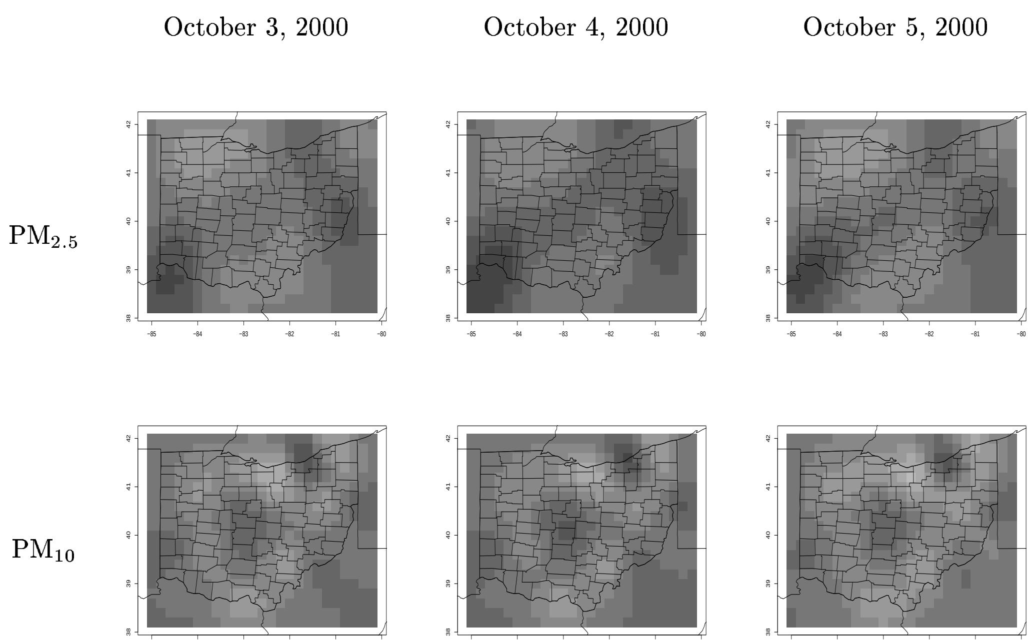

Figure 4: Interpolated PM concentration levels on three consecutive days based on the posterior means of the model parameters. Darker shades represent higher PM concentrations. Note that the shading scale for $PM_{2.5}$ and PM_{10} is not the same.

5 Discussion

The methodology developed in this paper provides a novel approach for characterizing the distribution of $PM_{2.5}$ and PM_{10} concentration levels over space and time. An alternative method might specify a simpler form for the cross-covariance structure of the different types of PM measurements directly and then place a constraint on the ordering of the PM reading (specifically, that $PM_{10} > PM_{2.5}$) at locations where both types of PM data are not observed. Such an approach, however, does not provide an interpretable cross-covariance structure. In addition, most likely the full conditional distributions of all of the model parameters could not be found in closed formed due to the ordering restrictions, and, as a result, model fitting complications could arise.

While our method is attractive in terms of computational issues related to model fitting and the interpretability of the cross-covariance structure, there are definitely some limitations of the approach which have been acknowledged in the paper. The identifiability issue resulting from the dynamic process convolution construction of the underlying spatial temporal processes (discussed in Section 3) makes our results subject to criticisms related to the subjectivity of our prior assumptions. In addition, the fact that we cannot model PM on the log scale means that our model cannot handle extreme observations and permits nonsensical predictions (PM concentration levels less than zero). Still, there are only a few extreme observations that probably are not informative about the overall patterns in the distribution of PM over space and time so omitting them is reasonable. Also, since the distribution of PM levels are, for the most part, not near zero, negative predictions are unlikely.

Despite the limitations of our model, we argue that the latent variable approach to specifying the cross-covariance structure of $PM_{2.5}$ and PM_{10} concentration levels is powerful. It would be nearly impossible to adequately characterize the nonstationary and anisotropic properties of the cross-covariance structure of the different types of PM measurements without being able to incorporate our understanding of the latent PM_{fine} and PM_{coarse} processes. In addition, the nature of the misaligned PM data is readily handled by the dynamic process convolution framework without the need for data imputation. Overall, we believe that our latent variable, dynamic process convolution model provides a coherent and interpretable framework for modeling multivariate space-time processes with unobserved latent components.

References

Calder, C. (2003). "Exploring latent structure in spatial temporal processes using dynamic process convolutions." Ph.D. thesis, Duke University, Durham, NC 27708.

Calder, C., Holloman, C., and Higdon, D. (2002). "Exploring Space-Time Structure in Ozone Concentration Using a Dynamic Process Convolution Model." In *Case Studies in Bayesian Statistics*, vol. VI, 165–176. Springer.

Cressie, N. (1993). *Statistics for Spatial Data*. New York: John Wiley.

Gelfand, A. and Sahu, S. (1999). "Indentifiability, improper priors, and Gibbs sampling for generalized linear models." *Jounal of the American Statistical Association*, 94, 445, 247–253.

Gelfand, A. and Smith, A. (1990). "Sampling-based approaches to calculating marginal densities." *Journal of the American Statistical Association*, 85, 410, 398–409.

Goldberg, M., Burnett, R., Bailar III, J., Brook, J., Bonvalot, Y., Tamblyn, R., Singh, R., and Valois, M. (2001). "The association between daily mortality and short-term effects of ambient air particle pollution in Montreal, Quebec." *Environmental Research*, A86, 12–25.

Higdon, D. (1998). "A process-convolution approach to modeling temperatures in the North Atlantic Ocean." *Journal of Environmental and Ecological Statistics*, 5, 173–190.

— (2002). "Space and space-time modeling using process convolutions." In *Quantitative Methods for Current Environmental Issues*, eds. C. Anderson, V. Barnett, P. C. Chatwin, and A. H. El-Shaarawi, 37–56. Springer Verlag.

Katsouyanni, K., Touloumi, G., Samoli, A., Le Tetre, A., Monopolis, Y., Rossi, G., Zmirou, D., Ballester, D., Boumghar, A., Anderson, H., Wojtyniak, B., Paldy, A., Braunstein, R., J., P., Schindler, C., and Schwartz, J. (2001). "Confounding and effect modification in the short-term effects of ambient particles on total mortality: results from 29 European Cities with the APHEA2 Project." *Epidemiology*, 12, 521–531.

Kern, J. (2000). "Bayesian process-convolution approaches to specifying spatial dependence structure." Ph.D. thesis, Duke University, Durham, NC 27708.

Laden, F., Neas, L., Dockery, D., and Schwartz, J. (2000). "Association of fine particulate matter from different sources with daily mortality in six U.S. cities." *Environmental Health Perspectives*, 941–947.

Lippman, M. (2000). *Environmental Toxicants*. 2nd ed. Wiley-Interscience.

Mar, T., Norris, G., Koenig, J., and Larson, T. (2000). "Assocations between air pollution and mortality in Phoenix, 1995-1997." *Environmental Health Perspectives*, 108, 347–353.

Ver Hoef, J. and Barry, R. (1998). "Constructing and Fitting Models for Cokriging and Multivariable Spatial Prediction." *Journal of Statistical Planning and Inference*, 69, 275–294.

West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. 2nd ed. New York: Springer-Verlag.