Published in final edited form as:

J R Stat Soc Ser C Appl Stat. 2008 April ; 57(2): 187–205. doi:10.1111/j.1467-9876.2007.00608.x.

Modeling mercury deposition through latent space-time processes

Ana G. Rappold, Alan E. Gelfand, and David M. Holland*

Abstract

This paper provides a space-time process model for total wet mercury deposition. Key methodological features introduced include direct modeling of deposition rather than of expected deposition, the utilization of precipitation information (there is no deposition without precipitation) without having to construct a precipitation model, and the handling of point masses at 0 in the distributions of both precipitation and deposition. The result is a specification that enables spatial interpolation and temporal prediction of deposition as well as aggregation in space or time to see patterns and trends in deposition.

We use weekly deposition monitoring data from the NADP/MDN (National Atmospheric Deposition Program/Mercury Deposition Network) for 2003 restricted to the eastern U.S. and Canada. Our spatio-temporal hierarchical model allows us to interpolate to arbitrary locations and, hence, to an arbitrary grid, enabling weekly deposition surfaces (with associated uncertainties) for this region. It also allows us to aggregate weekly depositions at coarser, quarterly and annual, temporal levels.

Keywords

Interpolation; misalignment; multivariate dynamic model, stationarity in space, stationarity in time

1 Introduction

Assessing risk from mercury is a substantial and growing problem. It is extremely complex, requiring high-resolution spatial information on mercury deposition. The contribution of this paper is to provide a space-time process model for total wet mercury deposition. Such deposition is defined as the mass transfer of dissolved gaseous, aerosol, or particulate mercury species by precipitation to the earth's surface. We seek to develop baseline spatial patterns of total mercury deposition for establishing linkages with broader environmental endpoints, and for assessing future progress on mercury reductions associated with legislated mercury emission reduction programs. Key methodological features introduced to achieve this objective include direct modeling of deposition rather than of expected deposition, the utilization of precipitation information without having to construct a precipitation model, the potential to accommodate misalignment between available deposition information and precipitation

* Ana G. Rappold is NRC Research Associate at U.S. Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC (Email: rappold.ana@epa.gov). Alan E. Gelfand is Professor in the Department of Statistical Science at Duke University, Durham, NC 27708-0251 (Email: alan@stat.duke.edu). David M. Holland is senior statistician, U.S. Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC (Email: holland.david@epa.gov). The authors would like to thank Bob Larson of NADP/MDN for providing the raw data with helpful comments and Margaret Polinkovsky for organizing the datasets.

Publisher's Disclaimer: The U.S. Environmental Protection Agency's Office of Research and Development partially collaborated in the research described here. Although it has been reviewed by U.S. EPA and approved for publication, it does not necessarily reflect the Agency's policies or views.

information, and the handling of point masses at 0 in the distributions of both precipitation and deposition. The result is a specification that enables (with full assessment of uncertainty) spatial interpolation and temporal prediction of deposition as well as aggregation in space or time to see patterns and trends in deposition.

In this sense we consequentially exceed the capability of currently-used techniques which apply traditional geostatistical methods to the very sparse monitoring data. Such interpolation methods, in the presence of zero inflated measurements, lead to underestimation of deposition and the related uncertainty. Hence, we address the urgent need for more refined prediction of mercury deposition, particularly in non-monitored sensitive ecosystems. Eventually, by integrating our mercury predictions with variables capturing sensitive populations, fish tissue, and methylation potential it becomes possible to map ecological risks associated with mercury deposition.

Mercury is a highly toxic and prevalent element in the environment. The use and release of mercury as a by-product of human activities has increased the amount of mercury in many parts of the environment including the atmosphere, lakes, and streams. Direct anthropogenic sources include fuel combustion, waste incineration, industrial processes, and metal refining. To reduce mercury emissions from power plants, the U.S. Environmental Protection Agency issued the Clean Air Mercury Rule in 2005 intended to reduce mercury emissions from coal-fired power plants by 70% when fully implemented. Natural emission sources of mercury, such as oceanic, terrestrial, and vegetative emissions, also influence the levels of atmospheric mercury. Origin of deposited mercury is a major source of uncertainty to atmospheric modelers (see Lindberg *et al.* (2007)). Debate over the global pool or local emitters is blurred by the complex mechanism of chemical transformations related to the oxidation of elemental Mercury.

In the atmosphere, Mercury is the most commonly found in its elemental form. However, it readily oxidizes to soluble forms. These organic and soluble inorganic salts are of the largest concern to health. Mercury from the inorganic salts become available for methylation by microorganisms in the soils and sediments after removal from the atmosphere by precipitation. Methylmercury is an organic compound that bio-accumulates in organisms and is a potent toxin to humans and wildlife. Human exposure arises primarily from eating fish containing methylmercury (<http://www.epa.gov/mercury/faq.htm>).

In this analysis, we use weekly deposition monitoring data from the National Atmospheric Deposition Program/Mercury Deposition Network (NADP/MDN) for 2003 restricted to the eastern U.S. and Canada (<http://nadp.sws.uiuc.edu/mdn>). Wet Deposition (ng/m^2) is calculated as the product of concentration (ng/L) and water volume (L) in the sampling bucket. Without precipitation, wet samples are not available; zero deposition is measured at locations with no precipitation during a measurement interval. When wet samples are available, the amount of mercury is directly related to the amount and type of precipitation. Most deposition occurs at the onset of rain/snow events. For example, strong, short-duration summer thunderstorms produce higher mercury deposition in comparison to extended periods of precipitation that bring little or no additional deposits. Therefore, without the joint analysis of precipitation and deposition it is difficult to discern trends and spatial patterns in deposition. To accommodate for zero values, deposition is often transformed by taking the logarithm of the observed value plus one. We find the arbitrary translation by one unit to be unattractive and instead model the distribution of deposition and precipitation as continuous non-negative variables with point masses at zero. Evidently, both the mass at 0 and the precipitation and deposition, when they occur, should have spatio-temporal structure.

Formally, this paper presents a hierarchical model that addresses the space-time relationship between the precipitation and deposition in the atmosphere, incorporates discontinuities for both processes at zero, and handles misalignment between the two processes. The key idea is the introduction of a third *latent* process which is modeled as an autoregressive Gaussian spatial process (in weekly time steps) and drives these two processes. A three-dimensional dynamic space-time process model results such that, for a given location and week, a positive value of the latent process implies non-zero precipitation and deposition for that location in that week. On the other hand, a negative value implies no precipitation, hence no deposition, at that location for that week. The resulting model directly enables inference about deposition in the presence of precipitation information. Moreover, the model can be extended to include other available sources of information such as computer modeled ambient deposition or other correlated chemical depositions such as SO_4 or NO_3 . The result is a $p + 1$ dimensional spatio-temporal process for p variables driven by a latent dynamic process. Hence, mercury deposition could be inferred in the presence of other sources of information.

Our spatio-temporal hierarchical model allows us to interpolate to arbitrary locations and, hence, to an arbitrary grid, enabling weekly deposition surfaces with associated uncertainties for the eastern U.S. and Canada. It also allows us aggregate weekly depositions to coarser, quarterly and annual, temporal levels. Weekly maps are of potential interest in learning about circulation of mercury; the coarser level summaries are of interest to regulatory bodies. We adopt a fully Bayesian approach and, upon fitting the model, obtain posterior distributions and thus uncertainty estimates associated with any of these interpolations or summaries.

At present, other than what we propose here, there is no formal modeling work on mercury deposition in the literature that we are aware of. However, there is by now a substantial literature on space-time modeling, in particular with regard to environmental exposure. Explicit space-time modeling of pollutants can be found in early work of, e.g., Guttorm *et al.* (1994), Haas (1995), and Carroll *et al.* (1997). Hierarchical Bayesian approaches for spatial prediction of air pollution have been developed in Brown *et al.* (1994), and Sun *et al.* (2000). Specific focus on ground level ozone appears in Cox and Chu (1992), Guttorm *et al.* (1994), and Carroll *et al.* (1997). More recent contributions include Zhu *et al.* (2003), Huerta *et al.* (2004), McMillan *et al.* (2005), and Sahu *et al.* (2006). Attention to PM_{10} and $PM_{2.5}$ concentrations is found in, e.g., Zidek *et al.* (2002), Cressie *et al.* (1999), Kibria *et al.* (1997), Shaddick and Wakefield (2002), Smith *et al.* (2003), and Sahu and Mardia (2005). Finally, Wikle (2003) provides an overview of hierarchical modeling in environmental science. The paper by Gelfand *et al.* (2005) provides the general framework for the sorts of dynamic spatial models we consider here. The book by Banerjee *et al.* (2004) provides a general summary of space-time modeling with further references therein.

The format of the paper is as follows. In section 2 we describe the available data sources for our deposition model. In Section 3 we detail the model including its associated dependence structure and how interpolation is implemented. In Section 4 we provide the analysis of the dataset using the proposed modeling. Finally, Section 5 offers a brief summary as well as an indication of some future work to be done in this setting.

2 The data

The National Atmospheric Deposition Program (NADP) is the longest continuously operating national monitoring network, collecting information about acidic chemicals in USA and Canada since 1977. The Mercury Deposition Network (MDN) began operating in 1995 with 13 stations and had expanded to approximately one hundred stations in U.S. and Canada. It is run through cooperative agreements of local, state, and federal organization which have adopted standardized methods for collection, chemical analysis, and quality assurances. All

stations are located in rural areas at least 10km away from the nearest major source of mercury emission. Wet deposition is defined by the product of measured concentration and water volume in the sampling instrument (ng/m^2).

In this study we use 2003 data from 45 (out of 56) stations in eastern U.S. and Canada, rated A or B¹ type quality for more than 50% of the weeks (see Figure 1). Eleven stations were excluded out of which seven (those with A and B rated measurements) were used for cross-validation. The 45 stations provided us with 2038 measurements (out of $52 \times 45 = 2340$, or 87%) and the remaining measurements were considered missing at random. The distributions of non-zero measured deposition on natural and log-scale, across time for all stations are summarized in Figure 2. Typically, between 5 – 13% of stations record zero weekly deposition and occasionally as many as 35% or as few as 0%. Seasonal variability in deposition occurs due to the seasonal variability in precipitation and emissions. Median deposition in the summer months can be up to 50% higher than in the winter months (see Figure 3).

3 Modeling

3.1 Basic modeling ideas

Indexing locations by s and weeks by t , let $\mathbf{Y}_t(s) = (Y_{1t}(s), Y_{2t}(s))'$ denote the observed precipitation and deposition respectively for week t , $t = 1, 2, \dots, T$ at location s , $s \in D$. The $Y_{lt}(s)$, $l = 1, 2$ are thought of as continuous positive random variables with a point mass at 0. Next, we introduce a trivariate latent process $\mathbf{V}_t(s) = (V_{0t}(s), V_{1t}(s), V_{2t}(s))'$ under which we define precipitation $Y_{1t}(s)$ as $Y_{1t}(s) = \exp(V_{1t}(s))1(V_{0t}(s) > 0)$ and mercury deposition $Y_{2t}(s)$ as $Y_{2t}(s) = \exp(V_{2t}(s))1(V_{0t}(s) > 0)$. In other words, $P(Y_{1t}(s) = 0) = P(Y_{2t}(s) = 0) = P(V_{0t}(s) \leq 0)$ so we have a coherent process specification. When $V_{0t}(s) \leq 0$ we have no precipitation, hence no deposition; when $V_{0t}(s) > 0$ there will be actual precipitation and hence, actual deposition.

There are numerous ways to specify the $\mathbf{V}_t(s)$ process. We adopt a conditional choice that prescribes $V_{0t}(s)$ as a driver of precipitation and deposition. In particular, suppose we envision a *large-scale* atmospheric state over the study area at time t , V_{0t} , which follows a mean-centered $AR(1)$ process, $V_{0t} - \mu_{0t} = \rho(V_{0,t-1} - \mu_{0,t-1}) + \delta_{0t}$ where μ_{0t} is a global weekly centering and $\phi \in (0, 1)$ is unknown. We then view $V_{0t}(s)$ as the local adjustment at location s and time t to this process, i.e., $V_{0t}(s) = V_{0t} + \zeta_t(s)$ where $\zeta_t(s)$ are independent mean 0 Gaussian processes. A roughly equivalent version of this specification (which reduces computation) becomes

$$V_{0t}(s) - \mu_{0t} = \phi(V_{0,t-1}(s) - \mu_{0,t-1}) + \sigma_{0t}z_{0t}(s). \quad (1)$$

Here, the $z_{0t}(s)$ are independent replicates of a mean 0, unit variance Gaussian process over D with correlation function ρ_0 . Then, given $V_{0t}(s)$, we specify the bivariate process for $(V_{1t}(s), V_{2t}(s))$ to be conditionally independent across t . So, altogether, $\mathbf{V}_t(s)$ is a dynamic Gaussian process as in Gelfand et al (2005); $V_{0t}(s)$ represents the “state” with (1) denoting the transition process while the bivariate surface realizations $(V_{1t}(s), V_{2t}(s))$ play the role of the “data”. In particular, we specify $f(V_{1t}(s), V_{2t}(s)|V_{0t}(s))$ in the form $f(V_{2t}(s)|V_{1t}(s), V_{0t}(s)) \times f(V_{1t}(s)|V_{0t}(s))$, i.e.,

$$\begin{aligned} V_{2t}(s) &= \mu_{2t} + \alpha^{(2|0)} V_{0t}(s) + \alpha^{(2|1)} V_{1t}(s) + \sigma_{2t} z_{2t}(s) \\ V_{1t}(s) &= \mu_{1t} + \alpha^{(1|0)} V_{0t}(s) + \sigma_{1t} z_{1t}(s) \end{aligned} \quad (2)$$

¹According to the QA/QC Protocols, A samples are fully qualified, B samples have minor problems but are used in summary statistics, and C samples are invalid data which should not be used.

where, again $z_{1t}(s)$ and $z_{2t}(s)$ are independent replicates of a mean 0, unit variance Gaussian process over D with correlation functions ρ_1 and ρ_2 respectively. Such conditional specification of bivariate or multivariate processes has been advocated in Royle and Berliner (1999) and is discussed more generally in Gelfand *et al.* (2004).

We note that the spatial process models for the residuals in (2) allows us to accommodate spatial misalignment in observed precipitation and deposition using ideas given in Chapter 6 of Banerjee *et al.* (2004). This becomes even more useful were we to consider additional information sources as mentioned in the Introduction. Also, the dependence structure associated with (1) and (2) can be calculated under stationarity in time, i.e., $\sigma_{1t} = \sigma_1, \sigma_{2t} = \sigma_2$. (See Rappold *et al.* (2006)). Finally, expressions for $\text{Cov}(Y_{lt}(s), Y_{l't'}(s')), l, l' = 1, 2$ can be written down using the joint distribution of $(V_{0t}(s), V_{1t}(s), V_{2t}(s))$ and $(V_{0t}(s'), V_{1t}(s'), V_{2t}(s'))$. The process is non-stationary and the expressions are analytically intractable. Evaluation is most easily accomplished using Monte Carlo integration based upon samples from this joint distribution. An illustrative example is provided in Rappold *et al.* (2006). As expected, correlation in the Y 's is weaker than that for the corresponding latent V 's.

Note that (1) and (2) provide a purely spatial (no nugget/pure error term) model for $(V_{1t}(s), V_{2t}(s))$. This is intentional. At sufficiently high spatial resolution, we envision continuous (perhaps, not differentiable) surfaces for realized precipitation and deposition and thus seek mean square continuous process realizations for $(V_{1t}(s), V_{2t}(s))$. We also envision the latent $V_{0t}(s)$ as mean square continuous. For appropriate choices of correlation function this will be the case. (See, e.g., Stein (1999), or Banerjee and Gelfand (2002).) The Matérn class with smoothness parameter in $(0, 1)$ provides this. For computational ease, in the sequel we adopt the exponential covariance structure (Matérn with smoothness equals .5), i.e.,

$$\rho_l(s - s'; \xi_l) = \exp(-\xi_l d), \quad l=0, 1, 2$$

where d is distance between the stations s and s' . In attempting to support this assumption, the best we can do is to create a variogram for the observed non zero log precipitation and non zero log depositions. The resulting variogram plots (not shown) suggest exponential correlation functions are plausible.

It is clear from (1) that changing σ_{0t} corresponds to rescaling $V_{0t}(s)$ (and μ_{0t}). However, in (2) $\alpha^{(1|0)}$ and $\alpha^{(2|0)}$ will not be separable from σ_{0t} . So, to ensure identifiability we set $\sigma_{0t} = 1, t = 1, 2, \dots, T$ while σ_{1t} and σ_{2t} remain unknown parameters. Together with the regression weights $\alpha = (\alpha^{(1|0)}, \alpha^{(2|1)}, \alpha^{(2|0)})'$ these parameters completely specify the cross-covariance structure of the tri-variate latent process $\mathbf{V}_t(s)$. We note that our definition of the trivariate process in (1) and (2) permits various model selection decisions - do we need temporal variation in the μ 's, do we need temporal or spatial or even spatio-temporal variation in the σ 's?

3.2 The full model specification

Let $\mathbf{V}_{lt}, l = 0, 1, 2$ denote the $\{V_{lt}(s_i)\}, i = 1, \dots, n$, let $\mathbf{V}_t = (\mathbf{V}_{0t}, \mathbf{V}_{1t}, \mathbf{V}_{2t}), t = 1, 2, \dots, T$ and finally let $\mathbf{V} = \{\mathbf{V}_t\}$. Let θ denote all of the parameters in the process model for $\mathbf{V}$, that is, $\theta = (\{\mu_{lt}\}, \{\sigma_{lt}\}, \alpha^{(1|0)}, \alpha^{(2|0)}, \alpha^{(2|1)}, \{\xi_{lt}\}, \mathbf{V}_0)$ where $\mathbf{V}_0$ is the initializing process realization. Then the first stage specification becomes

$$f(\mathbf{Y}|\mathbf{V}) = \prod_{t=1}^T \prod_{l=1}^2 \prod_{i=1}^n \delta_{\exp(V_{lt}(s_i))1(V_{0t}(s_i) > 0)} \quad (3)$$

where δ_x denotes a degenerate distribution with point mass at x . Next, $f(\mathbf{V}|\theta)$ will be expressed sequentially in time, i.e., $f(\mathbf{V}|\theta) = \prod_{t=1}^T f(\mathbf{V}_t|\mathbf{V}_{t-1}, \theta)$ whence the full model takes the form

$$f(\mathbf{Y}|\mathbf{V}) f(\mathbf{V}|\mathbf{V}_0, \theta) f(\mathbf{V}_0|\theta) f(\theta). \quad (4)$$

$V_{0t}(s)$ is initialized at $t = 0$ with a Gaussian process having mean μ , variance 1, and exponential correlation function with decay parameter ξ_0 .

Turning to the remaining components of θ , we assume a priori that μ_{lt} , $l = 0, 1, 2$ are independent and normally distributed with means 0 and variances 1000. The reciprocal conditional variances (σ_{lt}^{-2}) are assumed to follow Gamma prior $Ga(2, 1)$ while the α 's are assumed independent from a normal prior $N(0, 5)$. These priors are reasonably vague; the former provides an infinite variance for the σ^2 's while the latter constrains the α 's to essentially $(-15, 15)$, much wider than a least squares interval estimate of $\alpha^{(2|1)}$ using the positive depositions. Spatial correlation or decay parameters are known to be poorly identified with weak priors on the variances and so are fixed based on empirical evidence (see Sahu *et al.* (2006) in this regard). More about the choice of the ξ 's is given in Section 4. The temporal auto-regressive coefficient ϕ is given a uniform prior $U(0, 1)$.

3.3 Model Fitting and Spatial Interpolation

Inference is focused on the prediction of aggregated summaries of deposition over the eastern U.S. for quarterly and annual time periods. This is achieved by interpolating to a fairly dense grid over this region. Such interpolation is a post-model fitting activity, implemented one-to-one with posterior samples of $\mathbf{V}$ and θ . In particular, we do individual interpolation. For a new location s_0 , the normal distribution of $\mathbf{V}_t, \mathbf{V}_t(s_0)$ given $\mathbf{V}_{t-1}, \mathbf{V}_{t-1}(s_0)$ and θ provides the normal distribution $[\mathbf{V}_t(s_0)|\mathbf{V}_t, \mathbf{V}_{t-1}, \mathbf{V}_{t-1}(s_0), \theta]$. So, we can draw $\mathbf{V}_t(s_0)$ in sequence and retain $V_{2t}(s_0)$ to show interpolated deposition surfaces.

The model fitting itself is computationally demanding. An advantage of the conditional model specification in (1) and (2) is that it allows us to partition the likelihood for each time step and work with $3Tn \times n$ matrices rather than with the $3nT \times 3nT$ matrix required in the jointly specified model. The Gibbs sampler alternates between updating $\mathbf{V}$ and θ .

The components of θ are updated using the distributions indicated in Appendix AII. Here, we detail a bit more the updating of $\mathbf{V}$. By definition, a sampled realization of $\mathbf{V}(s)$ determines a realization of $\mathbf{Y}(s)$. On the other hand, for observed locations with non-zero measured precipitation and deposition we have $V_{1t}(s) = \log(Y_{1t}(s))$ and $V_{2t}(s) = \log(Y_{2t}(s))$, respectively with an associated constraint on $V_{0t}(s)$. For observed locations with zero precipitation and deposition, we have only a constraint on the associated $V_{0t}(s)$. $Y_{1t}(s)$ and/or $Y_{2t}(s)$ will be missing at some times for a given location. However, this only means that there is no constraint on the latent associated $V_t(s)$'s. We can still update and infer about these V 's through the dynamic model that specifies the $\mathbf{V}_t(s)$ process. Sampling of the latent (unobserved) components of $\mathbf{V}$ is done by taking advantage of the dynamic model structure in (1) and (2). That is, we update $\mathbf{V}_{0t}$ using the full conditional distribution, $[\mathbf{V}_{0t}|\mathbf{V}_{0,t-1}, \mathbf{V}_{0,t-1}, \theta, \mathbf{Y}]$. Then, we update $\mathbf{V}_{1t}, \mathbf{V}_{2t}$ given $\mathbf{V}_{0t}, \theta, \mathbf{Y}$ using the forms $[\mathbf{V}_{1t}|\mathbf{V}_{0t}, \theta, \mathbf{Y}]$ and $[\mathbf{V}_{2t}|\mathbf{V}_{1t}, \mathbf{V}_{0t}, \theta, \mathbf{Y}]$. Each of these distributions is a (constrained) multivariate normal.

At the interpolation stage, consider prediction of deposition at a new location $s^* \notin \mathbf{s} = (s_1, \dots, s_n)$ at time t . Since $[\mathbf{V}(s^*)|\mathbf{V}] = \int [\mathbf{V}(s^*)|\mathbf{V}, \theta][\mathbf{V}, \theta|\mathbf{Y}]$, Monte Carlo samples from $[\mathbf{V}(s^*)|\mathbf{V}]$ are obtained one-for-one using a draw from $[\mathbf{V}(s^*)|\mathbf{V}, \theta]$ with a sample from $[\mathbf{V}, \theta|\mathbf{Y}]$. Here, the latter distribution is, again, a multivariate normal. The sampling is done sequentially in time

and in the order above. Conveniently, at a given iteration, for each $l, l = 0, 1, 2$, the covariance matrices for the V 's at the set of observed locations is the same for all s^* and so may be computed once, inverted, and stored. We only need to introduce the $n \times 1$ vector of regression coefficients and a correlation with $V_{lt}(s)$ for each s^* . These also have to be computed once and stored, reducing the computational burden (Sahu *et al.* (2006)). Finally, predictive samples of the $Y_{lt}(s^*)$ are deterministically obtained from the $V_{lt}(s^*)$ using $Y_{lt}(s^*) = \exp(V_{lt}(s^*))1(V_{0t}(s^*) > 0)$.

Quarterly and annual summaries are obtained directly from the MCMC output in two stages. We create the predictive samples of Y 's at each t and aggregate to the quarterly and annual levels. Posterior distributions of aggregated surfaces are summarized by medians and widths of 95% equal tail credible intervals.

4 Analysis

It becomes computationally infeasible to update the decay parameters (ξ 's) within an MCMC algorithm. Moreover, the ξ 's and associated σ 's are weakly identified and can not be consistently estimated. (See, e.g., Zhang (2004).) Also, with exponential covariance functions, interpolation is sensitive to the value of $\sigma^2\xi$ but not to either one individually (Stein, 1999). Hence, we elect to fix the ξ 's at appropriate levels (as described in Appendix I) and adopt weak priors for the σ^2 's.

After fitting a variety of models to the data we made some modeling choices. As stated earlier, for the identifiability of scale in V_{0t} we fix $\sigma_{0t}^2 = 1$ for all time periods and allow σ_{1t}^2 and σ_{2t}^2 to vary with time. Time varying variances accommodate changes in seasonal variability and uncertainty due to missing observations. Posterior summaries of $\mu_{0t}, t = 1, \dots, 52$ revealed no evidence of seasonal trend so we chose a more parsimonious model setting $\mu_{0t} = \mu$. The posterior distribution of μ_0 has median 0.32 with 95% interval estimate (0.13, 0.53). For the α 's we find the following: $\alpha^{(10)}$ has median 2.70 with a 95% probability interval (2.49, 2.91), $\alpha^{(20)}$ has median -0.27 with a 95% probability interval (-0.48, -0.09), and $\alpha^{(21)}$ has median 0.66 with a 95% probability interval (0.62, 0.69).

Model based prediction intervals for deposition at the observed location are compared to the non-zero observations in Figure 4. The comparison is given on the log scale. The proportion of intervals containing the observed deposition is around 95%. Also approximately half of the predictions are greater than their corresponding observed values, and half are lower, suggesting no bias concerns with our model. Also, note the pattern of over-prediction for smaller depositions, under-prediction for larger ones. This is a manifestation of the shrinkage (or spatial smoothing) inherent in process models.

Figure 5 displays the posterior summaries of the time varying intercepts and variances, for precipitation and for mercury deposition. Considerable temporal variation is evident in all four panels. A particularly strong seasonal trend is picked up by the Hg intercept term μ_{2t} , even after adjustment for precipitation (Equation (2)), likely due to differences in emissions as well as the increase in atmospheric reaction taking place with increased summertime light and temperature. The seasonality in precipitation is similar but less strong, essentially following the observed data. Temporal variation in the variances appears to be somewhat associated with variation in mean levels and is also influenced by the number of missing or zero observations within a week.

Model-based interpolation maps on annual and quarterly levels converted to $\mu\text{g}/\text{m}^2$ are given in Figures 6 and 7. We observe lower deposition rates in the northeast in comparison to the rest of eastern seaboard. We attribute this result to the prevalence of coal-burning power plants

in conjunction with the availability of ozone for mercury oxidation in the southeast during the summer months. In comparison, the northeast is impacted by smaller emissions from the higher percentage of nuclear and oil-burning power plants in this region. The largest deposition in all seasons is found at the southern tip of Florida. This is largely due to tall thundering rain storms, occurring year-round which are capable of oxidizing additional mercury from the stratosphere as well as due to the availability of some oxidative agents factors found in the ocean waters such as Bromine.

As noted above, prediction under the proposed model extends to accommodate inclusion of additional sources of information, in our case, precipitation. In addition to the MDN measurements of precipitation we also have national weather station data on weekly precipitation (NWS). This is a far more extensive network, distinct from MDN, but uses the same measurement protocols. We summarize quarterly and annual deposition using national weather station data to illustrate the improved prediction that results from combining deposition and precipitation (Figures 8 - 10). The overall spatial distribution of deposition predicted conditional on the precipitation data is comparable to the inference without additional information. Most notable differences occur in the prediction away from the observed sites and the prediction uncertainty. Conditional on NWS data, predictions for the south-east region are on the order of $2\text{--}6\mu\text{g}/\text{m}^2$ higher annually than in predictions without the additional information. The improvement in prediction uncertainty can be up to four fold. The width of the intervals, in the region stretching from coastal South Carolina to northern Ohio, is in range $15\text{--}20\mu\text{g}/\text{m}^2$ while, conditional on the weather station data, uncertainty intervals over the same region span at $5\text{--}13\mu\text{g}/\text{m}^2$. Differences in prediction using the additional precipitation data is more striking when compared with the surfaces obtained via ordinary kriging (Figure 11). Confidence intervals as seen in Figure 11(b) grossly underestimate the uncertainty and under-predicts on average by $5\mu\text{g}/\text{m}^2$ and up to $10\mu\text{g}/\text{m}^2$ in some locations. The two regions where our model over-smooths is on southern tip of Florida and around Mobile Alabama. As previously mentioned this two locations, at times, have inordinately large depositions due to conducive meteorological and geographical conditions.

Additional benefit to the sharper inference obtained by the use of our modeling in conjunction with the National Weather Station data is for network design. Such design to choose the sample locations of large-scale monitoring networks is typically based on minimizing some measure of predictive variance (e.g. average or maximum prediction variance over a set of locations). The large reductions in predictive uncertainties achieved by conditioning on the extensive NWS precipitation network may reduce the need to add locations based on optimizing predictive variance design criteria. In turn, this would allow NADP/MDN to consider new types of network design that address the prediction of extreme mercury levels or better deeper investigation of “hot spot” areas. Additionally, reduced prediction uncertainties allow more accurate delineation of regions where mercury levels exceed a threshold value, and permit a useful evaluation of particular threshold values.

5 Summary and Future Work

We have provided the first fully model-based approach to assessing mercury deposition in the literature. The model is hierarchical incorporating a number of key features. It is fitted within the Bayesian framework, enabling suitable interpolation in space and aggregation in time. Application was made to what is currently the best available deposition dataset and a fairly detailed analysis was presented. Further analysis could include total integrated mercury *loading* over a specified area, resulting from process integration over the area. Such block accumulation can be handled within our framework.

The model specification could be enriched with the availability of suitable covariates. (For example, elevation is not useful since there is very little variation in elevation for the current monitoring stations.) Also, the covariance structure could be extended to allow non-stationarity, most easily through spatially varying process variances described say, through a trend surface.

There is available data on wet deposition of SO_4 and NO_3 through NADP/CASTNET. Preliminary investigation of association between these two depositions as well as with mercury deposition revealed weak dependence. However, as noted in the modeling development, it is possible to augment the dynamic specification to include this additional information, to examine what it contributes to prediction of mercury deposition. In a similar vein, there is considerable ongoing effort in the development of computer models for ambient mercury (Bullock and Brehme (2002)). The output from such modeling can be introduced, again to explain wet deposition, requiring a data fusion modeling component along with calibration of the model output. A useful but more challenging related task would be to assess total mercury deposition, wet plus dry.

Appendix

5.1 AI

As noted in Section 4, we elect to fix the ξ 's at appropriate levels and adopt weak priors for the σ^2 's. An approach for selecting the ξ 's is through a validation mean square error criterion (VMSE) evaluated on the set of observations not included in the model estimation along with the coverage probability for the observed stations. VMSE is computed as

$$VMSE = \frac{\sum_{t=1}^T \sum_{i=1}^{n'} (Y_{2t}(s_i) - \widehat{Y}_{2t}(s_i))^2 I(Y_{2t}(s_i))}{\sum_{t=1}^{n_T} \sum_{i=1}^{n'} I(Y_{2t}(s_i))}$$

We consider A and B rated observations from the excluded stations and calculate a model-based prediction estimate $\widehat{Y}_t(s')$. The prediction stations are FL32, MA01, MI48, NH05, ON10, ON11, and SC03. For decay parameters ξ_0 , ξ_1 and ξ_2 we consider the values 1/1500, 1/1000, 1/500, 1/250, and 1/100. By solving the relationship $\exp\{-\xi d\} \approx 0.05$, these values correspond to the approximate range of 4500, 3000, 1500, 750, and 300 km. We searched for the most appropriate values among the combinations of ξ'_s such that $\xi_0 \leq \xi_1 \leq \xi_2$ to preserve identifiability. The best combinations was $\xi_0 = 1/1500$, $\xi_1 = 1/250$, $\xi_2 = 1/100$. This choice of decay parameters gave us predictive coverage of 95.7% for precipitation and 95.4% for mercury deposition on the fitted data.

AI

We provide the full conditional distributions used for fitting the model in (1) and (2). First the full conditional distributions for latent components

$$\begin{aligned}
& -2 \log f(V_{0t}(s) | V_{0t-1}(s), V_{0t+1}(s), \theta) \propto \\
& \frac{1}{\sigma_0^2} (V_{0t}(s) - \mu_{0t} - \phi(V_{0t-1}(s) - \mu_{0t}))^T \sum_0^{-1} (V_{0t}(s) - \mu_{0t} - \phi(V_{0t-1}(s) - \mu_{0t})) \times \\
& \frac{1}{\sigma_{1t}^2} (V_{1t}(s) - \mu_{1t} - \alpha^{(1|0)} V_{0t}(s))^T \sum_1^{-1} (V_{1t}(s) - \mu_{1t} - \alpha^{(1|0)} V_{0t}(s)) \times \\
& \frac{1}{\sigma_{2t}^2} (V_{2t}(s) - \mu_{2t} - \alpha^{(2|0)} V_{0t}(s) - \alpha^{(2|1)} V_{1t}(s))^T \sum_2^{-1} \\
& (V_{2t}(s) - \mu_{2t} - \alpha^{(2|0)} V_{0t}(s) - \alpha^{(2|1)} V_{1t}(s)) \times \\
& \frac{1}{\sigma_0^2} (V_{0t+1}(s) - \mu_{0t} - \phi(V_{0t}(s) - \mu_{0t}))^T \sum_0^{-1} (V_{0t+1}(s) - \mu_{0t} - \phi(V_{0t}(s) - \mu_{0t}))
\end{aligned}$$

So, $V_{0t}(s) | V_{0,t-1}(s), V_{0,t+1}(s), \theta \sim N(Bb, B)$ where

$$\begin{aligned}
B &= \left(\frac{1+\phi^2}{\sigma_0^2} \sum_0^{-1} + \frac{\alpha^{(1|0)^2}}{\sigma_{1t}^2} \sum_1^{-1} + \frac{\alpha^{(2|0)^2}}{\sigma_{2t}^2} \sum_2^{-1} \right)^{-1}, \\
b &= \frac{b_0 + \phi^2 b_3}{\sigma_0^2} \sum_0^{-1} + \frac{b_1 \alpha^{(1|0)^2}}{\sigma_{1t}^2} \sum_1^{-1} + \frac{b_2 \alpha^{(2|0)^2}}{\sigma_{2t}^2} \sum_2^{-1} \\
b_0 &= \mu + \phi(V_{0t-1}(s) - \mu), b_1 = \frac{V_{1t}(s) - \mu_{1t}}{\alpha^{(1|0)}}, \\
b_2 &= \frac{V_{2t}(s) - \mu_{2t} - \alpha^{(2|1)} V_{1t}(s)}{\alpha^{(2|0)}}, \quad b_3 = \mu + \frac{V_{0t+1}(s) - \mu}{\phi}
\end{aligned}$$

and

$$\begin{aligned}
V_{1t}(s) | V_{0t}(s), \theta &\sim N\left(\mu_{1t} + \alpha^{(1|0)} V_{0t}(s), \sigma_{1t}^2 \sum_1^{-1}\right) \\
V_{2t}(s) | V_{1t}(s), V_{0t}(s), \theta &\sim N\left(\mu_{2t} + \alpha^{(2|0)} V_{0t}(s) + \alpha^{(2|1)} V_{1t}(s), \sigma_{2t}^2 \sum_2^{-1}\right)
\end{aligned}$$

Full conditional distributions for intercepts:

$$\begin{aligned}
\mu_{0t} | \dots &\sim N(Bb, B) \\
B^{-1} &= \frac{1}{B_0} + \sum_{t=1}^T (1 - \phi)^2 \mathbf{1}^T \sum_0^{-1} \mathbf{1}, \quad b = \sum_{t=1}^T \mathbf{1}^T \sum_0^{-1} (V_{0t}(s) - \phi V_{0t-1}(s)) \\
\mu_{lt} | \dots &\sim N(Bb, B) \quad l=1,2 \\
B^{-1} &= \frac{1}{B_0} \sum_{t=1}^T \mathbf{1}^T \sum_l^{-1} \mathbf{1} / \sigma_{lt}^2, \quad b = \mathbf{1}^T \sum_l^{-1} (V_{0t}(s) - m_l) / \sigma_{lt}^2
\end{aligned}$$

Full conditional distribution for ϕ

$$\begin{aligned}
\phi | \dots &\sim N(Bb, B) \mathbf{1}_{(0,1)} \\
B^{-1} &= \sum_{t=1}^T Z_t^T \sum_0^{-1} Z_t / \sigma_{0t}^2, \quad b = \sum_{t=1}^T Z_t^T \sum_0^{-1} E_t / \sigma_{0t}^2 \\
Z_t &= V_{0t-1} - \mu_0, \quad E_t = V_{0t} - \mu_0
\end{aligned}$$

full conditional distributions for regression coefficients α

$$\begin{aligned}
[\alpha^{(i/j)}] &\sim N(Bb, B) \\
B^{-1} &= \frac{1}{B_0^6} + \sum_{t=1}^T Z_{jt}^T \sum_0^{-1} Z_{jt} / \sigma_{it}^2 & b &= \sum_{t=1}^T Z_{jt}^T \sum_0^{-1} E_t / \sigma_{it}^2 \\
Z_{jt} &= V_{jt} \\
E_t &= V_{1t} - \mu_{1t} & i=1, j=0 \\
E_t &= V_{2t} - \mu_{2t} - \alpha^{(2|1)} V_{1t} & i=2, j=0 \\
E_t &= V_{2t} - \mu_{2t} - \alpha^{(2|0)} V_{0t} & i=2, j=1
\end{aligned}$$

Full conditional distribution for variance parameters σ_{lt} $l = 1, 2$

$$\begin{aligned}
\sigma_{lt}^2 &\sim IG \left(a + \frac{n}{2}, b + \frac{1}{2} (V_{lt} - m_l)^T \sum_i^{-1} (V_{it} - m_i) \right) \\
m_1 &= \mu_{1t} + \alpha^{(1|0)} V_{0t} \\
m_2 &= \mu_{2t} + \alpha^{(2|0)} V_{0t} + \alpha^{(2|1)} V_{1t}
\end{aligned}$$

References

- Banerjee S, Gelfand AE. Prediction, Interpolation and Regression for Spatially Misaligned Data Points. *Sankhya* 2002;64:227–245.
- Banerjee, S.; Carlin, BP.; Gelfand, AE. Hierarchical modeling and analysis for spatial data. Chapman & Hall; 2004.
- Brown PJ, Le ND, Zidek JV. Multivariate Spatial Interpolation and Exposure to Air Pollutants. *The Canadian Journal of Statistics* 1994;22:489–510.
- Bullock OR Jr. Brehme KA. Atmospheric mercury simulation using the CMAQ model: formulation, description and analysis of wet deposition results. *Atmospheric Environment* 2002;36:2135–2146.
- Carroll RJ, Chen R, George EI, Li TH, Newton HJ, Schmiediche H, Wang N. Ozone exposure and population density in Harris County, Texas. *Journal of the American Statistical Association* 1997;92:392–404.
- Cox WM, Chu SH. Meteorologically adjusted trends in urban areas, a probabilistic approach. *Atmospheric Environment* 1992;27:425–434.
- Cressie, N.; Kaiser, MS.; Daniels, MJ.; Aldworth, J.; Lee, J.; Lahiri, SN.; Cox, L. Spatial Analysis of Particulate Matter in an Urban Environment. In: Gmez-Hernandez, J.; Soares, A.; Froidevaux, R., editors. *GeoEnvII: Geostatistics for Environmental Applications*. Kluwer; Dordrecht: 1999. p. 41–52.
- Gelfand AE, Schmidt AM, Banerjee S, Sirmans CF. Nonstationary Multivariate Process Modeling through Spatially Varying Coregionalization (with discussion). *Test* 2004;2:1–50.
- Gelfand AE, Banerjee S, Gamerman D. Spatial Process Modelling for Univariate and Multivariate Dynamic Spatial Data. *Environmetrics* 2005;16:1–15.
- Guttorp P, Meiring W, Sampson PD. A Space-time Analysis of Ground-level Ozone Data. *Environmetrics* 1994;5:241–254.
- Haas TC. Local Prediction of a Spatio-Temporal Process With an Application to Wet Sulfate Deposition. *Journal of the American Statistical Association* 1995;90:1189–1199.
- Huerta G, Sansó B, Stroud JR. A spatiotemporal model for Mexico City ozone levels. *Journal of the Royal Statistical Society, Series C* 2004;53:231–248.
- Kibria B, Golam M, Sun L, Zidek JV, Le ND. Bayesian Spatial Prediction of Random Space-Time Fields With Application to mapping PM_{2.5} Exposure. *Journal of American Statistical Association* 2002;97:112–124.
- Le ND, Sun W, Zidek JV. Bayesian Multivariate Spatial Interpolation with Data Missing by Design. *Journal of Royal Statistical Society, Series B* 1997;59:501–510.

- Lindberg SB, Bullock RE, Dinghaus R, Engstrom D, Feng X, Fitzgerald W, Pirrone N, Prestbo E, Seigneur C. Panel on Source Attribution of Atmospheric mercury: A Synthesis of Progress and Uncertainties in Attribution the Sources of Mercury in Deposition. *Ambio* 2007;36:1, February 2007.
- Mardia, K.V.; Goodall, C. Spatial-Temporal Analysis of Multivariate Environmental Monitoring Data, Multivariate Environmental Statistics. In: Patil, G.P.; Rao, C.R., editors. Elsevier; Amsterdam: 1993. p. 347-386.
- McMillan N, Bortnick SM, Irwin ME, Berliner M. A hierarchical Bayesian model to estimate and forecast ozone through space and time. *Atmospheric Environment* 2005;39:1373-1382.
- Munthe J, Kindbom K, Kruger O, Peterson G, Pacyna J, Iverfeldt A. Examining source-receptor relationships for mercury in Scandinavia. *Water Air Soil Pollut. Focus* 1 2001;99:110.
- Rappold, AG.; Gelfand, AE.; Holland, DM. Technical Report. Institute of Statistics and Decision Science, Duke University; 2006. Modeling mercury deposition through latent space-time processes.
- Royle JA, Mark Berliner L. A Hierarchical Approach to Multivariate Spatial Modeling and Prediction. *Journal of Agricultural, Biological, and Environmental Statistics* 1999;4:29-56.
- Sahu SK, Mardia KV. A Bayesian Kriged-Kalman model for short-term forecasting of air pollution levels. *Journal of the Royal Statistical Society, Series C* 2005;54:223-244.
- Sahu SK, Gelfand AE, Holland DM. Spatio-temporal modeling of fine particulate matter. *Journal of Agricultural, Biological, and Environmental Statistics* 2006;11:61-86.
- Shaddick G, Wakefield J. Modelling Daily Multivariate Pollutant Data at Multiple Sites. *Journal of the Royal Statistical Society, Series C* 2002;51:351-372.
- Smith RL, Kolenikov SC, Cox LH. Spatio-Temporal Modeling of PM_{2.5} Data with Missing Values. *Journal of Geophysical Research-Atmospheres* 2003, 108 D24 9004, doi:10.29/2002JD002914.
- Stein, ML. Interpolation of Spatial Data: Some Theory for Kriging. New York: Springer-Verlag: 1999.
- Sun L, Zidek JV, Le ND, Ozkaynak H. Interpolating Vancouver's Daily Ambient PM₁₀ Field. *Environmetrics* 2000;11:651-663.
- Wikle CK. Hierarchical models in environmental science. *International Statistical Review* 2003;71:181-199.
- Zhang. Inconsistent Estimation and Asymptotically Equal Interpolations in Model-Based geostatistics. *Journal of the American Statistical Association* 2004;99:250-261.
- Zhu L, Carlin BP, Gelfand AE. Hierarchical regression with misaligned spatial data: relating ambient ozone and pediatric asthma ER visits in Atlanta. *Environmetrics* 2003;14:537-557.
- Zidek JV, Sun L, Le ND, Özkaynak H. Contending with Space-time interaction in the spatial prediction of pollution: Vancouver's hourly ambient PM_{2.5} field. *Environmetrics* 2002;13:595-613.

Figure 1. NADP-MDN stations in East US and Canada. Stations marked by circles are used in the analysis while stations marked by diamonds had more than half of the weeks in 2003 with missing values. Station data denoted with diamonds are used for cross-validation.

Figure 2.
Distribution of annual deposition on natural and log scale.

(a)

(b)

Figure 3.
Distribution of the observed mercury deposition by (a) weeks, (b) stations, in 2003.

Figure 4. Posterior intervals vs log observed mercury deposition. Zero mercury deposition readings are excluded.

Figure 5.
Posterior distribution of (a) μ_{1t} , (b) μ_{2t} , (c) σ_{1t}^2 , and (d) σ_{2t}^2 by weeks.

Figure 6.
Annual distribution of mercury deposition in $\mu\text{g}/\text{m}^2$ for the year 2003.

Figure 7.
Quarterly distribution of mercury deposition in $\mu\text{g}/\text{m}^2$ for the year 2003.

Figure 8. Prediction surfaces conditional on the National Weather Station precipitation data for the year 2003, (a) the annual distribution of mercury deposition in $\mu\text{g}/\text{m}^2$, (b) widths of 95% probability intervals.

Figure 9. Quarterly distribution of mercury deposition in $\mu\text{g}/\text{m}^2$ for the year 2003, conditional on the National Weather Station precipitation data.

Figure 10. Widths of 95% probability intervals for the quarterly mercury deposition in $\mu\text{g}/\text{m}^2$ for the year 2003, conditional on the National Weather Station precipitation data.

Figure 11.

(a) Kriged surface (b) widths of 95% confidence intervals (c) Difference of the predicted annual deposition between the inference conditional on the National Weather Station precipitation data and ordinary kriging.