



APPENDIX AVAILABLE ON THE HEI WEB SITE

Research Report 167

Assessment and Statistical Modeling of the Relationship Between Remotely Sensed Aerosol Optical Depth and PM_{2.5} in the Eastern United States

Christopher J. Paciorek and Yang Liu

Appendix D. Flexible Buffer Modeling Using Penalized Splines

Correspondence may be addressed to Dr. Christopher J. Paciorek, Department of Statistics, 367 Evans Hall, University of California, Berkeley, CA 94720; e-mail: paciorek@stat.berkeley.edu.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR-83234701 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

This document was reviewed by the HEI Health Review Committee but did not undergo the HEI scientific editing and production process.

© 2012 Health Effects Institute, 101 Federal Street, Suite 500, Boston, MA 02110-1817

D. Flexible Buffer Modeling Using Penalized Splines

D.1. Introduction

In fields such as exposure science, social epidemiology and health services research, interest often lies in understanding or accounting for the effect of proximity to sources of exposure to something that may influence an outcome of interest. Exposure might be to pollution sources, access to fast food or liquor outlets, or access to recreation or health centers, among others, while the outcome is often either a pollutant concentration or some measure of health or socioeconomic outcome (Sallis et al. 1990, Morland et al. 2002, Laraia et al. 2004, Farley et al. 2006). If there is at most a single exposure for each observation, one could include presence/absence of the exposure or distance to the source of exposure as a linear term in a regression model, or estimate the effect of distance to source as a smooth regression term (e.g., Wood 2006). Accounting for multiple sources is more difficult, and it is commonplace to use a geographic information system (GIS) to calculate a metric of exposure to sources (such as the sum of source strengths or number of sources) within a buffer of the location associated with each observation in a dataset (such as a person's residence) (e.g., Baxter et al. 2007, Yanosky et al. 2009). This requires choosing a buffer window size and ignores information on distance. An alternative is to use distance to the nearest source as the metric, which ignores the effect of remaining sources. Other approaches apply a weight function, e.g., a Gaussian decay kernel, to downweight values at larger distances; the methodology introduced here estimates the weight function flexibly based on the data.

Here we introduce a method that can estimate a smooth effect of distance to source, accounting for multiple sources and weighting by source strength, within the standard mixed model approach to penalized spline smoothing. The approach is appropriate when the effect of each source can be considered to be additive and the distance decay is radially symmetric. In keeping with the theme of this report, the method allows us to combine information about point emissions sources with other terms in a statistical modeling framework. We illustrate the method in the models developed in Sections 5 and 6, as well as a separate dataset on ultrafine (UF) and fine PM in a neighborhood in New York City, in Brooklyn.

D.2. Methods

D.2.1. Basic model

If we have a single source, we can represent a smooth effect of distance to source in the mixed model formulation of a penalized spline (Ruppert et al. 2003) as

$$f(d_i) = Z_i^T b \tag{D1}$$

where Z_i is a spline basis expansion, evaluated as a function of the distance, $d_i = \|s_i - s_0\|$, between the source, at location s_0 , and i th observation (the receptor) at s_i , and b is a set of basis coefficients. The coefficients are then assigned a random effects distribution, or equivalently, a prior distribution in a Bayesian context. Next suppose that we wish to account for

source strength, e , and that we can assume that the effect of the source on the concentration at arbitrary distance scales as $e \cdot f(d_i)$.

The key contribution is now to recognize that the linear structure of (D1) allows us to sum over source strengths and basis matrix values but retain the same basis coefficients, allowing us to represent the effect of multiple sources, $j = 1, \dots, J$, on a single receptor as

$$f(d_i) = \sum_j e_j Z_{ji}^T b = \left(\sum_j e_j Z_{ji}^T \right) b = \tilde{Z}_i^T b \quad (\text{D2})$$

for $d_i \equiv \{d_{i1}, \dots, d_{iJ}\}$ a vector of distances from receptor i to the J sources, with \tilde{Z}_i the i th row of a basis matrix constructed as the sum of basis matrices for individual sources, where each matrix is scaled by its source strength. Thus we can estimate b using the variety of approaches for estimating random effects based on \tilde{Z}_i and interpret (D1) as the smooth effect of distance to each source individually for a unit of emissions. \tilde{Z} can be precomputed as it is a function only of the distances between sources and receptors and the source strengths. The representation can be embedded in an additive or hierarchical model to account for the contribution of point emissions to pollution concentrations at arbitrary locations, with the distance decay estimated empirically.

In our applications we use a thin plate spline in one dimension, equivalent to a cubic radial basis function, using a basis matrix constructed as in Ruppert et al. (2003) and Crainiceanu et al. (2005):

$$Z = \left(\|d_i - \kappa_k\|_{k=1, \dots, K}^3 \right) \left(\|\kappa_k - \kappa_{k'}\|_{k, k'=1, \dots, K}^3 \right)^{-\frac{1}{2}}$$

with $\{\kappa_k\}$ is a set of K knots spaced over the range of distances present in a given domain. We specify the random effects distribution simply as $b \sim N(0, \tau^2 I)$. Distributions that penalize differences of neighboring coefficients, such as the Eilers and Marx (1996) approach are also possible, but when fit in standard random effects software these require specification of a more complicated covariance structure. For our mixed model formulation using the cubic radial basis function, our representation, $f(d_i) = \{f(d_{i1}, d_{iJ})\}$, also includes a fixed effect term (Ruppert et al.

2003, Crainiceanu et al. 2005), in this case simply $\left\{ \sum_j e_j d_{ji} \beta_d \right\}_{i=1, \dots, n} = \tilde{X}_d \beta_d$.

In practice, we include in the calculation of \tilde{Z}_i and $\tilde{X}_{d,i}$ only sources within d_{\max} of the receptor giving $J(i)$ sources for receptor i . This avoids estimating the decay function at distances for which we expect sources to have no effect on receptors and limits the computations involved in computing \tilde{Z} . Computation can also be limited by omitting very small sources. Having chosen d_{\max} , we choose knots on the interval $(0, d_{\max})$, with the inter-knot spacing increasing with distance to more easily capture sharp initial declines near to sources.

Note that our model assumes the effect of each additional unit of emissions is additive, contributing an independent incremental effect, and does not interact (e.g., chemically) with other emissions, although we note that the smooth function, f , to some extent accounts for interaction of emissions from an individual source, provided this interaction scales linearly with e_j . Thus the model is a simplification for some pollutants, depending on their physical and chemical dynamics.

Line and area sources may be represented as a fine grid of point sources. Given that the basis matrix can be precomputed, even a large number of grid points involves only a single slow calculation, done in advance of all model fitting. In Section D.3.2, we show an example for the effects of two highways on pollution in a neighborhood in New York City.

D.2.2. Monotonicity and decay constraints

In fitting $f(\cdot)$, we would like to ensure in most cases that the decay function declines monotonically (leaving aside the effect of source height in pollution contexts) and that at d_{\max} the function is zero. The former can be enforced using monotonic splines to build Z . The particular formulation we have considered is the monotonic b-spline, which consists simply of enforcing that $b_k > b_m$ for $k > m$ where the knot indices are ordered with increasing d and knots are equally spaced (He and Shi 1998). This constraint can be enforced in a Bayesian context within an MCMC by only proposing (or rejecting other proposals) coefficient values that respect monotonicity. Alternatively we transform the coefficients to ensure the constraint is satisfied with the new parameters being the log transformed differences in the coefficients for adjacent knots. This transformation helps move about the space in MCMC more efficiently than attempting to propose only sets of coefficients that respect monotonicity. To ensure that the emissions effect at d_{\max} is zero, simply place a knot at d_{\max} and fix the coefficient for that knot to be zero. Note that we have not found a simple way to enforce monotonicity when fitting in the random effect context.

For source effects that decay rapidly to zero, one concern is oversmoothing the function as it approaches zero and in concert undersmoothing the function at distances at which there is no effect of the source, where we estimate the function to be near zero. The basic problem is nonstationarity in the true decay function, with a sharp gradient at short distances and little or no gradient at longer distances, a feature that standard spline representations do not handle. Our experience in practice with ultrafine PM in the Brooklyn data is that we estimate a sharp drop up to distances of about 100m and then an essentially flat function, with wiggles that are likely an effect of undersmoothing. When we enforce monotonicity, the resulting estimate oversmooths the decay function as it approaches zero, with the estimated function flattening out more and more as the distance approaches d_{\max} . Thus, while monotonicity is conceptually appealing, the practical results have been discouraging. An ad hoc solution is to avoid monotonicity, estimate the decay function roughly using a large d_{\max} and, after identifying the likely point at which the decay nears zero, set d_{\max} near this value, so that one forces zero effect at larger distances. In our experience this has avoided much of the oversmoothing, while also avoiding undersmoothing at larger distances.

With regard to the decay to zero constraint when we do not force monotonicity, we enforce this a priori based on manipulation of the \tilde{Z} matrix. In particular, first construct \tilde{Z}_{\max} , which is the \tilde{Z} one would obtain if all sources within d_{\max} of a given receptor were placed exactly d_{\max} units away from every receptor. Then take $\tilde{Z}^* \equiv \tilde{Z} - \tilde{Z}_{\max}$, which ensures that the effect of a

source exactly d_{\max} units from a receptor is zero. Note that we also modify X_d similarly as $\tilde{X}_{d,i}^* = \tilde{X}_{d,i} - d_{\max} \sum_j e_j I(d_{ji} < d_{\max})$. Henceforth, I will refer to \tilde{X}_d^* and \tilde{Z}^* as simply \tilde{X} and \tilde{Z} .

D.2.3. Fitting the representation

To fit the model, if there were a single dominant source for each receptor, avoiding the summation in (D2), one could simply use `gam()` in R using the vector of emissions as the 'by' variable. This takes the form of a varying coefficient model in which the effect of emissions varies by distance. Instead, we can fit the model based on the mixed model representation as follows.

One option is to use MCMC in a fully Bayesian context in a conceptually straightforward manner. Alternatively, we would like to be able to fit the model in the standard mixed effects framework. If the source term (D2) enters into the mean in an additive fashion, we can fit using standard mixed effects software, precomputing \tilde{Z} and specifying that $b \sim N(0, \tau^2 I)$. However, often one fits a model on log-transformed outcomes, particularly when a pollutant is the outcome. Given that our approach represents the effect of multiple sources as adding over individual sources, this suggests that $f = \tilde{X}_d \beta_d + \tilde{Z}b$ should enter as $\log(\tilde{X}_d \beta_d + \tilde{Z}b)$. The first issue is that we have forced the effect to be zero for locations with no sources within d_{\max} , but we cannot take the log of zero. The second is that we cannot enter the log of a random effects term into standard software. Therefore, consider including in the mean for log outcome the term, $\log(\tilde{X}_d \beta_d + \tilde{Z}b + 1)$, such that for locations near no sources (for which $\tilde{X}_d \beta_d + \tilde{Z}b = 0$), the contribution to the log outcome is zero. Equivalently, exponentiating both outcome and the source term, we have that the source term simply multiplies the mean of the outcome by one. Next, note that we have the approximation, $\log(1+x) \approx x$ by Taylor expansion around zero. This suggests that in the log-transformed model, we can include $\tilde{X}_d \beta_d + \tilde{Z}b$ as simple linear terms.

Uncertainty about the fitted $f(\cdot)$ in the Bayesian context is easily estimated using samples of b from the posterior. For the mixed model formulation, we follow Ruppert et al. (2003, p. 103) and Wood (2006, p. 189) in using an empirical Bayes style approach, treating the estimated observation error variance and the variance component for b , τ^2 , as known, fixed at their point estimates. Consider estimating the joint covariance of $\hat{\theta} = \{\hat{\beta}, \hat{b}_{\text{full}}\}$ where $\hat{\beta}$ represents all the fixed effects coefficients in the model, including the linear term for the distance decay effect, β_d , and \hat{b}_{full} includes all the random effects in the model, including the coefficients for our distance decay smooth, b . Collecting the fixed and random effects design matrices as $C = (XZ)$, an estimate of $\text{Cov}(\hat{\theta})$ is

$$\left(\frac{C^T C + \hat{B}}{\hat{\sigma}^2} \right)^{-1}, \quad (\text{D3})$$

where $\hat{\sigma}^2$ is an estimate of the observation error variance and \hat{B} is a block diagonal matrix with the first block all zeroes, representing the prior precision (inverse variance) of β , and the second

block the inverse of the variance matrix for the random effects. In the simple setting with only random effects for our distance decay, the second block is a $K \times K$ matrix, $\frac{1}{\tau^2} I$. Empirical Bayes style pointwise intervals for $f(\cdot)$ on a fine grid can then be calculated by drawing samples, $\{\beta_d^*, b^*\}$, centered around $\{\hat{\beta}_d, \hat{b}\}$ with variance the appropriate subblock of $\text{Cov}(\theta)$ and calculating quantiles of $f_h = \tilde{X}_{d,h} \beta_d^* + \tilde{Z}_h^T b^*$ for d_h on a fine grid over $(0, d_{\max})$. To estimate prediction uncertainty for the collective influence of multiple sources at arbitrary location s_i , one can proceed in similar fashion using $\tilde{X}_{d,i} \beta_d^* + \tilde{Z}_i^T b^*$.

D.3. Examples

We illustrate the methodology in two examples. The first is the use of our representation to capture the effects of point source $\text{PM}_{2.5}$ emissions in our general modeling framework (Sections 5 and 6). The second is an analysis of ultrafine and fine PM in the Williamsburg neighborhood of Brooklyn, New York City, where a team of researchers led by Dr. Jonathon Levy and former graduate student Leonard Zwack at Harvard School of Public Health collected data using mobile backpack monitors over several weeks of field work in the vicinity of the Williamsburg Bridge and Brooklyn-Queens Expressway interchange (Zwack et al. 2011).

D.3.1. Modeling point source $\text{PM}_{2.5}$ emissions in the mid-Atlantic and eastern U.S.

Here β_d and b are fit using MCMC methods as part of our full MCMC for the models described in Section 6, with these coefficients integrated over in fitting the model as described in Appendix C. We considered point sources emitting more than 5 tons with $d_{\max} = 100$ km for our mid-Atlantic analyses, limiting to those emitting more than 10 tons with $d_{\max} = 50$ km in our eastern U.S. analyses to reduce computations. Given the drawbacks of the monotonic spline approach, outlined in Section D.2.2 and seen in the Brooklyn example, we have not enforced monotonicity. Note that in some our fits this has resulted in functions sufficiently non-monotone as to be of concern. Fig. D1 shows the monitor and point source locations providing information for the estimation.

To illustrate the decay functions we estimate, Fig. D2(a) shows the posterior mean for $f(\cdot)$ for each of 12 months, scaled to represent the effect of a single 1000 ton source as a function of distance to the receptor for the eastern U.S. analysis (Section 6). There is a suggestion of a yearly cycle with a more shallow decay during the summer months and steeper decay in the winter and transitional periods, although the 95% pointwise credible intervals indicated for a single month (March) as an example (Fig. D2(b)) are large enough that we cannot draw any definitive conclusion about this pattern. For a few of the months, some scientifically implausible non-monotonicity is evident as the distance gets larger than 20 km, but this falls within the uncertainty bands (as illustrated in Fig. D2(b)) and is relatively small in magnitude relative to the dropoff at small distances. With the number of terms in the model, concavity between the source term and other GIS-based covariates in the model make it difficult to conclude that the

estimated distance decay represents the real physical structure of decay in $PM_{2.5}$ away from the source, but the term plays a role in explaining variability in the $PM_{2.5}$ observations and the form of the estimated functions appears reasonable from a scientific perspective.

D.3.2. Modeling road source ultrafine PM and $PM_{2.5}$ emissions in Brooklyn

Data are available from mobile backpack-based monitoring providing approximately 5000 single minute averages in a several square kilometer area of Williamsburg, Brooklyn over 10 weekdays (9 a.m. to 5 p.m.) during a two week period, June 12-26, 2007. Interest focuses on the effects of two major roadways in the area: the Brooklyn-Queens Expressway (BQE) and the Williamsburg Bridge and its approach road. Weather data were available from a single fixed site on the edge of the area, near the Brooklyn base of the bridge.

Our basic model for the Brooklyn data was

$$\log Y_i = X_i^T \beta + g(s_i) + \log(f(d_i) + 1) + \varepsilon_i$$

where $g(s_i)$ accounts for large-scale spatial variation in pollution over the area, fit using a thin plate spline, and in $X_i^T \beta$ we include linear terms for wind speed, relative humidity, and temperature and dummy variables for individual days of observation. $f(d_i)$ represents the joint effect of multiple sources. More sophisticated modeling outside the scope of this report includes accounting for short-term temporal trends and residual temporal autocorrelation (Zwack et al., in prep.). To fit the model using mixed model software, we approximate the distance decay effect as

$$\log(f(d_i) + 1) \approx f(d_i) = \tilde{X}_{d,i}^T \beta_d + \tilde{Z}_i^T b$$

for $d_i = \{d_{i1}, \dots, d_{iJ(i)}\}$ (Section D.2.3).

Most observations in this dataset were with wind speeds less than two m/s; we subset the data to exclude observations under higher wind conditions to more plausibly model the effect of each source as being radially symmetric. We treat the expressway and bridge as line sources (each carriageway of these divided highways is treated as a separate line source), discretizing each line source as a series of point along the length of each roadway, separated by 10 m. We ignore traffic volume as volume is roughly constant over this daytime sampling period, and low volume periods often correspond to low vehicle speeds in the afternoon rush hour, suggesting that volume may not be a good proxy for emissions. Note that by using the log transformation, $f(d_i)$ can be interpreted as the percentage change in mean Y for a single 10m road segment on a single carriageway. For UF, we truncate original values larger than 150,000 and for $PM_{2.5}$, values larger than 75.

We fit the model using both MCMC and mixed model software, the latter in R using the function `gamm()` which calls `lme()` for the random effects fitting. Fig. D3(a,d) shows the estimated percentage change in pollution as a function of distance from a single 10 m segment of major roadway. As expected based on physical principles, the effect of UF drops off more sharply than for $PM_{2.5}$, with negligible effects beyond 100 m for UF. Note that for UF, the results for the random effects fit are very similar to those using MCMC, except the uncertainty is somewhat less, presumably because the estimated intervals condition on the fitted variance

components. For $PM_{2.5}$, the random effects approach estimates the variance component to be zero and the point estimates of the decay functions are different but within the context of wide uncertainty. Fig. D3(b,e) shows the combined effect of the distance decay and the spatial residual term on pollution on the log scale modeled over the entire neighborhood based on the MCMC fit, while Fig. D3(c,f) isolates the effect of the distance decay term. Note that in Fig. D3(c), the influence of the slight non-monotonicity seen in Fig. D3(a), summed over many 10 m segments, results in a unrealistic dip in the roadway effect (darkest blue areas). We also fit models using monotonic b-splines, described in Section D.2.2, but the estimated decay functions (not shown) oversmoothed in comparison to Fig. D3(a).

D.4. Discussion

We have presented a methodology for empirical estimation of source decay effects in a spatial context based on a spatial array of source and receptor locations. When radial symmetry is appropriate and with sufficient spatially distributed monitoring data, the model could serve in place of a deterministic model to estimate pollution dispersion. We believe the methodology may be more broadly useful than pollution exposure contexts, in particular for research in which the effect of an exposure, broadly defined, is expected to decay with distance. Social epidemiology and health services research are potential areas of application.

We note that in place of the estimated decay functions shown here, some sort of parameterized decay function might be used, but caution that such a parameterization would involve a nonlinear function of a decay parameter, without the ability to do the summation over sources in closed form, thereby increasing computation and requiring estimation via nonlinear least squares or other algorithms. In contrast, our approach can be fit with standard random effects software and the decay is determined from the data, without the need to decide upon a suitable decay function in advance.

We have considered a parameterization that enforces monotonicity, but this approach cannot be easily fit with random effects software and showed a tendency to oversmooth. However, non-monotone estimates of the decay function can cause undesirable artefacts in the estimated exposure surfaces when aggregated over multiple sources.

The major drawback of our statistical approach to estimating the source contributions from multiple sources is that the model does not account for the directional effects of wind in the pollution context. For this, one approach would be to use an asymmetric distance function rather than our spherically symmetric spline representation. Ideally one would have an asymmetric distance function that is relatively simply-parameterized and allows the distance decay to be sharp in the upwind direction and with a much longer dropoff in the downwind direction, with a parameterized contribution of wind direction and strength. One would also like a model that handles conditions of little or no wind and of strong winds. Clearly, ellipsoidal type representations such as bivariate Gaussian densities for correlated random variables will not do the trick because they are symmetric in the 'upwind' and 'downwind' directions.

If we were not concerned with changing wind speed and merely wanted to capture a long-term average effect of prevailing wind direction, one possibility would be to combine a spline representation of distance decay with a von Mises distribution, the extension of a Gaussian distribution to the circle, multiplying $Z_i^T b$ by $h(\theta_{ji})$ where θ_{ji} is the direction from source to

receptor and $h(\cdot)$ is the von Mises density. Unfortunately this immediately loses the key linearity property of (D1) that allows for summation over multiple sources. Further extension to introducing dependence on wind speed complicates matters further. Wind speed dependence might be induced by having the log of the von Mises variance be a regression on wind speed and possibly replacing the smooth distance decay function with a parameterized exponential or squared exponential decay in which the decay parameter varies with wind speed.

References

Baxter LK, Clougherty JE, Paciorek CJ, Wright RJ, Levy JI. 2007. Predicting residential indoor concentrations of nitrogen dioxide, fine particulate matter, and elemental carbon using questionnaire and geographic information system based data. *Atmospheric Environment* 41: 6561–6571.

Crainiceanu CM, Ruppert D, Wand MP. 2005. Bayesian analysis for penalized spline regression using WinBUGS. *Journal of Statistical Software* 14: 1–24.

Eilers PHC, Marx BD. 1996. Flexible smoothing with B-splines and penalties. *Statistical Science* 11: 89–102.

Farley TA, Mason K, Rice J, Habel JD, Scribner R, Cohen DA. 2006. The relationship between the neighbourhood environment and adverse birth outcomes. *Paediatric & Perinatal Epidemiology* 20: 188–200.

He X, Shi P. 1998. Monotone B-spline smoothing. *Journal of the American Statistical Association* 93: 643–650.

Laraia BA, Siega-Riz AM, Kaufman JS, Jones SJ. 2004. Proximity of supermarkets is positively associated with diet quality index for pregnancy. *Preventive Medicine* 39: 869–875.

Morland K, Wing S, Roux AD. 2002. The contextual effect of the local food environment on residents' diets: The atherosclerosis risk in communities study. *American Journal of Public Health* 92: 1761–1768.

Ruppert D, Wand MP, Carroll RJ. 2003. *Semiparametric Regression*. Cambridge University Press, Cambridge.

Sallis JF, Hovell MF, Hofstetter CR, Elder JP, Hackley M, Caspersen CJ, Powell KE. 1990. Distance between homes and exercise facilities related to frequency of exercise among San Diego residents. *Public Health Rep* 105: 179–85.

Wood SN. 2006. *Generalized Additive Models: An Introduction with R*. Chapman & Hall, Boca Raton.

Yanosky JD, Paciorek CJ, Suh HH. 2009. Predicting chronic fine and coarse particulate exposures using spatio-temporal models for the northeastern and midwestern U.S. *Environmental Health Perspectives* 117: 522–529.

Zwack, L, Paciorek CJ, Spengler J, Levy, J. 2011. Modeling spatial patterns of traffic-related air pollutants in complex urban terrain. *Environmental Health Perspectives* 119: 852-859.

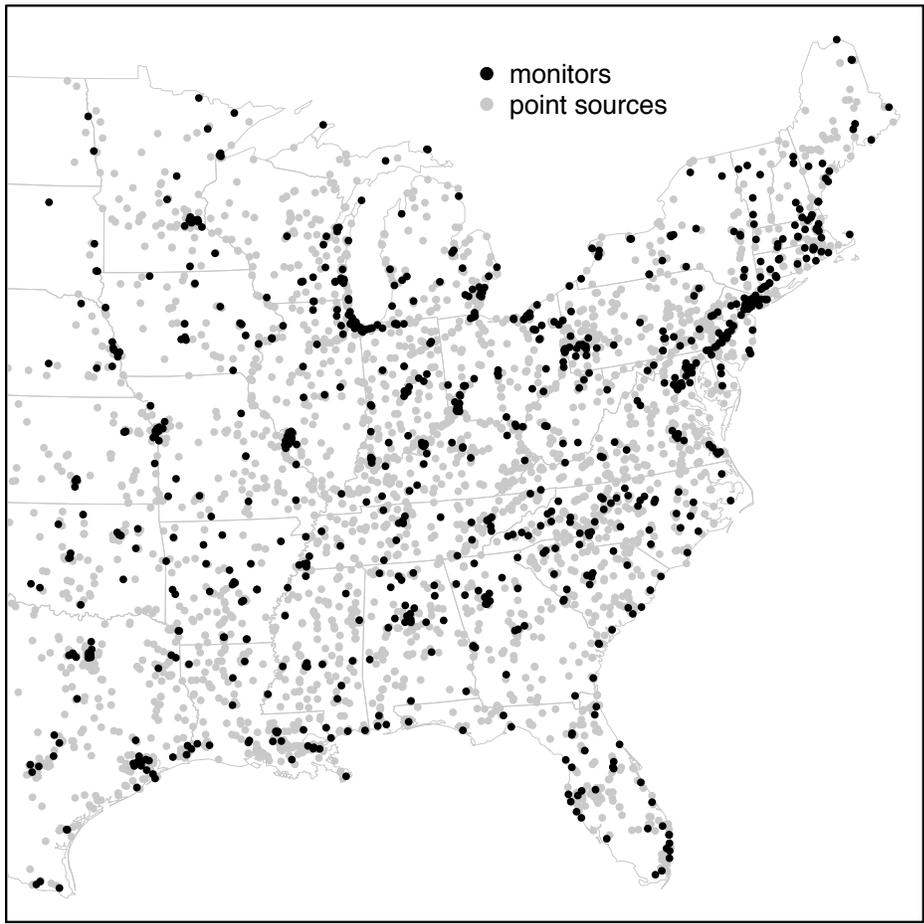


Figure D1. Locations of PM_{2.5} monitors reporting in 2001 (black) and point sources emitting more than 10 tons of primary PM_{2.5} emissions in 2002 based on the National Emissions Inventory (grey) in the eastern U.S.

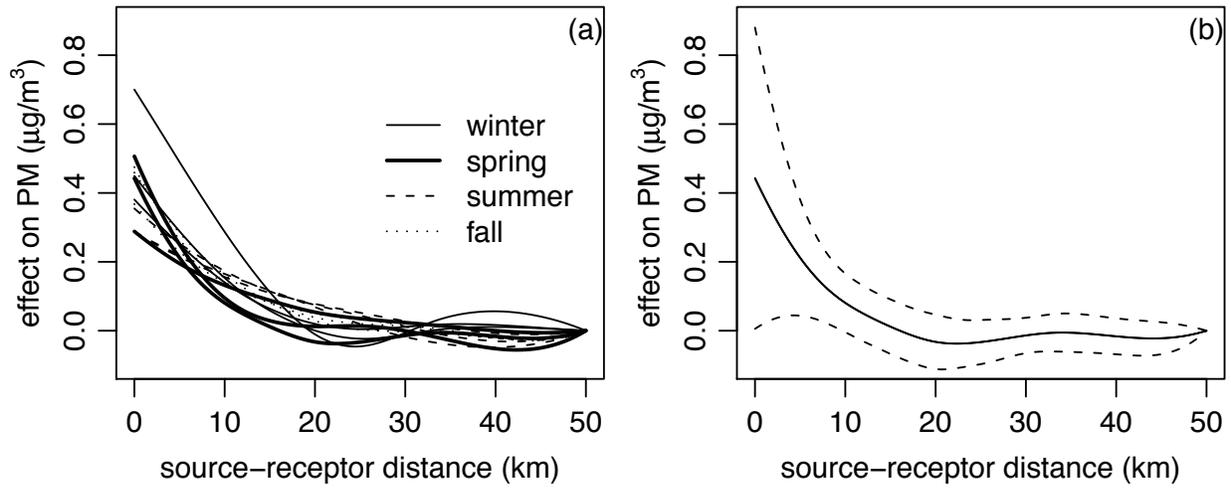


Figure D2. (a) Posterior means of the distance decay functions, $f(\cdot)$, scaled to estimate the effect of a single 1000 ton source as a function of source-receptor distance for each of 12 months, distinguished by season, for $\text{PM}_{2.5}$ in the eastern U.S. in 2001. (b) Example of 95% pointwise credible intervals for a single month (March); other months show a similar or moderately smaller amount of uncertainty.

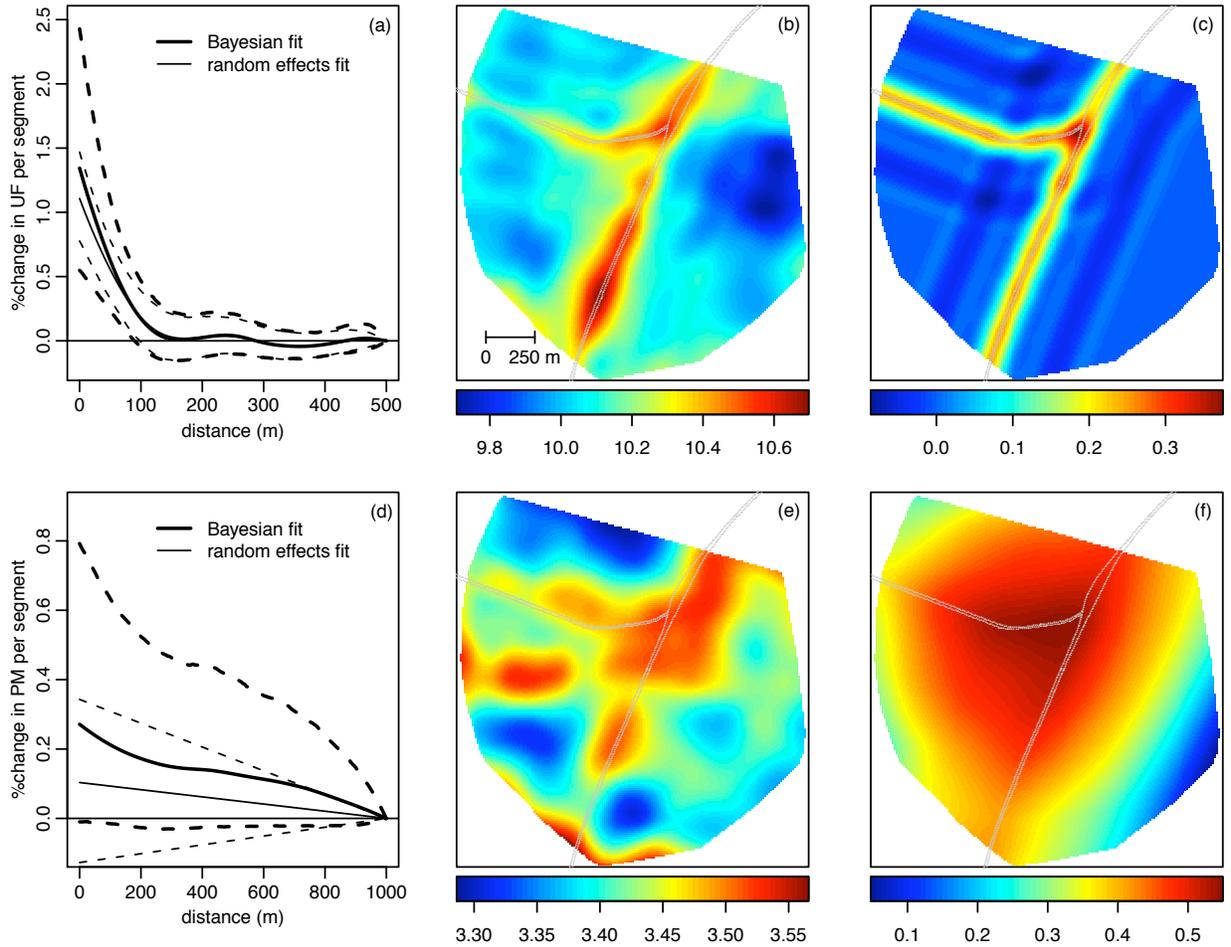


Figure D3. For the effects of large roadways in the Williamsburg section of Brooklyn, New York, posterior means and pointwise 95% credible intervals for the distance decay functions, $f(\cdot)$, for UF (a) and $PM_{2.5}$ (d), scaled to represent the percentage change in pollution from a single 10 m road segment as a function of source-receptor distance. Results based on the random effects fit using `gamm()` in R are shown in red (thick line in the greyscale version). Posterior mean log pollution surfaces in units of $\log(\text{count}/\text{cm}^3)$ for UF (b) and $\log \mu\text{g}/\text{m}^3$ for $PM_{2.5}$ (e). Posterior mean incremental influence of the distance decay effect on log pollution for UF (c) and $PM_{2.5}$ (f). In (b), (c), (e), and (f), the grey dots indicate the two carriageways for each of the Williamsburg Bridge and approach road (upper left of each subplot) and BQE (middle, slanting slightly from left to right).