Biostatistics (2015), **0**, 0, *pp.* 1–24
doi:10.1093/biostatistics/"Spatial SIMEX paper Alexeeff for Biostatistics R1"

Spatial measurement error and correction by spatial SIMEX in linear regression models when using predicted air pollution exposures

Stacey E. Alexeeff[1,*], Raymond J. Carroll[2], and Brent Coull[3]

[1]Institute for Mathematics Applied to Geosciences, National Center for Atmospheric Research, Boulder, CO.

[2]Department of Statistics, Texas A & M University, College Station, TX.

[3]Department of Biostatistics, Harvard School of Public Health, Boston, MA.

salexeeff@ucar.edu

Summary

Spatial modeling of air-pollution exposures has become widespread in air pollution epidemiology research as a way to improve exposure assessment. However, there are key sources of exposure model uncertainty when air pollution is modeled, including estimation error and model misspecification. We examine the use of predicted air pollution levels in linear health effect models under a measurement error framework. For the prediction of air pollution exposures, we consider a universal kriging framework, which may include land use regression terms in the mean function and a spatial covariance structure for the residuals. We derive the bias induced by estimation error and by model misspecification in the exposure model, and we find that a misspecified exposure model can induce asymptotic bias in the effect estimate of air pollution on health. We propose a new spatial SIMEX procedure, and we demonstrate that the procedure has good performance in correcting this asymptotic bias. We use a bootstrap procedure to estimate the standard errors in the spatial SIMEX method. We illustrate the spatial SIMEX

*To whom correspondence should be addressed.

approach in a study of air pollution and birthweight in Massachusetts.

Key words: air pollution; birthweight; environmental epidemiology; kriging; model uncertainty; spatial model.

1. Introduction

There is strong evidence in reviews of epidemiological studies that both short-term and long-term exposures to air pollution are related to cardiovascular morbidity and mortality (Brook *and others*, 2010). Spatial modeling of air-pollution levels using ordinary kriging and universal kriging methods is now commonplace in air pollution epidemiology research. Existing pollution monitoring networks are used to collect data on regional air pollution concentrations, and spatial prediction models are then used to estimate location-specific exposures at the home address of each subject in a study. However, air pollution levels are typically measured only at a small number of monitors, while the air pollution surface varies over space and the overall regional heterogeneity may be difficult to characterize. For example, ambient levels of $PM_{2.5}$ have been shown to vary considerably within a given city region, in part due to traffic sources (Brauer *and others*, 2003; Clougherty *and others*, 2008).

This measurement error setting where exposure is predicted at a set of exposure locations of interest that does not match the set of measured exposure locations is called *spatial misalignment* (Gryparis *and others*, 2009). The most common use of predicted exposures in a health effects analysis is the direct plug-in of the individual-specific exposure estimates. This approach treats the exposures as known, without acknowledgement of the uncertainty in the prediction process. Ignoring measurement error can lead to biased health effect estimates and overstated confidence in the resulting risk assessments (Carroll *and others*, 2006). The issue of spatial misalignment measurement error in air pollution epidemiology has received considerable attention in recent literature (Szpiro *and others*, 2011; Szpiro and Paciorek, 2013; Madsen *and others*, 2008; Lopiano *and others*, 2011, 2013; Gryparis *and others*, 2009; Peng and Bell, 2010; Bergen *and others*, 2013). However, the role of spatial exposure model misspecification has not formally been investigated.

The work presented here adds to this existing literature in two ways. First, almost all of these existing studies focus on the impact of uncertainty associated with estimation of unknown parameters in a known exposure model, not on the impact of model misspecification. In this work, we explicitly evaluate the impact of model misspecification, which in reality will always be an issue given the spatial-temporal complexity of pollutions emissions and the atmospheric fate and transport of these emissions. Second, we propose methods that can be implemented in settings in which the exposure prediction algorithm is a complex process that does not yield uncertainties for the exposure model parameter estimates. Examples of such approaches that have recently appeared in the environmental science literature include a daily kriging model implemented in ArcGIS that does not yield uncertainties for the kriging variance component estimates (Liao *and others*, 2006) and multi-stage model fitting and missing data imputation schemes that make it difficult to propagate the uncertainty associated with each stage through to the final exposure predictions (Kloog *and others*, 2012, 2014; Nordio *and others*, 2013).

This article investigates two factors of exposure estimation that may affect resulting health effect estimates: *estimation error* and *model misspecification*. In practice, spatial air pollution models are fit with sparse monitoring data. Hence, we examine the effects of estimation error in the kriging model parameters under small sample size. In addition, the underlying exposure model that generates air pollution levels in any given region is not known. Thus, we investigate the impact of model misspecification in the spatial model by omitting a spatial covariate. To correct for bias in the health effect estimates, we introduce a spatial version of the simulation extrapolation method (SIMEX). To our knowledge, our proposed spatial SIMEX procedure is the first treatment of SIMEX allowing for spatially correlated measurement errors.

The remainder of this paper is arranged as follows. In Section 2 we introduce our modeling framework and the specific exposure models of interest under universal kriging. In Section 3 we examine the bias analytically for each of the exposure models of interest and we derive the probability limits of the misspecified parameters and the resulting classical error variance. In Section 4 we propose a new spatial SIMEX correction method where correlated classical error is added to the exposure predictions to correct for bias. In Section 5 we present a simulation study to investigate the degree of bias induced by each of the exposure models of interest and to demonstrate the performance of our spatial SIMEX correction method. We then illustrate the spatial SIMEX correction in a study of air pollution and birthweight in the greater Boston area in Section 6. We end with a concluding discussion in Section 7.

2. Spatial Exposure Models in Air Pollution and Health Studies

2.1 *Model Framework*

Let the health effect model of interest be a simple linear regression model, $Y_i = \beta_0 + \beta_X X_i + \epsilon_i$, for each subject $i = 1, \ldots, n$, where Y_i is a continuous health outcome, X_i is the continuous air pollution exposure of interest, and ϵ_i are independent and identically distributed (i.i.d.) with mean 0 and variance σ_ϵ^2. The goal of the analysis is to estimate β_X, the parameter measuring the association between the health outcome and the air pollution exposure.

Let $\mathbf{X}$ denote the n-length vector of true unmeasured air pollution exposures corresponding to the subject's address location. Let $\mathbf{X}^*$ denote the m-length vector of measured air pollution levels at m monitor sites spread throughout the same geographic region, where $m \ll n$. We assume the setting of *spatial misalignment*, where the n subject address locations needed for the health model do not match any of the m monitor locations.

We consider a universal kriging model for the exposure which may include covariates as part of the mean model. Suppose the true pollution process, $\mathcal{X} = (\mathbf{X}, \mathbf{X}^*)$, is generated by a Gaussian Random Field, and that the realizations of this process take a parametric form. Specifically, denote the length K vector of parameters for the mean model as $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)$, and denote the length J vector of parameters for the variance model as $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_J)$, and let $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\psi})$. Then a realization of one surface follows a Multivariate Normal distribution,

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{X}^* \end{pmatrix} \sim \mathcal{N} \left\{ \begin{pmatrix} \boldsymbol{\mu}_X(\boldsymbol{\alpha}) \\ \boldsymbol{\mu}_{X^*}(\boldsymbol{\alpha}) \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{XX}(\boldsymbol{\psi}) & \boldsymbol{\Sigma}_{XX^*}(\boldsymbol{\psi}) \\ \boldsymbol{\Sigma}_{X^*X}(\boldsymbol{\psi}) & \boldsymbol{\Sigma}_{X^*X^*}(\boldsymbol{\psi}) \end{pmatrix} \right\}. \tag{2.1}$$

The kriging estimator for $\mathbf{X}$ conditional on the observed monitor data $\mathbf{X}^*$ is defined as

$$g_X(\boldsymbol{\theta}; \mathbf{X}^*) = \boldsymbol{\mu}_X(\boldsymbol{\alpha}) + \boldsymbol{\Sigma}_{XX^*}(\boldsymbol{\psi}) \boldsymbol{\Sigma}_{X^*X^*}^{-1}(\boldsymbol{\psi}) \left\{ \mathbf{X}^* - \boldsymbol{\mu}_{X^*}(\boldsymbol{\alpha}) \right\}. \tag{2.2}$$

We now consider two spatial exposure model scenarios in this study.

Scenario I: Universal kriging model. Consider equation (2.2) and assume that the variance $\boldsymbol{\Sigma}_{\mathcal{X}}(\boldsymbol{\psi})$ follows a Matérn family with variance parameters $\boldsymbol{\psi} = (\phi, \sigma^2, \nu)$, where ϕ is the range, ν is the smoothness, and σ^2 is the variance (see Appendix in Supplementary Material for explicit Matérn covariance function). Universal kriging assumes a spatial correlation structure for the variance, and allows the mean model to depend linearly on a set of covariates or to be constant. For the mean model we consider two scenarios. First, we consider a constant mean model, $\mu_{\mathcal{X}}(\boldsymbol{\alpha}) = \alpha$, which we refer to as Scenario IA. Second, we consider a mean model that depends linearly on covariates, $\mu_X(\boldsymbol{\alpha}) = \alpha_0 + \alpha_1 S_1 + \alpha_2 S_2$, where we assume covariates, S_1, S_2 represent spatially-varying land-use characteristics. This type of model is often called a "land-use regression" model, which we refer to as Scenario 1B. Land-use covariates for air pollution models include measures such as percentages of residential land, greenspace, and industry, population size, distances to major roads, and traffic intensity (Ross *and others*, 2007; Eeftens *and others*, 2012).

Scenario II: Misspecified universal kriging model. Scenario II considers a misspecified universal kriging model. We assume that the true exposure is generated under the universal kriging model defined above, and that the fitted exposure model is misspecified by omitting $\mathbf{S}_2$. Thus, the misspecified exposure model that is fit is

$$g_X\left(\boldsymbol{\theta}_N; \mathbf{X}^*\right) = \boldsymbol{\mu}_X\left(\boldsymbol{\alpha}_N\right) + \boldsymbol{\Sigma}_{XX^*}\left(\boldsymbol{\psi}_N\right) \boldsymbol{\Sigma}_{X^*X^*}^{-1}\left(\boldsymbol{\psi}_N\right) \left\{\mathbf{X}^* - \boldsymbol{\mu}_{X^*}\left(\boldsymbol{\alpha}_N\right)\right\} \tag{2.3}$$

where $\boldsymbol{\alpha}_N = (\alpha_{0,N}, \alpha_{1,N}, 0)$, $\boldsymbol{\mu}_{\mathcal{X}}(\boldsymbol{\alpha}) = \alpha_{0,N}\mathbf{1} + \alpha_{1,N}\mathbf{S}_1$, and $\boldsymbol{\psi}_N = \left(\phi_N, \sigma_N^2, \nu_N\right)$, with the subscript N denoting the naive parameters from the misspecified model. Here we also assume that $\mathbf{S}_1$ and $\mathbf{S}_2$ are each spatially correlated, generated from their own Gaussian Processes.

2.2 *Decomposition into Berkson and Classical error components*

We now review and extend the decomposition of exposure measurement error into Berkson and Classical components for each of these three scenarios. The mean and variance parameters can be estimated jointly via maximum likelihood. Let $\widehat{\boldsymbol{\theta}}$ be the vector of maximum likelihood estimates of $\boldsymbol{\theta}$ for the true model and let $\widehat{\boldsymbol{\theta}}_N$ be the vector of maximum likelihood estimates of $\boldsymbol{\theta}_N$ in the naive misspecified model. A general measurement error framework can be used to characterize the difference between the true unobserved exposures $\mathbf{X}$ and the predicted exposures $g_X(\widehat{\boldsymbol{\theta}};\mathbf{X}^*)$. The decomposition of errors into Berkson and classical type measurement error components follows the development of Gryparis *and others* (2009) and Szpiro *and others* (2011), where Gryparis *and others* (2009) considers a Bayesian Gaussian Process model with constant mean, and Szpiro *and others* (2011) considers a universal kriging model where the mean depends on several spatial covariates. We now extend this viewpoint to our models of interest defined in Secion 2.1.

For scenario I, the predicted exposures, $g_X(\widehat{\boldsymbol{\theta}};\mathbf{X}^*)$, have the form

$$g_X(\widehat{\boldsymbol{\theta}};\mathbf{X}^*) = \boldsymbol{\mu}_X\left(\widehat{\boldsymbol{\alpha}}\right) + \boldsymbol{\Sigma}_{XX^*}\left(\widehat{\boldsymbol{\psi}}\right)\boldsymbol{\Sigma}^{-1}_{X^*X^*}\left(\widehat{\boldsymbol{\psi}}\right)\left\{\mathbf{X}^* - \boldsymbol{\mu}_{X^*}(\widehat{\boldsymbol{\alpha}})\right\}. \tag{2.4}$$

We decompose the error into Berkson and classical components,

$$\mathbf{X} - g_X\left(\widehat{\boldsymbol{\theta}};\mathbf{X}^*\right) = \mathbf{X} - g_X\left(\boldsymbol{\theta};\mathbf{X}^*\right) + g_X\left(\boldsymbol{\theta};\mathbf{X}^*\right) - g_X\left(\widehat{\boldsymbol{\theta}};\mathbf{X}^*\right). \tag{2.5}$$

The Berkson error term, $\mathcal{U}_b = \mathbf{X} - g_X\left(\boldsymbol{\theta};\mathbf{X}^*\right)$, represents the difference between the true measurements and the expectation of $\mathbf{X}$ conditional on $\mathbf{X}^*$. The classical error term, $\mathcal{U}_c = g_X\left(\boldsymbol{\theta};\mathbf{X}^*\right) - g_X\left(\widehat{\boldsymbol{\theta}};\mathbf{X}^*\right)$, represents the difference between the true model and the estimated model. Thus, we can call this classical error *estimation error*. In scenario I, this kriging estimator fit under the correct model is the best linear unbiased predictor (BLUP) (Cressie, 1993).

For scenario II, the predicted exposures, $g_X(\widehat{\boldsymbol{\theta}}_N;\mathbf{X}^*)$, have the form

$$g_X\left(\widehat{\boldsymbol{\theta}}_N;\mathbf{X}^*\right) = \boldsymbol{\mu}_X\left(\widehat{\boldsymbol{\alpha}}_N\right) + \boldsymbol{\Sigma}_{XX^*}\left(\widehat{\boldsymbol{\psi}}_N\right)\boldsymbol{\Sigma}^{-1}_{X^*X^*}\left(\widehat{\boldsymbol{\psi}}_N\right)\left\{\mathbf{X}^* - \boldsymbol{\mu}_{X^*}(\widehat{\boldsymbol{\alpha}}_N)\right\}. \tag{2.6}$$

Decomposing the error into Berkson and classical components,

$$\mathbf{X} - g_X\left(\widehat{\boldsymbol{\theta}}_N;\mathbf{X}^*\right) = \underbrace{\mathbf{X} - g_X\left(\boldsymbol{\theta};\mathbf{X}^*\right)}_{\mathcal{U}_b} + \underbrace{g_X\left(\boldsymbol{\theta};\mathbf{X}^*\right) - g_X\left(\boldsymbol{\theta}_N;\mathbf{X}^*\right)}_{\mathcal{U}_c,\text{ model misspecification}} + \underbrace{g_X\left(\boldsymbol{\theta}_N;\mathbf{X}^*\right) - g_X\left(\widehat{\boldsymbol{\theta}}_N;\mathbf{X}^*\right)}_{\mathcal{U}_c,\text{ estimation error}}. \tag{2.7}$$

In Scenario II, there are two classical measurement error components. The first is attributed to choosing the incorrect model, and the second is due purely to estimation error of the parameters.

3. Analysis of bias in health effect estimates induced by exposure models

Now, with the measurement error framework established, we study the impact of measurement error in the predicted exposures on bias of the coefficient β_X representing the association between air pollution exposure and the health outcome. First we investigate bias in the case of estimation error only in Scenario I analytically, focusing on the small-sample bias properties not previously addressed in other studies. Next, we study the asymptotic bias in the case of model misspecification error in Scenario II by deriving the probability limits of the MLEs in the misspecified model and deriving the particular form of the classical error variance. Later, in Section 5, we will complement this analysis with a simulation study.

3.1 *Bias Analysis for Scenario I*

To study the estimation error bias in Scenario I, we introduce notation for the least squares estimators. Without loss of generality we assume centered variables. Let $M(\,\cdot\,;\mathbf{X}^*,\mathbf{Y})$ be the function for the least squares estimate of β_X given monitoring data, spatial covariates, and observed health outcomes. Our notation explicitly shows the dependence on $\mathbf{X}^*$ and $\mathbf{Y}$, but implicitly this also depends on $\mathbf{S}^*, \mathbf{S}$ as well. Specifically, define

$$M(\boldsymbol{\theta};\mathbf{X}^*,\mathbf{Y}) \equiv g_X(\boldsymbol{\theta};\mathbf{X}^*)^{\mathsf{T}}\mathbf{Y}/\left\{g_X(\boldsymbol{\theta};\mathbf{X}^*)^{\mathsf{T}}g_X(\boldsymbol{\theta};\mathbf{X}^*)\right\} \tag{3.8}$$

for $n\times 1$ vector $\mathbf{Y}$, $(J+K)\times n$ matrix $\mathbf{S}$, and $(J+K)\times 1$ vector $\boldsymbol{\theta}$. Then let $\widehat{\beta}_{X,\boldsymbol{\theta}}$ denote the least squares estimate of β_X based on the exposure model using the true parameters $\boldsymbol{\theta}$, so $\widehat{\beta}_{X,\boldsymbol{\theta}} = M(\boldsymbol{\theta};\mathbf{X}^*,\mathbf{Y})$. Similarly, let $\widehat{\beta}_{X,\widehat{\boldsymbol{\theta}}}$ denote the least squares estimate of β_X based on the exposure model using the estimated exposure model parameters $\widehat{\boldsymbol{\theta}}$, so $\widehat{\beta}_{X,\widehat{\boldsymbol{\theta}}} = M(\widehat{\boldsymbol{\theta}};\mathbf{X}^*,\mathbf{Y})$.

First, we note that the Berkson error component does not induce any bias in the estimate of β_X. Hence, any bias in the estimator comes from the classical error component. Using a second order Taylor expansion of $M(\widehat{\boldsymbol{\theta}};\mathbf{X}^*,\mathbf{Y})$ around $M(\boldsymbol{\theta};\mathbf{X}^*,\mathbf{Y})$ the approximate bias of $\widehat{\beta}_{X,\widehat{\boldsymbol{\theta}}}$ is

$$E\left\{\widehat{\beta}_{X,\boldsymbol{\theta}}-\widehat{\beta}_{X,\widehat{\boldsymbol{\theta}}}\right\}\approx\boldsymbol{\gamma}E\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)+\frac{1}{2}\text{trace}\left\{\boldsymbol{\Lambda}\mathsf{Var}\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)\right\}+\frac{1}{2}E\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^{\mathsf{T}}\boldsymbol{\Lambda}E\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right) \tag{3.9}$$

where $\boldsymbol{\gamma}=E\left[\frac{\partial}{\partial\theta}M(\boldsymbol{\theta};\mathbf{X}^*,\mathbf{Y},\mathbf{S})\right]^{\mathsf{T}}$ and $\boldsymbol{\Lambda}=E\left[\frac{\partial^2}{\partial\theta\partial\theta^{\mathsf{T}}}M(\boldsymbol{\theta};\mathbf{X}^*,\mathbf{Y},\mathbf{S})\right]$. Equation (3.9) illustrates that when the number of monitors m is large, then $E(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})\to 0$, resulting in an asymptotically unbiased estimator for β_X. However, in practice when m is small, then Jensen's inequality gives us the result that $E\left(\widehat{\beta}_{X,\widehat{\boldsymbol{\theta}}}\right)\neq\beta_X$ because $M(\cdot;\mathbf{X}^*,\mathbf{Y})$ is a nonlinear function of the exposure model covariance parameters. This is true even under the condition of an unbiased estimate of the covariance parameters, $E(\widehat{\boldsymbol{\theta}})=\boldsymbol{\theta}$. Specifically, following a similar argument to Zimmerman and Cressie (1992), Jensen's inequality says that if $M(\cdot)$ is strictly concave, then

$$E\left(\widehat{\beta}_{X,\widehat{\boldsymbol{\theta}}}\right)=E\{M(\widehat{\boldsymbol{\theta}};\mathbf{X}^*,\mathbf{Y})\}<M\{E(\widehat{\boldsymbol{\theta}};\mathbf{X}^*,\mathbf{Y})\}=M(\boldsymbol{\theta};\mathbf{X}^*,\mathbf{Y})=\beta_X. \tag{3.10}$$

Similarly, if $M(\cdot;\mathbf{X}^*,\mathbf{Y})$ is strictly convex, the resulting bias is upward and only linearity of $M(\cdot;\mathbf{X}^*,\mathbf{Y})$ yields an unbiased estimator. Supplementary Figure 1 shows an example of $M(\cdot;\mathbf{X}^*,\mathbf{Y})$ as a concave function for a particular choice of covariance matrix and set of covariates. Although we know that $M(\cdot;\mathbf{X}^*,\mathbf{Y})$ is a nonlinear function of the exposure model covariance parameters, its form depends on the spatial covariance function, the distances of the monitors from each other and the distances of the subject addresses from the monitors, all of which affect the matrix multiplications of the covariances to generate $g_X(\boldsymbol{\theta};\mathbf{X}^*)$. The scale of the nonlinear functions shown in Supplementary Figure S1 suggest that this bias may be small. Thus, in practice, we expect to see some small sample bias in $\widehat{\beta}_{X,\widehat{\boldsymbol{\theta}}}$ due to a small number of monitors m. We investigate this small sample bias further in a simulation study in Section 5.

3.2 *Bias Analysis for Scenario II*

Scenario II contains the additional component of error due to model misspecification, $g_X(\boldsymbol{\theta};\mathbf{X}^*)-g_X(\boldsymbol{\theta}_N;\mathbf{X}^*)$. To understand the bias induced by this component, we first consider the asymptotic behavior of the naive model parameter estimates. Following Wang *and others* (1998), the MLE's $\widehat{\boldsymbol{\theta}}_N$ will be the solutions to the score equations based on the Multivariate Normal likelihood for Equation (2.3), and thus will converge in probability to the solutions of the following equations:

$$E\left\{\mathbf{S}_N^{*\mathsf{T}}\mathbf{V}_N^{-1}(\mathbf{X}^* - \mathbf{S}_N^*\boldsymbol{\alpha}_N)\right\} = \mathbf{0} \tag{3.11}$$

$$\frac{1}{2}\left[E\left\{\left(\mathbf{X}^* - \mathbf{S}_N^*\boldsymbol{\alpha}_N\right)^{\mathsf{T}}\mathbf{V}_N^{-1}\frac{\partial \mathbf{V}_N}{\partial \psi_{l,N}}\mathbf{V}_N^{-1}\left(\mathbf{X}^* - \mathbf{S}_N^*\boldsymbol{\alpha}_N\right)\right\} - \mathrm{tr}\left\{\mathbf{V}_N^{-1}\frac{\partial \mathbf{V}_N}{\partial \psi_{l,N}}\right\}\right] = \mathbf{0} \tag{3.12}$$

where $\mathbf{S}_N^* = (\mathbf{1}, \mathbf{S}_1^*)$, the $m \times 2$ subset of spatial covariates in the misspecified model, $\mathbf{V}_N^{-1} = \Sigma_{X^*X^*}^{-1}(\boldsymbol{\psi}_N)$, and $l = (1,2,3)$ indexes the variance parameters. Solving these equations for $\boldsymbol{\theta}_N$ yileds the asymptotic relationship between $\boldsymbol{\theta}_N$ and $\boldsymbol{\theta}$, as shown in the Supplementary Material Appendix. Closed-form solutions exist for the naive model mean parameters, but for the variance parameters we derive equations which can be only be solved numerically. In general, the solutions depend on the joint distribution of the correlated spatial covariates.

4. SIMEX for correlated Berkson and Classical Errors

The simulation extrapolation method (SIMEX) has been developed as a flexible method to correct for the effect of classical measurement errors on the estimation of a parameter of interest (Cook and Stefanski, 1994). SIMEX is a functional method which uses resampling techniques and has several attractive properties, including placing minimal assumptions on the underlying distribution of the exposures. SIMEX has two steps: a simulation step (SIM), where simulated error is added to the mismeasured exposures in increasing amounts, and an extrapolation step (EX), where a trend is fit to the mean of the parameter estimates over the increasing error levels and extrapolated back to the case of no error. It has been suggested that SIMEX may be suitable for several exposures with correlated classical errors when the correlations of the errors are known or estimable (Carroll *and others*, 2006). We now present the spatial SIMEX procedure, an extension of SIMEX that allows the classical measurement errors to be correlated over space.

4.1 *Spatial SIMEX procedure*

The spatial SIMEX procedure is implemented as follows. Let R and L be given positive integers, and let $(\lambda_1, \ldots, \lambda_L)$ be an increasing sequence of nonnegative numbers starting with $\lambda_1 = 0$.

Simulation Step. For each $\lambda \in (\lambda_1, \ldots, \lambda_L)$ and $r = 1, \ldots R$, we generate a pseudo-dataset for the vector of exposures, $\mathbf{W}^{(r)}(\lambda) = \widehat{\mathbf{X}} + \sqrt{\lambda}\, \mathcal{U}_c^{(r)},$ where $\mathcal{U}_c^{(r)} \sim \mathcal{N}(0, \mathbf{\Sigma}_c)$. Adding the error creates pseudo datasets equal to the unbiased exposure plus an error component with covariance of $(1 + \lambda)\mathbf{\Sigma}_c$. This allows exploration of how the health effect parameter is biased as a function of increasing measurement error variance. For each λ and r, we estimate the parameter of interest $\beta_X^{(r)}(\lambda)$ by fitting the linear health model using the pseudo-dataset. Thus, $\widehat{\beta}_X^{(r)}(\lambda)$ estimates the association between the pseudo-exposures $\mathbf{W}^{(r)}(\lambda)$ and the outcome $\mathbf{Y}$.

Extrapolation Step. We obtain an estimate of $\widehat{\beta}_X$ for each λ by averaging over the R simulations, $\widehat{\beta}_X(\lambda) = \frac{1}{R}\sum_{r=1}^{R} \widehat{\beta}_X^{(r)}(\lambda)$. We then fit a trend to $\widehat{\beta}_X(\lambda)$ versus λ using a linear or quadratic model. The predicted value of this trend at $\lambda = -1$ is the spatial SIMEX corrected estimate of the parameter, $\widehat{\beta}_{X,\text{SIMEX}}$, estimating the health effect parameter under no measurement error.

Spatial SIMEX can be implemented with a bootstrap standard error estimate, following the same general bootstrap approach used for one-dimensional SIMEX (Carroll *and others*, 2006). To implement the bootstrap standard error, we first estimate $\widehat{\beta}_{X,\text{SIMEX}}$. Then, for $k = 1, ..., K$ bootstrap samples: (1) resample monitor locations with replacement, (2) fit the initial exposure model to the new sample of monitoring data, (3) predict the exposures at the health locations, and (4) repeat the entire SIMEX procedure using these new predictions to obtain $\widehat{\beta}_{X,\text{SIMEX}}^{(k)}$. The standard error estimate is then computed by the standard deviation of the K bootstrap SIMEX estimates, $\hat{s.e.}(\widehat{\beta}_{X,\text{SIMEX}}) = (K-1)^{-1}\sum_{k=1}^{K}\left(\widehat{\beta}_{X,\text{SIMEX}}^{(k)} - \widehat{\beta}_{X,\text{SIMEX}}\right)^2$.

In general, the asymptotic results of and Cook and Stefanski (1994) for the unbiasedness of the point estimate apply when (i) the bias in the naive estimator is a continuous function of the measurement error variance, (ii) the measurement error variance is known, and (iii) the true extrapolant function based on the bias function is known. In practice, the measurement error variance and the true extrapolant function are often unknown. Still, even an approximate exrapolant function can help reduce bias (Carroll *and others*, 2006). In addition, the degree of fit to the error-inflated parameters can be assessed, and if the trend in bias is unclear the number of SIMEX samples R can be increased as well as the number of λ's, L. The next subsection discusses how to estimate the classical measurement error variance needed for SIMEX.

4.2 *Estimation of spatial measurement error variance parameters*

The simulation step of spatial SIMEX relies on generating random samples of error from the multivariate classical error distribution. Based on the derivations in our bias analysis of Section 3, the only component of error that leads to asymptotic bias is the model misspecification component of classical error. Thus, only the model misspecification variance is needed to generate these pseudo-datasets in the SIMEX procedure to asymptotically correct for the bias. The classical error due to model misspecification has the distribution

$$\begin{aligned}
&\mathcal{U}_{c,\mathsf{model\ mis}} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_c),\\
&\qquad \mathbf{\Sigma}_c = \alpha_2^2 \mathbf{\Sigma}_{S_2 S_2} - \alpha_2 \mathbf{\Sigma}_{S_2 S_2^*} \mathbf{A}^\intercal - \alpha_2 \mathbf{A} \mathbf{\Sigma}_{S_2^* S_2} + \mathbf{B} \mathbf{\Sigma}_{X^* X^*} \mathbf{B}^\intercal + \mathbf{A} \mathbf{\Sigma}_{S_2^* S_2^*} \mathbf{A}^\intercal,\\
&\qquad \mathbf{A} = \mathbf{\Sigma}_{X X^*}(\boldsymbol{\psi}) \mathbf{\Sigma}^{-1}_{X^* X^*}(\boldsymbol{\psi}) - \mathbf{\Sigma}_{X X^*}(\boldsymbol{\psi}_N) \mathbf{\Sigma}^{-1}_{X^* X^*}(\boldsymbol{\psi}_N),\\
&\qquad \mathbf{B} = \mathbf{\Sigma}_{X X^*}(\boldsymbol{\psi}_N) \mathbf{\Sigma}^{-1}_{X^* X^*}(\boldsymbol{\psi}_N),
\end{aligned} \tag{4.13}$$

as showin in the Supplementary Material.

Equation (4.13) shows that $\mathbf{\Sigma}_c$ depends on the spatial covariances of the exposures under both the true parameters and the naive parameters as well as the spatial covariance of the unobserved spatial covariate $\boldsymbol{S}_2$. In practice, these spatial covariances would not be known and thus the true $\mathbf{\Sigma}_c$ would not be known. In that case, $\mathbf{\Sigma}_c$ can be approximated by using external validation data from held-out monitors. First, we fit the spatial exposure model and predict the exposure at the held-out monitor locations. Then, we subtract the predicted exposures form the observed exposures at the held-out monitor locations to obtain the residuals at those locations. We then fit a spatial model to the set of residuals to estimate the total spatial covariance. In practice, we use a Matérn covariance structure for the spatial model of the residuals.

One key issue issue which has been discussed in previous studies is the nonidentifiability of the proportion of measurement error that is Berkson versus classical in models with both Berkson and classical measurement error (Mallick *and others*, 2002; Li *and others*, 2007). While external validation data allows the estimation of the total error variance, the relative proportions of Berkson and classical errors cannot be determined. These previous studies consider the case where both the Berkson errors and the Classical errors are assumed to be i.i.d. Normal in one dimension. To deal with the identifiability issue in a practical application, the authors perform sensitivity analyses regarding the percentage of variance assumed to be classical versus Berkson. We take a similar approach, estimating the total spatial covariance $\widehat{\boldsymbol{\Sigma}}_{total}$ by using external validation data. We then use compute $\widehat{\boldsymbol{\Sigma}}_c = p\widehat{\boldsymbol{\Sigma}}_{total}$, where p represents the proportion the total spatial error attributable to Classical error, with the remaining error as Berkson.

5. Simulation Study

We conduct a simulation study to explore the degree of bias for small sample estimation error when the exposure model is correctly specified. We consider a small number of monitors, $m = 20$ and $m = 40$. In scenario IA, we assume a constant mean model, so all of the exposure variability comes from the spatially correlated residuals. First, we consider a smooth surface that satisfies the smoothness conditions needed in our Taylor expansion. For our Matérn covariance function, we chose covariance parameters $\phi = 0.2$ for the range and $\sigma^2 = 0.5$ for the variance, and we consider a smooth surface with a smoothness of $\nu = 3$, and a rough surface with a smoothness of $\nu = 1$, with examples shown in Figure S2 in the Supplement. Note that the a rough surface does not satisfy the smoothness conditions needed in our Taylor expansion because it only has first derivatives. Scenario IB assumes that the mean depends on two spatial covariates, and for the residuals we use the same parameter values for the smooth and rough surface as in Scenario I.

The results for the smooth and rough exposure surfaces for Scenario I are given in Table 1. For the smooth surface, even for small numbers of monitors we observe negligible bias. The rough surface is a particularly difficult case to estimate because we use a small number of monitors to estimate the form of a rough surface. We see a small amount of bias in the case of Scenario IA with only 20 monitors. Interestingly, the direction of this bias here is upward, which may be an artifact of using a particularly sparse dataset and a rough surface. As expected, the Berkson error underestimates the standard error, leading to insufficient coverage of the confidence intervals.

We conduct a simulation study to explore the degree of bias in the model misspecification scenario and evaluate performance of the spatial SIMEX correction method. To generate the exposure model, we use the same setup as Scenario IB with two spatial covariates and smooth residuals. To fit the exposure model under model misspecification, we omit the second spatial covariate. We also assume first that $\mathbf{\Sigma}_c$ is known, and in later simulations we relax this assumption.

The results for Scenario II are given in Table 2. We observe substantial bias toward the null. Results show that the bias in the health effect parameter is approximately corrected by spatial SIMEX. We see a difference in performance between the two extrapolation functions used for spatial SIMEX, where the linear extrapolation function under corrects the bias. In the simulation step of the spatial SIMEX procedure, we use the derived covariance matrix with true parameter values to sample random error to generate txhe pseudo-datasets. We implement the bootstrap standard error as described in Section 4.3 using $K = 200$ bootstrap resampling steps.

We conducted a simulation study to evaluate performance of the SIMEX correction method when $\mathbf{\Sigma}_c$ is estimated using the available monitoring data. In Section 4 we describe our approach to estimate $\mathbf{\Sigma}_c$ using held-out monitors. This simulation assumes held-out monitors are not available, so we remove one third of the available monitors to use as a held-out dataset. The estimate of $\mathbf{\Sigma}_c$ also depends on the assumed proportion of Classical to Berkson error. We know that most of the error is Classical, but there is also some correlated Berkson error. We choose proportions of 80% Classical error and 90% Classical error as realistic approximations. To examine how robust the spatial SIMEX estimator is to the choice of p, we also use more extreme values of 100% Classical error and 50% Classical error. The results of this simulation are given in Table 3.

We find that the spatial SIMEX procedure still works well even when approximating $\mathbf{\Sigma}_c$. The best performance is seen when p is 0.80. In the extreme case assuming 100% Classical error, the spatial SIMEX procedure using the quadratic extrapolation appears to over-correct the bias. In the other extreme case assuming 50% Classical error, both the linear and quadratic extrapolations under-correct the bias. Although there is sensitivity to the choice of p, all the spatial SIMEX estimates noticeably reduce the bias. There is slight under-coverage in the 95% CI estimates produced by spatial SIMEX across all choices of p.

In practice, the true exposure surface may exhibit non-Gaussian distributions. To explore the performance of our spatial SIMEX method when our Gaussian assumptions are not satisfied, we considered simulation settings where the spatial covariates were generated from spatial log-normal distributions, with results given in Supplementary Material, Table S1. Overall, results are as expected, where spatial SIMEX corrects adequately for bias or may over-correct, yielding effect estimates that could have slight upward bias. The standard errors in this scenario are over-estimated, leading to CIs that are wider than necessary and coverage greater than 99%.

6. Data Example: Association between air pollution and low birthweight

We applied our spatial SIMEX method to a study of birthweight and particulate matter exposure during pregnancy in Massachusetts. The objective of the study was to estimate the association between birthweight and $PM_{2.5}$ exposure during the second and third trimesters. The study population included all singleton live births in Massachusetts from the Massachusetts Birth Registry during 2008 (January 1 to December 31), a total of 70,340 births. Individual-level data on the mother and baby come from the Massachusetts Birth Registry. Confounders in the health model include maternal age, gestational age, number of cigarettes smoked during and before pregnancy, chronic conditions of mother or conditions of pregnancy (lung disease, hypertension, gestational diabetes and non-gestational diabetes), and socioeconomic measures (mother's race, mother's years of education, and the Kotelchuck index of adequacy of prenatal care utilization). Area-level socioeconomic status is controlled by census-tract median household income using data from the United States Census Bureau of 2000 for each census tract in Massachusetts. These covariates are consistent with the published literature on birthweight and particulate matter (Dadvand *and others*, 2013). Some studies also adjust for co-pollutant exposures, such as ozone, although the need for this may vary by region, where studies in the northeast have found similar effect sizes after this adjustment (Bell *and others*, 2007).

$PM_{2.5}$ measurements during 2007 and 2008 were obtained from 40 monitoring sites in Massachusetts as part of the EPA (Environmental Protection Agency) and IMPROVE (Interagency Monitoring of Protected Visual Environments) monitoring networks (Kloog *and others*, 2011). The residential address of each mother at time of birth was geocoded as described in Kloog *and others* (2012). To predict $PM_{2.5}$ at the home address of the mother for each birth, a universal kriging model is assumed, with a Matérn covariance structure for the residuals. The mean function for the kriging model includes a linear trend for three land use covariates: distance to primary highway, distance to known particulate matter emission source, and average traffic density, as described in Kloog *and others* (2011). Separate models are fit for each month using the monthly average $PM_{2.5}$ concentrations at the monitoring sites during 2007 and 2008. Exposures during the second and third trimesters of pregnancy are estimated by averaging the monthly $PM_{2.5}$ concentrations prior to the delivery date.

We fit linear health effect models for each exposure of interest, second trimester $PM_{2.5}$ and third trimester $PM_{2.5}$, adjusting for confounders. This model yields a naive effect estimate that is not corrected for measurement error. We also apply our proposed spatial SIMEX correction method using a quadratic extrapolation function and assuming that 80% of the correlated spatial error is classical. We use 50 SIMEX simulation steps to correct the bias and 50 bootstrap resampling steps to estimate the standard error. Further description of our implementation of spatial SIMEX for this application is given in the Supplementary Material.

Without correcting for measurement error, we find negative associations between birthweight and each $PM_{2.5}$ exposure, and the estimated effect size is larger when we apply spatial SIMEX. Specifically, the change in birthweight per 1 $\mu g/m^3$ second trimester $PM_{2.5}$ exposure is estimated to be -5.04 grams, 95% CI $(-8.02, -2.05)$, without accounting for measurement error. When corrected by spatial SIMEX, this association is estimated to be -7.90 grams per 1 $\mu g/m^3$ $PM_{2.5}$ exposure in the second trimester, 95% CI $(-8.20, -7.61)$. For the third trimester, the change in birthweight per 1 $\mu g/m^3$ $PM_{2.5}$ is estimated to be -3.49 grams, 95% CI $(-6.08, -0.89)$, without any measurement error correction. Applying spatial SIMEX, we estimate an association of -4.91 grams per 1 $\mu g/m^3$ third trimester $\mathrm{PM}_{2.5}$ exposure, 95% CI $(-5.17, -4.66)$. In the Supplementary Material, Figure S4 and Table S2, we report results from sensitivity analyses varying the assumed percentage of classical error and using both linear and quadratic extrapolation functions.

7. Discussion and Conclusions

In this paper, we have conducted a bias analysis of several key scenarios in exposure modeling of air pollution. We have shown that when the exposure model is misspecified by omitting an important covariate, notable downward bias in the health effect estimate can occur in addition to the underestimation of the standard errors. We have proposed a new spatial SIMEX approach to adjust for bias and standard error estimation in the presence of model misspecification. We demonstrated that this bias due to exposure model misspecification can be approximately corrected by this spatial SIMEX procedure. We have also shown analytically and via simulation the presence of small-sample bias due to estimation error in the case of a correctly specified exposure model, although the degree of bias in practice is typically negligible. Hence, this work has demonstrated that with respect to bias, model misspecification is a much bigger problem than parameter estimation.

Previous research in this area has suggested that using the plug-in estimator typically induces little bias, and authors have advocated for using the plug-in estimator to estimate the effect size and then adjusting the standard errors to account for the additional variability in using the exposure predictions (Szpiro *and others*, 2011; Madsen *and others*, 2008; Lopiano *and others*, 2011; Gryparis *and others*, 2009). However, those papers investigate bias in simulation studies primarily by fitting the correct exposure model used to generate the data. Our findings in Section 3 for the bias of exposure model scenario I are consistent with these previous studies, as we also found in simulations that the degree of bias is small when the correct exposure model is specified. Our findings are also consistent with a recent study investigating exposure model misspecification via simulation that illustrated bias induced in health effect estimates when universal kriging exposure models had poor fit (Alexeeff *and others*, 2014).

In practice, the underlying exposure model that generates air pollution levels in any given region is not exactly known. In addition, current approaches for correcting the standard errors of estimates also rely on the assumption that the exposure model is correctly specified (Szpiro *and others*, 2011; Madsen *and others*, 2008). We have approached this problem using both analytical methods and simulation studies, and we have presented a more thorough bias analysis than what has been considered in previous work. In particular, we have extended existing work to the case of model misspecification, which is important since exposure models are typically complex and no single statistical model is likely to be correct.

This work also points to a few practical considerations which are important in order to help with the implementation of this spatial SIMEX method in a realistic setting. As in other SIMEX procedures, the classical error variance is needed to generate the simulated re-measurements for the bias correction, and that variance is assumed to be known. In Section 3, we derived the particular form of the classical error variance for scenario II, but in practice the exact classical error variance would not be known and finding a way to estimate that variance may be difficult. The other key practical consideration in the estimation of the classical error variance is the nonidentifiability issue created by the mixture of Berkson and classical errors. Typically in measurement error problems, external validation data with measurements of both the true exposure and the mismeasured exposure are used to estimate the measurement error variance. However, previous studies looking at mixtures of independent and identically distributed Berkson and classical errors have noted that the amount of uncertainty that is Berkson versus classical is not identifiable (Mallick *and others*, 2002; Li *and others*, 2007). While external validation data would allow the estimation of the total error variance, the relative proportions of Berkson and classical errors cannot be determined. In Li *and others* (2007), the authors perform sensitivity analyses by considering a range of values for the percentage of variance assumed to be classical versus Berkson to deal with the identifiability issue.

Our work points several areas of future research interest. First, model misspecification in land use regression and kriging models for air pollution exposure should be examined further. We have considered one scenario of model misspecification, and we observed notable bias induced in that case. Many other scenarios of model misspecification may also be of interest to study, for example the case where the omitted covariate is actually a confounder of the other predictors. In addition, it would be helpful to study the practical issues of the implementation, including methods for estimating the classical error variance given the issue of identifiability, as mentioned above, as well as the robustness of spatial SIMEX to incorrect estimation of the classical error variance.

One advantage of SIMEX is that the general methodology can be adapted to cases in which the measurement error biases cannot be derived in closed-form, including logistic regression and Poisson regression (Carroll *and others*, 2006). SIMEX can also be extended to the multi-pollutant setting in which more than one exposure covariate is measured with measurement error (Carroll *and others*, 2006). In the multi-pollutant case, we can use the same cross-validation procedure to estimate prediction errors for all pollutants at each location. We can then fit a covariance structure model for these spatially correlated multivariate errors, such as the kronecker product of a multi-pollutant covariance matrix for prediction errors measured at the location and parametric (e.g. exponential) spatial correlation structures for predictions errors for a given pollutant measured at different locations. Therefore, our proposed spatial SIMEX approach would be applicable to the multi-pollutant setting as long as all the pollutants are jointly measured. In addition, SIMEX can be adapted when the measurement error itself follows a different form, for example multiplicative log-Gaussian errors (Eckert *and others*, 1997).

This work examines aspects of exposure modeling of air pollution for health effect studies and provides some insight into the role of estimation error and model misspecification in the estimation of health effects. Understanding these impacts when constructing land use regression and kriging models is of fundamental importance to studies of air pollution and health. In particular, these results should be taken into account when interpreting the results of air pollution epidemiology studies that use land use regression and kriging models for exposure estimation. The spatial SIMEX procedure provides one possible measurement error correction strategy which may be beneficial to correct for bias induced by model misspecification.

Acknowledgements

This work was supported by USEPA grant 834798 and NIH grants ES020871 and ES000002. This publication's contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US EPA. Further, US EPA does not endorse the purchase of any commercial products or services mentioned in the publication. Carroll's research was supported by a grant from the National Cancer Institute (R37-CA057030).

References

Alexeeff, Stacey E., Schwartz, Joel, Kloog, Itai, Chudnovsky, Alexandra, Koutrakis, Petros and Coull, Brent A. (2014). Consequences of kriging and land use regression for pm2.5 predictions in epidemiologic analyses: insights into spatial variability using high-resolution satellite data. *Journal of Exposure Science and Environmental Epidemiology*, 1–7.

Bell, M. L., Ebisu, K. and Belanger, K. (2007). Ambient air pollution and low birth weight in connecticut and massachusetts. *Environmental Health Perspectives* **115**, 1118–1124.

Bergen, S., Sheppard, L., Sampson, P.D., Kim, S.Y., Richards, M., Vedal, S., Kaufman, J.D. and Szpiro, A.A. (2013). A national prediction model for pm2.5 component exposures and measurement error-corrected health effect inference. *Environmental Health Perspectives* **121**, 1017–1025.

Brauer, Michael, Hoek, Gerard, van Vliet, Patricia, Meliefste, Kees, Fischer, Paul, Gehring, Ulrike, Heinrich, Joachim, Cyrys, Josef, Bellander, Tom, Lewne, Marie *and others*. (2003). Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology* **14**(2), 228–239.

Brook, Robert D., Rajagopalan, Sanjay, III, C. Arden Pope, Brook, Jeffrey R., Bhatnagar, Aruni, Diez-Roux, Ana V., Holguin, Fernando, Hong, Yuling, Luepker, Russell V., Mittleman, Murray A., Peters, Annette, Siscovick, David, Jr, Sidney C. Smith, Whitsel, Laurie *and others*. (2010). A.h.a. scientific statement. particulate matter air pollution and cardiovascular disease. *Circulation* **121**, 2331–2378.

Carroll, R. J., Ruppert, D., Stefanski, L. A. and Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*, 2nd edition. New York, New York: Chapman & Hall.

Clougherty, J.E., Wright, R.J., Baxter, L.K. and Levy, J.I. (2008). Land use regression modeling of intra-urban residential variability in multiple traffic-related air pollutants. *Environ Health* **7**(17).

Cook, J. R. and Stefanski, L. A. (1994). Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical Association* **89**, 1314–1328.

Cressie, Noel. (1993). *Statistics for Spatial Data*, Revised edition. Wiley Series in Probability and Statistics.

Dadvand, P, Parker, J, Bell, M.L., Bonzini, M, Brauer, M., Darrow, L.A., Gehring, U., Glinianaia, S.V., Gouveia, N., Ha, E., Leem, J.H., van den Hooven, E.H., Jalaludin, B., Jesdale, B.M., Lepeule, J., Morello-Frosch, R., Morgan, G.G., Pesatori, A.C., Pierik, F.H., Pless-Mulloli, T., Rich, D.Q., Sathyanarayana, S., Seo, J., Slama, R., Strickland, M., Tamburic, L., Wartenberg, D., Nieuwenhuijsen, M.J. *and others*. (2013). Maternal exposure to particulate air pollution and term birth weight: A multi-country evaluation of effect and heterogeneity. *Environmental Health Perspectives* **121**, 367–373.

Eckert, R.S., Carroll, R.J. and Wang, N. (1997). Transformations to additivity in measurement error models. *Biometrics* **53**, 262–272.

Eeftens, Marloes, Beelen, Rob, de Hoogh, Kees, Bellander, Tom, Cesaroni, Giulia, Cirach, Marta, Declercq, Christophe, Ddel, Audrius, Dons, Evi, de Nazelle, Audrey, Dimakopoulou, Konstantina, Eriksen, Kirsten, Falq, Grgoire, Fischer, Paul, Galassi, Claudia, Grauleviien, Regina, Heinrich, Joachim, Hoffmann, Barbara, Jerrett, Michael, Keidel, Dirk, Korek, Michal, Lanki, Timo, Lindley, Sarah, Madsen, Christian, Mlter, Anna, Ndor, Gizella, Nieuwenhuijsen, Mark, Nonnemacher, Michael, Pedeli, Xanthi, Raaschou-Nielsen, Ole, Patelarou, Evridiki, Quass, Ulrich, Ranzi, Andrea, Schindler, Christian, Stempfelet, Morgane, Stephanou, Euripides, Sugiri, Dorothea, Tsai, Ming-Yi, Yli-Tuomi, Tarja, Varr, Mihly J, Vienneau, Danielle, Klot, Stephanie von, Wolf, Kathrin, Brunekreef, Bert *and others*. (2012). Development of land use regression models for pm2.5, pm2.5 absorbance, pm10 and pmcoarse in 20 european study areas; results of the escape project. *Environmental Science and Technology* **46**(20), 11195–11205. PMID: 22963366.

Gryparis, A., Paciorek, C. J., Zeka, A., Schwartz, J. and Coull, B. A. (2009). Measurement error caused by spatial misalignment in environmental epidemiology. *Biostatistics* **10**, 258–274.

Kloog, I., Chudnovsky, A.A., Just, A.C., Nordio, Francesco, Koutrakis, P., Coull, B.A., Lyapustin, A., Wang, Y.J. and Schwartz, J. (2014). A new hybrid spatio-temporal model for estimating daily multi-year pm2.5 concentrations across northeastern usa using high resolution aerosol optical depth data. *Atmospheric Environment* **95**, 581–590.

Kloog, I., Koutrakis, P., Coull, B.A., Lee, H.J. and Schwartz, J. (2011). Temporally and spatially resolved pm2. 5 exposures for epidemiological studies using satellite aerosol optical depth measurements. *Atmospheric Environment* **45**, 6267–6275.

Kloog, Itai, Melly, Steven J, Ridgway, William L, Coull, Brent A and Schwartz, Joel. (2012*a*). Using new satellite based exposure methods to study the association between pregnancy pm2.5 exposure, premature birth and birth weight in massachusetts. *Environmental Health* **11**, 40.

Kloog, I., Nordio, Francesco, Coull, B.A. and Schwartz, J. (2012*b*). Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal pm2.5 exposures in the mid-atlantic states. *Environmental Science and Technology* **46**, 11913–11921.

Li, Y., Guolo, A., Hoffman, O. and Carroll, R.J. (2007). Shared uncertainty in measurement error problems, with application to nevada test site fallout data. *Biometrics* **63**, 1226–1236.

Liao, D.P., Peuquet, D.J., Duan, Y.K., Whitsel, E.A., Dou, J.W., Smith, R.L., Lin, H.M., Chen, J.C. and Heiss, G. (2006). Gis approaches for the estimation of residential-level ambient pm concentrations. *Environmental Health Perspectives* **114**, 1374–1380.

Lopiano, K.K., Young, L.J. and Gotway, C.A. (2011). A comparison of errors in variables methods for use in regression models with spatially misaligned data. *Statistical Methods in Medical Research* **20**, 29–47.

Lopiano, K.K., Young, L.J. and Gotway, C.A. (2013). Estimated generalized least squares in spatially misaligned regression models with berkson error. *Biostatistics* **14**, 737–751.

Madsen, L., Ruppert, D. and Altman, N.S. (2008). Regression with spatially misaligned data. *Environmetrics* **19**, 453–467.

Mallick, B., Hoffman, O. and Carroll, R.J. (2002). Semiparametric regression modeling with mixtures of berkson and classical error, with application to fallout from the nevada test site. *Biometrics* **58**, 13–20.

Nordio, Francesco, Kloog, I., Coull, B.A. and Schwartz, J. (2013). Estimating spatio-temporal resolved pm10 aerosol mass concentrations using modis satellite data and land use regression over lombardy, italy. *Environmental Science and Technology* **74**, 227–236.

Peng, R.D. and Bell, M.L. (2010). Spatial misalignment in time series studies of air pollution and health data. *Biostatistics* **11**(4), 720–740.

Ross, Zev, Jerrett, Michael, Ito, Kazuhiko, Tempalski, Barbara and Thurston, George D. (2007). A land use regression for predicting fine particulate matter concentrations in the new york city region. *Atmospheric Environment* **41**(11), 2255 – 2269.

Szpiro, A.A., Sheppard, L. and Lumley, T. (2011). Efficient measurement error correction with spatially misaligned data. *Biostatistics* **12**, 610–623.

Szpiro, Adam A. and Paciorek, Christopher J. (2013). Measurement error in two-stage analyses, with application to air pollution epidemiology. *Environmetrics* **24**(8), 501–517.

Table 1. Simulation results for smooth and rough exposure surfaces for Scenario I with different number of monitors m

Scenario	m	Exposure	Bias	empir SE	model SE	MSE	Coverage
IA Smooth	20	True X	0.001	0.082	0.076	0.007	94.200
IA Smooth	20	$g(\boldsymbol{\theta};\mathbf{X}^*)$	-0.003	0.108	0.078	0.012	86.400
IA Smooth	20	$g(\hat{\boldsymbol{\theta}};\mathbf{X}^*)$	0.004	0.111	0.080	0.012	85.200
IA Smooth	40	True X	-0.000	0.086	0.076	0.007	94.000
IA Smooth	40	$g(\boldsymbol{\theta};\mathbf{X}^*)$	0.001	0.088	0.076	0.008	93.000
IA Smooth	40	$g(\hat{\boldsymbol{\theta}};\mathbf{X}^*)$	0.002	0.088	0.076	0.008	93.000
IB Smooth	20	True X	0.001	0.044	0.043	0.002	94.990
IB Smooth	20	$g(\boldsymbol{\theta};\mathbf{X}^*)$	0.002	0.055	0.044	0.003	89.379
IB Smooth	20	$g(\hat{\boldsymbol{\theta}};\mathbf{X}^*)$	0.002	0.059	0.044	0.003	87.976
IB Smooth	40	True X	0.002	0.043	0.043	0.002	94.400
IB Smooth	40	$g(\boldsymbol{\theta};\mathbf{X}^*)$	0.003	0.044	0.043	0.002	93.600
IB Smooth	40	$g(\hat{\boldsymbol{\theta}};\mathbf{X}^*)$	0.003	0.045	0.043	0.002	93.800
IA Rough	20	True X	-0.001	0.057	0.055	0.003	93.865
IA Rough	20	$g(\boldsymbol{\theta};\mathbf{X}^*)$	-0.007	0.179	0.080	0.032	63.190
IA Rough	20	$g(\hat{\boldsymbol{\theta}};\mathbf{X}^*)$	0.058	0.220	0.089	0.052	57.055
IA Rough	40	True X	-0.000	0.057	0.054	0.003	94.990
IA Rough	40	$g(\boldsymbol{\theta};\mathbf{X}^*)$	0.002	0.104	0.067	0.011	80.962
IA Rough	40	$g(\hat{\boldsymbol{\theta}};\mathbf{X}^*)$	0.024	0.114	0.070	0.014	78.557
IB Rough	20	True X	0.002	0.038	0.038	0.001	94.400
IB Rough	20	$g(\boldsymbol{\theta};\mathbf{X}^*)$	0.002	0.040	0.038	0.002	93.600
IB Rough	20	$g(\hat{\boldsymbol{\theta}};\mathbf{X}^*)$	0.002	0.042	0.038	0.002	92.200
IB Rough	40	True X	0.002	0.038	0.038	0.001	95.000
IB Rough	40	$g(\boldsymbol{\theta};\mathbf{X}^*)$	0.002	0.038	0.038	0.001	95.000
IB Rough	40	$g(\hat{\boldsymbol{\theta}};\mathbf{X}^*)$	0.002	0.038	0.038	0.001	94.600

Wang, N., Lin, X., Gutierrez, R.G. and Carroll, R.J. (1998). Bias analysis and simex approach in generalized linear mixed measurement error models. *Journal of the American Statistical Association* **93**, 249–261.

Zimmerman, D. L. and Cressie, N. (1992). Mean squared prediction error in the spatial linear model with estimated covariance parameters. *Annals of the Institute of Statistical Mathematics* **44**, 27–43.

Table 2. Simulation results for Scenario II, misspecified exposure model, and correction by spatial SIMEX when spatial measurement error variance is known.

Scenario	m	Exposure	Bias	empirical SE	model SE	MSE	Coverage
II	50	True X	0.000	0.036	0.034	0.001	93.6
II	50	$g(\boldsymbol{\theta}; \mathbf{X}^*)$	0.000	0.036	0.034	0.001	93.6
II	50	$g(\hat{\boldsymbol{\theta}}_N; \mathbf{X}^*)$	-0.203	0.182	0.039	0.074	22.2
II	50	spatial SIMEX, linear	-0.072	0.211	0.241	0.050	90.6
II	50	spatial SIMEX, quad	0.026	0.254	0.285	0.065	91.9

Table 3. Simulation results for Scenario II, misspecified exposure model, and correction by spatial SIMEX when spatial measurement error variance parameters are estimated, for different proportions, p, of Classical error

Scenario	p	Exposure	Bias	empirical SE	model SE	MSE	Coverage
II		True X	0.001	0.036	0.034	0.001	93.6
II		$g(\boldsymbol{\theta}; \mathbf{X}^*)$	0.001	0.036	0.034	0.001	93.6
II		$g(\hat{\boldsymbol{\theta}}_N; \mathbf{X}^*)$	-0.200	0.180	0.039	0.072	22.4
II	1.00	spatial SIMEX, linear	-0.068	0.213	0.228	0.050	91.3
II	1.00	spatial SIMEX, quadratic	0.067	0.322	0.336	0.108	87.2
II	0.90	spatial SIMEX, linear	-0.075	0.211	0.228	0.050	91.1
II	0.90	spatial SIMEX, quadratic	0.044	0.308	0.331	0.097	87.8
II	0.80	spatial SIMEX, linear	-0.076	0.208	0.227	0.049	91.6
II	0.80	spatial SIMEX, quadratic	0.028	0.294	0.324	0.087	89.1
II	0.50	spatial SIMEX, linear	-0.112	0.200	0.225	0.053	89.6
II	0.50	spatial SIMEX, quadratic	-0.058	0.247	0.304	0.064	91.3