6-19-2008

Accounting for Errors from Predicting Exposures in Environmental Epidemiology and Environmental Statistics

Adam A. Szpiro
University of Washington, aszpiro@u.washington.edu

Lianne Sheppard

Thomas Lumley

Accounting for Errors from Predicting Exposures in Environmental Epidemiology and Environmental Statistics

Adam A. Szpiro, Lianne Sheppard, and Thomas Lumley

Department of Biostatistics, University of Washington

June 19, 2008

Abstract

In environmental epidemiology and related problems in environmental statistics, it is typically not practical to directly measure the exposure for each subject. Environmental monitoring is employed with a statistical model to assign exposures to individuals. The result is a form of exposure misspecification that can result in complicated errors in the health effect estimates if the exposure is naively treated as known. The exposure error is neither "classical" nor "Berkson", so standard regression calibration methods do not apply. We decompose the health effect estimation error into three components. First, the standard errors are too small if the exposure field is correlated, independent of variability in estimating the exposure field parameters. Second, the standard errors are too small because they do not account for variability in estimating the exposure field parameters. Third, there is a bias from using approximate exposure field parameters in place of the unobserved true ones. We outline a three-stage correction procedure to account separately for each of these errors. A key insight is that we can account for the second part of the error (sampling variability in estimating the exposure) by averaging over simulations from the part of the posterior exposure surface that is informative for the outcome. This amounts to averaging over samples of the posterior exposure model parameters, a procedure that we call "parameter simulation". One implication is that it is preferable to use a parametric correlation model (e.g., kriging) rather than a semi-parametric approximation. While the latter approach has been found to be effective in estimating mean exposure fields, it does not provide the needed decomposition of the posterior into informative and non-informative components. We illustrate the properties of our corrected estimators in a simulation study and present an example from environmental statistics. The focus of this paper is on linear health effect models with uncorrelated outcomes, but extensions to generalized linear models and correlated outcomes are possible.

Key words: air pollution, measurement error, universal kriging, exposure modeling, environmental epidemiology, environmental statistics

1 Introduction

Recently there has been significant interest in assessing the relationship between health outcomes and environmental exposures. Technical or logistical constraints typically preclude directly measuring the exposure for individual subjects. Limited environmental monitoring is employed instead, and a statistical model is constructed to assign exposures based on the available measurements. This methodology exploits the fact that exposures are often predictable based on Geographic Information System (GIS) covariates and, further, that they tend to exhibit spatial or spatio-temporal correlation structures that can be useful for interpolation.

An important example is evaluating the relationship between exposure to ambient air pollution and adverse health outcomes. Many studies have documented adverse effects of air pollution (Dockery et al. 1993; Samet et al. 2000; Pope et al. 2002), but until recently most studies have assigned exposures based on area-wide monitored concentrations. In order to improve exposure assessment, some recent studies have used predicted individual air pollution exposure. Prediction approaches include assigning the nearest monitor value to the participant's residential location (Miller et al. 2007; Basu et al. 2000; Ritz et al. 2006), regression based on GIS covariates (Brauer et al. 2003; Jerrett et al. 2005a), and interpolation by a geostatistical method such as kriging (Jerrett et al. 2005b; Kunzli et al. 2005) or semi-parametric smoothing (Gryparis et al. 2007; Kunzli et al. 2005).

Many environmental applications that do not involve human health effects are analogous from a statistical perspective. An example that we will return to later in this

paper involves assessing the relationship between chloride levels in streams and nearby watershed land cover (Madsen et al. 2007; Herlihy et al. 1998). Another application involves assessing the relationship between changes in annual rainfall and vegetation cover. This has been studied in the African Sahel, where remotely sensed vegetation levels are available everywhere but rainfall levels are only monitored at a limited number of locations. As in the case of air pollution epidemiology, several approaches to predicting precipitation levels at locations without measurements have been proposed (Los et al. 2006; Lindstron and Lindgren 2008).

A common feature is that little is known about the properties of the resulting inference for regression models based on estimated exposures. Kim et al. (2008) have shown that plugging in the mean exposure estimates from kriging tends to perform better than assigning exposures based on the nearest neighbor monitoring location, but the resulting errors are complicated and involve both biased point estimates and incorrect standard errors. Madsen et al. (2007) describes a methodology to correct the standard errors when the exposure is estimated with ordinary kriging. Their "krige and regress" approach has key features in common with ours, but the development has several limitations. First, although Madsen et al. (2007) combines two stages of standard error correction to separately account for spatial correlation in the exposure and variability in its estimation (as we do in Sections 4.1 and 4.2), the second stage of the correction does not seem to be required in their examples. Therefore, it is difficult to assess how well it will work when it is needed. Second, the examples do not suffer from bias and, perhaps for this reason, no attempt is made to address this potential source of error. Our examples show that there can be residual bias. Finally, the methodology is spe-

cific to ordinary kriging. It can probably be extended to universal kriging and other more flexible models, but this would involve significant calculations and approximations beyond what is described in the paper.

Gryparis et al. (2007) reviews the relevant measurement error literature and compares several strategies in a simulation study. They recommend using a fully Bayesian joint model for the exposure and outcome when computationally feasible. For larger data sets, this is generally not practical. It has also been noted by Gryparis et al. (2007), Best et al. (2006), and others that a joint Bayesian approach may result in poor estimation of the exposure surface because of a combination of model misspecification and/or outliers in the outcome variable. Therefore it is preferable to separate estimation of the exposure surface from inference about the outcome. As an alternative to joint Bayesian modeling, Gryparis et al. (2007) suggests applying standard regression calibration techniques that are based on the assumption of classical measurement error in the estimated exposure (Carroll et al. 1995). While this seems to work reasonably well in some of the simulation examples, it is not conceptually consistent with the actual error structure that results from using an estimated exposure surface based on modeling. Also, the implementation considered in Gryparis et al. (2007) requires holding out a sub-sample of monitoring data for calibration.

One of the interesting findings in Gryparis et al. (2007) is that a strategy they call "exposure simulation" performs extremely poorly, resulting in significant bias toward the null. The idea of "exposure simulation" is to plug multiple realizations from the posterior exposure surface into the health model and average the resulting effect estimates. This approach is appealing because it seems to provide a natural way of assessing the

variability in the effect size due to uncertainty in the exposure estimation. However, as Gryparis et al. (2007) show, posterior realizations of the exposure surface contain some variability that is entirely non-informative for the outcome, and adding this to the mean exposure estimate is equivalent to introducing classical measurement error.

An important contribution of our work is to clarify why "exposure simulation" fails and to describe an alternative that avoids the problem of bias toward the null. Essentially the idea is that we only want to simulate from the part of the posterior that is informative for the outcome. In a kriging model, it is clear that this should not include the nugget. Even though we regard the nugget as part of the exposure field (not measurement error), there is no useful information from the nugget at locations without measured exposure data. Thus, it is easy to see that including the nugget in "exposure simulation" is analogous to introducing classical measurement error. It turns out that the variability corresponding to the partial sill is also non-informative for the outcome. This is somewhat less intuitive, but one way of seeing it is to appreciate that conditional on the kriging model parameters, samples from the correlated residuals are independent of the measured exposure data and are therefore independent of the outcome.

We conclude that the correct approach is to average over multiple realizations of the conditional mean of the exposure field, where the conditioning is on the observed exposure data and on the model parameters. This amounts to averaging over multiple realizations from the posterior distribution for the exposure model parameters and plugging in the corresponding conditional mean exposure fields. We call this procedure "parameter simulation". There is no systematic bias toward the null of the type that arises in "exposure simulation" because we avoid introducing random fluctuations

unrelated to the outcome.

In addition to variability from estimation of the exposure surface, there are two additional sources of error. First, even if we know the exact values of the model parameters, naively plugging the mean exposure field into a health effect model gives standard errors that are too small if the exposure surface is correlated. We correct for this with a version of the "sandwich" estimator. Second, using estimated values of the exposure model parameters results in a bias of the effect estimate. This bias vanishes asymptotically if there is much exposure monitoring data (in contrast to the bias toward the null that results from full "exposure simulation"), but it can be important in realistic situations where the monitoring data is limited. We correct the bias by a form of regression calibration that is appropriate for the error structure.

In the remainder of this paper we describe in detail our strategy for using modeled exposures in a linear health effect model. We begin in Section 2 by introducing notation and setting out the problem. In Section 3 we analyze the relatively simple situation of an uncorrelated linear regression model for the exposure. Without some of the subtleties of a more general model, this case provides a clear motivation for "parameter simulation". In Section 4 we consider a more general exposure model that includes universal kriging and various spatio-temporal correlation structures as special cases. We describe the three sources of error that arise from ignoring estimation of the exposure variable, and we give a three-stage procedure for correcting these errors. In Section 5 we use simulation examples to illustrate the frequentist properties of our corrected effect estimates. We consider an example with publicly available data from environmental statistics in Section 6, and we conclude in Section 7 with a discussion.

2 Notation and Problem Setup

Consider an epidemiology study with N subjects and the corresponding $N \times 1$ vector of outcomes Y, $N \times 1$ vector of exposures X, and $N \times n$ matrix of covariates Z (potentially including an intercept). We assume

$$Y = X\beta_1 + Z\gamma + \varepsilon, \tag{1}$$

with unknown regression coefficient β_1 for the exposure and $n \times 1$ vector of regression coefficients γ for the covariates. Assume that ε is an $N \times 1$ random vector distributed as $N(0, \sigma_\varepsilon^2 I_N)$, where σ_ε^2 is unknown and I_N is the identity matrix.

The goal is inference for β_1, and this is routine if X is known without error. The ordinary least squares (OLS) point estimate is

$$\begin{pmatrix} \beta_1(X) \\ \gamma(X) \end{pmatrix} = \left[\begin{pmatrix} X \\ Z \end{pmatrix}^t \begin{pmatrix} X \\ Z \end{pmatrix} \right]^{-1} \begin{pmatrix} X \\ Z \end{pmatrix}^t Y, \tag{2}$$

and the corresponding classical standard error is

$$\begin{pmatrix} \sigma_{\beta_1,\text{class}}(X, \sigma_\varepsilon^2) \\ \sigma_{\gamma,\text{class}}(X, \sigma_\varepsilon^2) \end{pmatrix} = \sqrt{\sigma_\varepsilon^2 \text{diag}\left(\left[\begin{pmatrix} X \\ Z \end{pmatrix}^t \begin{pmatrix} X \\ Z \end{pmatrix} \right]^{-1} \right)}, \tag{3}$$

using for σ_ε^2 the method-of-moments estimate from the residuals.

We are interested in the situation where Y and Z are directly observed, but X needs to be estimated using a model. Assume that the related $N^* \times 1$ vector X^* is observed. Typically X^* is comprised of different samples from the same field as X, and we refer to N^* as the number of exposure monitors. Assume that X and X^* are jointly distributed

as

$$\begin{pmatrix} X \\ X^* \end{pmatrix} = \begin{pmatrix} S \\ S^* \end{pmatrix} \alpha + \begin{pmatrix} \eta \\ \eta^* \end{pmatrix}. \tag{4}$$

In this expression, S and S^* are known $N \times m$ and $N^* \times m$ dimensional matrices of covariates, α is an unknown $m \times 1$ vector of coefficients, and

$$\begin{pmatrix} \eta \\ \eta^* \end{pmatrix} \sim N\Big(0, \Sigma_{\eta,\eta^*}(\theta)\Big) \tag{5}$$

for a known positive definite matrix function $\Sigma_{\eta,\eta^*}(\cdot)$ and unknown parameter θ. Universal kriging is a special case if θ comprises the range, partial sill, and nugget parameters from a geostatistical model (Cressie 1993). More general spatio-temporal correlation structures also fit naturally in this framework (Banerjee et al. 2004). The remainder of this paper is concerned with the question of how to use an estimate of X based on the observed values of X^* to derive valid inference for β_1.

3 Linear Regression Exposure Model

We begin with a simple example that illustrates the key ideas we will develop in the following section for a more general model. Let the covariance function from equation (5) be

$$\Sigma_{\eta,\eta^*}(\sigma^2_\eta) = \sigma^2_\eta I_{(N+N^*)}$$

so that the joint vector of measured and unmeasured exposures is independent, conditional on the covariates S and S^*. A natural strategy is to estimate α from X^* and S^*, use this to derive an estimate $\hat{X}$ based on S, and then to make inference about β_1 by plugging $\hat{X}$ into equations (2) and (3).

The best possible situation is if N^* is arbitrarily large, in which case we can assume that α is known exactly. Then letting

$$\hat{X}_{\text{exact}} = S\alpha$$

results in pure Berkson error that can be ignored in estimating β_1 (Zeger et al. 2000). That is, plugging $\hat{X}_{\text{exact}}$ into the classical estimators in equations (2) and (3) gives correct frequentist inference for

$$\begin{aligned} \hat{\beta}_{1,\text{exact}} &= \beta_1(\hat{X}_{\text{exact}}) \\ \hat{\sigma}_{\beta_1,\text{exact}} &= \sigma_{\beta_1,\text{class}}(\hat{X}_{\text{exact}}, \hat{\sigma}^2_{\varepsilon,\text{exact}}), \end{aligned}$$

where $\hat{\sigma}^2_{\varepsilon,\text{exact}}$ is the method-of-moments estimate of σ^2_ε. The only difference from the case where we know X without error is that the sampling variability is greater, but this is correctly accounted for by $\hat{\sigma}_{\beta_1,\text{exact}}$. One way to see this is by noticing the equivalence to estimating β_1 in

$$\begin{aligned} Y &= (S\alpha)\beta_1 + Z\gamma + (\varepsilon + \eta\beta_1) \\ &= (S\alpha)\beta_1 + Z\gamma + \varepsilon', \end{aligned} \tag{6}$$

with independent errors

$$\varepsilon' \sim N(0, \sigma^2_{\varepsilon'} I_N),$$

where $\sigma^2_{\varepsilon'} = \sigma^2_\varepsilon + \sigma^2_\eta\beta_1$. It doesn't matter that β_1 appears in the mean and the variance since the covariance matrix is diagonal, and we can estimate $\sigma^2_{\varepsilon'}$ without regard to the contribution from β_1.

In reality, we typically have few exposure monitoring locations and need to estimate

α. It is convenient to use a Bayesian model. With the standard reference prior

$$p(\alpha, \sigma_\eta^2) \propto \sigma_\eta^{-2},$$

it is known (Box and Tiao 1992) that the posterior parameter $\tilde{\alpha}$ has a multivariate t-distribution with mean

$$\hat{\alpha} = (S^{*t} S^*)^{-1} S^{*t} X^*.$$

The naive "plug-in" approach is to let

$$\hat{X}_{\text{plug-in}} = S\hat{\alpha}$$

and plug this estimate into the classical estimators in equations (2) and (3) to obtain

$$\hat{\beta}_{1,\text{plug-in}} = \beta_1(\hat{X}_{\text{plug-in}})$$

and its associated standard error estimate $\hat{\sigma}_{\beta_1,\text{plug-in}}$. However, the standard error is evidently too small because it fails to account for sampling variability of $\hat{\alpha}$. It also turns out that $\hat{\beta}_{1,\text{plug-in}}$ is biased for estimating β_1. Note that both of these problems go away in the asymptotic limit for large N^*, but they can be important in applied settings where there is limited exposure monitoring data.

We can account for variability in the posterior mean $\hat{\alpha}$ by "parameter simulation", which entails averaging over simulated samples from the posterior random variable $\tilde{\alpha}$. We set $\tilde{X} = S\tilde{\alpha}$ and use this in the classical estimators from equations (2) and (3) to obtain

$$\begin{aligned} \tilde{\beta}_1 &= \beta_1(\tilde{X}) \\ &= \beta_1(S\tilde{\alpha}) \end{aligned}$$

and its associated standard error estimate $\tilde{\sigma}_{\beta_1}$. It is then natural to define

$$\hat{\beta}_{1,\text{sim}} = \tilde{E}(\tilde{\beta}_1)$$

and

$$\hat{\sigma}_{\beta_1,\text{sim}} = \sqrt{\tilde{\text{E}}(\tilde{\sigma}^2_{\beta_1}) + \tilde{\text{Var}}(\tilde{\beta}_1)}$$

where $\tilde{\text{E}}$ and $\tilde{\text{Var}}$ refer to the posterior expectation and variance, respectively.

There is an important distinction between "parameter simulation" and the "exposure simulation" approach described in Gryparis et al. (2006). Our $\tilde{X}$ does not sample from the full posterior for X. As noted in Gryparis et al. (2006), such sampling would involve adding random noise corresponding to η, and this noise would behave like classical measurement error. This would bias the estimate of β_1 toward the null in a way that does not vanish even in the asymptotic limit with a large number of exposure monitors. Instead, we only sample from the part of the posterior for X that is informed by the data through the posterior estimate of the parameter α.

There is still the potential for a bias in $\hat{\beta}_{1,\text{sim}}$ which we can correct by a form of regression calibration. Since $\hat{\beta}_{1,\text{exact}}$ is an unbiased estimator for β_1, the residual bias in $\hat{\beta}_{1,\text{sim}}$ is

$$\begin{aligned}
\pi &= \text{E}(\hat{\beta}_{1,\text{sim}} - \hat{\beta}_{1,\text{exact}}) \\
&= \text{E}(\hat{\beta}_{1,\text{sim}} - \hat{\beta}_{1,\text{plug-in}}) + \text{E}(\hat{\beta}_{1,\text{plug-in}} - \hat{\beta}_{1,\text{exact}}) \\
&= \text{E}\Big[\tilde{E}(\beta_1(S\tilde{\alpha})) - \beta_1(S\hat{\alpha})\Big] + \Big[\text{E}(\beta_1(S\hat{\alpha})) - \beta_1(S\alpha)\Big] \qquad (7) \\
&= \text{E}\Big[\tilde{E}(\beta_1(S\tilde{\alpha})) - \beta_1(\tilde{E}(S\tilde{\alpha}))\Big] + \Big[\text{E}(\beta_1(S\hat{\alpha})) - \beta_1(E(S\hat{\alpha}))\Big],
\end{aligned}$$

where the expectation denoted by E is over the frequentist sampling distribution. Notice

that in fourth line above we use that $S\hat{\alpha} = \tilde{E}(S\tilde{\alpha})$ is an unbiased estimator for $S\alpha$.

The Bernstein-von Mises theorem (Le Cam and Yang 2000) implies that for almost every realization of X^*, the posterior distribution of $\tilde{\alpha}$ is approximately equal to the sampling distribution of $\hat{\alpha}$. Based on this approximation, we define an estimate of the bias by

$$\begin{aligned} \hat{\pi} &= 2\Big[\tilde{E}(\beta_1(S\tilde{\alpha})) - \beta_1(\tilde{E}(S\tilde{\alpha}))\Big] \\ &= 2(\hat{\beta}_{1,\text{sim}} - \hat{\beta}_{1,\text{plug-in}}) \end{aligned}$$

and then define the calibrated estimator

$$\begin{aligned} \hat{\beta}_{1,\text{sim-cal}} &= \hat{\beta}_{1,\text{sim}} - \hat{\pi} \\ &= 2\hat{\beta}_{1,\text{plug-in}} - \hat{\beta}_{1,\text{sim}}. \end{aligned}$$

The calibration procedure can change the variability, but we neglect this and set

$$\hat{\sigma}_{\beta_1,\text{sim-cal}} = \hat{\sigma}_{\beta_1,\text{sim}}.$$

It appears from our simulation examples in Section 5 that the correct value of $\hat{\sigma}_{\beta_1,\text{sim-cal}}$ is slightly smaller than $\hat{\sigma}_{\beta_1,\text{sim}}$. This phenomenon will be the subject of future study. Overall, our simulation examples show that the pair $\hat{\beta}_{1,\text{sim-cal}}$ and $\hat{\sigma}_{\beta_1,\text{sim-cal}}$ has good inference properties, with accurate coverage and minimal bias.

4 General Exposure Model

In this section we return to the more general exposure model from Section 2 that allows for spatial or spatio-temporal correlation. We need to estimate $\hat{\theta}$ in addition to $\hat{\alpha}$ in

order to derive an estimate $\hat{X}$ from equation (4), but beyond that the situation is very similar. The naive approach would be to treat the estimate $\hat{X}$ as if it were the true exposure and to derive $\hat{\beta}_1 = \beta_1(\hat{X})$ and $\hat{\sigma}_{\beta_1} = \sigma_{\beta_1,\text{class}}(\hat{X}, \hat{\sigma}^2_\varepsilon)$, where $\hat{\sigma}^2_\varepsilon$ is a method-of-moments estimate for σ^2_ε. There are three problems with ignoring the errors from estimating X:

A. The standard error $\hat{\sigma}_{\beta_1}$ is too small because it does not account for correlation in X conditional on θ and α (i.e., even if the parameters are known, the error is not strictly Berkson).

B. The standard error $\hat{\sigma}_{\beta_1}$ is too small because it does not account for variability in estimating θ and α.

C. The point estimate $\hat{\beta}_1$ is biased because it is based on estimated values for θ and α. The bias can be in either direction, depending on details of the true exposure model.

Notice that item (A) does not pertain for the uncorrelated linear exposure model discussed in the previous section, but items (B) and (C) pertain in that situation similarly to here. Thus, our strategy in what follows is to first address (A) in Section 4.1 by deriving correct standard errors when θ and α are known. Then, following the approach take in the previous section, we proceed in Sections 4.2 and 4.3 to address the standard error and bias problems in (B) and (C). The only other difference from the previous section is that an additional approximation is needed to justify the regression calibration for item (C).

4.1 Standard Error Correction (Stage I)

If the parameters α and θ are known exactly, it is natural to define

$$\begin{aligned} \hat{X}_{\text{exact}} &= \mathrm{E}(X|X^*, \alpha, \theta) \\ &= S\alpha + \Sigma_{\eta|\eta^*}(\theta)\Sigma_{\eta^*}^{-1}(\theta)(X^* - S^*\alpha). \end{aligned} \tag{8}$$

It follows from the properties of multivariate normal distributions that

$$(X - \hat{X}_{\text{exact}})|\hat{X}_{\text{exact}} = \Lambda \sim N(0, \Sigma_\Lambda(\theta)) \tag{9}$$

where

$$\begin{aligned} \Sigma_\Lambda(\theta) &= \mathrm{Cov}(X|X^*, \alpha, \theta) \\ &= \Sigma_\eta(\theta) - \Sigma_{\eta|\eta^*}(\theta)\Sigma_{\eta^*}^{-1}(\theta)\Sigma_{\eta^*|\eta}(\theta). \end{aligned} \tag{10}$$

In the above expressions Σ_η, Σ_{η^*}, and $\Sigma_{\eta^*|\eta}$ are derived from sub-matrices of Σ_{η,η^*}. This is similar to Berkson error and suggests that using $\hat{X}_{\text{exact}}$ in place of X will not introduce bias in estimating β_1. As in (6) we can reformulate the regression problem for Y in terms of $\hat{X}_{\text{exact}}$

$$\begin{aligned} Y &= \hat{X}_{\text{exact}}\beta_1 + Z\gamma + (\varepsilon + \Lambda\beta_1) \\ &= \hat{X}_{\text{exact}}\beta_1 + Z\gamma + \varepsilon', \end{aligned} \tag{11}$$

where

$$\varepsilon' \sim N(0, \Sigma_{\varepsilon'})$$

for

$$\Sigma_{\varepsilon'} = \Sigma_\varepsilon + \beta_1^2\Sigma_\Lambda(\theta). \tag{12}$$

Unlike in equation (6), the elements of ε' are neither independent nor homoscedastic and classical standard errors based on assuming a diagonal covariance matrix are not correct (Carroll et al. 1995, page 63).

One approach to accounting for the non-i.i.d. error structure is to use generalized least squares (GLS). This seems to work well in the ordinary kriging examples of Madsen et al. (2007) where it is called "krige-and-regress". Their methodology also includes a correction to the standard error to account for variability in estimating X. Gryparis et al. (2006) also consider a GLS methodology, but without the additional standard error correction. Both papers note that since β_1 appears in the mean and in the non-diagonal covariance matrix that determines weights, GLS can not be applied in a straightforward way and there is the potential for numerical difficulties.

Instead of using GLS, we adopt the philosophy of generalized estimating equations (Liang and Zeger 1986). That is, we estimate

$$\hat{\beta}_{1,\text{exact}} = \beta_1(\hat{X}_{\text{exact}}) \tag{13}$$

by OLS using (2) and then derive standard errors from the "sandwich" estimator (Carroll et al. 1995, Section A.3)

$$\begin{pmatrix} \sigma_{\beta_1,\text{sand}}(X, \Sigma_\varepsilon) \\ \sigma_{\gamma,\text{sand}}(X, \Sigma_\varepsilon) \end{pmatrix} = \Big[\text{diag}\Big(A(X)^{-1}B(X, \Sigma_\varepsilon)A(X)^{-1}\Big)\Big]^{1/2} \tag{14}$$

with $\hat{X}_{\text{exact}}$ plugged in for X and $\hat{\Sigma}_{\varepsilon',\text{exact}}$ plugged in for Σ_ε, where $\hat{\Sigma}_{\varepsilon',\text{exact}}$ is an estimate of the now non-diagonal covariance matrix of ε' in equation (12). The "bread" in the

"sandwich" estimator is defined by

$$A(X) = \begin{pmatrix} X \\ Z \end{pmatrix}^t \begin{pmatrix} X \\ Z \end{pmatrix}$$

and the "meat" is

$$B(X, \Sigma_\varepsilon) = \begin{pmatrix} X \\ Z \end{pmatrix}^t \Sigma_\varepsilon \begin{pmatrix} X \\ Z \end{pmatrix}.$$

What remains is to construct $\hat{\Sigma}_{\varepsilon',\text{exact}}$ as an approximation to (12). Since θ is assumed to be known, we can derive $\Sigma_\Lambda(\theta)$ from (10) and plug in $\hat{\beta}_{1,\text{exact}}$. As an approximation to σ^2_ε, we take difference between the variance of residuals from (11) and the average of the diagonal entries in $\Sigma_\Lambda(\theta)\hat{\beta}^2_{1,\text{exact}}$. That is

$$\hat{\sigma}^2_{\varepsilon,\text{exact}} = \sigma^2_\varepsilon(\hat{X}_{\text{exact}}, \hat{\beta}_{1,\text{exact}}, \hat{\gamma}_{\text{exact}}, \theta)$$

where $\hat{\gamma}_{\text{exact}} = \gamma(\hat{X}_{\text{exact}})$ and

$$\sigma^2_\varepsilon(X, \beta_1, \gamma, \theta) = \frac{1}{N}\sum_{i=1}^{N}\left[Y_i - \begin{pmatrix} X \\ Z \end{pmatrix}\begin{pmatrix} \beta_1 \\ \gamma \end{pmatrix}_i\right]^2 - \beta_1^2 \frac{1}{N}\sum_{i=1}^{N}\left[\Sigma_\Lambda(\theta)\right]_{ii}. \qquad (15)$$

This is one of several reasonable method-of-moments estimators for σ^2_ε in this setting. It is also possible to estimate σ^2_ε using likelihood or Bayesian methods. Another alternative is to estimate $\Sigma_{\varepsilon'}$ non-parametrically from the data using a "weighted empirical adaptive variance estimator" (WEAVE) or similar adaptation of generalized estimating equations to spatial or spatio-temporal correlation structures (Lumley and Heagerty 1999). Extensions to accommodate correlated errors in ε are also possible.

To summarize in the case where θ and α are known exactly, the Berkson-like error structure means that the OLS estimator $\hat{\beta}_{1,\text{exact}}$ defined by (2), (8), and (13) is unbi-

ased. Based on (11) we get a standard error estimate $\hat{\sigma}_{\beta_1,\text{exact}}$ by using the "sandwich" form (14), replacing X by $\hat{X}_{\text{exact}}$ and Σ_ε by

$$\hat{\Sigma}_{\varepsilon',\text{exact}} = \hat{\sigma}^2_{\varepsilon,\text{exact}} I_N + \Sigma_\Lambda(\theta)\hat{\beta}^2_{1,\text{exact}}.$$

The simulation examples in Section 5 show that the estimators defined in this subsection have good frequentist properties.

4.2 Standard Error Correction (Stage II)

The development in the previous subsection assumes that θ and α are known exactly. This is not the case, but they can be estimated based on the model for X^* implied by (4) and (5)

$$X^* = S^*\alpha + \eta^*, \quad \eta^* \sim N(0, \Sigma_{\eta^*}(\theta)).$$

It is convenient to use a Bayesian model with prior distributions on θ and α and then to derive posterior random variables $\tilde{\theta}$ and $\tilde{\alpha}$.

A simple approach would be to derive "plug-in" estimators $\hat{\beta}_{1,\text{plug-in}}$ and $\hat{\sigma}_{\beta_1,\text{plug-in}}$ that incorporate the standard error correction from Section 4.1. We would do this by replacing θ and α in the derivations of $\hat{\beta}_{1,\text{exact}}$ and $\hat{\sigma}_{\beta_1,\text{exact}}$ with their posterior means

$$\hat{\theta} = \tilde{\text{E}}(\tilde{\theta}), \qquad \hat{\alpha} = \tilde{\text{E}}(\tilde{\alpha})$$

and replacing $\hat{X}_{\text{exact}}$ with

$$\hat{X}_{\text{plug-in}} = \text{E}(X|X^*, \hat{\alpha}, \hat{\theta}).$$

However, as we remarked in item (B) above, $\hat{\sigma}_{\beta_1,\text{plug-in}}$ is too small because it fails to account for the sampling variability of $\hat{\theta}$ and $\hat{\alpha}$.

We can address this problem by "parameter simulation", which entails averaging over simulated samples from the posterior random variables $\tilde{\theta}$ and $\tilde{\alpha}$. Define a random variable $\tilde{X}$ having the posterior distribution of conditional means of X by

$$\tilde{X} = \mathrm{E}(X|X^*, \tilde{\alpha}, \tilde{\theta}),$$

and $\tilde{\beta}_1$ by plugging $\tilde{X}$ into the OLS estimator (2)

$$\tilde{\beta}_1 = \beta_1(\tilde{X}).$$

The corresponding posterior standard error from the sandwich estimator (14) is

$$\tilde{\sigma}_{\beta_1} = \sigma_{\beta_1,\mathrm{sand}}(\tilde{X}, \tilde{\Sigma}_{\varepsilon'})$$

where

$$\tilde{\Sigma}_{\varepsilon'} = \sigma^2_{\varepsilon}(\tilde{X}, \tilde{\beta}_1, \tilde{\gamma}, \tilde{\theta})I + \Sigma_{\Lambda}(\tilde{\theta})\tilde{\beta}_1^2$$

with $\tilde{\gamma} = \gamma(\tilde{X})$. It is then natural to define

$$\hat{\beta}_{1,\mathrm{sim}} = \tilde{E}(\tilde{\beta}_1)$$

and

$$\hat{\sigma}_{\beta_1,\mathrm{sim}} = \sqrt{\tilde{\mathrm{E}}(\tilde{\sigma}^2_{\beta_1}) + \tilde{\mathrm{V}}\mathrm{ar}(\tilde{\beta}_1)},$$

where $\tilde{\mathrm{E}}$ and $\tilde{\mathrm{V}}\mathrm{ar}$ refer to the posterior expectation and variance, respectively.

As in Section 3, the distinction from the "exposure simulation" approach that is found to perform poorly in Gryparis et al. (2006) is important. Our $\tilde{X}$ does not sample from the full posterior for X, but only from the conditional mean of the posterior that is informed by X^* and estimates of θ and α. Therefore, our approach does not

introduce the kind of bias toward the null that would be characteristic of adding classical measurement error. Instead, sampling from $\tilde{\theta}$ and $\tilde{\alpha}$ inflates the standard error to properly account for uncertainty from parameter estimation. While there is still residual bias in $\hat{\sigma}_{\beta_1,\text{sim}}$, it vanishes asymptotically for large N^* since θ and α can be estimated arbitrarily well. This does not occur with "exposure simulation". The next subsection describes how we correct for the residual bias using a form of regression calibration.

4.3 Bias Correction

Although $\hat{\beta}_{1,\text{sim}}$ is asymptotically unbiased for large N^* corresponding to the situation with many exposure monitors, in practice N^* tends to be relatively small. So we use a form of regression calibration to correct the residual bias. Since $\hat{\beta}_{1,\text{exact}}$ is an unbiased estimator for β_1, similarly to (7) we can write the bias of $\hat{\beta}_{1,\text{sim}}$ as

$$\begin{aligned} \pi &= \mathrm{E}\Big[\tilde{E}(\beta_1(\tilde{X})) - \beta_1(\hat{X}_{\text{plug-in}})\Big] + \Big[\mathrm{E}(\beta_1(\hat{X}_{\text{plug-in}})) - \beta_1(\hat{X}_{\text{exact}})\Big] \\ &\approx \mathrm{E}\Big[\tilde{E}(\beta_1(\tilde{X})) - \beta_1(\tilde{\mathrm{E}}\tilde{X})\Big] + \Big[\mathrm{E}(\beta_1(\hat{X}_{\text{plug-in}})) - \beta_1(\mathrm{E}(\hat{X}_{\text{plug-in}}))\Big] \end{aligned} \tag{16}$$

The approximate equality in the second line comes from the assumption that

$$\mathrm{E}(\hat{X}_{\text{plug-in}}) \approx \hat{X}_{\text{exact}}.$$

The corresponding relationship in (7) holds exactly, but in this case the potentially nonlinear role of θ in $E(X|X^*, \alpha, \theta)$ and details of the choices for priors on θ and α make it an approximation.

Finally, we assume that for almost every realization of X^*, the posterior distribution of $\tilde{X}$ is approximately equal to the sampling distribution of $\hat{X}_{\text{plug-in}}$. For independent data this would would be a consequence of the Bernstein-von Mises theorem. Le Cam

and Yang (2000) give abstract conditions for a Bernstein-von Mises theorem in terms of local asymptotic Normality that show independence is not necessary, and our experience with spatial and spatio-temporal modeling suggests that the result is at least approximately true, although large N^* may be required for the asymptotics to dominate.

As in Section 3, we define an estimate of the bias by

$$\begin{aligned} \hat{\pi} &= 2\Big[\tilde{E}(\beta_1(\tilde{X})) - \beta_1(\tilde{E}(\tilde{X}))\Big] \\ &= 2(\hat{\beta}_{1,\text{sim}} - \hat{\beta}_{1,\text{plug-in}}) \end{aligned}$$

and then define the calibrated estimator

$$\hat{\beta}_{1,\text{sim-cal}} = 2\hat{\beta}_{1,\text{plug-in}} - \hat{\beta}_{1,\text{sim}}. \tag{17}$$

This procedure generally works well at reducing the bias in our simulations. It may be possible to improve on the calibration by refining the approximation in the second line of equation (16) for specific covariance models.

We neglect changes in variability from the calibration and set

$$\hat{\sigma}_{\beta_1,\text{sim-cal}} = \hat{\sigma}_{\beta_1,\text{sim}}.$$

It appears from our simulation examples in Section 5 that the correct value of $\hat{\sigma}_{\beta_1,\text{sim-cal}}$ is slightly smaller than $\hat{\sigma}_{\beta_1,\text{sim}}$. This phenomenon will be the subject of future study.

5 Simulations

5.1 Linear exposure model

We begin with a simulation study in the context of the linear regression exposure model from Section 3. The number of subjects is set at $N = 1010$, and we consider varying

numbers of exposure monitors $N^* = 25$, 50, 100, or 200. The exposure model is

$$X = 2 + 8S + 3N(0, I_N)$$

where I_N is the $N \times N$ identity matrix and the covariate S is random with a uniform distribution on $[0, 1]$. The model for the health outcome is

$$Y = 2 + X + 2N(0, I_N),$$

so the parameter of interest is $\beta_1 = 1$. Results from 5000 simulations are shown in Table 1.

As expected, using an estimated exposure based on the exact parameters gives correct inference for $\hat{\beta}_{1,\text{exact}}$ with larger standard errors than if the true exposure were known (pure Berkson error). Plugging in an exposure based on estimated parameters and treating it as known gives $\hat{\beta}_{1,\text{plug-in}}$ and standard errors $\hat{\sigma}_{\beta_1,\text{plug-in}}$ that are too small, resulting in significantly less than nominal coverage. If we attempt to correct this by using "exposure simulation", the estimate of β_1 is biased toward the null regardless of the number of exposure monitors. Instead, we successfully correct the standard errors by doing "parameter simulation" and averaging over simulated samples from the posterior distribution of the exposure model parameters to get $\hat{\beta}_{1,\text{sim}}$ and $\hat{\sigma}_{\beta_1,\text{sim}}$. Notice that for $N^* = 25$, the expected value of $\hat{\sigma}_{\beta_1,\text{sim}}$ does not agree with the standard deviation of $\hat{\sigma}_{\beta_1,\text{sim}}$ but the coverage for 95% confidence intervals is nearly nominal. This suggests that $\hat{\sigma}_{\beta_1,\text{sim}}$ generally performs well but that outliers affect the estimated mean values over the set of 5000 simulations. Although bias is not a significant contributor to the error (i.e., $\text{SD}(\hat{\beta}_1)$ is very close to the RMSE), regression calibration reduces the small residual bias that is present (at least for $N^* \geq 50$).

5.2 Universal kriging model

We now consider simulations using exposures that follow a universal kriging structure. In a 500×500 box we randomly select locations for $N = 410$ subjects and $N^* = 25$, 50, or 100 exposure monitors. The universal kriging exposure model is

$$X = 2 + \alpha_1 S + \eta$$

with uncorrelated random covariate S drawn from a uniform distribution on $[0, 1]$. The error term η has an exponential variogram structure with parameters

$$\theta = (\phi, \psi, \tau),$$

where ϕ is the inverse range, ψ is the partial sill, and τ is the nugget (Cressie 1993). The model for the health outcome is the same as in Section 5.1.

We consider four sets of exposure model parameters, corresponding to combinations of *short* and *long* correlation ranges and *strong* and *weak* dependence on the covariate S. The parameter specifications are given in Table 2, and example realizations are shown in Figure 1. Notice that since the covariate S has no spatial structure, strong dependence on S has the appearance of adding spatial roughness. However, it is part of the mean model so it affects the inference differently than having a short correlation range.

For each of twelve scenarios corresponding to $N^* = 25$, 50, or 100 and the four classes of spatial exposure field, we simulate 5000 Monte-Carlo runs and estimate β_1 and σ_{β_1} according to the methodology described in Section 4. To estimate parameters in the exposure model, we fit Bayesian kriging using the spBayes package for R (Finley

et al. 2007; R Development Core Team 2008) with diffuse inverse gamma priors $\Gamma^{-1}(2,5)$ on ψ and τ and with a uniform prior $U(0.003, 0.3)$ on ϕ^{-1} . We run a single chain for ten thousand iterations and discard the first two thousand iterations as "burn-in". We graphically verify for a small subset of Monte-Carlo runs that the chain has good mixing and convergence properties. The posterior distribution is approximated by 100 samples from the MCMC output.

The results for the four combination of *short* and *long* correlation ranges and *strong* and *weak* dependence on the covariate S are shown in Tables 3 - 6. The findings are generally consistent with the theory developed in Section 4. Using estimated exposures based on the exact parameters (which are not known in practice) gives unbiased estimates for β_1, but the standard errors are too small resulting in less than nominal coverage for 95% confidence intervals. As expected, this discrepancy is more pronounced in the exposure fields with long ranges since these have more spatial correlation. Applying the correction from Section 4.1 to obtain $\hat{\sigma}_{\beta_1,\text{exact}}$ inflates the standard errors to appropriately account for spatial correlation and results in nominal coverage for the confidence intervals.

The "plug-in" estimator $\hat{\beta}_{1,\text{plug-in}}$ obtained by treating the estimated kriging parameters as known and plugging in the corresponding exposure field is slightly biased. More importantly, the corresponding confidence intervals based on $\hat{\sigma}_{\beta_1,\text{plug-in}}$ (which already incorporate the standard error correction from Section 4.1) are too narrow. Using "exposure simulation" in an attempt to correct the standard errors leads to significant bias toward the null. We successfully correct the standard errors to account for parameter estimation by employing "parameter simulation" and averaging over simulated

samples from the posterior parameter distribution and deriving $\hat{\beta}_{1,\text{sim}}$ and $\hat{\sigma}_{\beta_1,\text{sim}}$ as in Section 4.1. The coverage of the corresponding 95% confidence intervals is close to nominal, but this procedure actually introduces additional bias. In some scenarios there is a significant difference between the standard deviation of $\hat{\beta}_{1,\text{sim}}$ and its root-mean-square-error (RMSE), indicating that the bias is an important contributor to the overall error. In most of our simulation scenarios the bias is away from the null, but in some cases it is toward the null.

The regression calibration of Section 4.3 gives $\hat{\beta}_{1,\text{sim,cal}}$ which tends to have less bias than $\hat{\beta}_{1,\text{sim}}$. In particular, the RMSE of $\hat{\beta}_{1,\text{sim,cal}}$ is generally smaller than the RMSE of $\hat{\beta}_{1,\text{sim}}$, and it also agrees well with the standard deviation of $\hat{\beta}_{1,\text{sim,cal}}$. The regression calibration works better for the exposure fields with *strong* covariate dependence than for those with *weak* covariate dependence. This is not surprising since the *strong* covariate dependence cases are closer to the uncorrelated exposure model of Section 3, and the calibration procedure's validity requires fewer assumptions in that setting.

6 Example

An example application is the problem considered by Madsen et al. (2007) of estimating the association between the log of chloride levels in streams

$$Y = \text{log\_cl}$$

and the logit of the fraction of forestation in the local watershed

$$X = \text{logit\_for}.$$

We do not include additional covariates, so the assumed model is

$$Y = X\beta_1 + Z\gamma + \varepsilon,$$

where Z is a vector of ones and γ is the intercept. Although it is plausible that ε has a spatial correlation structure, for the purposes of this analysis it is assumed to be uncorrelated. We utilize data collected by the Environmental Protection Agency in the Mid-Atlantic Highlands region of the eastern United States during the years 1993-1996. The data are freely available on the internet (U.S EPA Environmental Monitoring and Assessment Program 1999).

For a subset of $N = 157$ streams the dataset includes only the outcome Y. Where multiple measurements are available from different times, we use the earliest time. For another $N^* = 337$ streams the dataset includes the exposure X^* and the outcome Y^* (we have restricted to streams with percent forestation strictly between 0% and 100% in order to allow for the logit transformation). All of the locations are shown on a map in Figure 2. Since we actually have full data at $N^* = 337$ streams, we can examine the association between chloride levels and forestation without having to estimate the exposure. Simple linear regression gives $\hat{\beta}_{1,\text{true}} = -0.278$ with a standard error of $\hat{\sigma}_{\beta_1,\text{true}} = 0.026$. A scatterplot is shown in Figure 3.

In the remainder of the example, we assume that Y^* is not available and regard X^* as exposure monitoring data as in Section 2. We need to predict the exposure vector X at the $N = 157$ locations with measured outcomes Y in order to estimate β_1. We do this using a universal kriging model with a linear trend in latitude and longitude. An empirical binned variogram for the residuals of X^* is shown in Figure 4. Distances

are calculated using a flat-earth approximation, with one degree of latitude equal to 111.3 km and one degree of longitude equal to 85.9 km.

Based on this empirical variogram we assume an exponential covariance model with the same parameterization as in Section 5.2, and we estimate the posterior distribution of the parameters in a Bayesian setting using the same methodology. The only difference is that in this case we use a uniform prior $U(0.001, 0.3)$ on the inverse range parameter ϕ. The posterior means (standard deviations) of the variogram parameters are as follows:

$$\begin{aligned} \text{Range}^{-1}\ (\phi) &= 0.033(0.011) \\ \text{Partial sill}\ (\psi) &= 3.01(0.58) \\ \text{Nugget}\ (\tau) &= 1.14(0.36). \end{aligned}$$

The Bayesian kriging model can be used to estimate X at the $N = 157$ locations without exposure data. Results of using the methods in Section 4 to make inference on β_1 are given in the top half of Table 7. The estimated value of β_1 is approximately -0.24 for all of our methods, which is close to $\hat{\beta}_{1,\text{true}} = -0.278$ that was obtained with true exposure and outcome data. The standard error for the plug-in estimator without any correction is 0.0786, and the first stage correction from Section 4.1 has little effect. This is consistent with the fact that we estimate little spatial structure in the exposure field (range of approximately 30 km). The "parameter simulation" standard error correction from Section 4.2 accounts for uncertainty in estimating parameters in the exposure model and slightly increases the standard error to 0.0835. There is little change from the bias correction of Section 4.3, potentially because $N^* = 337$ is relatively large and the bias vanishes asymptotically. Overall these results are consistent with our

expectations, although they suggest that $N^* = 337$ is enough exposure monitors to justify ignoring the errors described in items (B) and (C) of Section 4. As expected, "exposure simulation" results in a seriously biased estimate for β_1 of -0.0808.

As a sensitivity analysis, we evaluate what would happen if there were only $N^* = 157$ exposure monitors. Results are given in the bottom half of Table 7, averaged over 1000 randomly selected subsets of monitor locations. Overall the mean estimate of β_1 is approximatley -0.30, which is different from the value obtained using the full set of exposure monitors but is still close to $\hat{\beta}_{1,\text{true}} = -0.278$. Now there is a significant inflation of the standard error from "paramter simulation" compared to simply plugging in the posterior mean parameters, and there is also a slight change in the estimate of β_1 from the bias correction of Section 4.3. This suggests that with only $N^* = 157$ exposure monitors it is important to account for the errors described in items (B) and (C) of Section 4. In particular, ignoring the uncertainty from estimating the exposure model parameters results in standard error estimates $\hat{\sigma}_{\beta_1,\text{plug-in}}$ that are on average 25% smaller than the corrected values $\hat{\sigma}_{\beta_1,\text{sim}}$.

7 Discussion

The question of how to use estimated exposures in a regression model is of increasing interest in air pollution epidemiology and other applications of environmental epidemiology and statistics. There is a growing recognition that ignoring the uncertainty in exposure induces serious and complicated errors in the resultant effect estimates (Kim et al. 2008). It is also clear that the errors do not fit into the typical categories of "clas-

sical" or "Berkson" measurement error, and standard regression calibration methods do not apply (Carroll et al. 1995). While there has been some progress in the recent literature (Madsen et al. 2007; Gryparis et al. 2007), to our knowledge this is the first systematic description of the three sources of error that can occur along with general methods for correcting each of them.

A key insight is that incorrect standard errors resulting from variability in estimation of the exposure can be accounted for by averaging over simulations from the part of the posterior exposure field that is informative for the outcome. This amounts to "parameter simulation", which entails simulating from the posterior parameters and averaging effect estimates derived from the corresponding conditional mean exposure fields. There are important implications for how environmental exposures should be modeled. Universal kriging and related spatio-temporal methodologies that explicitly model the residual correlation with a parametric covariance lend themselves to our approach. Some authors have suggested using semi-parametric smoothing as a less computationally demanding alternative (Gryparis et al. 2006; Kunzli et al. 2005). This can yield good estimates of the mean exposure field, but there is no natural way to identify the part of the posterior exposure field that is informative for the health outcome. Thus, it is not clear how to account for variability in the exposure estimates. Our methodology applies for a general parametric correlated exposure model where the posterior is Gaussian conditional on the parameters. In future research we will explore how to isolate the informative part of the posterior for non-Gaussian models.

In the interest of clarity, we have focused on a relatively simple health effect model, restricting to linear regression for a continuous and uncorrelated outcome. Extension

to a correlated outcome would require a modification of the standard error correction in Section 4.2. The method of moments estimate of the standard deviation for ε will no longer work, but it seems reasonable to use likelihood based methods if the correlation in ε can be described parametrically or to use the WEAVE covariance estimator (Lumley and Heagerty 1999). Extension to generalized linear models is also of interest. The theory is more complicated, but we expect that our methods will work well, especially in the typical environmental epidemiology situations where the effect size is small.

Finally, we remark that the bias correction of Section 4.3 involves approximations that could be improved for specific exposure models. In particular, the approximate equality in the second line of equation (16) results from nonlinearity of the covariance as a function of its parameters. In future work we will look at improving the approximation with higher order corrections for specific covariance models. This would come at the expense of the simple formula in equation (17). In the examples we have considered in this paper, the bias is a relatively minor contributor to the error and the simple correction works reasonably well.

8 Acknowledgments

Research described in this manuscript has been funded wholly or in part by the United States Environmental Protection Agency through R831697 to the University of Washington. It has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

References

S. Banerjee, B. P. Carlin, and A. E. Gelfand. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman and Hall, CRC, 2004.

R. Basu, T. J. Woodruff, J. D. Parker, L. Saulnier, and K. C. Schoendorf. Particulate air pollution and mortality: Findings from 20 u.s. cities. *New England Journal of Medicine*, 343(24):1742–1749, 2000.

N. Best, D. Spiegelhalter, D. Lunn, G. Graham, and B. Neuenschwander. Cutting 'feedback' in Bayesian full probability models. In *IceBUGS Conference*, 2006.

G. E. P. Box and G. C. Tiao. *Bayesian Inference in Statistical Analysis*. Wiley, 1992.

M. Brauer, G. Hoek, P. van Vliet, K. Meliefste, P. Fischer, U. Gehring, J. Heinrich, J. Cyrys, T. Bellander, M. Lewne, and B. Brunekreef. Estimating long-term average particulate air pollution concentrations: Application of traffic indicators and geographic information systems. *Epidemiology*, 14(2):228–239, 2003.

R. J. Carroll, D. Ruppert, and L. A. Stefanski. *Measurement Error in Nonlinear Models: A Modern Perspective, Second Edition*. Chapman and Hall/CRC, 1995.

N. A. C. Cressie. *Statistics for Spatial data*. John Wiley and Sons, New York, 1993.

D. W. Dockery, C. A. Pope, X Xu, J. D. Spangler, J. H. Ware, M. E. Fay, B. G. Ferris, and F. E. Speizer. An association between air pollution and mortality in six cities. *New England Journal of Medicine*, 329(24):1753–1759, 1993.

A. O. Finley, S. Banerjee, and B. P Carlin. spBayes: An R package for univariate

and multivariate hierarchical point-referenced spatial models. *Journal of Statistical Software*, 19(4):1–24, 2007.

D. Freedman. Wald lecture: On the Bernstein-von Mises theorem with infinite-dimensional parameters. *Annals of Statistics*, 27(4):1119–1141, 1999.

A. Gryparis, C. J. Paciorek, A. Zeka, J. Schwartz, and B. A. Coull. Measurement error caused by spatial misalignment in environmental epidemiology. *Harvard University Biostatistics Working Paper Series*, (59), 2006.

A. Gryparis, B. A. Coull, J. Schwartz, and H. H. Suh. Semiparametric latent variable regression models for spatio-temporal modeling of mobile source particles in the greater Boston area. *Journal of the Royal Statistical Society, Series C*, 56(2):183–209, 2007.

A. T. Herlihy, J. T. Stoddard, and C. B. Johnson. The relationship between stream chemistry and watershed land cover in the Mid Atlantic region, U.S. *Water, Air, and Soil Pollution*, 105:377–386, 1998.

M. Jerrett, A. Arain, P. Kanaroglou, B. Beckerman, D. Potoglou, T. Sahsuvaroglu, J. Morrison, and C. Giovis. A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology*, 15: 185–204, 2005a.

M. Jerrett, R. T. Burnett, R. Ma, C. A. Pope, D. Krewski, K. B. Newbold, G. Thurston, Y. Shi, N. Finkelstein, E. E. Calle, and M. J. Thun. Spatial analysis of air pollution mortality in Los Angeles. *Epidemiology*, 16(6):727–736, 2005b.

S. Y. Kim, L. Sheppard, and H. Kim. Influences of prediction approaches for spatially-dependent air pollution exposure on health effect estimation. *Submitted*, 2008.

N. Kunzli, M. Jerrett, W. J. Mack, B. Beckerman, L. LaBree, F. Gilliland, D. Thomas, J. Peters, and H. N. Hodis. Ambient air pollution and atherosclerosis in Los Angeles. *Environmental Health Perspectives*, 113(2):201–206, 2005.

L. Le Cam and G. L. Yang. *Symptotics in Statistics*. Springer, 2000.

K-Y Liang and S. L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22, 1986.

J. Lindstron and F. Lindgren. A Gaussian Markov random field model for total yearly precipitation over the African Sahel. *Lund University Mathematical Statistics Series*, 2008.

S. O. Los, G. P. Weedon, P. R. J. North, J. D. Kaduk, C. M. Taylor, and P. M. Cox4. An observation-based estimate of the strength of rainfall-vegetation interactions in the Sahel. *Geophsyical Research Letters*, 33:L16402, 2006.

T. Lumley and P. Heagerty. Weighted empirical adaptive variance estimators for correlated data regression. *Journal of the Royal Statistical Society, Series B*, 61(Part 2): 459–477, 1999.

L. Madsen, D. Ruppert, and N. S. Altman. Regression with spatially misaligned data. *Environmetrics*, in press, 2007.

K. A. Miller, D. S. Sicovick, L. Sheppard, K. Shepherd, J. H. Sullivan, G. L. Anderson,

and J. D. Kaufman. Long-term exposure to air pollution and incidence of cardiovascular events in women. *New England Journal of Medicine*, 356(5):447–458, 2007.

C. A. Pope, R. T. Burnett, M. J. Thun, E. E. Calle, K. Ito D. Krewski, and G. D. Thurston. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association*, 9(287):1132–1141, 2002.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. URL `http://www.R-project.org`. ISBN 3-900051-07-0.

B. Ritz, M. Wilhelm, and Y. Zhao. Air pollution and infant death in southern California, 19892000. *Pediatrics*, 118(2):493–502, 2006.

J. M. Samet, F. Dominici, F. Curriero, I. Coursac, and S. L . Zeger. Particulate air pollution and mortality: Findings from 20 u.s. cities. *New England Journal of Medicine*, 343(24):1742–1749, 2000.

U.S EPA Environmental Monitoring and Assessment Program. Mid-Atlantic Streams 1993-96 Data Sets, 1999. `http://www.epa.gov/emap/html/data/surfwatr/data/mastreams/9396/wchem/chmval.txt`, `http://www.epa.gov/emap/html/data/surfwatr/data/mastreams/9396/location/watchr.txt`.

S. L. Zeger, D. Thomas, F. Dominici, J. M. Samet, J. Schwartz, D. Dockery, and A. Cohen. Exposure measurement error in time-series studies of air pollution: Concepts and consequences. *Environmental Health Perspectives*, 108(59):419–423, 2000.

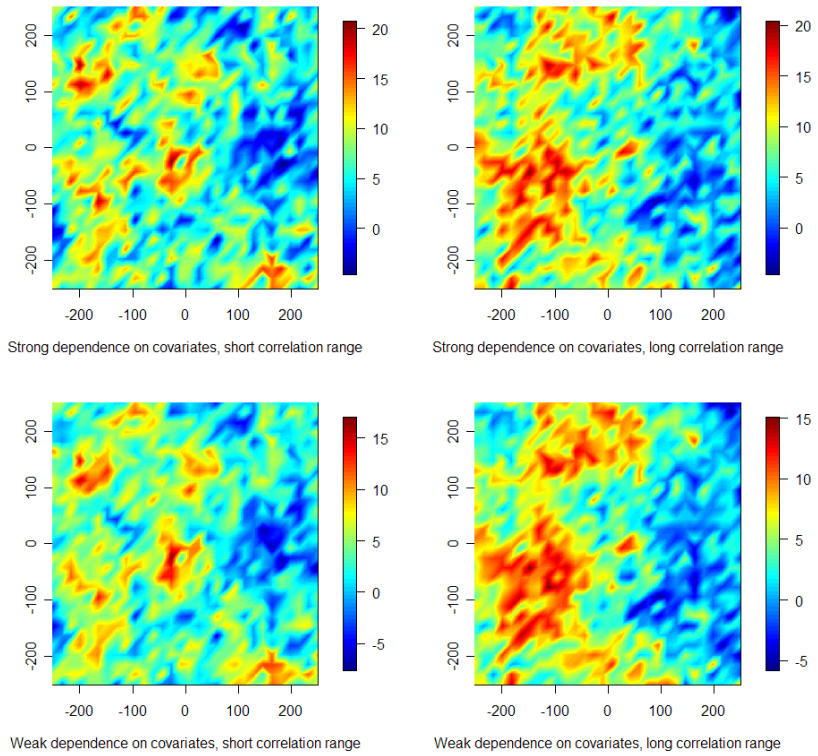

Figure 1: Examples of simulated spatially correlated exposure surfaces. Strong dependence on covariates and short correlation range (Top, Left); Strong dependence on covariates and long correlation range (Top, Right); Weak dependence on covariates and short correlation range (Bottom, Left); Weak dependence on covariates and long correlation range (Bottom, Right).

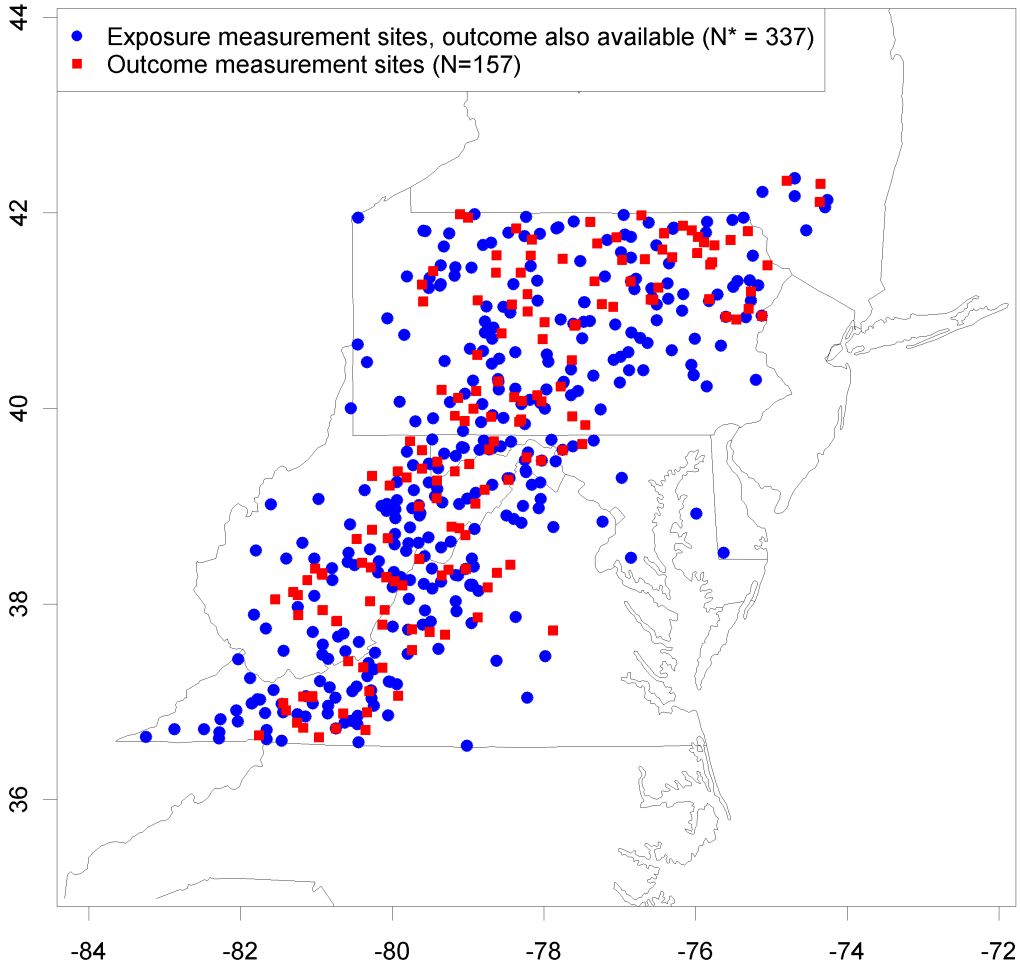

Figure 2: Locations at which full or partial stream data are available.

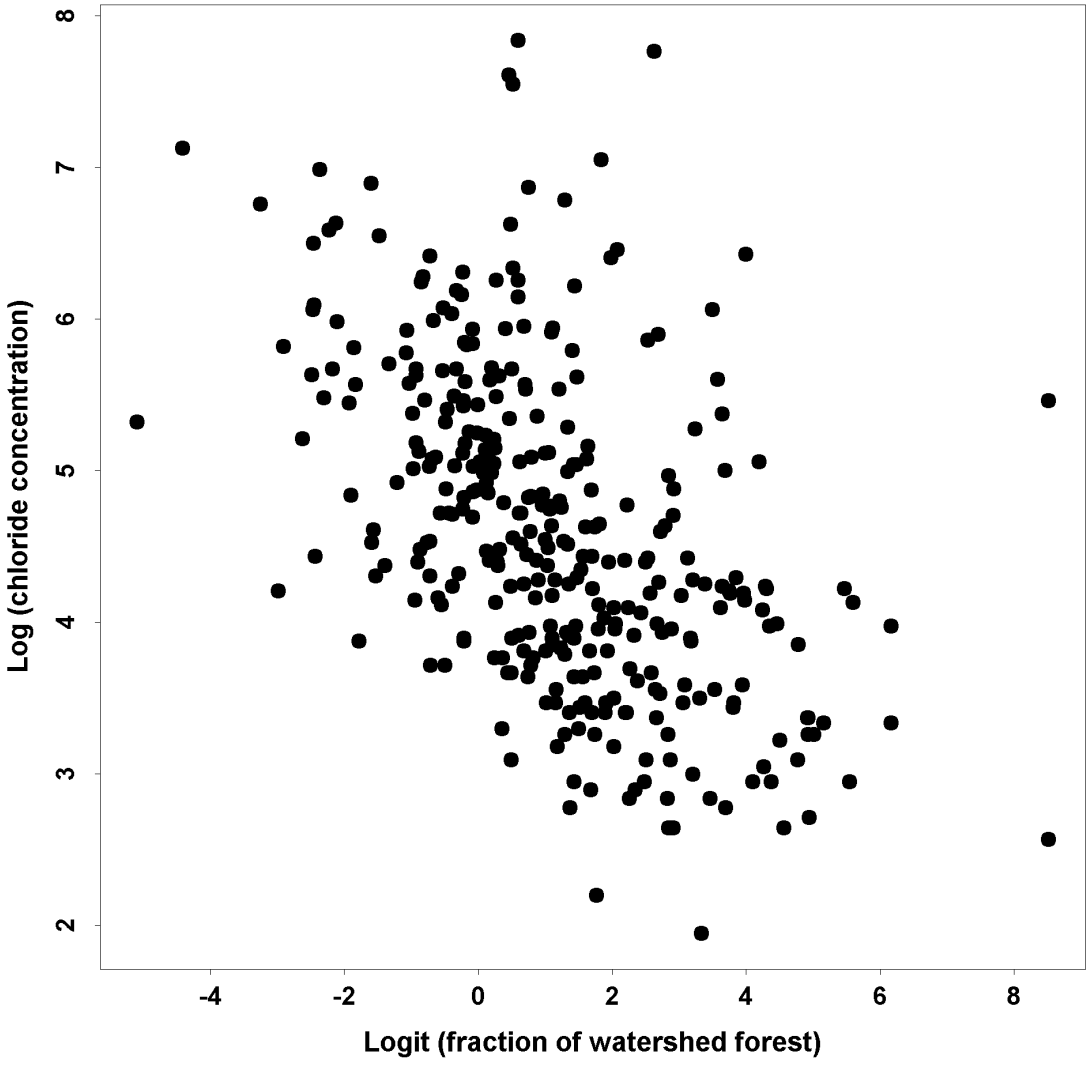

Figure 3: Scatterplot of log(chloride concentration) vs. logit(fraction of watershed forest) for the $N^* = 337$ streams at which both values are available. Simple linear regression gives $\hat{\beta}_{1,\text{true}} = -0.278$ with a standard error of $\hat{\sigma}_{\beta_1,\text{true}} = 0.026$.

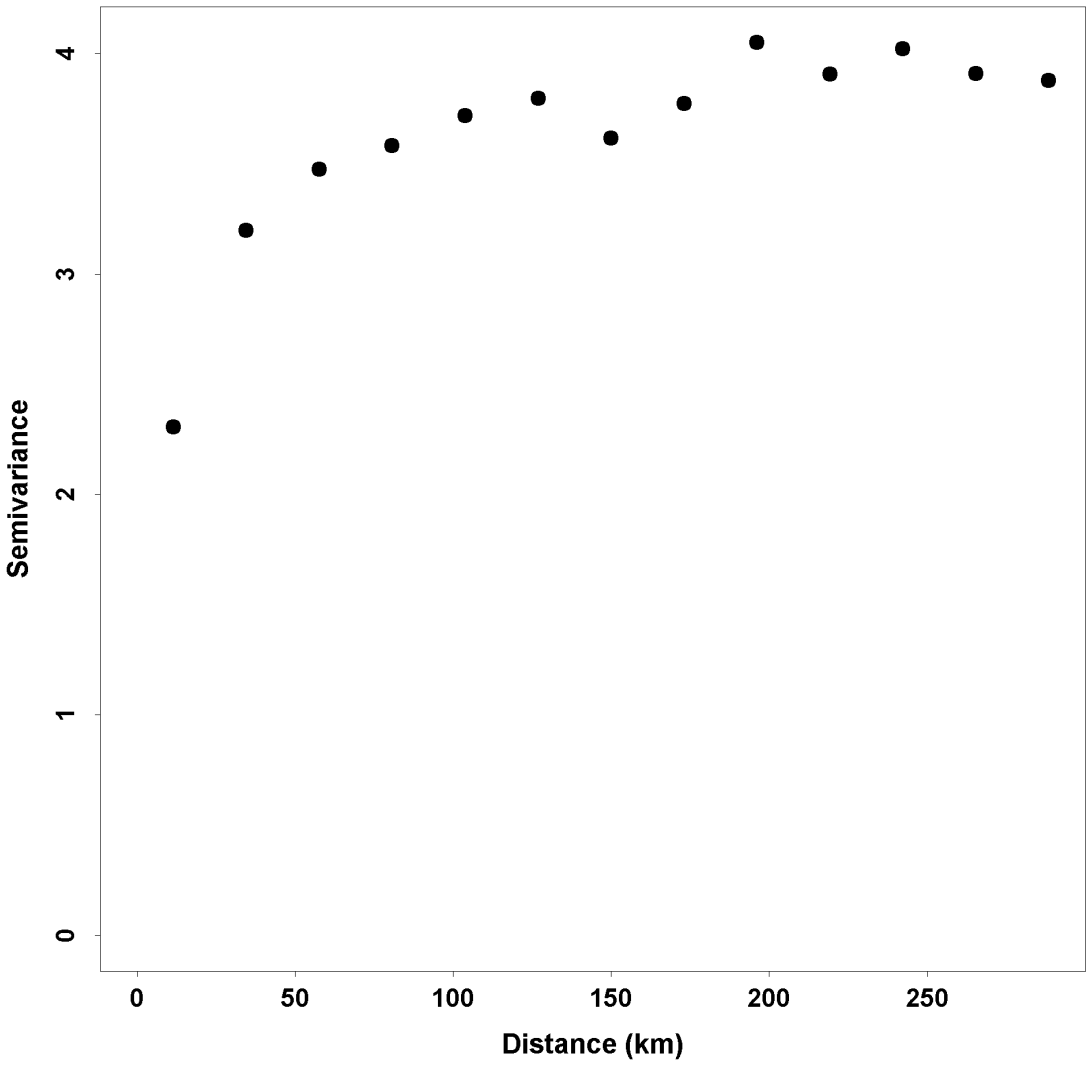

Figure 4: Empirical binned variogram for residuals of logit(fraction of watershed forest) with fitted linear trend in degrees latitude and longitude. Calculated using the $N^* = 337$ streams at which forestation data are available.

		$\mathrm{E}(\hat{\beta}_1) - \beta_1$	$\mathrm{SD}(\hat{\beta}_1)$	RMSE	$\mathrm{E}(\hat{\sigma}_{\beta_1})$	Coverage
$N^* = 20$	True exposure	−0.0002	0.0165	0.0165	0.0166	0.9478
	Exact parameters $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0002	0.0487	0.0487	0.0491	0.9474
	Plug-in parameters $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.1470	1.3743	1.3830	0.0582	0.2448
	Parameter simulation $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.5211	14.4393	14.4472	11.4845	0.9522
	Parameter simulation (bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	0.2272	14.4766	14.4769	11.4845	0.9354
	Exposure simulation	−0.6574	0.0920	0.6638	0.1029	0.0000
$N^* = 50$	True exposure	−0.0002	0.0164	0.0164	0.0166	0.9554
	Exact parameters $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	−0.0003	0.0490	0.0490	0.0491	0.9538
	Plug-in parameters $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0413	0.2357	0.2392	0.0512	0.3778
	Parameter simulation $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.0984	0.5927	0.6008	0.5664	0.9530
	Parameter simulation (bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	−0.0158	0.5431	0.5432	0.5664	0.9386
	Exposure simulation	−0.6399	0.0569	0.6424	0.0693	0.0000
$N^* = 100$	True exposure	0.0001	0.0167	0.0167	0.0166	0.9470
	Exact parameters $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0004	0.0495	0.0495	0.0491	0.9448
	Plug-in parameters $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0195	0.1519	0.1531	0.0501	0.5084
	Parameter simulation $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.0411	0.1638	0.1688	0.1631	0.9602
	Parameter simulation (bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	−0.0021	0.1421	0.1421	0.1631	0.9486
	Exposure simulation	−0.6348	0.0419	0.6362	0.0565	0.0000
$N^* = 200$	True exposure	0.0000	0.0167	0.0167	0.0166	0.9502
	Exact parameters $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	−0.0001	0.0487	0.0487	0.0491	0.9538
	Plug-in parameters $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0070	0.1074	0.1076	0.0495	0.6492
	Parameter simulation $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.0164	0.1107	0.1119	0.1105	0.9540
	Parameter simulation (bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	−0.0024	0.1050	0.1050	0.1105	0.9504
	Exposure simulation	−0.6302	0.0313	0.6310	0.0498	0.0000

Table 1: Simulation results for the linear exposure model with no correlation. Statistics from 5000 Monte-Carlo runs, 100 posterior samples for simulated parameters..

		α_1	ϕ^{-1}	ψ	τ
"Strong" dependence on covariates	"Short" correlation range	8.0	40.0	9.0	4.0
	"Long" correlation range	8.0	120.0	9.0	4.0
"Weak" dependence on covariates	"Short" correlation range	2.0	40.0	9.0	4.0
	"Long" correlation range	2.0	120.0	9.0	4.0

Table 2: Parameters for different types of spatial fields in universal kriging simulation. Regression coefficient (α_1), variogram range (ϕ^{-1}), partial sill (ψ), and nugget (τ).

		$\mathrm{E}(\hat{\beta}_1) - \beta_1$	$\mathrm{SD}(\hat{\beta}_1)$	RMSE	$\mathrm{E}(\hat{\sigma}_{\beta_1})$	Coverage
$N^* = 20$	True exposure	0.0005	0.0232	0.0232	0.0233	0.9536
	Exact parameters	0.0023	0.0898	0.0898	0.0793	0.9194
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0023	0.0898	0.0898	0.0896	0.9530
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0872	0.4348	0.4434	0.1267	0.4726
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.1113	0.3605	0.3772	0.6377	0.9434
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	0.0631	0.5989	0.6022	0.6377	0.9274
	Exposure simulation	−0.8675	0.0811	0.8713	0.1249	0.0004
$N^* = 50$	True exposure	−0.0002	0.0234	0.0234	0.0233	0.9490
	Exact parameters	0.0006	0.0859	0.0858	0.0722	0.8988
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0006	0.0859	0.0858	0.0852	0.9466
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0227	0.2082	0.2094	0.0921	0.6388
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.0569	0.2282	0.2351	0.2478	0.9526
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	−0.0115	0.1957	0.1961	0.2478	0.9404
	Exposure simulation	−0.6207	0.0600	0.6235	0.0923	0.0000
$N^* = 100$	True exposure	−0.0005	0.0233	0.0233	0.0233	0.9518
	Exact parameters	0.0012	0.0773	0.0773	0.0658	0.9032
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0012	0.0773	0.0773	0.0776	0.9498
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0076	0.1357	0.1359	0.0795	0.7630
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.0260	0.1415	0.1438	0.1462	0.9610
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	−0.0109	0.1328	0.1333	0.1462	0.9468
	Exposure simulation	−0.5669	0.0528	0.5694	0.0797	0.0000

Table 3: Simulation results for the universal kriging model with *strong* dependence on covariates and *short* correlation range. Statistics from 5000 Monte-Carlo runs, 100 posterior samples for simulated parameters, $N = 410$.

		$\mathrm{E}(\hat{\beta}_1) - \beta_1$	$\mathrm{SD}(\hat{\beta}_1)$	RMSE	$\mathrm{E}(\hat{\sigma}_{\beta_1})$	Coverage
$N^* = 20$	True exposure	0.0007	0.0245	0.0246	0.0243	0.9482
	Exact parameters	0.0003	0.1059	0.1059	0.0654	0.7786
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0003	0.1059	0.1059	0.1047	0.9442
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.1267	0.3887	0.4088	0.1195	0.4970
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.1612	0.3574	0.3920	0.5368	0.9582
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	0.0922	0.4744	0.4832	0.5368	0.9504
	Exposure simulation	−0.5969	0.0826	0.6026	0.1272	0.0034
$N^* = 50$	True exposure	0.0001	0.0247	0.0247	0.0244	0.9488
	Exact parameters	0.0004	0.0852	0.0852	0.0583	0.8194
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0004	0.0852	0.0852	0.0860	0.9502
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0375	0.1718	0.1758	0.0868	0.7242
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.0613	0.1854	0.1953	0.1920	0.9698
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	0.0137	0.1629	0.1635	0.1920	0.9606
	Exposure simulation	−0.5034	0.0766	0.5092	0.0931	0.0002
$N^* = 100$	True exposure	0.0000	0.0246	0.0246	0.0244	0.9482
	Exact parameters	0.0006	0.0723	0.0723	0.0537	0.8562
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0006	0.0723	0.0723	0.0717	0.9462
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0073	0.1102	0.1105	0.0712	0.8116
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.0208	0.1147	0.1165	0.1128	0.9530
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	−0.0061	0.1073	0.1074	0.1128	0.9470
	Exposure simulation	−0.4365	0.0708	0.4422	0.0739	0.0000

Table 4: Simulation results for the universal kriging model with *strong* dependence on covariates and *long* correlation range. Statistics from 5000 Monte-Carlo runs, 100 posterior samples for simulated parameters, $N = 410$.

		$\mathrm{E}(\hat{\beta}_1) - \beta_1$	$\mathrm{SD}(\hat{\beta}_1)$	RMSE	$\mathrm{E}(\hat{\sigma}_{\beta_1})$	Coverage
$N^* = 20$	True exposure	0.0004	0.0274	0.0274	0.0275	0.9498
	Exact parameters	0.0060	0.2948	0.2948	0.1868	0.7844
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0060	0.2948	0.2948	0.2976	0.9354
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	−0.2121	0.5293	0.5702	0.3017	0.6050
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	−0.2997	0.3569	0.4660	0.8188	0.9132
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	−0.1245	0.7711	0.7810	0.8188	0.8830
	Exposure simulation	−0.9293	0.0481	0.9305	0.1039	0.0000
$N^* = 50$	True exposure	−0.0003	0.0277	0.0277	0.0275	0.9484
	Exact parameters	0.0024	0.2125	0.2125	0.1369	0.7932
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0024	0.2125	0.2125	0.2122	0.9448
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	−0.0052	0.3514	0.3514	0.2343	0.7992
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.0259	0.3361	0.3370	0.5974	0.9560
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	−0.0364	0.3983	0.3999	0.5974	0.9414
	Exposure simulation	−0.8685	0.0564	0.8703	0.0987	0.0000
$N^* = 100$	True exposure	−0.0006	0.0273	0.0273	0.0275	0.9508
	Exact parameters	0.0013	0.1579	0.1579	0.1097	0.8266
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0013	0.1579	0.1579	0.1584	0.9446
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0146	0.2371	0.2375	0.1678	0.8642
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.0640	0.2600	0.2677	0.3235	0.9658
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	−0.0348	0.2264	0.2290	0.3235	0.9466
	Exposure simulation	−0.7942	0.0661	0.7969	0.0918	0.0000

Table 5: Simulation results for the universal kriging model with *weak* dependence on covariates and *short* correlation range. Statistics from 5000 Monte-Carlo runs, 100 posterior samples for simulated parameters, $N = 410$.

		$\mathrm{E}(\hat{\beta}_1) - \beta_1$	$\mathrm{SD}(\hat{\beta}_1)$	RMSE	$\mathrm{E}(\hat{\sigma}_{\beta_1})$	Coverage
$N^* = 20$	True exposure	0.0007	0.0296	0.0296	0.0293	0.9478
	Exact parameters	−0.0030	0.2904	0.2903	0.1199	0.5954
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	−0.0030	0.2904	0.2903	0.2897	0.9128
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0629	0.5399	0.5435	0.3267	0.7650
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	−0.0428	0.3860	0.3883	0.8673	0.9606
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	0.1687	0.7697	0.7879	0.8673	0.9470
	Exposure simulation	−0.8675	0.0811	0.8713	0.1249	0.0004
$N^* = 50$	True exposure	0.0000	0.0298	0.0298	0.0293	0.9448
	Exact parameters	0.0025	0.1876	0.1876	0.0939	0.6756
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0025	0.1876	0.1876	0.1857	0.9368
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0690	0.2838	0.2921	0.2013	0.8850
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.1070	0.2925	0.3114	0.4295	0.9778
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	0.0310	0.2957	0.2973	0.4295	0.9712
	Exposure simulation	−0.7394	0.1160	0.7484	0.1085	0.0000
$N^* = 100$	True exposure	0.0001	0.0297	0.0297	0.0293	0.9468
	Exact parameters	0.0021	0.1344	0.1344	0.0810	0.7632
	Exact parameters (variance corrected) $(\hat{\beta}_{1,\text{exact}}, \hat{\sigma}_{\beta_1,\text{exact}})$	0.0021	0.1344	0.1344	0.1336	0.9468
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	0.0168	0.1764	0.1772	0.1352	0.8992
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	0.0460	0.1923	0.1977	0.2024	0.9638
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	−0.0123	0.1660	0.1664	0.2024	0.9532
	Exposure simulation	−0.6392	0.1165	0.6497	0.0884	0.0000

Table 6: Simulation results for the universal kriging model with *weak* dependence on covariates and *long* correlation range. Statistics from 5000 Monte-Carlo runs, 100 posterior samples for simulated parameters, $N = 410$.

		$\hat{\beta}_1$	$\hat{\sigma}_{\beta_1}$
$N^* = 337$	Plug-in parameters	-0.2365	0.0786
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	-0.2365	0.0793
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	-0.2389	0.0830
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	-0.2340	0.0830
	Exposure simulation	-0.0787	0.0621
$N^* = 157$	Plug-in parameters	-0.2968	0.1085
	Plug-in parameters (variance corrected) $(\hat{\beta}_{1,\text{plug-in}}, \hat{\sigma}_{\beta_1,\text{plug-in}})$	-0.2968	0.1120
	Parameter simulation (variance corrected) $(\hat{\beta}_{1,\text{sim}}, \hat{\sigma}_{\beta_1,\text{sim}})$	-0.2920	0.1539
	Parameter simulation (variance and bias corrected) $(\hat{\beta}_{1,\text{sim,cal}}, \hat{\sigma}_{\beta_1,\text{sim,cal}})$	-0.3015	0.1539
	Exposure simulation	-0.0554	0.0692

Table 7: Results for the stream data example for $N = 157$ outcome measurements and exposures estimated by universal kriging using $N^* = 337$ and $N^* = 157$ exposure monitors. Results for $N^* = 157$ are mean values from 1000 random subsets of exposure monitors.