An Autoregressive Point Source Model for Spatial Processes

Jacqueline M. Hughes-Oliver, Tae-Young Heo, Sujit K. Ghosh

Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203

(hughesol@stat.ncsu.edu; theo@stat.ncsu.edu; sghosh@stat.ncsu.edu)

Abstract

We suggest a parametric modeling approach for nonstationary spatial processes driven by point sources. Baseline near-stationarity, which may be reasonable in the absence of a point source, is modeled using a conditional autoregressive Markov random field. Variability due to the point source is captured by our proposed autoregressive point source model. Inference proceeds according to the Bayesian hierarchical paradigm, and is implemented using Markov chain Monte Carlo in WinBUGS. The parametric approach allows a formal test of effectiveness of the source. Application is made to a real dataset on electromagnetism measurements in a field containing a metal pole.

KEY WORDS: Bayesian inference; Correlation nonstationarity; Heterogeneity; Hierarchical model; Random effect; Variance nonstationarity.

1 INTRODUCTION

Stochastic processes across spatial domains are often modeled by decomposing them into trend and stationary error processes. However, it has become increasingly clear that an assumption of stationarity of the error process is driven more by mathematical convenience than by reality, and more practitioners are choosing to replace this convenient assumption with the more realistic assumption of nonstationarity; see, for example, Sampson and Guttorp (1992), Haas (1995), Cressie and Majure (1997), Higdon, Swall and Kern (1999), Fuentes (2002), and Wikle (in press). Nonstationary error models are difficult to develop and estimate because deviations from stationarity can occur in many ways. They require direct or indirect methods for describing variance and correlation as they change over the index set of the process, while guaranteeing that linear transformations of the process will have non-negative variances. Auxiliary information, if available, can ease the difficulty associated with modeling nonstationary error processes.

In this paper, we consider the case where a point source impacts the stochastic process of interest. Auxiliary information provided by a point source and the resulting impact on the covariance function of a process has received only minimal attention. In epidemiology, relatively recent activities have led to a large body of literature relating point sources to disease "hot spots." The primary goal is to test whether or not clusters of disease cases in space and/or time are significant, after accounting for chance variation. In their nomenclature, a "focused" test is used to detect significant clustering around a point source exposure to a pollutant that is presumed to increase disease risk. These efforts typically model disease incidence using the Cox process based on inhomogeneous Poisson processes. A few key references are Diggle and Rowlingson (1994), Bithell (1995), Lawson (1995, 2001), Lawson and Waller (1996), and Diggle, Morris, Elliot, and Shaddick (1997). Hughes-Oliver, Lu, Davis, and Gyurcsik (1998), Hughes-Oliver, Gonzalez-Farias, Lu, and Chen (1998), and Hughes-

Oliver and Gonzalez-Farias (1999) discuss the effects of point sources on a deposition process in semiconductor manufacturing and measuring electromagnetism in a field.

We propose a hierarchical Bayesian approach to an extension of the process decomposition model suggested by Hughes-Oliver and Gonzalez-Farias (1999). There are three major benefits afforded by the Bayesian paradigm. First, it is able to account for uncertainty in parameter estimates when evaluating prediction uncertainty; frequentist approaches are highly criticized for their inability to deal with this issue. Second, Bayesian inference does not require derivation of asymptotic properties of estimators because results are based entirely on simulated observations from relevant posterior distributions. Third, the Bayesian paradigm allows the incorporation of prior information that may not be as easily incorporated in a frequentist approach. There are also benefits of the process decomposition model. This model provides a parametric method for decomposing the observed process into a trend surface, a baseline error process, and an additional error process (associated with the point source) that may be viewed as a "shock" to the baseline. Nonstationarity is easily modeled by this additional error process since no restrictions are placed on the forms of their covariance functions.

We also introduce a newly created autoregressive point source (ARPS) process capturing the effect of a point source on the error component of a process. This process is attractive for at least three reasons. First, it allows site-specific variances to be a function of proximity to source, where no restriction is placed on definition of distance; that is, any relevant distance metric may be used. Second, site-to-site correlations may be a function of proximity to source. Third, and possibly most attractive, this model is parametric and lends itself to testing the statistical significance of point source impacts on the error component of a process. We use the data from Hughes-Oliver and Gonzalez-Farias (1999) to illustrate our proposed method for modeling the effect of a point source.

Section 2 contains a description of the process decomposition approach. Section 3 introduces our approach, which we call the ARPS process, for modeling the effect of a point source on the covariance structure. The full, overall covariance structure is explored in Section 4. Section 5 contains modeling details for electromagnetism in a field containing a metal pole. Concluding remarks are presented in Section 6.

2 PROCESS DECOMPOSITION

Suppose $\{Y(s) : s \in D \subseteq R^2\}$ is the stochastic process of interest and D is the index set, such that $Y(s)$ is the response at site s. That is, $s = (x, y)$, where x may be longitude and y may be latitude. Covariates for this process are given as $\{X(s) : s \in D \subseteq R^2$, X is q-dimensional$\}$. These q covariates may be fixed or random and will often include distance to the point source located at P. Given covariates, the $Y(\cdot)$ process is decomposed into trend and error as

$$Y(s) = \mu(s) + Z(s), \tag{1}$$

where $\mu(s) = f(X(s), \beta)$ (the trend) is the expectation of $Y(s)$ conditioned on covariates $X(s)$ at site s and $Z(s)$ (the error) has zero expectation. Effect of the point source on trend or large-scale variation is modeled by judicious choice of the function $f(\cdot, \cdot)$. This is commonly done using a linear or non-linear function where the functional form of $f(\cdot, \cdot)$ is completely known and only the vector β is unknown. Generalized additive models (GAMs, Hastie and Tishirani 1990) may also be used as a first-stage approximation of the trend assuming errors are independent and identically distributed; see Holland, De Oliveira, Cox, and Smith (2000) for an application of GAMs to environmental data.

It is also possible, however, that the point source affects the covariance or small-scale variation; see Hughes-Oliver et al. (1998a, 1998b), and Hughes-Oliver and Gonzalez-Farias

(1999) for discussions and datasets. This effect leads to covariances that are functions of the distances (and possibly angles) between the sites and source. The resulting nonstationarity of $Y(s)$, where the covariance is site-dependent, is modeled using $Z(s)$.

Suppose that $Z(s)$ is the result of a baseline zero-mean process $Z_0(\mathbf{s})$ that has been "shocked" by the zero-mean point source process $Z_1(s)$ in such a way that

$$Z(s) = Z_0(s) + Z_1(s). \tag{2}$$

Suppose further that processes $Z_0(\cdot)$ and $Z_1(\cdot)$ are independent and that $\Sigma_i(\cdot,\cdot)$ is the covariance function for $Z_i(\cdot)$, that is, $\Sigma_i(s,t) = \mathrm{Cov}(Z_i(s), Z_i(t))$. The covariance function for $Z(\cdot)$ is then $\Sigma(s,t) = \Sigma_0(s,t) + \Sigma_1(s,t)$.

Any valid covariance may be assigned to the Σ_is and this will automatically guarantee that the covariance of $Z(\cdot)$ is valid, thus linear transformations of the $Z(\cdot)$ process will have nonnegative variances. The choice of Σ_i should be motivated by the data in such a way that the baseline process is well represented and parameters of Σ_1 measure the strength of importance of the source. Specific features may be built into the Σ_is in a variety of ways. Suppose the effect of source P is such that response is the same at all sites equi-distant from P. Then the required two-dimensional covariance $\Sigma_1(s,t)$ may be built from a one-dimensional covariance $r_1(d_s, d_t)$, where $d_s \equiv ||s - P||$ is the Euclidean distance between site s and the source. This is because $\{Z_1(s) : s \in D \subseteq R^2\}$ may be written as a process $\{Z_1^*(d_s) : d_s \in R\}$ on the real line. Another approach to building $\Sigma_1(s,t)$ is to explore the local behavior of the process around the source by means of neighbor relationships. In the next section, we propose a hierarchical model for Z_1.

3 AUTOREGRESSIVE POINT SOURCE MODELING

Building on the process decomposition model in equations (1) and (2), we propose using an autoregressive model to capture the effect of the point source. For this point source at P, we create a sequence of concentric regions R_k for $k = 1, \ldots, r$, such that

$$P \in R_r \text{ and } R_1 \subseteq (R_1 \cup R_2) \subseteq \cdots \subseteq (\cup_{k=1}^{r} R_k) = D,$$

where $R_k \cap R_{k'} = \phi$, for all $k \neq k'$. Next, we assume that the (random) spatial effect due to point source P for the region R_k is captured by α_k. In other words, $\{R_1, \ldots, R_r\}$ forms a partition of the index set D and the process Z_1 is constant across each of these regions. Finally, the random effects $\alpha_1, \ldots, \alpha_r$ arise as an $AR(1)$ process. This assumption is based on the fact that the spatial effect due to region R_k depends on the two neighboring regions R_{k-1} and R_{k+1}. More specifically, the error process caused by the point source P may be written as

$$\begin{aligned} Z_1(s) &= \sum_{k=1}^{r} \alpha_k I(s \in R_k) \\ \alpha_1 &\sim N(0, \sigma_1^2), \qquad \sigma_1^2 > 0 \\ \alpha_k | \alpha_{k-1} &\sim N(\psi \alpha_{k-1}, \sigma_2^2), \qquad \sigma_2^2 > 0, \qquad k = 2, \cdots, r. \end{aligned} \tag{3}$$

In the ensuing discussion, we will refer to the model in (3) as ARPS$(\psi, \sigma_1^2, \sigma_2^2)$, where ARPS stands for *autoregressive point source*.

To understand the impact of (3) on the covariance Σ_1, let $\mathbf{Z}_1 = (Z_1(s_1), \ldots, Z_1(s_n))'$ be the vector of responses from the point source process over all n sampling sites. Then

$\Sigma_1 \equiv \text{Var}(\mathbf{Z}_1)$ has (i,j) element

$$\sum_{k=1}^{r}\sum_{m=1}^{r} I(s_i \in R_k)I(s_j \in R_m)\text{Cov}(\alpha_k, \alpha_m)$$

where, assuming $m \geq k$,

$$\text{Cov}(\alpha_k, \alpha_m) = \psi^{m-k}\Big[\sigma_1^2\psi^{2(k-1)} + \sigma_2^2\sum_{q=0}^{k-2}\psi^{2q}\Big].$$

Suppose $\mathbf{Z}_1^\dagger$ contains the reordered entries of $\mathbf{Z}_1$ such that $\mathbf{Z}_1^\dagger = [\mathbf{Z}_{1_1}, \mathbf{Z}_{1_2}, \cdots, \mathbf{Z}_{1_r}]'$, where $\mathbf{Z}_{1_k}$ corresponds to the vector of all n_k sites in region R_k. Then, the variance-covariance matrix for $\mathbf{Z}_1^\dagger$ is the $n \times n$ matrix

$$\Sigma_1^\dagger = \text{Var}(\mathbf{Z}_1^\dagger) = \begin{bmatrix} \delta_{1,1}J_{n_1\times n_1} & \delta_{1,2}J_{n_1\times n_2} & \delta_{1,3}J_{n_1\times n_3} & \cdots & \delta_{1,r}J_{n_1\times n_r} \\ \delta_{1,2}J_{n_2\times n_1} & \delta_{2,2}J_{n_2\times n_2} & \delta_{2,3}J_{n_2\times n_3} & \cdots & \delta_{2,r}J_{n_2\times n_r} \\ \delta_{1,3}J_{n_3\times n_1} & \delta_{2,3}J_{n_3\times n_2} & \delta_{3,3}J_{n_3\times n_3} & \cdots & \delta_{3,r}J_{n_3\times n_r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \delta_{1,r}J_{n_r\times n_1} & \delta_{2,r}J_{n_r\times n_2} & \delta_{3,r}J_{n_r\times n_3} & \cdots & \delta_{r,r}J_{n_r\times n_r} \end{bmatrix},$$

where J is a matrix with all elements equal to 1, $\delta_{1,1} = \sigma_1^2$, $\delta_{k,k} = \psi^2\delta_{k-1,k-1} + \sigma_2^2$ for $k = 2, 3, \ldots, r$, and $\delta_{k,m} = \psi\delta_{k,m-1}$ for $k = 1, \ldots, r-1$ and $m = k+1, \ldots, r$. If $\psi = \pm 1$, the variance-covariance matrix $\text{Var}(\mathbf{Z}_1^\dagger)$ takes a simpler form with $\delta_{k,m} = \sigma_1^2 + (k-1)\sigma_2^2$ if $\psi = 1$ and $\delta_{k,m} = (-1)^{m-k}[\sigma_1^2 + (k-1)\sigma_2^2]$ if $\psi = -1$, for $k \leq m$. Note that $\text{Var}(\mathbf{Z}_1^\dagger)$ is always singular. Also note that there is relatively little information available for estimating σ_1^2, so in a Bayesian framework one needs to either specify a very informative prior for σ_1^2 or specify a relationship between σ_1^2 and one or more of the other parameters; we develop this idea more fully in Section 4.

The model in (3) allows for two types of nonstationarity that we loosely call variance nonstationarity and correlation nonstationarity. Variance nonstationarity is nonconstant variance and correlation nonstationarity is correlation that is not a function only of site-to-site distance. If $\sigma_1^2 > \sigma_1^2\psi^2 + \sigma_2^2$, then variance increases as sites are further from the source. On the other hand, variance is constant when $\sigma_1^2 = \sigma_1^2\psi^2 + \sigma_2^2$ and decreases when $\sigma_1^2 < \sigma_1^2\psi^2 + \sigma_2^2$. If $\psi \neq 0$, then correlation is a function of both site-to-site and site-to-source distances from P, no matter what relationships exist between σ_1^2, σ_2^2, and ψ. Consequently, we test the impact of point source P on small-scale variability by testing the null hypothesis that

$$\psi = 0 \qquad \text{and} \qquad \sigma_1^2 = \sigma_2^2$$

against an appropriate alternative; under these conditions, the process exhibits neither variance nonstationarity nor correlation nonstationarity. Using slightly different notation, let

$$\eta = \frac{\sigma_1^2}{\sigma_1^2\psi^2 + \sigma_2^2}.$$

Then

- $\eta > 1$ and $\psi = 0$ implies there is no correlation and the point source causes a single increase in variance from R_2 to R_1 (variance is the same in $R_2, \ldots, R_r$);
- $\eta > 1$ and $\psi \neq 0$ implies variance increases as sites are further from source P and correlation is a function of site-to-source distances;
- $\eta = 1$ and $\psi = 0$ implies that variance is constant and there is no correlation, i.e., the point source does not affect small-scale variability of the process Y;
- $\eta = 1$ and $\psi \neq 0$ implies that variance is constant and correlation is a function of site-to-source distances;

- $\eta < 1$ and $\psi = 0$ implies there is no correlation and the point source causes a single decrease in variance from R_2 to R_1 (variance is the same in $R_2, \ldots, R_r$); and
- $\eta < 1$ and $\psi \neq 0$ implies variance decreases as sites are further from source P and correlation is a function of site-to-source distances.

We can thus test the impact of point source P on small-scale variability by testing the null hypothesis that $\psi = 0$ and $\eta = 1$.

4 RESTRICTIONS ON THE PARAMETER SPACE

Process decomposition, in particular equation (2), separates the overall error process into two component processes, one representing the baseline (when no source acts on the system) and the other representing the point source. Point source process Z_1 was discussed in Section 3 with a comment made that restrictions are necessary for one of its parameters. This need is discussed here.

The baseline process Z_0 is assumed to include a measurement error component having covariance $\Sigma_{00} = \sigma_e^2 I_n$ and a residual component having covariance Σ_{01}. In other words, using Σ_0 to represent $\mathrm{Var}(\mathbf{Z}_0)$, where $\mathbf{Z}_0 = (Z_0(s_1), \ldots, Z_0(s_n))'$, we get

$$\Sigma_0 = \Sigma_{00} + \Sigma_{01} = \sigma_e^2 I_n + \Sigma_{01},$$

and consequently

$$\Sigma = \sigma_e^2 I_n + \Sigma_{01} + \Sigma_1$$

is the overall covariance matrix of the Y process.

Let $\mathbf{Z}_0^\dagger = [\mathbf{Z}_{0_1}, \mathbf{Z}_{0_2}, \cdots, \mathbf{Z}_{0_r}]'$ follow the order of $\mathbf{Z}_1^\dagger$ as described in Section 3. If the Z_0

process is iid (making Σ_{01} the zero matrix) and Z_1 is ARPS$(1, \sigma_1^2, \sigma_2^2)$, then

$$\begin{aligned}\Sigma^\dagger &= \sigma_e^2 I_n + \Sigma_1^\dagger \\ &= \sigma_e^2 I_n + \begin{bmatrix} \sigma_1^2 J_{n_1\times n_1} & \sigma_1^2 J_{n_1\times n_2} & \sigma_1^2 J_{n_1\times n_3} & \cdots & \sigma_1^2 J_{n_1\times n_r} \\ \sigma_1^2 J_{n_2\times n_1} & (\sigma_1^2+\sigma_2^2) J_{n_2\times n_2} & (\sigma_1^2+\sigma_2^2) J_{n_2\times n_3} & \cdots & (\sigma_1^2+\sigma_2^2) J_{n_2\times n_r} \\ \sigma_1^2 J_{n_3\times n_1} & (\sigma_1^2+\sigma_2^2) J_{n_3\times n_2} & (\sigma_1^2+2\sigma_2^2) J_{n_3\times n_3} & \cdots & (\sigma_1^2+2\sigma_2^2) J_{n_3\times n_r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_1^2 J_{n_r\times n_1} & (\sigma_1^2+\sigma_2^2) J_{n_r\times n_2} & (\sigma_1^2+2\sigma_2^2) J_{n_r\times n_3} & \cdots & [\sigma_1^2+(r-1)\sigma_2^2] J_{n_r\times n_r} \end{bmatrix}.\end{aligned}$$

In this case, the scale parameter σ_e^2 always appears with scale parameter σ_1^2 as $\sigma_e^2 + \sigma_1^2$. To improve identifiability, we can place a restriction on either σ_e^2, σ_1^2, or both. A commonly used restriction in time series analysis is one that guarantees stationarity of the one-dimensional autoregressive process, namely $\sigma_1^2 = \sigma_2^2/(1-\psi^2)$. This restriction would force $\eta = 1$ and consequently not allow variances of the Z_1 process to differ based on distance from source. Finding this unacceptable, we settled on the alternative restriction $\sigma_1^2 = \sigma_e^2$. This is quite reasonable given that σ_e^2 represents measurement variability of the baseline process Z_0. Region R_1, whose variance is σ_1^2, is farthest from the source and is thus least affected by that source. Depending on the strength of impact of the source and the relative size of study region D, we argue that R_1 should exhibit features very similar to the baseline process Z_0, thus supporting the restriction $\sigma_1^2 = \sigma_e^2$. We use this restriction throughout.

5 MODELING ELECTROMAGNETISM IN A FIELD CONTAINING A METAL POLE

5.1 Dataset and Candidate Models

To illustrate some details of our approach, we now describe a real dataset from Hughes-Oliver and Gonzalez-Farias (1999) in which a point source acts as a catalyst for the response.

Quoting these authors (pp. 63–64), "measurements [of electromagnetism] are taken at sites falling on a regular grid, as shown in [Figure 1], where the sites are one meter apart in both the vertical and horizontal directions. ... Electromagnetism is expected to be fairly constant across the field, but an existing metal pole affects the measuring device so that the constant pattern in the field is not observable. It is in this sense that we consider the metal pole to be a *point source*. ... Electromagnetism appears to be a function only of distance to the point source, and because the contours are approximately circular, there is no apparent need for rotating or rescaling the axes."

[Figure 1 about here.]

[Figure 2 about here.]

Let $Y(s)$ denote the following transformed value of electromagnetism at site s, $g(s)$:

$$Y(s) = \frac{46300 - g(s)}{4630}.$$

This simple rescaling moderates extreme behavior in our Markov chain Monte Carlo (MCMC) techniques discussed below. Another view of the data is given in Figure 2, where both electromagnetism and transformed electromagnetism are plotted as a function of distance from the point source. This figure clearly suggests that the mean and variance of $Y(s)$ are decreasing functions of distance from the point source. Using $d(s)$ to represent Euclidean distance between s and source P, we model

$$Y(s) = \mu(s) + Z(s),$$

where

$$\mu(s) = \beta_0 + \frac{\beta_1}{d(s)}$$

describes the trend surface and

$$Z(s) = Z_{00}(s) + Z_{01}(s) + Z_1(s)$$

describes the error process. In this error process, Z_{00} is a measurement error process and Z_{01} captures the effect of spatial proximity in the baseline process; Z_{00} and Z_{01} together represent the baseline process Z_0 discussed in Sections 2 and 4. The process Z_1 captures the effect of the point source, based on $r = 10$ concentric regions from the metal pole. ARPS regions $R_1, \ldots, R_{10}$ are indicated in Figure 2, along with boxplots of the transformed electromagnetism, y, within these regions. The processes Z_{00}, Z_{01}, and Z_1 all have mean zero and are assumed independent, conditioned on the mean function $\mu(\cdot)$.

We begin by assuming normality of the conditional Y process. Normality can easily be replaced by a fat-tailed distribution like the Student's t, but this would require an adjustment in interpretation of the variance-covariance matrix by incorporating a scale factor; we do not pursue this here. The Z_{00} process is assumed to follow a normal distribution with constant variance; the Z_{01} process is assumed to follow a normal conditional autoregressive (CAR) model; and the Z_1 process follows an ARPS model. More specifically, suppose D is an index set containing sites $s_1, \ldots, s_n$ on the regular grid of Figure 1, where $n = 160$. Then, the Bayesian hierarchical model may be written as follows:

$$\begin{aligned} Y(s_j) &\sim \text{Normal}(\mu(s_j), \sigma_e^2), \quad j = 1, \ldots, n \\ \mu(s_j) &= \beta_0 + \frac{\beta_1}{d(s_j)} + Z_{01}(s_j) + Z_1(s_j) \\ Z_{01}(\cdot) &\sim \text{CAR}(\rho, \sigma_c^2) \\ Z_1(\cdot) &\sim \text{ARPS}(\psi, \sigma_1^2 = \sigma_e^2, \sigma_2^2). \end{aligned} \tag{4}$$

The measurement error variance is σ_e^2. While the CAR$(1, \sigma_c^2)$ process is more popular, we prefer the more general CAR(ρ, σ_c^2) process (Sun, Tsutakawa, Kim, and He 2000) for several reasons. First, because $-1 \leq \rho \leq 1$, the former is actually a special case of the latter. Second, the value of ρ indicates the strength of spatial correlation. And third, when $-1 < \rho < 1$, CAR(ρ, σ_c^2) yields a positive definite covariance structure for Z_{01}, unlike the semidefinite covariance offered by CAR$(1, \sigma_c^2)$. The CAR(ρ, σ_c^2) process specifies conditional distributions as

$$Z_{01}(s_j)|\{Z_{01}(s_k), j \neq k\} \sim \text{Normal}\Big(\sum_{k=1}^{n} b_{jk} Z_{01}(s_k), \epsilon_j^2\Big),$$

where $b_{jj} = 0$, $b_{jk} = (\rho/N_j)$I(locations s_j and s_k are neighbors), N_j is the number of neighbors of s_j, $\epsilon_j^2 = \sigma_c^2/N_j$, and σ_c^2 is a scale parameter measuring smoothness. The overall variance of the Z_{01} process is obtained from $\Sigma_{01}^{-1} = M^{-1}(I - \rho B)$, where $B = ((b_{jk}))_{j,k=1}^n$ and M is a diagonal matrix with elements $\epsilon_1^2, \ldots, \epsilon_n^2$. Carlin and Banerjee (2002) argue that negative smoothness parameters are not desirable, so they use $0 < \rho < 1$.

We use proper priors for all hyper-parameters, $\theta = (\beta_0, \beta_1, \sigma_e^2, \rho, \sigma_c^2, \psi, \sigma_2^2)$, as outlined below:

$$\begin{aligned}
\beta_0 &\sim \text{Normal}(0, 10000) \\
\beta_1 &\sim \text{Normal}(0, 10000) \\
\sigma_e^2 &\sim \text{Inverse Gamma}(0.001, 0.001) \\
\rho &\sim \text{Uniform}(0, 1) \\
\sigma_c^2 &\sim \text{Inverse Gamma}(.01, .01) \\
\psi &\sim \text{Uniform}(-0.1, 2) \\
\sigma_2^2 &\sim \text{Inverse Gamma}(0.001, 0.001).
\end{aligned}$$

The six modeling scenarios considered here are summarized in Table 1. Model 1 recognizes no spatial patterns in small-scale variability, neither in the baseline process nor due to the point source; Model 1 uses ordinary least squares estimation. Models 2, 4, and 6 all include point source covariance models. For Model 2, no spatial proximity modeling is used for the baseline process. Models 3 and 4 both fit the singular CAR model, while Models 5 and 6 fit the more general CAR model.

[Table 1 about here.]

Posterior distributions are obtained using MCMC in software WinBUGS Version 1.3 (Spiegelhalter, Thomas, and Best 2000). This version of WinBUGS provides a function for CAR(ρ, σ_c^2), but only for deterministic ρ. We hard-coded the cases where ρ was allowed to be random, namely, Models 5 and 6. WinBUGS offers many useful convergence measures that can be conveniently summarized using the CODA module (Best, Cowles, and Vines 1996) running under R or SPlus.

Using this Bayesian hierarchical approach, we implement an MCMC method to sample from the full conditional distributions of all these parameters. While all hyperpriors are proper, Models 2, 4, and 6 contain the singular ARPS and Models 3 and 4 contain the singular CAR($1, \sigma_c^2$). Presence of the measurement error (ME) component, however, ensures that all posterior distributions are proper. Concerns raised by Sun, Tsutakawa, Speckman (1999), Sun et al. (2000), and Sun, Tsutakawa, He (2001) are not applicable here. In Table 2, we report summaries of the posterior densities for each of the six models, and in Table 3 we report model diagnostics. Results are based on 100,000 iterations after a burn-in period of 30,000 iterations. Several checks were made on convergence of the MCMC, including

running three separate chains using different starting values and observing Gelman and Rubin's (1992) diagnostic measure R; all values were very close to unity, indicating good mixing of the chains.

For all models, the $2.5th$ posterior percentile of β_1 exceeds 0.68, thus suggesting that electromagnetism increases at sites further from the pole. Posterior regions for β_0, β_1, and σ_e^2 overlap in significant ways across all six models. The posterior mean of σ_e^2 under Models 1 and 2 are almost double the values under other models, thus suggesting that the residuals of Models 1 and 2 contain spatial variation that is later captured by the CAR process. The fitted ARPS process is most impressive when the CAR process is not modeled, as expected. In the presence of the strong spatial structure created by a degenerate CAR$(1, \sigma_c^2)$, ARPS parameters are closer to their null values, with posterior means of 1.147 for ψ (null value is 0) and 0.5262 for η (null value is 1) in Model 4. When strength of spatial dependence is estimated using CAR(ρ, σ_c^2) with random ρ, as in Model 6, posterior means are 1.332 for ψ, 0.3801 for η, and 0.4643 for ρ.

[Table 2 about here.]

[Table 3 about here.]

5.2 Model Comparison

Many techniques have been suggested for choosing among competing models. Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC), and Bayes factors are among the most often used, but these methods fail in our models because of the random effects (Spiegelhalter, Best, Carlin, and van der Linde 2002). Bayes factors are not interpretable with a CAR prior for spatial effects (Han and Carlin 2001). This CAR prior creates n additional parameters that cannot be counted as n free parameters, thus making AIC and

BIC inapplicable. For our model comparisons, we use the Deviance Information Criterion (DIC) proposed by Spiegelhalter et al. (2002). Sum of squared errors (SSE) is used as a secondary model comparison criterion. SSE due to prediction is computed as

$$\text{SSE} = \text{E}\Big\{\sum_{j=1}^{n}[Y(s_j) - \widehat{Y}(s_j)]^2 | Y(\cdot)\Big\},$$

where the expectation is taken with respect to the posterior predictive distribution of $\widehat{Y}(\cdot)$. See Gelfand and Ghosh (1998) for a decision-theoretic justification of this quantity as a model-choice criterion.

DIC, like AIC and BIC, balances model adequacy against model complexity. Model adequacy is measured by the posterior mean $\overline{D} = E_{\theta|y}[D(\theta)]$ of the deviance $D(\theta)$, where $\theta = (\beta_0, \beta_1, \sigma_e^2, \rho, \sigma_c^2, \psi, \sigma_2^2)$ implies the level of hierarchy being considered, $D(\theta) = -2\log f(y|\theta) + 2\log[h(y)]$, and $h(y)$ is unaffected by the model. Model complexity is measured by the effective numbers of parameters, $p_D = \overline{D} - D(\overline{\theta})$, where $D(\overline{\theta})$ is the deviance evaluated at the posterior mean, $\overline{\theta}$, of the parameters. Consequently, DIC $= \overline{D} + p_D$ and small values are desirable. DIC is effective in high-dimensional or very complicated models because of the method used for determining the effective number of parameters.

In Table 3, we report DIC, and its components, for each of the six models. Multiple runs yield differences of at most 1.9 in DIC. We also report p_θ, the total number of parameters, and the posterior mean of SSE for all six models. Models 1 and 2 are clearly inadequate. Model 4 is better than Models 1 and 2 but is not as good as Models 3, 5, and 6. Model 6 is the best model by all measures.

Model 6 has many interesting features. While the CAR process enforces unequal variances only to address edge effects caused by differing numbers of neighbors, the ARPS process explicitly fits unequal variances as suggested by the data in Figure 2. Figure 3(a)

shows fitted variances (obtained as the diagonal elements of Σ evaluated at the posterior mean, $\overline{\theta}$, of the parameters) for Models 3, 5, and 6. The curve for Model 6 is much more variable than the curves for Models 3 and 5. Figure 3(b) shows fitted variances only for Model 5, and it is clear that these fitted variances are not consistent with the patterns in the data, as displayed in Figure 2. Corner sites (site numbers 1, 20, 141, and 160) have largest variances, followed by sites adjacent to corners, followed by other edge sites, and finally followed by interior sites. Figure 3(c) shows fitted variances only for Model 3, and, again, these fitted variances are clearly not consistent with the data. In both Models 3 and 5, fitted variances are a function of number of neighbors and they increase away from the pole. There are much stronger spatial patterns and smoothness in Model 3 ($\rho = .99$ and $\sigma_c^2 = .0018$ in CAR) than in Model 5 ($\rho = .85$ and $\sigma_c^2 = .0022$ in CAR).

[Figure 3 about here.]

A second interesting feature of Model 6 is the relationship between correlation and site-to-source distances. Figure 4 shows correlation clouds from Σ evaluated at $\overline{\theta}$ as a function of site-to-site and site-to-source distances for Models 3, 5, and 6. Figure 4(a) shows that for Model 5 correlations between sites closest to the pole (ARPS region R_{10}) are in general smaller than correlations between sites farthest from the pole (ARPS region R_1). There is, however, no clear separation between these groups of correlations. For Model 3, Figure 4(b) shows a greater degree of separation of correlations according to distance from pole, but the separation is still small. On the other hand, Model 6, as shown in Figure 4(c), clearly has correlations that are a function of distance to pole. Moreover, correlations between sites in R_{10} are largest.

[Figure 4 about here.]

Third, focusing only on the ARPS process, we consider the one-step correlations from

regions R_1 to R_{10}, defined as

$$\begin{aligned} \text{Corr}(\alpha_1, \alpha_2) &= \psi\sqrt{\frac{\sigma_1^2}{\sigma_1^2\psi^2 + \sigma_2^2}}, \\ \text{Corr}(\alpha_k, \alpha_{k+1}) &= \psi\sqrt{\frac{\sigma_1^2\psi^{2(k-1)} + \sigma_2^2\sum_{q=0}^{k-2}\psi^{2q}}{\sigma_1^2\psi^{2k} + \sigma_2^2\sum_{q=0}^{k-1}\psi^{2q}}} \qquad \text{for } k = 2, 3, \ldots, 9. \end{aligned}$$

Evaluating at $\overline{\theta}$ for Model 6, we obtain $\text{Corr}(\alpha_k, \alpha_{k+1})$, $k = 1, 2, \ldots, 9$ as: 0.790, .908, .954, .975, .987, .993, .996, .998, .999. These numbers change quickly from $k = 1$ to $k = 3$ and are primarily responsible for the differences seen in Figure 4(c).

Finally, as evidenced by the posterior density intervals for ψ and η given in Table 2, as well as additional results not presented here, there is clear evidence of statistical significance of the ARPS model. More specifically, the posterior interval for ψ does not contain the value 0 and the posterior interval for η does not contain the value 1.

6 CONCLUDING REMARKS

Our newly proposed ARPS process can improve models of stochastic phenomena across spatial domains because it takes advantage of auxiliary information provided by point sources. The result is more simplified, interpretable, parametric, yet realistic nonstationary modeling of error processes. By restricting the initial variance σ_1^2 of the ARPS process to equal the measurement error variance σ_e^2, we strengthen the interpretation that ARPS captures extra variation due to the point source, and sites far from this source are well represented by the baseline error process. Additionally, we comment on the interplay between the ARPS process, the commonly used CAR$(1, \sigma_c^2)$ process, and the more desirable CAR(ρ, σ_c^2) process. This is all done in a Bayesian hierarchical framework where key objectives include simple tests for, and clear interpretations of, the impact of a point source.

A complete analysis of electromagnetism in a field containing a metal pole clearly shows

the advantages of our approach. With ARPS, we capture the observed trend of decreasing variances away from the pole; this finding is consistent with Hughes-Oliver and Gonzalez-Farias (1999). We also achieve greater flexibility in modeling correlations within and across regions of various distances from the pole. Posterior density regions for parameters ψ and η clearly do not contain null values, thus implying statistical significance of the pole's impact on small-scale variability of the electromagnetism process. Model selection is based on both the DIC and SSE.

There are, however, several issues that are currently being investigated. One such issue deals with dynamic regionalization for the ARPS process, where we observe a spatio-temporal process that can possibly alter the effect of a source over time. This is particularly relevant in processes affect by meteorological conditions. Another issue under current investigation is simultaneously accounting for the impact of multiple sources. Ranking the impact of different sources is particularly difficult when sites are affected by multiple sources in very different ways. The field of source apportionment (see, for example, Henry 1997; Henry, Spiegelman, Collins, and Park 1997; Park, Guttorp, and Henry 2001; Christensen and Sain 2002; and Park, Spiegelman, and Henry 2002) ranks the impact of multiple sources by effectively partitioning the observed variance matrix into factors that represent different categories of sources. Bayesian hierarchical modeling of multiple ARPS processes is a fundamentally different approach from the standard techniques of source apportionment. This will be the subject of future manuscripts.

REFERENCES

Best, N., Cowles, M.K., and Vines, K. (1996), *Convergence Diagnostics and Output Analysis Software for Gibbs Sampling Output*, Cambridge: MRC Biostatistics Unit, Institute of Public Health.

Bithell, J. F. (1995), "The Choice of Test for Detecting Raised Disease Risk Near a Point Source," *Statistics in Medicine*, 14, 2309–2322.

Carlin, B. P., and Banerjee, S. (2002), "Hierarchical Multivarite CAR Models for Spatio-Temporally Correlated Survival Data," (with discussion) in *Bayesian Statistics 7*, eds. J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West, Oxford: Oxford University Press.

Christensen, W. F., and Sain, S. R. (2002), "Accounting for Dependence in a Flexible Multivariate Receptor Model," *Technometrics*, 44, 328–337.

Cressie, N. A. C., and Majure, J. J. (1997), "Spatio-Temporal Statistical Modeling of Livestock Waste in Streams," *Journal of Agricultural, Biological, and Environmental Statistics*, 2, 24–47.

Diggle, P., Morris, S., Elliot, P., and Shaddick, G. (1997), "Regression Modelling of Disease Risk in Relation to Point Sources," *Journal of the Royal Statistical Society Series A–Statistics in Society*, 160, 491–505.

Diggle, P. J., and Rowlingson, B. S. (1994), "A Conditional Approach to Point Process Modeling of Elevated Risk," *Journal of the Royal Statistical Society Series A–Statistics in Society*, 157, 433–440.

Fuentes, M. (2002), "Spectral Methods for Nonstationary Spatial Processes," *Biometrika*, 89, 197–210.

Gelfand, A. E., and Ghosh, S. K. (1998), "Model Choice: A Minimum Posterior Predictive Loss Approach," *Biometrika*, 85, 1–11.

Gelman, A., and Rubin, D. B. (1992), "Inference From Iterative Simulation Using Multiple Sequences" (with discussion), *Statistical Science*, 7, 457–511.

Haas, T. C. (1995), "Local Predictions of a Spatio-Temporal Process With an Application to Wet Sulphate Deposition," *Journal of the American Statistical Association*, 90, 1189–1199.

Han, C., and Carlin, B. (2001), "MCMC Methods for Computing Bayes Factors: a Comparative Review. *Journal of the American Statistical Association*, 96, 1122–1132.

Hastie, T. J., and Tibshirani, R. J. (1990), *Generalized Additive Models*, New York: Chapman and Hall.

Henry, R. C. (1997), "History and Fundamentals of Multivariate Air Duality Receptor Models," *Chemometrics and Intelligent Laboratory Systems*, 37, 37–42.

Henry, R. C., Speigelman, C. H., Collins, J. F., and Park, E. S. (1997), "Reported Emissions of Organic Gases are not Consistent With Observations," *Proceedings of the National Academy of Sciences, USA*, 94, 6596–6599.

Higdon, D., Swall, J., and Kern, J. (1999), "Non-Stationary Spatial Modeling," in *Bayesian Statistics 6*, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Oxford: Oxford University Press, 761–768.

Holland, D. M., De Oliveira, V., Cox, L. H., and Smith, R. L. (2000), "Estimation of Regional Trends in Sulfur Dioxide Over the Eastern United States," *Environmetrics*, 11, 373–393.

Hughes-Oliver, J. M., Gonzalez-Farias, G., Lu, J. C., and Chen, D. (1998), "Parametric Nonstationary Correlation Models," *Statistics & Probability Letters*, 40, 267–268.

Hughes-Oliver, J. M., Lu, J. C., Davis, J. C., and Gyurcsik, R. S. (1998), "Achieving Uniformity in a semiconductor Fabrication Process Using Spatial Modeling," *Journal of the American Statistical Association*, 93, 36-45.

Hughes-Oliver, J. M., and Gonzalez-Farias, G. (1999), "Parametric Covariance Models for Shock-Induced Stochastic Processes," *Journal of Statistical Planning and Inference*, 77, 51–72.

Lawson, A. B. (1995), "MCMC Methods for Putative Pollution Source Problems in Environmental Epidemiology," *Statistics in Medicine*, 14, 2473–2485.

Lawson, A. B. (2001), *Statistical Methods in Spatial Epidemiology*, New York: Wiley.

Lawson, A. B., and Waller, L. A. (1996), "A Review of Point Pattern Methods for Spatial Modelling of Events Around Sources of Pollution," *Environmetrics*, 7, 471–487.

Park, E. S., Guttorp, P., and Henry, R. C. (2001), "Multivariate Receptor Modeling for Temporally Correlated Data by Using MCMC," *Journal of the American Statistical Association*, 96, 1171–1183.

Park, E. S., Spiegelman, C. H., and Henry, R. C. (2002), "Bilinear Estimation of Pollution Source Progiles and Amounts by Using Multivariate Receptor Models," *Environmetrics*, 13, 775–809.

Sampson, P. D., and Guttorp, P. (1992), "Nonparametric Estimation of Nonstationary Spatial Covariance Structure," *Journal of the American Statistical Association*, 87, 108-119.

Spiegelhalter, D. J., Thomas, A., and Best, N. G. (2000) *WinBUGS Version 1.3 User Manual*. Cambridge: Medical Research Council Biostatistics Unit. (Available from `http://www.mrc-bsu.cam.ac.uk/bugs`.)

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002), "Bayesian Measures of Model Complexity and Fit." *Journal of the Royal Statistical Society, Series B*, 64(4), 583–639.

Sun, D., Tsutakawa, R. K., Kim, H., and He, Z.Q. (2000), "Spatio-Temporal Interaction With Disease Mapping," *Statistics in Medicine*, 19, 2015–2035.

Sun, D., Tsutakawa, R. K., and He, Z.Q. (2001), "Propriety of Posteriors With Improper Priors in Hierarchical Linear Mixed Models," *Statistica Sinica*, 11, 77–95.

Sun, D., Tsutakawa, R. K., and Speckman, P. L.(1999), "Posterior Distribution of Hierarchical Models Using CAR(1) Distributions," *Biometrika*, 86, 341-350

Wikle, C.K. (in press) "Hierarchical Models in Environmental Science," *International Statistical Review.*

List of Figures

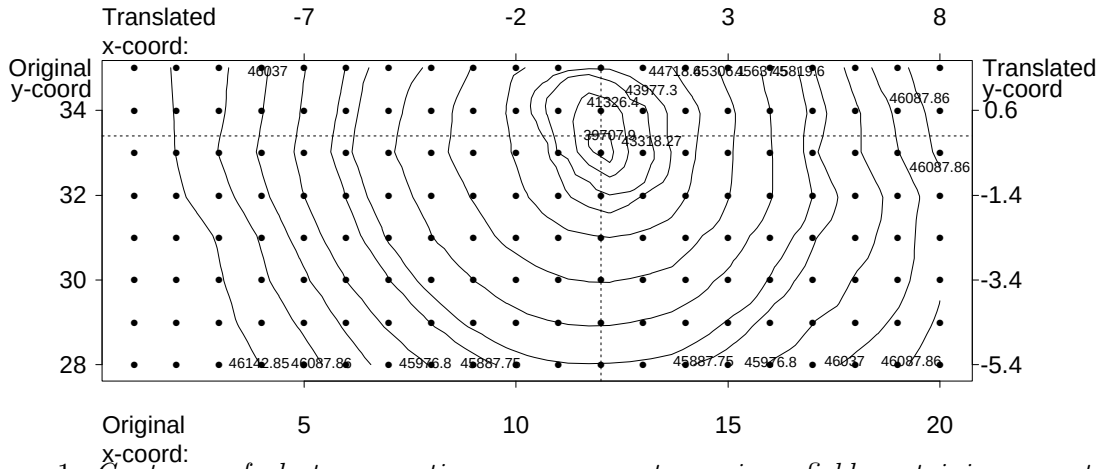

Figure 1: *Contours of electromagnetism measurements, g, in a field containing a metal pole. The measurement sites* ($\bullet$) *fall on a regular grid, with spacings of one meter in both directions. The metal pole is located at* $(12, 33.4)$ *in the original coordinate system, and at* $(0, 0)$ *in the translated coordinate system. The contours represent the following sample percentiles:* .7, 1, 2.5, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90.

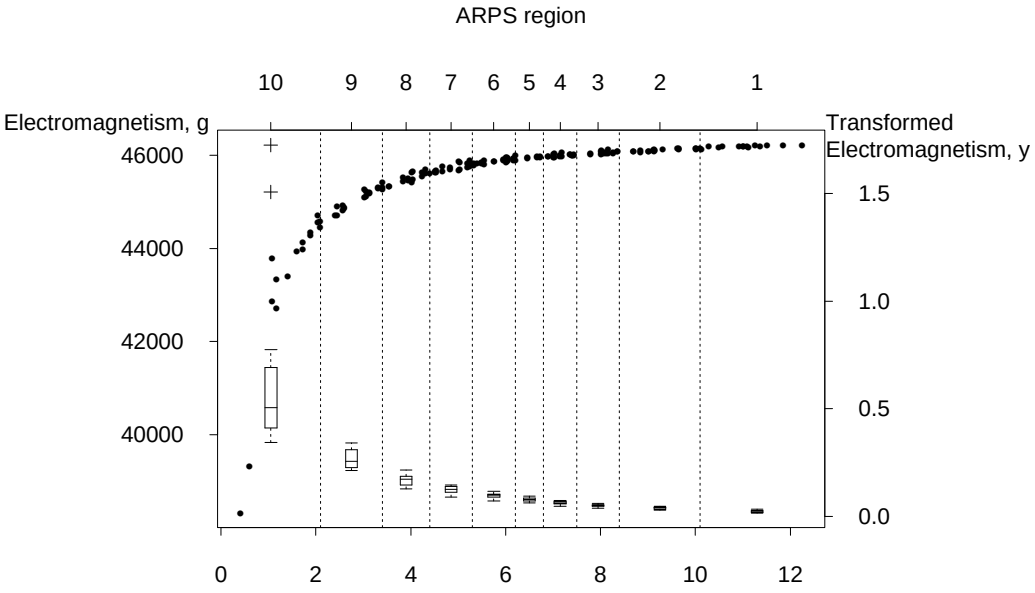

Figure 2: *Electromagnetism (plotted as bullets) as a function of distance between measurement sites and the metal pole. Locations of ARPS regions $R_1, \ldots, R_{10}$ and boxplots of transformed electromagnetism within these regions are also shown.*

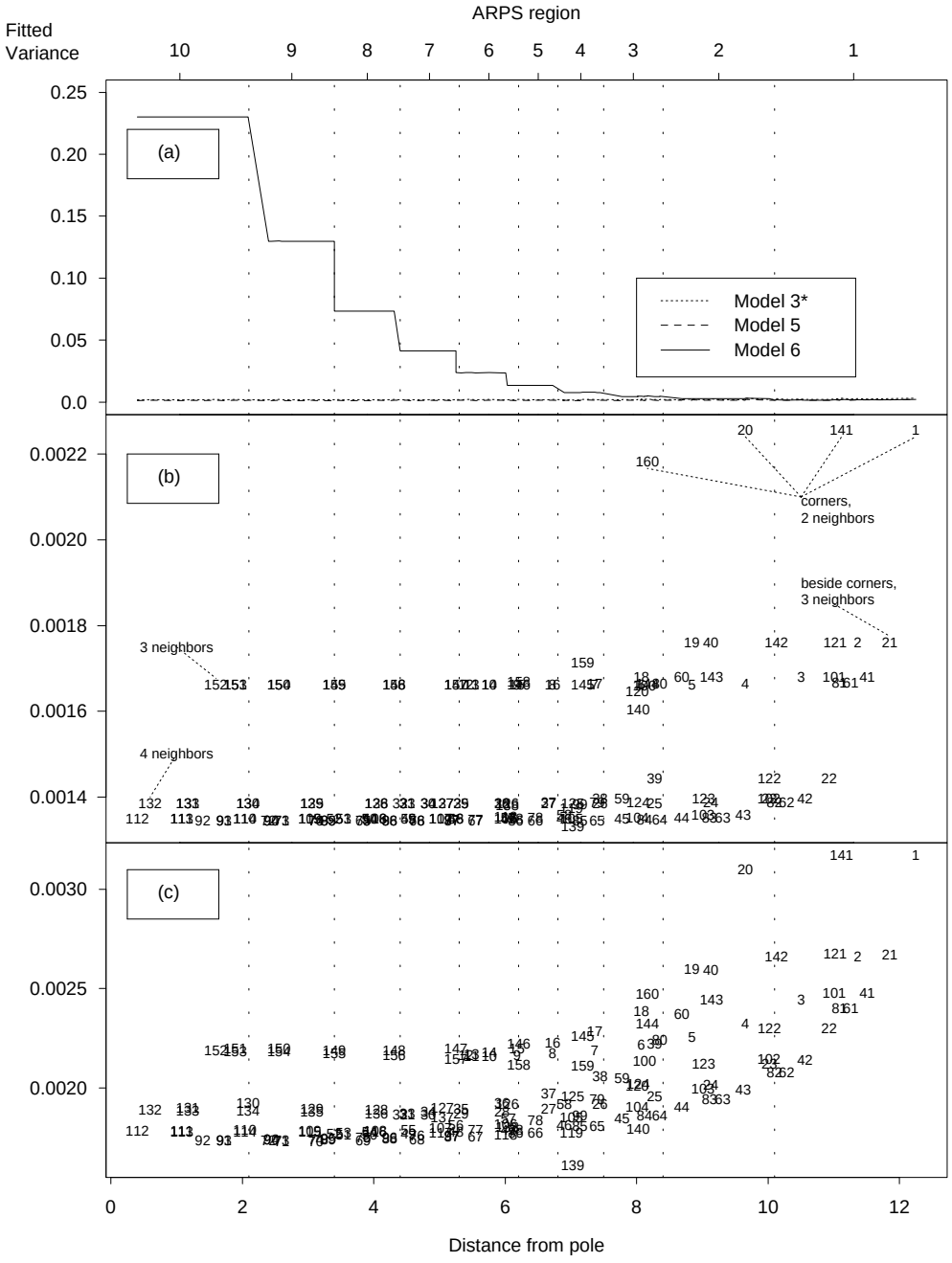

Figure 3: *Fitted variances from Models 3 (modified), 5, and 6. A modification is necessary for Model 3 because this fitted model is singular. For this model, we replace* $\rho = 1$ *with* $\rho = 0.99$ *to obtain invertible matrices that yield variances. Panel (a) shows variances for all models, while panel (b) shows variances only for Model 5 and panel (c) shows variances only for modified Model 3. Observation sites numbers are used as plotting symbols in panels (b) and (c), where site 1 is in the bottom left corner of Figure 1 and site 160 is in the top right corner.*

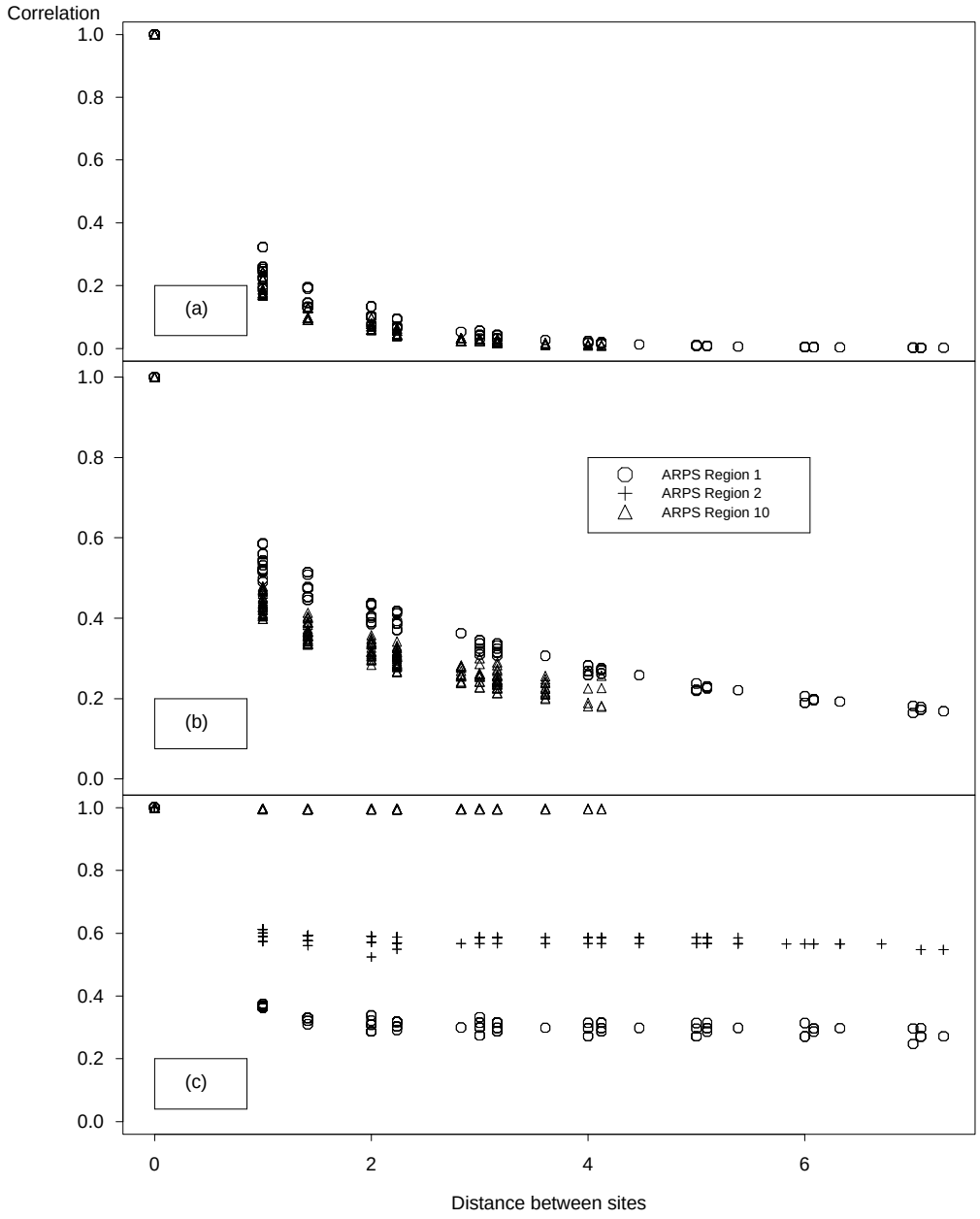

Figure 4: *Fitted correlations from Models 3 (modified), 5, and 6. A modification is necessary for Model 3 because this fitted model is singular. For this model, we replace* $\rho = 1$ *with* $\rho = 0.99$ *to obtain an invertible matrix that yields correlations. Panels (a), (b), and (c) show correlations for Models 5, 3, and 6, respectively. All panels display correlations between all pairs of sites where both sites are either in ARPS Region* R_1 *or* R_{10}*. Panel (c) also displays correlations between all pairs of sites where both sites are in ARPS Region* R_2*.*

List of Tables

Table 1: Modeling Cases for Electromagnetism in a Field

Model*	CAR$(1, \sigma_c^2)$	CAR(ρ, σ_c^2)	ARPS$(\psi, \sigma_e^2, \sigma_2^2)$
1, trend+ME			
2, trend+ME+ARPS			x
3, trend+ME+CAR$_1$	x		
4, trend+ME+CAR$_1$+ARPS	x		x
5, trend+ME+CAR$_\rho$		x	
6, trend+ME+CAR$_\rho$+ARPS		x	x

Table 2: Summaries of Posterior Densities from Modeling Electromagnetism

Parameter	Mean	MCerror	2.5%	Median	97.5%
Model 1					
β_0	-0.0400	1.397E-5	-0.0477	-0.0400	-0.0324
β_1	0.8059	3.904E-5	0.7848	0.8059	0.8271
σ_e^2	0.001326	3.143E-7	0.001062	0.001315	0.001653
Model 2					
β_0	-0.0211	2.796E-5	-0.0299	-0.0211	-0.0123
β_1	0.7281	1.036E-4	0.6985	0.7281	0.7579
σ_e^2	0.001011	3.097E-7	0.000806	0.001003	0.001266
ψ	1.330	3.213E-3	0.5619	1.3430	1.9280
η	0.4737	1.955E-3	0.2265	0.4259	1.0580
σ_2^2	0.000586	1.710E-6	0.000179	0.000472	0.001671
Model 3					
β_0	-0.0224	4.356E-5	-0.0312	-0.0224	-0.0137
β_1	0.7339	1.712E-4	0.7012	0.7340	0.7660
σ_e^2	0.000599	1.211E-6	0.000330	0.000593	0.000906
σ_c^2	0.001775	3.568E-6	0.001051	0.001725	0.002776
Model 4					
β_0	-0.0191	4.190E-5	-0.0282	-0.0191	0.0100
β_1	0.7201	1.637E-4	0.6866	0.7202	0.7534
σ_e^2	0.000665	1.173E-6	0.000392	0.000661	0.000959
σ_c^2	0.001454	3.927E-6	0.000815	0.001398	0.002420
ψ	1.1470	4.478E-3	0.2223	1.182	1.866
η	0.5262	3.107E-3	0.2037	0.4479	1.384
σ_2^2	0.000588	1.498E-6	0.000180	0.000476	0.001671
Model 5					
β_0	-0.0292	1.067E-4	-0.0433	-0.0295	-0.0127
β_1	0.7633	2.263E-4	0.7267	0.7632	0.8001
σ_e^2	0.000596	1.779E-6	0.000279	0.000588	0.000962
ρ	0.8527	8.813E-4	0.5397	0.8817	0.9934
σ_c^2	0.002221	6.378E-6	0.001143	0.002147	0.003705
Model 6					
β_0	-0.0204	5.191E-5	-0.0307	-0.0204	-0.0098
β_1	0.7252	1.719E-4	0.6928	0.7253	0.7574
σ_e^2	0.000554	1.527E-6	0.000265	0.000549	0.000878
ρ	0.4643	1.339E-3	0.0330	0.4630	0.9254
σ_c^2	0.002090	5.583E-6	0.001116	0.002031	0.003378
ψ	1.3320	2.858E-3	0.5457	1.3470	1.9250
η	0.3801	1.299E-3	0.1706	0.3494	0.7814
σ_2^2	0.000592	1.645E-6	0.000180	0.000477	0.001690

Table 3: Diagnostic Measures from Modeling Electromagnetism

	Criteria					
Model	DIC(Rank)	p_D	p_θ	$D(\overline{\theta})$	$\overline{D}$	SSE(Rank)
1, trend+ME	-606(6)	3	3	-612	-609	0.208(6)
2, trend+ME+ARPS	-642(5)	10	15	-662	-652	0.158(5)
3, trend+ME+CAR_1	-660(2)	82	164	-824	-742	0.093(3)
4, trend+ME+CAR_1+ARPS	-648(4)	74	176	-796	-722	0.103(4)
5, trend+ME+CAR_ρ	-656(3)	89	165	-834	-745	0.092(2)
6, trend+ME+CAR_ρ+ARPS	-664(1)	90	177	-845	-755	0.086(1)