

**ESTIMATION OF REGIONAL TRENDS IN SULFUR DIOXIDE
OVER THE EASTERN UNITED STATES**

David M. Holland¹, Victor De Oliveira², Lawrence H. Cox¹, and
Richard L. Smith³

*¹U.S. Environmental Protection Agency
Office of Research and Development
Research Triangle Park, NC 27711-0001*

*²Department of Scientific Computing and Statistics
Simón Bolívar University
Caracas, Venezuela*

*³University of North Carolina
Department of Statistics
Chapel Hill, NC 27599-3260*

ESTIMATION OF REGIONAL TRENDS IN SULFUR DIOXIDE OVER THE EASTERN UNITED STATES

Abstract. Emission reductions were mandated in the Clean Air Act Amendments of 1990 with the expectation of concomitant reductions in ambient concentrations of atmospherically-transported pollutants. To evaluate the effectiveness of the legislated emission reductions using monitoring data, this paper proposes a two-stage approach for the estimation of regional trends and their standard errors. In the first stage, a generalized additive model (*GAM*) is fitted to airborne sulfur dioxide (SO_2) data at each of 35 sites in the eastern United States to estimate the form and magnitude of the site-specific trend (defined as *percent total change*) from 1989 to 1995. This analysis is designed to adjust the SO_2 data for the influences of meteorology and season. In the second stage, the estimated trends are treated as samples with site-dependent measurement error from a Gaussian random field with a stationary covariance function. Kriging methodology is adapted to construct spatially-smoothed estimates of the true trend for three large regions in the eastern U.S. Finally, a Bayesian analysis with Markov Chain Monte Carlo (MCMC) methods is used to obtain regional trend estimates and their standard errors, which take account of the estimation of the unknown covariance parameters as well as the stochastic variation of the random fields. Both spatial estimation techniques produced similar results in terms of regional trend and standard error.

Key word index: regional trend, generalized additive models, kriging, Markov Chain Monte Carlo

1. INTRODUCTION

The implementation of the Clean Air Act (CAA), from its passage in 1970 to the 1990 amendments, has always required an assessment of the effects of atmospherically-transported pollutants on the environment. The 1990 amendments included new requirements that will appreciably reduce sulfur dioxide (SO₂) emissions. The total emission release expected in 2010 is 8.95 million tons per year, a reduction of 10 million tons per year from the amount projected to be released without controls. These reductions are intended to reduce public health risks and to protect sensitive ecosystems. Ground-level concentrations of SO₂ depend on the proximity to source(s), prevailing meteorology, and source-receptor relationships. The chemical and physical interactions of atmospheric processes and SO₂ emissions produce data patterns that show large spatial and temporal variability (U.S. EPA 1998a), making it difficult to evaluate the extent to which the emission control programs achieved the goal of reduced atmospheric concentrations.

The estimation of trends in airborne concentrations has been the subject of many investigations since the implementation of national monitoring networks in the late 1970's. Shreffler and Barnes (1996) used a linear model to estimate trends (percent change per year) in daily atmospheric concentrations of SO₂ adjusted for meteorological factors, ozone, and seasonal cycles at six locations in the northeastern U.S. They estimated the median reduction from 1977 to 1989 to be 46%, which is not commensurate with an estimated 25% decline in SO₂ emissions over this period. This result may reflect changes in nearby sources of SO₂. Lefohn and Shadwick (1991) estimated trends in ozone, SO₂ and nitrogen dioxide at rural sites in the U.S. using non-parametric methods. Flaum *et al.* (1996) estimated trends in hourly ozone concentrations after

removing the influences of temperature and other meteorological variables at eight sites in the eastern U.S. They found that adjusting for meteorology was critically important for evaluating the impact of emission controls on ambient ozone levels. Bloomfield *et al.* (1993) studied ozone concentrations and meteorology in rural areas surrounding Chicago, Illinois. They used a parametric non-linear model to describe the non-linear and non-additive dependence of network-typical ozone values on meteorological variables. All of these studies focused on developing models either for site-specific trends or models for trend in a summary statistic that represents a network typical value. They consider neither trend estimation at un-monitored sites nor estimation of regional trends, which requires a model for the spatial variability of site-specific trends.

In recent years, the focus of environmental policy has shifted toward regional-scale strategies which require regional estimates of trend for both their development and subsequent evaluation. In an effort to provide meaningful regional trend information, this paper describes a two-stage modeling approach to estimate regional trends in airborne concentrations of SO_2 that have been adjusted for concomitant changes in meteorology and season. Concentrations of SO_2 are strongly affected by prevailing meteorological conditions, but the precise form of this dependence is not clear, so we used a *generalized additive model (GAM)* to describe the dependence of SO_2 on meteorological conditions, seasonality, and time at each monitoring site. The first stage of this analysis focuses on using the time component of the *GAM* to estimate site-specific trend where trend is defined as a *percent total change* that is a measure of the total change in SO_2 concentrations that have been adjusted for the effects of meteorology and seasonal cycles.

The second stage of analysis uses an extension of kriging methodology to predict smoothed surfaces of trend based on the site-specific estimates from the first stage. For this, trends are assumed to vary over a geographic region of interest as a realization of a Gaussian random field, and maximum likelihood estimation is used to fit the model. Then, averages of trend (defined as regional trend) over geographic regions of interest can be estimated. The analysis uses a form of hierarchical model, and differs from traditional kriging methodology by making explicit allowance for the sampling variability in the estimates from the first stage. This analysis is applied to airborne SO_2 concentration (ug/m^3) data measured at 35 rural long-term monitoring sites in the eastern U.S. that are part of the Clean Air Act Status and Trends Monitoring Network (CASTNet) (U.S. EPA 1998a). Since kriging variances do not account for the estimation of unknown parameters in the covariance function, Bayesian techniques are also used to estimate prediction variances. These two approaches are compared to quantify the effect of ignoring the uncertainty of the covariance parameters on inference about regional trends. Finally, estimates of regional trend in SO_2 are compared to corresponding changes in SO_2 emissions for three large regions in the eastern U.S. to evaluate the impact of reduced emissions on measured SO_2 concentrations.

This paper is organized as follows. In Section 2, we summarize the monitoring data, Section 3 describes the estimation of sites-specific trends and prediction of a trend surface using kriging, and Section 4 presents the kriging and Bayesian approaches for estimating regional trends and their standard errors. Finally, results of the analysis are given in Section 5 and Section 6 provides the conclusions.

2. DATA

CASTNet began measuring air and deposition variables in 1987 at a variety of sites within the conterminous U.S., and, by 1989 most of the sites were operational. Monitoring locations were selected according to strict siting criteria designed to avoid undue influence from point and area emission sources, and local activities (*e.g.* agriculture). Continuous measurements of meteorological variables including temperature ($^{\circ}\text{C}$), wind speed ($m\ s^{-1}$), and wind direction (*degrees clockwise from north*) were summarized hourly at each site, and weekly measurements of SO_2 concentrations were obtained from filter pack measurements. It was necessary to summarize hourly meteorological data on the same scale as the SO_2 measurements, *viz.* weekly. These meteorological summaries were calculated by averaging all hourly meteorological variables between 10 *am* and 5 *pm* across the week to characterize conditions during periods of atmospheric mixing, and reflect regional, large-scale flows of air pollution across areas covered by CASTNet. To be included in this analysis, a site's data record (January 1, 1989 to September 30, 1995) must include at least 80% of all days with concurrent SO_2 and meteorological data. Adherence to this data-completeness criterion produced 35 long-term monitoring sites (Holland *et al.* 1999) suitable for trend analysis in the eastern U.S. The accuracy and precision of CASTNet monitoring data are detailed by Clarke *et al.* (1997). Time-series plots of the data observed at three sites are shown in Figure 1. The two sites located in close proximity to major SO_2 emissions point sources (Alhambra, IL and Lykens, OH) , show strong evidence of seasonal cycles in the data. Typically SO_2 is higher in the winter compared to the summer, when SO_2 is converted to other atmospheric pollutants. The remote, background site in Ashland, ME shows

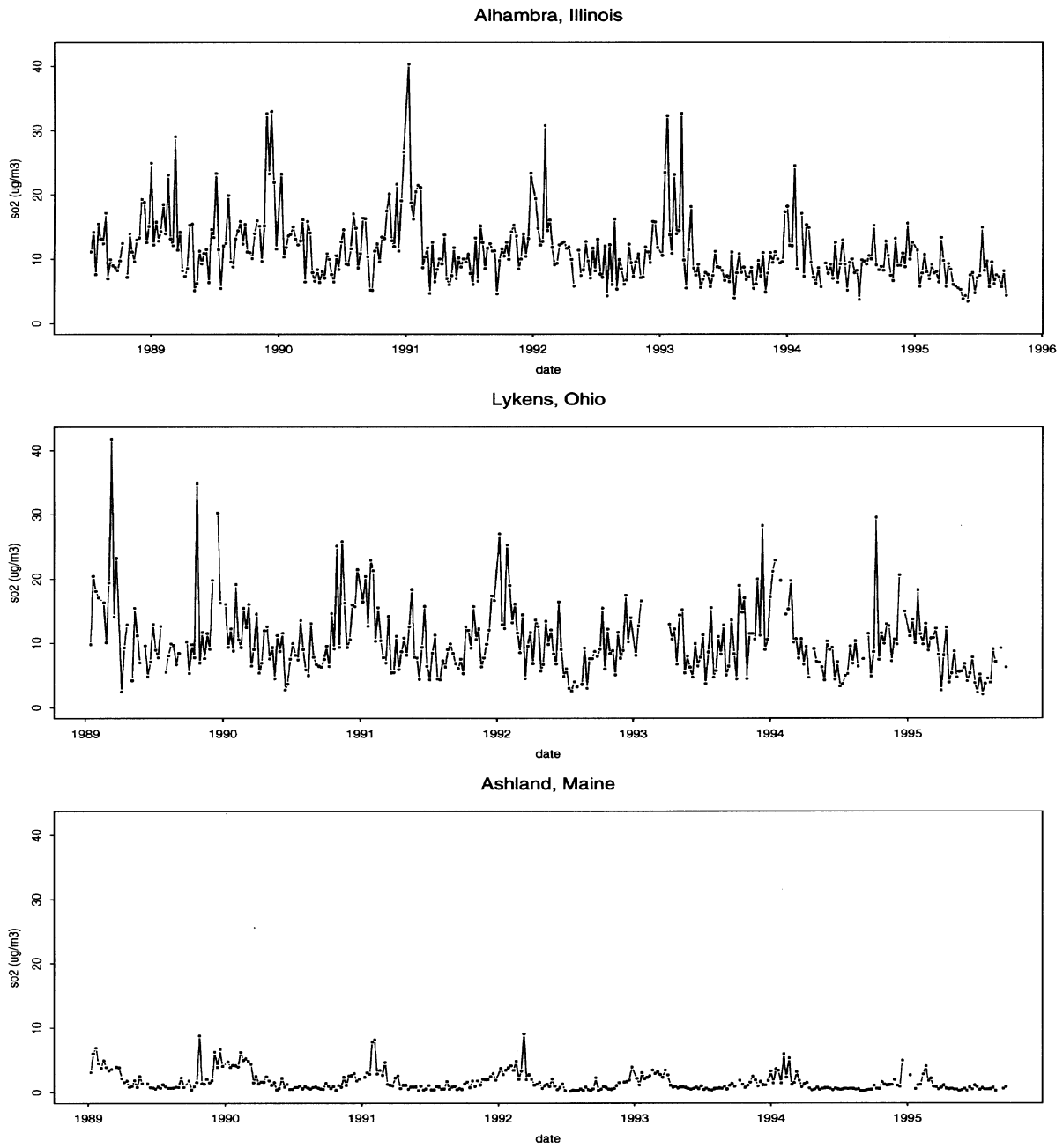


Figure 1. Time-series plot of SO₂ concentrations ($\mu\text{g}/\text{m}^3$) at three CASTNet sites.

much lower SO₂ concentrations compared to the other two sites, but seasonal cycles are still

evident.

3. TWO-STAGE MODEL FOR REGIONAL TREND

3.1 First Stage

The first-stage of analysis uses *GAM* to model the relationship between the logarithm of SO₂ concentrations and prevailing meteorological conditions, seasonal effects, and time. The *GAM* approach offers an adaptive method for regression modeling using non-parametric scatterplot smoothers. Rather than requiring *a priori* specification of the model form, this approach lets the data suggest the form of the model, which results in a highly flexible technique for describing non-linear relationships. Several smooth terms can be fit simultaneously through iterative use of scatterplot smoothers, assuming that the underlying functions are reasonably smooth. Both smoothing spline functions and local regression models (usually referred to as a *loess* fit) can be used to model these relationships in the *S-PLUS*¹ programming environment. Hastie and Tibshirani (1990) present a detailed account of the theory and application of *GAM*. Davis and Speckman (1999) used *GAM* to forecast maximum ozone levels as a function of meteorology in the Houston area and Holland *et al.* (1999) used *GAM* to estimate trend in SO₂ concentrations that were adjusted for the effects of meteorology and seasonality. The model presented by Holland *et al.* (1999) is summarized below and is used to obtain the first-stage

¹*S-PLUS* is the commercial version of the S language and is distributed by Statistical Sciences, Inc., a division of Mathsoft.

estimates of trend and variance at each of the 35 CASTNet monitoring sites.

The relationship between logarithmically-transformed weekly concentrations of SO₂ and meteorology, seasonal cycles, and time (trend) was examined at each site s_l , $l=1,\dots,35$, using a *GAM* of the form,

$$\begin{aligned} \log(SO_{2(ijl)}) = & \mu_l + g_{1(l)}(week_i) + g_{2(l)}(year_{ij}) + g_{3(l)}(temperature_{ijl}) \\ & + g_{4(l)}(u_{ijl}, v_{ijl}) + \varepsilon_{ijl}, \end{aligned} \quad (1)$$

where $SO_{2(ijl)}$ refers to the measured SO₂ concentration in the i^{th} week of the j^{th} year at the l^{th} site, $g_{k(l)}(\cdot)$, $k=1,\dots,4$, are smooth functions of the corresponding covariates, and $\varepsilon_{ijl} \sim N(0, \sigma_l^2)$ are independent random variables. Linearity ($g_{k(l)}(x) = \beta_{k(l)}x$) remains a special case. The variable ‘year’ is time measured in decimal years (Julian date of the mid-point of the sampling period divided by 365) starting in 1989, temperature is average weekly temperature (°C), u is the average east-west component of wind (calculated as $-\{\text{wind speed}\} \times \text{sine}(\text{wind direction})$), and v is the average north-south component of wind (calculated as $-\{\text{wind speed}\} \times \text{cosine}(\text{wind direction})$). For each site, the model was fit by solving a system of equations using the Gauss-Seidel iterative method, also known as *backfitting* (Hastie and Tibshirani 1990), obtaining estimates of $\hat{\mu}_l$ and $\hat{g}_{k(l)}$. The natural logarithm of SO₂, instead of untransformed SO₂, was modeled as an additive function of the explanatory variables because this transformation produced residuals that were in closer agreement with the distributional properties of the errors in model (1). The inclusion or exclusion of each variable in the model and the type of smoothing method used were determined via a step-wise search implemented in *S-PLUS*. In this procedure, an ordered regimen of candidate models is evaluated in terms of the Akaike Information Criterion (AIC) (see Akaike 1973). When this criterion cannot be decreased by choice of an alternative model, the selection process terminates.

To provide an evaluation of trend consistent with EPA reporting of trend in observed SO₂ concentrations, the following approach was used. EPA (U.S. EPA 1998b) reports *percent total change* or trend based on ratios of annual averages of SO₂ concentrations. For two annual periods (m, n), trend is defined as $100(\text{mean}_n/\text{mean}_m - 1)$ for $m < n$. Given our goal of estimating trend in SO₂ concentrations adjusted for seasonal and meteorological effects ($\exp(\hat{g}_{2(t)}(\text{year}_{ij}))$), we use the following estimate for trend based on the EPA approach: $100\{E[\exp(\hat{g}_{2(t)}(\text{year}_{in}))] / E[\exp(\hat{g}_{2(t)}(\text{year}_{im}))] - 1\}$. For each site location s_l , we assume $\hat{g}_{2(t)}(\text{year}_{ij})$, the estimated effect due to time, has a normal distribution with parameters μ_{jl} and σ_{jl}^2 . Then, $E[\exp(\hat{g}_{2(t)}(\text{year}_{ij}))] = \exp(\mu_{jl} + \sigma_{jl}^2/2)$, and hence site-specific trend ($Z(s_l)$) may be defined as

$$Z(s_l) = 100 [e^{\mu_{nl} - \mu_{ml}} e^{(\sigma_{nl}^2 - \sigma_{ml}^2)/2} - 1]. \quad (2)$$

Although the resulting trend cannot be directly related to SO₂ emissions *per se*, it seems likely that emission changes would be the dominant effect reflected in the trend.

We also are interested in assessing the statistical significance of trend in SO₂ concentrations. It is difficult to compute the variance of this change directly because the *S-PLUS* implementation of fitting a non-parametric smoothing function for year is limited to returning an approximation to the diagonal term of the variance-covariance matrix of the smoothed predictions (Chambers and Hastie 1992). Further, it is not clear how to extend this approximation to estimate the rest of the variance-covariance matrix. Given this condition, we implement a jackknife procedure to provide estimates of trend and variances. In the jackknife approach, n (where $n=81$ is the number of months in the data record) model-based estimates of $Z(s_l)$ were obtained by deleting one month of SO₂ concentrations at a time, fitting model (1), and replacing

the population quantities with the sample estimates of the parameters in (2). Using these n estimates, quantities known as *pseudo-values* were used to obtain the jackknife estimates of both trend ($\tilde{Z}(s_i)$) and variance ($\tilde{\sigma}_i^2$) at each site (see Efron and Tibshirani 1993). We use these estimates as reasonable measures of site-specific trend and variance, but note that other measures of trend associated with other types of statistical models could also be used as input to the second stage of regional trend analysis to be described later.

3.2 Second Stage

To make inference about the trend at un-monitored sites as well as regional trends, we apply an extension of kriging analysis that allows for the errors of estimation in the first stage of the analysis. We assume that the smooth, unobserved surface of trend varies over the eastern U.S. as a realization of a Gaussian random field $\{Z(s), s \in \mathbb{R}^2\}$, with

$$E\{Z(s)\} = \sum_{j=1}^q \beta_j f_j(s); \quad Cov\{Z(s), Z(u)\} = \alpha K_\theta(\|s-u\|), \quad (3)$$

where β is a q -vector of unknown regression parameters, $f_1(s), \dots, f_q(s)$ are known functions, $\alpha = Var\{Z(s)\}$, $K_\theta(\cdot, \cdot)$ is a correlation function in \mathbb{R}^2 parameterized by $\theta \in \Theta$, and $\|s-u\|$ denotes the Euclidean distance between sites s and u . For each of the principal correlation functions (exponential, spherical and Gaussian) considered in this work, θ is the range parameter.

For each monitoring site s_i , we assume that the true trend $Z(s_i)$ is unobserved and unknown, but that it is estimated by $\tilde{Z}(s_i)$, along with its standard error, from the *GAM* in the

first stage of the analysis. The two variables are related by the equation

$$\tilde{Z}(s_l) = Z(s_l) + e_l, \quad (4)$$

where $e_l \sim N(0, \tilde{\sigma}_l^2)$ are interpreted as measurement errors, independent of the random field $Z(\cdot)$, $l=1, \dots, 35$. Note that each monitoring site has its own ‘‘nugget effect’’, $\tilde{\sigma}_l^2$, that are assumed known and equal to the jackknifed variance estimates of trend obtained in the first stage by fitting model (1). For now, we assume e_1, \dots, e_{35} to be independent random variables. This is justified, as an approximation, by the asymptotic normality of the *GAM* estimates.

Equations (3) and (4) together form a hierarchical model for the ‘‘trend data’’,

$\tilde{\mathbf{Z}} = (\tilde{Z}(s_1), \dots, \tilde{Z}(s_{35}))'$, but because both parts of the hierarchy are assumed normally distributed,

we can proceed directly to the likelihood of the model parameters, which is given by

$$L(\beta, \alpha, \theta; \tilde{\mathbf{Z}}) = \left(\frac{1}{2\pi}\right)^{35/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\tilde{\mathbf{Z}} - X\beta)' \Sigma^{-1}(\tilde{\mathbf{Z}} - X\beta)\right\} \quad (5)$$

where X is a known $35 \times q$ matrix of known regressors, $X_{ij} = f_j(s_i)$, and $\Sigma = \alpha V(\theta) + S$,

$V(\theta)_{ij} = K_\theta(\|s_i - s_j\|)$ and $S = \text{diag}(\tilde{\sigma}_1^2, \dots, \tilde{\sigma}_{35}^2)$. An alternative modeling strategy which allows for spatial correlation among the measurement errors could be used. Under this model, the diagonal S would be replaced with a general covariance matrix $S = \text{Var}(\underline{e})$. The model parameters β , α , and θ are estimated by maximum likelihood; see Cressie (1993) and Mardia and Marshall (1984) for details about fitting and asymptotic properties of ML estimators.

We estimate the trend $Z(s_o)$ at any location s_o using the best linear unbiased predictor

(BLUP), given by

$$\hat{Z}(s_o) = x_o' \hat{\beta} + \tau_o' \Sigma^{-1} (\tilde{Z} - X \hat{\beta}) \quad (6)$$

where $x_o' = (f_1(s_o), \dots, f_q(s_o))$, $\hat{\beta} = (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} \tilde{Z}$, and $\tau_o' = (Cov\{Z(s_o), \tilde{Z}(s_1)\}, \dots, Cov\{Z(s_o), \tilde{Z}(s_{35})\})$. The mean-squared prediction error (MSPE) of $\hat{Z}(s_o)$ is given by

$$\sigma^2(s_o) = \alpha - \tau_o' \Sigma^{-1} \tau_o + (x_o - X' \Sigma^{-1} \tau_o)' (X' \Sigma^{-1} X)^{-1} (x_o - X' \Sigma^{-1} \tau_o). \quad (7)$$

The right-hand sides of equations (6) and (7) depend on the covariance parameters which are unknown. So in practice, we use the estimated (or empirical) BLUP and MSPE obtained by replacing α and θ with their estimates. This tends to underestimate the prediction uncertainty. To remedy this, we later consider a Bayesian analysis to evaluate the effect of ignoring the uncertainty of the covariance parameters on inference about regional trends.

4. REGIONAL TREND ESTIMATION

Let R be a geographical subregion over which trend estimation is desired. A natural measure of trend over R is given by

$$\frac{1}{|R|} \int_R Z(s) ds, \quad (8)$$

which we define as the regional trend of R ; $|R|$ is the size of the region. This integral can be approximated by

$$Z_R = \frac{1}{N} \sum_{i=1}^N Z(s_{oi}), \quad (9)$$

where $\{s_{o1}, \dots, s_{oN}\} \subseteq R$ form a grid of points discretizing the region R .

4.1 Kriging Procedures

The optimal predictor of Z_R is given by

$$\hat{Z}_R^{krig} = E(Z_R | \tilde{Z}, \hat{\alpha}, \hat{\theta}) = \frac{1}{N} \sum_{i=1}^N \hat{Z}(s_{oi}), \quad (10)$$

with variance

$$\sigma_R^{2krig} = Var(Z_R | \tilde{Z}, \hat{\alpha}, \hat{\theta}) \cong \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N Cov\{Z(s_{oi}), Z(s_{oj}) | \tilde{Z}, \hat{\alpha}, \hat{\theta}\}, \quad (11)$$

where $\hat{\alpha}, \hat{\theta}$ are the ML estimates, $\hat{Z}(s_{oi})$ is given by (6), and

$$\begin{aligned} Cov\{Z(s_{oi}), Z(s_{oj}) | \tilde{Z}, \hat{\alpha}, \hat{\theta}\} &= \hat{\alpha} K_{\hat{\theta}}(\|s_{oi} - s_{oj}\|) - \hat{\tau}_{oi}' \Sigma^{-1} \hat{\tau}_{oj} + \\ & (x_{oi} - X' \hat{\Sigma}^{-1} \hat{\tau}_{oi})' (X' \hat{\Sigma}^{-1} X)^{-1} (x_{oj} - X' \hat{\Sigma}^{-1} \hat{\tau}_{oj}), \end{aligned} \quad (12)$$

where x_{oi} and τ_{oi} are defined analogously as x_o and τ_o . In equation (12), ML estimates of α and θ are used to obtain $\hat{\tau}_{oi}, \hat{\tau}_{oj}$, and $\hat{\Sigma}^{-1}$.

4.2 Bayesian Procedures

Because the covariance parameters are treated as known when in fact they are estimated, the empirical MSPE is expected to underestimate the true prediction variance. This defect is not easily remedied within the maximum likelihood estimation framework, but is easier to handle if we adopt a Bayesian perspective to compute the posterior distribution of the covariance parameters. To introduce this, we use the model defined in (5) with prior density (Handcock and Stein 1993)

$$\pi(\beta, \alpha, \theta) \propto \frac{\pi(\theta)}{\alpha}. \quad (13)$$

Then we have that $\pi(\beta, \alpha, \theta | \tilde{Z}) \propto L(\beta, \alpha, \theta; \tilde{Z}) \pi(\beta, \alpha, \theta)$ and after integrating this posterior with respect to β we obtain

$$\pi(\alpha, \theta | \tilde{Z}) \propto |\Sigma|^{-1/2} |X' \Sigma^{-1} X|^{-1/2} \exp\left\{-\frac{1}{2}(\tilde{Z} - X\hat{\beta})' \Sigma^{-1}(\tilde{Z} - X\hat{\beta})\right\} \frac{\pi(\theta)}{\alpha}. \quad (14)$$

An obvious way to implement this is to take $\pi(\theta)$ to be an improper prior, but there is a problem with this because such a choice leads to an improper posterior. For a standard geostatistical model without measurement error ($S=0$), this has been shown by Berger, De Oliveira, and Sanso (paper forthcoming), while in the case we are considering with measurement error, the likelihood tends to a positive constant as θ tends to infinity, so if the prior distribution is improper, the posterior distribution will be as well. Therefore, we must employ a proper prior for θ . The inverse gamma (*IG*) family encompasses a wide variety of distributional shapes and we therefore adopt a prior distribution $\pi(\theta) = IG(a, b)$ with mean $b/(a-1)$ for $a > 1$.

This posterior distribution is not amenable to analytical treatment, so we use Markov chain Monte Carlo (MCMC) methods to draw a sample from the posterior distribution (14). This sample will be used to make inference about α and θ as well as regional trends; see Gilks, Richardson, and Spiegelhalter (1996) for details on MCMC methods. For this, note that none of the full conditional distributions of (14) are of standard form, so we use a Metropolis-Hastings algorithm where we update the parameters α and θ one at a time. For proposal distributions, we use uniform distributions centered at the current values with scales chosen so that the empirical acceptance rate is approximately 0.5.

With the sample $\{(\alpha_i, \theta_i); i=1, \dots, m\}$ from the posterior distribution (14), we can estimate regional trends and their standard errors in a way that does not require the covariance parameters to be known and accounts for their uncertainty. For this we use the optimal Bayesian predictor of Z_R , which from the decomposition of the expectation can be written as

$$\hat{Z}_R^{Bayes}(s_o) = E(Z_R | \tilde{Z}) \cong \frac{1}{Nm} \sum_{r=1}^N \sum_{i=1}^m E(Z(s_{or}) | \tilde{Z}, \alpha_i, \theta_i), \quad (15)$$

where $E(Z(s_{or}) | \tilde{Z}, \alpha, \theta)$ is given by equation (6), and

$$\sigma_R^{2Bayes} = Var(Z_R | \tilde{Z}) \cong \frac{1}{N^2} \sum_{r=1}^N \sum_{t=1}^N Cov\{Z(s_{or}), Z(s_{ot}) | \tilde{Z}\}. \quad (16)$$

The above expression is computed using the decomposition of the covariance

$$\begin{aligned} \text{Cov}\{Z(s_{or}), Z(s_{ot})|\tilde{Z}\} &= E\{\text{Cov}\{Z(s_{or}), Z(s_{ot})|\tilde{Z}, \alpha, \theta|\tilde{Z}\} \\ &+ \text{Cov}\{E(Z(s_{or})|\tilde{Z}, \alpha, \theta), E(Z(s_{ot})|\tilde{Z}, \alpha, \theta)|\tilde{Z}\} \end{aligned} \quad (17)$$

Since the $E(Z(s_{or})|\tilde{Z}, \alpha, \theta)$ and $\text{Cov}\{Z(s_{or}), Z(s_{ot})|\tilde{Z}, \alpha, \theta\}$ are explicit functions of α and θ given respectively by (10) and (12), the covariance in (17) can be approximated using the MCMC sample of these parameters.

It should be pointed out that the Bayesian analysis given here is not a “fully” Bayesian approach. The latter would apply Bayesian analysis to both stages of the analysis, including the initial *GAM* stage. The MCMC approach given here treats the results of the *GAM* analysis as given, and applies Bayesian methodology only to the spatial part. The main motivation behind pursuing a Bayesian analysis was to account for the uncertainty of the spatial model parameters in calculating standard errors for the kriging analysis, and this is achieved by the analysis which has been given. A fully Bayesian analysis would allow us to examine the adequacy of these calculations in more detail, but this would be considerably more computationally intensive than what has been given here, and has not so far been attempted.

4.3 Clustering of Sites

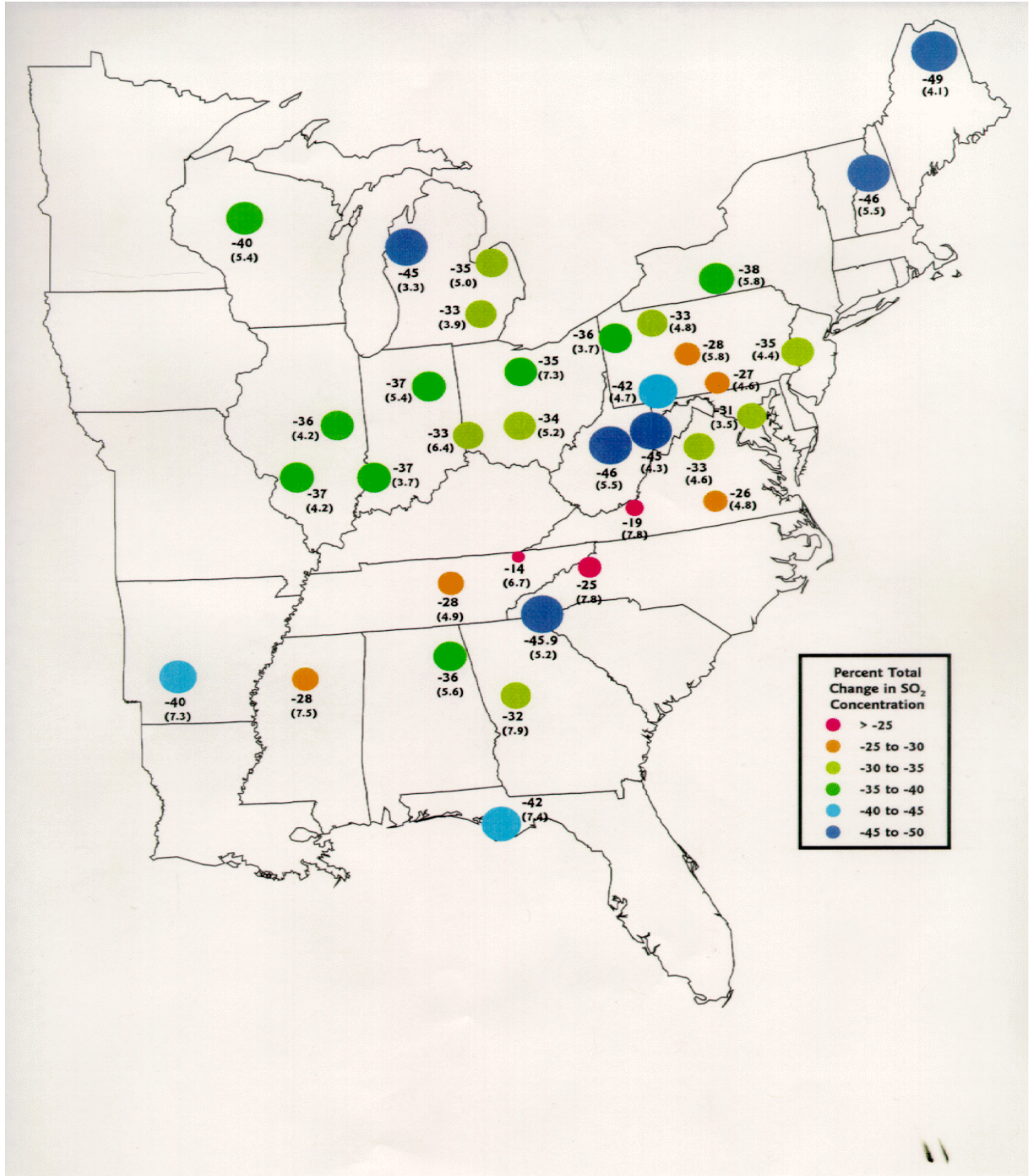
Based on a *k-means* clustering procedure (Hartigan and Wong 1979), the CASTNet sites were aggregated into three broad clusters of sites. Boundaries around each of these clusters were used to define the geographic areas over which inference about regional trends was done. Sites within a cluster show similar data patterns among themselves, while sites from different clusters

show comparatively distinct data patterns. In the *k-means* analysis the objective is to form the clusters in a way that the total within-cluster sum of squares is minimized. After an initial estimate for the cluster centroids, sites are moved from one cluster to another based on the cluster centroid, and then cluster centroids are recalculated. This process continues until no additional switching of sites among clusters reduces the total of within-cluster sum of squares.

5. RESULTS

5.1 Site-Specific Trends

Jackknifed site-specific estimates of trend were all negative (*i.e.*, SO₂ levels are decreasing) and could be statistically differentiated (.05 significance level) from a zero percent change at all sites (see Figure 2). The exponential variance factor in equation (2) was found to have minimal effect on $\tilde{Z}(s_l)$ and $\tilde{\sigma}_l^2$ (less than 1% for all sites) when compared to estimates without this factor. The functions $g_1(\textit{week})$ and $g_2(\textit{year})$ were included in the model at all sites based on the model selection procedure, and smoothing functions were fit to both of these variables at all sites. For 33 sites, $g_3(\textit{temperature})$ was included in the model and smoothing functions were fit at 32 of these sites. At 32 sites, SO₂ concentrations were dependent on a smoothed bivariate surface of (u, v) wind components. Site-specific model R^2 values ranged between 0.45 to 0.79. For 27 of the 35 sites, R^2 values exceeded 0.6, providing good *de facto* evidence that the site models are accounting for the seasonal and meteorological influences affecting SO₂ concentrations. Quantile-quantile plots of the residuals showed a non-linear



relationship between Gaussian quantiles of -2 to 2, indicating a normal distribution for 95% of the population. Autocorrelation in the site-specific model residuals was examined to determine if the model error structure is misspecified. For 4 of 35 sites, significant temporal correlation among the residuals was present based on 95% confidence limits for an independent series with $\rho=0$. All of these significant correlations were relatively low, approximately 0.2. Overall, these results support the assumption of uncorrelated residuals in model (1) as a reasonable approximation.

5.2 Model Determination for $Z(s)$

Based on an exploratory data analysis of the relationship between monitoring site location (latitude and longitude) and magnitude of trend, we evaluated two forms for $E\{Z(s)\}$: one constant and the other: $\beta_1 + \beta_2 x + \beta_3 y$; $s=(x,y)$. For each of these mean functions, we considered three commonly used correlation functions: exponential, Gaussian, and spherical (Cressie 1993).

The Gaussian correlation function

$$K_{\theta}(s_i, s_j) = \exp\left(-\frac{\|s_i - s_j\|^2}{\theta}\right), \quad (18)$$

in which $\theta > 0$ is a parameter controlling the spatial range of the correlation function and $\|\cdot\|$ is the Euclidean distance between two monitoring sites, produced the highest likelihood function evaluated at the parameter estimates among the three candidate correlation models, regardless of which of the mean functions is used. Although the likelihood function using the Gaussian correlation was greater in comparison to the other correlation functions, the differences among

the likelihood values for the three correlation functions were relatively small (less than 1.5). Further support for this correlation function was found after evaluation of the Matérn class of correlation functions. The Matérn class is characterized by two parameters: a scale parameter controlling the range of correlation and a smoothness parameter for controlling the smoothness of the random field. As the smoothness parameter goes to infinity, the Gaussian correlation is a limiting case of the Matérn class (Handcock and Stein 1993). In this analysis, the ML estimate of the smoothness parameter iteratively moved towards infinity corresponding to a model that is in practice equivalent to the Gaussian correlation. When including an additional nugget effect in $K(\theta)$ separate from the measurement error effect defined in equation (9), its estimated value was zero while the estimates of the other parameters were equivalent to those obtained through estimation with no nugget. Using the likelihood ratio test, the hypothesis that the process has a constant mean function is not rejected at the 0.05 significance level (p -value is about 0.09), so we selected the constant mean function over the alternative.

A further elaboration of this model was considered assuming spatial correlation among the e_1, \dots, e_{35} . The jackknifed series of trend estimates at each site was used to estimate the sample correlation matrix of e_1, \dots, e_{35} . Most of these correlations were less than 0.4, but some were as high as 0.7. For the model with constant mean and Gaussian correlation, replacing the diagonal matrix S with a general or non-diagonal covariance matrix $S = \text{Var}(\underline{e})$ provided the best fit in terms of the AIC and BIC criteria (see Table 1). This model was also tested for geometric anisotropy where the process is not isotropic in the original space, but in a linearly-transformed space (2-dimensional transformation) that corresponds to a combination of stretching and compressing the coordinates. As shown in Table 1, this model offers no improvement in fit when compared to the

fitting criteria values for the other models. Thus, our final selected model has a constant mean, Gaussian correlation with no nugget, and a non-diagonal S matrix. One potential disadvantage in estimating all elements of the non-diagonal S matrix is the possibility of increased estimation error in the other parameters of the model. This does not appear to occur here since a direct comparison of the two S matrices indicates that the non-diagonal matrix provides a better fit.

Table 1. Comparison of covariance fits for two forms of $E\{Z(s)\}$ and $S=\text{Var}(\underline{e})$ using a Gaussian correlation model.

Covariance Model	-(log-likelihood)	number of parameters (p)	AIC ¹	BIC ²
constant mean / diagonal $\text{Var}(\underline{e})$	87.0	3	180.0	184.7
linear mean / diagonal $\text{Var}(\underline{e})$	84.6	5	179.3	187.1
constant mean / general $\text{Var}(\underline{e})$	83.2	3	172.5	177.1
linear mean / general $\text{Var}(\underline{e})$	82.0	5	174.0	181.8
constant mean / general $\text{Var}(\underline{e})^3$	82.0	5	174.0	181.7
linear mean / general $\text{Var}(\underline{e})^3$	80.8	7	175.6	186.5

¹ Akaike Information Criterion: $-2(\log\text{-likelihood})+2p$

² Bayesian Information Criterion: $-2(\log\text{-likelihood})+p\log(n)$

³ Geometric Anisotropic Covariance

A rather vague prior specification for $\pi(\theta)$ is obtained by setting $a = 2$ (so having an infinite a priori variance) and $b=1.73$ in the inverted gamma distribution (see Figure 3). The Markov Chain is generated by sampling from equation (14) using a Metropolis-Hastings step. Figure 4 shows the estimated marginal posterior distributions of α and θ , obtained from the MCMC sample. A limited sensitivity analysis of the posterior distributions of α and θ indicated that the posterior of α was not very sensitive to the choice of a and b . However, the posterior

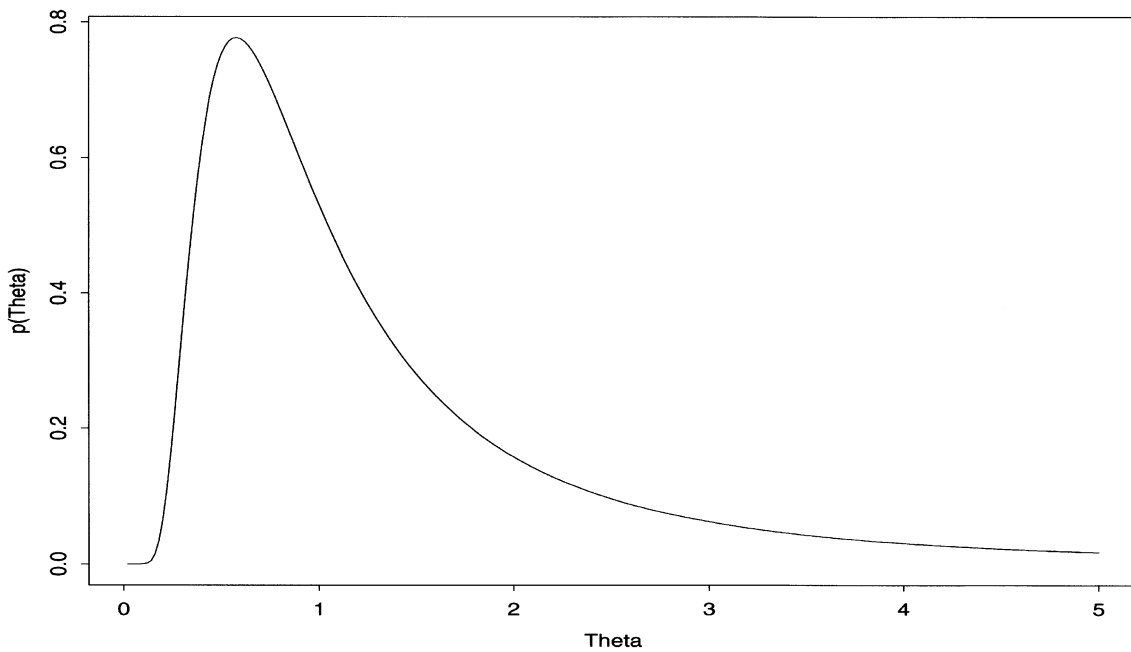


Figure 3. Plot of prior distribution for θ , $\pi(\theta)$, for inverted gamma distribution parameters $a=2$ and $b=1.73$.

distribution of θ showed more probability mass for small values of theta when both a and b were close to zero. An examination of other estimated posterior distributions for values of $a > 1$ and $b > 1$ indicated that using $a=2$ and $b=1.73$ provided a reasonable choice for the prior distribution.

These hyperparameter values produce a comparatively flat prior distribution that imparts some sense of being “noninformative”.

The predicted surfaces for trend and standard error (based on estimators in (6) and (7)) are shown in Figures 5 and 6. Overall, trend predictions vary from -35% to -45% over much of the eastern U.S. Increasing and decreasing gradients of this predicted surface appear to be consistent with the sites showing the highest and lowest trends in SO_2 concentrations. The

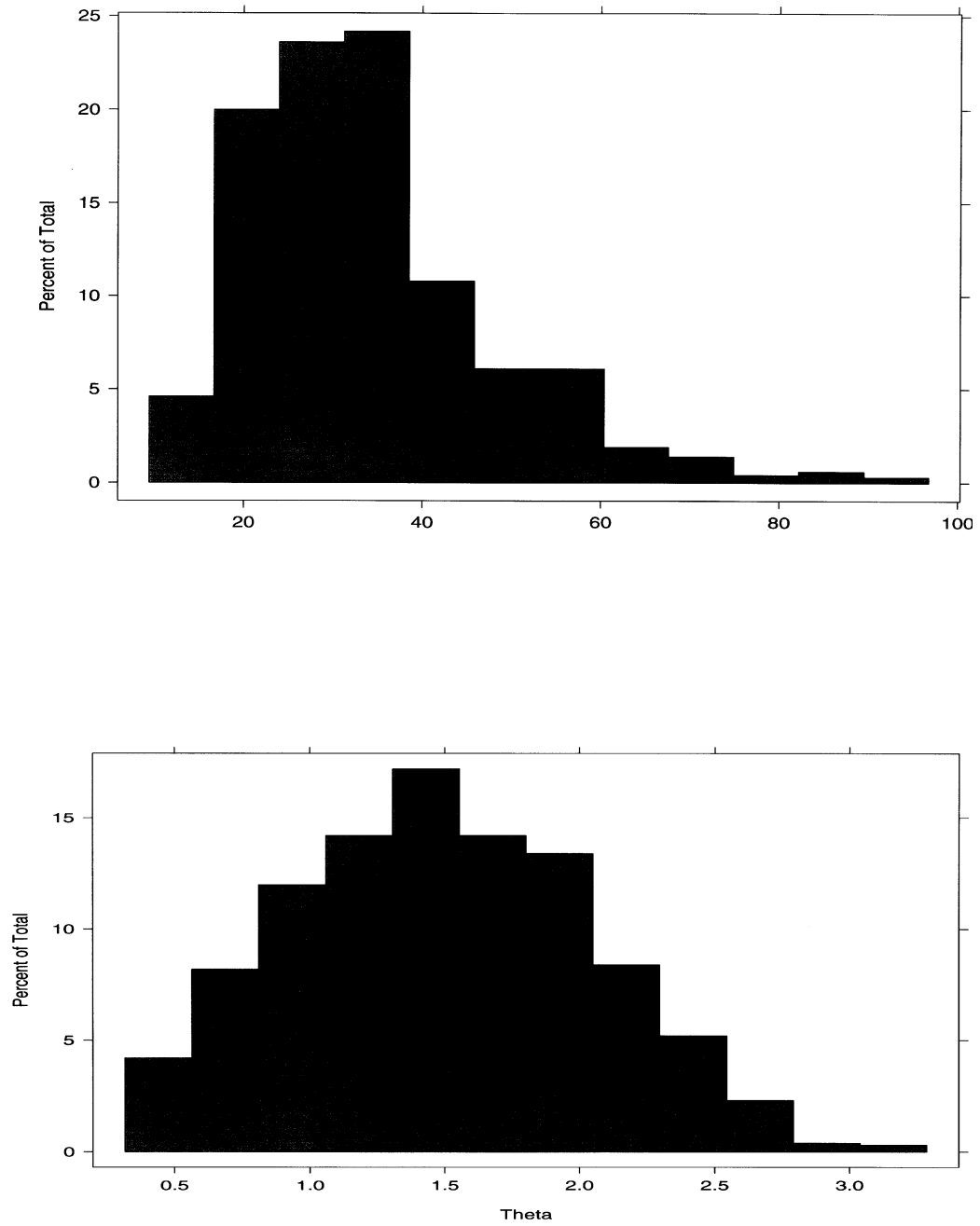


Figure 4. Estimated marginal posterior distribution of α (%²) and θ (degrees).

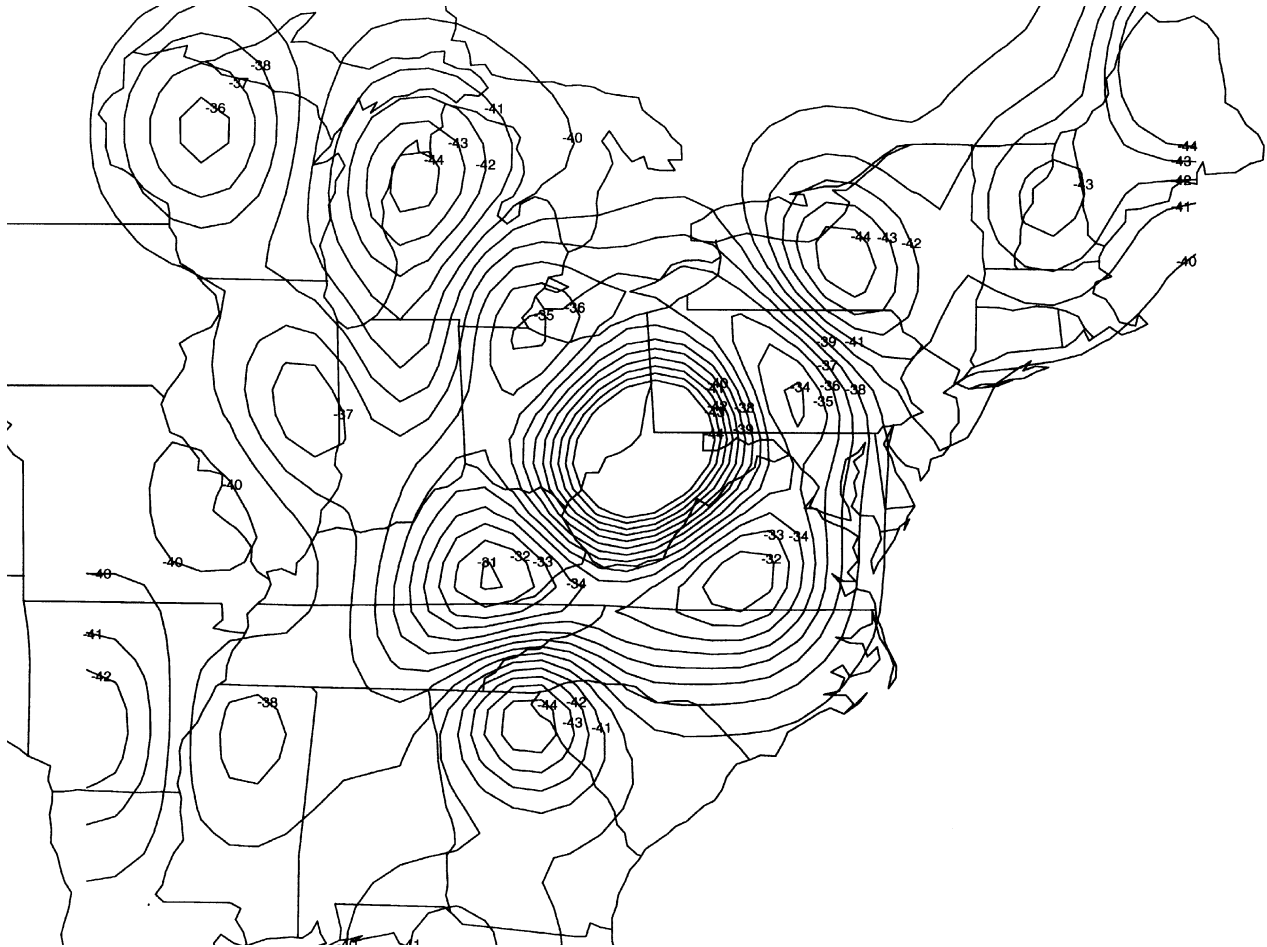


Figure 5. Smoothed surface of trend (*percent total change*) in SO_2 concentrations ($\mu\text{g}/\text{m}^3$) based on covariance model with a constant mean, Gaussian correlation, and general $S=\text{Var}(\underline{e})$

predicted surface in the Ohio River Valley shows a nearly uniform surface varying from -33% to -37%.

5.3 Regional Trend Estimation

The delineation of the boundaries of the regions was based on identifying geographic

areas of coherent seasonally-adjusted SO_2 data (based on *k-means* analysis of section 4.3) and matching these regions to regions administered by the U.S. EPA. Regional EPA administrators

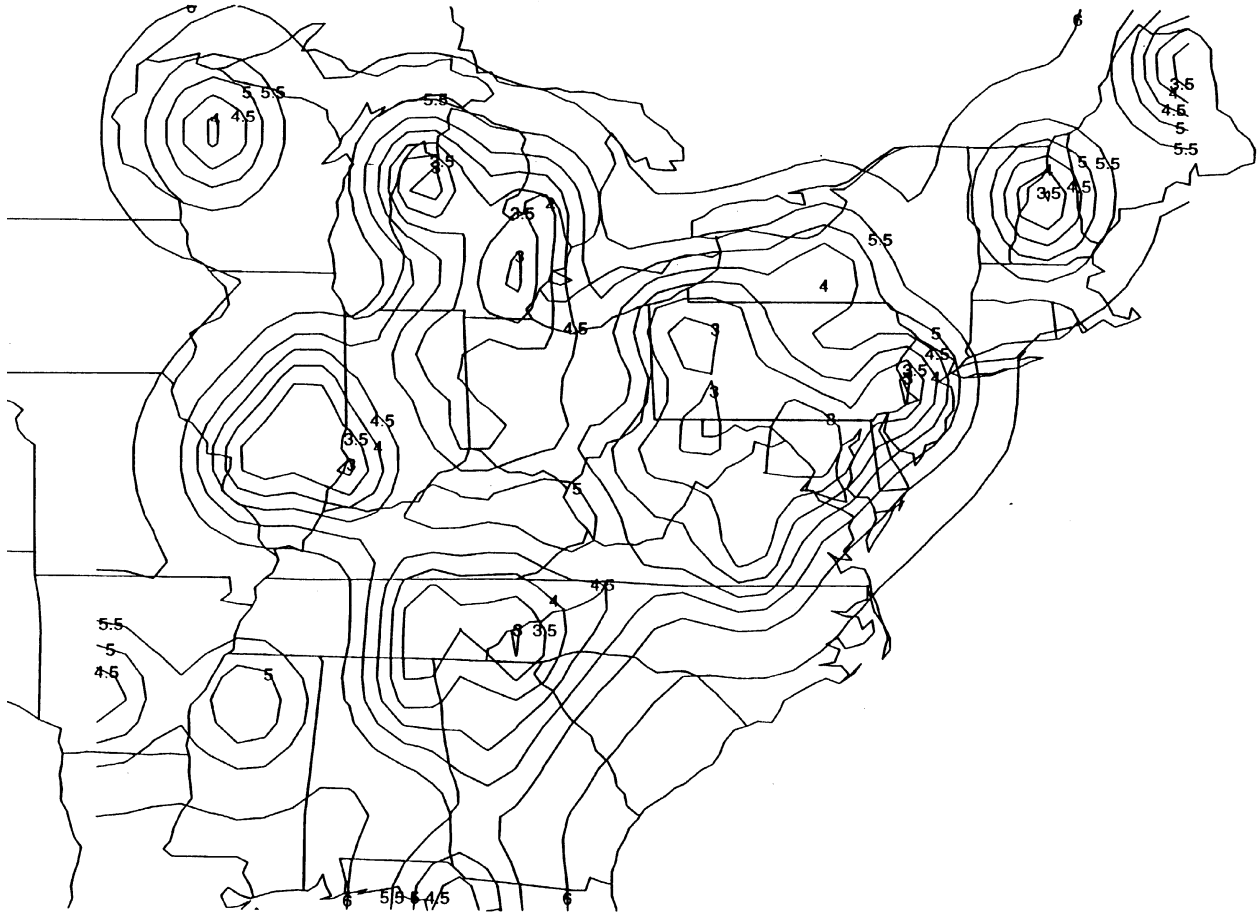


Figure 6. Surface of standard error of trend (%) based on a covariance model with a constant mean, Gaussian correlation, and general $S=\text{Var}(e)$.

are responsible for developing and implementing programs for integrated environmental protection activities and conducting effective enforcement and compliance programs. Thus,

regional trend estimates that adhere as closely as possible to these regions would have the most impact on contributing to the efficiency of environmental management within each EPA region. Given the generally uniform distribution of sites across states in the Midwest region, this region was defined to include Illinois, Indiana, Ohio, Michigan, and Wisconsin. This region is nearly identical to EPA Region V, the difference being the exclusion of Minnesota. The Southern region was defined to include Virginia, North Carolina, South Carolina, Tennessee, Georgia, Alabama, Mississippi, Arkansas, and the northern area of Florida. This region covers almost all of the states in EPA Region IV. However, Kentucky is not included and Region IV does not include Virginia. The Mid-Atlantic region covers all of West Virginia, Maryland, Delaware, New Jersey, Pennsylvania, and southern New York. These areas are part of EPA Regions II and III.

Although the AIC and BIC calculations clearly favor a covariance model with non-diagonal S , a sensitivity of this modeling assumption was performed using a Gaussian correlation and constant mean. For two types of covariance, one with diagonal S and the other with non-diagonal S , kriging predictions and standard errors were calculated at the grid nodes of a lattice superimposed over each of the three regions (see Figure 7) using the estimators given in equations (6) and (7). Regional estimates of trend and its variance were obtained from the estimators in equations (10) and (11). Similarly, the Bayesian approach was used to obtain the regional estimates for comparison to the kriging results. In the MCMC analysis, 1000 samples were drawn from the posterior distribution given in equation (14).

For each covariance model described above, there is little difference between the kriging and Bayesian estimates for regional trends (see Table 2). The similar trend estimates may be explained by recognizing that kriging predictions are generally robust against specification of the

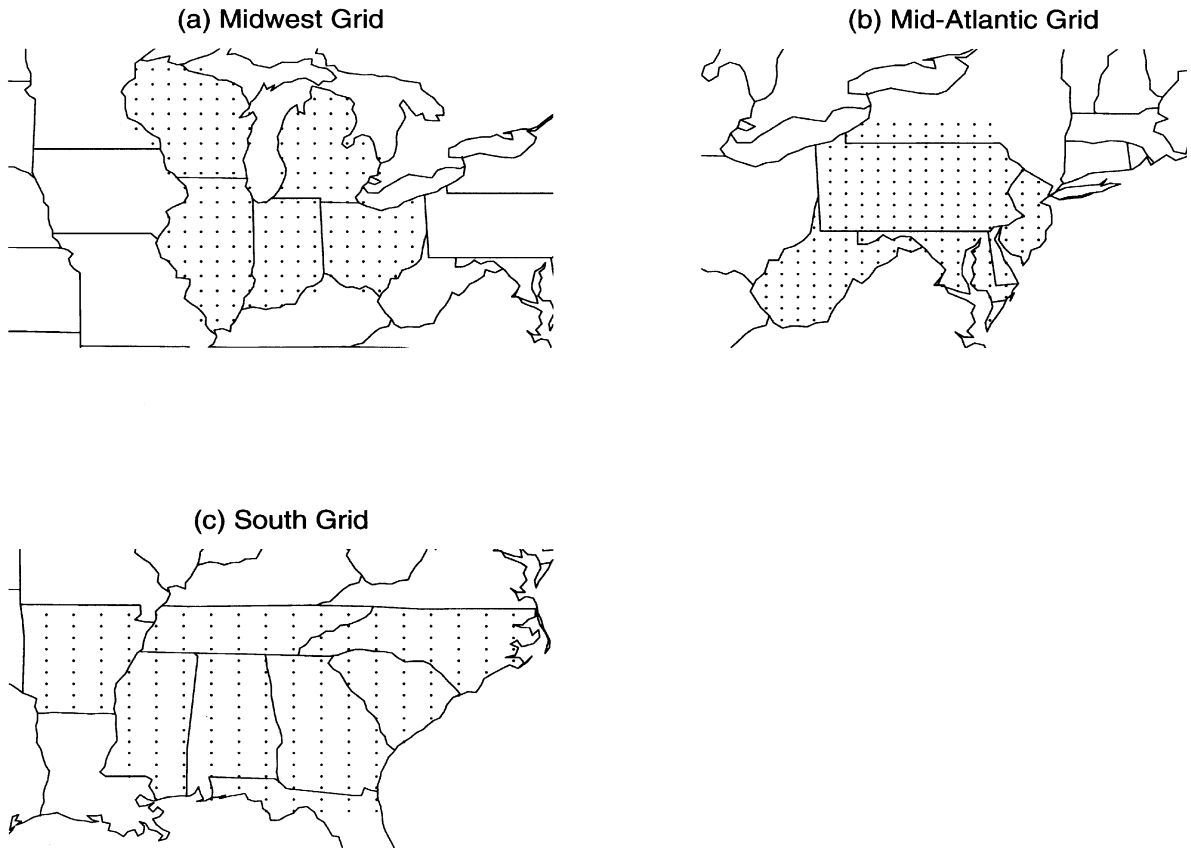


Figure 7. Lattices superimposed over three regions. Region (a) is the Midwest, (b) is the Mid-Atlantic, and (c) is the South.

covariance parameters and we are averaging over large geographic regions. Although we expected the Bayesian analysis to result in a significant increase in standard errors, we found that this did not occur. The Bayesian estimates of standard errors were just slightly larger than the kriging standard errors, or even slightly smaller in some cases. The latter result is undoubtedly an artifact, apparently caused by the fact that estimates of α tended to be slightly smaller in the Bayesian analyses than the ML estimation analyses, but the importance of this conclusion is that the standard errors are not changed very much by switching to Bayesian analysis. In the Mid-

West, results (using non-diagonal S) were obtained for two additional lattices, one with approximately twice as many grid nodes and one with approximately one-half the nodes of the lattice used in Table 2. Varying the spatial resolution of the lattice had little effect on the estimates of regional trend and standard error.

Table 2. Comparison of regional trend estimates (%) and their standard errors (%) using the kriging and Bayesian Methods.

Region	Kriging ¹	Bayesian ¹	Kriging ²	Bayesian ²
Mid-West	-36.56 ³ (1.37)	-36.28 (1.31)	-38.85 ¹ (2.28)	-38.45 (2.31)
Mid-Atlantic	-35.41 (1.35)	-35.58 (1.26)	-39.95 (2.27)	-39.39 (2.40)
South	-35.71 (1.65)	-35.78 (1.43)	-39.31 (2.45)	-38.85 (2.49)

¹ Covariance model with Gaussian correlation, constant mean, diagonal S .

² Covariance model with Gaussian correlation, constant mean, non-diagonal S .

³ The top number in each cell is the estimate of regional trend, the standard error is given in parenthesis.

Ignoring the off-diagonal elements of the S matrix does not result in a significant change in the estimates of regional trend, but trend estimates for this model have smaller standard error. A kriging analysis using restricted maximum likelihood (REML) estimates instead of ML estimates provided almost equivalent estimates of trend, with slightly larger standard errors.

Another example of how trend results can differ over a wider selection of models (including the previous models) is shown in Figure 8. This figure shows 95% prediction intervals for four predictands (average of all 35 stations, plus the three regional grids) and for eight models

(every possible combination of exponential/Gaussian correlation, constant/linear trend, and diagonal/non-diagonal S matrix). Also shown on each plot is the sample mean of the first stage GAM trend estimates within each region. In general, there is more smoothing in the models with non-diagonal S , but this is compensated for by wider prediction intervals. The models with diagonal S give a closer fit to the sample averages of GAM trend, but this should not be interpreted to mean that this model should be chosen over models with non-diagonal S . Based on the AIC and BIC calculations, the covariance with Gaussian correlation, constant mean, and non-diagonal S is still our recommended model. The comparisons in Figure 8 are intended to give some indication of the sensitivity of our final conclusions to these assumptions.

In our analysis, we have used degrees of latitude and longitude as a measure of distance, rather than a measure of real distance. To investigate whether this makes any difference for our recommended model, we also performed an analysis using nautical miles (nm). It was assumed as an approximation that 1° latitude = $60\ nm$ and 1° longitude = $45\ nm$. We note several results from this analysis: (1) although the likelihood values are better in the nm model than in the degrees model, the difference is too small to indicate a significant improvement; (2) the comparisons among different models for spatial trend and covariance function are exactly the same in both models; and (3) the changes in estimated regional averages are very close to those given in Table 2.

For our analysis, annual SO_2 emission estimates from both area and point sources were used for comparison to regional trends in SO_2 concentrations. Area source estimates are available at the county level for each year from 1989-1995. This approach produces a small inconsistency between the final complete year of emissions data (1995) and the final year of SO_2 data that

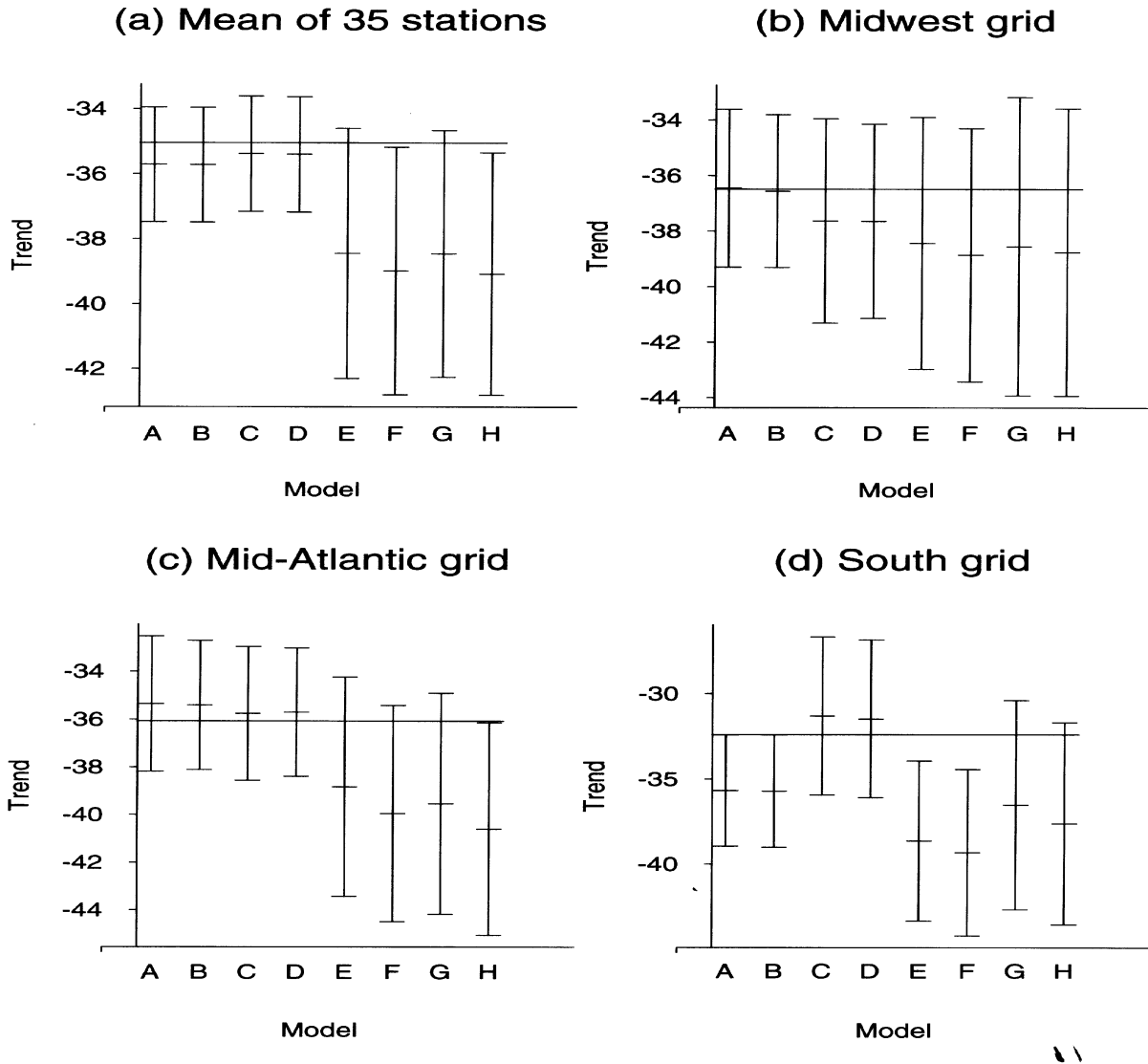


Figure 8. Comparison of regional average predictions for eight covariance models: A, covariance based on exponential correlation function, constant mean, diagonal S ; B, covariance based on Gaussian correlation function, constant mean, diagonal S ; C, covariance based on exponential correlation function, linear mean, diagonal S ; D, covariance based on Gaussian correlation function, linear mean, diagonal S . E-H as in A-D, but include non-diagonal S . Predictions are calculated for (a) overall mean of 35 stations, (b) Midwest grid (c) Mid-Atlantic grid, (c) South grid. In each case, the mean prediction and 95% prediction error bounds (non-Bayesian calculation) are shown for each of the eight models. The horizontal line on each plot represents the sample mean of all 35 stations (plot(a)) or of the monitoring sites within each region (plots(b)-(d)).

ended on September 30, 1995. Total SO_2 emission changes in the Mid-West, Mid-Atlantic, and

South were calculated to be -28%, -18 and -10, respectively (U.S. EPA 1998b). For all regions, the trends in SO₂ concentrations are much larger than the corresponding change in emissions. This disparity may be explained by the following discussion. The largest contributor of SO₂ emissions is fuel combustion, accounting for 85% of the total in 1997 (U.S. EPA 1998c). The electric utility industry accounts for most of the fuel combustion emissions, in particular the coal-burning power plants have been the largest contributor to SO₂ emissions. The national reductions from 1994-95 in SO₂ emissions and ambient concentrations of SO₂ are due mainly to Phase I implementation of the CAA. Phase I compliance has significantly reduced emissions from the participating utilities which are the largest emitters of SO₂. Since by design CASTNet sites are "rural" sites, they are most likely to monitor regional air quality masses, and for SO₂ this means mostly electric utility emissions. For the period, 1994-95, Phase I emissions decreased by approximately 29%. Including the much smaller decrease that occurred prior to 1994, the overall (1989-1995) decrease in Phase I emissions is similar to the estimates of regional trends reported here.

6. CONCLUSIONS

The increasing focus of environmental policy decisions on regional-scale solutions has led to a corresponding increase in the importance of estimating pollution trends over regional-scale landscapes. Where in the past, environmental policy-makers may have been content with site-specific estimates of trends from a monitoring network, they now expect the characterization

of pollution trends over broad geographic areas so that the effectiveness of emission control strategies can be evaluated with respect to pollution-sensitive receptors, whether human or ecological. This paper presents methodology to obtain spatially smooth estimates of regional trend and its standard error, and applies a Bayesian analysis with MCMC methods to account for the extra variability induced when parameters of the spatial covariance function are estimated. Although there is evidence that trend varies by location over the eastern US, these approaches produced similar results in terms of broad regional trend and standard error. For these data, the additional uncertainty from the estimation of the covariance parameters did not significantly increase the regional standard error. However, for other applications this source of variability may be more important. For future research, we anticipate extending the hierarchical models analysis to allow for the non-Gaussian distribution of errors in the first stage, and a fully Bayesian implementation, including the estimation of a posterior distribution of trend parameters at each of the sites.

Disclaimer

The U.S. Environmental Protection Agency, through its Office of Research and Development, funded and performed the research described here. This paper has been subjected to peer and administrative review and has been approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Acknowledgements

This research was supported in part by Contract QT-RT-97-001532 between the U.S. Environmental Protection Agency and the National Institute of Statistical Sciences (NISS). Richard Smith was partly funded by NSF grant DMS-9705166. While a post-doctoral researcher at NISS, Victor De Oliveira was partly funded by NSF grant DMS-9208758. We would like to thank Jerry Sacks, Peter Bloomfield, Peter Principe, and Joseph Sickles for their comments, which helped to clarify and improve the manuscript.

References

- Akaike, H. (1973). 'Information Theory and an Extension of the Maximum Likelihood Principle'. in Petrov, B. N., and Csàki, F. (eds), *Second International Symposium on Information Theory*, Akademia Kiadó, Budapest, 267-281.
- Bloomfield, P., Royle, A. J., and Yang, Q. (1993). 'Rural Ozone and Meteorology: Analysis and Comparison with Urban Ozone'. National Institute of Statistical Sciences, Technical Report No. 5.
- Chambers, J. M. and Hastie, T. J. (Eds.) (1992). *Statistical Models in S*, Wadsworth & Brooks/Cole Computer Science Series.
- Clarke, J. F., Edgerton, E. S., and Martin, B. E. (1997). 'Dry deposition calculations for the Clean Air Status and Trends Network'. *Atmos. Envir.*, **31**, 3667-3678.
- Cressie, N. (1993). *Statistics for Spatial Data*, Revised edition, John Wiley.
- Davis, J. M. and Speckman, P. (199). 'A model for predicting maximum and 8-hr average ozone in Houston'. *Atmos. Envir.* **33**, 2487-2500.
- Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*, Chapman and Hall.
- Flaum, J. B., Rao, S. T., and Zurbenko, I. G. (1996). 'Moderating the influence of meteorologic conditions on ambient ozone concentrations'. *J. Air & Waste Manage. Assn.*, **46**, 3546.
- Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*, Chapman & Hall
- Handcock, M. S. and Stein, M. L. (1993). 'A Bayesian analysis of kriging'. *Technometrics* **35**, 403-410.
- Hartigan, J. A. and Wong, M. A. (1979). 'A k-means clustering algorithm'. *Appl. Stat.* **28**, 100-108.
- Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*, Chapman and Hall.
- Holland, D. M., Principe, P. P., and Sickles, J. E., II, (1999). 'Trends in atmospheric concentrations of sulfur and nitrogen species in the Eastern United States for 1989-1995'. *Atmos. Envir.* **33**, 37-49.
- Lefohn, A. S., and Shadwick, D. S. (1991). 'Ozone, sulfur dioxide, and nitrogen dioxide trends at rural sites located in the United States'. *Atmos. Envir.*, **25A**, 491-501.
- Mardia, K. V. and Marshall, R. J. (1984). 'Maximum likelihood estimation of models for residual covariance in spatial regression'. *Biometrika* **71**, 135-146.
- Shreffler, J. H. and Barnes, M. H., Jr. (1996). 'Estimation of trends in atmospheric concentrations of sulfate in the northeastern United States'. *J. Air & Waste Manage. Assn.* **46**, 621-630.

U. S. Environmental Protection Agency (1998a). *Clean Air Status and Trends Network (CASTNet) Deposition Summary Report (1989-1995)*. EPA/600/R-98/027. Office of Research and Development, Research Triangle Park, NC.

U. S. Environmental Protection Agency (1998b). *National Air Pollutant Emissions Trends Update: 1970-1997*. EPA/454/E-98-007. Office of Air Quality Planning and Standards, Research Triangle Park, NC.

U. S. Environmental Protection Agency (1998c). *National Air Quality and Emissions Trend Report, 1997*. EPA/454/R-98-016. Office of Air Quality Planning and Standards, Research Triangle Park, NC.