High-Resolution Space–Time Ozone Modeling for Assessing Trends

Sujit K. SAHU, Alan E. GELFAND, and David M. HOLLAND

This article proposes a space–time model for daily 8-hour maximum ozone levels to provide input for regulatory activities: detection, evaluation, and analysis of spatial patterns and temporal trend in ozone summaries. The model is applied to the analysis of data from the state of Ohio that contains a mix of urban, suburban, and rural ozone monitoring sites. The proposed space–time model is autoregressive and incorporates the most important meteorological variables observed at a collection of ozone monitoring sites as well as at several weather stations where ozone levels have not been observed. This misalignment is handled through spatial modeling. In so doing we adopt a computationally convenient approach based on the successive daily increments in meteorological variables. The resulting hierarchical model is specified within a Bayesian framework and is fitted using Markov chain Monte Carlo techniques. Full inference with regard to model unknowns as well as for predictions in time and space, evaluation of annual summaries, and assessment of trends are presented.

KEY WORDS: Dynamic model; Forecasting/prediction; Markov chain Monte Carlo; Misalignment; Spatial variability; Stationarity.

1. INTRODUCTION

In this article we develop a new spatial–temporal model for predicting spatial patterns and associated uncertainties in ozone concentrations and for detecting long-term trends. Here, we use data observed from 1997 to 2004 for the state of Ohio at ozone monitoring sites located in a variety of urban, suburban, and rural settings. We use several important meteorological variables observed at some of the ozone monitoring sites and also at sites where ozone has not been observed (e.g., several airports in Ohio and neighboring states). We address the spatial and temporal misalignment of the pollution and meteorological data through spatial modeling. This allows us to develop a disaggregated model that takes a novel form by relating current-day ozone concentration to previous day-ozone (autoregressive part), an annual intercept term, an incremental effect due to meteorology, and a spatially correlated error term. As we clarify in Section 3.1, this avoids the potentially contentious issue of direct modeling of meteorology and focuses on modeling successive daily increments for a set of meteorological variables, a more straightforward task. In all of this we infer about latent "true" ozone levels used to define the ozone standard (discussed later), recognizing that observed levels may introduce missingness, bias error, and measurement error. The current National Ambient Air Quality Standards (NAAQS) for ozone is met if the 3-year rolling average of the annual fourth highest daily maximum 8-hour true average ozone concentration is less than 80 parts per billion (ppb); see, for example, *epa.gov/air/criteria.html*.

By modeling daily ozone concentrations, we can easily aggregate to any desired temporal summary of ozone, particularly the summary underlying the ozone air quality standard and, hence, study trends in such summaries. Note that, unlike other models that seek to examine trends (discussed later), we learn about trend without having to assume any functional form for it. By using space–time modeling, we can interpolate or predict ozone levels at any location in the state, again to whatever desired temporal summary. As a result, we achieve the most highly resolved (with regard to both space and time) analysis of ozone yet developed. A byproduct of our high-resolution modeling is the potential to link predicted true ozone concentrations to adverse health outcomes (see, e.g., Bell, McDermott, Zeger, Samet, and Dominici 2004). Daily predictions of true average ozone at arbitrary locations, which we can provide, offer a source of information for modeling such linkage. Moreover, by capturing uncertainty at such resolution and implementing inference within the Bayesian framework, we immediately obtain the uncertainty associated with any aggregation.

Lastly, space–time process modeling for ozone levels achieves a perhaps less appreciated benefit with regard to investigating extremes, such as the annual fourth highest daily average. Non-model-based interpolation of extremes, as would be done with monitoring data using standard software packages to create spatial surfaces, will tend to smooth them out, resulting in underestimation of the extent of noncompliance. The space–time dependence structure associated with process modeling is more effective in retaining the extremes of the latent ozone surfaces. (See Sec. 5 in this regard.)

More specifically, this article develops and illustrates several notions of site-specific summaries and trend surfaces over a large spatial domain. We consider spatial patterns in the annual fourth highest daily maximum 8-hour ozone concentrations and in the 3-year rolling averages, defined previously, across Ohio. Spatial patterns for the 3-year rolling averages can be examined for overall changes across the period 1997–2004 to assess trends in the ozone surface over time periods when emission reductions have been in place. Further, our modeling approach can be used to study site-specific trends by adjusting predicted ozone concentrations for meteorological effects where those have been observed. In this regard, we could attempt direct spatiotemporal modeling of extreme levels, for instance, as in the recent work of Gilleland and Nychka (2005). However, extremes need not be our only interest; the proposed high-resolution modeling enables more general assessment of ozone patterns in space and time.

Space–time modeling of air pollutants, ground-level ozone concentrations in particular, has attracted recent attention; see, for example, Guttorp, Meiring, and Sampson (1994) and Carroll et al. (1997). In recent years, hierarchical Bayesian approaches

Sujit K. Sahu is Senior Lecturer, School of Mathematics, Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, U.K. (E-mail: *S.K.Sahu@maths.soton.ac.uk*). Alan E. Gelfand is Professor, Institute of Statistics and Decisions, Duke University, Durham, NC 27708 (E-mail: *alan@stat.duke.edu*). David M. Holland is Senior Statistician, U.S. Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC 27711 (E-mail: *holland.david@epa.gov*).

© 0 American Statistical Association
Journal of the American Statistical Association
???? 0, Vol. 0, No. 00, Applications and Case Studies
DOI 10.1198/00

for spatial prediction of air pollution have been developed; see, for example, Brown, Le, and Zidek (1994), Sahu and Mardia (2005), Sahu, Gelfand, and Holland (2006) and the references therein. McMillan, Bortnick, Irwin, and Berliner (2005) proposed a regime switching model for ozone forecasting using meteorological variables as covariates and they illustrated this model using data from April to September in 1999 over a spatial domain covering Lake Michigan. They did not explicitly model the meteorological variables but their method required them as input possibly obtained from weather forecast data. They worked with data projected to a grid and then introduced a "nearest neighbor" spatial model; as a result, interpolation was precluded. With one year of data, their methodology was suitable for short-term forecasting of ozone but they were unable to investigate trends.

Cox and Chu (1992) used a generalized linear model approach, assuming a conditional Weibull distribution for ozone concentrations given meteorology, to estimate trends in daily maximum ozone levels. Porter, Rao, Zurbenko, Dunker, and Wolff (2001) reported on the estimation of trends in ozone concentrations adjusted for meteorological variables at individual monitoring sites. These authors used a moving-average, Kolmogorov–Zurbenko filter to separate a baseline component of log-transformed ozone consisting of long-term trend and seasonal variation from short-term weather variation. Cocchi, Fabrizi, and Trivisano (2005) followed the approach of Huang and Smith (1999) by using a tree-based partitioning of daily maximum ozone concentrations and assumed these maxima were Weibull distributed. The trend of ozone maxima was evaluated at a single site in Italy in terms of the sequence of yearly variations of medians within groups having homogeneous meteorology. A comprehensive overview of statistical methods for the statistical adjustment of ground-level ozone was given by Thompson, Reynolds, Cox, Guttorp, and Sampson (2001). Huerta, Sanso, and Stroud (2004) modeled hourly readings of concentrations of ozone jointly with air temperature for data from Mexico City. Their approach used a dynamic linear model with seasonal harmonics that enabled simultaneous forecasting of ozone and air temperature. Zhu, Carlin, and Gelfand (2003) related ambient ozone and pediatric asthma emergency room visits in Atlanta using hierarchical regression methods for spatially misaligned data. Finally, Wikle (2003) provided an overview of hierarchical modeling in environmental science.

The remainder of this article is organized as follows. Section 2 presents pertinent exploratory analyses of the data in order to facilitate model development. Our proposed model is developed in Section 3. Bayesian prediction methods and development of trend analysis are detailed in Section 4. Model-based analyses are provided in Section 5. A brief summary and future issues to explore are given in Section 6. An Appendix contains the computational details.

2. EXPLORATORY ANALYSIS

We model daily maximum 8-hour ozone concentration data obtained from $n = 53$ sites in the state of Ohio for our analysis. We have ozone data from $n_1 = 50$ National Air Monitoring Stations/State and Local Air Monitoring Stations (NAMS/SLAMS; *epa.gov/cludygxb/programs/namslam.html*) and $m_1 = 3$ Clean Air Status and Trends Network (CASTNET; *epa.gov/castnet*) sites. Most of the NAMS/SLAMS sites are located in or around the big cities, whereas the CASTNET sites operate in mostly rural areas. The NAMS/SLAMS network does not record meteorological data, whereas the CASTNET sites do. In addition, we have meteorological data from $m_2 = 9$ weather stations, which are mostly located near airports. Thus, we have ozone data from $n = n_1 + m_1 = 53$ sites and meteorological data from $n_2 = m_1 + m_2 = 12$ sites. All 62 sites are plotted in Figure 1 (three CASTNET sites numbered 1, 2, and 3 in the figure are overlapping). Note that there is at least one meteorological station near every cluster of ozone monitoring sites. In fact, two meteorological stations outside the state of Ohio have been kept precisely to achieve this purpose.

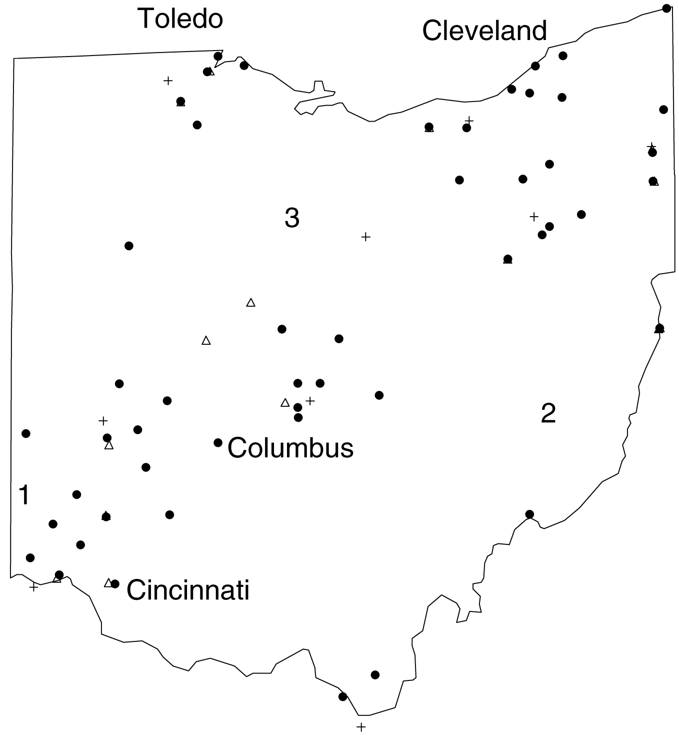

Figure 1. Ozone monitoring sites and meteorological sites in Ohio. The 50 NAMS/SLAMS sites are plotted as points; the sites numbered 1, 2, and 3 are three CASTNET sites; sites denoted by + are meteorological sites (two are outside Ohio); 15 validation sites are shown by the symbol △.

We consider data for $r = 8$ years from 1997 to 2004, inclusive. In each year we have data for $T = 169$ days covering the high ozone season from April 15 to September 30. However, 7,832 (=10.93%) of the total $N = nrT = 71{,}656$ are missing. In particular, about 50% of the data (roughly 4 years) were missing in eight sites; some of these sites started gathering data from 2001. In fact, in the years 1997–2000 the percentages of missing values were 22.94, 19.73, 16.70, and 13.61, respectively.

The boxplot of ozone values by year are plotted in Figure 2, which shows the overall levels. The overall level goes up in 1998, comes down to the lowest levels in 2000, and then rises again, but comes down in the year 2004. This pattern is also seen in the annual fourth highest daily maximum 8-hour average concentration levels as well; see the top panel of Figure 3. The bottom panel of Figure 3 plots the 3-year rolling averages, and we observe evidence of nonattainment (true ozone values greater than 80) in most of the sites in our study period.

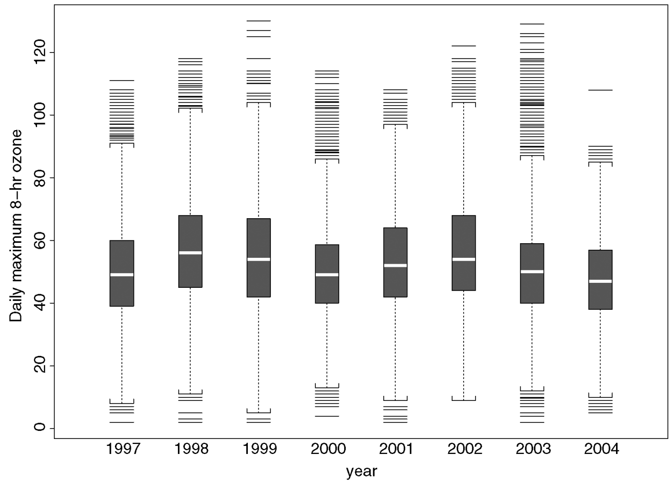

Figure 2. Boxplot of the daily maximum 8-hour ozone levels by year.

Standard multiple regression methods (stepwise, forward, and backward selection) were used to choose the most important meteorological variables to include in our model. The four ($=p$) most important variables are found to be maximum daily temperature in degrees Celsius, relative average humidity, and wind speeds in the morning and in the afternoon. McMillan et al. (2005) also included these four variables and two additional variables, average station pressure and wind direction in their work. However, given the four variables mentioned previously, we did not find these additional variables to be significant in our spatiotemporal analysis for data taken over the eight years 1997–2004. All the $n_2rTp = 64{,}896$ values of the successive daily increments of the four meteorological variables were used for our analysis. The time series plots (not included) of these variables are all centered around 0 and they do not show any autocorrelation, making those amenable to the independence assumption made in Section 3.1.

Data from 15 sites (in addition to the 53 modeling sites) have been set aside for validation purposes; these sites are also plotted in Figure 1. They are not included for modeling because about 70.46% observations were missing. In particular, there were only 5,991 available values out of the possible 20,282 ($=15rT$) observations.

Histograms and normal quantile–quantile (Q–Q) plots were plotted on the three measurement scales: original, logarithmic, and square root. The data on the original scale are surprisingly symmetric, but high variability would lead to negative fitted and predicted ozone concentrations. The log scale introduces negative skewness. The square root scale seems most attractive in terms of both symmetry and stabilizing the variance so that there are no negative fitted or predicted ozone values. This is in accord with other work in modeling air pollutants; see, for example, Sahu et al. (2006).

3. MODEL DEVELOPMENT

We use the notation $Z_l(\mathbf{s}, t)$ to denote the observed *square root* ozone concentration at location $\mathbf{s}$ in year l on day t. We have $t = 1, \ldots, T = 169$ and $l = 1, \ldots, 8$. We model data from $n = 53$ stations, denoted by $\mathbf{s}_1, \ldots, \mathbf{s}_n$, all within Ohio. Further, let $O_l(\mathbf{s}, t)$ denote the *true* value corresponding to $Z_l(\mathbf{s}, t)$. Let $x_{lj}(\mathbf{s}, t)$ and $\delta_{lj}(\mathbf{s}_i, t)$ denote, respectively, the value of the jth meteorological variable and the increment, $j = 1, \ldots, p$ in year l on day t. That is, $\delta_{lj}(\mathbf{s}_i, t) = x_{lj}(\mathbf{s}_i, t) - x_{lj}(\mathbf{s}_i, t-1)$. We shall use the

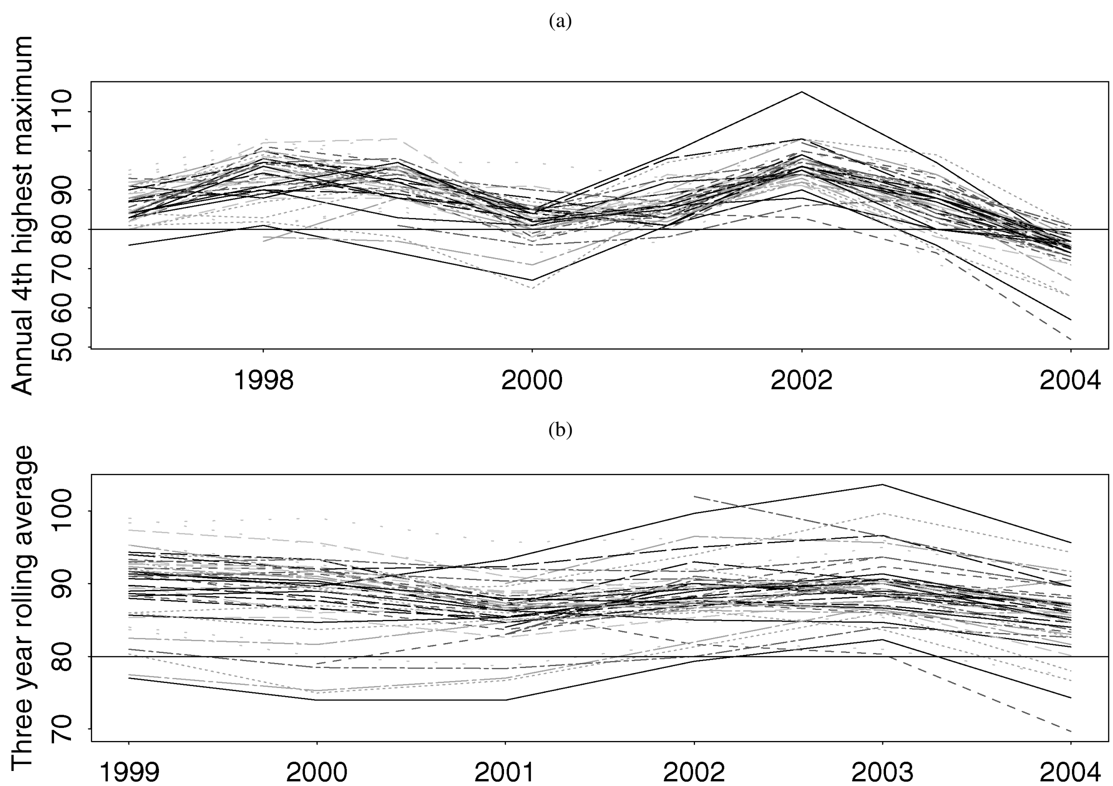

Figure 3. Annual fourth highest daily maximum ozone levels at 53 data sites in panel (a) and 3-year rolling averages in panel (b).

following vector notation: $\mathbf{Z}_{lt} = (Z_l(\mathbf{s}_1, t), \ldots, Z_l(\mathbf{s}_n, t))'$, $\mathbf{O}_{lt} = (O_l(\mathbf{s}_1, t), \ldots, O_l(\mathbf{s}_n, t))'$, $\mathbf{x}_l(\mathbf{s}_i, t) = (x_{l1}(\mathbf{s}_i, t), \ldots, x_{lp}(\mathbf{s}_i, t))'$, and $\boldsymbol{\delta}_l(\mathbf{s}_i, t) = \mathbf{x}_l(\mathbf{s}_i, t) - \mathbf{x}_l(\mathbf{s}_i, t-1)$.

To handle missingness along with potential bias and measurement error, we assume:

$$Z_l(\mathbf{s}_i, t) = O_l(\mathbf{s}_i, t) + \epsilon_l(\mathbf{s}_i, t), \qquad i = 1, \ldots, n, t = 1, \ldots, T, \tag{1}$$

where $\epsilon_l(\mathbf{s}_i, t)$ is a white-noise process, specifically assumed to follow $\mathrm{N}(0, \sigma_\epsilon^2)$ independently. Thus, σ_ϵ^2 is the so-called nugget effect. The Gaussian error assumption may be a concern due to occasional large excursions in ozone concentration levels. The use of square root transformation helps in this regard. However, it is possible to use a non-Gaussian error model for the $\epsilon_l(\mathbf{s}_i, t)$'s such as a t process. Regardless, preliminary residual analysis suggests that it is plausible to take σ_ϵ^2 to be homogeneous in space and time.

Next, we turn to the modeling for $O_l(\mathbf{s}, t)$. There is high autocorrelation between ozone measurements on successive days; hence, we include an autoregressive term in our model. We also introduce a global (not site specific) annual intercept parameter. However, the residuals after fitting such a model will show significant local variation. From the Introduction, we anticipate that this arises primarily due to changes in local meteorological conditions. However, we also introduce space–time random effects to allow for other unobserved but consequential local variables, enabling spatiotemporally varying intercepts. Thus, we assume that

$$O_l(\mathbf{s}, t) = \rho O_l(\mathbf{s}, t-1) + \xi_l + \boldsymbol{\delta}_l'(\mathbf{s}, t)\boldsymbol{\beta} + \eta_l(\mathbf{s}, t), \qquad t = 2, \ldots, T, \tag{2}$$

where $\eta_l(\mathbf{s}, t)$ is a spatially correlated error term, $\rho O_l(\mathbf{s}, t-1)$ is the autoregressive term with $0 < \rho < 1$, ξ_l is the global annual intercept in year l, and $\boldsymbol{\delta}_l'(\mathbf{s}, t)\boldsymbol{\beta}$ is the local adjustment to $O_l(\mathbf{s}, t)$ arising due to the increments in meteorological variables $\mathbf{x}_l(\mathbf{s}, t)$. In principle, nonlinear functions of change in meteorology could be employed. However, these may be hard to interpret, and out-of-sample model validation suggests that our flexible model is adequate.

Clarification of the dynamic model defined by (1) and (2) may be helpful. We are modeling true ozone dynamically to suggest that ozone differentials are explained by meteorology differentials. The meteorology differentials are not modeled dynamically. Another approach, arguably more demanding and more open to criticism, would be to build a dynamic weather model and then treat true ozone at time t to be conditionally independent given the weather at time t. Expressed in different terms, for us, $O_l(\mathbf{s}, t-1)$ serves as a proxy for many other unobserved explanatory variables for ozone concentration levels.

The autoregressive models require an initial condition for $O_l(\mathbf{s}, 1)$, the first value in year l. We assume the following model:

$$O_l(\mathbf{s}, 1) = \mu_l + \gamma_l(\mathbf{s}), \tag{3}$$

where $\gamma_l(\mathbf{s})$ is the additional regional effect in year l at site $\mathbf{s}$ over a global level μ_l.

Note that we could instead adopt a "random walk" model for $O_l(\mathbf{s}, t)$, for example,

$$O_l(\mathbf{s}, t) - \mathbf{x}_l'(\mathbf{s}, t)\boldsymbol{\beta} = O_l(\mathbf{s}, t-1) - \mathbf{x}_l'(\mathbf{s}, t-1)\boldsymbol{\beta} + \eta_l(\mathbf{s}, t),$$

that is,

$$O_l(\mathbf{s}, t) = O_l(\mathbf{s}, t-1) + \boldsymbol{\delta}_l'(\mathbf{s}, t-1)\boldsymbol{\beta} + \eta_l(\mathbf{s}, t). \tag{4}$$

This model corresponds to setting $\rho = 1$ in (2) and eliminates the need for the ξ_l's. However, we find the fixing of ρ to be unsatisfactory. [Indeed, model comparison using model choice and validation showed considerably poorer performance for (4) compared with (2).] In essence, the $\rho = 1$ model is nonstationary, yielding prediction/forecasting that is explosive in time. As a noteworthy aside, inference regarding the $\boldsymbol{\beta}$'s is essentially the same in (2) and (4). Intuitively, using $O_l(\mathbf{s}, t-1)$ to explain $O_l(\mathbf{s}, t)$ with a 45° line through the origin or with a more flexible line would not be expected to much affect how the meteorology variables explain $O_l(\mathbf{s}, t)$. Empirically, it is observed in comparing the two fitted models.

A second alternative is to change the right side of (4) to $\rho(O_l(\mathbf{s}, t-1) - \mathbf{x}_l'(\mathbf{s}, t-1)\boldsymbol{\beta})$ in the spirit of (2). However, we can see that this model does not permit us to work with incremental meteorology; it would force us to model the meteorology, which we seek to avoid. (Again, see Sec. 3.1.) So, in the following discussion, we confine ourselves to the specifications in (2) and (3).

Now we write the preceding models using vectors and matrices to facilitate computation. The first model equation is obtained from (1):

$$\mathbf{Z}_{lt} = \mathbf{O}_{lt} + \boldsymbol{\epsilon}_{lt}, \qquad l = 1, \ldots, r, t = 1, \ldots, T, \tag{5}$$

where $\boldsymbol{\epsilon}_{lt} = (\epsilon_l(\mathbf{s}_1, t), \ldots, \epsilon_l(\mathbf{s}_n, t))'$. Let $\mathbf{1}$ be the vector of dimension n with all elements unity and $\boldsymbol{\gamma}_l = (\gamma_l(\mathbf{s}_1), \ldots, \gamma_l(\mathbf{s}_n))'$. From (3) and (2) we have, respectively,

$$\mathbf{O}_{l1} = \boldsymbol{\gamma}_l + \mu_l \mathbf{1}, \qquad l = 1, \ldots, r, \tag{6}$$

$$\mathbf{O}_{lt} = \xi_l \mathbf{1} + \rho \mathbf{O}_{lt-1} + F_{lt}\boldsymbol{\beta} + \boldsymbol{\eta}_{lt}, \qquad l = 1, \ldots, r, t = 2, \ldots, T. \tag{7}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$, $\boldsymbol{\eta}_{lt} = (\eta_l(\mathbf{s}_1, t), \ldots, \eta_l(\mathbf{s}_n, t))'$, and

$$F_{lt} = \begin{pmatrix} \boldsymbol{\delta}_l'(\mathbf{s}_1, t) \\ \boldsymbol{\delta}_l'(\mathbf{s}_2, t) \\ \vdots \\ \boldsymbol{\delta}_l'(\mathbf{s}_n, t) \end{pmatrix} = \begin{pmatrix} \delta_{l1}(\mathbf{s}_1, t) & \delta_{l2}(\mathbf{s}_1, t) & \cdots & \delta_{lp}(\mathbf{s}_1, t) \\ \delta_{l1}(\mathbf{s}_2, t) & \delta_{l2}(\mathbf{s}_2, t) & \cdots & \delta_{lp}(\mathbf{s}_2, t) \\ \vdots & \vdots & \vdots & \vdots \\ \delta_{l1}(\mathbf{s}_n, t) & \delta_{l2}(\mathbf{s}_n, t) & \cdots & \delta_{lp}(\mathbf{s}_n, t) \end{pmatrix}.$$

For the measurement error in (5) we assume that $\boldsymbol{\epsilon}_{lt} \sim \mathrm{N}(\mathbf{0}, \sigma_\epsilon^2 I_n)$, $l = 1, \ldots, r$, $t = 1, \ldots, T$, independently, where $\mathbf{0}$ is the vector with all elements 0 and I_n is the identity matrix of order n. For the spatially correlated error we assume that $\boldsymbol{\eta}_{lt} \sim \mathrm{N}(\mathbf{0}, \Sigma_\eta)$, $l = 1, \ldots, r$, $t = 2, \ldots, T$ independently, where Σ_η has elements $\sigma_\eta(i, j) = \sigma_\eta^2 \rho(\mathbf{s}_i - \mathbf{s}_j; \phi_\eta)$. We take $\rho(\mathbf{s}_i - \mathbf{s}_j; \phi_\eta) = \rho(d_{ij}, \phi_\eta) = \exp(-\phi_\eta d_{ij})$, where d_{ij} is the distance between sites $\mathbf{s}_i$ and $\mathbf{s}_j$, $i, j = 1, \ldots, n$. (The use of an isotropic covariance function for the residual process in an autoregressive, local meteorology-adjusted model seems reasonable. Of course, alternate choices could be examined.) We acknowledge the simplification associated with choosing the exponential covariance structure; however, other members of the Matèrn family of covariance functions can be chosen.

Finally, we assume that $\boldsymbol{\gamma}_l \sim \mathrm{N}(\mathbf{0}, \Sigma_l)$, $l = 1, \ldots, r$ independently, where $\Sigma_l = \sigma_l^2 \Sigma_\gamma$ and Σ_γ has elements $\Sigma_\gamma(i,j) = \rho_\gamma(\mathbf{s}_i - \mathbf{s}_j; \phi_\gamma)$. As before, we assume that $\rho_\gamma(\mathbf{s}_i - \mathbf{s}_j; \phi_\gamma) = \exp(-\phi_\gamma d_{ij})$. The parameters ϕ_η and ϕ_γ are determined using cross-validation as discussed in Section 5.1.

3.1 Specification for $\boldsymbol{\delta}_l(\mathbf{s}, t)$

It is a highly complex problem to model a multidimensional meteorological variable over a large spatial domain for a number of years. Numerical models based on a large number of input parameters are often implemented on a supercomputer to produce many aspects of climate forecasting. It is beyond a reasonable scope for our work to attempt to replicate such climate models, to attempt dynamic modeling of the meteorological variables $\mathbf{x}_l(\mathbf{s}, t)$ at the unobserved sites. Instead, we specify spatially correlated but temporally independent models for the increments $\boldsymbol{\delta}_l(\mathbf{s}, t)$.

In particular, recall that we have only observed the p-dimensional increments in meteorological variables, $\boldsymbol{\delta}_l(\mathbf{s}, t)$, in each year l and on each day t in n_2 sites of which m_1 are CASTNET sites and m_2 weather stations. We order the sites so that the first n_1 are NAMS/SLAMS sites where $\boldsymbol{\delta}_l(\mathbf{s}, t)$ has not been observed, the next m_1 sites are the CASTNET sites, and the last m_2 sites are weather stations.

Based on our exploratory analysis, as mentioned in Section 2, we assume that each of these $\boldsymbol{\delta}_l(\mathbf{s}, t)$ is independently normally distributed with zero mean. The p components of $\boldsymbol{\delta}_l(\mathbf{s}, t)$, however, will have correlation with each other. In addition, we expect them to be spatially associated with spatial decay that may vary with component. Hence, we need to specify and estimate correlation structures between components, $k \neq k' = 1, \ldots, p$, $\delta_{lk}(\mathbf{s}_i, t)$ and $\delta_{lk'}(\mathbf{s}_j, t)$ for any given year l and day t. We assume that the correlation structure is not influenced by the true ozone values (or their transformations), $O_l(\mathbf{s}, t)$. Hence, in order to estimate the parameters describing the correlation structure of $\boldsymbol{\delta}_l(\mathbf{s}, t)$, we only use the observations $\boldsymbol{\delta}_l(\mathbf{s}, t)$ observed at n_2 sites over all the years $l = 1, \ldots, r$ and the days $t = 1, \ldots, T$. We now specify the correlation structure and discuss its estimation.

The correlation structure within the p components of $\boldsymbol{\delta}_l(\mathbf{s}, t)$ at any given $\mathbf{s}$, l, and t can be described, without loss of generality, by a $p \times p$ lower triangular matrix A, say, where $A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p)$. (This is the so-called coregionalization matrix discussed in, e.g., Gelfand, Kim, Sirmans, and Banerjee 2004.) Let $\rho_k(\mathbf{s}_i - \mathbf{s}_j; \phi_k)$ denote the correlation between $\delta_{lk}(\mathbf{s}_i, t)$ and $\delta_{lk}(\mathbf{s}_j, t)$. For convenience, we adopt the exponential covariance structure, that is, $\rho_k(\mathbf{s}_i - \mathbf{s}_j; \phi_k) = \exp(-\phi_k d_{ij})$, $k = 1, \ldots, p$, where d_{ij} is the distance between the sites $\mathbf{s}_i$ and $\mathbf{s}_j$. As a result, we obtain the cross-covariance function between $\boldsymbol{\delta}_l(\mathbf{s}_i, t)$ and $\boldsymbol{\delta}_l(\mathbf{s}_j, t)$, $\mathrm{cov}(\boldsymbol{\delta}_l(\mathbf{s}_i, t), \boldsymbol{\delta}_l(\mathbf{s}_j, t)) \equiv C(\mathbf{s}_i - \mathbf{s}_j) = \sum_{k=1}^p \rho_k(\mathbf{s}_i - \mathbf{s}_j; \phi_k) T_k$, where $T_k = \mathbf{a}_k \mathbf{a}_k'$. Let $\boldsymbol{\delta}_{lt} = ((\boldsymbol{\delta}_{lt}^{(1)})', (\boldsymbol{\delta}_{lt}^{(2)})')'$, where $\boldsymbol{\delta}_{lt}^{(1)} = (\boldsymbol{\delta}_l'(\mathbf{s}_1, t), \ldots, \boldsymbol{\delta}_l'(\mathbf{s}_{n_1}, t))'$ and $\boldsymbol{\delta}_{lt}^{(2)} = (\boldsymbol{\delta}_l'(\mathbf{s}_{n_1+1}, t), \ldots, \boldsymbol{\delta}_l'(\mathbf{s}_{n_1+n_2}, t))'$. The covariance matrix of $\boldsymbol{\delta}_{lt}$ is

$$\Sigma(\boldsymbol{\delta}) = \begin{pmatrix} AA' & C(\mathbf{s}_1 - \mathbf{s}_2) & \cdots & C(\mathbf{s}_1 - \mathbf{s}_{n_1+n_2}) \\ C(\mathbf{s}_2 - \mathbf{s}_1) & AA' & \cdots & C(\mathbf{s}_2 - \mathbf{s}_{n_1+n_2}) \\ \vdots & \vdots & \vdots & \vdots \\ C(\mathbf{s}_{n_1+n_2} - \mathbf{s}_1) & C(\mathbf{s}_{n_1+n_2} - \mathbf{s}_2) & \cdots & AA' \end{pmatrix}.$$

By partitioning,

$$\Sigma(\boldsymbol{\delta}) = \begin{pmatrix} \Sigma_{11}(\boldsymbol{\delta}^{(1)}) & \Sigma_{12}(\boldsymbol{\delta}^{(1)}, \boldsymbol{\delta}^{(2)}) \\ \Sigma_{21}(\boldsymbol{\delta}^{(2)}, \boldsymbol{\delta}^{(1)}) & \Sigma_{22}(\boldsymbol{\delta}^{(2)}) \end{pmatrix}.$$

We have

$$\boldsymbol{\delta}_{lt}^{(2)} \sim \mathrm{N}(\mathbf{0}, \Sigma_{22}), \tag{8}$$

$$\boldsymbol{\delta}_{lt}^{(1)} | \boldsymbol{\delta}_{lt}^{(2)} \sim \mathrm{N}\big(\Sigma_{12}\Sigma_{22}^{-1}\boldsymbol{\delta}_{lt}^{(2)}, \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\big) \tag{9}$$

for $l = 1, \ldots, r$ and $t = 1, \ldots, T$, where we have dropped the arguments for Σ for ease of notation. We also note that from (9) it is easy to obtain the distribution of $\boldsymbol{\delta}_l(\mathbf{s}', t) | \boldsymbol{\delta}_{lt}^{(2)}$ for any arbitrary location $\mathbf{s}'$, which we shall require for prediction purposes in Section 4.

Equation (8) provides the likelihood specification for estimating the elements in the lower triangular matrix A and the parameters ϕ_k, $k = 1, \ldots, p$. Let $\boldsymbol{\nu}$ denote these parameters and let $\mathbf{u}$ denote the observations $\boldsymbol{\delta}_{lt}^{(2)}$, $l = 1, \ldots, r$, $t = 1, \ldots, T$. We assume $\mathrm{N}(0, 10^4)$ for each element in the lower triangle of A and the uniform prior distribution $\mathrm{U}(.001, .1)$ for each ϕ_k. [This range was adequate to capture the rates of decay in $\boldsymbol{\delta}_{lk}(\mathbf{s}, t)$, $k = 1, \ldots, p$.] The likelihood (8) and these prior specifications are used to obtain the posterior distribution of $\boldsymbol{\nu}$ given $\mathbf{u}$.

We run the Metropolis–Hastings algorithm to sample from this posterior distribution of $\boldsymbol{\nu}$ as follows. Let S_{22} denote the sample covariance matrix of order $n_2 p \times n_2 p$ obtained from the $rT = 8 \times 169 = 1{,}352$ realizations $\boldsymbol{\delta}_{lt}^{(2)}$. The starting values for the elements of A are obtained by taking the first 4×4 submatrix of Cholesky decomposition of S_{22}. The starting values of the ϕ parameters are all chosen to be .005. The jump sizes for the Metropolis algorithm are tuned to have 40–50% acceptance rates. For our data the proposal variance of the normal proposal distribution for the elements of A is found to be .5 to have the desired acceptance rate. (The actual observed rate was 47.45%.) The ϕ parameters are sampled on the log scale with a uniform proposal distribution with jump size .02. Note that this algorithm can be run both before or with the main Gibbs sampler for ozone model fitting. Finally, we can compare the model-based estimate (say, the posterior mean) of Σ_{22} with the observed covariance matrix S_{22} to check the quality of model fit. Omitting details, whether we use a histogram of differences or conventional trace or determinant criteria, the suggestion is that the model fits very well. Moreover, we validate the ozone model extensively in Section 5, again producing very satisfactory results.

3.2 Joint Posterior Details

Define $N = nrT$ and $M = nr(T-1)$ and let $\boldsymbol{\vartheta}_{lt} = \xi_l \mathbf{1} + \rho \mathbf{O}_{lt-1} + F_{lt}\boldsymbol{\beta}$ for $l = 1, \dots, r$ and $t = 2, \dots, T$. Further, let $\boldsymbol{\theta}$ denote all the parameters, μ_l, σ_l^2, ξ_l, $l = 1, \dots, r$, $\boldsymbol{\beta}$, ρ, σ_ϵ^2, and σ_η^2. Let $\mathbf{w}$ denote all the augmented data, $\mathbf{o}_{lt}$, $\boldsymbol{\delta}_{lt}^{(1)}$, and the missing data, denoted by $z_l^*(\mathbf{s}_i, t)$, for $i = 1, \dots, n$, $l = 1, \dots, r$, $t = 1, \dots, T$, and let $\mathbf{z}$ denote all the nonmissing data $z_l(\mathbf{s}_i, t)$ for $i = 1, \dots, n$, $l = 1, \dots, r$, $t = 1, \dots, T$. The log of the posterior distribution, denoted by $\log \pi(\boldsymbol{\theta}, \mathbf{w}|\mathbf{u}, \mathbf{z})$, can be written as

$$
\begin{aligned}
&-\frac{N}{2}\log(\sigma_\epsilon^2) - \frac{1}{2\sigma_\epsilon^2}\sum_{l=1}^{r}\sum_{t=1}^{T}(\mathbf{Z}_{lt} - \mathbf{O}_{lt})'(\mathbf{Z}_{lt} - \mathbf{O}_{lt}) \\
&\quad - \frac{M}{2}\log(\sigma_\eta^2) - \frac{1}{2\sigma_\eta^2}\sum_{l=1}^{r}\sum_{t=2}^{T}(\mathbf{O}_{lt} - \boldsymbol{\vartheta}_{lt})'\Sigma_\eta^{-1}(\mathbf{O}_{lt} - \boldsymbol{\vartheta}_{lt}) \\
&\quad - \frac{rT}{2}\log|\Sigma(\boldsymbol{\delta})| - \frac{1}{2}\sum_{l=1}^{r}\sum_{t=1}^{T}\boldsymbol{\delta}_{lt}'\Sigma(\boldsymbol{\delta})^{-1}\boldsymbol{\delta}_{lt} \\
&\quad + \log(\pi(\rho, \boldsymbol{\beta}, \sigma_\epsilon^2, \sigma_\eta^2)) \\
&\quad + \sum_{l=1}^{r}\Bigg[\log(\pi(\mu_l)) + \log(\pi(\sigma_l^2)) + \log(\pi(\xi_l)) \\
&\qquad\qquad - \frac{1}{2}\log|\Sigma_l| - \frac{1}{2}\boldsymbol{\gamma}_l'\Sigma_l^{-1}\boldsymbol{\gamma}_l\Bigg],
\end{aligned}
$$

where $\pi(\mu_l), \pi(\sigma_l^2), \pi(\xi_l)$, and $\pi(\rho, \boldsymbol{\beta}, \sigma_\epsilon^2, \sigma_\eta^2)$ are the prior distributions. We assume that a priori μ_l and ξ_l are independent normally distributed with means 0 and variances 10^4. The autoregressive coefficient ρ is specified the $\mathrm{N}(0, 10^4)I(0 < \rho < 1)$ prior distribution. The Bayesian model specification is completed by the further prior assumptions: $\boldsymbol{\beta} \sim \mathrm{N}(\mathbf{0}, 10^4 I_p)$, $\frac{1}{\sigma_\epsilon^2} \sim \mathrm{G}(a, b)$, $\frac{1}{\sigma_\eta^2} \sim \mathrm{G}(a, b)$, and $\frac{1}{\sigma_l^2} \sim \mathrm{G}(a, b)$, $l = 1, \dots, r$, independently, where the distribution $\mathrm{G}(a, b)$ has mean a/b. In our implementation we take $a = 2$ and $b = 1$ to have a proper prior specification for each of these variance components.

4. PREDICTION DETAILS

4.1 Predicting Ozone at a New Location

Spatial prediction at location $\mathbf{s}'$ and time t' is based on the predictive distribution of $Z_l(\mathbf{s}', t')$ given in the model equations (1), (2), and (3). These models allow us to interpolate the spatial surface at any time point $t' \geq 1$ in a given year. According to (1), for a new location $\mathbf{s}'$ at time t', $Z_l(\mathbf{s}', t')$, has the distribution:

$$Z_l(\mathbf{s}', t') \sim \mathrm{N}(O_l(\mathbf{s}', t'), \sigma_\epsilon^2), \tag{10}$$

where, for $t' = 1$,

$$O_l(\mathbf{s}', 1) = \gamma_l(\mathbf{s}') + \mu_l \tag{11}$$

and for $t' > 1$, $O_l(\mathbf{s}', t') = \rho O_l(\mathbf{s}', t'-1) + \boldsymbol{\delta}_l'(\mathbf{s}', t')\boldsymbol{\beta} + \eta_l(\mathbf{s}', t')$. From this it is clear that $O_l(\mathbf{s}', t')$ can only be sequentially determined using all the previous $O_l(\mathbf{s}', t)$ up to time t'. Hence, we introduce the notation $\mathbf{O}_l(\mathbf{s}, [t])$ to denote the vector $(O_l(\mathbf{s}, 1), \dots, O_l(\mathbf{s}, t))'$ for $t \geq 1$.

The posterior predictive distribution of $Z_l(\mathbf{s}', t')$ is obtained by integrating over the unknown quantities in (10) with respect to the joint posterior distribution, that is,

$$
\begin{aligned}
&\pi(Z_l(\mathbf{s}', t')|\mathbf{u}, \mathbf{z}) \\
&\quad = \int \pi\big(Z_l(\mathbf{s}', t')|O_l(\mathbf{s}', [t']), \sigma_\epsilon^2\big) \\
&\qquad \times \pi\big(O_l(\mathbf{s}', [t'])|\gamma_l(\mathbf{s}'), \boldsymbol{\delta}_l'(\mathbf{s}', t'), \boldsymbol{\theta}, \mathbf{w}\big) \\
&\qquad \times \pi(\boldsymbol{\delta}_l'(\mathbf{s}', t')|\mathbf{u}, \boldsymbol{\nu})\pi(\gamma_l(\mathbf{s}')|\boldsymbol{\theta})\pi(\boldsymbol{\theta}, \mathbf{w}|\mathbf{u}, \mathbf{z})\pi(\boldsymbol{\nu}|\mathbf{u}) \\
&\qquad \times dO_l(\mathbf{s}', [t'])\, d\boldsymbol{\delta}_l'(\mathbf{s}', t')\, d\gamma_l(\mathbf{s}')\, d\boldsymbol{\theta}\, d\mathbf{w}\, d\boldsymbol{\nu}.
\end{aligned}
\tag{12}
$$

When using MCMC methods to draw samples from the posterior, the predictive distribution (12) is sampled by composition; draws from the posterior distributions, $\pi(\boldsymbol{\theta}, \mathbf{w}|\mathbf{u}, \mathbf{z})$ and $\pi(\boldsymbol{\nu}|\mathbf{u})$, enable draws from the preceding component densities. Details are provided in the following discussion.

In (12) we need to generate the random variables $\gamma_l(\mathbf{s}')$, $\boldsymbol{\delta}_l(\mathbf{s}', t')$, and $\mathbf{O}_l(\mathbf{s}', t')$ conditional on the posterior samples at the observed locations $\mathbf{s}_1, \dots, \mathbf{s}_n$ and at the time points $1, \dots, T$. To draw samples from $\boldsymbol{\delta}_l(\mathbf{s}', t')$, we use the conditional distribution $\pi(\boldsymbol{\delta}_l(\mathbf{s}', t')|\mathbf{u}, \boldsymbol{\nu})$. This distribution is similar to (9); see Section 3.1 for more details. Once $\boldsymbol{\delta}_l(\mathbf{s}', t')$ has been drawn we draw $O_l(\mathbf{s}', t')$ from its conditional distribution given all the parameters, data, and $O_l(\mathbf{s}', [t'-1])$. For $t' = 1$ we need to sample $\gamma_l(\mathbf{s}')$ for each l. For this we have

$$\begin{pmatrix} \gamma_l(\mathbf{s}') \\ \boldsymbol{\gamma}_l \end{pmatrix} \sim \mathrm{N}\left[\begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix}, \sigma_l^2 \begin{pmatrix} 1 & \Sigma_{\gamma,12} \\ \Sigma_{\gamma,21} & \Sigma_\gamma \end{pmatrix}\right],$$

where $\Sigma_{\gamma,12}$ is $1 \times n$ with the ith entry given by $\sigma_\gamma(\mathbf{s}_i - \mathbf{s}') = \exp(-\phi_\gamma d(\mathbf{s}_i, \mathbf{s}'))$, where $d(\mathbf{s}_i, \mathbf{s}')$ is the distance between the sites $\mathbf{s}_i$ and $\mathbf{s}'$ and $\Sigma_{\gamma,21} = \Sigma_{\gamma,12}'$. Therefore,

$$\gamma_l(\mathbf{s}')|\boldsymbol{\theta} \sim \mathrm{N}\big(\Sigma_{\gamma,12}\Sigma_\gamma^{-1}\boldsymbol{\gamma}_l, \sigma_l^2(1 - \Sigma_{\gamma,12}\Sigma_\gamma^{-1}\Sigma_{\gamma,21})\big). \tag{13}$$

Analogous to (7), we obtain, for $t' > 1$,

$$
\begin{aligned}
\begin{pmatrix} O_l(\mathbf{s}', t') \\ \mathbf{O}_{lt'} \end{pmatrix} \sim \mathrm{N}\Bigg[&\begin{pmatrix} \xi_l + \rho O_l(\mathbf{s}', t'-1) + \boldsymbol{\delta}_l'(\mathbf{s}', t')\boldsymbol{\beta} \\ \xi_l \mathbf{1} + \rho \mathbf{O}_{lt'-1} + F_{lt'}\boldsymbol{\beta} \end{pmatrix}, \\
&\sigma_\eta^2 \begin{pmatrix} 1 & \Sigma_{\eta,12} \\ \Sigma_{\eta,21} & \Sigma_\eta \end{pmatrix}\Bigg],
\end{aligned}
$$

where $\Sigma_{\eta,12}$ is $1 \times n$ with the ith entry given by $\sigma_\eta(\mathbf{s}_i - \mathbf{s}') = \exp(-\phi_\eta d(\mathbf{s}_i, \mathbf{s}'))$, where $d(\mathbf{s}_i, \mathbf{s}')$ is the distance between the sites $\mathbf{s}_i$ and $\mathbf{s}'$ and $\Sigma_{\eta,21} = \Sigma_{\eta,12}'$. Hence,

$$O_l(\mathbf{s}', t')|\gamma_l(\mathbf{s}'), \boldsymbol{\delta}_l'(\mathbf{s}', t'), \mathbf{O}_{lt'}, \boldsymbol{\theta}, \mathbf{w} \sim \mathrm{N}(\chi, \Lambda), \tag{14}$$

where $\Lambda = \sigma_\eta^2(1 - \Sigma_{\eta,12}\Sigma_\eta^{-1}\Sigma_{\eta,21})$ and

$$
\begin{aligned}
\chi = {}& \xi_l + \rho O_l(\mathbf{s}', t'-1) + \boldsymbol{\delta}_l'(\mathbf{s}', t')\boldsymbol{\beta} \\
&+ \Sigma_{\eta,12}\Sigma_\eta^{-1}(\mathbf{O}_{lt'} - \xi_l \mathbf{1} - \rho \mathbf{O}_{lt'-1} - F_{lt'}\boldsymbol{\beta}).
\end{aligned}
$$

In summary, we implement the following algorithm to predict $Z_l(\mathbf{s}', t')$.

1. Draw a sample $\boldsymbol{\theta}^{(j)}, \boldsymbol{\nu}^{(j)}, j \geq 1$, from the posterior distribution.
2. Draw $\gamma_l^{(j)}(\mathbf{s}')$ using (13).
3. Draw $\boldsymbol{\delta}_l^{(j)}(\mathbf{s}', t')$ from the distribution $\pi(\boldsymbol{\delta}_l(\mathbf{s}', t')|\mathbf{u}, \boldsymbol{\nu}^{(j)})$.

4. Draw $\mathbf{O}_l^{(j)}(\mathbf{s}, [t'])$ sequentially; that is, first obtain $O_l^{(j)}(\mathbf{s}', 1)$ from (11) and then draw $O_l^{(j)}(\mathbf{s}', 2)$ using (14) and iterate.
5. Finally draw $Z_l^{(j)}(\mathbf{s}', t')$ from $\mathrm{N}(O_l^{(j)}(\mathbf{s}', t'), \sigma_\epsilon^{2(j)})$.

The ozone concentration on the original scale is the square of $Z_l^{(j)}(\mathbf{s}', t')$. If we want the predictions of the smooth ozone concentration process without the nugget term, we simply omit the last step in the preceding algorithm and square the realizations $\mathbf{O}_l^{(j)}(\mathbf{s}, t')$. We use the median of the MCMC samples and the lengths of the 95% intervals to summarize the predictions. The median as a summary measure preserves the one-to-one relationships between summaries for O and Z and for O^2 and Z^2.

4.2 Ozone Summaries

We now develop a methodology for assessing trends in ozone summaries. We investigate these trends using the true ozone process $O_l(\mathbf{s}, t)$. Recall that we model ozone levels on the square root scale; hence, to return to the original scale, we use $O_l^2(\mathbf{s}, t)$ where appropriate.

The true annual fourth highest daily maximum 8-hour average ozone concentration, denoted by $f_l(\mathbf{s})$, is given by the fourth highest value of the series $O_l^2(\mathbf{s}, 1), \dots, O_l^2(\mathbf{s}, T)$ in any given year $l, l = 1, \dots, r$. The summaries of the posterior predictive realizations $f_l^{(j)}(\mathbf{s}), j \ge 1$, are used for predictions of the annual fourth highest daily maximum 8-hour average ozone concentration (and to obtain their uncertainties).

The 3-year rolling average of the annual fourth highest daily maximum 8-hour average ozone concentration is obtained by averaging $f_l(\mathbf{s})$ over 3 successive years and assigning the average to the final year of averaging. Thus, the 3-year rolling average of the annual fourth highest daily maximum 8-hour average ozone concentration in year l is given by

$$g_l(\mathbf{s}) = \frac{f_{l-2}(\mathbf{s}) + f_{l-1}(\mathbf{s}) + f_l(\mathbf{s})}{3}, \qquad l = 3, \dots, r.$$

Again we obtain posterior predictive samples $g_l^{(j)}(\mathbf{s})$ from MCMC iterations and thereby get the prediction summary values along with their uncertainties. We can also estimate the probability of nonattainment at a site $\mathbf{s}$ and in year l, denoted by $P(g_l(\mathbf{s}) > 80)$, by averaging the indicator functions $I(g_l^{(j)}(\mathbf{s}) > 80)$ over j.

We can obtain the meteorology-adjusted levels only for the sites where we have observed the values $\mathbf{x}_l(\mathbf{s}, t)$ of the meteorological variables. These levels are obtained from the residuals $O_l(\mathbf{s}, t) - \mathbf{x}_l'(\mathbf{s}, t)\boldsymbol{\beta}$ using

$$h_l(\mathbf{s}) = \frac{1}{T}\sum_{t=1}^{T}\{O_l(\mathbf{s}, t) - \mathbf{x}_l'(\mathbf{s}, t)\boldsymbol{\beta}\}^2, \qquad l = 1, \dots, r.$$

The posterior predictive realizations $h_l^{(j)}(\mathbf{s})$ are summarized to obtain the adjusted levels at site $\mathbf{s}$ in year l. Note that $O_l(\mathbf{s}, t) - \mathbf{x}_l'(\mathbf{s}, t)\boldsymbol{\beta}$ is the natural definition of the locally adjusted level although, in the presence of $O_l(\mathbf{s}, t-1)$ as in (2), this become less clear. However, because the $\boldsymbol{\beta}$'s obtained under the model in (2) and (3) are very similar to those obtained under the model in (4) and because adjusting for meteorology in (4) immediately takes the natural form, we propose the use of this summary. The unadjusted levels are given by $u_l(\mathbf{s}) = \frac{1}{T}\sum_{t=1}^{T} O_l^2(\mathbf{s}, t)$, $l = 1, \dots, r$.

We evaluate relative percentage change between 1997 and 2004 as

$$c_{97,04}^{\text{adj}}(\mathbf{s}) = 100 \times \frac{h_8(\mathbf{s}) - h_1(\mathbf{s})}{h_1(\mathbf{s})}$$

and

$$c_{97,04}^{\text{unadj}}(\mathbf{s}) = 100 \times \frac{u_8(\mathbf{s}) - u_1(\mathbf{s})}{u_1(\mathbf{s})}.$$

Again averaging over posterior predictive realizations produces the desired inference. Percentage change for any other pair of years can be handled similarly.

5. ANALYSIS

5.1 Model Checking

Under weak prior distributions it is not possible to estimate all the parameters in the covariance structure, σ_ϵ^2, σ_η^2, ϕ_η, and ϕ_γ, consistently; see, for example, Zhang (2004), Sahu et al. (2006), and the references therein. Hence, we use the set-aside validation data from 15 stations to select the two decay parameters ϕ_η and ϕ_γ. The variance components are estimated using MCMC. Let $\hat{Z}_l^2(\mathbf{s}_i^*, t)$ denote the model-based validation estimate for $Z_l^2(\mathbf{s}_i^*, t)$, where $\mathbf{s}_i^*$ denotes the ith validation site. Again recall that we model ozone on the square root scale. The validation mean squared error (VMSE) is given by

$$\text{VMSE} = \frac{1}{n_v}\sum_{i=1}^{15}\sum_{l=1}^{r}\sum_{t=1}^{T}\big(Z_l^2(\mathbf{s}_i^*, t) - \hat{Z}_l^2(\mathbf{s}_i^*, t)\big)^2 I(Z_l(\mathbf{s}_i^*, t)),$$

where $I(Z_l(\mathbf{s}_i^*, t)) = 1$ if $Z_l(\mathbf{s}_i^*, t)$ has been observed and 0 otherwise, and n_v is the total number of available observations at the 15 validation sites. For our dataset $n_v = 5{,}991$ as mentioned in Section 2. We searched for the optimal values in a two-dimensional grid composed of the values .004, .005, .01, and .05. The pair of values $\phi_\eta = .005$ and $\phi_\gamma = .05$, provided the smallest estimated VMSE. The VMSE increases hugely if the values of ϕ_η and ϕ_γ are interchanged. However, the VMSE is not sensitive to the choice of the decay parameters near these best values. As a result, although it is possible to further refine the grid in a neighborhood of the best value, we do not explore beyond our grid here.

As mentioned previously we have performed validation for all 5,991 available observations in the 15 hold-out sites. Overall, 94.73% of the 95% prediction intervals contain the actual observations, and about 50.4% of the predictions are greater than the actual observations. Figure 4 shows the validation plot for a randomly chosen site. To enhance readability, we only show the validations for 1 out of every 14 days. The validations indicate that the model does not appear to introduce any bias in prediction and performs very well for out-of-sample predictions. The predicted annual surfaces discussed in the next section also validate the model.

We have performed the usual model diagnostics for checking model adequacy using various residual plots. For example, the fitted versus residual plot (not shown) does not show any curvature or pattern and confirms that the homoscedasticity assumption is acceptable; there were only a few extreme values.

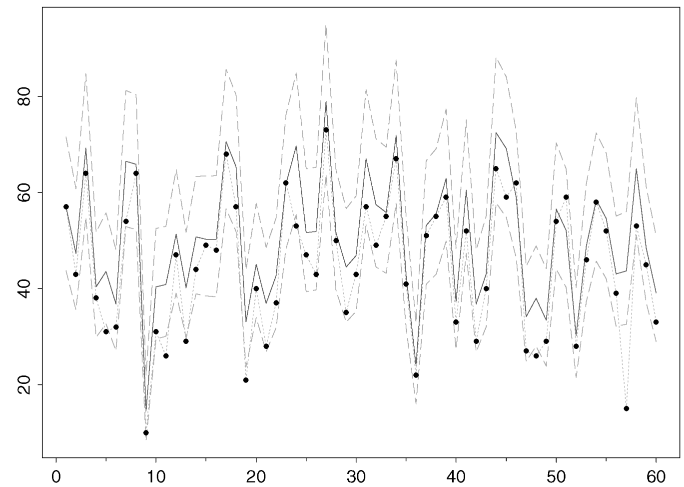

Figure 4. Validation plot for a randomly chosen hold-out site. The observed data are plotted as points. The validation predictions are plotted as a solid line, and the 95% equal-tailed prediction intervals are plotted as a broken line.

The residuals plotted against the year (not shown) do not reveal any anomalies. The variability of the residuals for different years are approximately constant, and, in fact, the ratio of the maximum to the minimum variance is less than 2.

5.2 Results and Interpretation

The point and interval estimates of the model parameters are given in Tables 1 and 2. We found strong dependence among successive-day ozone concentrations (estimate of $\rho = .7783$). Except for wind speed, all meteorological variables were found to be significantly related to ozone concentrations. The estimates of the variance components σ_ϵ^2 and σ_η^2 show that more variation is explained by the spatiotemporal effects than by the pure error process $\epsilon_l(\mathbf{s}, t)$.

The estimates of $\mu_1, \ldots, \mu_8$ (see Table 2) show that changes in the global ozone level as defined in our model are similar to that in Figure 2. Due to the inclusion of the autoregressive term $\rho O_l(\mathbf{s}, t-1)$ for $t > 1$ in (2), the estimates of $\xi_1, \ldots, \xi_8$ (see Table 2) do not show this pattern. The estimates of $\sigma_1^2, \ldots, \sigma_8^2$ (see Table 2) capture (significantly) differing levels of variability between years. The ratio of the maximum variance in 1998 to the minimum in 2001 is more than 3; this is also evident in the boxplots provided in Figure 2.

We now summarize the different types of trend information that can be realized from this modeling approach as described in Section 4.2. The annual fourth highest daily maximum true ozone values are plotted by linearly interpolating the predictions at 289 gridded locations in Ohio (see Fig. 5) with the observed fourth highest daily maxima at the monitoring sites superimposed on the predictive surface. We find excellent agreement among the predicted and observed maximum values. Quantification of this agreement can be found by calculating the root mean squared error (RMSE) between the observations and predictions closest to the monitoring sites. The RMSEs, in units of ppb, for the years 1997–2004 are 3.4, 5.1, 4.9, 4.3, 3.7, 4.6, 5.0, and 4.5, respectively. Thus, the model is predicting the maxima within a range of 3–5 ppb on average. Figure 6 shows the lengths of the 95% prediction intervals. As expected, these intervals are larger in nonmonitored areas compared with monitored areas. The majority of NOx and volatile organic compound (VOC) emissions in the eastern United States come from three sources: mobile sources, industrial processes, and large electric utilities. Mobile sources and electric utilities accounted for 78% of annual NOx emissions in 2004; see U.S. Environmental Protection Agency (2005). From 1997 to 2004, annual NOx emissions have decreased by 25% in the eastern United States, and similarly VOC emissions have decreased by 21%. Figure 5 shows decreasing patterns of true ozone levels across time that might be attributed to reduced levels of ozone precursor emissions.

Table 1. Estimation of the parameters

Parameter	Mean	Standard deviation	95% CI
ρ	.7783	.0030	(.7723, .7842)
β_1 (temperature)	.1069	.0021	(.1029, .1109)
β_2 (humidity)	−.0126	.0004	(−.0134, −.0118)
β_3 (wind speed a.m.)	−.0025	.0030	(−.0083, .0033)
β_3 (wind speed p.m.)	−.0120	.0025	(−.0170, −.0074)
σ_ϵ^2	.0486	.0007	(.0460, .0487)
σ_η^2	.3235	.0046	(.3149, .3326)

NOTE: CI stands for equal-tailed credible intervals.

Model-based interpolated maps of the 3-year rolling averages of the annual fourth highest daily maximum 8-hour true ozone concentrations (Fig. 7) are given for the years 1999–2004. We define the year 1999 as the rolling average for 1997–1999, and similarly for the other rolling averages. As with the annual patterns, we find good agreement with the data superimposed on the plot. The RMSEs are 3.7, 4.2, 3.6, 3.7, 3.6, and 3.5, respectively. These RMSEs show somewhat better predictions for the 3-year averages in comparison to the annual maxima. Also, these patterns reflect the reduced emission levels over this time period. The increase in the 3-year rolling averages that include 2002 can be attributed to the above-normal ozone-forming conditions for that year. In 2002, temperatures were above normal and precipitation was below normal in the Northeast. These maps of true ozone concentrations (O in our model) suggest regions of nonattainment with the current NAAQS for ozone of 80 ppb based on the 3-year rolling averages. Uncertainty in this inference is given by the lengths of the prediction intervals (Fig. 8). To quantify the extent of nonattainment across Ohio, we developed maps of the probabilities of exceeding the ozone NAAQS. Using a nominal probability level of .8, almost all of Ohio was found to exceed this probability level for all rolling averages except for 2004 where we see improved air quality conditions in southern Ohio (Fig. 9).

Trends in meteorology-adjusted ozone predictions at 12 monitoring sites in Ohio, along with trends in unadjusted predictions, are shown in Figure 10. Again we see high ozone unadjusted predictions in 2002, in comparison to the smoother adjusted predictions at and around this time period. From the adjusted predictions, we see an overall decreasing pattern in ozone that is not easily discerned in the plot of the unadjusted predictions. Figure 11 illustrates the spatial pattern of trend at monitoring sites defined as a relative difference (%) of adjusted and unadjusted predictions for 1997–2004. These

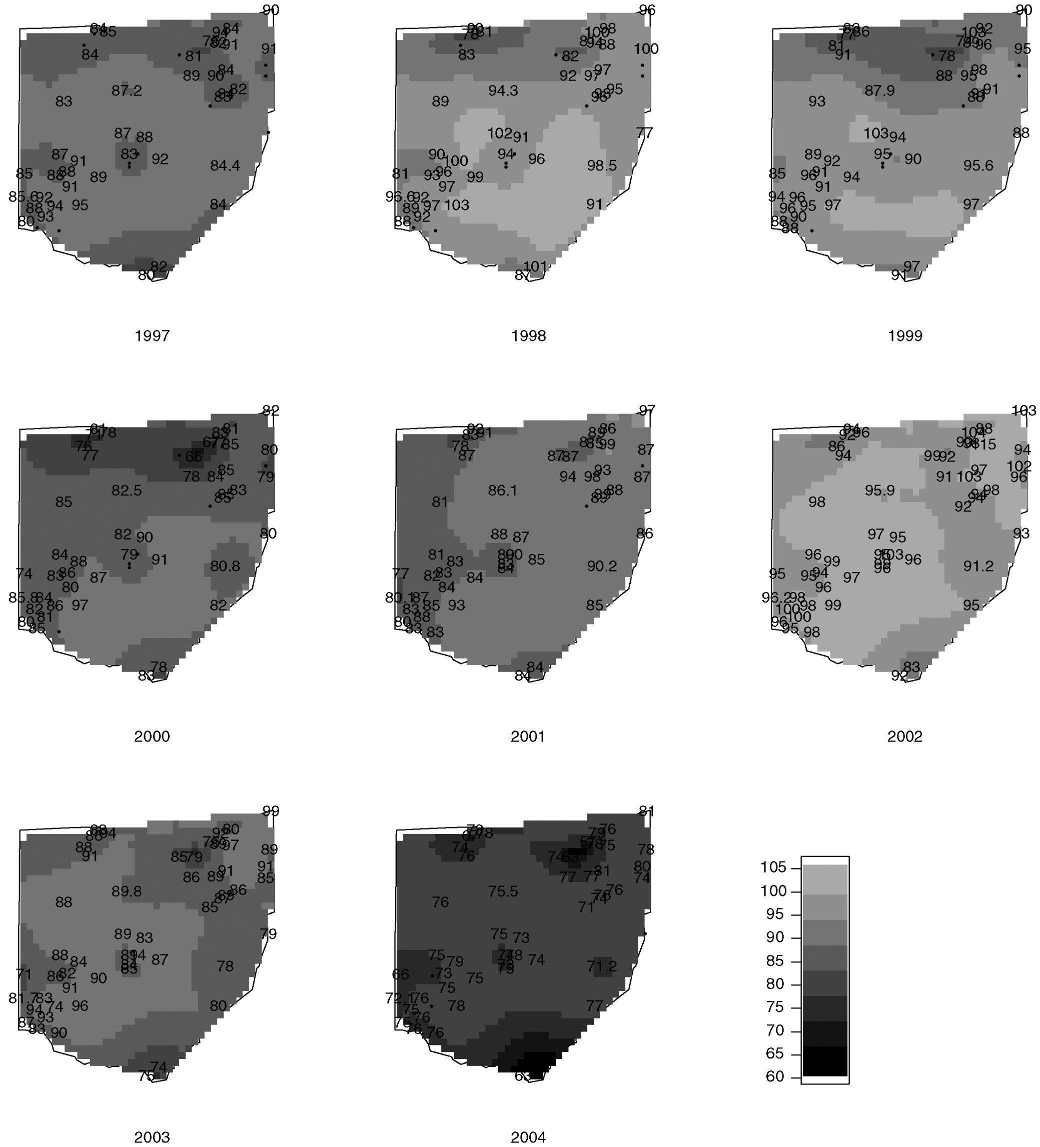

Figure 5. Model-based interpolation of the true annual fourth highest maximum ozone levels for 8 years. Observed data are superimposed on the plots.

trends are all negative, some significant, according to the 95% predictive intervals based on the MCMC replications. Given our definition of trend, we see more significant reductions based on the meteorology-adjusted ozone predictions. Only 6 of 12 sites show significant reductions on the unadjusted scale, while 11 out of 12 show significant reductions on the adjusted scale. Although, from a public health perspective, all that matters is realized ozone concentration levels, clarifica-

Table 2. Estimation of μ_l, σ_l^2, and ξ_l, $l = 1, \ldots, 8$

μ_l			σ_l^2			ξ_l		
Mean	Standard deviation	95% CI	Mean	Standard deviation	95% CI	Mean	Standard deviation	95% CI
7.30	.10	(7.10, 7.51)	.24	.06	(.15, .37)	1.52	.04	(1.45, 1.59)
6.28	.14	(5.98, 6.56)	.42	.10	(.27, .65)	1.61	.04	(1.54, 1.68)
6.18	.14	(5.88, 6.44)	.40	.09	(.25, .61)	1.60	.04	(1.52, 1.66)
7.17	.09	(6.98, 7.35)	.20	.05	(.12, .31)	1.53	.04	(1.46, 1.60)
6.55	.08	(6.39, 6.70)	.13	.03	(.08, .20)	1.58	.04	(1.51, 1.65)
7.35	.09	(7.17, 7.54)	.20	.05	(.12, .30)	1.63	.04	(1.56, 1.70)
8.42	.08	(8.26, 8.58)	.15	.04	(.10, .23)	1.52	.04	(1.46, 1.59)
7.09	.11	(6.88, 7.32)	.28	.06	(.18, .42)	1.48	.04	(1.41, 1.55)

NOTE: CI stands for equal-tailed credible intervals.

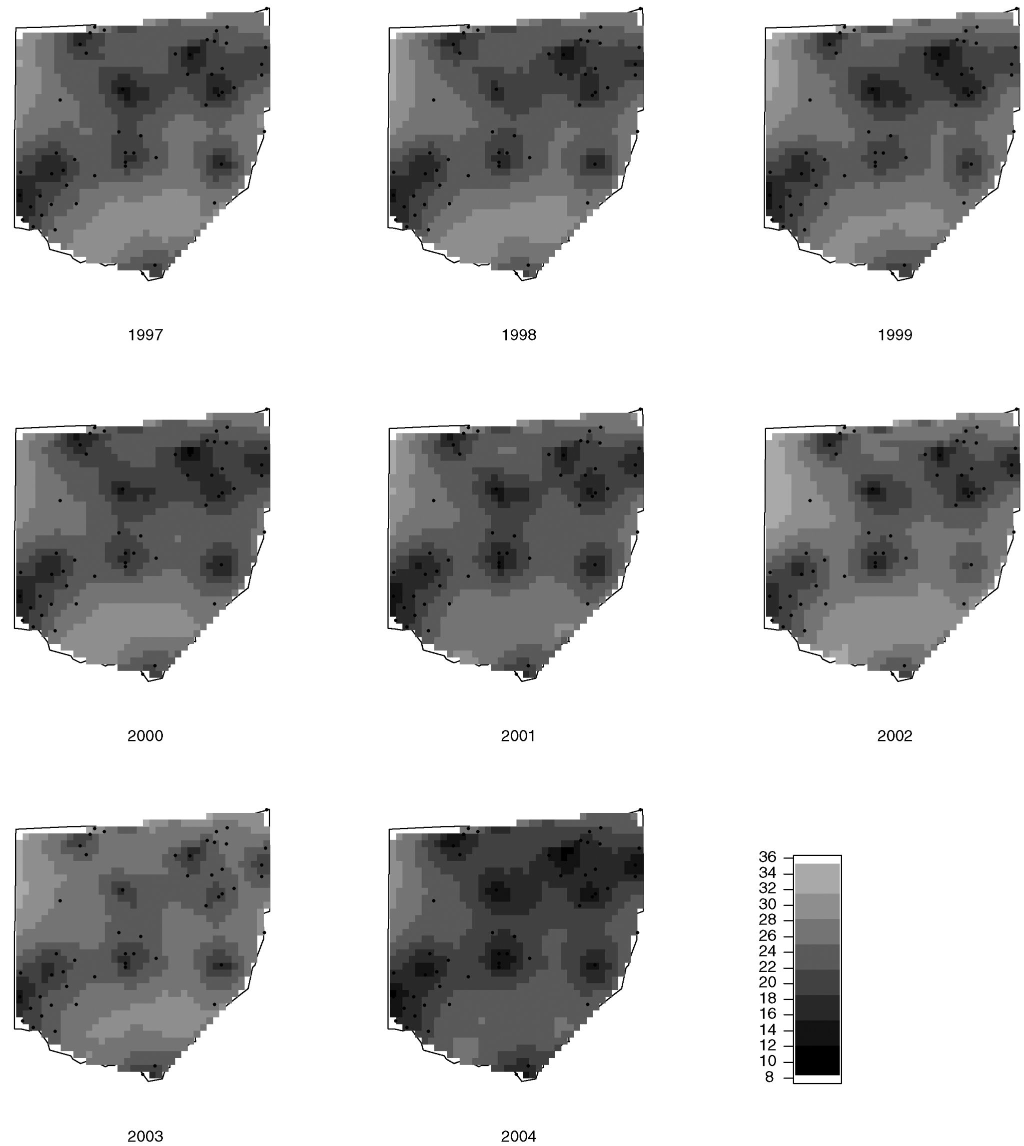

Figure 6. Lengths of 95% intervals of the true annual fourth highest maximum ozone levels for 8 years.

tion of trends in the nonmeteorological component is informative.

6. DISCUSSION

We have formulated a model for assessing ozone levels at point-level spatial resolution and daily temporal resolution. We have shown how to use this model to perform standard prediction but, more interestingly, to provide summaries of annual fourth highest daily average ozone levels and also summaries in the form of 3-year rolling averages of these fourth highest daily averages. Moreover, we can attach uncertainty to all of these predictions, derived from the model fitting. We are also able to demonstrate the benefit of fitting models when interpolating extremes as opposed to interpolating the observations themselves. This contrasts with the case for summarizing averages.

As the U.S. EPA continues with its ozone control program, it will be necessary to further refine and update statistical analyses of trend. In future work, we plan to investigate spatially varying coefficients (see, e.g., Gelfand et al. 2003) in the incremental meteorology model, imagining that the effect of different meteorological variables might be different for different parts of the state. We also plan to extend the analysis to, at the least, the eastern portion of the United States. This will dramatically increase the number of sites for both ozone measurements and meteorology data. (The current article handles 142,543 observations altogether.) Approximate computation will be required. In particular, approximate process representations (see, e.g., Xia and Gelfand 2006, or Paciorek 2007) will be employed.

DISCLAIMER

The U.S. Environmental Protection Agency's Office of Research and Development partially collaborated in the research described here. Although it has been reviewed by the EPA and approved for publication, it does not necessarily reflect the agency's policies or views.

APPENDIX: DISTRIBUTIONS FOR GIBBS SAMPLING

Conditional Distributions for σ_ϵ^2, σ_η^2, $\mathbf{O}_{lt}$, ρ, and $\boldsymbol{\beta}$

Any missing value $Z_l(\mathbf{s}, t)$ is to be sampled from $\mathrm{N}(O_l(\mathbf{s}, t), \sigma_\epsilon^2)$, $l = 1, \ldots, r$, $t = 1, \ldots, T$. Straightforward calculation yields the fol-

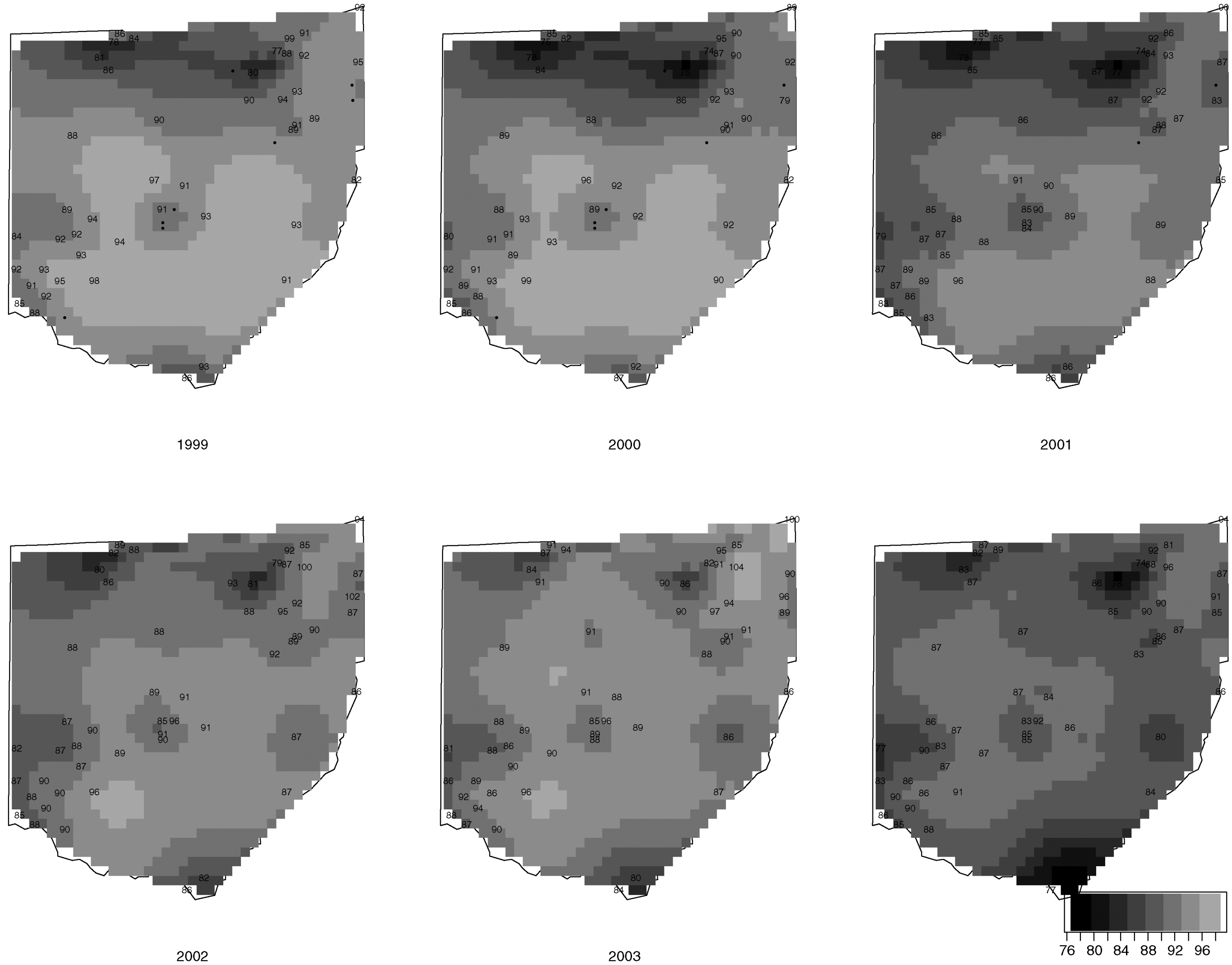

Figure 7. Model-based interpolation of the 3-year rolling averages of the true annual fourth highest maximum ozone levels. Observed data are superimposed.

lowing complete conditional distributions:

$$\frac{1}{\sigma_\epsilon^2} \sim \mathrm{G}\left(\frac{N}{2}+a, b+\frac{1}{2}\sum_{l=1}^{r}\sum_{t=1}^{T}(\mathbf{Z}_{lt}-\mathbf{O}_{lt})'(\mathbf{Z}_{lt}-\mathbf{O}_{lt})\right),$$

$$\frac{1}{\sigma_\eta^2} \sim \mathrm{G}\left(\frac{M}{2}+a, b+\frac{1}{2}\sum_{l=1}^{r}\sum_{t=2}^{T}(\mathbf{O}_{lt}-\boldsymbol{\vartheta}_{lt})'\Sigma_\eta^{-1}(\mathbf{O}_{lt}-\boldsymbol{\vartheta}_{lt})\right).$$

Note that we do not need to sample from $\mathbf{O}_{l1}$ because we have the identity (6). Let $Q_\eta = \Sigma_\eta^{-1}$. The full conditional distribution of $\mathbf{O}_{lt}$ is $\mathrm{N}(\Lambda_{lt}\boldsymbol{\chi}_{lt}, \Lambda_{lt})$, where

$$\Lambda_{lt}^{-1} = \frac{I_n}{\sigma_\epsilon^2} + (1+\rho^2)Q_\eta,$$

$$\boldsymbol{\chi}_{lt} = \frac{\mathbf{Z}_{lt}}{\sigma_\epsilon^2} + Q_\eta\{\rho\mathbf{O}_{lt-1} + F_{lt}\boldsymbol{\beta} + \xi_l\mathbf{1} + \rho(\mathbf{O}_{lt+1} - F_{lt+1}\boldsymbol{\beta} - \xi_l\mathbf{1})\},$$

when $1 < t < T$, and, for $t = T$,

$$\Lambda_{lt}^{-1} = \frac{I_n}{\sigma_\epsilon^2} + Q_\eta, \qquad \boldsymbol{\chi}_{lt} = \frac{\mathbf{Z}_{lt}}{\sigma_\epsilon^2} + Q_\eta(\xi_l\mathbf{1} + \rho\mathbf{O}_{lt-1} + F_{lt}\boldsymbol{\beta}).$$

The full conditional distribution of ξ_l is $\mathrm{N}(\Lambda\chi, \Lambda)$, where

$$\Lambda^{-1} = (T-1)\mathbf{1}'Q_\eta\mathbf{1} + 10^{-4},$$

$$\chi = \sum_{t=2}^{T}\mathbf{1}'Q_\eta(\mathbf{O}_{lt} - \rho\mathbf{O}_{lt-1} - F_{lt}\boldsymbol{\beta}).$$

The full conditional distribution of ρ is $\mathrm{N}(\Lambda\chi, \Lambda)$, where

$$\Lambda^{-1} = \sum_{l=1}^{r}\sum_{t=2}^{T}\mathbf{O}_{lt-1}'Q_\eta\mathbf{O}_{lt-1} + 10^{-4},$$

$$\chi = \sum_{l=1}^{r}\sum_{t=2}^{T}\mathbf{O}_{lt-1}'Q_\eta(\mathbf{O}_{lt} - \xi_l\mathbf{1} - F_{lt}\boldsymbol{\beta}),$$

restricted in the interval (0, 1). The full conditional distribution of $\boldsymbol{\beta}$ is $\mathrm{N}(\Lambda\boldsymbol{\chi}, \Lambda)$, where

$$\Lambda^{-1} = \sum_{l=1}^{r}\sum_{t=2}^{T}F_{lt}'Q_\eta F_{lt} + 10^{-4}I_p,$$

$$\boldsymbol{\chi} = \sum_{l=1}^{r}\sum_{t=2}^{T}F_{lt}'Q_\eta(\mathbf{O}_{lt} - \xi_l\mathbf{1} - \rho\mathbf{O}_{lt-1}).$$

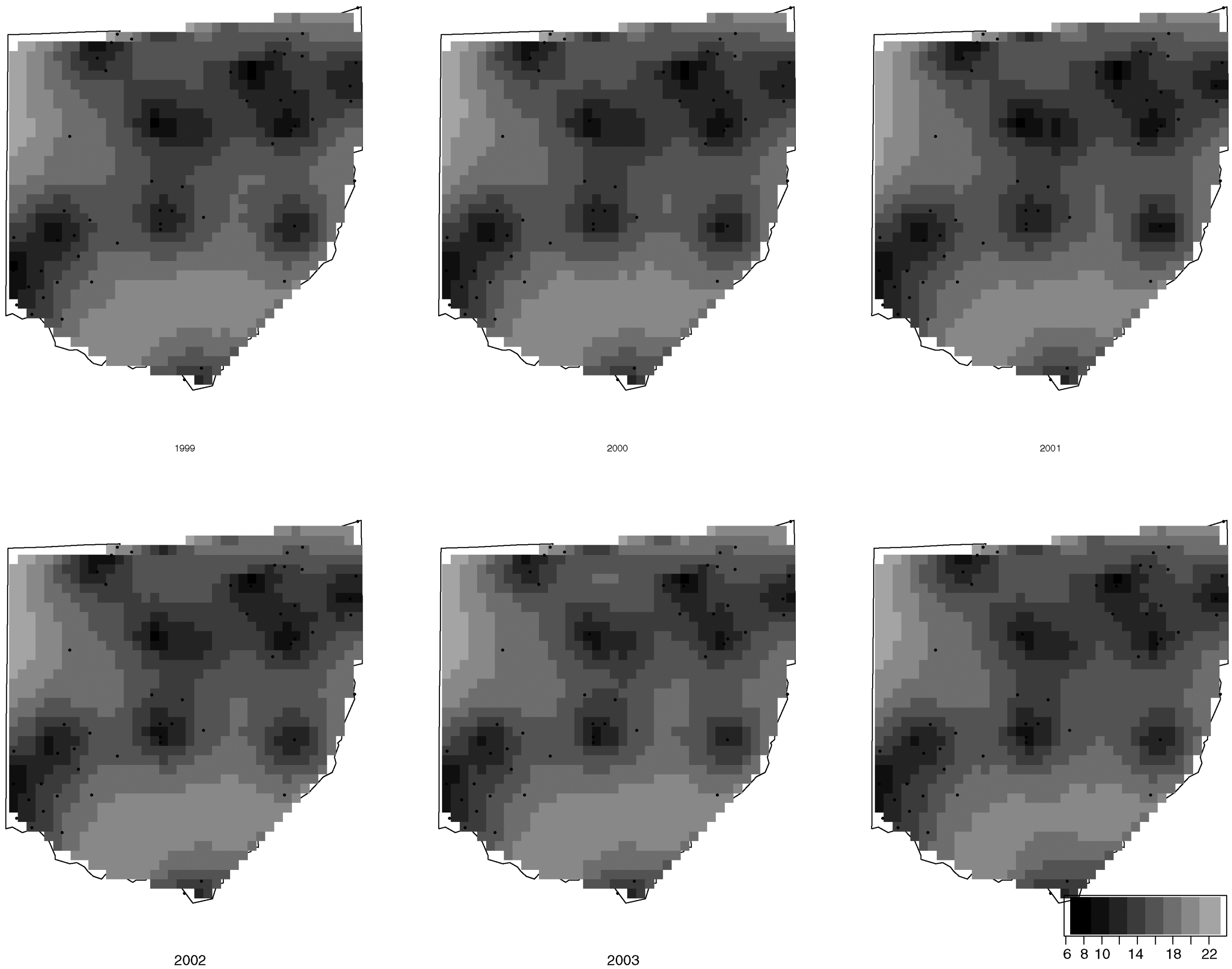

Figure 8. Lengths of 95% intervals of the 3-year rolling averages of the true annual fourth highest maximum ozone levels.

Conditional Distribution of $\boldsymbol{\delta}_{lt}^{(1)}$

We obtain the likelihood contribution for $\boldsymbol{\delta}(\mathbf{s}_i, t)$, $i = 1, \ldots, n_1$, as follows. We have

$$F_{lt}\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}'\boldsymbol{\delta}_l(\mathbf{s}_1, t) \\ \boldsymbol{\beta}'\boldsymbol{\delta}_l(\mathbf{s}_2, t) \\ \vdots \\ \boldsymbol{\beta}'\boldsymbol{\delta}_l(\mathbf{s}_n, t) \end{pmatrix}$$

$$= \begin{pmatrix} \boldsymbol{\beta}' & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\beta}' & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\beta}' \end{pmatrix} \begin{pmatrix} \boldsymbol{\delta}_l(\mathbf{s}_1, t) \\ \boldsymbol{\delta}_l(\mathbf{s}_2, t) \\ \vdots \\ \boldsymbol{\delta}_l(\mathbf{s}_n, t) \end{pmatrix}$$

$$= \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \boldsymbol{\delta}_{lt},$$

where X_1 is $n_1 \times n_1 p$ and X_2 is $(n - n_1) \times (n - n_1)p$. Let $\boldsymbol{\delta}_{lt}^{(12)} = (\boldsymbol{\delta}_l'(\mathbf{s}_{n_1+1}, t), \boldsymbol{\delta}_l'(\mathbf{s}_{n_1+2}, t), \ldots, \boldsymbol{\delta}_l'(\mathbf{s}_n, t))'$. Let us partition Q_η as follows:

$$Q_\eta = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix},$$

where Q_{11} is $n_1 \times n_1$ and Q_{22} is $n_2 \times n_2$ and we suppress the symbol η for convenience. Define $\mathbf{a}_{lt} = \mathbf{O}_{lt} - \xi_l \mathbf{1} - \rho \mathbf{O}_{lt-1}$ and partition $\mathbf{a}_{lt} = ((\mathbf{a}_{lt}^{(1)})', (\mathbf{a}_{lt}^{(2)})')'$, where $\mathbf{a}_{lt}^{(1)}$ is $n_1 p \times 1$. Now

$$\begin{aligned} &(\mathbf{a}_{lt} - F_{lt}\boldsymbol{\beta}) Q_\eta (\mathbf{a}_{lt} - F_{lt}\boldsymbol{\beta}) \\ &\quad = \begin{pmatrix} \mathbf{a}_{lt}^{(1)} - X_1 \boldsymbol{\delta}_{lt}^{(1)} \\ \mathbf{a}_{lt}^{(2)} - X_2 \boldsymbol{\delta}_{lt}^{(2)} \end{pmatrix}' Q_\eta \begin{pmatrix} \mathbf{a}_{lt}^{(1)} - X_1 \boldsymbol{\delta}_{lt}^{(1)} \\ \mathbf{a}_{lt}^{(2)} - X_2 \boldsymbol{\delta}_{lt}^{(2)} \end{pmatrix} \\ &\quad = \big(\mathbf{a}_{lt}^{(1)} - X_1 \boldsymbol{\delta}_{lt}^{(1)}\big)' Q_{11} \big(\mathbf{a}_{lt}^{(1)} - X_1 \boldsymbol{\delta}_{lt}^{(1)}\big) \\ &\qquad + 2\big(\mathbf{a}_{lt}^{(1)} - X_1 \boldsymbol{\delta}_{lt}^{(1)}\big)' Q_{12} \big(\mathbf{a}_{lt}^{(2)} - X_2 \boldsymbol{\delta}_{lt}^{(2)}\big) + C, \end{aligned}$$

where C is free of $\boldsymbol{\delta}_{lt}^{(1)}$. Now from (9) we have

$$\begin{aligned} \boldsymbol{\delta}_{lt}^{(1)} | \boldsymbol{\delta}_{lt}^{(2)} &\sim \mathrm{N}\big(\Sigma_{12}\Sigma_{22}^{-1}\boldsymbol{\delta}_{lt}^{(2)}, \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\big) \\ &\equiv \mathrm{N}(\boldsymbol{\zeta}_{lt}, \Sigma_\delta), \quad \text{say.} \end{aligned}$$

Thus, the conditional posterior distribution of $\boldsymbol{\delta}_{lt}^{(1)}$ is $\mathrm{N}(\Lambda \boldsymbol{\chi}_{lt}, \Lambda)$, where

$$\Lambda^{-1} = \Sigma_\delta^{-1} + X_1' Q_{11} X_1$$

and

$$\boldsymbol{\chi}_{lt} = \Sigma_\delta^{-1}\boldsymbol{\zeta}_{lt} + X_1'\big\{Q_{11}\mathbf{a}_{lt}^{(1)} + Q_{12}\big(\mathbf{a}_{lt}^{(2)} - X_2 \boldsymbol{\delta}_{lt}^{(12)}\big)\big\}.$$

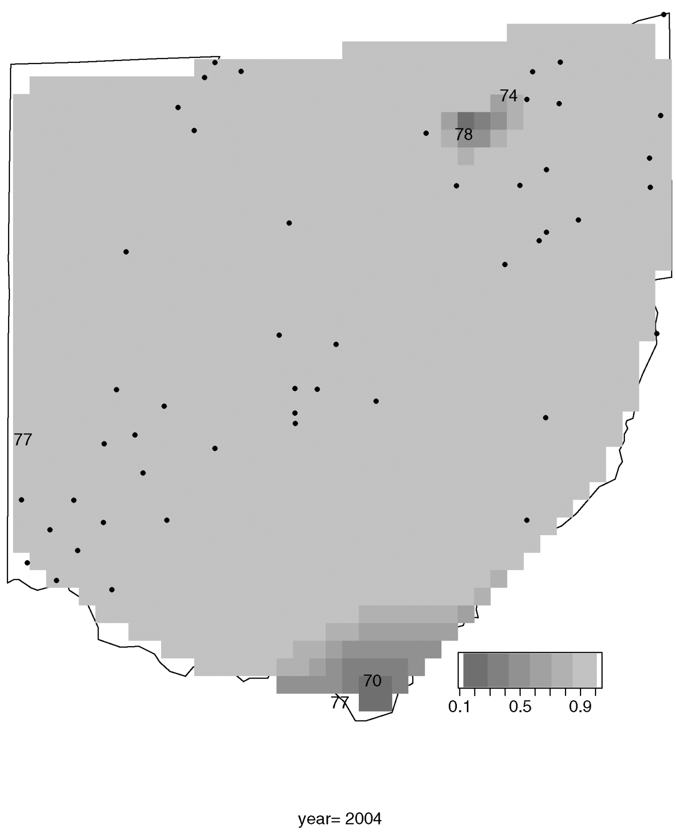

Figure 9. Probability that 3-year rolling averages of the true annual fourth highest maximum ozone levels exceed 80 for the year 2004. Observed 3-year averages that are less than 80 are superimposed on the plot.

Conditional Distributions for $\boldsymbol{\gamma}_l$, μ_l, and σ_l^2

The conditional posterior distribution of $\boldsymbol{\gamma}_l$ will come from

$$\begin{aligned}\mathbf{Z}_{l1} &= \boldsymbol{\gamma}_l + \mu_l \mathbf{1} + \boldsymbol{\epsilon}_{l1},\\ \mathbf{O}_{l2} &= \rho(\boldsymbol{\gamma}_l + \mu_l \mathbf{1}) + F_{l2}\boldsymbol{\beta} + \boldsymbol{\eta}_{l2},\\ \boldsymbol{\gamma}_l &\sim \mathrm{N}(\mathbf{0}, \Sigma_l).\end{aligned}$$

Hence, the conditional posterior distribution of $\boldsymbol{\gamma}_l$ is $\mathrm{N}(\Lambda_l \boldsymbol{\chi}_l, \Lambda_l)$, where

$$\Lambda_l^{-1} = \frac{I_n}{\sigma_\epsilon^2} + \Sigma_l^{-1} + \rho^2 Q_\eta$$

and

$$\boldsymbol{\chi}_l = \rho Q_\eta (\mathbf{O}_{l2} - \rho\mu_l \mathbf{1} - \xi_l \mathbf{1} - F_{l2}\boldsymbol{\beta}) + (\mathbf{Z}_{l1} - \mu_l \mathbf{1})/\sigma_\epsilon^2.$$

We also have the conditional distribution:

$$\frac{1}{\sigma_l^2} \sim \mathrm{G}\left(\frac{n}{2} + a, b + \frac{1}{2}\boldsymbol{\gamma}_l' \Sigma_\gamma \boldsymbol{\gamma}_l\right).$$

The conditional posterior distribution of μ_l is $\mathrm{N}(\chi_l, \lambda_l)$, where

$$\chi_l = \lambda_l^{-1}\left(\frac{\mathbf{1}'(\mathbf{Z}_{l1} - \boldsymbol{\gamma}_l)}{\sigma_\epsilon^2} + \rho \mathbf{1}' Q_\eta (\mathbf{O}_{l2} - \rho\boldsymbol{\gamma}_l - \xi_l \mathbf{1} - F_{l2}\boldsymbol{\beta})\right)$$

and

$$\lambda_l^{-1} = \frac{n}{\sigma_\epsilon^2} + \rho^2 \mathbf{1}' Q_\eta \mathbf{1} + 10^{-4}.$$

[Received May 2006. Revised September 2006.]

REFERENCES

Bell, M. L., McDermott, A., Zeger, S. L., Samet, J. M., and Dominici, F. (2004), "Ozone and Short-Term Mortality in 95 US Urban Communities, 1987–2000," *Journal of the American Medical Association*, 292, 2372–2378.

Brown, P. J., Le, N. D., and Zidek, J. V. (1994), "Multivariate Spatial Interpolation and Exposure to Air Pollutants," *The Canadian Journal of Statistics*, 22, 489–510.

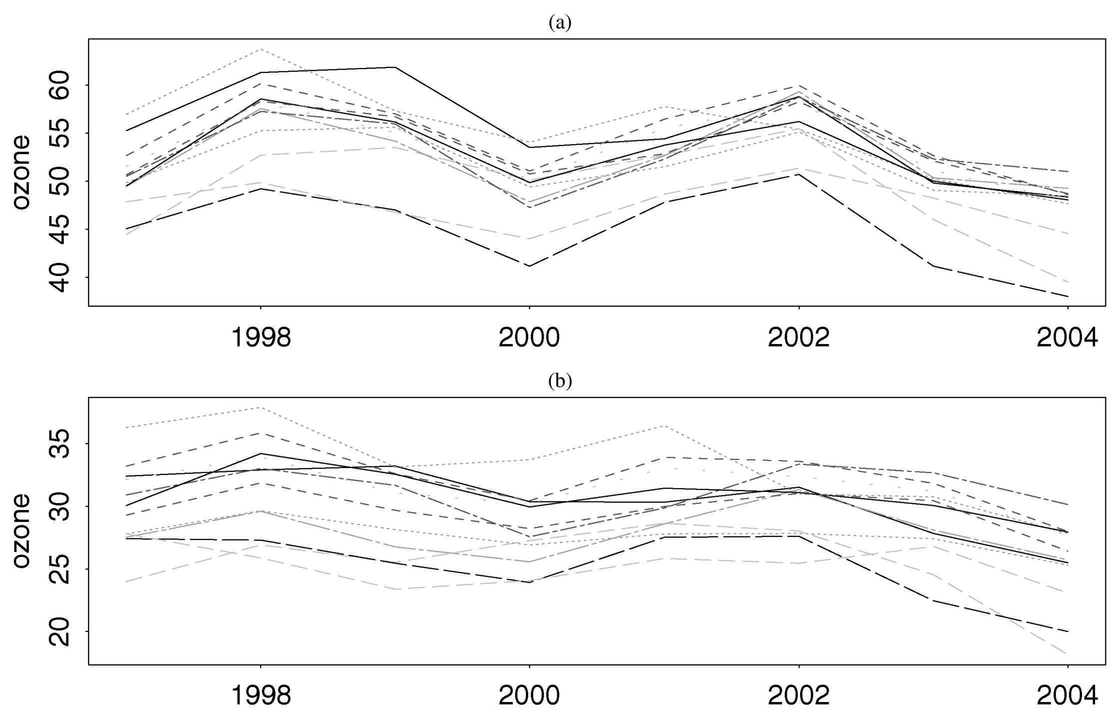

Figure 10. Trends in ozone levels at the 12 sites where meteorological variables have been observed: Panel (a) for the unadjusted and (b) for the adjusted trends.

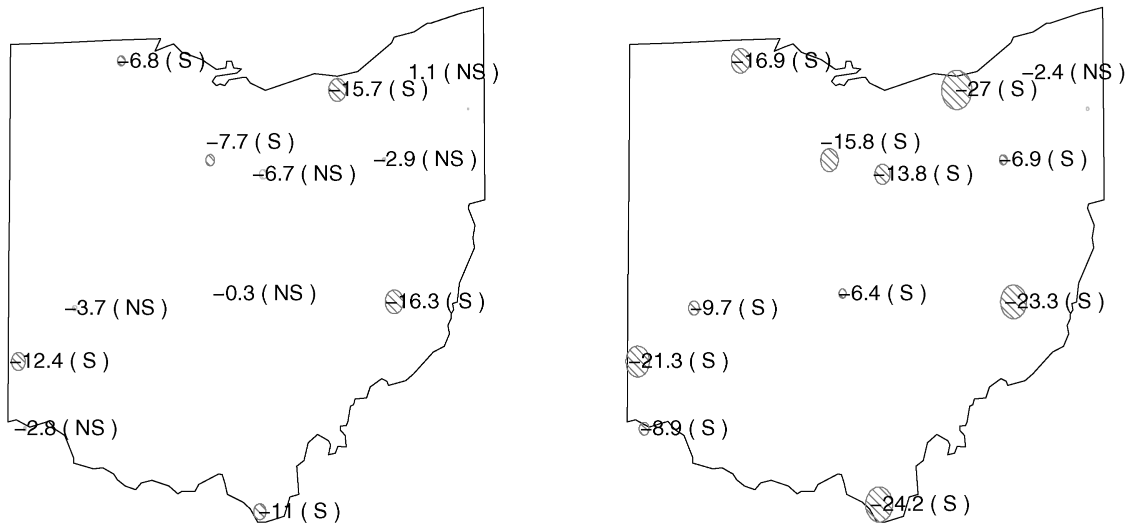

Figure 11. Relative percentage trends (RPTs) in the year 2004 (base year = 1997) at the 12 sites where meteorological variables have been observed. Significant values are labeled by (S), and nonsignificant values are labeled by (NS). The radii of the plotted circles are proportional to the RPTs labeled in the plots. The left panel is for the unadjusted ones, and the right panel is for the meteorology-adjusted ones.

Carroll, R. J., Chen, R., George, E. I., Li, T. H., Newton, H. J., Schmiediche, H., and Wang, N. (1997), "Ozone Exposure and Population Density in Harris County, Texas," *Journal of the American Statistical Association*, 92, 392–404.

Cocchi, D., Fabrizi, E., and Trivisano, C. (2005), "A Stratified Model for the Assessment of Meteorologically Adjusted Trends of Surface Ozone," *Environmental and Ecological Statistics*, 12, 1195–1208.

Cox, W. M., and Chu, S. H. (1992), "Meteorologically Adjusted Trends in Urban Areas, a Probabilistic Approach," *Atmospheric Environment*, 27, 425–434.

Gelfand, A. E., Kim, H.-J., Sirmans, C. F., and Banerjee, S. (2003), "Spatial Modeling With Spatially Varying Coefficient Processes," *Journal of the American Statistical Association*, 98, 387–396.

Gelfand, A. E., Schmidt, A. M., Banerjee, S., and Sirmans, C. F. (2004), "Nonstationary Multivariate Process Modelling Through Spatially Varying Coregionalization" (with discussion), *Test*, 2, 1–50.

Gilleland, E., and Nychka, D. (2005), "Statistical Models for Monitoring and Regulating Ground Level Ozone," *Environmetrics*, 16, 535–546.

Guttorp, P., Meiring, W., and Sampson, P. D. (1994), "A Space–Time Analysis of Ground-Level Ozone Data," *Environmetrics*, 5, 241–254.

Huang, L. S., and Smith, R. L. (1999), "Meteorologically-Dependent Trends in Urban Ozone," *Environmetrics*, 10, 103–118.

Huerta, G., Sanso, B., and Stroud, J. R. (2004), "A Spatiotemporal Model for Mexico City Ozone Levels," *Journal of the Royal Statistical Society*, Ser. C, 53, 231–248.

McMillan, N., Bortnick, S. M., Irwin, M. E., and Berliner, M. (2005), "A Hierarchical Bayesian Model to Estimate and Forecast Ozone Through Space and Time," *Atmospheric Environment*, 39, 1373–1382.

Paciorek, C. J. (2007), "Computational Techniques for Spatial Logistic Regression With Large Datasets," *Computational Statistics and Data Analysis*, in press.

Porter, P. S., Rao, S. T., Zurbenko, I. G., Dunker, A. M., and Wolff, G. T. (2001), "Ozone Air Quality Over North America: Part II. An Analysis of Trend Detection and Attribution Techniques," *Journal of Air & Waste Management Association*, 51, 283–306.

Sahu, S. K., Gelfand, A. E., and Holland, D. M. (2006), "Spatio-Temporal Modeling of Fine Particulate Matter," *Journal of Agricultural, Biological, and Environmental Statistics*, 11, 61–86.

Sahu, S. K., and Mardia, K. V. (2005), "A Bayesian Kriged–Kalman Model for Short-Term Forecasting of Air Pollution Levels," *Journal of the Royal Statistical Society*, Ser. C, 54, 223–244.

Thompson, M. L., Reynolds, J., Cox, L. H., Guttorp, P., and Sampson, P. D. (2001), "A Review of Statistical Methods for the Meteorological Adjustment of Tropospheric Ozone," *Atmospheric Environment*, 35, 617–630.

U.S. Environmental Protection Agency (2005), "Evaluating Ozone Control Programs in the Eastern United States: Focus on the NO*x* Budget Trading Program, 2004," EPA-454-K-05-001, U.S. EPA, Office of Air and Radiation, Washington, DC.

Wikle, C. K. (2003), "Hierarchical Models in Environmental Science," *International Statistical Review*, 71, 181–199.

Xia, G., and Gelfand, A. E. (2006), "Stationary Process Approximation for the Analysis of Large Spatial Datasets," technical report, Duke University, ISDS.

Zhang, H. (2004), "Inconsistent Estimation and Asymptotically Equal Interpolations in Model-Based Geostatistics," *Journal of the American Statistical Association*, 99, 250–261.

Zhu, L., Carlin, B. P., and Gelfand, A. E. (2003), "Hierarchical Regression With Misaligned Spatial Data: Relating Ambient Ozone and Pediatric Asthma ER Visits in Atlanta," *Environmetrics*, 14, 537–557.