A class of nonseparable and nonstationary spatial temporal covariance functions *

Montserrat Fuentes, Li Chen, and Jerry M. Davis

Spectral methods are powerful tools to study and model the dependency structure of spatial temporal processes. However, standard spectral approaches as well as geostatistical methods assume separability and stationarity of the covariance function; these can be very unrealistic assumptions in many settings. In this work, we introduce for the first time a general and flexible parametric class of spatial temporal covariance models, that allows for lack of stationarity and separability by using a spectral representation of the process. This new class of covariance models has a unique parameter that indicates the strength of the interaction between the spatial and temporal components; it has the separable covariance model as a particular case. We introduce an application with ambient ozone air pollution data provided by the U.S. Environmental Protection Agency.

*Fuentes is an associate professor in the Statistics Department at North Carolina State University (NCSU). E-mail: fuentes@stat.ncsu.edu. Chen is a research associate in the Center for Integrating Statistical and Environmental Science (CISES) at the University of Chicago, Chicago, IL 60637. Davis is a professor emeritus in the Department of Marine Earth and Atmospheric Sciences at NCSU. Fuentes has been partially supported by two National Science Foundation grants DMS 0353029 and DMS 0706731. Chen has been supported by an EPA STAR Cooperative Agreement #R-82940201-0 awarded to the University of Chicago. *Key words:* nonseparability, nonstationarity, spatial temporal models, ambient ozone, spectral density.

1 Introduction

Ozone (O_3) is a widely studied atmospheric pollutant because of its adverse health effects and its impact on certain agricultural crops. In addition, it is also known to lead to the breakdown of certain categories of materials. Ozone is a secondary pollutant and is formed by photochemical processes which involve nitrogen oxides (NOX) and volatile organic compounds VOCs). A complete description of the chemical processes involving ozone can be found in Seinfeld and Pandis (1998). Because of the documented effects on human health caused by ozone, the U.S. Environmental Protection Agency (U.S. EPA) has implemented control strategies to reduce ozone levels in the atmosphere. A detailed look at these various control programs at the Federal and local level can be found in U.S. National Research Council (2004). These control strategies have focused on the reduction of NOX and VOC emissions. A recent report, U.S. EPA (U.S. EPA, 2005), which focuses on the eastern US, indicates that there have been significant reductions in NOX and VOC emissions in this part of the country over the period from 1997 to 2004. As a direct result of the reductions in emissions, ozone levels have also gradually decreased in the eastern US. It should also be noted that ozone levels in any given year and any given region are strongly dependent on meteorological conditions. Thus reduced ozone levels in a given region and a given year are likely to be dependent on meteorological conditions as well as emissions.

There is an extensive literature on the application of various statistical methods to the analysis and prediction of ozone. Most of these papers analyze the data separately at each geographic location or they assume stationary models. Thompson et al. (2001) reviewed those articles that applied various statistical techniques to the forecasting of ozone levels,

the estimation of ozone time trends, and an understanding of the fundamental mechanisms responsible for the production of ozone. They classified these techniques into three main categories: regression analysis, extreme value analysis, and space-time analysis. A number of the more recent papers on the application of statistical techniques to the analysis of ozone have been reviewed in Zheng et al. (2007). The emphasis here will be on research work which emphases the spatial temporal aspects of ozone variability.

Three recent papers (Huerta, et al. (2004), McMillan, et al. (2005) and Zheng et al. (2007)) have explored the spatial temporal aspects of ozone modeling. Huerta, et al. (2004) use a dynamic linear model for ambient ozone. McMillan et al. (2005) use a nearest-neighbor spatial model. Zheng et al. (2007) use a dynamic linear model and a generalized additive model to explain ozone trends. A recent paper by Sahu (2007) presents a very elegant approach, using ozone differentials to explain spatial temporal patterns for ozone in Ohio. In their model, the authors add a correlated error structure using a linear model of coregionalization that has a stationary covariance model. In this application we have a larger geographic domain, so we need to go beyond this assumption of stationarity (defined below). In a space-time regression context the use of dynamic linear models (as in all the papers just referred), in principle, could overcome the problem of modeling complex space-time dependency structures, avoiding then, the issue of having to propose general space-time parametric models for the covariance function. However, simple linear dynamic models (DLM) with only white noise error components might not capture complex space-time structures. The hybrid approach described in the application presented in this paper, using DLM with space-time (nonstationary and nonseparable) correlated parametric covariance functions may provide an optimal modeling framework.

The difficulty of modeling the spatial temporal structure of ambient ozone concentrations can be overcome by using separable processes. The assumption of separability for spatial temporal processes offers a simplified representation of any variance-covariance matrix, and consequently, some remarkable computational benefits. A spatial temporal process $Z(\mathbf{x}, t)$, where $\mathbf{x}$ represents space and t represents time, is separable if $\text{cov}\{Z(\mathbf{x}, t), Z(\mathbf{x}', t')\} = C_S(\mathbf{x}, \mathbf{x}')C_T(t, t')$, where $C_S(\cdot)$ is a purely spatial covariance and $C_T(\cdot)$ is a temporal covariance. For a separable process, the spatial and temporal structures can be modeled separately. Therefore, many techniques that have been developed and successfully implemented in time series analysis and geostatistics can be used with this subclass of separable spatial temporal processes. Another advantage of assuming separability is the computational efficiency, since the spatial temporal covariance matrix can be written as the Kronecker product of two smaller matrices. But with ozone data, separability is not generally a realistic assumption. Nonseparable spatial temporal covariance models have been proposed by Cressie and Huang (1999), Gneiting (2002), and Stein (2005). Wikle et al. (1998) provide a review of Bayesian hierarchical space-time modeling for environmental processes. In this paper, we introduce a new class of nonseparable covariance models, which has the separable model as a particular case. In this new class of nonseparable stationary covariance models, there is a unique parameter to indicate the strength of the dependency between the spatial and temporal components is. This parameter could be used to construct a parametric test for separability.

We also extend this new model to the nonstationary case. Define $\mathbf{s} = (\mathbf{x}, t)$. If $Z(\mathbf{s})$ has finite

second moments, its mean function is constant and the covariance function of Z only depends on the distance vector between $\mathbf{s}$ and $\mathbf{s}'$ (i.e., $\mathbf{s}-\mathbf{s}'$), then $Z(\mathbf{s})$ is a weakly stationary spatial temporal process. It is also called a second-order stationary spatial temporal process. If $Z(\mathbf{s})$ is weakly stationary, $\text{cov}\{Z(\mathbf{s}), Z(\mathbf{s}')\} = C(\mathbf{s}-\mathbf{s}')$, where $C(\cdot)$ is a valid covariance function. This implies that the relationship between the values of the process at two locations only depends on the vector distance between these two locations. This assumption of stationarity is not realistic for ozone data, especially when the geographic domain of interest is large, as it is in our application. There is an extensive literature on nonstationary models, e.g. Sampson and Guttorp, (1992), Nychka et al.,(2002), Higdon et al., (1999), Fuentes (2002). Fuentes (2005a), and Fuentes (2005b) introduce nonparametric tests for stationary and separability respectively, using spectral methods. In this paper, we introduce a nonstationary model for space-time data that is also nonseparable. Our approach has some similarities to Higdon's method, and that will be discussed. In this paper we introduce for the first time a parametric nonstationary and nonseparable covariance model. Our model is flexible and general enough to explain the complex space-time dependency structures of ozone data.

The models we propose here are based on spectral representations of the spatial temporal processes.

In section 2 we describe the data used in the application presented in this paper. In section 3 we introduce a spectral representation of stationary spatial temporal processes. In section 4 we propose a new class for nonseparable stationary covariance models. The general model for nonseparability and nonstationarity is introduced in section 5. Section 6 presents a model for the large scale structure of a space-time process. In section 7 we apply the new statistical

models to ozone data. We conclude with a discussion in Section 8.

2 Data description

All of the ozone and meteorological data used in this study came from the U.S. EPA's Clean Air Status and Trends Network (CASTNET) monitoring stations. A complete description of this network can be found at www.epa.gov/castnet. The CASTNET sampling period is one week long and runs from Tuesday morning to Tuesday morning. All the meteorological data and the ozone values are available hourly. The original dataset spanned the time from 1997 to 2004. There are 49 CASTNET stations in the spatial domain used in this study (see Figure 1). In this study only the months from May to September 2002 (153 days) were used. This period constitutes the ozone season for most locations. It is a period dominated by long days with abundant solar radiation input. It is also a time of year when stagnating high pressure systems are more frequent. These systems are characterized by clear, very warm days with light winds, conditions ideal for the formation of ozone.

The final dataset included the following variables: temperature at 9 m (daily maximum and minimum values and the daily average from 10AM to 5PM), temperature at 2 m (daily maximum and minimum, and the daily average from 10AM to 5PM), relative humidity at 9 m (daily average from 10AM to 5PM), dew point temperature at 9 m calculated from the relative humidity and temperature (daily average from 10AM to 5PM), solar radiation (daily peak value, average from 10AM to 5PM, and sum over 10AM to 5PM) atmospheric stability calculated based on the bulk Richardson number (daily average from 10AM to 5PM and from 8PM to 5AM, see Stull (1988)), wind speed (daily average from 6AM to 9AM, average

from 10AM to 5PM, and average from 8PM to 5AM), precipitation (daily total), and ozone (the maximum 1-hour average for each day and the maximum 8-hour average for each day). The calculation of the 8-hour average ozone values is based on EPA's National Ambient Air Quality Standards. See www.epa.gov/ttn/amtic/files/ambient/criteria/reldocs/guidefin.pdf The meteorological variables which are most important in the spatial temporal modeling of the 8-hr average ozone are the daily minimum and average temperature at two meters, daily maximum and average solar radiation, daily average relative humidity and precipitation. These are the covariates used in our analysis.

3 A spectral representation for stationary spatial temporal processes

Consider $\{Z(\mathbf{x},t) : \mathbf{x} \in D \subset \mathbb{R}^d, t \in T \subset \mathbb{R}\}$ a stationary random field with spatial temporal covariance $C(\mathbf{x},t)$, then we can represent the process in the form of the following Fourier-Stieltjes integral (see Yaglon (1987), Christakos (1992), Stein (1999) for example):

$$Z(\mathbf{x},t) = \int_{\mathbb{R}^d} \int_{\mathbb{R}} \exp(i\boldsymbol{\omega}^T\mathbf{x} + i\tau t) dY(\boldsymbol{\omega},\tau). \tag{1}$$

where Y is a random function that has uncorrelated increments with complex symmetry except for the constraint, $dY(\boldsymbol{\omega},\tau) = dY^c(-\boldsymbol{\omega},-\tau)$, needed to ensure $Z(\mathbf{x},t)$ is real-valued. Y^c denotes the conjugate of Y. Using the spectral representation of Z and proceeding formally,

$$C(\mathbf{x},t) = \int_{\mathbb{R}^d} \int_{\mathbb{R}} \exp(i\boldsymbol{\omega}^T\mathbf{x} + i\tau t) F(d\boldsymbol{\omega}, d\tau), \tag{2}$$

where the function F is a positive finite measure and is called the spectral measure or spectrum for Z. The spectral measure F is the mean square value of the process Y,

$$E\{|Y(\boldsymbol{\omega},\tau)|^2\} = F(\boldsymbol{\omega},\tau).$$

It is easy to see that for any finite positive measure F, the function given in (2) is positive definite. If F has a density with respect to Lebesgue measure, it is the spectral density, f, which is the Fourier transform of the spatial temporal covariance function:

$$f(\boldsymbol{\omega},\tau) = \frac{1}{(2\pi)^{d+1}} \int_{\mathbb{R}^d} \int_{\mathbb{R}} \exp(-i\boldsymbol{\omega}^T\mathbf{x} - i\tau t) C(\mathbf{x},t) d\mathbf{x} dt, \tag{3}$$

and the corresponding covariance function is given by

$$C(\mathbf{x},t) = \int_{\mathbb{R}^d} \int_{\mathbb{R}} \exp(i\boldsymbol{\omega}^T\mathbf{x} + i\tau t) f(\boldsymbol{\omega},\tau) d\boldsymbol{\omega} d\tau. \tag{4}$$

When $f(\boldsymbol{\omega},\tau) = f^{(1)}(\boldsymbol{\omega}) f^{(2)}(\tau)$, we obtain

$$\begin{aligned} C(\mathbf{x},t) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}} \exp(i\boldsymbol{\omega}^T\mathbf{x} + i\tau t) f^{(1)}(\boldsymbol{\omega}) f^{(2)}(\tau) d\boldsymbol{\omega} d\tau \\ &= \int_{\mathbb{R}^d} \exp(i\boldsymbol{\omega}^T\mathbf{x}) f^{(1)}(\boldsymbol{\omega}) d\boldsymbol{\omega} \int_{\mathbb{R}} \exp(i\tau t) f^{(2)}(\tau) d\tau \\ &= C^{(1)}(\mathbf{x}) C^{(2)}(t), \end{aligned}$$

which means the corresponding spatial temporal covariance is separable.

A common parametric spatial (or spatial temporal) covariance function is the Matérn given by

$$C(\mathbf{x}) = \frac{\pi^{d/2}\gamma}{2^{\nu-1}\Gamma(\nu + \frac{d}{2})\alpha^{2\nu}} (\alpha|\mathbf{x}|)^\nu \mathcal{K}_\nu(\alpha|\mathbf{x}|),$$

where $\mathcal{K}_\nu(\alpha|\mathbf{x}|)$ is a modified Bessel function (see Gradshteyn and Ryzhik (2007), for example). The parameter α^{-1} explains the rate of decay of the spatial correlation, γ is a

scale parameter. The parameter ν measures the degree of smoothness of the process Z. The corresponding spectral density is given by

$$f(\boldsymbol{\omega}) = \gamma(\alpha^2 + |\boldsymbol{\omega}|^2)^{-\nu - d/2}.$$

In geostatistics, is common to reparameterize the Matérn covariance, and use instead the following reparameterization:

$$\nu, \quad \rho = 2\nu^{1/2}/\alpha, \quad \text{and } \sigma^2 = \frac{\gamma\Gamma(\nu)\pi^{d/2}}{\Gamma(\nu + d/2)\alpha^{2\nu}}, \tag{5}$$

ν (same as before), the parameter ρ measures how the correlation decays with distance; generally this parameter is called the range. The parameter σ^2 is the variance of the process Z, i.e, $\sigma^2 = \text{var}(Z(\mathbf{x}, t))$, the covariance parameter σ^2 is usually referred to as the sill.

4 A new class of nonseparable stationary covariance models

We propose the following spatial temporal spectral density that has a separable model as a particular case,

$$f(\boldsymbol{\omega}, \tau) = \gamma(\alpha^2\beta^2 + \beta^2|\boldsymbol{\omega}|^2 + \alpha^2\tau^2 + \epsilon|\boldsymbol{\omega}|^2\tau^2)^{-\nu}, \tag{6}$$

where γ, α and β are positive, $\nu > \frac{d+1}{2}$ and $\epsilon \in [0, 1]$. The function in (6) is a valid spectral density (see Appendix A).

In the representation (6), the parameter α^{-1} (spatial range) explains the rate of decay of the spatial correlation, β^{-1} (temporal range) explains the rate of decay for the temporal correlation, and γ is a scale parameter proportional to the variance σ^2 (sill parameter)

of the process. The sill parameter is $\sigma^2 = \text{var}\{Z(\mathbf{x}, t)\} = \gamma\{\int_{\mathbb{R}^2}\int_{\mathbb{R}} f(\boldsymbol{\omega}, \tau)/\gamma d\boldsymbol{\omega} d\tau\}$, The parameter ν (smoothness) measures the degree of smoothness of the process Z. The parameter ϵ (parameter of separability) indicates the interaction between the spatial and temporal components. We define the covariance vector parameter, $\theta = (\sigma^2, \alpha^{-1}, \beta^{-1}, \nu, \epsilon)$. Next, we discuss two particular cases of this new model for spatial temporal dependency structures presented in (6).

- $\underline{\epsilon = 1.}$ When $\epsilon = 1$, Equation (6) can be written as

$$f(\boldsymbol{\omega}, \tau) = \gamma(\alpha^2\beta^2 + \beta^2|\boldsymbol{\omega}|^2 + \alpha^2\tau^2 + |\boldsymbol{\omega}|^2\tau^2)^{-\nu} = \gamma(\alpha^2 + |\boldsymbol{\omega}|^2)^{-\nu}(\beta^2 + \tau^2)^{-\nu}.$$

  Therefore the corresponding spatial temporal covariance is separable. Moreover, in the expression of this covariance, both the spatial component and the temporal component are Matérn-type covariances, see Figure 2. In Figure 2 we observe ridges along the lines where the spatial and temporal lags are 0. For most physical processes, these ridges are not realistic, they are an artifact due to the assumption of separability.

- $\underline{\epsilon = 0.}$ When $\epsilon = 0$, $f(\boldsymbol{\omega}, \tau) = \gamma(\alpha^2\beta^2 + \beta^2|\boldsymbol{\omega}|^2 + \alpha^2\tau^2)^{-\nu}$.

  This function is an extension of the traditional Matérn (Matérn, 1960) spectral density. It treats time as an additional coordinate, but it does have a different rate of decay for time, the parameters α^{-1} and β^{-1} explain respectively, the rate of decay of the spatial and temporal components, see Figure 3.

For $\epsilon \in (0, 1)$, we calculate analytically the spatial temporal covariance that corresponds to

the spectral density in (6),

$$\begin{aligned} C(\mathbf{x},t) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}} \exp(i\boldsymbol{\omega}^T\mathbf{x}+i\tau t) f(\boldsymbol{\omega},\tau) d\boldsymbol{\omega} d\tau \\ &= \int_{\mathbb{R}} \exp(i\tau t) g(\mathbf{x},\tau) d\tau, \end{aligned} \tag{7}$$

where $g(\mathbf{x},\tau) = \int_{\mathbb{R}^d} \exp(i\boldsymbol{\omega}^T\mathbf{x}) f(\boldsymbol{\omega},\tau) d\boldsymbol{\omega}$. The function $g(\mathbf{x},\tau)$ is available from the integration, therefore $C(\mathbf{x},t)$ can be computed by numerically carrying out a one-dimensional Fourier transformation of g. This can be quickly approximated using the fast Fourier transform, this is the approach adopted in the application presented in this paper. The expression of g is given by

$$g(\mathbf{x},\tau) = \frac{\pi^{d/2}\gamma}{2^{\nu-\frac{d}{2}-1}\Gamma(\nu)} (\beta^2+\epsilon\tau^2)^{-\nu} \left(\frac{|\mathbf{x}|}{\theta(\tau)}\right)^{\nu-\frac{d}{2}} \mathcal{K}_{\nu-\frac{d}{2}}(\theta(\tau)|\mathbf{x}|),$$

where $\theta(\tau) = \sqrt{\frac{\alpha^2(\beta^2+\tau^2)}{\beta^2+\epsilon\tau^2}}$. When $\gamma = \alpha = \beta = d = 1$ and $\nu = 3/2$, contour plots of the corresponding separable spatial temporal covariances with ϵ =0.1, 0.2, 0.3, 0.5, 0.7, 0.8, 0.9 are given in Figure 4. The ridges in the covariance function get sharper as ϵ increases.

In summary, in this new class of spectral density/covariance functions the parameter ϵ controls the degree of lack separability by explaining the potential interaction between the dependency structure of the spatial and temporal components.

To incorporate the lack of stationarity in this model, we allow the nonseparable spectral function f to be space-dependent, as explained in the next Section.

5 Nonstationary models for space-time covariances

In the previous sections, we introduced the spectral representation of stationary spatial temporal processes, and proposed a new class of nonseparable stationary spectral densities. Now we generalize the idea to nonstationary spatial temporal processes.

Here, we define a generalized class for spatial temporal nonseparability and nonstationarity. In this class, the spectral representation itself and the corresponding spectral distribution function (or spectral density) can change slowly in space and time. Let $Z(\mathbf{x}, t)$ be a general space-time process, $\mathbf{x} \in D \subseteq \mathbb{R}^d$ and $t \in T \subseteq \mathbb{R}$, we use the following representation

$$Z(\mathbf{x}, t) = \int_{\mathbb{R}^d} \int_{\mathbb{R}} \exp(i\boldsymbol{\omega}^T \mathbf{x} + i\tau t) dY_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega}, \tau). \tag{8}$$

We are going to assume throughout this section that the mean of the process is zero. We capture the potential lack of stationarity by allowing the spectral process Y associated with Z to change in space and time.

The general spectral representation proposed in (8) provides an ideal framework to model complex space-time dependent structures. Next, we present a flexible model that correspond to processes with this type of spectral representation.

Mixture of local spectrums

A particular case of the general representation in (8) is when the lack of stationarity is explained by allowing the amplitude of the spectral process Y to be space-time dependent.

This means

$$Z(\mathbf{x},t) = \int_{\mathbb{R}^d}\int_{\mathbb{R}} \exp(i\boldsymbol{\omega}^T\mathbf{x} + i\tau t)\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega},\tau)dY_0(\boldsymbol{\omega},\tau) \tag{9}$$

where $\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega},\tau)$ is a space-time dependent amplitude function, and Y_0 is a space-time invariant Wiener process, which satisfies the relation

$$E\left[Y_0(\boldsymbol{\omega},\tau)Y_0^c(\boldsymbol{\omega}',\tau')\right] = \delta(\boldsymbol{\omega}-\boldsymbol{\omega}')\delta(\tau-\tau')$$

where δ is the delta Dirichlet function. This is a space-time version of the evolutionary spectrum presented by Priestley (1965). We assume the functions $\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega},\tau)$ satisfy the condition

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}} |\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega},\tau)|^2 d\boldsymbol{\omega}d\tau < \infty \tag{10}$$

for all $\mathbf{x}$ and t. The functions $\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega},\tau)$ must also satisfy $\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega},\tau) = \phi^c_{\mathbf{x},\mathbf{t}}(-\boldsymbol{\omega},-\tau)$ to ensure $Z(\mathbf{x},t)$ is real-valued.

Then, it is easy to see that the covariance function, C, of the process $Z(\mathbf{x},t)$ is given by the formula

$$\begin{aligned} & \operatorname{cov}\{Z(\mathbf{x}_1,t_1), Z(\mathbf{x}_2,t_2)\} \\ = \;& C(\mathbf{x}_1,t_1;\mathbf{x}_2,t_2) \\ = \;& \int_{\mathbb{R}^d}\int_{\mathbb{R}} \exp\{i\boldsymbol{\omega}^T(\mathbf{x}_1-\mathbf{x}_2)\}\exp\{i\tau(t_1-t_2)\}\phi_{\mathbf{x}_1,t_1}(\boldsymbol{\omega},\tau)\phi^c_{\mathbf{x}_2,t_2}(\boldsymbol{\omega},\tau)d\boldsymbol{\omega}d\tau \end{aligned} \tag{11}$$

In particular

$$\operatorname{var}\{Z(\mathbf{x},t)\} = C(\mathbf{x},t;\mathbf{x},t) = \int_{\mathbb{R}^d}\int_{\mathbb{R}} |\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega},\tau)|^2 d\boldsymbol{\omega}d\tau \tag{12}$$

so that condition (10) is necessary for the variance of $Z(\mathbf{x},t)$ to be finite at all $\mathbf{x}$ and t, i.e. for the existence of a covariance function $C(\mathbf{x}_1,t_1;\mathbf{x}_2,t_2)$.

The representation in (9) may be interpreted as a representation of the process Z in the form of a superposition of sinusoidal oscillations with different frequencies $\boldsymbol{\psi} = (\boldsymbol{\omega}, \tau)$ and random amplitudes $\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega}, \tau)$ varying over space and time, i.e. $Z(\mathbf{x}, t)$ is a superposition of amplitude modulated random oscillations. According to this interpretation, (12) describes the distribution of the "total power" of the process $Z(\mathbf{x}, t)$ at location $\mathbf{x}$ and time t over the frequencies $(\boldsymbol{\omega}, \tau)$, hence the contribution from the frequency $(\boldsymbol{\omega}, \tau)$ is $|\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega}, \tau)|^2 d\boldsymbol{\omega} d\tau$. Therefore the function $F_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega}, \tau)$ defined by the relation

$$dF_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega}, \tau) = |\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega}, \tau)|^2 d\boldsymbol{\omega} d\tau \tag{13}$$

will be called the spatial temporal spectral distribution function of the process Z, and $f_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega}, \tau) = |\phi_{\mathbf{x},\mathbf{t}}(\boldsymbol{\omega}, \tau)|^2$ is the spatial temporal spectral density of Z.

There exist different representations of the form (8) for a spatial temporal process Z, each representation is based on a different family of $\phi_{\mathbf{s}}(\boldsymbol{\psi})$ functions, where we write $\mathbf{s} = (\mathbf{x}, \mathbf{t})$ and $\boldsymbol{\psi} = (\boldsymbol{\omega}, \tau)$ to simplify the notation. This problem is similar to the selection of a basis for a vector space.

Apart from that, it would not be physically meaningful to interpret $\boldsymbol{\psi}$ as the frequency in all cases. In the physical theory of oscillations the function $A_{\mathbf{s}}(\boldsymbol{\psi}) = \phi_{\mathbf{s}}(\boldsymbol{\psi}) \exp(i\mathbf{s}^T\boldsymbol{\psi})$ is said to describe the amplitude modulated oscillation of frequency $\boldsymbol{\psi}$ only if the "amplitude" $\phi_{\mathbf{s}}(\boldsymbol{\psi})$ is slowly varying compared to $\exp(i\mathbf{s}^T\boldsymbol{\psi})$ function, i.e. if the Fourier transform of $\phi_{\mathbf{s}}(\boldsymbol{\psi})$ as a function of $\mathbf{s}$ includes mainly frequencies much lower than $\boldsymbol{\psi}$. It is even often assumed that this transform must be concentrated in a neighborhood of zero frequency. We restrict the

permissible variability of the function $\phi_{\mathbf{s}}(\boldsymbol{\psi})$ of $\mathbf{s}$ by considering only functions $\phi_{\mathbf{s}}(\boldsymbol{\psi})$ that admit a generalized Fourier representation

$$\phi_{\mathbf{s}}(\boldsymbol{\psi}) = \int_{\mathbb{R}^{d+1}} e^{i\mathbf{s}^T\boldsymbol{\sigma}} dH_{\boldsymbol{\psi}}(\boldsymbol{\sigma}) \tag{14}$$

with $|dH_{\boldsymbol{\psi}}(\boldsymbol{\sigma})|$ having its maximum at $\boldsymbol{\sigma} = \mathbf{0}$ for any fixed $\boldsymbol{\psi}$. This condition guarantees that the Fourier transform of $\phi_{\mathbf{s}}(\boldsymbol{\psi})$, as a function of $\mathbf{s}$, includes mainly frequencies much lower than any $\boldsymbol{\psi}$, as has been suggested by Priestley in the time series context. Since $\phi_{\mathbf{s}}(\boldsymbol{\psi})$ is a slowly varying function of space and time, it is clear that the process Z may be regarded as being "approximately stationary" within subregions in our spatial temporal domain. If, however, we examine the behavior of Z within two subregions which are sufficiently far apart, we could find that although Z is practically stationary in both subregions, the spectral distribution function of the two "portions" of Z will, in general, be different (i.e., the spectral distribution of the power of Z varies on space and time). Since the functions $\phi_{\mathbf{s}}(\boldsymbol{\psi}) = 1$ clearly satisfy the conditions to be imposed on $\phi_{\mathbf{s}}(\boldsymbol{\psi})$, the representation (8) certainly includes all the spatial temporal stationary processes having a finite variance.

A particular case is when ϕ is a separable function of spatial frequency and temporal frequency, i.e.

$$\phi_{\mathbf{s}}(\boldsymbol{\psi}) = \phi^{(1)}_{\mathbf{x}}(\boldsymbol{\omega})\phi^{(2)}_{t}(\tau). \tag{15}$$

When ϕ is of the form (15) the spatial temporal process is separable, since

$$\begin{aligned}
&\text{cov}\{Z(\mathbf{x}_1,t_1),Z(\mathbf{x}_2,t_2)\}\\
&\quad=\int_{\mathbb{R}^{d+1}}\exp\{i\boldsymbol{\omega}^T(\mathbf{x}_1-\mathbf{x}_2)\}\exp\{i\tau^T(t_1-t_2)\}\phi^{(1)}_{\mathbf{x_1}}(\boldsymbol{\omega})\phi^{(2)}_{t_1}(\tau)(\phi^{(1)}_{\mathbf{x_2}}(\boldsymbol{\omega}))^c(\phi^{(2)}_{t_2}(\tau))^c d\boldsymbol{\omega}d\tau\\
&\quad=\int_{\mathbb{R}^{d}}\exp\{i\boldsymbol{\omega}^T(\mathbf{x}_1-\mathbf{x}_2)\}\phi^{(1)}_{\mathbf{x_1}}(\boldsymbol{\omega})(\phi^{(1)}_{\mathbf{x_2}}(\boldsymbol{\omega}))^c d\boldsymbol{\omega}\int_{\mathbb{R}}\exp\{i\tau^T(t_1-t_2)\}\phi^{(2)}_{t_1}(\tau)(\phi^{(2)}_{t_2}(\tau))^c d\tau\\
&\quad=C^{(1)}(\mathbf{x}_1,\mathbf{x}_2)C^{(2)}(t_1,t_2).
\end{aligned}\tag{16}$$

where $C^{(1)}$ and $C^{(2)}$ are spatial and temporal covariance functions.

We propose a more general model for ϕ that has the separable model in (15) as a particular case. We model ϕ as a mixture of local spectral (amplitude) functions,

$$\phi_{\mathbf{s}}(\boldsymbol{\psi})=\sum_{j=1}^{k}K(\mathbf{s}-\mathbf{s}_j)\phi_{\mathbf{s}_j}(\boldsymbol{\psi})\tag{17}$$

where each $\phi_{\mathbf{s}_j}(\boldsymbol{\psi})$ function explains the spatial temporal structure of Z in a neighborhood of $\mathbf{s}_j$. $K(\mathbf{s}-\mathbf{s}_i)$ is a kernel function.

Locally (in a neighborhood of $\mathbf{s_j}$), we use the nonseparable parametric model proposed in (6) for $\phi_{\mathbf{s}_j}$, that has a separable model as a particular case. We can write $\phi_{\mathbf{s}_j}(\boldsymbol{\omega},\tau)$ as

$$\phi_{\mathbf{s}_j}(\boldsymbol{\omega},\tau)=\gamma_j(\alpha_j^2\beta_j^2+\beta_j^2\|\boldsymbol{\omega}\|^2+\alpha_j^2|\tau|^2+\epsilon_j\|\boldsymbol{\omega}\|^2|\tau|^2)^{-(\nu_j+d+1)/2}.\tag{18}$$

Again, the parameter α_j^{-1} explains the rate of decay of the spatial correlation component. For the temporal correlation, the rate of decay is explained by β_j^{-1}, and γ_j is a scale parameter. The parameter $\nu_j>0$ measures the degree of smoothness of the process Z at $\mathbf{s_j}$, the higher the value of ν_j the smoother Z would be. If $\epsilon_j=0$ we have a $d+1$ dimensional Matérn

type model with different spatial and temporal ranges, which takes into account the change in units from the spatial to the temporal domain. If $\epsilon_j = 1$ we have a separable model,

$$\phi_{\mathbf{s}_j}(\boldsymbol{\psi}) = \gamma_j(\alpha_j^2 + \|\boldsymbol{\omega}\|^2)^{-(\nu_j+d)/2}(\beta_j^2 + |\tau|^2)^{-(\nu_j+d)/2}.$$

Then, the corresponding (local) covariance is separable (as in (16)).

Hence, the spectral model for a nonseparable and nonstationary spatial temporal process $Z(\mathbf{s}) = Z(\mathbf{x}, t)$ is given by

$$Z(\mathbf{s}) = \int_{\mathbb{R}^{d+1}} \exp(i\boldsymbol{\psi}^T\mathbf{s}) \left(\sum_{j=1}^{k} K(\mathbf{s} - \mathbf{s}_j)\phi_{\mathbf{s}_j}(\boldsymbol{\psi}) \right) dY_0(\boldsymbol{\psi}), \tag{19}$$

where $\phi_{\mathbf{s}_j}(\boldsymbol{\psi}) = \phi_{\mathbf{s}_j}(\boldsymbol{\omega}, \tau)$ which is defined in (18). The value of k is the number of local (stationary) spectral functions, it can be estimated from the data using the Akaike information criterion (AIC, Akaike, 1974) or the Bayesian information criterion (BIC, Schwarz, 1978), $\mathbf{s}_j$ corresponds to the center of the subregion which gives the j-th local stationary component, $K(\mathbf{s} - \mathbf{s}_j)$ represents the contribution from the j-th local stationary components. If the kernel function K is a separable function of space and time, and $\epsilon_j = 1$ for $j = 1, \cdots, k$, the process $Z(\mathbf{s})$ is separable and nonstationary.

Higdon, Swall and Kern (1999) considered other classes of nonstationary models. The development of Higdon *et al.* relies heavily on the Gaussian form of kernel function and it is not clear how restrictive this is. The approach proposed here has a quite different motivation, but it also includes Higdon's model as a particular case. In model (9), if we treat the amplitude function, $A_{\mathbf{x}}(\boldsymbol{\omega})$, defined as $A_{\mathbf{x}}(\boldsymbol{\omega}) = \phi_{\mathbf{x}}(\boldsymbol{\omega}) \exp(i\mathbf{x}^T\boldsymbol{\omega})$, as a positive kernel function and Y_0 as a Wiener process, we obtain Higdon's representation. However, our representation is more general, $A_{\mathbf{x}}$ is a complex function and Y_0 a process with orthogonal increments (not

necessarily a Wiener process). The alternative representation of a nonstationary process introduced by (Fuentes, 2002) as a mixture/convolution of independent local stationary processes is not reducible to the one introduced here, because of the assumption of independency between the local processes, if the processes are allowed to be dependent then that model is a particular case of the more general one introduced here.

6 Spatial temporal trend for ozone

In the previous sections we have introduced some new models for space-time covariance functions. In this section, we introduce spatial temporal models for the mean structure of a space-time process, using covariates that have space-time dynamic coefficients.

In general, a spatial temporal process can be represented by

$$Z(\mathbf{x}, t) = \mu(\mathbf{x}, t) + \epsilon(\mathbf{x}, t)$$

where $\mathbf{x} \in D \subseteq \mathbb{R}^d$ and $t \in T \subseteq \mathbb{R}$. The function $\mu(\mathbf{x}, t)$ represents the mean surface.The residual term $\epsilon(\mathbf{x}, t)$ is a zero mean space-time correlated error that explains the spatial temporal short scale dependency structure. The statistical models proposed in the previous section are used to explain the spatial temporal correlation structure of $\epsilon(\mathbf{x}, t)$.

Here, we represent the large scale structure (e.g., trend surface) of Z using a space-time dynamic statistical model:

$$\mu(\mathbf{x}, t) = \sum_{i=1}^{m} \beta_i(\mathbf{x}, t) f_i(\mathbf{x}, t), \tag{20}$$

where $\{f_i\}_i$ are m covariates (e.g. weather data) of interest with coefficients β_i that vary in space and time.

We model the dynamic coefficients β_i using a hierarchical model in terms of an overall time component $\gamma_{i,t}$ and a space-time process $\gamma_i(\mathbf{x}, t)$,

$$\beta_i(\mathbf{x}, t) = \gamma_{i,t} + \gamma_i(\mathbf{x}, t)$$

where

$$\gamma_{i,t} = \gamma_{i,t-1} + u(t),$$

and

$$\gamma_i(\mathbf{x}, t) = \gamma_i(\mathbf{x}, t-1) + \eta(\mathbf{x}, t).$$

η and u are independent white noise processes. Here we present both $\gamma_{i,t}$ and $\gamma_i(\mathbf{x}, t)$ in the form of an (intrinsic) autoregression model. More general forms can be used. In our application, the components $\gamma_i(\mathbf{x}, t)$ are kept constant within subregions of stationarity, so $\gamma_i(\mathbf{x}, t) = \gamma_{ij,t}$ for all $\mathbf{x}$ in subregion j.

By letting the coefficients β_i to be random, the correlations among the coefficients are added to the final (marginal) models. To avoid lack of identifiability, one possible solution is to model the process η as a mean-zero spatial temporal process and treat the residual term ϵ as a white noise process. Alternatively, we could simplify the dynamic model for β_i and let the residual term ϵ have a spatial temporal correlation structure. The later solution is the one addopted in this paper. We simplify the model for the large scale structure by representing β_i in terms of two components. One is an overall effect, and the other component explains the variation within subregions of stationarity (as defined in Section 5). Therefore, we represent

the trend surface in our application as

$$\mu(\mathbf{x},t)=\sum_{i=1}^{m}[\gamma_{i,t}+\gamma_{ij,t}]f_i(\mathbf{x},t),\ \text{ if }\ \mathbf{x}\in S_j, \tag{21}$$

where $S_j, j=1,\cdots,k$, are the subregions of stationarity used in the model for $\epsilon(\mathbf{x},t)$ (as in Section 5).

7 Application

MODELING OZONE.

We apply the nonseparable and nonstationary model introduced in this paper to observed ozone from CASTNet (Clear Air Status and Trends Network). We model the ozone data, Z, using the representation given in the previous section, $Z(\mathbf{x},t)=\mu(\mathbf{x},t)+\epsilon(\mathbf{x},t)$. The goal of this applications section is to examine the relationship between large scale synoptic weather patterns and maximum ozone levels. Our modelling framework is uniquely designed to provide detailed patterns of ozone, which can be directly compared with meteorological conditions at individual weather stations. The level of cloud cover at these weather stations is a crtical factor in the production of ozone.

As previously mentioned, we use CASTNet daily 8-hr maximum ozone data for 49 locations east of the Mississippi River (Figure 1) for the 2002 ozone season (1 May to 30 September). The percentage of missing ozone data is 2.69%. For convenience, we refer to the 8-hr maximum ozone simply as ozone. The units for ozone are ppb (parts per billion). A major advantage of using the CASTNet data is that there are meteorological observations and chemical observations at the same location. However, there are no upper air meteorological

observations, which have been shown to be important covariates in the modeling of ozone (Eder et al. (1994) and Davis et al. (1998)).

Empirical analysis suggested that normality was a reasonable assumption. Thus all computations were done using ozone in the original scale. The spatial domain that we are looking is fairly large; thus, one could not expect spatial stationarity across this domain. To overcome this problem the domain has been divided into sub-regions of stationarity using K-means cluster analysis. Figure 1 shows the final subregion configuration for ozone based on the K-means custering procedure. These subregions (or clusters) are similar to those found in Zheng et al. (2007), which used the same data set for the period 1997 to 2004. Each of these subregions has unique weather conditions over the period that is being examined; however, during the summer months the spatial variability is much less than that in the winter months. The number of subregions was determined using BIC.

We model the trend using the model in (21), where the covariates f_i are: daily minimum and average temperature at two meters, daily maximum and average solar radiation, daily average relative humidity and precipitation. Ozone is a secondary pollutant that depends on the photolysis of nitrogen dioxide for its formation. As a result, ozone formation is dependent on a number of meteorology variables that have an impact on the photolysis rate. The meteorology variables are measured at the CASTNet locations. Zheng et al.(2007) found that ozone values were most strongly correlated with maximum temperature and average daytime relative humidity. For many sites, the contribution of solar radiation was also important, but to a lesser extent. In this study, the temperature at 2 meters above the ground, solar radiation, relative humidity and precipitation were found to be important. The Henry's Law (Seinfeld and Pandis, 2006) value for ozone indicates that the dissolution

of ozone in cloud water droplets is very small and thus ozone is only weakly scavenged by rain. Precipitation is probably a surrogate for cloud cover, which would limit the photolysis rate.

PRIOR DISTRIBUTIONS.

In terms of the prior distributions for the coefficients of the weather and the other covariates, we use normal priors, $N(0, 0.01)$ (0.01 is the precision). For the uncorrelated error term component (nugget effect), we use a zero-mean normal prior, where the hyperprior for the standard deviation is a Unif(1,3). We introduce knowledge about the precision of the instruments used to measure ozone in determining this prior distribution for the ozone measurement error effect.

In terms of the prior distributions the covariance vector parameter, θ (Section 4). The prior for ϵ is a mixture of a point mass distribution at zero, a point mass distribution at one, and a $Unif(0, 1)$, using a multinomial logit prior (Gelfand et al., 2005) for the weights of the mixture. The spatial range (km), the temporal range ($days$) and the sill parameter (ppb^2) have uniform priors with support $(0, 300)$, $(0, 30)$ and $(0, 100)$, respectively. The smoothness parameter has a uniform prior with support $(0, 4)$. These hyperpriors are chosen in a non-informative fashion compatible with the data to ensure acceptable MCMC convergence.

POSTERIOR DISTRIBUTIONS.

Figure 5(a) shows the posterior mean for the overall trend, $\sum_{i=1}^{m} \gamma_{i,t} f_i(\mathbf{x}, t)$ (where the parameters are estimated by their posterior means), for the period being examined, while Figure 5(b) shows the variation in the trend for each subregion j, $\sum_{i=1}^{m} \gamma_{ij,t} f_i(\mathbf{x}, t)$ (using posterior means of the parameters). From Figure 5(a), it appears that the peak period for ozone is in July; although, strong peaks in ozone can be found at other times (e.g., the peak that

occurred in early September). From the plot for the variation in each subregion, one can see that for subregion 4 there is little variation from the overall trend, while for subregions 2 and 3 there is positive variation. Subregions 1 and 5 show negative variation.

In Table 1, we have a summary of the posterior distribution for the covariance vector parameter θ (as defined in Section 4). The posterior distribution for the sill for each subregion is also shown in Figure 6. The smallest sill values occur in subregion 5, which is to be expected given the uniformity in the weather conditions in Florida. The other subregions show greater variability which is expected since summer weather conditions would be more variable in those regions than in Florida.

The posterior distribution for the smoothness parameter is shown in Figure 7. There is not a lot of variation in the smoothness parameter between subregions.

In terms of the (epsilon) parameter that controls the potential lack of separability between space and time, we present its posterior summary in Table 1. An epsilon value of one indicates complete spatial temporal separability. This implies that the spatial structure does not change with time nor the temporal structure with space. Subregion 1 comes closest to meeting the separability criteria, while subregion 4 is most distant from an epsilon value of one. From a meteorological standpoint, one would have thought that subregion 5 would be the most likely to exhibit separability because of the spatial and temporal consistency of the weather conditions during the ozone season. In that subregion, one would not expect to see much spatial variation with time or much temporal variation with space.

OZONE MAPPING.

We examine the 8-hr maximum ozone in some detail for the period 16-23 June. This was a period of increasing ozone levels. Figures 8 and 9 show the mean of the posterior predictive

Table 1: The table below presents the mean and standard deviation of posterior distributions of the parameters

	Subregion 1	Subregion 2	Subregion 3	Subregion 4	Subregion 5
Sill	5.35 (2.2)	12.81 (6.1)	6.49 (4.3)	2.73 (1.1)	0.08 (0.03)
Spatial range	16.28 (7.3)	25.29 (9.2)	18.29 (10)	9.01 (2.3)	8.58 (2.6)
Temporal range	0.75 (2.4)	1.57 (3)	1.27 (3.3)	0.56 (0.9)	0.54 (0.2)
Smoothness	2.3 (0.1)	2.2 (0.1)	2.0 (0.1)	1.5 (0.09)	2.0 (0.2)
Separability	0.8 (0.1)	0.3 (0.05)	0.6 (0.05)	0.1 (0.04)	0.5 (0.24)

distribution for the ozone, and they illustrate how the ozone levels changed over that period of time. Figures 10 and 11 show how the synoptic scale meteorological conditions changed during the same period of time. We present the weather maps for June 16 and June 20 at 7am (local time) as an illustration, but similar maps were examined for the other days during our time window. The lines on Figures 10 and 11 are for sea-level pressure (hPa). Each station is plotted with detailed information. The wind bard shows the direction from which the wind is coming along with the speed (see Ahrens (2006) for a detailed description of the station information). The circle in the middle of the plot indicates cloud cover information. Temperature (F) is to the upper left of this circle, while the dew point temperature is to the lower left. Flow around the low pressure area is counterclockwise, while around the high pressure area the flow is clockwise. In a general sense, since ozone production relies on the photolysis process, cloud cover will have a major influence on ozone levels. Davis and Speckman (1999) found cloud cover to be an important covariate in the development of an

ozone prediction model for Houston. It should be noted that loud cover is only one factor in determining local ozone levels.

On 16 June (Figure 10), the region is dominated by a low pressure area located over the Great Lakes. The locations in and around this low pressure system and around the stationary front in south Florida have mostly overcast conditions. It is also cloudy in central Tennessee. Ozone levels are all low in all these areas. The low ozone levels in central Tennessee stand out in comparison to the areas to the south, where there is less cloud cover. There are some small mostly cloud free areas along the North Carolina coast and in portions of central Virginia and Maryland. Ozone levels are higher in these areas.

Over the next three days high pressure is in control of most, but not all, of the region. A stationary front in the southeast brought cloudy conditions to Florida and some of the states bordering it. Cloud cover persisted in portions of the northeast for much of this period. Observed ozone levels were low along the southeast coast particularly on the 18th and 19th as a result of the stationary front. As usual there were localized areas of high or low ozone depending on local conditions. During this period, temperatures ranged from the low 50s in the northern portions of the study region to the mid to upper 70s in Florida.

The map for 20 June (Figure 11) shows an approaching cold front from the west as the high pressure system moved eastward. By 21 June, a portion of the front had become stationary. As the high pressure area over the northern Great Lakes moved eastward the eastern portion of the front dropped to the south and generally maintained that position for the rest of the study period although there was some northward movement on the 23rd. After the 21st, the front is near the northern boundary of our study region. Conditions for the formation of ozone were generally favorable in the northern portion of the study region and appeared

to become increasing more unfavorable in the southern portion of the region. Cloud cover persisted along the southeast coast and inland during much of this period. These conditions contributed to the low ozone levels in this region.

CALIBRATION.

In Figure 12, predicted ozone levels for a monotoring station (41.84° latitude, −72.01° longitude), chosen randomly, which was not part of the original data analysis are compared with observed data. Our model is clearly able to capture the daily variations in maximum 8-hr ozone with regard to the amplitude and phase of the observed ozone cycle. The predicted ozone levels are 94.8% of the time within the 95% prediction bands. Similar results were obtained at other sites. We conclude our model is well calibrated.

8 Discussion

We have introduced a flexible spatial temporal model that allows for lack of stationarity and separability of the covariance function. This model includes a parameter that explains the potential lack of separability, the parameter is estimated from the data. By studying if this parameter is significantly different than 1 (separable case), using its posterior distribution or a likelihood-based approach, we can obtain a parametric test for separability. Our model is used to study spatial temporal patterns of ozone, taking into account the effect of different weather covariates. The model seems to be flexible enough to characterize complex spatial temporal dependency structures of ambient ozone.

Appendix A

Proposition: **The function in (6) is a valid spectral density.**

Proof:

We have that $f(\boldsymbol{\omega}, \tau) > 0$ everywhere, and $f(\boldsymbol{\omega}, \tau) \leq \gamma(\alpha^2\beta^2 + \beta^2|\boldsymbol{\omega}|^2 + \alpha^2\tau^2)^{-\nu}$. Thus, since

$$\begin{aligned} & \int_{\mathbb{R}^d} \int_{\mathbb{R}} \exp(i\boldsymbol{\omega}^T\mathbf{x} + i\tau t)\gamma(\alpha^2\beta^2 + \beta^2|\boldsymbol{\omega}|^2 + \alpha^2\tau^2)^{-\nu} d\boldsymbol{\omega} d\tau \\ = & \frac{\pi^{\frac{d+1}{2}}\gamma}{2^{\nu-\frac{d+1}{2}-1}\Gamma(\nu)\alpha^{2\nu-d}\beta^{2\nu-1}} \left(\alpha\sqrt{(\frac{\beta}{\alpha}t)^2 + |\mathbf{x}|^2}\right)^{\nu-\frac{d+1}{2}} \times \\ & \mathcal{K}_{\nu-\frac{d+1}{2}}\left(\alpha\sqrt{(\frac{\beta}{\alpha}t)^2 + |\mathbf{x}|^2}\right), \end{aligned} \tag{22}$$

$\int_{\mathbb{R}^d} \int_{\mathbb{R}} \exp(i\boldsymbol{\omega}^T\mathbf{x} + i\tau t) f(\boldsymbol{\omega}, \tau) d\boldsymbol{\omega} d\tau$ exists.

References

Ahrens, C.D. (2003). *Meteorology Today; An Introduction to Weather, Climate, and the Environment.* New York: Thomson-Brooks/Cole. 544pp.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* **19**, (6): 716723.

Christakos, G. (1992). *Random field models in Earth sciences.* San Diego: Academic Press.

Cressie, N., Huang, H.C., (1999). Classes of nonseparable, spatio-temporal stationary covariance functions. *J. Amer. Statist. Assoc.,* **94**, 1330 - 1340.

Davis, J.M., Eder, B.K., Nychka, D., and Yang, Q. (1998). Modeling the effects of meteorology on ozone in Houston using subregion analysis and generalized additive models. *Atmospheric Environment,* **32**, 2505-2520.

Davis, J.M. and Speckman, P. (1999). A model for predicting maximum and 8 h average ozone in Houston. *Atmospheric Environment,* **33**, 2487-2500.

Eder, B.K., Davis, J.M. and Bloomfield, P. (1994). An automated classification scheme designed to better elucidate the dependence of ozone on meteorology. *Journal of Applied Meteorology,* 33, 1182-1199.

Fuentes, M. (2005a). A formal test for nonstationarity of spatial stochastic processes. *Journal of Multivariate Analysis,* **96**, 30-55

Fuentes, M. (2005b). Testing for separability of spatial temporal covariance functions. *Journal of Statistical Planning and Inference,* **136**, 447-466.

Fuentes, M. (2002). Periodogram and other spectral methods for nonstationary spatial processes. *Biometrika,* **89**, 197-210.

Gelfand A.E., Kottas A., and MacEachern S.N. (2005). Bayesian Nonparametric Spatial Modeling with Dirichlet Process Mixing. *Journal of the American Statistical Association*, **100**, 1021–1035.

Gneiting, T., (2002). Nonseparable, stationary covariance functions for space-time data. *J. Amer. Statist. Assoc.,* **97**, 590-600.

Gradshteyn, I.S., and Ryzhik, I.M. (2007).*Table of Integrals, Series, and Products.* Alan Jeffrey and Daniel Zwillinger (eds.) Seventh edition. Academic Press, MA.

Higdon, D., Swall, J. and Kern, J. (1999), Non-stationary spatial modeling. In *Bayesian Statistics 6*, eds. J.M. Bernardo *et al.*, Oxford University Press, pp. 761–768.

Huerta, G., Sanso, B., and Stroud, J.R. (2004). A spatiotemporal model for Mexico City ozone levels. *Journal of the Royal Statistical Society, Series 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.

Matérn, B. (1986). *Spatial Variation.* Lecture Notes in Statistics, Number 36. Springer Varlag, New York. (Second edition: originally published in 1960).

Nychka, D., Wikle, C.K., and Royle, J.A., (2002). Multiresolution models for nonstationary spatial covariance functions. *Statistical Modelling: An International Journal,* **2**, 315-331.

Priestley, M. B. (1965). Evolutionary spectral and non-stationary processes. *Journal of the Royal Statistical Society. Series B.* **27**, 204-237.

Sahu, S.K., Gelfand, A.E., and Holland, D.M. (2007). High resolution space-time ozone modeling for assessing trends. *Journal of the American Statistical Society,* in press.

Sampson, P.D., and Guttorp, P. (1992), Nonparametric estimation of nonstationary spatial covariance structure. *J. Amer. Statist. Assoc.* **87**, 108-119.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics,* 6(2), 461-464.

Seinfeld, J.H., and Pandis, S.N. (2006). Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. Wiley-Interscience, NJ.

Stein, M. L. (1999). *Interpolation of Spatial Data: some theory for kriging.* Springer-Verlag, New York.

Stein, M. L., 2005. Space-time covariance functions. *J. Amer. Statist. Assoc.,* **100**, 310-321.

Stull, R.B. (1988). *An introduction to boundary layer meteorology.* Boston: Kluwer Academic Publishers.

Thompson, M.L., Reynolds, J., Cox, L.H., Guttorp, P., and Sampson, P.D. (20010). A review of statistical methods for the meteorologial adjustment of tropospheric ozone. *Atmospheric Environment,* **35**, 617-630.

U.S. Environmental Protection Agency, (2005). Evaluating ozone control programs in the eastern United States. EPA454-K-05-001, Washington, D.C.

U.S. National Research Council, (2004). Air Quality Management in the United States. National Academy Press, Washington, D.C.

Wikle, C. K., Berliner, L. M. and Cressie, N. (1998). Hierarchical Bayesian space-time models. *Environmental and Ecological Statistics,* **5**, 117 - 154.

Yaglom, A.M. (1987). *Correlation theory of stationary and related random functions I.* Springer-Verlag, New York.

Zheng, J., Swall, J., COX, W.M., and Davis, J.M. (2007). Interannual variation in meteorological adjusted ozone levels in the eastern United State: A comparison of two approaches. *Atmospheric Environment,* **41**(4) 705-716.

Figure 1: Map with the locations of the ozone monitoring stations. The stations are classified into 5 subregions of stationarity.

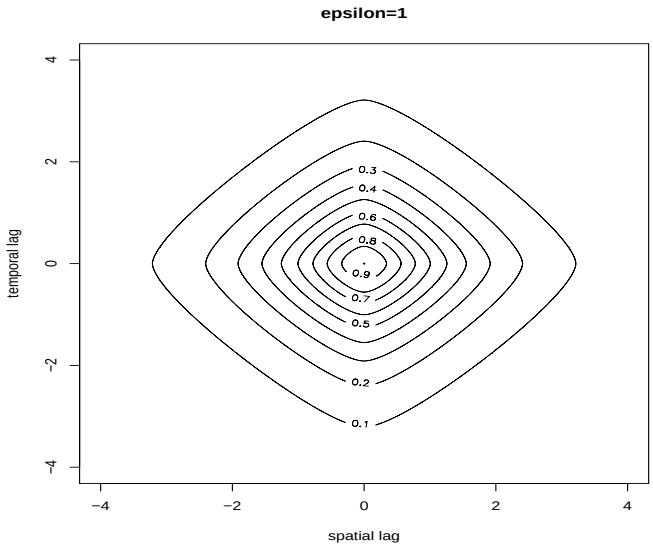


Figure 2: The contour plot for a separable spatial temporal covariance, with $\epsilon = 1$, $\gamma = \alpha = \beta = d = 1$ and $\nu = 3/2$.

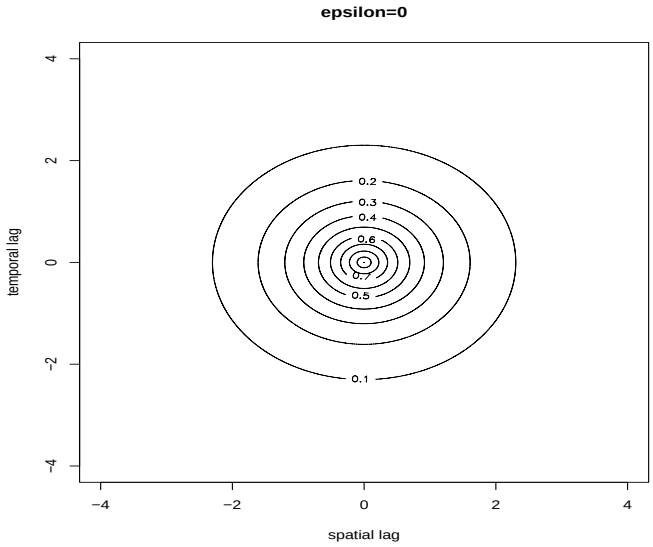


Figure 3: The contour plot for a nonseparable spatial temporal covariance, with $\epsilon = 0$, $\gamma = \alpha = \beta = d = 1$ and $\nu = 3/2$.

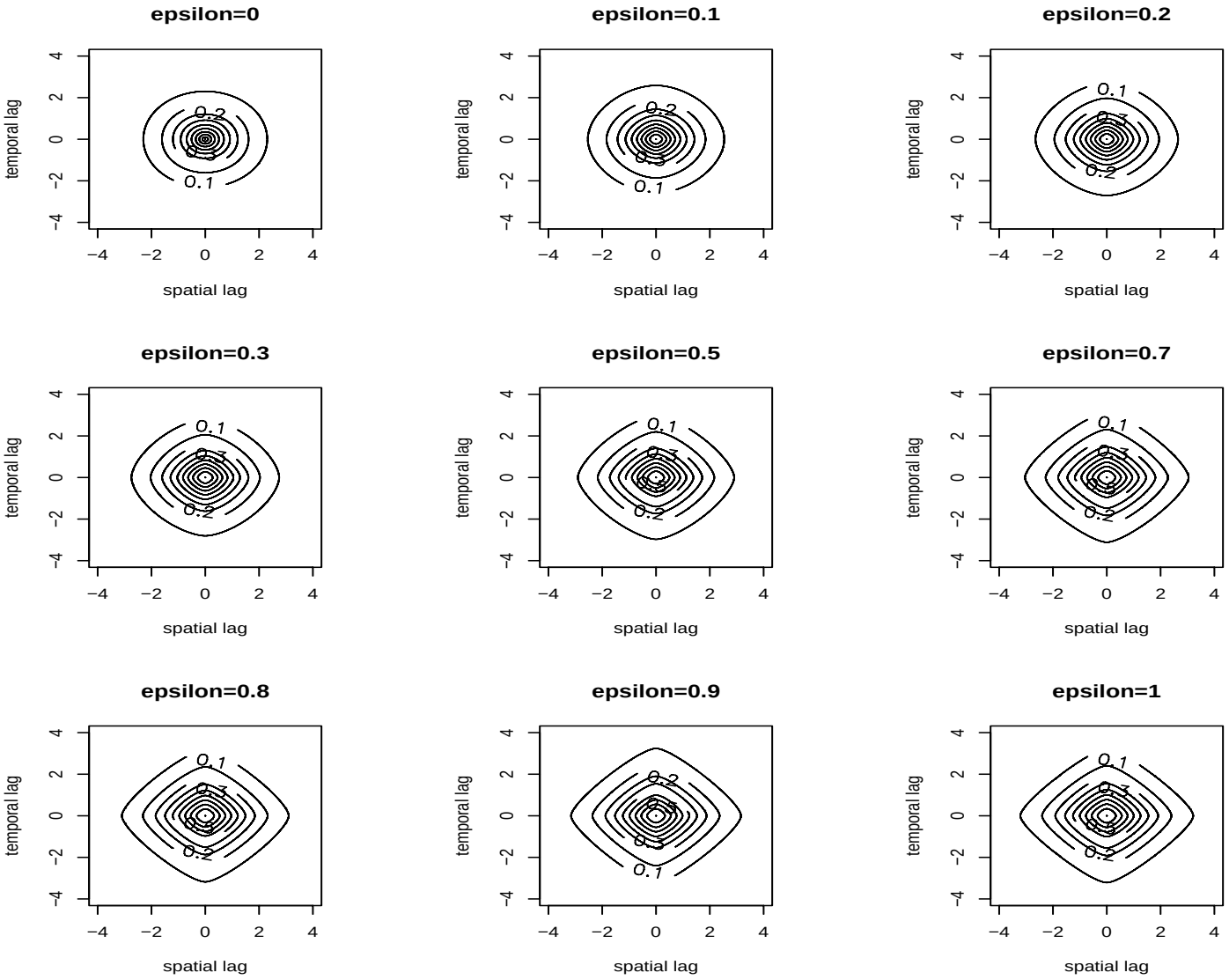


Figure 4: Contour plots for some models of spatial temporal covariances.

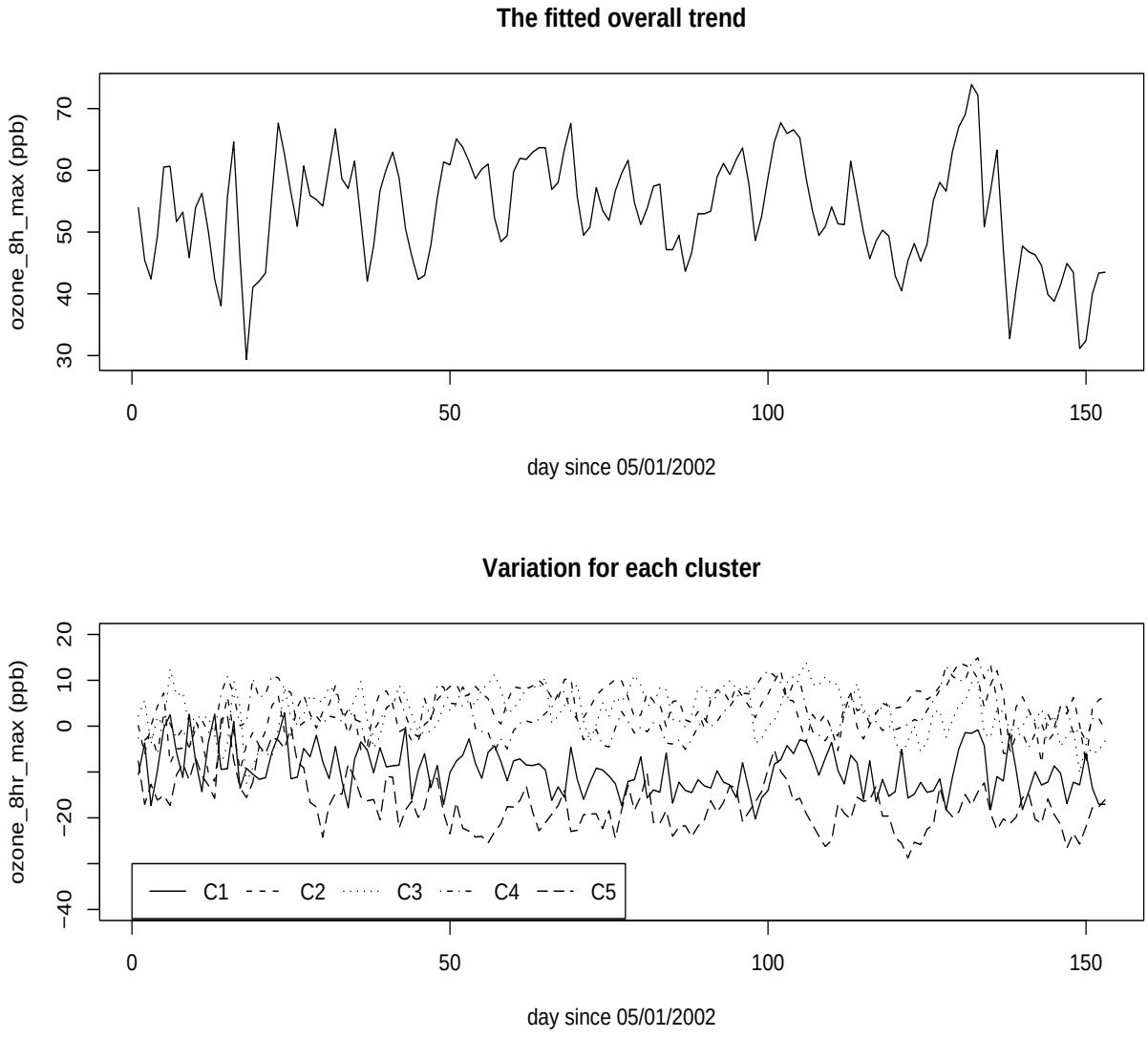


Figure 5: Posterior mean of the overall trend for the ozone values and the variation within each subregion.

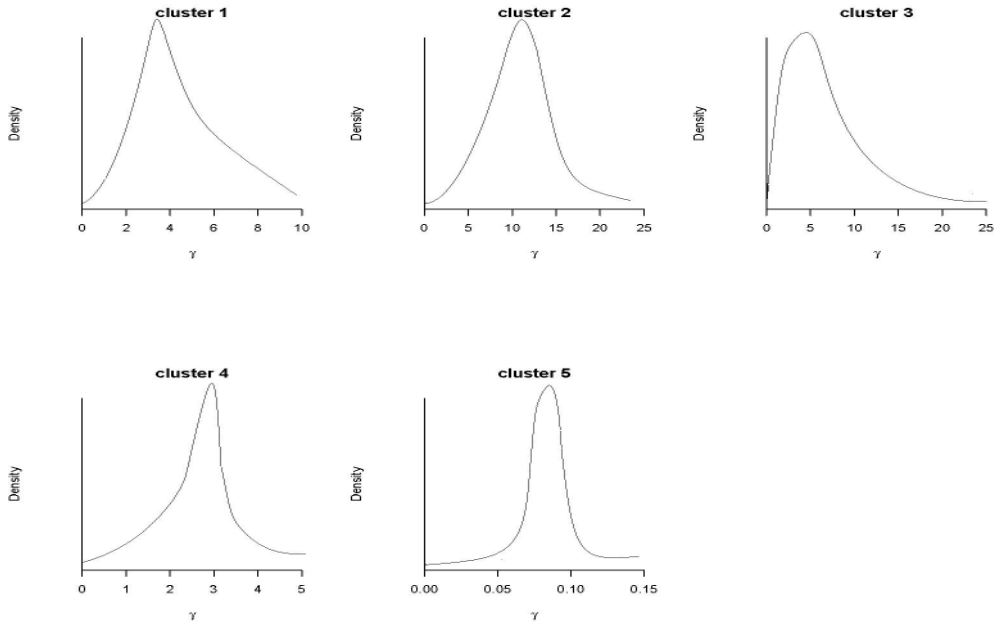


Figure 6: Posterior distribution for the sill parameter for each subregion of stationarity.

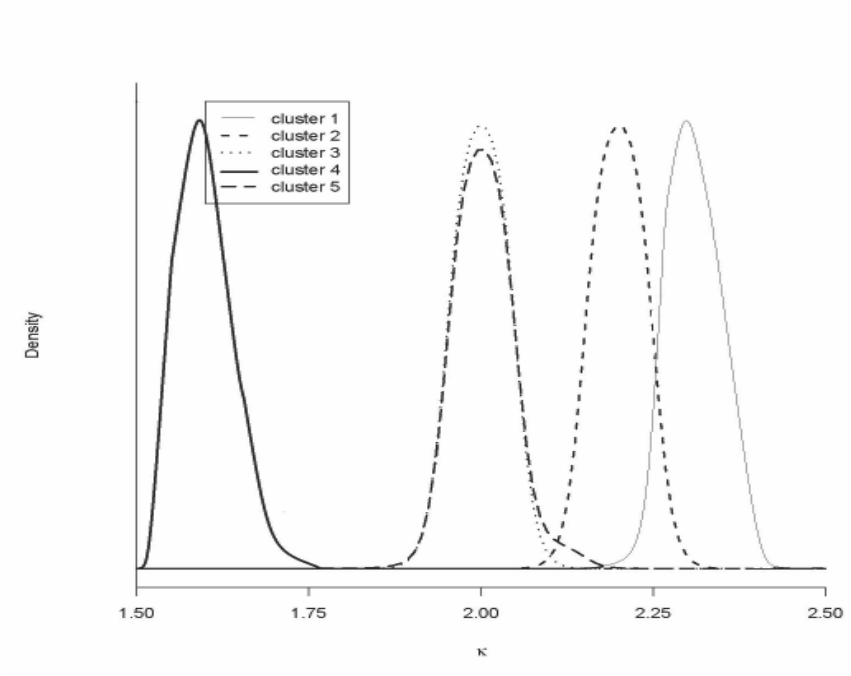
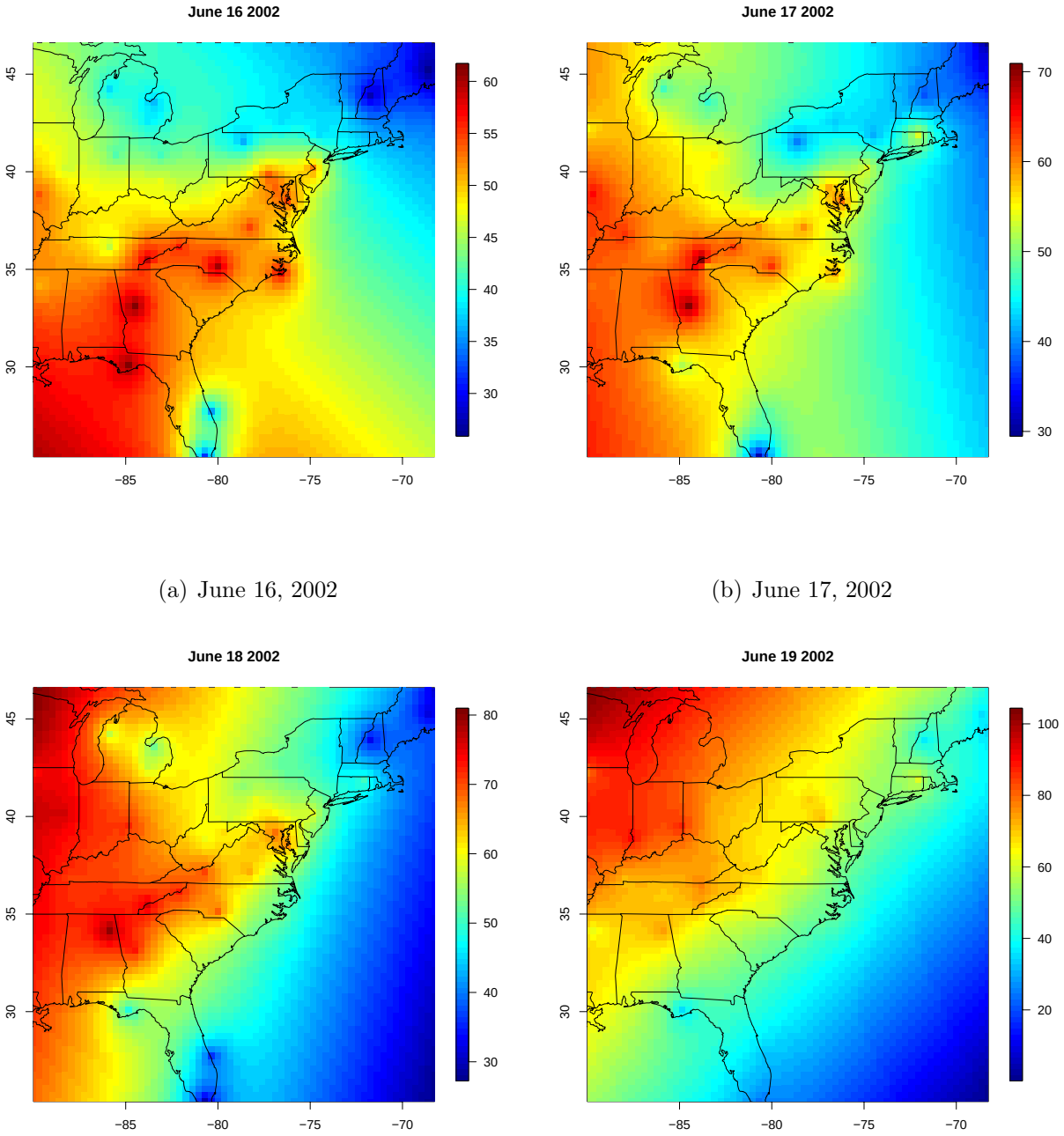


Figure 7: Posterior distribution for the smoothness parameter for each subregion of stationarity.



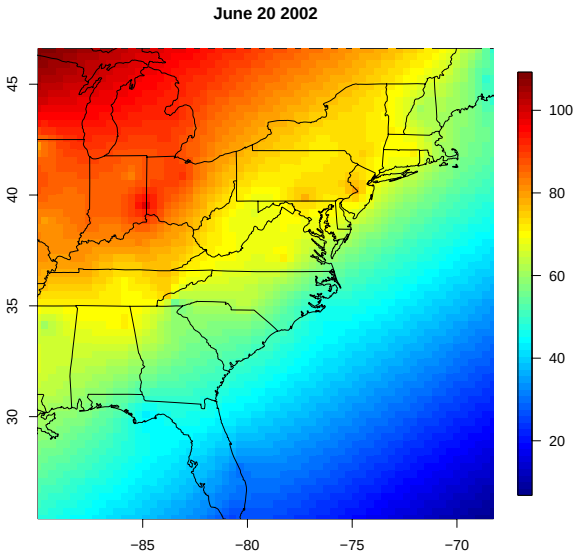
(a) June 16, 2002

(b) June 17, 2002

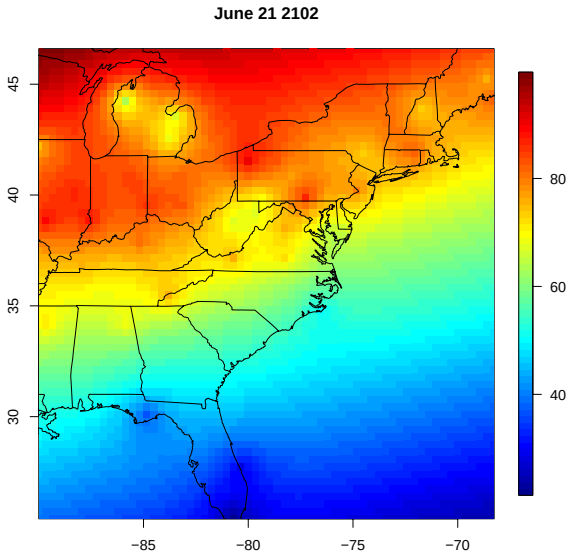
(c) June 18, 2002

(d) June 19, 2002

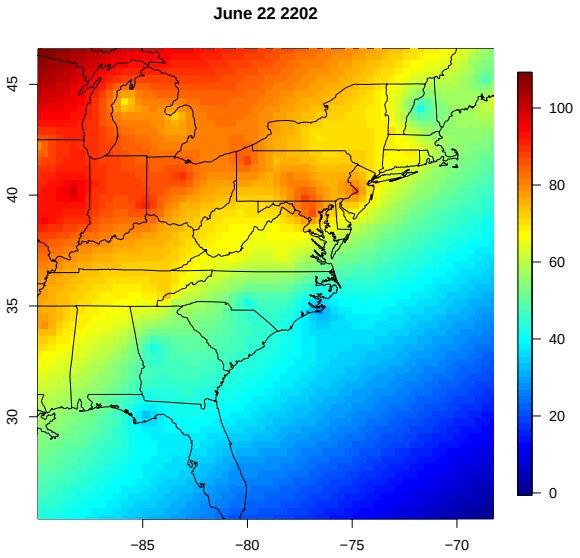
Figure 8: Ozone predictive surfaces, using the mean of the posterior predictive distribution.



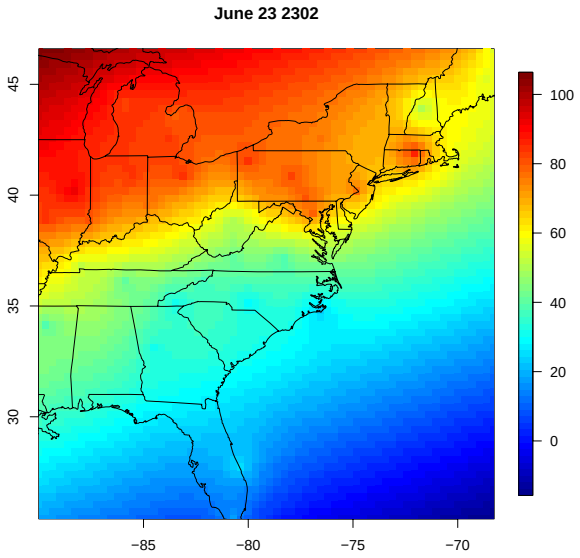
(a) June 20, 2002



(b) June 21, 2002



(c) June 22, 2002



(d) June 23, 2002

Figure 9: Ozone predictive surfaces, using the mean of the posterior predictive distribution.

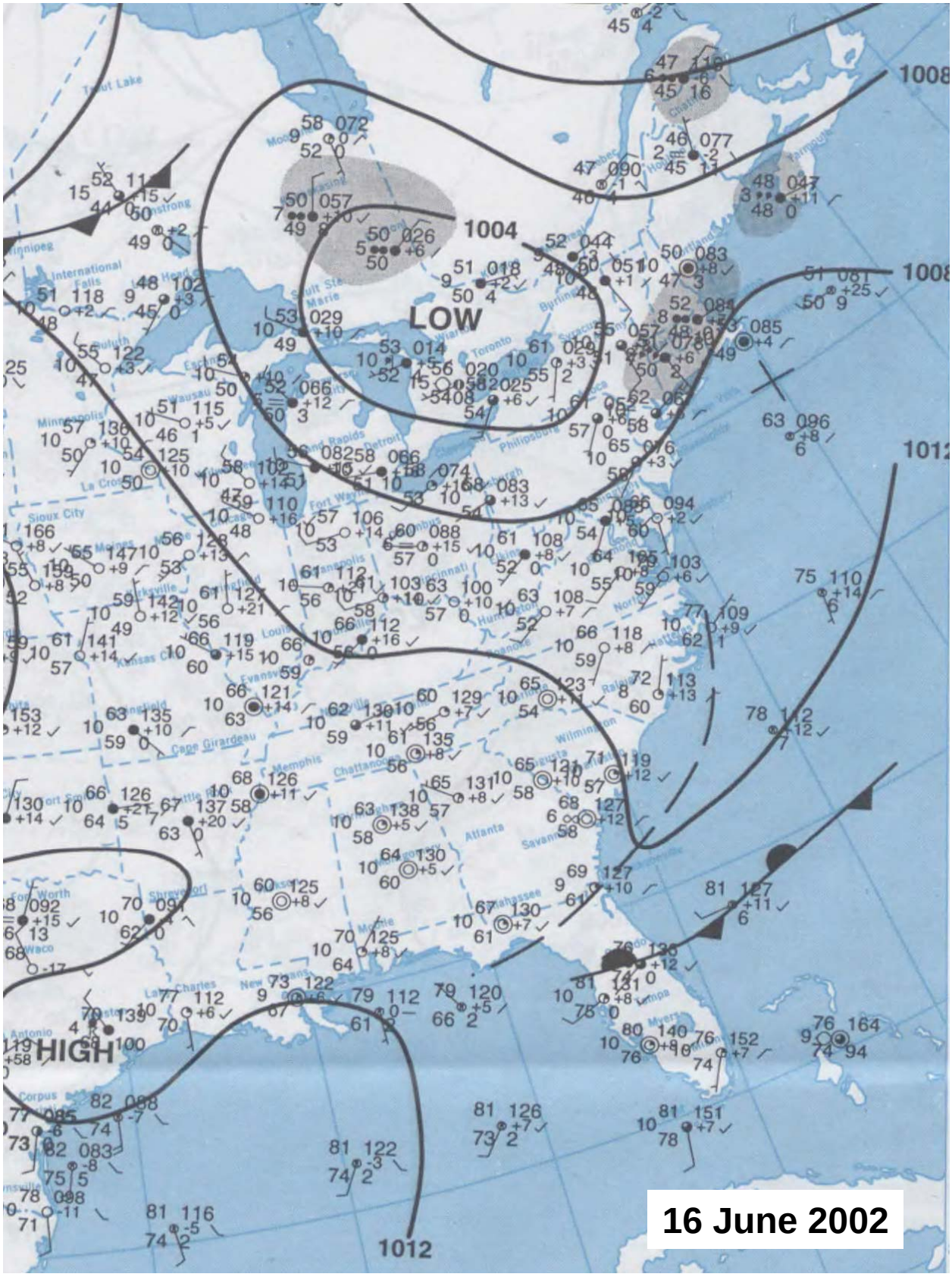


Figure 10: weather map.

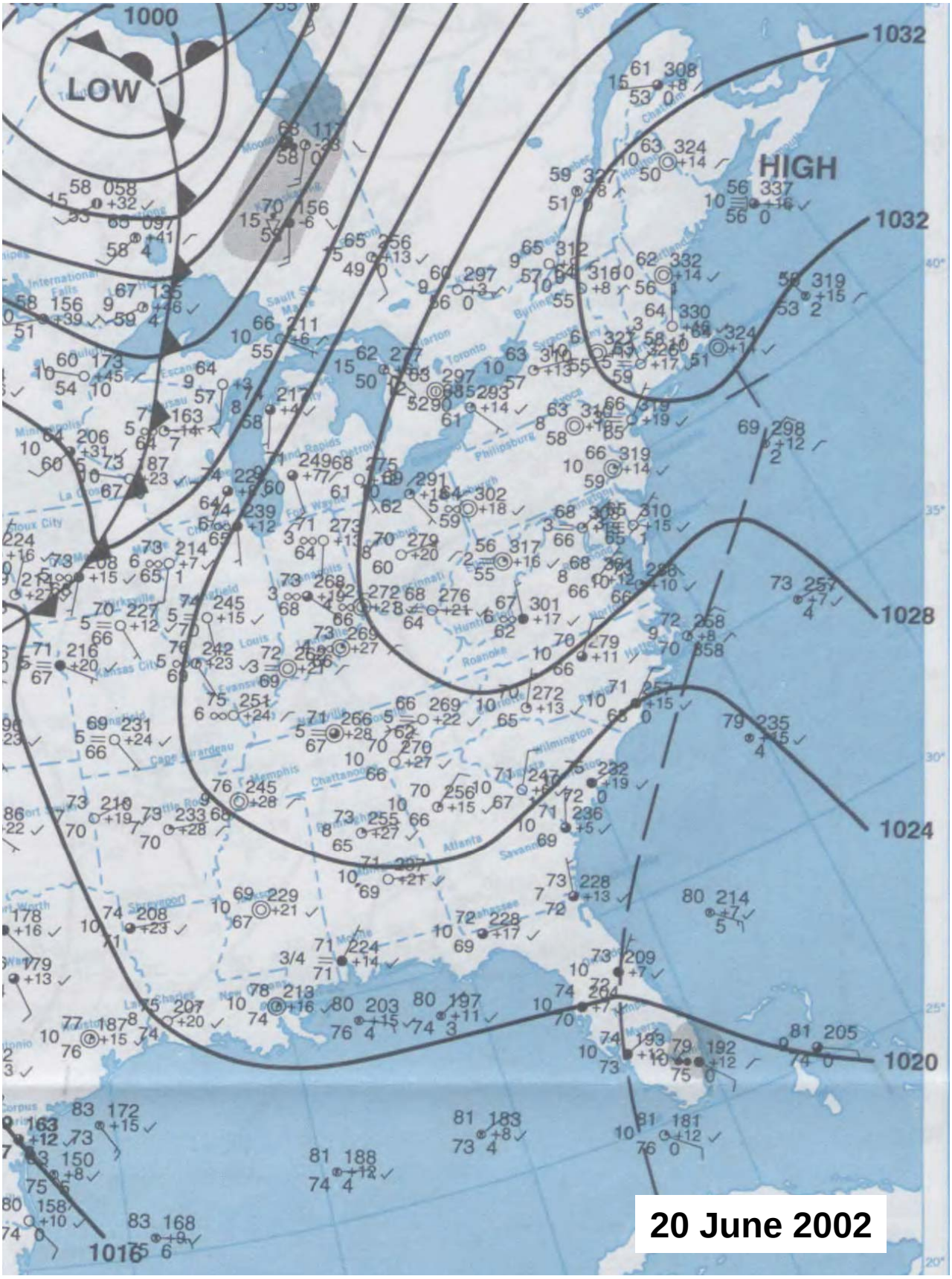


Figure 11: weather map.

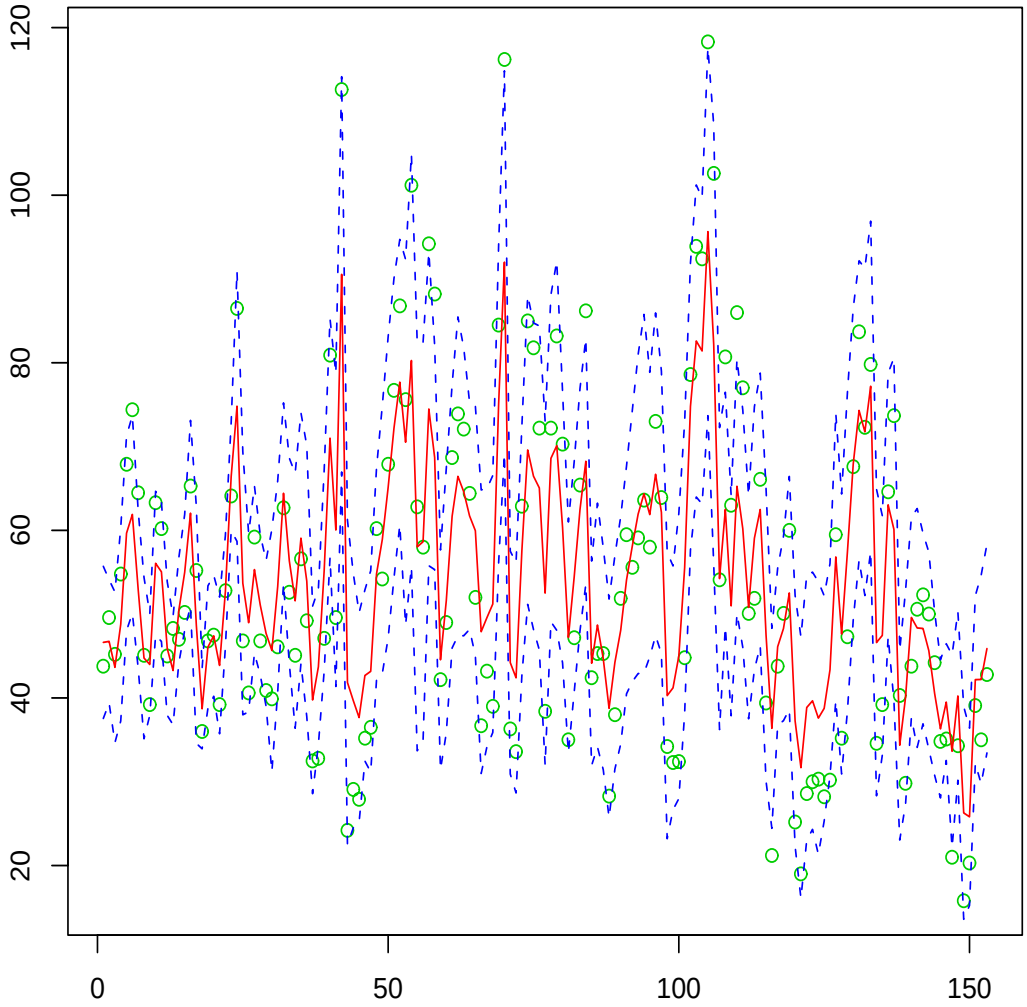


Figure 12: Validation plot at a randomly selected hold-out monitoring station (latitude=41.84 degrees, longitude= -72.01 degrees, site in CT). The observations are plotted as points, the predicted values (mean of the posterior predictive distribution) are plotted as solid lines and the 95% prediction intervals are plotted as a dashed-line.