A Spatio-Temporal Model for Mexico City Ozone Levels

Gabriel Huerta

Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131 USA. E-mail: *ghuerta@stat.unm.edu,* URL: *www.stat.unm.edu/∼ghuerta*

Bruno Sansó†

Centro de Estadística y Software Matemático, Universidad Simón Bolívar, Caracas, Venezuela and Department of Applied Mathematics and Statistics, University of California Santa Cruz, USA.
E-mail: *bruno@ams.ucsc.edu,* URL: *www.ams.ucsc.edu/∼bruno*

Jonathan R. Stroud

Department of Statistics, The Wharton School, University of Pennsylvania, 465 Jon M. Huntsman Hall, 3730 Locust Walk, Philadelphia, PA 19104 USA. E-mail: *stroud@wharton.upenn.edu,*
URL: *stat.wharton.upenn.edu/∼stroud*

Abstract

We consider hourly readings of ozone concentrations over Mexico City and propose a model for spatial as well as temporal interpolation and prediction. The model is based on regressing the observed readings on air temperature. Two harmonic components are added to account for the main periodicities that ozone presents during a given day and that are not explained by the temperature covariate. The model incorporates spatial covariance structure for the observations and the parameters that define the harmonic components. Using the dynamic linear model framework, we show how to compute smoothed means and predictive values for ozone. We illustrate the methodology on data from September, 1997.

Key Words: Tropospheric ozone; spatio-temporal modelling; Bayesian inference; Markov chain Monte Carlo; state-space models; Kriging; exponential variogram.

1. Introduction

The study of levels of tropospheric ozone is important for understanding and improving air quality in major urban areas. Environmental experts and authorities have a special interest in ozone because of its impact in diminishing health, deteriorating materials and damaging vegetation. According to environmental standards, pure air should contain less than 1% ozone and exceedingly high levels may cause eye irritation and aggravate respiratory and cardiovascular diseases.

In this paper, we focus on analyzing tropospheric ozone for Mexico City, one of the most polluted cities in the world. Located at the bottom of a valley, with approximately 20 million inhabitants, Mexico City has maintained high levels of pollution during recent years mainly due to huge amounts of motor vehicle and industrial activity. In 1986, city authorities recognized the magnitude of the problem and installed a network of monitoring stations to measure ozone, carbon monoxides and hydrocarbons. The network is named Red Automática de Monitoreo Ambiental de la Ciudad de México (R.A.M.A.). Currently, the stations that form part of the R.A.M.A. operate 365 days per year with short periods of interruptions for calibration of the measuring instruments. Each station takes measures of pollutants automatically, second by second, and the corresponding averages per hour are

†*Address for correspondence:* Bruno Sansó, Applied Mathematics and Statistics, Baskin School of Engineering, University of California, 1156 High Street, Santa Cruz CA 95064 USA.

reported to the public. In general, the units of the measurements are in parts per billion (ppb), that is, the concentration of the substance in a volume, where the volume is divided into one billion parts.

In this paper, we consider the spatio-temporal analysis of ozone time series obtained at some of the stations of the R.A.M.A. As a covariate, we use temperature which is measured only at some stations. Our goal is to propose a statistical model that forecasts temporally and interpolates spatially hourly ozone levels. Although ozone is the variable of interest, the large amounts of missing values for temperature leads us to built a space-time model also for the covariate. We elaborate our models within the Bayesian paradigm using dynamic linear models as in West and Harrison (1997) to account for temporal non-stationarities in the data. We use an efficient Markov chain Monte Carlo method to produce forecasts and spatial maps.

Previous analyses of ground-level ozone data for multiple sites, modelled jointly, appear in the paper by Carroll, Chen, George, Li, Newton, Schmiediche and Wang (1997), which uses a spatially homogeneous and temporally stationary space-time model to study ozone exposure in Texas. Their model includes temperature, wind speed and wind direction as covariables. Also, Guttorp, Meiring and Sampson (1994) built a space-time model for tropospheric ozone via the spatial deformation method of Sampson and Guttorp (1992), and placed it in a temporal framework by adding a stationary AR process at each site. As a continuation of this work, Guttorp, Meiring and Sampson (1998) present an approach to estimate hourly grid-cell surface ozone concentrations in Northern California based on observations from point monitoring sites in space for assessment of a deterministic model. This modelling approach leads to the estimation of a non-separable space time correlation structure which is spatially non-stationary, but it involves separate estimation of the temporal and spatial parts of the model. On the other hand, there is work that considers multiple sites but modelled separately. For example, Rao, Zurbenko, Neagu, Porter, Ku and Henry (1997) and Milanchus, Rao and Zurbenko (1998) consider an iterative moving-average filter that decomposes ozone into a baseline, trend and a seasonal variation site by site. An extensive and critical review of different approaches of meteorological adjustment and spatio-temporal estimation of ozone are discussed in Thompson, Reynolds, Cox and Guttorp (1999). Other approaches to space-time modelling appear in Stroud, Müller and Sansó (2001), Sansó and Guenni (2000), Tonellato (1997), Wikle, Berliner and Cressie (1999), Berliner, Royle, Wikle and Milliff (1999), Mardia, Goodall, Redfern and Alonso (1998), among others. More recently, Shaddick and Wakefield (2002) proposed a multivariate spatio-temporal model of four pollutants measured daily at eight monitoring sites in London. The approach presented in this paper is similar to ours in the sense that their models are developed within the dynamic linear modelling framework. However, their models are for daily, not hourly data and assume that the effect of the potential covariates is constant in time.

The paper is organized as follows. In Section 2, we describe the data under study. In Section 3, we find the relevant periodicities of the ozone series using a standard Bayesian regression tool and discuss on the analyses of the data, site by site. In Section 4, we present our space-time model for ozone, and a brief description of the Markov chain Monte Carlo method (MCMC) used to fit the model appears in Section 5. We leave the details of the algorithm for an appendix. In Section 6, we present the results based on the model, and Section 7 has some conclusions and extensions of the current model.

2. Data description

We consider hourly averages of ozone in ppb measured during 1997 at 19 different monitoring stations scattered irregularly in Mexico City. For 10 of these 19 stations, we also have hourly measurements corresponding to temperature in degrees centigrade. In Figure 1 we show the geographical location of the 19 stations over a topographical map of Mexico City. The scale of the x and y axes of the map is in kilometres, with the origin located in the main square of Mexico City, known as the *Zocalo*. The small box denotes the interpolation region for the space-time model of Section 5. The large box represents a $3600km^2$ region centered at *Zocalo* and that contains the Mexico City metropolitan area. The dotted contour lines correspond to altitude ranging from 2200 to 2700 metres.

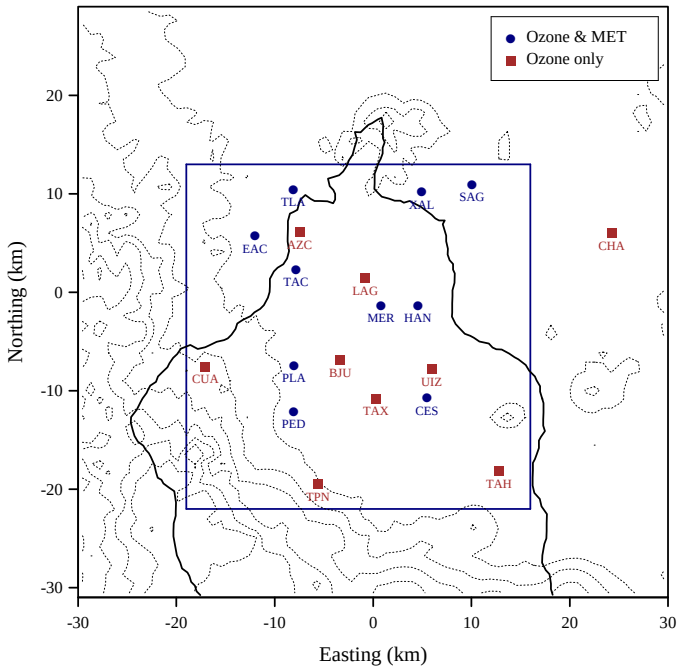

Fig. 1. Locations of 19 monitoring stations. Circles represent stations where ozone and temperature are measured. Squares represent stations where only ozone is measured. The dotted lines represent topographical contours; the solid line represents the political boundaries of Mexico, D.F.; and the inner box indicates our modelling region.

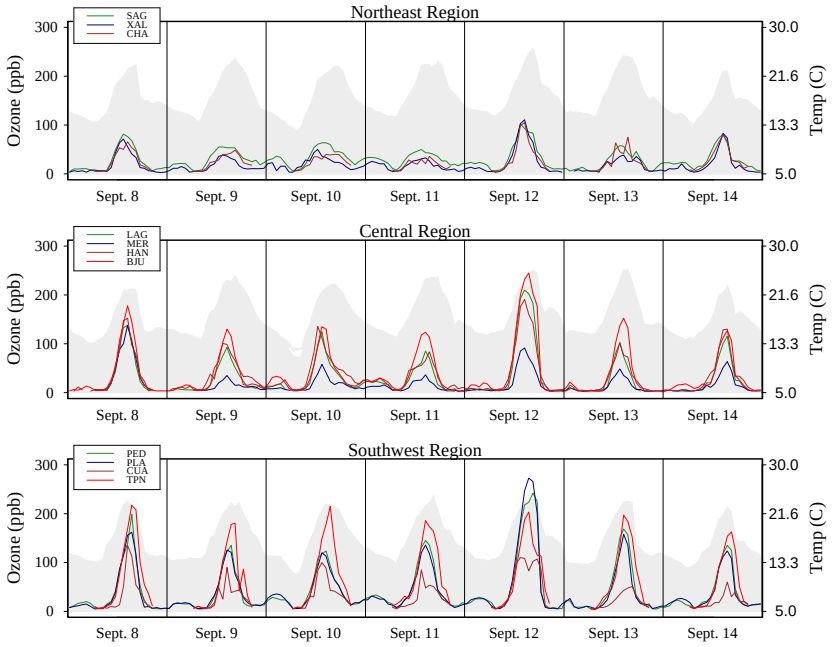

Fig. 2. Ozone and temperature time series for September 8-September 14.

Most of the data that we consider for our model appear in Figure 2. The ozone time series of 11 of the stations, from September 8 to September 14 of 1997, were classified by region and plotted with the average temperature levels as a shade. The left y-axis scale is for ppb and the right y-axis is for degrees centigrade. In general, we visualise a diurnal cycle of ozone and a very high peak during the early afternoon hours, between 1 pm and 4 pm. This high peak can be associated with the daily maximum temperature and the motor-vehicular activity in the city during the morning and early afternoon hours. Also, there is a smaller but frequent nocturnal peak. We do not notice any obvious weekly patterns or weekend effects but there are changes from one day to the other that suggest that, even after considering daily cycles, there is lack of stationarity in the series.

Figure 2 illustrates some of the spatial variability of ozone in Mexico City. In fact, a study of the hourly means over all the year of 1997 for each of the 19 monitoring stations reveals that the variability of the mean level across stations is important and that the cyclical behaviour of the series differs in amplitude according to location.

As is usual for ozone measurements, the distribution of the data has an asymmetric shape that suggests the use of a transformation previous to building models based on the assumption of a normal distribution for the error. The two most common transformations in the literature for ozone data are the square root and the natural logarithm. Thompson, Reynolds, Cox and Guttorp (1999) report a summary of the transformations used by different authors that analyze ozone series. In this paper, we consider the square root transformation for the data. This is supported by analyses of the distribution of the observed values as well as the behaviour of the residuals of the models that we propose in the following sections.

3. Univariate Analysis

An accurate specification of the cyclical behaviour of the ozone series is a key feature for modelling such data. To make inferences on periodicities we used the *Bayesian periodogram* introduced by Bretthorst

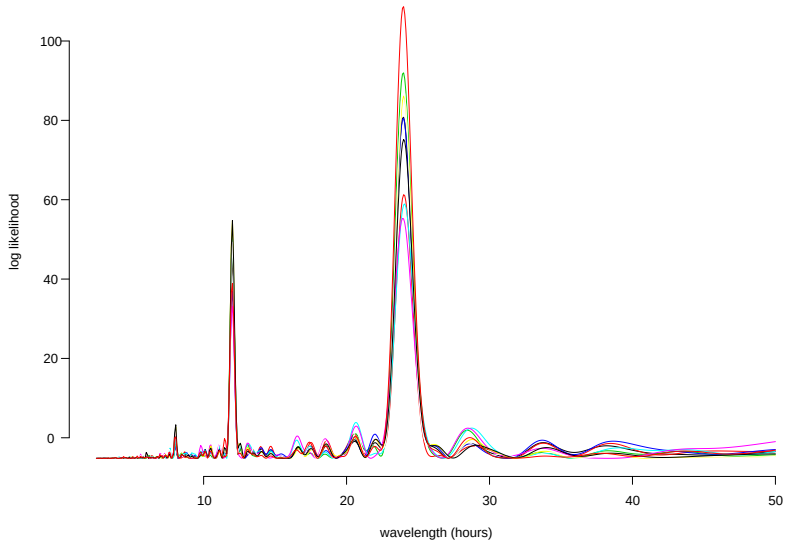

Fig. 3. Bayesian periodogram of ozone concentrations at ten monitoring stations in Mexico City, during the first two weeks of September 1997.

(1988). The Bayesian periodogram is defined as the marginal log-likelihood of the regression model

$$Y_t = a\ \cos(2\pi t/\lambda) + b\ \sin(2\pi t/\lambda) + \epsilon_t,$$

with a non-informative or *reference* prior $p(a, b, \sigma^2) \propto 1/\sigma^2$, where t indexes time, $\epsilon_t \sim N(0, \sigma^2)$ and λ is the underlying periodicity or wavelength of the process Y_t.

Figure 3 shows the Bayesian periodograms for the square-root ozone time series for the month of September 1997 and taken at the monitoring stations that measure temperature. The range of values for λ shown in the Figure is between 0 and 50 hours. The general pattern of all the periodograms is similar with a peak corresponding to a daily cycle with wavelengths of 24 hours and a peak corresponding to a harmonic cycle with a wavelength of 8 hours. Some of the stations present a smaller peak at eight hours. We think this small peak may be related to early morning reactions between VOC and NO_X. We also evaluated the periodograms for values of λ greater than 50 hours and could not find any other relevant peaks. Specifically, we could not find any cycles associated with weekend effects.

Our initial modelling of the data consisted of relating the square-root ozone readings to the temperature using a time series model fitted station by station. Many of the modelling strategies adopted for the spatio-temporal case are the product of the univariate analysis. Given the apparent non-stationarity of the data, we considered univariate dynamic regressions. This has as response variable the square root of ozone Y_t, the temperature $Z_t, t = 1, \ldots, T$ and two sin/cos terms with periodicities $2\pi/24$ and $2\pi/12$ favoured by the Bayesian periodogram.

The model can be written as

$$\begin{aligned} Y_t &= S_t'\alpha_t + Z_t\gamma_t + \varepsilon_t, & \varepsilon_t &\sim N(0, V) \\ \alpha_t &= \alpha_{t-1} + \boldsymbol{\omega}_{1t}, & \boldsymbol{\omega}_{1t} &\sim N(0, \mathbf{W}_{1t}) \\ \gamma_t &= \gamma_{t-1} + \omega_{2t}, & \omega_{2t} &\sim N(0, W_{2t}) \end{aligned}$$

where $S_t' = (\cos(\pi t/12), \sin(\pi t/12), \cos(\pi t/6), \sin(\pi t/6))$ and $\alpha_t' = (\alpha_{1t}, \alpha_{2t}, \alpha_{3t}, \alpha_{4t})$. The parameter γ_t is a scalar, ε_t, $\boldsymbol{\omega}_{1t}$ and ω_{2t} are assumed independent of each other. The variance of ε_t is unknown but equal for all t, whilst the variance of $\boldsymbol{\omega}_{1t}$ and ω_{2t} is modelled with discount factors (see West and Harrison, 1997). We disregard harmonics with higher frequencies since they do not appear relevant in the periodogram analysis of the data.

The model described above is a special case of a dynamic linear model as presented in West and Harrison (1997) and can be fitted using the well known updating and filtering equations, which are related to the Kalman filter. We fitted the model to each one of the ten stations with measurements of temperature. We used a fairly flat initial prior for the parameters at time $t = 0$ and a global discount factor of 0.97 for the evolution covariances.

We found that this dynamic regression has a reasonable predictive performance, particularly for 24 hour-step ahead forecasts. Also, we notice that the estimates for γ_t are quite variable across time but relatively less variable across stations. On the other hand, the component parameters α_t are quite variable across station and time, but roughly, are such that the corresponding phase is constant across stations whilst the amplitude varies substantially.

We fitted dynamic models that did not include the two harmonic components, hoping that the periodicities in the data would be captured by the cyclical behaviour of the temperature, but the model residuals presented very high autocorrelations or severe non-normality. Similarly, we observed that the temperature coefficient is significant for most time steps and that dropping temperature resulted in a severe lack of fit. On the other hand the addition of an eight hour harmonic, which is hinted by the periodogram, produced no significant improvements in the fit, nor did the addition of atmospheric pressure or wind as covariates.

We studied the spatial behaviour of the residuals from the station by station fits, plotting the empirical variograms and fitting models with geometric anisotropy that depended on the direction of the wind. None of the models we considered produced substantial improvements over the model that consisted of a Gaussian field with isotropic exponentially decaying covariance.

4. Spatio-Temporal Model

Let Y_{it} denote the observed square root ozone concentration, for each station i and time t and let Z_{it} be the temperature at time t and station i, $t = 1, \ldots T$, $i = 1, \ldots, n$. Let α_{it} denote a 4-dimensional vector of seasonal coefficients for station i corresponding to a seasonal component $S_t^{'}$ consisting of sine and cosine terms as introduced in the previous section.

The general space-time model that we propose for ozone data is given by

$$Y_{it} = \beta_t^y + S_t^{'}\alpha_{it} + Z_{it}\gamma_t + \varepsilon_{it}$$

where the errors, ε_{it}, are assumed spatially correlated with a Gaussian distribution. Here β_t^y defines a canonical spatial trend and $Z_{it}\gamma_t$ is the effect of the covariate on Y_{it}. Note that the coefficient related to the covariate is assumed equal for all stations, whilst each station has its own set of seasonal parameters. This is justified by the results obtained with the univariate time series analyses. Clearly, this general space-time model has a very large number of parameters, since at each time t there are 4 parameters for each station, a common one for the covariate, plus the parameters that define the spatial correlation.

A substantial reduction in the number of parameters is achieved by assuming that the amplitudes of each cyclical component are different, but the phases are very similar between stations and almost constant in time. This assumption is supported by the univariate models fitted station by station.

We can thus consider a modification of the model given by

$$Y_{it} = \beta_t^y + S_t^{'}(\mathbf{a})\alpha_{it} + Z_{it}\gamma_t + \varepsilon_{it}^y;$$

but now $\alpha_{it} = (\alpha_{1it}, \alpha_{2it})$ is a vector of dimension 2 and $S_t^{'}(\mathbf{a}) = (S_{1t}(a_1), S_{2t}(a_2))$ where $S_{jt}(a_j) = \cos\left(\frac{\pi jt}{12}\right) + a_j \sin\left(\frac{\pi jt}{12}\right)$, $j = 1, 2$ defines the cyclical component in the data and $\mathbf{a} = (a_1, a_2)'$.

Thus $\alpha_{jit}^2(1+a_j^2)$ is the amplitude of the j-th periodicity of the i-th station at time t, and $\tan^{-1}(a_j)$ is its phase, which is independent of i and t.

In vector form, our modified space-time model can be written as:

$$\mathbf{Y}_t = 1_n\beta_t^y + S_{1t}(a_1)\boldsymbol{\alpha}_{1t} + S_{2t}(a_2)\boldsymbol{\alpha}_{2t} + \mathbf{Z}_t\gamma_t + \boldsymbol{\epsilon}_t^y$$

where $\mathbf{Y}_t^{'} = (Y_{1t}, \ldots, Y_{nt})$, $1_n^{'} = (1, \ldots, 1)$, $\boldsymbol{\alpha}_{it}^{'} = (\alpha_{i1t}, \alpha_{i2t}, \ldots, \alpha_{int})$, $i = 1, 2$ and $\boldsymbol{\epsilon}_t^y = (\varepsilon_{1t}^y, \varepsilon_{2t}^y, \ldots, \varepsilon_{nt}^y)$.

Additionally, we assume that $\boldsymbol{\epsilon}_t^y \sim N(0, \sigma_y^2\mathbf{V}^y)$ with $\mathbf{V}^y = \exp(-\mathbf{D}/\lambda_y)$, $\mathbf{D}$ the matrix of distances in kilometers between stations, $\exp(\mathbf{D})$ denotes the exponentiation of each of its components and λ_y is a positive range parameter. That is, we assume that the errors are normally distributed with an exponential covariance function.

Furthermore, the parameter evolutions of the model are defined as

$$\begin{aligned}
\beta_t^y &= \beta_{t-1}^y + \omega_t^y, & \omega_t^y &\sim N(0, \sigma_y^2\tau_y^2) \\
\boldsymbol{\alpha}_{it} &= \boldsymbol{\alpha}_{it-1} + \boldsymbol{\omega}_t^{\alpha_i}, & \boldsymbol{\omega}_t^{\alpha_i} &\sim N(0, \sigma_y^2\tau_i^2\mathbf{W}^{\alpha_i}) \\
\gamma_t &= \gamma_{t-1} + \omega_t^{\gamma}, & \omega_t^{\gamma} &\sim N(0, \sigma_y^2\tau_\gamma^2)
\end{aligned}$$

where τ_y^2, τ_γ^2, $\tau_i^2, i = 1, 2$ are fixed constants, and $\mathbf{W}^{\alpha_i} = \exp(-\mathbf{D}/\lambda_i)$ is the evolution variance-covariance matrix, which depends on the spatial range parameters, $\lambda_i, i = 1, 2$. Therefore, we are incorporating spatial dependency on the parameters that determine the seasonal terms.

Notice that the above model assumes that the observations for temperature are known at each time and station so it is a model for $\mathbf{Y}_t$ conditional on $\mathbf{Z}_t$. Values of Z_{it} at arbitrary locations and times are needed to interpolate the values of Y_{it} both temporally and spatially and they may not be available. Then, we also formulate a spatio-temporal model on Z_{it} defined as

$$Z_{it} = \beta_t^z + h_i\eta_t + (1, 0, 1, 0)\boldsymbol{\delta}_t + \boldsymbol{\epsilon}_{it}^z,$$

with β_t^z a canonical spatial trend on Z_{it}, h_i the height of station i, η_t the altitude coefficient at time t, $\boldsymbol{\delta}_t = (\delta_{1t}, \delta_{2t}, \delta_{3t}, \delta_{4t})$, and $\boldsymbol{\epsilon}_{it}^z$ is the error.

In vector form, the model is

$$\mathbf{Z}_t = 1_n\beta_t^z + \mathbf{H}'\eta_t + \mathbf{E}'\boldsymbol{\delta}_t + \boldsymbol{\epsilon}_t^z$$

where $\mathbf{Z}_t^{'} = (Z_{1t}, Z_{2t}, \ldots, Z_{nt})$; $1_n^{'} = (1, 1, \ldots, 1)$, $\mathbf{H} = (h_1, h_2, \ldots, h_n)$, $\mathbf{E}'$ is a $n \times 4$ dimensional matrix where all the rows are equal to the vector $(1, 0, 1, 0)$ and $\boldsymbol{\epsilon}_t^z = (\boldsymbol{\epsilon}_{1t}^z, \ldots, \boldsymbol{\epsilon}_{nt}^z)$ is the vector of errors with $\boldsymbol{\epsilon}_t^z \sim N(0, \sigma_z^2\mathbf{V}^z)$ where $\mathbf{V}^z = \exp(-\mathbf{D}/\lambda_z)$.

A univariate time series analysis of the temperatures reveals that there are time changing differences in mean values due to the height of the station. On the other hand the seasonality is fairly homogeneous in space. Therefore, we are considering a model for $\mathbf{Z}_t$ with the same seasonal effect for all locations.

Additionally, the parameter evolutions are defined as

$$\begin{aligned}
\beta_t^z &= \beta_{t-1}^z + \omega_t^z, & \omega_t^z &\sim N(0, \sigma_z^2\tau_z^2), \\
\eta_t &= \eta_{t-1} + \omega_t^\eta, & \omega_t^\eta &\sim N(0, \sigma_z^2\tau_\eta^2), \\
\boldsymbol{\delta}_t &= G\boldsymbol{\delta}_{t-1} + \boldsymbol{\omega}_t^\delta, & \boldsymbol{\omega}_t^\delta &\sim N(0, \sigma_z^2\tau_\delta^2\mathbf{I})
\end{aligned}$$

where τ_z^2, τ_η^2 and τ_δ^2 are fixed constants, G is 4×4 block diagonal matrix with blocks of the form

$$G_j = \begin{pmatrix} \cos(\pi j/12) & \sin(\pi j/12) \\ -\sin(\pi j/12) & \cos(\pi j/12) \end{pmatrix}, \; j = 1, 2.$$

which defines the seasonal-component on $\mathbf{Z}_t$, $\mathbf{I}$ is a 4×4 identity matrix.

The model is completed by specifying a prior distribution on the range parameters of the covariance matrix, the scale parameters at the observation level and the phase parameters $\mathbf{a}$. A particular prior specification is presented in Section 6.

In summary, our new spatio-temporal model for ozone is defined by the two equations

$$\begin{aligned} \mathbf{Y}_t &= 1_n\beta_t^y + S_{1t}(a_1)\boldsymbol{\alpha}_{1t} + S_{2t}(a_2)\boldsymbol{\alpha}_{2t} + \mathbf{Z}_t\gamma_t + \boldsymbol{\epsilon}_t^y \\ \mathbf{Z}_t &= 1_n\beta_t^z + \mathbf{H}'\eta_t + \mathbf{E}'\boldsymbol{\delta}_t + \boldsymbol{\epsilon}_t^z \end{aligned}$$

This model is in the dynamic linear model (DLM) or state-space form notation of West and Harrison (1997). Thus, conditional on $\mathbf{a}$ and all the hyperparameters that define the covariance structure at both observational and evolution levels, the filtering and recurrence equations of the DLM produces predictive values and retrospective inferences for the observed data. Formal Bayesian inference on some of the hyperparameters, for example, the range parameters λ_y and λ_z, leads to the *Forward-Filtering Backward-Sampling* algorithm as we describe now.

5. MCMC algorithm and spatial interpolation

Under the space-time model just presented, posterior, predictive and interpolation analyses are available via MCMC methods. The structure of relevant conditional distributions is briefly outlined here. Further details appear in Appendix A. First we give some notation. Define $\boldsymbol{\theta}_t^y = (\beta_t^y, \boldsymbol{\alpha}_{1t}', \boldsymbol{\alpha}_{2t}', \gamma_t)$ as the ozone state vector, and $\boldsymbol{\theta}_t^z = (\beta_t^z, \eta_t, \boldsymbol{\delta}_t')$ as the temperature state vector at time t, and set $\boldsymbol{\theta}^y = (\boldsymbol{\theta}_0^y, \ldots, \boldsymbol{\theta}_T^y)$, $\boldsymbol{\theta}^z = (\boldsymbol{\theta}_0^z, \ldots, \boldsymbol{\theta}_T^z)$. $\mathbf{Y} = (Y_1, \ldots, Y_T)$ and $\mathbf{Z} = (Z_1, \ldots, Z_T)$. Let the superscripts "$o$" and "$u$" denote the observed and unobserved data, respectively. Posterior inferences are then based on summarising the joint posterior distribution:

$$p(\mathbf{Y}^u, \boldsymbol{\theta}^y, \mathbf{a}, \sigma_y^2, \lambda_y, \mathbf{Z}^u, \boldsymbol{\theta}^z, \sigma_z^2, \lambda_z \mid \mathbf{Y}^o, \mathbf{Z}^o)$$

Our MCMC is a block sampling scheme in which we iteratively simulate each component of the joint posterior distribution. Considering blocks of parameters is a key issue to obtain convergence and computational efficiency. Related work for Markov random field models appears in Rue (2001) and Knorr-Held and Rue (2002). The full conditional distributions are given below. The prior distributions for the unknown parameters are given in Section 6.

5.1. Sampling the state vectors, error variances and range parameters

Using the state-space equations for $\mathbf{Y}_t$ and $\mathbf{Z}_t$, we sample $(\boldsymbol{\theta}^y, \sigma_y^2, \lambda_y)$ and $(\boldsymbol{\theta}^z, \sigma^z, \lambda_z)$ as blocks. Dropping the y and z sub-indexes and the conditioning on all other parameters and observations, we note that $p(\boldsymbol{\theta}, \sigma^2, \lambda) = p(\boldsymbol{\theta}|\sigma^2, \lambda)p(\sigma^2|\lambda)p(\lambda)$. We can obtain samples of $p(\boldsymbol{\theta}|\sigma^2, \lambda)$ using Forward Filtering Backward Sampling. That is, we apply the filtering recurrences for the DLMs defined for the response and the covariate and then sample each element of $\boldsymbol{\theta}$ recursively from $t = T, \ldots, 1$ as in Carter and Kohn (1994) or Frühwirth-Schnatter (1994). $p(\sigma^2|\lambda)$ corresponds to an inverse gamma distribution and $p(\lambda)$ can be easily evaluated for each λ. The details appear in Appendix A. We then propose the following Metropolis-Hastings scheme:

- Choose a jumping distribution for λ, say, $q(\lambda^*|\lambda)$.
- Sample $(\boldsymbol{\theta}^*, \sigma^{2*}, \lambda^*) \sim p(\boldsymbol{\theta}|\sigma^{2*}, \lambda^*)p(\sigma^{2*}|\lambda^*)q(\lambda^*|\lambda)$
- Let
$$\alpha = \min\left\{1, \frac{p(\lambda^*)q(\lambda|\lambda^*)}{p(\lambda)q(\lambda^*|\lambda)}\right\}$$
- Accept $(\boldsymbol{\theta}^*, \sigma^{2*}, \lambda^*)$ with probability α.

Note that the acceptance ratio α does not depend on $\boldsymbol{\theta}^*$ and σ^{2*}, so we can avoid the backward sampling of $\boldsymbol{\theta}$ if λ^* is not accepted. The former accounts for a substantial reduction in the computational cost of the MCMC. Our implementation of the algorithm considers a transition density $q(\lambda^*|\lambda)$ defined on the $\log(\lambda)$ scale that follows a normal distribution centred on the last sampled value.

5.2. Conditional for missing values

Since we are assuming that the error terms $\boldsymbol{\epsilon}_t^y$ and $\boldsymbol{\epsilon}_t^z$ are normally distributed, given all the model parameters, the joint distributions of $(\mathbf{Y}_t^o, \mathbf{Y}_t^u)$ and $(\mathbf{Z}_t^o, \mathbf{Z}_t^u)$ are multivariate normal respectively. Then, the conditional distributions of $\mathbf{Y}_t^u$ given $\mathbf{Y}_t^o$ and $\mathbf{Z}_t^u$ given $\mathbf{Z}_t^o$ are also normal with moments of well known form.

5.3. Conditional for constant phase parameters

Note that the constant phase parameter $\mathbf{a}$ only appears in the model equation for $\mathbf{Y}_t$. Given $\boldsymbol{\theta}^y$, σ_y^2 and λ_y, it can be easily shown that the conditional likelihood for $\mathbf{a}$ is given by the regression model

$$\mathbf{K}_t = a_1 \sin\left(\frac{\pi t}{12}\right)\boldsymbol{\alpha}_{1t} + a_2 \sin\left(\frac{\pi t}{6}\right)\boldsymbol{\alpha}_{2t} + \boldsymbol{\epsilon}_t^y$$

where $\mathbf{K}_t = \mathbf{Y}_t - 1_n\beta_t^y - \cos(\pi t/12)\,\boldsymbol{\alpha}_{1t} - \cos(\pi t/6)\,\boldsymbol{\alpha}_{2t} - \mathbf{Z}_t\gamma_t$ is computed at the current value of $\boldsymbol{\theta}^y$. Then, the conditional likelihood function for $\mathbf{a}$ has a bivariate normal kernel. Computations are then particularly easy if the prior is a conjugate bivariate normal or a standard reference prior $p(\mathbf{a}) \propto 1$.

5.4. Spatial interpolation

Let $\mathbf{s}$ denote an unobserved site; Y_t^s and Z_t^s are the unobserved square-root ozone and temperature at such site at time t. Spatial interpolation requires producing samples of (Y_t^s, Z_t^s) from its posterior distribution. Obtaining such a sample is done by first sampling Z_t^s from

$$p(Z_t^s|\beta_t^z, \eta_t, \boldsymbol{\delta}_t, \sigma_z^2, \lambda_z, \mathbf{Z}_t),$$

and then sampling the parameters $(\alpha_{1t}^s, \alpha_{2t}^s)$ that define the seasonal term at site $\mathbf{s}$ from

$$p(\alpha_{it}^s|\alpha_{it-1}^s, \boldsymbol{\alpha}_{it}, \boldsymbol{\alpha}_{it-1}),\ i = 1, 2.$$

Finally, we sample Y_t^s from

$$p(Y_t^s|\alpha_{1t}^s, \alpha_{2t}^s, Z_t^s, \beta_t^y, \boldsymbol{\alpha}_{1t}, \boldsymbol{\alpha}_{2t}, \gamma_t, \sigma_y^2, \lambda_y, \mathbf{Y}_t, \mathbf{Z}_t).$$

Since the state vector $\boldsymbol{\theta}_z$ is independent of location, the conditional distribution of Z_t^s given $\boldsymbol{\theta}_z$ and $\mathbf{Z}_t$ is derived from the observation equation for $(Z_t^s, \mathbf{Z}_t)$ and follows a Normal distribution. To sample $\alpha_{it}^s, i = 1, 2$ we consider the augmented state vectors $\boldsymbol{\alpha}_{it}^* = (\boldsymbol{\alpha}_{it}^s, \boldsymbol{\alpha}_{it}), i = 1, 2$ and derive its distribution from the corresponding evolution equation. Given a realization of $(\alpha_{1t}^s, \alpha_{2t}^s)$, the conditional for Y_t^s is derived from the joint distribution of the pair $(Y_t^s, \mathbf{Y}_t)$, as obtained from the observation equation. Notice that the simulation needed for the spatial interpolation can be implemented after running the MCMC that uses only the data observed at the monitoring stations.

6. Results

We fitted our space-time model to the data that appear in Figure 2 with the purpose of obtaining forecasts and kriged estimates of ozone. The hyperparameters of the prior distributions were chosen under the assumption that the evolution range is larger than the observation range. This is because

Table 1. Posterior medians and 95% intervals for non-dynamic parameters.

Quantile	σ_y^2	λ_y	a_1	a_2	σ_z^2	λ_z
Median	1.27	6.63	2.47	9.83	0.89	6.99
2.5%	1.15	5.60	2.14	7.43	0.81	5.76
97.5%	1.40	7.84	2.85	13.26	0.98	8.37

harmonics are spatially more correlated than the errors in the observation equation, since the seasonality is fairly similar across locations. Also, we assumed that the prior for the observation variance is larger for temperature than for ozone, reflecting the fact that the temperature shows more variability than the square root of ozone.

The results below are based on the following prior specifications. For the observation variogram parameters, we choose inverse gamma priors: $\lambda_y \sim IG(1,5)$ and $\sigma_y^2 \sim IG(2,.01)$ for ozone, and $\lambda_z \sim IG(1,.5)$, and $\sigma_z^2 \sim IG(1,.25)$ for temperature. For the evolution equation, we fix the range and variance parameters at the following values: $\lambda_1 = 25$, $\lambda_2 = 25$, $\tau_y^2 = .02$, $\tau_1^2 = .0002$, $\tau_2^2 = .0004$, and $\tau_\gamma^2 = .0002$ for ozone, and $\tau_\eta^2 = 18$ and $\tau_z^2 = .004$, $\tau_\delta^2 = .04$ for temperature. The prior hyperparameters were selected based on variogram fits to the univariate residuals. The evolution parameters were selected after an extensive simulation study, where we required the simulated data to closely match spatial and seasonal patterns in the historical data.

We complete the specification with priors on the initial states. For ozone, we set $\boldsymbol{\theta}_0^y \sim N(\mathbf{m}_0^y, \mathbf{C}_0^y)$ with $\mathbf{m}_0^y = (2.85, -.75\,1_n', .08\,1_n', .01)'$, and $\mathbf{C}_0^y = \text{block diag}\,(1, .01\,\mathbf{I}, .01\,\mathbf{I}, .001)$, with the intercept centred around the mean square-root level, the harmonic parameters centred at values that match the bimodal diurnal cycle, and the temperature coefficient centred at a slightly positive value. We choose a fairly large variance for all components except for the harmonics. Finally, for temperature, we set $\boldsymbol{\theta}_0^z \sim N(\mathbf{m}_0^z, \mathbf{C}_0^z)$, with $\mathbf{m}_0^z = (18.5, 0, -2.7, -5, 0, 0)'$, and $\mathbf{C}_0^z = \text{diag}(10, 1, 10, 10, 10, 10)$, with the intercept centred at the September average, and the harmonics calibrated to capture the daily cycle.

Under these prior specifications, we ran the MCMC for 25000 iterations after a burn-in period of 1000 iterations, and collected all samples for posterior inference. Table 1 reports the posterior median as well as the extremes for a 95% posterior probability interval for the static parameters in the model.

Figure 4 shows retrospective medians and predictive values of ozone for 3 monitoring stations: Xalostoc, Benito Juárez and Pedregal. From September 8 until September 13, medians and 95% probability bands (solid lines) are plotted with the hourly ozone observations (dots). Also, we present forecast median with the 95% predictive probability intervals (solid lines) and the actual observed values of ozone (open circles) for the 12 hours of September 14. In the retrospective sense, the model represents the cyclical patterns and non-stationarities of the data adequately. Also, the median and predictive intervals for September 14 look in accordance with the actual data. Note that our results are reported on the original scale and not on square-root units.

Maps of median ozone levels for September 8, 1997, from 11 am to 6 pm appear in Figure 5. The maps are based on a 25×25 interpolation grid over which we apply the spatial interpolation algorithm presented in Section 6. The median was computed across MCMC samples. The resulting map seems to be consistent with the cyclical behavior of the data and theories about the dispersion of ozone in Mexico City. The pollutant builds at around 12 p.m., the peak hours are between 2 and 4 p.m. The levels seem to dissipate by 6 p.m.

Figure 6 shows the corresponding 95% range maps for ozone on September 8. The maps illustrate three features of the model. First, the overall uncertainty levels change drastically from hour to hour from the peak uncertainty at 3 p.m. when ozone levels are highest, to a low at 6 a.m. when levels are lowest. Second, as expected, uncertainty is lowest near the monitoring stations. Finally, the range contours are in general spherically shaped, due to our assumption of an isotropic covariance function.

Figure 7 shows hourly means and intervals for temperature (in degrees Celsius) at the three stations over the 168 hour period. For the first 144 hours, we show the smoothed (retrospective) means and corresponding 95% intervals, based on the data up to time 144. For the last 24 hours, we plot

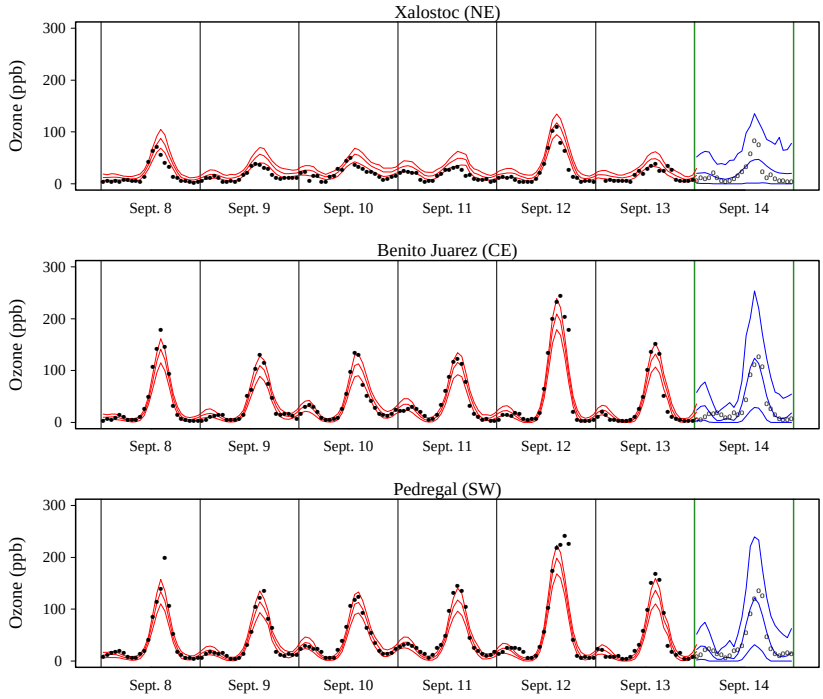

Fig. 4. Ozone data, retrospective medians and predictive values with 95% probability bands for 3 monitoring stations.

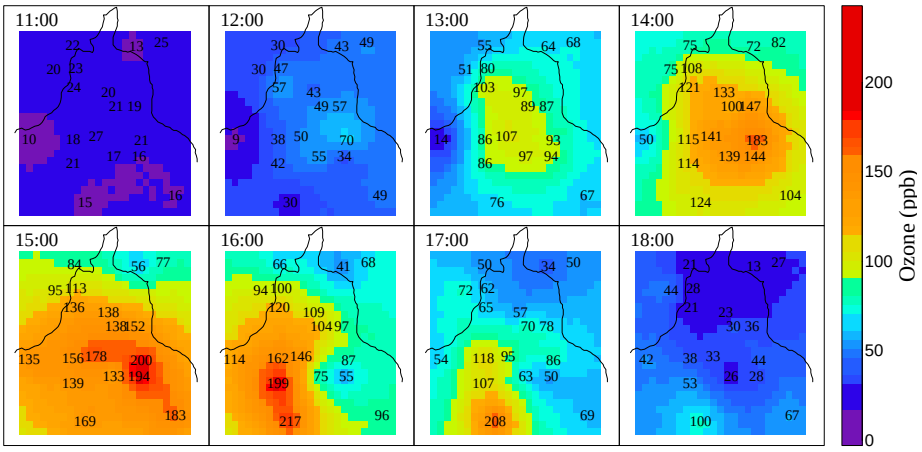

Fig. 5. Retrospective median ozone maps for September 8, 1997 based on 5000 MCMC samples, using a 1.4km^2 (25×25 pixel) mapping grid. The values written on the maps represent the observed data at that time and station.

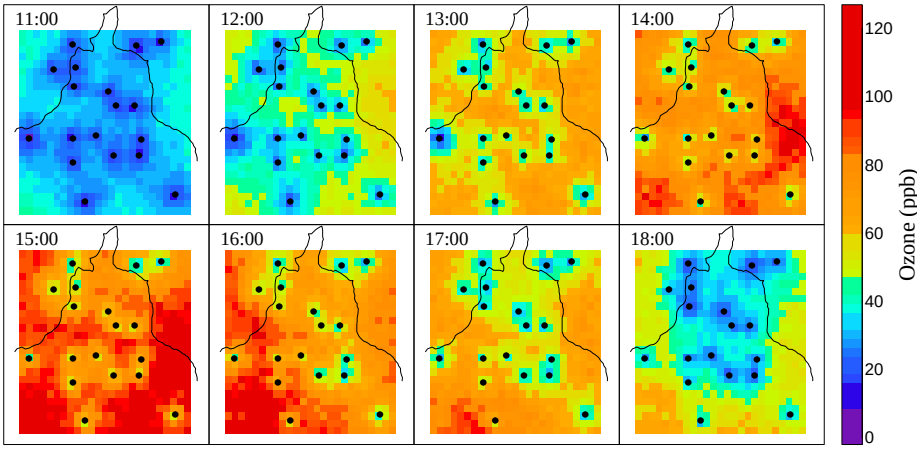

Fig. 6. Retrospective 95% range maps of ozone for September 8, 1997.

the predictive means and 95% intervals, again based on first 144 hours of data. The estimation of the altitude coefficient in the model for temperature reveals that there is significant effect of such a variable, which is more pronounced during the night. From Figure 7 and analogous plots for the rest of the stations, we observe that the model slightly underpredicts the maximum temperature of some stations, and overpredicts the minimum in other. We started our modelling by considering that the temperature was constant over all stations for each time step and found that the current model improves the fit of the ozone fields, which is our goal. We feel that further elaboration on the model for temperature will not produce results that are substantially better than the ones obtained.

In center panel of Figure 7, we show smoothed and predictive intervals at Benito Juárez, a station which does not take temperature measurements. This plot highlights a key feature of our modelling framework, namely that we are able to impute missing values and quantify uncertainty for either variable (temperature or ozone), at any specified location in the modelling domain.

6.1. Discussion

A difficult task for spatio-temporal models with a non-trivial number of sites and time steps is to assess the validity of the model. We have not been able to find in the literature a comprehensive and systematic approach for spatio-temporal model validation. This is probably due to the difficulties that entails dealing with massive amounts of data and take into account time and space. We concentrate our analysis on the study of the predictive behaviour of the model for which, given the Bayesian nature of our approach, we have a full probabilistic description.

To assess the spatial predictive capability of our model we took a "leave-one-out" approach with each one of the 19 stations. As an example, Figure 8 shows the posterior predictive densities at 2 p.m. for September 10, 1997. The labels on the x-axis correspond to the actual observed values, which were not used in the model fit. We observed that in most cases the actual observations fall in the central range of the corresponding predictive density, indicating compatibility between the model predictions and the observations. Additionally, the posterior forecast densities corresponding to 2 p.m. of September 14, 1997 are shown in Figure 9. In this case, the data of all the stations was used to compute the forecasts. Once again, the actual observed values fall within the range of the predictive densities which confirms our comments in the previous section with respect to the forecast intervals presented in Figure 4.

Kim, Shephard and Chib (1998) propose a method to assess the validity of a time series model using one step ahead predictions. We consider a modification of their method to the spatial case

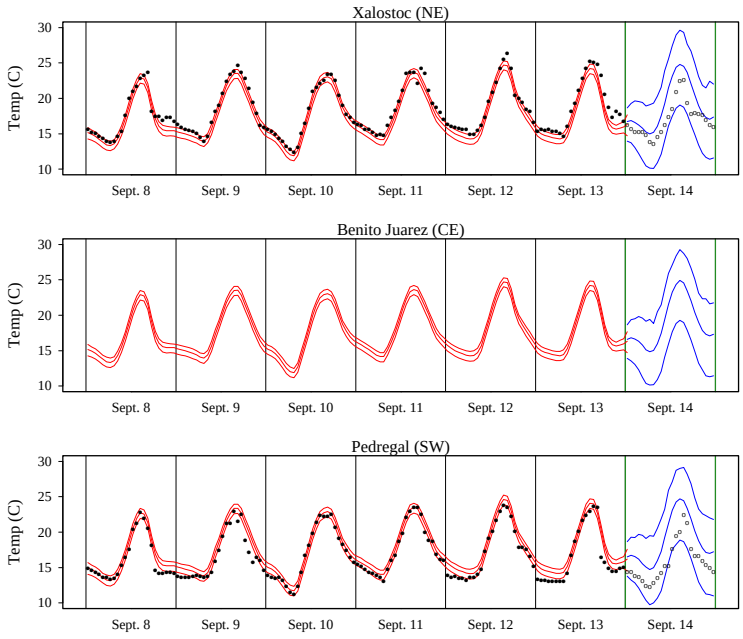

Fig. 7. Temperature data, retrospective means and predictive values with 95% probability bands for 3 monitoring stations.

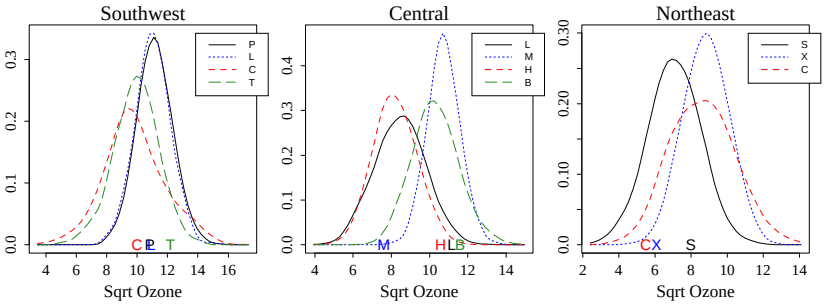

Fig. 8. Posterior predictive densities at 2 pm on September 10 for the "leave-one-out" analysis. Each panel corresponds to a different geographical region. Stations are identified with the first letter of their names. Letters on the x axis correspond to actual observations (not used in the predictions).

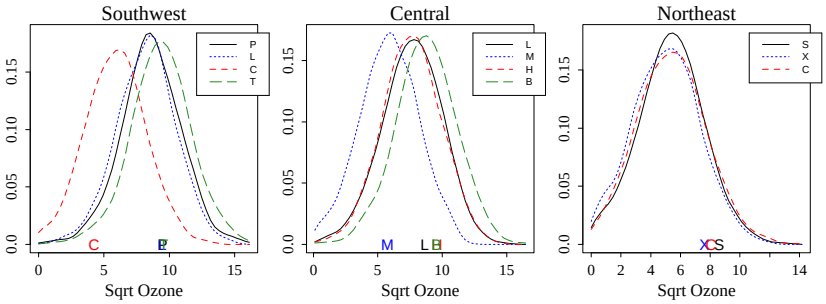

Fig. 9. Posterior forecast densities at 2pm on September 14. Each panel corresponds to a different geographical region. Stations are identified with the first letter of their names. Letters on the x axis correspond to actual future observations.

by considering the prediction produced by the model for a station that has been left out of the analysis. The idea is to compute the predictive probability that a 'leave-one-out' forecast is below its corresponding observed value. In other words, for an omitted station i, we obtained an estimate of $u_{it} = Pr[Y^*_{it} \leq Y_{it}|Y_{-i}]$, where the prediction is denoted with a star and Y_{-i} corresponds to the data without station i, by averaging across MCMC iterations. It can be shown that if the model is correctly specified, the sequence of u_{it} are independent and they are uniformly distributed on the interval $(0, 1)$. We explored the 19 possible sequences of u_t values using qq-plots and observed acceptable results in most cases. The worst behaviour was observed for Merced (MER), for which the distributional assumptions were completely unsatisfied. A careful analysis of the data from this station reveals anomalously low values of ozone in relation to the neighbouring stations. We hypothised a calibration problem, but, since similar discrepancies were not observed during the rest of the year, we suppose that there was either a transient failure in the measurement process or a very localised effect that our model is not able to capture.

A byproduct of the leave-one-out predictive analysis is the availability of draws from the posterior distributions of the parameters given by the 19 different data sets. These allowed us to perform a posterior robustness analysis with respect to the presence of each one of the stations. The marginal posterior distribution of most parameters was fairly insensitive to the removal of a station. The exceptions were the range parameters λ_y and the phase parameter a_1 for which boxplots are shown in Figure 10. We observed that removing Merced (MER) produced an important increase in λ_y. This is consistent with our previous comment regarding this station. On the other hand, removing Tlalpan (TPN) produced a significant decrease in a_1, something that is probably due to the fact that the downwind position of this station sets it slightly out of phase with the others.

6.2. Measurement error

Given the anomalies that we observed in some stations, particularly Merced, we thought of including a specific component in the model to explain measurement error. Ozone measurement devices are usually calibrated to a precision of 3 to 5 ppb for, roughly, the whole measurable range. Our model considers the data after a square root transformation, we thus notice that a first order Taylor expansion yields $\sqrt{\mu + v} \approx \sqrt{\mu} + 1/(2\sqrt{\mu})v$ where μ is the mean level and ν the measurement error. We assume that 5 ppb corresponds to one standard deviation of each of the minute by minute readings of the instrument. Since our observations correspond to one hour averages and the smallest reading amounts to 3 ppb, an upper bound of the SD of the error term in the previous equation is $5/(2 \cdot \sqrt{3 \cdot 3600}) = 0.02406$. So, under the assumption that the measurement errors are uncorrelated and that there is no bias due to miscalibration of the instruments, the measurement error is negligible.

A related issue is that of incorporating a 'nugget' effect in the model to capture small range

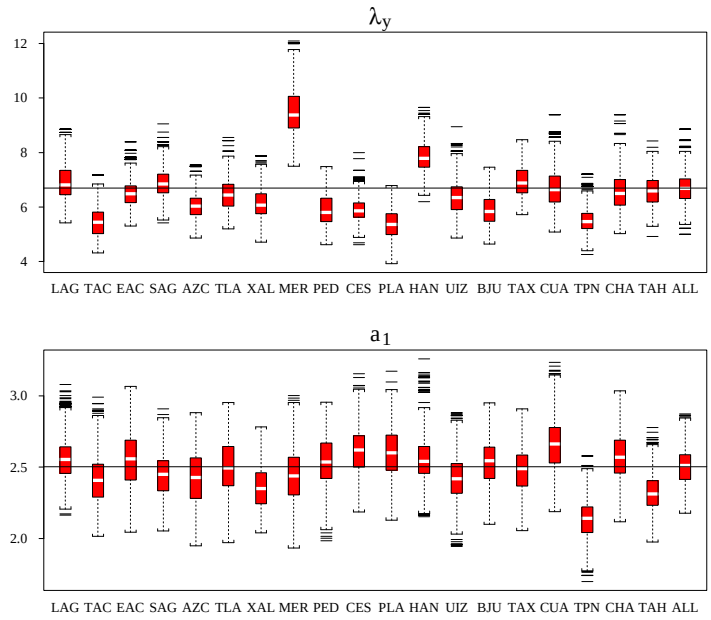

Fig. 10. Boxplots of samples from the the marginal posterior distributions of λ_y (upper panel) and a_1 (lower panel) obtained by leaving one station out. The horizontal axis denotes the omitted station. The far right boxplot represents the posterior with all stations included, and the horizontal line represents the posterior mean when all stations are included.

variabilities. This can be done by adding an extra random term or, equivalently, a third layer to the current model. We feel that our model already incorporates random effects for location, temperature and periodicity. A nugget effect will produce predictions that are less smooth, which is probably realistic, but comes at the cost of dealing with identifiability problems and adding parameters to an already complex model. Cressie and Wikle (1998) advocate the use of a three-layer model which focuses on a direct specification of the physical dynamics. This is our preferred path for elaborations over the present approach.

7. Conclusions and Extensions

In this paper, we introduced a new spatio-temporal model for hourly ozone levels in Mexico City that allows for spatial interpolation and prediction. The model is formulated within the state-space framework and includes a set of time-varying Fourier coefficients to account for the periodicity in the data. Also, the model considers uncertainty on any missing values of ozone and the covariates which in the present version, is only the temperature. We produce posterior inference with an efficient block MCMC simulation algorithm that samples the spatial covariance and error variance parameters unconditional to the state-vectors. The results obtained with the model seem reasonable for short-term forecasting and spatial interpolation, as can be assessed from the predictive analysis.

On the other hand, the actual model is only empirical and does not consider transport or chemical reactions related to ozone. Currently, we are trying to obtain second by second data, wind information and other covariables such as NO_X or VOC to incorporate into the model. We believe that this extra information combined with scientific priors based on climatology or photochemistry can produce a space-time model that not only predicts and interpolates adequately, but also has some physical interpretation.

Acknowledgements

We express our thanks to Cristina Ortuño of Departamento del Distrito Federal in Mexico City for providing the data analysed in this paper. We also thank Michael Stein, John Frederick and Jeff Gaffney for helpful comments and suggestions. G. Huerta and B. Sansó were partially funded by the CONACyT (Mexico) - CONICIT (Venezuela) international exchange grant PI-2000000834. J. Stroud was funded by NSF group infrastructure grant (GIG) 97-09696 at the University of Chicago. This work was also partially funded by the CONICIT grant G97-000592 and the CONACyT grant J34413-E.

Appendix A

A1. Marginal likelihood for range and variance parameters

The models presented in this paper have the general DLM or state space form where

$$p(\mathbf{Y}_t|\boldsymbol{\theta}_t, \sigma^2, \lambda) = N(\mathbf{Y}_t|\mathbf{F}_t'\boldsymbol{\theta}_t, \sigma^2\mathbf{V}_\lambda) \text{ and } p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}) = N(\boldsymbol{\theta}_t|G\boldsymbol{\theta}_{t-1}, \sigma^2\mathbf{W}_\lambda),$$

with $\boldsymbol{\theta}_t$ denoting the state vector and λ a non-negative parameter that defines the covariances $\mathbf{V}_\lambda$ and $\mathbf{W}_\lambda$, σ^2 a common scale factor, $p(\boldsymbol{\theta}_0|D_0) = N(\boldsymbol{\theta}_0|\mathbf{m}_0, \mathbf{C}_0)$. Denoting D_t the information available until time t and assuming that

$$p(\lambda, \sigma^2) = p(\lambda)p(\sigma^2)$$

then the joint posterior distribution for the state vectors, the variance and the range parameter is

$$p(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_T, \sigma^2, \lambda|D_T) = p(\lambda)p(\sigma^2)p(\boldsymbol{\theta}_0|D_0)\prod_{t=1}^{T} p(\mathbf{Y}_t|\boldsymbol{\theta}_t, \sigma^2, \lambda)p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, \sigma^2, \lambda).$$

Using Bayes' theorem and the Markov structure of the model, the joint posterior can be written as

$$\begin{aligned} p(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_T, \sigma^2, \lambda|D_T) &= p(\lambda)p(\sigma^2)\prod_{k=1}^{T} p(\boldsymbol{\theta}_{T-k}|\boldsymbol{\theta}_{T-k+1}, \sigma^2, \lambda, D_T)\prod_{t=1}^{T} p(\mathbf{Y}_t|\sigma^2, \lambda, D_{t-1}) \\ &= p(\boldsymbol{\theta}|\sigma^2, \lambda, D_T)p(\sigma^2|\lambda, D_T)p(\lambda|D_T) \end{aligned}$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_T)$.

Given that $p(\mathbf{Y}_t|\sigma^2, \lambda, D_{t-1}) = N(\mathbf{f}_t, \sigma^2\mathbf{Q}_t)$, where $\mathbf{f}_t$ and $\mathbf{Q}_t$ are obtained with the forward filtering equations in (West and Harrison, 1997, Chapter 16). Integrating out $\boldsymbol{\theta}$ we obtain that the posterior of (λ, σ^2) can be expressed as

$$\begin{aligned} p(\lambda, \sigma^2|D_T) &\propto p(\lambda)\, p(\sigma^2)\prod_{t=1}^{T} p(\mathbf{Y}_t|\sigma^2, \lambda, D_{t-1}) \\ &\propto p(\lambda)\, p(\sigma^2)\prod_{t=1}^{T} |\mathbf{Q}_t|^{-1/2}(1/\sigma^2)^{T/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{t=1}^{T}(\mathbf{Y}_t - \mathbf{f}_t)'\mathbf{Q}_t^{-1}(\mathbf{Y}_t - \mathbf{f}_t)\right) \quad . \end{aligned}$$

Thus, if we adopt an inverse gamma prior with parameters a_σ and b_σ for σ^2, we have that $p(\sigma^2|\lambda, D_T)$ is an inverse gamma distribution with shape parameter $T/2 + a_\sigma$ and scale parameter $\sum_{t=1}^{T}(\mathbf{Y}_t - \mathbf{f}_t)'\mathbf{Q}_t^{-1}(\mathbf{Y}_t - \mathbf{f}_t)/2 + b_\sigma$.

Additionally, if we integrate out σ^2 with this prior, then

$$p(\lambda|D_T) \propto p(\lambda)\prod_{t=1}^{T} |\mathbf{Q}_t|^{-1/2}\left(\sum_{t=1}^{T}(\mathbf{Y}_t - \mathbf{f}_t)'\mathbf{Q}_t^{-1}(\mathbf{Y}_t - \mathbf{f}_t) + b_\sigma\right)^{-T/2+a\sigma} \quad .$$

Note that both $\mathbf{Q}_t$ and $\mathbf{f}_t$ depend on λ since the forward filtering equations involve $\mathbf{V}_\lambda$ and $\mathbf{W}_\lambda$.

Finally, since

$$p(\boldsymbol{\theta}|\sigma^2, \lambda, D_T) = \prod_{k=1}^{T} p(\boldsymbol{\theta}_{T-k}|\boldsymbol{\theta}_{T-k+1}, \sigma^2, \lambda) \ ,$$

and each of the terms in the product is a Normal distribution, a sample of $\boldsymbol{\theta}$ can be obtained recursively moving backwards from $\boldsymbol{\theta}_T$ to $\boldsymbol{\theta}_1$.

References

Berliner, M., Royle, A., Wikle, C. and Milliff, R. (1999) Bayesian methods in the atmospheric sciences. In *Bayesian Statistics 6* (eds. J. Bernardo, J. Berger, A. Dawid and A. Smith), 83–100. Oxford: Oxford University Press.

Bretthorst, L. (1988) *Bayesian Spectral Analysis and Estimation.* New York: Springer Verlag.

Carroll, R., Chen, R., George, E., Li, T., Newton, H., Schmiediche, H. and Wang, N. (1997) Ozone exposure and population density in Harris County, Texas. *J. Am. Statist. Ass.*, **92**, 392–415.

Carter, C. and Kohn, R. (1994) Gibbs sampling for state space models. *Biometrika*, **81**, 541–53.

Cressie, N. and Wikle, C. (1998) Strategies for dynamic space-time statistical modeling: Discussion of The Kriged Kalman Filter by Mardia, Goodall, Redfern and Alonso. Test, 7, **2**, 257-64.

Frühwirth-Schnatter, S. (1994) Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, **15**, 183–102.

Guttorp, P., Meiring, W. and Sampson, P. (1994) A space-time analysis of ground-level ozone data. *Environmetrics*, **5**, 241–254.

— (1998) Space-time estimation of grid-cell hourly ozone levels for assessment of a deterministic model. *J. Env. Ecol. Statist.*

Kim, S., Shephard, N. and Chib, S. (1998) Stochastic volatility: likelihood inference and comparison with arch models. *Rev. Econ. Stud.*, **65**, 361–393.

Knorr-Held, L. and Rue, H. (2002) On block updating in Markov random field models for disease mapping (to appear). *Scand. J. Statist.*

Mardia, D., Goodall, C., Redfern, E. and Alonso, F. (1998) The kriged Kalman filter. *Test*, **7**, 217–285.

Milanchus, M., Rao, T. and Zurbenko, I. (1998) Evaluating the effectiveness of ozone management efforts in the presence of meteorological variability. *Journal of the Air and Waste Management Association*, **48**, 201,215.

Rao, S., Zurbenko, I., Neagu, R., Porter, P., Ku, J. and Henry, R. (1997) Space and time scales in ambient ozone data. *Bulletin of the American Meteorlogical Society*, **78**, 2153–2166.

Rue, H. (2001) Fast sampling of Gaussian Markov random fields. *J. R. Statist. Soc. B*, **63**, 325–338.

Sampson, P. and Guttorp, P. (1992) Nonparametric estimation of nonstationary spatial covariance structure. *J. Am. Statist. Ass.*, **87**, 108–119.

Sansó, B. and Guenni, L. (2000) A non-stationary multisite model for rainfall. *J. Am. Statist. Ass.*, **95**, 1064–1089.

Shaddick, G. and Wakefield, J. (2002) Modelling daily multivariate pollutant data at multiple sites (to appear). *J. R. Statist. Soc. C.*

Stroud, J., Müller, P. and Sansó, B. (2001) Dynamic models for spatio-temporal data. *J. R. Statist. Soc. B*, **63**, 673–689.

Thompson, M., Reynolds, J., Cox, L. and Guttorp, P. (1999) A review of statistical methods for the meteorological adjustment fo tropospheric ozone. *Tech. rep.*, N.R.C.S.E, University of Washington and E.P.A.

Tonellato, S. (1997) Bayesian dynamic linear models for spatial time series. *Tech. Rep. Rapporto di ricerca 5/1997*, Dipartimento di Statistica – Università Ca' Foscari di Venezia, Venice, Italy.

West, M. and Harrison, P. (1997) *Bayesian Forecasting and Dynamic Models*. New York: Springer, 2nd edn.

Wikle, C., Berliner, M. and Cressie, N. (1999) Hierarchical Bayesian space-time models. *J. Env. Ecol. Statist.*, **5**, 117–154.