A NON-HOMOGENEOUS POISSON MODEL WITH SPATIAL ANISOTROPY APPLIED TO OZONE DATA FROM MEXICO CITY

Eliane R. Rodrigues[1]

Instituto de Matemáticas

Universidad Nacional Autónoma de México

Area de la Investigación Científica

Mexico DF 04510, Mexico

Tel.: ++ (52) (55) 5622-4523

Fax: ++ (52) (55) 5616-0348

E-mail: eliane@math.unam.mx

Dani Gamerman

Departamento de Estatística

Instituto de Matemática

Universidade Federal do Rio de Janeiro, Brazil

Mario H. Tarumoto

Departamento de Estatística

Faculdade de Ciências e Tecnologia

Universidade Estadual Júlio de Mesquita Filho, Brazil

Guadalupe Tzintzun

Instituto Nacional de Ecología y Cambio Climático

Secretaría de Medio Ambiente y Recursos Naturales, Mexico

[1]Corresponding Author.

ABSTRACT

In this work we consider a non-homogenous Poisson model to study the behaviour of the number of times that a pollutant's concentration surpasses a given threshold of interest. In order to account for the possible correlation between measurements in different sites, a spatial dependence is imposed on the parameters of the Poisson intensity function. Due to the nature of the region of interest, an anisotropic model is used. Estimation of the parameters of the model is performed using the Bayesian point of view via Markov chain Monte Carlo (MCMC) algorithms. We also consider prediction of the days in which exceedances of the threshold might occur at sites where measurements cannot be taken. This is obtained by spatial interpolation using the information provided by the sites where measurements are available. The prediction procedure allows for estimation of the behaviour of the mean function of the non-homogeneous Poisson process associated with those sites. The models considered here are applied to ozone data obtained from the monitoring network of Mexico City.

Key words: Spatial models; non-homogeneous Poisson models; anisotropic models; Bayesian inference; MCMC methods; spatial interpolation.

1 Introduction

Due to the serious impact that pollutants in general have on human health, it is a very important issue to study their behaviour. Even though large cities are the ones most affected by this problem, it is worth mentioning that high levels of pollution may also occur in medium and small size cities. That depends on the car fleet, the presence of polluting industries and/or agricultural sites, among other factors (see for instance Linkens 2010).

Among the many pollutants that may affect human health we have ozone. It is well known that for ozone concentration levels above 0.11 parts per million (0.11ppm) sensitive parts of the population (e.g., elderly and newborn) may experience serious health deterioration (see, for example, Bell et al. 2004, 2005; Cifuentes et al. 2001; Dockery et al. 1992; Galizia and Kinney 1999; Gauderman et al. 2004; Gouveia and Fletcher 2000; Martins et al. 2002; WHO 2006). Therefore, it is very important to be able to study, understand and predict the behaviour of that particular pollutant.

Remark. Since we will be dealing only with ozone, from now on we are going to omit the "ppm" unit of measure.

Depending on the type of questions that are sought to be answered, different methodologies may be used to study the behaviour of ozone as well as other pollutants. Among the many studies analysing ozone pollution we have, for instance, Álvarez et al. (2005) and Larsen et al. (1990), using Markov chain models to study the problem of estimating the probability that the ozone concentration will belong to a certain interval some day into the future, given the intervals where it belongs in the present and in the more recent past; Achcar et al. (2008, 2011a), considering non-homogeneous Poisson process to study the problem of estimating the probability of having a certain number of ozone exceedances in a given time interval of interest; Achcar et al. (2011b), using volatility models to analyse the behaviour of ozone in Mexico City from the point of view of weekly average variability; Villaseñor-Alva and González-Estrada (2010), using compound Poisson models to study the behaviour of the maximum measurement in a cluster of ozone exceedances as well as the length of this cluster; Loomis et al. (1996), using time series analysis to analyse the relation of ozone exposure and mortality in Mexico City; Guardani et al. (2003), using multivariate analysis to study the behaviour of ozone in different areas of a city with an application to data from São Paulo, Brazil; Smith (1989) and Raftery (1989), using extreme value theory in order to study the behaviour of the maximum ozone measurements; and Javits (1980), using homogeneous Poisson models to estimate the probability that ozone exceedances occur a certain number of times in a year, given the rate at which they should occur in that one year period. However, in all those works only the temporal aspect of the problem was taken into account.

It is important to point out that, in some cases, the location of a monitoring station can also be of importance. As an example, we may have the case where the wind direction may intervene. Hence, it is possible that stations along the wind's path, may produce measurements with similar behaviour, even though those stations might be far apart. Therefore, it would also be interesting to include a spatial component in addition to the temporal one. In that direction we have, for instance, Huerta and Sansó (2007), in which a spatial dependence was assumed for the parameters of an extreme value distribution; Huerta et al. (2004), where a linear model with spatial dependence is assumed for ozone measurements. Those two works present applications to ozone data from the monitoring network of Mexico City, Mexico. In Szpiro et al. (2010), a spatio-temporal model is used to study the variation of air pollution data with an application to nitrogen oxides data from the city of Los Angeles, USA. Paez and Gamerman (2003), consider the space-time effects in the concentration of pollutants and focus on data from the city of Rio de Janeiro, Brazil. Another work related to the subject is that of Shaddick et al. (2013), where spatial models are used to study nitrogen dioxide concentration in an area formed by fifteen countries of the European Union. We also have Castro-Morales et al. (2013), using a space model with spatial deformation to study the behaviour of sulfur dioxide in the eastern part of the United States as well as the behaviour of the minimum temperature in Rio de Janeiro, Brazil. In addition to those works we also have Sahu et al. (2007), where space-time ozone modelling is used to assess trends in that pollutant's concentration.

In the present work we also consider a spatio-temporal model to study the behaviour of ozone in Mexico City. One of the interests here, resides in estimating the probability that an environmental threshold is surpassed by the ozone concentration a certain number of times in a time interval of interest. Even though this type of question has been analysed extensively in the past using non-homogenous Poisson models (see, for instance, Achcar et al. 2008, 2011a), the analysis then was performed considering only the temporal aspect of the model. Furthermore, estimation of the parameters was made separately for each region of the Metropolitan Area of Mexico City.

The novelty of this work is that a spatial component is included in addition to the temporal one. Hence, the parameters of the non-homogeneous Poisson model depend on the location of the monitoring sites as well as on the locations of those where measurements are not available, but where we are interested in knowing the ozone behaviour there.

Additionally, the presence of anisotropy is also allowed.

Another interest here, is to know the behaviour of ozone in specific locations representing sites of interest such as, a hospital or an area where it is under consideration to be the location of a new hospital, and where measurements are not taken at present. In relation to that, the interest resides in estimating the number of ozone exceedances of a given threshold as well as the days in which they might occur.

Therefore, in order to analyse in full the problem considered here, in addition to non-homogeneous Poisson models, we also assume a Gaussian process for the parameters of the Poisson intensity function to allow for a correlation between any given pair of sites. Since the prevailing wind direction in Mexico City is from northeast to southwest and also from north to south, we also consider an anisotropic model for those correlations. By doing so, we may have an idea of how measurements from one station may provide information regarding measurements taken in a different station. That information may be useful when we need to estimate the behaviour of ozone (or any other type of pollutant) in sites where no monitoring stations are present. The Gaussian process representation allows for information from sites where measurements are taken, to be carried over to sites where it is not possible to have them. Thus, another useful information that can be obtained here, concerns the behaviour of the mean number of exceedances in a site where we may not observe the measurements directly.

This work is organised as follows. In Section 2 the description of the mathematical setting is presented. The Bayesian formulation of the model is given in Section 3. In Section 4 we apply the model, described in previous sections, to ozone data obtained from the monitoring network of Mexico City. Finally, in Section 5, some comments about the results are given. In an Appendix, before the list of references, we present the posterior density functions considered in the MCMC algorithm used to estimate the parameters of the model.

2 Description of the mathematical model

The mathematical model used here may be described as follows. Assume that ozone measurements are obtained from $N_O \geq 1$ monitoring stations. Also, assume that there are $N_U \geq 0$ locations at which we cannot measure the ozone concentration directly (i.e., no monitoring stations are placed there). These locations are the ones where we would like to estimate the probability that the ozone concentration surpasses a given threshold a certain number of times in a given time interval as well as the mean number of ozone exceedances of the threshold. In addition to that, we would also like to predict the days in which those exceedances might occur.

Let $K_i \geq 0$ indicate the number of times that the threshold of interest is surpassed during the time interval $[0, T_i)$, $T_i > 0$, in the site i of interest. Denote by $\mathbf{D}_i = \{d_{1,i}, d_{2,i}, \ldots, d_{K_i,i}\}$ the set of those exceedance days.

Remarks. 1. Note that when we are considering sites where measurements are actually taken, the values of K_i and the sets $\mathbf{D}_i$ are known. However, when we are dealing with sites where we are not able to take measurements, K_i and $\mathbf{D}_i$ are unknown and need to be predicted.

2. Even though we are allowing the value of T_i be different for different values of i, in the application made here, we are taking them all equal. Hence, we have $T_i = T$, $i = 1, 2, \ldots, N_O + N_U$

Therefore, in the present case we have that the set of observed data is represented by $\mathbf{D}^{(O)} = \{\mathbf{D}_1^{(O)}, \mathbf{D}_2^{(O)}, \ldots, \mathbf{D}_{N_O}^{(O)}\}$. We are going to use the notation $\mathbf{D}^{(U)} = \{\mathbf{D}_1^{(U)}, \mathbf{D}_2^{(U)}, \ldots, \mathbf{D}_{N_U}^{(U)}\}$ to indicate the set of exceedance days at locations where measurements may not be taken. This set of days, need to be predicted.

In order to estimate the probabilities of interest, we assume that the number of exceedances in a given time interval $[0, t)$, $t \geq 0$, follows a non-homogeneous Poisson process with rate and mean functions $\lambda(t) > 0$ and $m(t) = \int_0^t \lambda(s)\, ds$, $t \geq 0$, respectively. We allow for different non-homogeneous processes at different sites. Therefore, following in that direction, let $M_t^{(i)} \geq 0$, indicate the number of times that the ozone concentration surpasses the threshold of interest at site i during the time interval $[0, t)$, $t \geq 0$. Hence, we assume that $M^{(i)} = \{M_t^{(i)} : t \geq 0\}$ is a non-homogeneous Poisson process with rate and mean functions indicated by $\lambda^{(i)}(t)$ and $m^{(i)}(t)$, respectively. Therefore, we have that

$$P(M^{(i)}(t+s) - M^{(i)}(t) = k) = \frac{[m^{(i)}(t+s) - m^{(i)}(t)]^k}{k!} \exp\left(-[m^{(i)}(t+s) - m^{(i)}(t)]\right), \quad (1)$$

with $k = 0, 1, \ldots$ and $t, s \geq 0$.

We assume that $\lambda^{(i)}(t)$, $t \geq 0$, has a Weibull form with parameters α_i and β_i, i.e., $\lambda^{(i)}(t) = (\alpha_i/\beta_i)(t/\beta_i)^{\alpha_i - 1}$, giving $m^{(i)}(t) = (t/\beta_i)^{\alpha_i}$, $t \geq 0$, $\alpha_i, \beta_i \in (0, \infty)$. Hence, even though the forms of the rate functions are the same, we allow the parameters to depend on the particular sites. Thus, we indicate by $(\alpha_i^{(O)}, \beta_i^{(O)})$, $i = 1, 2, \ldots, N_O$, the parameters of the rate functions that are related to the sites where measurements can be directly taken, and we indicate by $(\alpha_j^{(U)}, \beta_j^{(U)})$, $j = 1, 2, \ldots, N_U$, the corresponding parameters related to the sites where measurements cannot be taken directly.

Remark. This form of the rate function was used before to study the problem in the temporal framework only (Achcar et al. 2008, 2011a) and is commonly used in survival analysis and reliability theory. Since in the present case we are also interested on the number of ozone exceedances in a time interval of interest, and since in the temporal framework, when studying the behaviour of ozone for any given region of Mexico City, the rate function considered here proved to be suitable, we have decided to use it in the present situation as well.

Therefore, the vector of parameters related to the Poisson rate functions are $\boldsymbol{\theta}^{(O)}$ and $\boldsymbol{\theta}^{(U)}$, where $\boldsymbol{\theta}^{(O)} = ((\alpha_i^{(O)}, \beta_i^{(O)});\, i = 1, 2, \ldots, N_O)$ and $\boldsymbol{\theta}^{(U)} = ((\alpha_j^{(U)}, \beta_j^{(U)});\, j = 1, 2, \ldots, N_U)$, are the vectors of parameters related to the observable and unobservable sites, respectively. Estimation of the parameters will be made using the Bayesian point of view.

Remark. Even though there is a difference in the notation, when it comes to the application of the results, the indices "O" and "U" will be dropped occasionally. The cases of known and unknown data will be differentiated by the index associated to the site.

3 A Bayesian formulation of the problem

In this section, we present the Bayesian formulation of the model considered here. The distribution of interest is the posterior distribution of the vector of parameters. Note that for a parameter $\boldsymbol{\theta}$ its posterior distribution is given by $P(\boldsymbol{\theta} \,|\, \mathbf{D}) \propto L(\mathbf{D} \,|\, \boldsymbol{\theta})\, P(\boldsymbol{\theta})$, where $L(\mathbf{D} \,|\, \boldsymbol{\theta})$ is the likelihood function of the model and $P(\boldsymbol{\theta})$ is the prior distribution of $\boldsymbol{\theta}$.

Here, we have that,

$$L(\mathbf{D}^{(O)} \,|\, \boldsymbol{\theta}^{(O)}, \boldsymbol{\theta}^{(U)}) = L(\mathbf{D}^{(O)} \,|\, \boldsymbol{\theta}^{(O)}) = \prod_{i=1}^{N_O} L(\mathbf{D}_i^{(O)} \,|\, \boldsymbol{\theta}^{(O)}) = \prod_{i=1}^{N_O} L(\mathbf{D}_i^{(O)} \,|\, \alpha_i^{(O)}, \beta_i^{(O)}),$$

where, since we have by hypothesis, a non-homogeneous Poisson model, it follows that (Cox and Lewis 1966; and Lawless 1982)

$$L(\mathbf{D}_i^{(O)} \,|\, \alpha_i^{(O)}, \beta_i^{(O)}) = \left(\prod_{k=1}^{K_i} \lambda^{(i)}(d_{k,i} \,|\, \boldsymbol{\theta}^{(O)}) \right) \exp\left[-m^{(i)}(T_i \,|\, \boldsymbol{\theta}^{(O)}) \right], \tag{2}$$

where we use $\lambda^{(i)}(t \,|\, \boldsymbol{\theta})$ and $m^{(i)}(t \,|\, \boldsymbol{\theta})$ to make it explicit the dependence of the rate and mean functions on the corresponding parameter $\boldsymbol{\theta}$. Given the Weibull assumption for the rate function of the Poisson process, we have that,

$$L(\mathbf{D}^{(O)} \,|\, \boldsymbol{\theta}^{(O)}, \boldsymbol{\theta}^{(U)}) = \prod_{i=1}^{N_O} \left[\left(\frac{\alpha_i^{(O)}}{\beta_i^{(O)}} \right)^{K_i} e^{-\left(T_i/\beta_i^{(O)}\right)^{\alpha_i^{(O)}}} \left(\prod_{k=1}^{K_i} \left[\frac{d_{k,i}}{\beta_i^{(O)}} \right]^{\alpha_i^{(O)}-1} \right) \right]. \tag{3}$$

Independent Gaussian processes are assumed to rule the spatial variation of the parameters $\boldsymbol{\alpha} = (\alpha_i^{(O)}, \alpha_j^{(U)};\, i = 1, 2, \ldots, N_O,\, j = 1, 2, \ldots, N_U)$ and $\boldsymbol{\beta} = (\beta_i^{(O)}, \beta_j^{(U)};\, i = 1, 2, \ldots, N_O,\, j = 1, 2, \ldots, N_U)$ in their logarithmic scale. We assume prior independence between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. We also assume that $\log \boldsymbol{\alpha} = (\log \alpha_i^{(O)}, \log \alpha_j^{(U)};\, i = 1, 2, \ldots, N_O,\, j = 1, 2, \ldots, N_U)$ and $\log \boldsymbol{\beta} = (\log \beta_i^{(O)}, \log \beta_j^{(U)};\, i = 1, 2, \ldots, N_O,\, j = 1, 2, \ldots, N_U)$ have as prior distributions multivariate normal distributions with mean vectors

$$\boldsymbol{\mu}^{\alpha} = (\mu^{\alpha_i^{(O)}}, \mu^{\alpha_j^{(U)}};\, i = 1, 2, \ldots, N_O,\, j = 1, 2, \ldots, N_U)$$

and

$$\boldsymbol{\mu}^{\beta} = (\mu^{\beta_i^{(O)}}, \mu^{\beta_j^{(U)}};\, i = 1, 2, \ldots, N_O,\, j = 1, 2, \ldots, N_U),$$

and with variance-covariance matrices Σ^α and Σ^β, respectively. (The forms of Σ^α and Σ^β will be specified later.) Let $\Sigma^\alpha = \left(\nu_{ij}^\alpha\right)_{i,j=1,2,\ldots,N_O+N_U}$ and $\Sigma^\beta = \left(\nu_{ij}^\beta\right)_{i,j=1,2,\ldots,N_O+N_U}$ indicate the composition of those variance-covariance matrices. Hence,

$$\nu_{ij}^\alpha = \mathrm{Cov}(\log(\alpha_i), \log(\alpha_j)), \quad i, j = 1, 2, \ldots, N_O + N_U.$$

Similar notation is used for ν_{ij}^β, $i, j = 1, 2, \ldots, N_O + N_U$.

Let $\|s_i - s_j\|$ denote the Euclidean distance between sites i and j. Here, s_i indicates the UTM coordinates of site i. Define

$$R = \begin{pmatrix} \cos\psi_a & -\sin\psi_a \\ \sin\psi_a & \cos\psi_a \end{pmatrix} \quad \text{and} \quad X = \begin{pmatrix} 1 & 0 \\ 0 & \psi_r^{-1} \end{pmatrix},$$

where ψ_a is the anisotropy angle and $\psi_r > 1$ is the anisotropy ratio (Cressie 1991; Diggle and Ribeiro 2007; Schmidt and Rodríguez 2010). Consider the following matrix $A = R \times X$ and define $d(s_k) = A \times s_k$, with s_k the UTM coordinates of the site k, i.e., $d(s_k)$ is the vector of coordinates of the position of site k in the new space obtained by the rotation matrix R and the deformation given by X.

Remark. The use of a rotation and deformation of the UTM coordinates, is in order to allow for the possible influence of the wind directions and also to account for Mexico City's geographic configuration.

Consider the following forms for the variance-covariance between two parameters $\log\alpha$ related to sites i and j, $i, j = 1, 2, \ldots, N_O + N_U$,

$$\nu_{ij}^\alpha = \sigma_{\alpha_i}\,\sigma_{\alpha_j}\,\exp\left(-\phi_\alpha\,\|d(s_i) - d(s_j)\|\right), \tag{4}$$

and between two parameters $\log\beta$

$$\nu_{ij}^\beta = \sigma_{\beta_i}\,\sigma_{\beta_j}\,\exp\left(-\phi_\beta\,\|d(s_i) - d(s_j)\|\right), \tag{5}$$

where ϕ_α and ϕ_β are parameters that need to be estimated, and σ_{α_k} and σ_{β_k} are the standard deviations of $\log \alpha_k$ and $\log \beta_k$, respectively, $k = 1, 2, \ldots, N_O + N_U$. Let $\boldsymbol{\sigma}_\alpha = (\sigma_{\alpha_i};\, i = 1, 2, \ldots, N_O + N_U)$ and $\boldsymbol{\sigma}_\beta = (\sigma_{\beta_i};\, i = 1, 2, \ldots, N_O + N_U)$ indicate the vector of those standard deviations.

Remark. Note that, if the interest is in estimating the correlation between each possible pair of n sites, we could consider each of the $\binom{n}{2}$ pairs as parameters to be estimated. However, as the value of n increases, the number of parameters to be estimated also increases, and rapidly. Hence, in an attempt to reduce that number of parameters, and at the same time preserve the researcher's belief in how the correlation coefficients might behave, some models have been proposed. Among them, are the so-called isotropic and anisotropic models. In an isotropic model, only the distance between two sites are consider to contribute to the correlation coefficient between these two sites. In an anisotropic model, besides the distance between those sites, the contribution of the angle between them, with respect to some origin point, is also taken into account. Due to Mexico City's geographic configuration, the anisotropic version of the models seems to be a suitable one.

Therefore, the complete vector of parameters to be estimated in the present formulation, consists of the parameters related to the spatial variation, parameters related to the Poisson rate function at observed and unobserved sites, and also unknown measurements at unobservable sites. These unknown quantities can be all stacked together in the following vector of parameters

$$\overline{\boldsymbol{\theta}} = (\boldsymbol{\theta}^{(O)}, \boldsymbol{\theta}^{(U)}, \boldsymbol{\mu}^{\alpha}, \boldsymbol{\mu}^{\beta}, \boldsymbol{\sigma}_\alpha, \boldsymbol{\sigma}_\beta, \phi_\alpha, \phi_\beta, \psi_a, \psi_r, \mathbf{D}^{(U)}).$$

Consider the notation, $\boldsymbol{\alpha}^{(O)} = (\alpha_i^{(O)}, i = 1, 2, \ldots, N_O)$, $\boldsymbol{\alpha}^{(U)} = (\alpha_i^{(U)}, i = 1, 2, \ldots, N_U)$,

$\boldsymbol{\mu}^{\alpha^{(O)}} = (\mu^{\alpha_i^{(O)}}, i = 1, 2, \ldots, N_O)$, $\boldsymbol{\mu}^{\alpha^{(U)}} = (\mu^{\alpha_i^{(U)}}, i = 1, 2, \ldots, N_U)$, and

$$
\begin{aligned}
\Sigma^{\alpha^{(O)},\alpha^{(O)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha) &= \left(\mathrm{Cov}(\alpha_i^{(O)}, \alpha_j^{(O)})\right)_{i,j=1,2,\ldots,N_O} \\
\Sigma^{\alpha^{(U)},\alpha^{(U)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha) &= \left(\mathrm{Cov}(\alpha_i^{(U)}, \alpha_j^{(U)})\right)_{i,j=1,2,\ldots,N_U} \\
\Sigma^{\alpha^{(O)},\alpha^{(U)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha) &= \left(\mathrm{Cov}(\alpha_i^{(O)}, \alpha_j^{(U)})\right)_{i=1,2,\ldots,N_O; j=1,2,\ldots,N_U} \\
\Sigma^{\alpha^{(U)},\alpha^{(O)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha) &= \left(\mathrm{Cov}(\alpha_i^{(U)}, \alpha_j^{(O)})\right)_{i=1,2,\ldots,N_U; j=1,2,\ldots,N_O}.
\end{aligned}
$$

Therefore, the posterior distribution of the vector of parameters may be written as

$$
\begin{aligned}
P(\overline{\boldsymbol{\theta}} \,|\, \mathbf{D}^{(O)}) \;\propto\; & L(\mathbf{D}^{(O)} \,|\, \boldsymbol{\theta}^{(O)})\, P(\boldsymbol{\alpha}^{(O)} \,|\, \boldsymbol{\mu}^{\alpha^{(O)}}, \phi_\alpha, \psi_a, \psi_r, \boldsymbol{\sigma}_\alpha) \\
& P(\boldsymbol{\alpha}^{(U)} \,|\, \boldsymbol{\alpha}^{(O)}, \boldsymbol{\mu}^{\alpha}, \phi_\alpha, \psi_a, \psi_r, \boldsymbol{\sigma}_\alpha) \\
& P(\boldsymbol{\beta}^{(O)} \,|\, \boldsymbol{\mu}^{\beta^{(O)}}, \phi_\beta, \psi_a, \psi_r, \boldsymbol{\sigma}_\beta)\, P(\boldsymbol{\beta}^{(U)} \,|\, \boldsymbol{\beta}^{(O)}, \boldsymbol{\mu}^{\beta}, \phi_\beta, \psi_a, \psi_r, \boldsymbol{\sigma}_\beta) \qquad (6) \\
& P(\boldsymbol{\mu}^{\alpha})\, P(\boldsymbol{\mu}^{\beta})\, P(\phi_\alpha)\, P(\phi_\beta)\, P(\psi_a)\, P(\psi_r)\, P(\boldsymbol{\sigma}_\alpha)\, P(\boldsymbol{\sigma}_\beta) \\
& P(\mathbf{D}^{(U)} \,|\boldsymbol{\alpha}^{(U)}, \boldsymbol{\beta}^{(U)}),
\end{aligned}
$$

where $L(\mathbf{D}^{(O)} \,|\, \boldsymbol{\theta}^{(O)})$ is given by (3),

$$
\begin{aligned}
P(\boldsymbol{\alpha}^{(O)} \,|\, \boldsymbol{\mu}^{\alpha^{(O)}}, \phi_\alpha, \psi_a, \psi_r, \boldsymbol{\sigma}_\alpha) \propto & \frac{\left|\Sigma^{\alpha^{(O)},\alpha^{(O)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)\right|^{-1/2}}{|\prod_{i=1}^{N_O} \alpha_i^{(O)}|} \qquad (7) \\
\exp\left(-\frac{\left[\log(\boldsymbol{\alpha}^{(O)}) - \boldsymbol{\mu}^{\alpha^{(O)}}\right]^t \left(\Sigma^{\alpha^{(O)},\alpha^{(O)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)\right)^{-1} \left[\log(\boldsymbol{\alpha}^{(O)}) - \boldsymbol{\mu}^{\alpha^{(O)}}\right]}{2}\right)&,
\end{aligned}
$$

and

$$P(\boldsymbol{\beta}^{(O)} \mid \boldsymbol{\mu}^{\beta^{(O)}}, \phi_\beta, \psi_a, \psi_r, \boldsymbol{\sigma}_\beta) \propto \frac{\left|\Sigma^{\beta^{(O)},\beta^{(O)}}(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)\right|^{-1/2}}{|\prod_{i=1}^{N_O} \beta_i^{(O)}|} \tag{8}$$
$$\exp\left(-\frac{\left[\log(\boldsymbol{\beta}^{(O)}) - \boldsymbol{\mu}^{\beta^{(O)}}\right]^t \left(\Sigma^{\beta^{(O)},\beta^{(O)}}(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)\right)^{-1} \left[\log(\boldsymbol{\beta}^{(O)}) - \boldsymbol{\mu}^{\beta^{(O)}}\right]}{2}\right).$$

The parameters $\boldsymbol{\alpha}^{(U)}$ and $\boldsymbol{\beta}^{(U)}$ corresponding to sites where data are not available, will have distribution given by the respective conditional distributions given $\boldsymbol{\alpha}^{(O)}$ and $\boldsymbol{\beta}^{(O)}$. In that case we have that $\log(\boldsymbol{\alpha}^{(U)})$ will have a multivariate normal distribution with mean vector $\overline{\mu}^{\boldsymbol{\alpha}^{(U)}}$ and variance-covariance matrix $\overline{\Sigma}^{\alpha^{(U)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)$ given by

$$\overline{\mu}^{\boldsymbol{\alpha}^{(U)}} = \boldsymbol{\mu}^{\boldsymbol{\alpha}^{(U)}} + \Sigma^{\alpha^{(U)},\alpha^{(O)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha) \left[\Sigma^{\alpha^{(O)},\alpha^{(O)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)\right]^{-1} \left(\log \boldsymbol{\alpha}^{(O)} - \boldsymbol{\mu}^{\alpha^{(O)}}\right) \tag{9}$$

and

$$\overline{\Sigma}^{\alpha^{(U)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha^{(U)}) = \Sigma^{\alpha^{(U)},\alpha^{(U)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha) \tag{10}$$
$$-\Sigma^{\alpha^{(U)},\alpha^{(O)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha) \left[\Sigma^{\alpha^{(O)},\alpha^{(O)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)\right]^{-1} \Sigma^{\alpha^{(O)},\alpha^{(U)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)$$

The same methodology and similar notation are used in the case of $\boldsymbol{\beta}^{(U)}$, but now using $\boldsymbol{\beta}^{(U)}$ instead of $\boldsymbol{\alpha}^{(U)}$.

We are going to assume that the components of $\boldsymbol{\mu}^\alpha$ and $\boldsymbol{\mu}^\beta$ are independent, and that the prior distributions $P(\boldsymbol{\mu}^\alpha)$ and $P(\boldsymbol{\mu}^\beta)$ are the product of appropriate normal distributions that will be specified when applying the model to the data.

Regarding the parameters ϕ_α and ϕ_β, we take (see Schmidt and Rodríguez 2010) $P(\phi)$ an inverse Gamma function $\mathrm{IG}(a, b)$, $a > 2$, i.e.,

$$P(\phi) = \frac{b^a\, e^{-b/\phi}}{\Gamma(a)\, \phi^{a+1}}, \quad \phi > 0, \tag{11}$$

with possibly different hyperparameters when considering ϕ_α and ϕ_β. We assume that $P(\psi_a)$ is a uniform distribution on $(0, \pi)$ and that $P(\psi_r)$ is a Pareto distribution with hyperparameters (c, d), i.e.,

$$P(\psi_r) = c \, \frac{d^c}{\psi_r^{c+1}}, \quad \psi_r > d. \tag{12}$$

Remarks. 1. In the present study we are going to consider σ_{α_i} and σ_{β_i}, $i = 1, 2, \ldots, N_O + N_U$, known and they will be kept fixed. Hence, the prior distributions $P(\boldsymbol{\sigma}_\alpha)$ and $P(\boldsymbol{\sigma}_\beta)$ will assume value one for the given fixed vectors of values and they will be zero otherwise.

2. Since $\psi_r > 1$, we take $d = 1$. Note that the variance of the Pareto distribution is finite for $c > 2$. That value will be specified later. In the case of the inverse Gamma distribution, in order to have an infinite variance we may fix $a > 2$. Additionally, in order to have the prior mean of ϕ such that the practical range (correlation $= 0.05$) is reached at half of the maximum distance between sites, we consider (see Schmidt and Rodríguez 2010)

$$-\log(0.05) = \frac{\phi \max \|d(s) - d(s')\|}{2}, \tag{13}$$

where the maximum is over all pair of sites s and s'.

3. We would like to call attention to the fact that the hypothesis reflected in (13) is specific to the mean of the prior distribution of ϕ. Also, note that the hyperparameters of this prior distribution are such that we have an infinite variance. Therefore, even though (13) might represent a restriction on the mean, the infinite variance allows for a wide range of possibilities for the values that ϕ might assume.

Looking at the expression for the posterior distribution $P(\overline{\boldsymbol{\theta}} \mid \mathbf{D}^{(O)})$, it is possible to notice that it is analytically intractable and approximations must be sought. Hence, inference will be made using a sample obtained from the posterior distribution of the parameters using a Gibbs sampling algorithm (Gamerman and Lopes 2006; Gelfand and Smith 1990; Robert and Casella 1999; Smith and Roberts 1993).

The hyperparameters of the prior distributions that were not specified here, will be when we apply the model to the data.

4 Application to ozone data from the monitoring network of Mexico City

Here, we apply the model, described in previous sections, to the ozone data from the monitoring network of the Metropolitan Area of Mexico City. The Metropolitan Area is divided into five regions, Northeast (NE), Northwest (NW), Centre (CE), Southeast (SE) and Southwest (SW), and monitoring stations are placed throughout the city. However, not all of them measure ozone. Additionally, in 2011 some of the active stations that measured ozone were disabled and others were added to the network. Hence, in order to illustrate the application of the model presented here, we are going to consider only a subset of stations that measure ozone and that are active nowadays.

The stations considered in the present study are: San Agustín (SAG) and Chapingo (CHA) in the Northeast region of the city; ENEP-Acatlán (EAC) and Tlalneplantla (TLA) in the Northwest region; Merced (MER) in the central region, UAM-Iztapalapa (UIZ) and Tláhuac (TAH) in the Southeast region; and finally Pedregal (PED), Cuajimalpa (CUA) and Coyoacán (COY) in the Southwest region of the city. Those stations were considered as stations where the measurements are available, i.e., they will represent the sites where measurements are observable. We have also included the station Plateros (PLA) located in the Southwest region, but for the sake of our prediction exercise, that station is going to be considered a site where measurements are not available. The reason for choosing a site in that region in particular, is that this region is where, in general, higher levels of ozone occur. We will show results for a single site with unavailable data. Simultaneous prediction for a collection of sites can be obtained in the same manner. The difference being that we would have to sample α and β, in the case of unobservable sites, using a multivariate normal distribution instead of a univariate one. The location of the stations considered here, are shown in Figure 1.

Figure 1 about here.

The data considered in the present analysis are the daily maximum ozone measurements obtained from 01 January 2005 to 31 December 2009. Measurements are taken every minute at the monitoring station and the averaged hourly results are reported. The daily maximum ozone measurement at a given station is the maximum of the 24-h hourly averaged measurements. The averaged maximum daily ozone levels in stations SAG, CHA, EAC, TLA, MER, UIZ, TAH, PED, CUA and COY, during the observational period considered here were, respectively, 0.072, 0.075, 0.082, 0.077, 0.084, 0.09, 0.083, 0.099, 0.089, and 0.096, with respective standard deviations of 0.026, 0.024, 0.031, 0.029, 0.032, 0.031, 0.029, 0.037, 0.033, and 0.034.

Figures 2 and 3 present the plots of the daily maximum ozone level for all stations where measurements are considered to be observable as well as the case of station PLA.

Figure 2 about here.

Figure 3 about here.

It is possible to see from Figures 2 and 3, that even though in all stations there are days in which low concentration levels are detected, in most of the cases we have high ones (above 0.11ppm), specially when we consider stations located in region SW.

The threshold considered in the present work is 0.11, since that is the value specified by the Mexican ozone standard (NOM 2002). During the observational period considered here, this threshold was surpassed in 129, 116, 319, 223, 365, 492, 346, 708, 457 and 621 days in stations SAG, CHA, EAC, TLA, MER, UIZ, TAH, PED, CUA and COY, respectively. Those values correspond to the respective values of K for those stations. We also have that $T_i = T = 1826$, $i = 1, 2, \ldots, N_O + N_U$. Additionally, $N_O = 10$ and, for the purpose of illustration, we are taking $N_U = 1$, corresponding to the station PLA.

The prior distributions as well as the results obtained are described in the next three subsections. In the first subsection, we present the prior distributions for each parameter as well as their hyperparameters. In the second subsection, we present the values of the estimated parameters of the rate functions of the non-homogeneous Poisson processes for observable and unobservable sites as well as the estimated parameters present in the spatial part of the model. In the third subsection, we present the predicted values of the number of exceedances that might occur at site PLA as well as the estimated days when exceedances may have occurred.

4.1 Prior distributions

The normal prior distribution of each component of the vector $\boldsymbol{\mu}^{\alpha}$ corresponding to the stations where data are available, has mean -0.5 and variance 0.7. In the case of station

PLA, the normal prior distribution of the component of $\boldsymbol{\mu}^{\alpha}$ corresponding to it, has mean -3 and variance 0.5. In the case of the vector $\boldsymbol{\mu}^{\beta}$, the normal prior distribution of each component corresponding to stations where data are available has mean 1.6 and variance one. In the case of station PLA, the normal prior distribution has mean -1.3 and variance 0.2.

Remarks. 1. The hyperparameters of the prior distributions of $\boldsymbol{\mu}^{\alpha}$ and $\boldsymbol{\mu}^{\beta}$ were obtained using information provided by previous studies where only the temporal part was taken into account (see Achcar et al. 2008, 2011a). In those works, the daily maximum ozone measurements for each region of Mexico City were used. Also, in those works, even thought the period of time considered might intercept part of the period of time considered here, the data used before correspond to regional measurements obtained from stations in a given region, and not data corresponding to each individual station in separate.

2. Note that we need to have the parameters α assuming values mostly in $(0, 1)$ in order to be compatible with the behaviour expected by the data, but we must also allow for values larger than one.

The hyperparameters of the Pareto prior distribution of the parameter ψ_r are $c = 3$ and $d = 1$. The prior distribution of ψ_a is as specified in Section 3, i.e., it is a uniform distribution on $(0, \pi)$.

When taking into account the parameters ϕ_{α} and ϕ_{β}, they will have as their prior distributions, inverse Gamma distributions IG(2.5, b_{α}) and IG(3, b_{β}), respectively, where the hyperparameters b_{α} and b_{β} are the solutions of the optimisation problem (13) corresponding to each particular ϕ. The standard deviations, $\boldsymbol{\sigma}_{\alpha}$ and $\boldsymbol{\sigma}_{\beta}$ are given as follows. In the case of components in $\log \boldsymbol{\alpha}$, we have $\sigma_{\alpha_i} = 0.73$ for all values of i. The case of the components of $\log \boldsymbol{\beta}$, the behaviour is more heterogeneous. Hence, $\boldsymbol{\sigma}_\beta$ has value 0.33 in the components corresponding to the stations EAC, SAG, CHA, MER, CUA, PED, COY and PLA. In the case of the components corresponding the the stations TLA, TAH and UIZ we have σ_β equal to 0.22.

Remark. The difference in the values of the hyperparameters of the prior distributions for different sites, is due to the fact that the information used to consider them was based on the specific information of the region where the station is located. The selection of those values was also based on the mean value and standard deviation of the parameters related to the model where the maximum ozone measurements for each region were used.

In the case of the vector $\boldsymbol{\alpha}^{(O)}$ its prior distribution is a log-normal distribution with parameters $\boldsymbol{\mu}^{\alpha^{(O)}}$ and $\Sigma^{\alpha^{(O)},\alpha^{(O)}}$. Similar procedure is applied to the vector $\boldsymbol{\beta}^{(O)}$, but now using $\boldsymbol{\mu}^{\beta^{(O)}}$ and $\Sigma^{\beta^{(O)},\beta^{(O)}}$.

4.2 Estimation results

Estimation of the parameters is performed using a sample of size 5000 obtained with a MCMC algorithm (described in the Appendix) consisting of five chains, after a burn-in period of 30000 steps and with a sampling gap of 20.

The estimated mean values of the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as well as their respective standard deviations and 95% credible intervals are given in Table 1 for each station considered.

Table 1 about here.

Looking at Table 1 we may see that when considering the parameter $\boldsymbol{\alpha}$, stations in the same geographic region have similar behaviour. Furthermore, the α corresponding to stations in different regions may also present similar behaviour. It is possible to see that, with the exception of station in regions NE and CE and also station PLA in region SW, the value of the parameter α is in the interval (0.8, 0.84). In the case of stations in regions NE and CE, the value is between 0.69 and 0.77. These values imply that the rate at which ozone exceedances occur is decreasing in time. Hence, occurrence of ozone exceedances of the threshold 0.11 is becoming less frequent.

Even though the values of α corresponding to stations in the same region are similar, the value of β may be very different. The value of β related to station TLA is more than twice the value of β related to station EAC. Therefore, the coefficient of the rate function is more than twice larger in station EAC than in station TLA. (Note that $\hat{\lambda}^{(EAC)}(t) = 0.69\, t^{0.81}$ and that $\hat{\lambda}^{(TLA)}(t) = 0.34\, t^{0.84}$.) The values of α related to stations UIZ and TAH (in region SE) are similar as are the values of β. Hence, their rate functions behave in a similar way.

The different values of β implies that, even though exceedances rates are decreasing, they decrease in different forms. The larger (smaller) the value of β, the smaller (larger) the value of $\lambda(\cdot)$ for given fixed values of α and $t \geq 0$. Therefore, the model considered here, also capture the differences in the behaviour of different monitoring sites.

Table 2 presents the estimated means of the vectors $\boldsymbol{\mu}^{\alpha}$ and $\boldsymbol{\mu}^{\beta}$ as well as their respective standard deviations and 95% credible intervals for all stations considered. In all cases, the values of μ^{α} are negative and the values of μ^{β} are in the interval (-0.81, 1.22). Even though the 95% credible intervals of the $\boldsymbol{\mu}^{\alpha}$ are very wide, we would like to point out that the values of $\boldsymbol{\mu}^{\alpha}$ correspond to only one of the parameters of the mean and variance of the log-normal distribution of the vector $\boldsymbol{\alpha}$. The extremes in the credible intervals provide a parameter that allows the sampling of suitable values of $\boldsymbol{\alpha}$. Since the fit of the estimated accumulated mean to the data is good in all cases (see Figures 4, 5, and 6), the width of the confidence interval seems not to have a significant influence on the estimated values of $\boldsymbol{\alpha}$.

Table 2 about here.

The estimated mean values of ψ_r, ψ_a, ϕ_α, and ϕ_β as well as their respective standard deviations and 95% credible intervals are given in Table 3. The anisotropic angle ψ_a and the anisotropic ratio ψ_r^{-1} are approximately $\pi/2$ and 0.84, respectively. That indicates that there is a slight north-south deformation of the space where the stations are located (when considering the transformed space). That corresponds to stations located in the east-west direction in the original set of coordinates. Hence, monitoring sites in that direction are more likely to have the parameters of the models influencing each other. The influence of ϕ_α and ϕ_β is minimal.

Table 3 about here.

Figures 4 and 5 show the estimated and observed accumulated means for each of the monitoring stations considered here. (Here, the term accumulate mean means the function $m(\cdot)$ - either estimated or observed - evaluated at each point in time where an ozone exceedance occurred.) Figure 4 shows the plots for stations located in regions NW, NE, and CE, and Figure 5 shows the equivalent plots for stations located in regions SE and SW with the exception of the assumed non-observable station PLA. Figure 4 shows an almost perfect fit of the estimated to the observed mean in almost all cases. The exception being station MER located in region CE (last plot in the figure). In that case we may see that perhaps a model allowing the presence of a change-point or a different rate function could be considered. Figure 5 shows a good fit of the estimated to the observed accumulated mean. However, for stations CUA and COY, perhaps a model either with a different rate function or with the presence of change-points could provide a better fit.

Figure 4 about here.

Figure 5 about here.

4.3 Prediction results

Using the estimated parameters of the model, we may predict the number of exceedances of the threshold of interest as well as the days in which they occur at any location. That is performed as follows. Taking the estimated values of α and β for the rate function corresponding to the site PLA, we substitute them in $\lambda^{(PLA)}(t)$ and, using the non-homogeneous Poisson probability with respect to that rate function, we obtain a sample of size 5000 of the number of exceedances K_{PLA}.

The estimated value of K for site PLA is 504 with a standard deviation of 22.79 and 95% credible interval (459, 549). (Note that the actually observed value of K for site PLA is 518.) The K estimated days where a surpassing of 0.11 may have occurred are then obtained as follows. It is a well known fact (Cox and Lewis 1966) that, given a time interval $[0, T]$ where there are K occurrences of events of a non-homogeneous Poisson process with continuous mean function $m(t)$, the times $t_1, t_2, \ldots, t_K$, when those events occurred are distributed as order statistics of a sample with distribution $F(t) = m(t)/m(T)$, $t \in [0, T]$.

Therefore, we just have to generate K values from $F(\cdot)$ (where in our case the parameters α and β are replaced by their estimates) and order them in an increasing order. The resulting values are the estimated days at which exceedances of the threshold of interest may occur.

Figure 6 shows the plots of the accumulated means in the case of station PLA. Both estimated accumulated means predict well the behaviour of the actual observed accumulated mean.

Figure 6 about here.

Remark. In order to have the plot of the observed accumulated mean in the case of the station PLA, we have used the real observed data.

4.4 Correlation coefficients

Associate stations EAC, TLA, SAG, CHA, MER, UIZ, TAH, PED, CUA, COY and PLA with the natural numbers $i = 1, 2, \ldots, 11$, respectively. Indicate by $\rho_{\boldsymbol{\alpha}}$ and $\rho_{\boldsymbol{\beta}}$ the correlation matrices of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively, obtained using the generated values for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ provided by the MCMC algorithm. Those matrices are given as follows (due to symmetry we only report the matrices in the upper triangle form).

$$\rho_{\boldsymbol{\alpha}} = \begin{pmatrix} 1.0 & 0.32 & 0.02 & -0.014 & 0.13 & 0.045 & 0.009 & 0.14 & 0.19 & 0.104 & -0.013 \\ & 1.0 & 0.092 & 0.005 & 0.13 & 0.057 & 0.017 & 0.074 & 0.077 & 0.058 & -0.007 \\ & & 1.0 & 0.13 & 0.13 & 0.091 & 0.036 & 0.029 & 0.015 & 0.068 & -0.015 \\ & & & 1.0 & 0.034 & 0.066 & 0.041 & 0.012 & -0.002 & 0.007 & 0.0008 \\ & & & & 1.0 & 0.25 & 0.042 & 0.13 & 0.088 & 0.27 & -0.026 \\ & & & & & 1.0 & 0.15 & 0.112 & 0.037 & 0.207 & 0.012 \\ & & & & & & 1.0 & 0.078 & 0.006 & 0.089 & -0.007 \\ & & & & & & & 1.0 & 0.24 & 0.397 & -0.04 \\ & & & & & & & & 1.0 & 0.15 & -0.022 \\ & & & & & & & & & 1.0 & -0.023 \\ & & & & & & & & & & 1.0 \end{pmatrix}$$

and

$$\rho_{\boldsymbol{\beta}} = \begin{pmatrix} 1.0 & 0.15 & -0.036 & -0.035 & 0.062 & -0.081 & 0.049 & 0.15 & 0.103 & 0.064 & -0.0076 \\ & 1.0 & 0.083 & -0.046 & -0.043 & 0.086 & 0.0061 & 0.157 & 0.023 & -0.055 & -0.071 \\ & & 1.0 & -0.0003 & 0.061 & 0.136 & 0.047 & -0.0013 & 0.049 & 0.013 & 0.024 \\ & & & 1.0 & -0.13 & 0.032 & -0.028 & -0.04 & -0.049 & 0.011 & 0.0595 \\ & & & & 1.0 & -0.071 & 0.096 & 0.11 & 0.038 & 0.155 & -0.076 \\ & & & & & 1.0 & 0.007 & -0.011 & 0.138 & 0.099 & -0.0082 \\ & & & & & & 1.0 & -0.038 & -0.026 & 0.103 & -0.078 \\ & & & & & & & 1.0 & 0.224 & 0.209 & -0.149 \\ & & & & & & & & 1.0 & 0.133 & -0.067 \\ & & & & & & & & & 1.0 & -0.036 \\ & & & & & & & & & & 1.0 \end{pmatrix}$$

Looking at the matrix $\rho_{\boldsymbol{\alpha}}$, which corresponds to the parameters of the Poisson model that dictates how the rate function decreases/increases, we may see that with the exception of the stations CHA (in region NE) and UIZ (in region SE), the parameter α

corresponding to the station PLA is negatively correlated with the parameters α of the remaining stations. However, we may see that in all cases these correlation coefficients are of order 10^{-2} or less. That can also be said about most of the correlation coefficients related to stations other than PLA. Among the exceptions we have that the strongest correlations are those between the parameters α related to the stations PED and COY (value 0.397) both in region SW; stations TLA and EAC (value 0.32) both in region NW; stations MER, in region CE, and station COY, in region SW (value 0.27); stations MER, in region CE, and station UIZ, in region SE (value (0.25); stations PED and CUA (value 0.24) both in region SW; and stations UIZ, in region SE, and station COY, in region SW (value 0.207). Looking at Figure 1, we may see that larger correlation coefficients are present in stations sharing the same region and when in different regions, they are either geographically close to each other or in the wind's path. Similar analysis can be performed for the parameters β.

5 Discussion

In this work, we have used a spatial model applied to the parameters of a non-homogeneous Poisson process in order to study the behaviour of ozone in Mexico City. The aim was to impose a spatial relationship among the parameters of the non-homogeneous Poisson processes, which is used to count the number of times that a given threshold is surpassed in a time interval of interest. Using that information, the aim was to infer the behaviour of the process in sites where measurements could not be taken directly.

The methodology considered here, provides a mechanism for estimating the number of times that an environmental threshold would be surpassed in a site where we are not able to measure a pollutant's concentration. That is made using information provided by sites where measurements are actually observable.

In order to estimate the parameters present in the model, we have used a Gibbs sampling algorithm applied to their posterior distribution. We can observe by looking at Figures 4, 5, and 6, that the fit provided by the estimated accumulated means to the observed ones is, in general, good. Therefore, in this particular situation the methodology considered here predicts well the behaviour of ozone in sites where measurements are not available as well as describe well that behaviour at sites where measurements are available.

Remark. Note that even though in the model considered here, the data used correspond to the exceedance of a threshold by the ozone concentration, when we analyse the fit of the model to the observed data, we consider the accumulated means. This is made because when the means fit well, then the values of the probability (1) which is used to predict the number of exceedances (value of K) in a time interval $(t, t+s)$ in a given site, and the values of the one used to predicted exceedance days, dependent on the mean function. Hence, if the fit between observed and estimated means is good, then the prediction of the values of K and the surpassing days $d_1, d_2, \ldots, d_K$, have a large probability of being good as well.

Looking at Figures 4, 5, and 6, it seems that the estimated means fit well the observed ones. If one wants to see, in numerical terms, how well those fit are, then one could obtain the mean and standard deviation of the absolute value of the differences between the observed and estimated accumulated means. Hence, going in that direction, the means of the differences are 7.9, 7.11, 11.47, 10.61, 25.02, 14.91, 3.23, 16.08, 17.69, 22.32, and 20.43, in the cases of sites EAC, TLA, SAG, CHA, MER, UIZ, TAH, PED, CUA, COY, and PLA, respectively. The corresponding standard deviations are 5.6, 4.62, 6.04, 6.73, 12.79, 8.06, 10.1, 10.09, 13.38, 14.67, and 10.27. These values corroborate what the graphical comparison indicates. For instance, in the case of sites MER, CUA, and COY the fit was good but not as good as in other cases. That is reflected in the large mean of the differences and also by the corresponding standard deviation. Similar analysis can be performed for the other sites. Therefore, perhaps in the case of MER and other similar behaving sites, a model with the presence of change-points could be used in order to improve the estimations. We also have that in the case of station PLA, the mean of the differences of means is among the largest. However, recall that the observed data in that station was not taken into account when we estimated the parameters of the model. Also, the exceedance days are result of prediction.

If we look at the plots of the differences between the different pairs of accumulated means, we may see in which period of time the estimated values either over- or under-estimate the observed values. In Figures 7 and 8, we have the plots of $\overline{m}(\cdot) = \hat{m}(\cdot) - m(\cdot)$, where $m(\cdot)$ and $\hat{m}(\cdot)$ are the observed and the estimted accumulated means, respectively. The values of $\overline{m}(\cdot)$ were evaluated at points t where an ozone exceedances occurred in the particular site.

Figure 7 about here.

Figure 8 about here.

Looking at Figures 7 and 8, we may notice that there is a large variability in the ways the estimated means behave with respect to the observed ones. In general, the estimated mean under-estimate the observed one. In the beginning of the observational period we have, in general, an over-estimation. The under-estimation at the end of the observational period, could be considered an aid to the environmental authorities. If the estimated results reflect a bad situation, then in reality the situation could be worse. if the estimated results reflect a good situation, then perhaps an analysis of the real situation should be made in order to see how bad/good it is.

The effect of the anisotropic parameters in the deformation of the space are presented in Figure 9. The strongest effect on the transformed set of coordinates is the one given by the rotation of the map. The shrinkage is small (recall that the value corresponding to the shrinkage in the transformed set of coordinates is $\psi_r^{-1} = 0.84$). Therefore, it seems that for the chosen set of monitoring stations, the transformed space and the original one differ mostly by the rotation of the Metropolitan Area by an angle of $\pi/2$.

Figure 9 about here.

The rotation angle of $\pi/2$ indicates that the north-south direction is the one where deformation of the space should be made. That is compatible with the fact that there are influence of winds coming from the north and northeast to the south of the city. However, since the deformation parameter is 0.84, the deformation is small. Therefore, even though there is an evidence that the north-south wind direction produces some effect in the way sites might be correlated, that effect seems to be small.

We also can see (Figure 6) that the estimated days of exceedances also reflect well what happens in reality. Hence, the methodology considered here can be applied successfully to estimate the ozone behaviour at sites where measurements are not possible to obtain.

Remark. Note that even though we are only considering sites where all measurements are present, we may also include sites where only partial observations are available. These locations could be treated as unobserved sites and data could be simulated using the posterior distributions in the same manner as in the case of station PLA.

Note that in the case of station PLA, the estimated value of K, when compared to the real value, is reasonable. Therefore, using the estimated mean in the non-homogeneous Poisson probability, provide a reasonable estimate of the number of exceedances in a give time interval.

Acknowledgements

The authors thank two anonymous reviewers and an Associate Editor for their careful reading of this work, and for their comments and suggestions that helped to improve the presentation of the results. ERR and MHT were partially funded by the project PAPIIT-IN102713-3 of the Dirección General de Apoyo al Personal Académico of the Universidad Nacional Autónoma de México, Mexico. ERR is also grateful to the Departments of Statistics of the Universidade Estadual Paulista "Júlio de Mesquita Filho" – Campus Presidente Prudente and of the Universidade Federal do Rio de Janeiro, both in Brazil, for support and hospitality. MHT thanks the Instituto de Matemáticas of the Universidad Nacional Autónoma de México, Mexico, for support and hospitality. DG was supported by a grant from CNPq-Brazil.

Appendix

In this section we present the expressions for the full conditional posterior distributions from which the parameters were sampled. We will denote by $\overline{\boldsymbol{\theta}}_{(-\eta)}$ the complete set of parameters without the component η. Whenever we are able to sample directly from the full conditional posterior distributions that will be made. However, when that is not possible, sampled values will be obtained through a Metropolis-Hastings step.

As follows we have the full conditional posterior distributions of the parameters ψ_a and ψ_r. It is possible to see that those expressions are very complex and sampling directly from them might be complicated. Therefore, in each case we use a Metropolis-Hastings algorithm to obtain the values used in the sample. In that case, proposed values are generated using the respective prior distributions and they are either accepted or rejected according to the acceptance probability of the Metropolis-Hastings procedure.

$$\begin{aligned} P(\psi_a \,|\, \overline{\boldsymbol{\theta}}_{(-\psi_a)}, \mathbf{D}) \;\propto\;& |\Sigma^{\alpha}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)|^{-1/2} \left(\Sigma^{\beta}(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)\right)^{-1/2} \\ & \exp\left[-\frac{1}{2}\,(\log\boldsymbol{\alpha} - \boldsymbol{\mu}^{\boldsymbol{\alpha}})^t\,(\Sigma^{\alpha}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha))^{-1}\,(\log\boldsymbol{\alpha} - \boldsymbol{\mu}^{\boldsymbol{\alpha}})\right] \\ & \exp\left[-\frac{1}{2}\,\left(\log\boldsymbol{\beta} - \boldsymbol{\mu}^{\boldsymbol{\beta}}\right)^t\,\left(\Sigma^{\beta}(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)\right)^{-1}\,\left(\log\boldsymbol{\beta} - \boldsymbol{\mu}^{\boldsymbol{\beta}}\right)\right]. \end{aligned}$$

$$\begin{aligned} P(\psi_r \,|\, \overline{\boldsymbol{\theta}}_{(-\psi_r)}, \mathbf{D}) \;\propto\;& |\Sigma^{\alpha}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)|^{-1/2}\,|\Sigma^{\beta}(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)|^{-1/2}\,\frac{1}{\psi_r^{c+1}} \\ & \exp\left[-\frac{1}{2}\,(\log\boldsymbol{\alpha} - \boldsymbol{\mu}^{\boldsymbol{\alpha}})^t\,(\Sigma^{\alpha}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha))^{-1}\,(\log\boldsymbol{\alpha} - \boldsymbol{\mu}^{\boldsymbol{\alpha}})\right] \\ & \exp\left[-\frac{1}{2}\,\left(\log\boldsymbol{\beta} - \boldsymbol{\mu}^{\boldsymbol{\beta}}\right)^t\,\left(\Sigma^{\beta}(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)\right)^{-1}\,\left(\log\boldsymbol{\beta} - \boldsymbol{\mu}^{\boldsymbol{\beta}}\right)\right]. \end{aligned}$$

The full conditional posterior distribution of the blocks corresponding to the parameters $\boldsymbol{\mu}^{\alpha}$ and $\boldsymbol{\mu}^{\beta}$ are given as follows.

$$P(\boldsymbol{\mu}^{\alpha} \,|\, \overline{\boldsymbol{\theta}}_{(-\boldsymbol{\mu}^{\alpha})}, \mathbf{D}) \;\propto\; \exp\left[-\frac{1}{2}\,(\log\boldsymbol{\alpha} - \boldsymbol{\mu}^{\boldsymbol{\alpha}})^t\,(\Sigma_{\alpha}(\psi_a, \psi_r, \phi_{\alpha}, \boldsymbol{\sigma}_{\alpha}))^{-1}\,(\log\boldsymbol{\alpha} - \boldsymbol{\mu}^{\boldsymbol{\alpha}})\right]$$

$$\exp\left[-\frac{1}{2}\sum_{i=1}^{N_O+N_U}\left(\frac{\mu^{\alpha_i} - a_i}{b_i}\right)^2\right].$$

$$P(\boldsymbol{\mu}^{\beta} \,|\, \overline{\boldsymbol{\theta}}_{(-\boldsymbol{\mu}^{\beta})}, \mathbf{D}) \;\propto\; \exp\left[-\frac{1}{2}\,\left(\log\boldsymbol{\beta} - \boldsymbol{\mu}^{\boldsymbol{\beta}}\right)^t\,\left(\Sigma^{\beta}(\psi_a, \psi_r, \phi_{\beta}, \boldsymbol{\sigma}_{\beta})\right)^{-1}\,\left(\log\boldsymbol{\beta} - \boldsymbol{\mu}^{\boldsymbol{\beta}}\right)\right]$$

$$\exp\left[-\frac{1}{2}\sum_{i=1}^{N_O+N_U}\left(\frac{\mu^{\beta_i} - c_i}{d_i}\right)^2\right].$$

In each step of the Gibbs sampling algorithm, each component of the vector $\boldsymbol{\mu}^{\alpha}$ is proposed according to its prior distribution. Then, each component is separately either accepted or rejected in a individual acceptance-rejection method with the acceptance probability given by the Metropolis-Hastings step. The vector $\boldsymbol{\mu}^{\beta}$ will also be constructed in similar manner

In the case of the parameters ϕ_{α} and ϕ_{β}, the corresponding full conditional posterior distributions are,

$$P(\phi_{\alpha} \,|\, \overline{\boldsymbol{\theta}}_{(-\phi_{\alpha})}, \mathbf{D}) \;\propto\; |\Sigma^{\alpha}(\psi_a, \psi_r, \phi_{\alpha}, \boldsymbol{\sigma}_{\alpha})|^{-1/2}\;\frac{e^{-\mu_{\phi_{\alpha}}/\phi_{\alpha}}}{\phi_{\alpha}^{a_1+1}}$$

$$\exp\left[-\frac{1}{2}\,(\log\boldsymbol{\alpha} - \boldsymbol{\mu}^{\boldsymbol{\alpha}})^t\,(\Sigma^{\alpha}(\psi_a, \psi_r, \phi_{\alpha}, \boldsymbol{\sigma}_{\alpha}))^{-1}\,(\log\boldsymbol{\alpha} - \boldsymbol{\mu}^{\boldsymbol{\alpha}})\,|\right]$$

and

$$P(\phi_\beta \,|\, \overline{\boldsymbol{\theta}}_{(-\phi_\beta)}, \mathbf{D}) \;\; \propto \;\; |\Sigma^\beta(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)|^{-1/2} \; \frac{e^{-\mu_{\phi_\beta}/\phi_\beta}}{\phi_\beta^{a_2+1}}$$

$$\exp\left[-\frac{1}{2}\,\left(\log \boldsymbol{\beta} - \boldsymbol{\mu}^{\boldsymbol{\beta}}\right)^t\,\left(\Sigma^\beta(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)\right)^{-1}\,\left(\log \boldsymbol{\beta} - \boldsymbol{\mu}^{\boldsymbol{\beta}}\right)\right].$$

Due to the complexity of the expressions involved, a Metropolis-Hastings algorithm will also be used to generate the sampled values. Proposed values are produced using the respective prior distributions.

In the case of the blocks corresponding to the parameters $\boldsymbol{\alpha}^{(O)}$, $\boldsymbol{\alpha}^{(U)}$, $\boldsymbol{\beta}^{(O)}$ and $\boldsymbol{\beta}^{(U)}$ the full conditional posterior distributions are given, respectively, by

$$P(\boldsymbol{\alpha}^{(O)}, \boldsymbol{\alpha}^{(U)} \,|\, \overline{\boldsymbol{\theta}}_{(-\boldsymbol{\alpha})}, \mathbf{D}) \propto P(\boldsymbol{\alpha}^{(U)} \,|\, \boldsymbol{\alpha}^{(O)})\, P(\boldsymbol{\alpha}^{(O)} \,|\, \overline{\boldsymbol{\theta}}_{(-\boldsymbol{\alpha})}, \mathbf{D})$$

$$= P(\boldsymbol{\alpha}^{(U)} \,|\, \boldsymbol{\alpha}^{(O)}, \overline{\boldsymbol{\theta}}_{(-\boldsymbol{\alpha})}, \mathbf{D}) \times$$

$$\times \exp\left[-\sum_{i=1}^{N_O}\left(\frac{T_i}{\beta_i}\right)^{\alpha_i}\right]\left[\prod_{i=1}^{N_O}\prod_{k=1}^{K_i}\frac{d_{k,i}^{\alpha_i-1}}{\beta_i^{\alpha_i-1}}\right]\frac{|\Sigma^{\alpha^{(O)},\alpha^{(O)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)|^{-1/2}}{\left(\prod_{i=1}^{N_O}\alpha_i\right)} \times \left[\prod_{i=1}^{N_O}\left(\frac{\alpha_i}{\beta_i}\right)^{K_i}\right]$$

$$\times \exp\left[-\frac{1}{2}\left(\log \boldsymbol{\alpha}^{(O)} - \boldsymbol{\mu}^{\boldsymbol{\alpha}^{(O)}}\right)^t\left(\Sigma^{\alpha^{(O)},\alpha^{(O)}}(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)\right)^{-1}\left(\log \boldsymbol{\alpha}^{(O)} - \boldsymbol{\mu}^{\boldsymbol{\alpha}^{(O)}}\right)\right]$$

and

$$P(\boldsymbol{\beta}^{(O)}, \boldsymbol{\beta}^{(U)} \,|\, \boldsymbol{\theta}_{(-\boldsymbol{\beta})}, \mathbf{D}) \propto P(\boldsymbol{\beta}^{(U)} \,|\, \boldsymbol{\beta}^{(O)})\, P(\boldsymbol{\beta}^{(O)} \,|\, \overline{\boldsymbol{\theta}}_{(-\boldsymbol{\beta})}, \mathbf{D})$$

$$= P(\boldsymbol{\beta}^{(U)} \,|\, \boldsymbol{\beta}^{(O)}, \boldsymbol{\theta}_{(-\boldsymbol{\beta})}, \mathbf{D}) \times$$

$$\times \frac{|\Sigma^{\beta^{(O)},\beta^{(O)}}(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)|^{-1/2}}{\left(\prod_{i=1}^{N_O} \beta_i\right)} \exp\left[-\sum_{i=1}^{N_O} \left(\frac{T_i}{\beta_i}\right)^{\alpha_i}\right] \left[\prod_{i=1}^{N_O} \prod_{k=1}^{K_i} \frac{d_{k,i}^{\alpha_i - 1}}{\beta_i^{\alpha_i - 1}}\right] \times \left[\prod_{i=1}^{N_O} \left(\frac{\alpha_i}{\beta_i}\right)^{K_i}\right]$$

$$\times \exp\left[-\frac{1}{2}\left(\log \boldsymbol{\beta}^{(O)} - \boldsymbol{\mu}^{\boldsymbol{\beta}^{(O)}}\right)^t \left(\Sigma^{\beta^{(O)},\beta^{(O)}}(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)\right)^{-1} \left(\log \boldsymbol{\beta}^{(O)} - \boldsymbol{\mu}^{\boldsymbol{\beta}^{(O)}}\right)\right],$$

where $P(\boldsymbol{\alpha}^{(U)} \,|\, \boldsymbol{\alpha}^{(O)}, \boldsymbol{\theta}_{(-\boldsymbol{\alpha})}, \mathbf{D})$ is the multivariate log-normal distribution having its parameters given by (9) and (10). Similar comment is valid for $P(\boldsymbol{\beta}^{(U)} \,|\, \boldsymbol{\beta}^{(O)}, \boldsymbol{\theta}_{(-\boldsymbol{\beta})}, \mathbf{D})$.

Here, the generation of the values was made in two steps. First, using a Metropolis-Hastings algorithm, we generate $\boldsymbol{\alpha}^{(O)}$ and $\boldsymbol{\beta}^{(O)}$. In both cases, the proposed values are produced using the respective log-normal prior distributions. After obtaining the values of $\boldsymbol{\alpha}^{(O)}$ and $\boldsymbol{\beta}^{(O)}$, the generation of the parameter $\boldsymbol{\alpha}^{(U)}$ is performed using the log-normal distribution with parameters $\overline{\boldsymbol{\mu}}^{\alpha^{(U)}}$ and $\overline{\Sigma}^{\alpha^{(U)}}(\cdot)$, given by the conditional distribution of $\boldsymbol{\alpha}^{(U)}$ given $\boldsymbol{\alpha}^{(O)}$. Similar procedure is used in the case of $\boldsymbol{\beta}^{(U)}$, but now with $\overline{\boldsymbol{\mu}}^{\beta^{(U)}}$ and $\overline{\Sigma}^{\beta^{(U)}}(\cdot)$

Remark. Note that the case of the parameters $\boldsymbol{\alpha}^{(U)}$ and $\boldsymbol{\beta}^{(U)}$, we have a closed form for their distributions. However, in the case of $\boldsymbol{\alpha}^{(O)}$ and $\boldsymbol{\beta}^{(O)}$, the marginal conditional distribution do not have a closed form. Hence, the need of performing a Metropolis-Hastings step. The proposed values are generated using the respective prior distributions and each component is either accepted or rejected using an acceptance/rejection method.

If the vector of parameters $\boldsymbol{\sigma}_\alpha$ and $\boldsymbol{\sigma}_\beta$ are also considered as parameters to be estimated, their respective posterior distributions are,

$$P(\boldsymbol{\sigma}_\alpha \,|\, \overline{\boldsymbol{\theta}}_{(-\boldsymbol{\sigma}_\alpha)}, \mathbf{D}) \propto |\Sigma^\alpha(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)|^{-1/2} \, \frac{e^{-m/\sigma_\alpha}}{\sigma_\alpha^{n+1}}$$

$$\exp\left[-\frac{1}{2}\, (\log \boldsymbol{\alpha} - \boldsymbol{\mu}^{\boldsymbol{\alpha}})^t \, |\Sigma^\alpha(\psi_a, \psi_r, \phi_\alpha, \boldsymbol{\sigma}_\alpha)|^{-1} \, (\log \boldsymbol{\alpha} - \boldsymbol{\mu}^{\boldsymbol{\alpha}})\right]$$

and

$$P(\boldsymbol{\sigma}_\beta \,|\, \overline{\boldsymbol{\theta}}_{(-\boldsymbol{\sigma}_\beta)}, \mathbf{D}) \propto |\Sigma^\beta(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)|^{-1/2} \, \frac{e^{-w/\sigma_\beta}}{\sigma_\beta^{z+1}}$$

$$\exp\left[-\frac{1}{2}\left(\log \boldsymbol{\beta} - \boldsymbol{\mu}^{\boldsymbol{\beta}}\right)^t \, |\Sigma^\beta(\psi_a, \psi_r, \phi_\beta, \boldsymbol{\sigma}_\beta)|^{-1} \, \left(\log \boldsymbol{\beta} - \boldsymbol{\mu}^{\boldsymbol{\beta}}\right)\right].$$

Again, we need a Metropolis-Hastings step. The proposed values are obtained using the respective prior distributions and the acceptance/rejection of them are given by the respective Metropolis-Hastings acceptance probability.

The distribution, at an unobservable site i, of the number of days at which a surpass occurs is given by,

$$P(K_i \,|\, \boldsymbol{\theta}^{(U)}) \propto \frac{\left[m(T_i \,|\, \boldsymbol{\theta}^{(U)})\right]^{K_i}}{K_i!} \exp\left[-m(T_i \,|\, \boldsymbol{\theta}^{(U)})\right]$$

$$= \frac{1}{K_i!} \left(\frac{T_i}{\beta_i^{(U)}}\right)^{K_i\, \alpha_i^{(U)}} \exp\left[-\left(\frac{T_i}{\beta_i^{(U)}}\right)^{\alpha_i^{(U)}}\right],$$

and given K_i, the elements in $\mathbf{D}_i^{(U)}$ are distributed as the order statistics of a sample obtained from $F(t) = m^{(i)}(t)/m^{(i)}(T_i)$, $t \in [0, T_i]$, $i = 1, 2, \ldots, N_U$.

References

[1] Achcar JA, Fernández-Bremauntz AA, Rodrigues ER, Tzintzun G (2008) Estimating the number of ozone peaks in Mexico City using a non-homogeneous Poisson model. Environmetrics 19:469-485. doi:10.1002/env.890.

[2] Achcar JA, Rodrigues ER, Tzintzun G (2011a) Using non-homogeneous Poisson models with multiple change-points to estimate the number of ozone exceedances in Mexico City. Environmetrics 22:1-12. doi:10.002/env.1029.

[3] Achcar JA, Rodrigues ER, Tzintzun G (2011b) Using stochastic volatility models to analyse weekly ozone averages in Mexico City. Environ. Ecol. Stat. 18:271-290. doi:10.1007/s10651-010-0132-1

[4] Álvarez LJ, Fernández-Bremauntz AA, Rodrigues ER, Tzintzun G (2005) Maximum a posteriori estimation of the daily ozone peaks in Mexico City. Journal of Agricultural, Biological, and Environmental Statistics 10:276-290. doi:10.1198/108571105X5917.

[5] Bell ML, McDermont A, Zeger SL, Samet JM, Dominici F (2004) Ozone and short-term mortality in 95 US urban communities, 1987-2000. Journal of the American Medical Society 292:2372-2378.

[6] Bell ML, Peng R, Dominici F (2005) The exposure-response curve for ozone and risk of mortality and the adequacy of current ozone regulations. Environmental Health Perspectives 114:532-536.

[7] Castro-Morales FE, Gamerman D, Paez MS (2013) State space models with spatial deformation. Environmental and Ecological Statistics 20:191-214. doi: 10.1007/s10651-012-0215-2.

[8] Cifuentes L, Borja-Arbuto VH, Gouveia N, Thurston G, Davis DL (2001) Assessing the health benefits of urban air pollution reduction associated with climate change mitigation (2000-2020): Santiago, São Paulo, Mexico City and New York City. Environmental Health Perspectives 109:419-425.

[9] Cox DR, Lewis PA (1966) Statistical analysis of series events. Methuen. UK.

[10] Cressie NA (1991) Statistics for spatial data. John Wiley and Sons. USA.

[11] Diggle PJ, Ribeiro PJ Jr (2007) Model-based geostatistics. Springer. New York, USA.

[12] Dockery DW, Schwartz J, Spengler JD (1992), Air pollution and daily mortality: association with particulates and acid aerosols. Environmental Research 59:362-373.

[13] Galizia A, Kinney PL (1999) Long-term residence in areas of high ozone: association with respiratory health in a nationwide sample of nonsmoking adults. Environmental Health 99:675-679.

[14] Gamerman D, Lopes HF (2006) MCMC statistical simulation for Bayesian inference. Chapman and Hall. Boca Raton. USA.

[15] Gauderman WJ, Avol E, Gililand F, Vora H, Thomas D, Berhane K, McConnel R, Kuenzli N, Lurmman F, Rappaport E, Margolis H, Bates D, Peter J (2004) The effects of air pollution on lung development from 10 to 18 years of age. The New England Journal of Medicine 351:1057-1067. doi:10.1056/NEJMoa040610.

[16] Gelfand AE, Smith AFM (1990) Sampling-based approaches to calculating marginal densities. J. Amer. Statist. Assoc. 85:398-409. doi:10.1080/01621459.1990.10476213.

[17] Gouveia N, Fletcher T (2000) Time series analysis of air pollution and mortality: effects by cause, age and socio-economics status. Journal of Epidemiology and Community Health 54:750-755.

[18] Guardani R, Aguiar JL, Nascimento CAO, Lacava CIV, Yanagi Y (2003) Ground-level ozone mapping in large urban areas using multivariate analysis: application to the São Paulo Metropolitan Area. J. Air and Waste Management Assoc. 53:553-559.

[19] Huerta G, Sansó B, Stroud JR (2004) A spatiotemporal model for Mexico City ozone levels. Applied Statistics 53:231-248.

[20] Huerta G, Sansó B (2007) Time-varying models for extreme values. Environmental and Ecological Statistics 14:285-299. doi: 10.1007/s10651-007-0014-3.

[21] Javits JS (1980) Statistical interdependencies in the ozone national ambient air quality standard. Journal of Air Pollution Control Association 30:58-59. doi:10.1080/00022470.1980.10465918.

[22] Larsen LC, Bradley RA, Honcoop GL (1990) A new method of characterizing the variability of air quality-related indicators. In: Air and Waste Management Association, International Specialty Conference, Tropospheric Ozone and the Environment. Los Angeles, California Air and Waste Management Series, Pittsburgh, Penn., USA.

[23] Lawless JF (1982) Statistical models and methods for lifetime data. John Wiley and Sons. USA.

[24] Likens G (Lead Author) (2010) Environmental Protection Agency (Content source). (Davis W, Zaikowski L, Nodvin SC - topic eds). Acid rain. In: Cutler J (ed) Encyclopedia of Earth. Cleveland (Washington, D.C.: Environmental Information Coalition, National Council for Science and the Environment). (First published in the Encyclopedia of Earth October 9, 2006; Last revised January 2, 2010; Retrieved January 22, 2010). (Available from http://www.eoearth.org/article/Acid_rain.)

[25] Loomis D, Borja-Arbuto VH, Bangdiwala SI, Shy CM (1996) Ozone exposure and daily mortality in Mexico City: a time series analysis. Health Effects Institute Research Report 75:1-46.

[26] Martins LC, de Oliveira Latorre MRD, Saldiva PHN, Braga ALF (2002) Air pollution and emergency rooms visit due to chronic lower respiratory diseases in the elderly: an ecological time series study in São Paulo, Brazil. Journal of Occupational and Environmental Medicine 44:622-627.

[27] NOM (2002) Modificación a la Norma Oficial Mexicana NOM-020-SSA1-1993. Diario Oficial de la Federación. 30 October 2002. Mexico. (In Spanish.)

[28] Paez MS, Gamerman D (2003) Study of the space-time effects in the concentration of airborne pollutants in the Metropolitan Region of Rio de Janeiro. Environmetrics 14:387-408.

[29] Raftery AE (1989) Are ozone exceedance rate decreasing?, Comment of the paper "*Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone*" by R. L. Smith". Statistical Sciences 4:378–381.

[30] Robert CP, Casella G (1999) Monte Carlo statistical methods. Springer. New York. USA.

[31] Sahu SK, Gelfand AE, Holland DM (2007) High resolution space-time ozone modeling for assessing trends. Journal of the American Statistical Association 120:1221-1234

[32] Schmidt AM, Rodríguez MA (2010) Modelling multivariate counts varying continuously in space. In: Bernardo JM, Bayani MJ, Berger JO, David AP, Heckerman D, Smith AFM, West M (eds) Bayesian Inference 9. Oxford University Press, pp 1-20.

[33] Shaddick G, Yan H, Salway R, Vienneau D, Kounall D, Briggs D (2013) Large-scale Bayesian spatial modelling of air pollution for policy support. Journal of Applied Statistics 40:777-794.

[34] Smith RL (1989) Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. Statistical Sciences 4:367-393.

[35] Smith AFM, Roberts GO (1993) Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods (with discussion). J. Roy. Statist. Soc. Ser. B 55:3-23.

[36] Szpiro AA, Sampson PD, Sheppard L, Lumley T, Adar S.D, Kaufman JD (2010) Predicting intra-urban variation in air pollution concentrations with complex spatio-temporal dependencies. Environmetrics 21:606-631.

[37] Villaseñor-Alva JA, González-Estrada E (2010) On modelling cluster maxima with applications to ozone data from Mexico City. Environmetrics 21:528-540.

[38] WHO (2006) Air Quality Guidelines-2005. Particulate matter, ozone, nitrogen dioxide and sulfur dioxide. World Health Organization Regional Office for Europe, EU.

Tables caption

- **Table 1**. Estimated mean values, standard deviation (indicated by SD) and 95% credible intervals of the parameters of the non-homogeneous Poisson model for all station considered. The letters inside the parenthesis, beside the name of a station, represent the region where it is located.

- **Table 2**. Estimated mean values, standard deviation (indicated by SD) and 95% credible intervals of the parameters μ^{α} and μ^{β} for all station considered. The letters inside the parenthesis, beside the name of a station, represent the region where it is located.

- **Table 3**. Estimated mean values, standard deviation (indicated by SD) and 95% credible intervals of the parameters of the anisotropy model.

Table 1

Station		Mean	SD	95% Credible Interval
EAC (NW)	α	0.81	0.04	(0.73, 0.88)
	β	1.45	0.48	(0.66, 2.58)
TLA (NW)	α	0.84	0.05	(0.76, 0.94)
	β	2.98	1.11	(1.4, 5.7)
SAG (NE)	α	0.77	0.05	(0.68, 0.89)
	β	3.51	1.57	(1.35, 7.3)
CHA (NE)	α	0.69	0.05	(0.6, 0.78)
	β	1.73	0.83	(0.59, 3.72)
MER (CE)	α	0.76	0.04	(0.68, 0.85)
	β	0.82	0.35	(0.29, 1.66)
UIZ (SE)	α	0.84	0.03	(0.78, 0.9)
	β	1.07	0.29	(0.58, 1.77)
TAH (SE)	α	0.8	0.04	(0.73, 0.88)
	β	1.18	0.44	(0.55, 2.27)
PED (SW)	α	0.82	0.03	(0.77, 0.9)
	β	0.63	0.22	(0.36, 1.24)
CUA (SW)	α	0.8	0.04	(0.73, 0.88)
	β	0.86	0.32	(0.4, 1.68)
COY (SW)	α	0.8	0.03	(0.74, 0.86)
	β	0.57	0.19	(0.3, 0.99)
PLA (SW)	α	0.87	0.054	(0.2, 2.25)
	β	1.43	0.57	(0.52, 2.73)

Table 2

Station		Mean	SD	95% Credible Interval
EAC (NW)	μ^α	- 0.61	0.47	(-1.51, -0.3)
	μ^β	0.33	0.44	(-0.56, 1.18)
TLA (NW)	μ^α	- 0.48	0.48	(-1.4, -0.45)
	μ^β	1.04	0.4	(0.27, 1.84)
SAG (NE)	μ^α	- 0.45	0.48	(-1.4, -0.51)
	μ^β	1.22	0.5	(0.25, 2.21)
CHA (NE)	μ^α	- 0.47	0.49	(-1.43, -0.48)
	μ_β	0.57	0.53	(-0.48, 1.59)
MER (CE)	μ^α	- 0.65	0.47	(-1.57, -0.27)
	μ^β	-0.22	0.5	(-1.22, 0.76)
UIZ (SE)	μ^α	- 0.49	0.47	(-1.4, -0.46)
	μ^β	0.11	0.34	(-0.59, 0.75)
TAH (SE)	μ^α	- 0.44	0.5	(-1.4, -0.55)
	μ^β	0.18	0.4	(-0.6, 0.99)
PED (SW)	μ^α	- 0.97	0.49	(-1.9, -0.03)
	μ^β	-0.77	0.45	(-1.64, 0.15)
CUA (SW)	μ^α	- 0.67	0.49	(-1.63, -0.32)
	μ^β	-0.2	0.46	(-1.11, 0.72)
COY (SW)	μ^α	- 0.59	0.08	(-1.8, -0.06)
	μ^β	-0.75	0.44	(-1.57, 0.15)
PLA (SW)	μ^α	-1.96	0.43	(-2.85, -1.22)
	μ^β	-0.81	0.19	(-1.21, -0.46)

Table 3

	Mean	SD	95% Credible Interval
ψ_a	$1.58 \approx \pi/2$	0.9	(0.09, 3.07)
ψ_r^{-1}	0.84	0.13	(0.51, 0.995)
ϕ_α	1.07E-04	9.23E-05	(3.49E-05, 2.89E-04)
ϕ_β	1.92E-04	1.13E-04	(8.11E-05, 4.115E-04)

Figures caption

- **Figure 1**. Metropolitan Area of Mexico City with the location of the monitoring stations considered in the present study. Stations represented by ■ are the ones whose data were considered available and the one represented by • is the one whose data were considered unavailable. The thicker lines represent the boundary of the several regions in which the Metropolitan Area is divided. The thinner lines represent the boundary of Mexico City.

- **Figure 2** Ozone measurements in each of the monitoring station in regions NE, NW, and CE from 01 January 2005 to 31 December 2009.

- **Figure 3** Ozone measurements in each of the monitoring station in regions SE and SW, including station PLA, from 01 January 2005 to 31 December 2009.

- **Figure 4**. Estimated (dashed line) and observed (solid line) accumulated means for each of the monitoring stations located in regions NW, NE, and CE.

- **Figure 5**. Estimated (dashed line) and observed (solid line) accumulated means for each of the monitoring stations located in regions SE and SW with the exception of the assumed non observable station PLA.

- **Figure 6**. Observed and estimated accumulated means in the case of station PLA. The observed accumulated mean is represented by the solid line, the estimated accumulated mean obtained using the estimated parameters α and β is represented by the dashed solid line, and the estimated accumulated mean obtained using the estimated days of exceedance of the threshold 0.11 is represented by the dotted line.

- **Figure 7** Differences between the estimated and observed accumulated means for stations in regions NE, NW, and CE from 01 January 2005 to 31 december 2009.

- **Figure 8** Differences between the estimated and observed accumulated means for stations in regions SE and SW, including station PLA, from 01 January 2005 to 31 december 2009.

- **Figure 9**. Metropolitan Area of Mexico City with the location of the monitoring stations considered here using the coordinates in the transformed space. The new coordinates were obtained using the rotation and the deformation matrices with the estimated parameters of the model.