A non-homogeneous Poisson model and a reversible-jump MCMC algorithm to estimate the probability of occurrences of air pollution exceedances

Eliane R. Rodrigues[a,1], Mario H. Tarumoto[b], Juan A. Cruz-Juárez[c], Hortensia J. Reyes-Cervantes[c] and Guadalupe Tzintzun[d]

[a]*Instituto de Matemáticas, Universidad Nacional Autónoma de México, Area de la Investigación Científica, Circuito Exterior - Ciudad Universitaria, Ciudad de México 04510, Mexico*

[b]*Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista "Júlio de Mesquita Filho", Rua Roberto Simonsen, 305 - Centro Educacional, 19060-900 - Pres. Prudente - SP, Brazil*

[c]*Facultad de Ciencias Físico-Matemáticas, Benemérita Universidad Autónoma de Puebla, Edificio FM3-203, Ciudad Universitaria, Colonia San Manuel, 72570 Puebla, Mexico*

[d]*Instituto Nacional de Ecología y Cambio Climático, Secretaría de Medio Ambiente y Recursos Naturales, Blvd. Adolfo Ruiz Cortines, 4209, Colonia Jardines en la Montaña – Delegación Tlalpan, 14210 Ciudad de México, Mexico*

Abstract

We consider a reversible-jump Markov chain Monte Carlo algorithm to estimate the number of change-points, their locations and the parameters of the rate and mean functions in a non-homogeneous Poisson model. The model and algorithm are applied to ozone and particulate matter data obtained from the Mexico City monitoring network. Results are compared to those obtained in previous works using different approaches. Whereas in some cases the estimated change-points are placed in similar locations in others they differ in both number and locations. Additionally, in some cases different behaviours of the rate function are detected when compared to those given by previous studies.

Key words: Bayesian inference, Environmental exceedances, Non-homogeneous Poisson model, Reversible-jump MCMC

[1]Corresponding Author. E-mail: eliane@math.unam.mx

1 Introduction

Inhabitants of many cities around the world suffer from exposure to high levels of pollution. The adverse effects caused by this exposure can be very serious depending on the pollutant and its concentration. Among the many pollutants affecting the population's heath of a given city are ozone (O_3), particulate matter with diameter smaller than 10 microns (PM_{10}) and those with diameter smaller than 2.5 microns ($PM_{2.5}$). The hazardous effects of long-term exposure to those pollutants are also well known. Ozone may cause eyes irritation and upper respiratory system problems (Bell et al., 2004; Kelly, 2003; Loomis et al., 1996; Likens, 2010; WHO, 2006). Besides causing visibility problems and damage to the vegetation (EPA, 2018), if a population is exposed for long periods of time to high concentrations of PM_{10} and $PM_{2.5}$ there may be an increase in the risk of cardiovascular disease, lung cancer and other cardiopulmonary disorder (see, for instance, EPA, 2018; Feng et al., 2016; Janssen et al., 2013; Mauderly and Oberdörster, 1997; Xing et al., 2016; Thurston, 1996; Itô and Thurston, 1996; WHO, 2006; among others).

The interest here resides in comparing the results given by previous works (Achcar et al., 2008, 2011; Rodrigues et al., 2011, Súarez-Sierra et al., 2019, 2022) with those given by a reversible-jump Markov chain Monte Carlo (RJ-MCMC) algorithm (Green, 1995; Carlin and Chib, 1995) in terms of estimating of the number of change-points their locations and the parameters of the rate functions in a non homogeneous Poisson model used to study the behaviours of ozone, PM_{10} and $PM_{2.5}$ data obtained from the Mexico City monitoring network. The behaviour we are interested is related to the occurrences of exceedances of the corresponding environmental thresholds for those pollutants, as well as the locations of points in time where the data may change behaviour under a non-homogeneous Poisson models with change-points. In the present work we consider a RJ-MCMC to estimate the parameters present in the model. Reversible-jump MCMC algorithms have also been used to estimate the number and location of change-points in Poisson models with applications to air pollution data as in, for instance, Gyarmati-Szabó et al. (2010). In that work the rate function was a step function and a RJ-MCMC algorithm was used to estimate the location and number of change-points, as well as the values of the constant functions between change-points. The algorithm was applied

to nitrogen oxides (NO_2 and NO) and carbon monoxide (CO) data from the city of Leeds in the United Kingdom. The present work differ from that of Gyarmati-Szabó et al. (2010) when we consider a non-homogeneous Poisson model with a Weibull rate function and change-points and use the algorithm to estimate the parameters of this rate function, as well as the number and locations of the change-points. The algorithm is an adaptation of that given by Green (1995) and that given by Álvarez et al. (2006) where instead of proposing new parameters using functions of the parameters present in the model at a given iteration, we propose them by generating directly from appropriate distributions. That simplifies the form of the acceptance probabilities.

This work is organised as follows. In Section 2, we present the mathematical and the Bayesian formulations of the model. Section 3 gives the description of the reversible-jump MCMC algorithm. In Section 4 the model and algorithm are applied to Mexico City ozone, PM_{10} and $PM_{2.5}$ measurements. Section 5 presents a discussion of the results and in Section 6 we conclude. In an appendix, placed after the section References, some plots and additional information used in the main text are given.

2 The mathematical and Bayesian models

Denote by $[0, T]$ $(T > 0)$ the observational period where measurements were taken. Let $N_t \geq 0$ be the random variable recording the number of times that a pollutant's concentration exceeds its corresponding environmental threshold L $(L > 0)$ in the time interval $[0, t)$, $t \geq 0$. Assume that there are $K \geq 0$ days in which exceedances occurred during $[0, T]$, denote by $d_1, d_2, \ldots, d_K$ these exceedances days, and by $\mathbf{D} = \{d_1, d_2, \ldots, d_K\}$ the observed data. Let $N = \{N_t : t \geq 0\}$ be 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 take $\lambda(\cdot)$ of the Weibull form, i.e., $\lambda(t) = (\alpha/\sigma)\,(t/\sigma)^{\alpha-1}$ where $m(t) = (t/\sigma)^{\alpha}$ is the associated mean function, $t \geq 0$; $\alpha, \sigma > 0$.

Remark. The Weibull rate function is chosen over other possibilities because depending on the value of the parameter α, the function $\lambda(\cdot)$ may present either a decreasing, a constant or an increasing behaviour in addition to the fact that this rate function has a simple form and

the analysis is made easier.

For $M \geq 0$ a known and fixed natural number, let $I \in \mathcal{S} = \{0, 1, \ldots, M\}$ be the number of change-points present in the model and denote them by $\tau_1, \tau_2, \ldots, \tau_I$. If $I = 0$, then no change-points are present. The variables I and τ_i, $i = 1, 2, \ldots, I$ are considered parameters that need to be estimated. Take $\tau_0 = 0$ and $\tau_{I+1} = T$. Let $\lambda_i(t)$ denote the rate function between the change-points τ_{i-1} and τ_i, $i = 1, 2, \ldots, I+1$, i.e., $\lambda_i(t) = (\alpha_i/\sigma_i)(t/\sigma_i)^{\alpha_i - 1}$ with $m_i(t) = (t/\sigma_i)^{\alpha_i}$ the corresponding mean function where α_i and σ_i are parameters that also need to be estimated, $i = 1, 2, \ldots, I+1$. Thus, in the case of multiple change-points, the vector of parameters is $\boldsymbol{\phi} = (I, \boldsymbol{\tau}, \boldsymbol{\theta})$, where $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_I)$ and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_{I+1})$, with $\boldsymbol{\theta}_i = (\alpha_i, \sigma_i)$, $i = 1, 2, \ldots, I+1$. If $I = 0$, then the vector of parameters simplifies to $\boldsymbol{\phi} = \boldsymbol{\theta} = (\alpha, \sigma)$. We assume that given $\boldsymbol{\tau}$ and I, the parameters $\alpha_1, \ldots, \alpha_{I+1}, , \sigma_1, \ldots, \sigma_{I+1}$ are independent a priori and that given I, $\boldsymbol{\tau}$ depends a priori only on I. Hence, given that we have I change-points, the vector $\boldsymbol{\phi}$ is an element of the state space $\mathcal{C}_I = \{I\} \times \mathbb{R}_+^I \times \mathbb{R}_+^{I+1} \times \mathbb{R}_+^{I+1} = \{I\} \times \mathbb{R}_+^{3I+2}$, $I \geq 1$. When $I = 0$, the state space simplifies to $\mathcal{C}_0 = \{0\} \times \mathbb{R}_+ \times \mathbb{R}_+ = \{0\} \times \mathbb{R}_+^2$. These subspaces belong to the general space $\mathcal{C} = \cup_{I=0}^{M} \mathcal{C}_I$.

We follow the Bayesian point of view to estimate the parameters present in the model (Gamerman and Lopes, 2000). Hence, we use the fact that for $P(\boldsymbol{\theta} \,|\, \mathbf{D})$ and $P(\boldsymbol{\theta})$, respectively, the prior and posterior distributions of the vector of parameters $\boldsymbol{\theta}$ of a model describing $\mathbf{D}$, and $L(\mathbf{D} \,|\, \boldsymbol{\theta})$ the likelihood function, we have $P(\boldsymbol{\theta} \,|\, \mathbf{D}) \propto L(\mathbf{D} \,|\, \boldsymbol{\theta})\, P(\boldsymbol{\theta})$. In the present case, the form of the posterior distribution is

$$P(\boldsymbol{\phi} \,|\, \mathbf{D}) \;\; \propto \;\; P(I, \boldsymbol{\tau}, \boldsymbol{\alpha}, \boldsymbol{\sigma} \,|\, \mathbf{D}) = L(\mathbf{D} \,|\, I, \boldsymbol{\tau}, \boldsymbol{\alpha}, \boldsymbol{\sigma})\, P(\boldsymbol{\alpha} \,|\, \boldsymbol{\tau}, I)\, P(\boldsymbol{\sigma} \,|\, \boldsymbol{\tau}, I)\, P(\boldsymbol{\tau} \,|\, I)\, P(I)$$

where $P(\boldsymbol{\alpha} \,|\, \boldsymbol{\tau}, I)$, $P(\boldsymbol{\sigma} \,|\, \boldsymbol{\tau}, I)$, $P(\boldsymbol{\tau} \,|\, I)$ and $P(I)$ are the prior distributions of $\boldsymbol{\alpha}$, $\boldsymbol{\sigma}$, $\boldsymbol{\tau}$ and I, respectively, given as follows.

Given $\boldsymbol{\tau}$ and I, the parameters α_i and σ_i will have gamma prior distributions Gamma(a_1, a_2) and Gamma(b_1, b_2), respectively, with a_1 and b_1 the shape and a_2 and b_2 the scale parameters; $i = 1, 2, \ldots, I+1$; as in Green (1995), Gyarmati-Szabó et al. (2010) and Ai (2012), $\boldsymbol{\tau}$ will have as prior distribution the distribution of the even labelled order statistics of $(2\,I+1)$ independent points uniformly distributed in $[0, T]$ (see, for instance, Arnold et al., 2008). The number of

change-points will have as its prior distribution a truncated Poisson defined on the set $\mathcal{S}$ with parameter λ, i.e., $P(I) \propto (\lambda^I/I!)\, 1_{\mathcal{S}}(I)$, where $1_A(x) = 1$, if $x \in A$ and is zero otherwise.

The hyperparameters a_1, a_2, b_1, b_2 and λ are considered known and will be specified when the model is applied to the data. Estimation of the parameters will be performed using samples drawn from the respective posterior distributions using a RJ-MCMC algorithm programmed in R.

When the algorithm is in a version of the model with $I \geq 1$ change-points, the likelihood function is of the form (Yang and Kuo, 2001; Achcar et al., 2011)

$$\begin{aligned} L(\mathbf{D} \,|\, \boldsymbol{\phi}) \;\; \propto \;\; & \left(\prod_{i=1}^{N_{\tau_1}} \lambda_1(d_i) \right) \exp\left[-m_1(\tau_1)\right] \\ & \left[\prod_{j=2}^{I} \left(\prod_{i=N_{\tau_{j-1}}+1}^{N_{\tau_j}} \lambda_j(d_i) \right) \exp\left(-[m_j(\tau_j) - m_j(\tau_{j-1})]\right) \right] \\ & \left(\prod_{i=N_{\tau_I}+1}^{K} \lambda_{I+1}(d_i) \right) \exp\left(-\left[m_{I+1}(T) - m_{I+1}(\tau_I)\right]\right), \end{aligned} \tag{1}$$

with N_{τ_i} representing the number of exceedance days before the change-point τ_i, $i = 1, 2, \ldots, I$. When $I = 0$, expression (1) simplifies to $L(\mathbf{D} \,|\, \boldsymbol{\theta}) = \left[\prod_{i=1}^{K} \lambda(d_i)\right] \exp\left[-m(T)\right]$ (Cox and Lewis, 1966; Lawless, 1982).

In order to estimate the number of change-points present in the model best describing the behaviour of the data, we need to use some selection criteria. Besides using the sample generated by the RJ-MCMC algorithm to obtain the mean and the mode of the distribution of the number of change-points, we will also use the graphical criterion which assess the fit of the estimated to the observed means, the Bayesian information criterion (BIC), the marginal likelihood function (ML) - see for instance Raftery (1996) - the sum of the absolute difference (SAD) between the accumulated observed and estimated means, and their plots. The BIC, ML and SAD are defined as follows. The BIC is given by $\text{BIC} = -2 \sum_{i=1}^{n} \log(L[\mathbf{D} \,|\, \boldsymbol{\theta}_i)] + p \, \log(n)$, where p is the length of $\boldsymbol{\theta}_i$ and n is the sample size - see, for example, Schwarz (1978) and Akaike (1977, 1978). The smaller the value the more adequate is the model. The ML of Model l may be approached by a Monte Carlo estimate, denoted by V_l, and given by $\hat{V}_l = \frac{1}{M'} \sum_{i=1}^{M'} L\left(\mathbf{D} \,|\, \boldsymbol{\theta}_i^{(l)}\right)$ where M'

is the size of the simulated Markov chain Monte Carlo sample and $\boldsymbol{\theta}_i^{(l)}$, $i = 1, 2, \ldots, M'$, is the sample generated when Model l is used. The model with the highest V_l is the chosen model. The SAD is given by SAD $= \sum_t |m(t) - \hat{m}(t)|$, with $\hat{m}(\cdot)$ the estimated value of the observed $m(\cdot)$ which in our case correspond, respectively, to the estimated mean function obtained using the respective estimated parameters, and the accumulated observed mean. According to this criterion the selected model is the one with the smallest SAD value.

3 A reversible-jump MCMC algorithm

In order to start the description of the algorithm, first define

$$b_I = c \min\left\{1, \frac{P(I+1)}{P(I)}\right\} \quad \text{and} \quad d_I = c \min\left\{1, \frac{P(I-1)}{P(I)}\right\}$$

with $c > 0$ a constant such that $b_I + d_I < 1$ and $P(\cdot)$ the prior distribution of the number of change-points. Note that since $I \in \{0, 1, 2, \ldots, M\}$, we have $d_0 = b_M = 0$. Additionally, by definition the reversibility condition $b_I\, P(I) = d_{I+1}\, P(I+1)$ is satisfied.

When we have I change-points, three types of moves may be proposed: either the increase by one of the number of change-points (birth) with probability b_I, the decrease by one of the number of change-points (death) with probability d_I, and change in the position of one of the change-points (replacement) with probability $r_I = 1 - b_I - d_I$. The possible moves are described as follows.

(a) Birth move: if a birth move is proposed, then

i. As in Green (1995), Gyarmati-Szabó et al. (2010) and Ai (2012), generate a new τ' using a uniform distribution $U(0, T)$. This new change-point will belong to an interval $[\tau_{k-1}, \tau_k)$ for some $k \in \{1, 2, \ldots, I+1\}$ with probability one. Denote by $\boldsymbol{\tau}' = (\tau'_1, \tau'_2, \ldots, \tau'_{I+1})$ the new vector of change-points where $\tau'_j = \tau_j$, $j = 1, 2, \ldots, k-1$, $\tau'_k = \tau'$, $\tau'_j = \tau_{j-1}$, $j = k+1, k+2, \ldots, I+1$, with the appropriate adaptation in the case where either the first or the last subinterval is chosen.

ii. With probability 1/2 choose either $[\tau'_{k-1}, \tau'_k)$ or $[\tau'_k, \tau'_{k+1})$ to host the new α' and σ' which are generated using their corresponding prior distributions. If $[\tau'_{k-1}, \tau'_k)$ is selected, then relabel the coordinates of the new vector of parameters of the rate function as follows: $\alpha'_j = \alpha_j$ and $\sigma'_j = \sigma_j$, $j = 1, 2, \ldots, k-1$, $\alpha'_k = \alpha'$ and $\sigma'_k = \sigma'$, $\alpha'_j = \alpha_{j-1}$ and $\sigma'_j = \sigma_{j-1}$, $j = k+1, k+2, \ldots, I+2$. Let $\boldsymbol{\alpha}' = (\alpha'_1, \alpha'_2, \ldots, \alpha'_{I+2})$ and $\boldsymbol{\sigma}' = (\sigma'_1, \sigma'_2, \ldots, \sigma'_{I+2})$ be the rate functions new vector of parameters. In the case $[\tau'_k, \tau'_{k+1})$ is selected, the relabelling is made in a similar way.

Indicate by $\boldsymbol{\phi}' = (I+1, \boldsymbol{\tau}', \boldsymbol{\alpha}', \boldsymbol{\sigma}')$ the new overall proposed vector of parameters. In this case the acceptance probability is

$$A(\boldsymbol{\phi}, \boldsymbol{\phi}') = \min\left\{1, \frac{L(\mathbf{D} \,|\, I+1, \boldsymbol{\tau}', \boldsymbol{\alpha}', \boldsymbol{\sigma}')}{L(\mathbf{D} \,|\, I, \boldsymbol{\tau}, \boldsymbol{\alpha}, \boldsymbol{\sigma})} \frac{P(\boldsymbol{\alpha}' \,|\, \boldsymbol{\tau}', I+1)}{P(\boldsymbol{\alpha} \,|\, \boldsymbol{\tau}, I)} \frac{P(\boldsymbol{\sigma}' \,|\, \boldsymbol{\tau}', I+1)}{P(\boldsymbol{\sigma} \,|\, \boldsymbol{\tau}, I)} \right.$$
$$\left. \frac{P(\boldsymbol{\tau}' \,|\, I+1)}{P(\boldsymbol{\tau} \,|\, I)} \frac{P(I+1)}{P(I)} \frac{Q(\boldsymbol{\phi}', \boldsymbol{\phi})}{Q(\boldsymbol{\phi}, \boldsymbol{\phi}')} \mathbf{J} \right\}$$

where $\mathbf{J}$ is the Jacobian of the transformation which is equal to one, with

$$\frac{P(\boldsymbol{\tau}' \,|\, I+1)}{P(\boldsymbol{\tau} \,|\, I)} = \frac{(2I+3)\, 2\, (I+1)}{T^2} \frac{(\tau' - \tau_{k-1})\, (\tau_k - \tau')}{(\tau_k - \tau_{k-1})}$$

(see, for instance, Arnold et al., 1992; Green, 1995; Gyarmat et al., 2010; Ai, 2012), and where $Q(\cdot, \cdot)$ is the proposal distribution which is such that

$$\frac{Q(\boldsymbol{\phi}', \boldsymbol{\phi})}{Q(\boldsymbol{\phi}, \boldsymbol{\phi}')} = \frac{1}{B} \frac{d_{I+1} \frac{1}{I+1}}{b_I \frac{1}{T}}$$

where $B = (1-a)\, P(\alpha'_k \,|\, \boldsymbol{\tau}', I+1)\, P(\sigma'_k \,|\, \boldsymbol{\tau}', I+1) + a\, P(\alpha'_{k+1} \,|\, \boldsymbol{\tau}', I+1)\, P(\sigma'_{k+1} \,|\, \boldsymbol{\tau}', I+1)$ with $a = 0$ if the newly generated α' and σ' are placed in the interval $[\tau'_{k-1}, \tau'_k)$ and is equal to one if they are placed in the new $(k+1)$th interval $[\tau'_k, \tau'_{k+1})$, and $P(\cdot \,|\, \boldsymbol{\tau}', I+1)$ and $P(\cdot \,|\, \boldsymbol{\tau}, I)$ the corresponding prior distributions of the parameters of the rate functions.

(b) Death move: if a death move is proposed, then

i. Select an index $k \in \{1, 2, \ldots, I\}$ uniformly and remove τ_k. Denote by $\boldsymbol{\tau}' = (\tau'_1, \tau'_2, \ldots, \tau'_{I-1})$ the new vector of change-points where $\tau'_j = \tau_j$, $j = 1, 2, \ldots, k-1$, $\tau'_j = \tau_{j+1}$,

$j = k, k+1, \ldots, I-1$, with the appropriate adaptation in the case where either the first or the last change-points is chosen.

ii. With probability 1/2 choose either α_k and σ_k or α_{k+1} and σ_{k+1} to be removed. If α_k and σ_k are removed, then the coordinates of the new vector of parameters of the rate function are as follows: $\alpha'_j = \alpha_j$ and $\sigma'_j = \sigma_j$, $j = 1, 2, \ldots, k-1$, $\alpha'_j = \alpha_{j+1}$ and $\sigma'_j = \sigma_{j+1}$, $j = k, k+1, \ldots, I$. Let $\boldsymbol{\alpha}' = (\alpha'_1, \alpha'_2, \ldots, \alpha'_I)$ and $\boldsymbol{\sigma}' = (\sigma'_1, \sigma'_2, \ldots, \sigma'_I)$ be the rate function new vector of parameters. If α_{k+1} and σ_{k+1} are chosen, then the relabelling is performed in a similar way.

Indicate by $\boldsymbol{\phi}' = (I-1, \boldsymbol{\tau}', \boldsymbol{\alpha}', \boldsymbol{\sigma}')$ the new overall proposed vector of parameters. The acceptance probability of the present move is

$$A(\boldsymbol{\phi}, \boldsymbol{\phi}') = \min\left\{1, \frac{L(\mathbf{D} \,|\, I-1, \boldsymbol{\tau}', \boldsymbol{\alpha}', \boldsymbol{\sigma}')}{L(\mathbf{D} \,|\, I, \boldsymbol{\tau}, \boldsymbol{\alpha}, \boldsymbol{\sigma})} \frac{P(\boldsymbol{\alpha}' \,|\, \boldsymbol{\tau}', I-1)}{P(\boldsymbol{\alpha} \,|\, \boldsymbol{\tau}, I)} \frac{P(\boldsymbol{\sigma}' \,|\, \boldsymbol{\tau}', I-1)}{P(\boldsymbol{\sigma} \,|\, \boldsymbol{\tau}, I)} \right.$$
$$\left. \frac{P(\boldsymbol{\tau}' \,|\, I-1)}{P(\boldsymbol{\tau} \,|\, I)} \frac{P(I-1)}{P(I)} \frac{Q(\boldsymbol{\phi}', \boldsymbol{\phi})}{Q(\boldsymbol{\phi}, \boldsymbol{\phi}')} \mathbf{J} \right\}$$

where $\mathbf{J}$ is the Jacobian of the transformation which is equal to one, with

$$\frac{P(\boldsymbol{\tau}' \,|\, I-1)}{P(\boldsymbol{\tau} \,|\, I)} = \frac{T^2}{(2I+1)\, 2\, I} \frac{(\tau_{k+1} - \tau_{k-1})}{(\tau_{k+1} - \tau_k)\,(\tau_k - \tau_{k-1})}$$

(see, for instance, Arnold et al., 1992; Green, 1995; Gyarmat et al., 2010; Ai 2012), where $Q(\cdot, \cdot)$ is the proposal distribution which is such that

$$\frac{Q(\boldsymbol{\phi}', \boldsymbol{\phi})}{Q(\boldsymbol{\phi}, \boldsymbol{\phi}')} = C \frac{b_{I-1}\, \frac{1}{T}}{d_I\, \frac{1}{I}},$$

with $C = (1-a)\, P(\alpha_k \,|\, \boldsymbol{\tau}, I)\, P(\sigma_k \,|\, \boldsymbol{\tau}, I) + a\, P(\alpha_{k+1} \,|\, \boldsymbol{\tau}, I)\, P(\sigma_{k+1} \,|\, \boldsymbol{\tau}, I)$ with $a = 0$ if the deleted coordinates are α_k and σ_k and is equal to one if the deleted coordinates are α_{k+1} and σ_{k+1}, with $P(\cdot \,|\, \boldsymbol{\tau}, I)$ and $P(\cdot \,|\, \boldsymbol{\tau}', I-1)$ the corresponding prior distributions of the parameters of the rate functions.

(c) Replacement move: if the replacement of one change-point, then

i. As in Green (1995), Gyarmati-Szabó et al. (2010) and Ai (2012), select an index $k \in \{1, 2, \ldots, I\}$ uniformly and generate a new τ' using a uniform distribution

$U(\tau_{k-1}, \tau_{k+1})$. Denote by $\boldsymbol{\tau}' = (\tau'_1, \tau'_2, \ldots, \tau'_I)$ the new vector of change-points where $\tau'_j = \tau_j,\ j = 1, 2, \ldots, k-1,\ \tau'_k = \tau',\ \tau'_j = \tau_j,\ j = k+1, k+2, \ldots, I$, with the appropriate adaptation in the case where either the first or the last change-points is chosen.

ii. Generate new α' and σ' using their corresponding prior distributions. With probability $1/2$ choose either α_k and σ_k or α_{k+1} and σ_{k+1} to be replaced. If α_k and σ_k are chosen, then the coordinates of the new vector of parameters of the rate function are as follows: $\alpha'_j = \alpha_j$ and $\sigma'_j = \sigma_j,\ j = 1, 2, \ldots, k-1,\ \alpha'_k = \alpha'$ and $\sigma'_k = \sigma'$, $\alpha'_j = \alpha_j$ and $\sigma'_j = \sigma_j,\ j = k+1, k+2, \ldots, I+1$. Let $\boldsymbol{\alpha}' = (\alpha'_1, \alpha'_2, \ldots, \alpha'_{I+1})$ and $\boldsymbol{\sigma}' = (\sigma'_1, \sigma'_2, \ldots, \sigma'_{I+1})$ be the rate function new vector of parameters. If α_{k+1} and σ_{k+1} are chosen, then the relabelling is performed in a similar way.

Indicate by $\boldsymbol{\phi}' = (I, \boldsymbol{\tau}', \boldsymbol{\alpha}', \boldsymbol{\sigma}')$ the new overall vector of parameters. The acceptance probability is of the present move is

$$A(\boldsymbol{\phi}, \boldsymbol{\phi}') = \min \left\{ 1, \frac{L(\mathbf{D} \,|\, I, \boldsymbol{\tau}', \boldsymbol{\alpha}', \boldsymbol{\sigma}')}{L(\mathbf{D} \,|\, I, \boldsymbol{\tau}, \boldsymbol{\alpha}, \boldsymbol{\sigma})} \frac{P(\boldsymbol{\alpha}' \,|\, \boldsymbol{\tau}', I)}{P(\boldsymbol{\alpha} \,|\, \boldsymbol{\tau}, I)} \frac{P(\boldsymbol{\sigma}' \,|\, \boldsymbol{\tau}', I)}{P(\boldsymbol{\sigma} \,|\, \boldsymbol{\tau}, I)} \right.$$

$$\left. \frac{P(\boldsymbol{\tau}' \,|\, I)}{P(\boldsymbol{\tau} \,|\, I)} \frac{P(I)}{P(I)} \frac{Q(\boldsymbol{\phi}', \boldsymbol{\phi})}{Q(\boldsymbol{\phi}, \boldsymbol{\phi}')} \mathbf{J} \right\}$$

where $\mathbf{J}$ is the Jacobian of the transformation which is equal to one, with

$$\frac{P(\boldsymbol{\tau}' \,|\, I)}{P(\boldsymbol{\tau} \,|\, I)} = \frac{(\tau_{k+1} - \tau'_k)\,(\tau'_k - \tau_{k-1})}{(\tau_{k+1} - \tau_k)\,(\tau_k - \tau_{k-1})}$$

(see, for instance, Arnold et al., 1992; Green, 1995; Gyarmat et al., 2010; Ai, 2012), and where $Q(\cdot, \cdot)$ is the proposal distribution which is such that

$$\frac{Q(\boldsymbol{\phi}', \boldsymbol{\phi})}{Q(\boldsymbol{\phi}, \boldsymbol{\phi}')} = D \frac{\frac{1}{\tau_{k-1} - \tau_{k+1}}\, r_I \frac{1}{I}}{\frac{1}{\tau_{k-1} - \tau_{k+1}}\, r_I \frac{1}{I}}$$

where

$$D = \frac{(1-a)\,P(\alpha_k \,|\, \boldsymbol{\tau}, I)\,P(\sigma_k \,|\, \boldsymbol{\tau}, I) + a\,P(\alpha_{k+1} \,|\, \boldsymbol{\tau}, I)\,P(\sigma_{k+1} \,|\, \boldsymbol{\tau}, I)}{(1-a)\,P(\alpha'_k \,|\, \boldsymbol{\tau}', I)\,P(\sigma'_k \,|\, \boldsymbol{\tau}', I) + a\,P(\alpha'_{k+1} \,|\, \boldsymbol{\tau}', I)\,P(\sigma'_{k+1} \,|\, \boldsymbol{\tau}', I)}$$

with $a = 0$ if the replaced coordinates are α_k and σ_k and is one if α_{k+1} and σ_{k+1} are replaced, with $P(\cdot \,|\, \boldsymbol{\tau}', I)$ and $P(\cdot \,|\, \boldsymbol{\tau}, I)$ the corresponding prior distributions.

4 Application to the Mexico City data

The model and algorithm are applied to three sets of data obtained from the Mexico City monitoring network (`http://www.aire.cdmx.gob.mx`). We consider ozone, PM_{10} and $PM_{2.5}$ measurements. In all cases we take the Mexico City metropolitan area overall daily maximum measurements (given in parts per million - ppm - in the case of ozone and in microgram per cubic meter - $\mu g/m^3$ - in the cases of PM_{10} and $PM_{2.5}$). We have considered measurements obtained from 01 January 1995 to 31 December 2019 in the cases of O_3 and PM_{10} and from 01 January 2004 to 31 December 2019 in the case of $PM_{2.5}$. That gives us 9131 ozone and PM_{10} observations and 5844 $PM_{2.5}$ measurements. (The shorter length of the $PM_{2.5}$ observational period is due to the fact that $PM_{2.5}$ started being systematically monitored only in mid-2003).

Minute by minute measurements are taken in each station of the monitoring network and the hourly averaged result is reported every hour. The ozone daily maximum measurement in a given station is the maximum of the hourly averaged results reported at that station. Ozone overall daily maximum measurement for the metropolitan area is the maximum of all daily maxima recorded in all stations of the monitoring network. In the cases of PM_{10} and $PM_{2.5}$, the daily measurements in a given station are the average results over the 24-hour period of the corresponding pollutant at that station. The PM_{10} and $PM_{2.5}$ overall daily maximum measurements for the metropolitan area are the maxima of the daily averaged results for all stations in the network. The thresholds considered are those specified by NOM (2014a, 2014b), i.e., 0.095 ppm for ozone, 75 $\mu g/m^3$ for PM_{10} and 45 $\mu g/m^3$ for $PM_{2.5}$. During the observed period there were 6706, 5211 and 1028 ozone, PM_{10} and $PM_{2.5}$ exceedances, respectively.

Different values of λ were considered as hyperparameters for the prior distribution of the number of change-points I. Hence, in the case of ozone we take $\lambda = 1, 1.5, 2$; we have $\lambda = 0.9, 1, 1.2$ in the case of PM_{10} and $\lambda = 0.9, 1.2, 1.5$ in the case of $PM_{2.5}$. In all cases we take $M = 8$. Hence, the state space of the number of change-points I is $\mathcal{S} = \{0, 1, 2, \ldots, 8\}$. When ozone data are considered, the hyperparameters of the prior distributions of the parameters α are $a_1 = 177$ and $a_2 = 1/180$ and, in the cases of the parameters σ, are $b_1 = 30$ and $b_2 = 1/32$. If we consider the PM_{10} data we have $a_1 = 3855.119$, $a_2 = 1/4037.43$, $b_1 = 60.885$ and

$b_2 = 1/56.40477$. In the case of $PM_{2.5}$ the hyperparameters are $a_1 = 989.0532$, $a_2 = 1/1199.82$, $b_1 = 19.78535$ and $b_2 = 1/16.15016$. In all cases we take $c = 0.5$.

Remark. The values of the hyperparameters in the case of ozone were obtained from results given in previous studies where subsets of the datasets considered here were used (Súarez-Sierra et al. 2019; Rodrigues et al. 2019; Achcar et al. 2011). In the cases of and PM_{10} and $PM_{2.5}$ we have used information provided by the model when one change-point was assumed.

4.1 Results

Estimation of the parameters were made using samples of size 40000 obtained from five chains after a burn-in period of 20000 iterations taking every 10th generated values. In Table 1, given in the Appendix A, we have the values of the BIC, ML and SAD for each model and pollutant, and in Table 2, also in the Appendix A, we have the distribution of the numbers of change-points, as well as their mean values estimated using the RJ-MCMC generated samples. Additionally, in Figures 1 and 2, given in the Appendix B, we have, respectively, the plots of the accumulated observed and estimated means and the plots of their absolute differences.

Looking at Table 1, we see that in the case of ozone, of the three criteria, two choose the model with $\lambda = 1$ with two change-points. The third criterion selects the model with $\lambda = 1.5$ with one change-point. If we consider the results when the PM_{10} data are used, two criteria choose the model with $\lambda = 0.9$ with eight change-points. The third criterion prefers the model with $\lambda = 1$ and eight change-points. In the case of $PM_{2.5}$ the model with $\lambda = 1.5$ is chosen by two of the criteria. One of them prefers the model with three change-points and the other with eight. The third criterion chooses the model with $\lambda = 0.9$ with eight change-points.

If wee look at the results given by Table 2, we see that the estimated means of the number of change-points are 3 and 4 for all values of λ in the cases of ozone and PM_{10}, respectively, and if we consider the $PM_{2.5}$ data, then we have that the estimated values are 4 in the cases of $\lambda = 0.9, 1.2$ and it could be either 4 or 5 for $\lambda = 1.5$. The values of the modes coincide with the estimated means in the cases of PM_{10} and $PM_{2.5}$ and all values of λ, with the case of $\lambda = 0.9$ and PM_{10} closely followed by the model with three change-points. In the case of ozone, the

estimated number of change-points using the mode is 2.

If we turn our attention to the plots in Figures 1 and 2, in the case of ozone (see top three plots), the model with two change-points (continuous blue lines) provides the best fit for all values of λ with the model with three change-points (continuous pink lines) and $\lambda = 1.5$ also giving a good fit. If we consider the case of PM_{10} (see plots in the second row of the figures), results are a little fuzzy. We may have that the model with either three (continuous pink lines) or eight (dotted blue lines) change-points is suitable for all values of λ with the exception of the case where $\lambda = 1$ in which the preferred model is the one with seven change-points (dashed green line). Similar disperse results are found when we look at the third rows of the figures where we have the plots when $PM_{2.5}$ data are used. In this case we have that almost all models have a good fit, with the exception of the case where $\lambda = 0.9$ in which case by looking at Figure 1, we have that the model with eight change-points (dotted blue line) does not provide a good overall fit and in Figure 2, we have that the smallest difference could be given by the model with either three (continuous pink lines) or four (dashed black lines).

Therefore, taking into to account the selection made using the different criteria and selecting those preferred by more criteria and using parsimony, we may choose the model with $\lambda = 1$ with two change-points in the case of ozone, the model with $\lambda = 0.9$ and three change-points when we consider the PM_{10} data and, in the case of $PM_{2.5}$, we could use the model with $\lambda = 1.5$ with four change-points. The estimated values of the parameters in each chosen model are given in Table 3 in the Appendix A.

Looking at Table 3 we see that the change-points were located in the years 2003 and 2009 in the case of ozone; one in 1995 and two in 2015 if we consider the PM_{10} measurements and one in 2004 and 2011 and two in 2019 in the case of $PM_{2.5}$. We see that the two last detected change-points in the cases of PM_{10} and $PM_{2.5}$ are very close to each other. We could not find a real reason to why these change-points are so close to each other since in those periods of time there were no major changes that would justify their proximity. Note that even though the values of α in all cases are very close to each other, with the exception of the $PM_{2.5}$ data when we move from before to after the second change-point, they decrease as we move through the change-points. That indicates that the speed at each exceedances occur decreases slightly from

change-point to change-point. In the case of $PM_{2.5}$ the values of α restart to go from larger to smaller once we go from before to after the third change-point. Hence, we return to the case of a slowing down in the speed of occurrence of $PM_{2.5}$ exceedances.

In order to see how the selected estimated rate functions behave, in Figure 3 given in the Appendix B, we have their plots for all pollutants. Looking at Figure 3 we see that with the exception of the plots associated with the $PM_{2.5}$ (bottom plot), in spite of having close values of α, when we move through the change-points, in the cases of ozone (top left plot) and PM_{10} (top right plot) we see that there are small but significant differences in the values of the rate functions in the time intervals between change-points.

If we consider the selected models using the mode of the distribution of the number of change-points, we see, by looking at Table 2, that the selected models are $\lambda = 1.5$ and two change-points (with $\tau_1 = 3259$ and $\tau_2 = 5312$) in the case of ozone, $\lambda = 0.9$ and four change-points in the cases of PM_{10} (with $\tau_1 = 102$, $\tau_2 = 4188$, $\tau_3 = 7729$ and $\tau_4 = 7790$) and $PM_{2.5}$ (with $\tau_1 = 103$, $\tau_2 = 2884$, $\tau_3 = 5670$ and $\tau_4 = 5775$). In these cases the locations of the change-points were in the years 2002 and 2009 in the case of ozone; one in 1995 and 2006 and two in 2016 in the case of PM_{10}; and one in 2004 and 2011 and two in 2019 if we consider the $PM_{2.5}$ data. Note that, with the exception of the case where the PM_{10} data are used in which case an extra change-points located in the year 2006 is detected, the locations of the change-points do not differ much from the ones obtained using the criteria considered here to select the suitable model.

5 Discussion

When we compare the results obtained using the RJ-MCMC algorithm with the results produced by previous works where an empirical Bayes approach was used, we have the following. In Súarez-Sierra et al. (2019, 2022) where we have an observational period ranging from 01 January 2004 to 31 August 2015 for all three pollutants, selected models were the model with zero, one and three change-points when the ozone, PM_{10} and $PM_{2.5}$ data were used, respectively. The change-points locations were in the year 2011 in the case of PM_{10} and in the years

2005, 2008 and 2013 if we consider the $PM_{2.5}$ data.

In the case of PM_{10} the RJ-MCMC algorithm detects two change-points (which could be considered as only one due to their proximity and the lack of reasons for having a change in the behaviour of the data) at the end of the observational period considered in Súarez-Sierra et al. (2019), but it misses the one located in the year 2011. The other estimated change-points given by the present algorithm are located outside that observational period.

When we consider the $PM_{2.5}$ data the number of change-points differs when the empirical Bayes approach and the RJ-MCMC algorithm are used. However, since the two last change-points estimated by the RJ-MCMC algorithm are very close to each other, they could be considered as just one. Even though their locations vary they are not that different when we use both approaches. In the case of the estimated parameters α of the rate functions the RJ-MCMC algorithm does not detect the increasing behaviour between the second and third change-points as obtained in Súarez-Sierra et al. (2022).

When we compare the results related to the ozone data, we have that in Súarez-Sierra et al. (2019) the empirical Bayes approach did not detect any change-points during the time span ranging from 01 January 2004 to 31 August 2015. However, the RJ-MCMC algorithm has detected two change-points during that time span.

In another time frame, consider the results given by Rodrigues et al. (2019) where ozone data from 01 January 1995 to 31 December 2010 were used. In that work a spatial component was considered and data from ten stations were taken into account. Estimation of the parameters were also made using an empirical Bayes approach. In that work, two change-points were detected for the data of all stations. They were located in the years 2001 and 2007 which differ by only two years from those obtained using the RJ-MCMC algorithm.

Consider now the results given in Achcar et al. (2011). In that work ozone data obtained from 01 January 1990 to 31 December 2005 were used. The selected model detected the presence of one change-point whose location is in the year 2001. This coincides with the first of the estimated change-points in Rodrigues et al. (2019) and it is not far away from the first change-point obtained using the present algorithm.

6 Conclusion

In the present work we have considered a reversible-jump MCMC algorithm to estimate the parameters in a non-homogeneous Poisson model in the presence of change-points, as well as their number and locations. The model and algorithm were applied to ozone, PM_{10} and $PM_{2.5}$ data from Mexico City. Several criteria were used to select the model that best fit each dataset. Results were mixed. Models with two, three and four change-points were detected when the ozone, PM_{10} and $PM_{2.5}$ data were used, respectively. Note that even though some of the selection criteria have chosen the model with eight change-points, looking at Table 2 (Appendix A) we see that this model was visited by the algorithm only a small proportion of the iterations. Hence, the selection of this model by some of the criteria might be associated with the large number of parameters.

Even though in all the previous works and in the present, the number and locations of the change-points are more or less compatible, the difference lies in the values of the estimated parameters of the rate functions. In a more localised time span, changes in the behaviour of the rate function are detected, i.e., in some cases we have changes from increasing to decreasing and vice-versa. However, when the global picture is used some subtleties are not reflected in the results.

We would like to point out that in 2021 new regulations were published considering tighter threshold for ozone, PM_{10} and $PM_{2.5}$ (NOM 2021a, 2021b). In those regulations the threshold for ozone changed to 0.09 ppm when the one hour average measurements are used. In the cases of the 24-hour PM_{10} and $PM_{2.5}$ averages, a stead decrease in the threshold values will be applied until the stricter values established by the World Heath Organisation 2005 guidelines for the year 2025 are reached.

Acknowledgements

Part of this work is from JACJ's Ph.D. Thesis developed at the Benemérita Universidad Autónoma de Puebla, Puebla, México. JACJ, ERR and HJRC thank CONAHCyT-Mexico.

This work was partially developed while ERR was in an academic visit at the Department of Statistics of the Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP) – Campus Presidente Prudente, Brazil. ERR is grateful to the Department of Statistics of UNESP, for support and hospitality during the development of this work.

References

[1] J. A. Achcar, A. A. Fernández-Bremauntz, E. R. Rodrigues, and G. Tzintzun. Estimating the number of ozone peaks in Mexico City using a non-homogeneous Poisson model. *Environmetrics*, 19:469-485, 2008.

[2] J. A. Achcar, E. R. Rodrigues, and G. Tzintzun. Using non-homogeneous Poisson models with multiple change-points to estimate the number of ozone exceedances in Mexico City. *Environmetrics*, 22:1-1, 2011. http://dx.doi.org/10.1002/env.1029

[3] L. J. Álvarez, N. L. Garcia, and E. R. Rodrigues. Comparing the performance of a reversible jump Markov chain Monte Carlo algorithm for DNA sequences alignment. *Journal of Statistical Computation and Simulation*, 76:567-584, 2006.

[4] J. Ai. *Reversible-jump MCMC methods in Bayesian statistics*. Master Dissertation in Statistics. University of Leeds. UK, 2012

[5] H. Akaike. On the entropy maximization principle. In *Proceedings of the Symposium on Applications of Statistics*, P. R. Krishaiah (ed). North-Holland, Amsterdan, pages 27-47, 1977

[6] H. Akaike. On newer statistical approaches to parameter estimation and structure determination. *International Federation of Automatic Control*, 3:1877-1884, 1978

[7] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *A first course in order statistics*. SIAM Classics in Applied Mathematics 54. Society for Industrial and Applied Mathematics. Philadelphia, USA, 2008

[8] M. L. Bell, A. McDermott, S. L. Zeger, J. M. Samet, and F. Dominici. Ozone and short-term mortality in 95 US urban communities, 1987-2000. *Journal of the American Medical Society* 292:2372-2378, 2004. DOI:10.1001/jama.292.19.2372

[9] B.P. Carlin and S. Chib. Bayesian model choice via Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society Series B*, 57:473–484, 1995.

[10] D. R. Cox and P. A. W. Lewis. *Statistical analysis of series events*. Chapman and Hall, U.K. 1966.

[11] E.P.A. US Environmental Protection Agency. Available at `www.epa.gov/pm-pollution/health-and-environmental-effects-particulate-matter-pm`. (Assessed on 06 June 2018).

[12] S. Feng, D. Gaoa, F. Liao, F. Zhou, and X. Wang. The health effects of ambient $PM_{2.5}$ and potential mechanisms. *Ecotoxicology and Environmental Safety Review*, 128:67–74, 2016.

[13] D. Gamerman and H. F. Lopes. *Markov Chain Monte Carlo Stochastic Simulation for Bayesian Inference*. Second Edition. Chapman and Hall/CRC, USA, 2006.

[14] P. J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82:711-732, 1995.

[15] J. Gyarmati-Szabó, L. V. Bogachev, and H. Chen. Modelling threshold exceedances of air pollution concentrations via non-homogeneous Poisson process with multiple change-points. *Atmospheric Environment*, 45:5493-5503, 2011.

[16] K. Itô and G. D. Thurston. Daily PM10/mortality associations: an investigation of at-risk subpopulations. *Journal of Exposure Analysis and Environmental Epidemiology*, 6:79–95, 1996.

[17] N. A. H. Janssen, P. Fischer, M. Marra, C. Ameling, and F. R. Cassee. Short-term effects of $PM_{2.5}$, PM_{10} and $PM_{2.5-10}$ on daily mortality in the Netherlands. *Sci. Total Environ.*, 463-464:20-26, 2013. DOI: 10.1016/j.scitotenv.2013.05.062

[18] F. Kelly. Oxidative stress: its role in air pollution and adverse health effects. *Occupational and Environmental Medicine*, 60:612-616, 2003. DOI: 10.1136/oem.60.8.612

[19] J. F. Lawless. *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons. NewYork,1982.

[20] G. Likens (Lead Author). Environmental Protection Agency (Content source). (Topic eds., W. Davis, L. Zaikowski and S. C. Nodvin). *Acid rain*. Encyclopedia of Earth (ed., J. Cutler). 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), 2010 (available from `http://www.eoearth.org/article/Acidrain`).

[21] D. Loomis, V. H. Borja-Arbuto, S. I. Bangdiwala, and C. M. Shy. Ozone exposure and daily mortality in Mexico City: a time series analysis. *Health Effects Institute Research Report*, 75:1-46, 1996.

[22] J. Mauderly and G. Oberdörster. Current understanding of the health effects of particles and characteristics that determined dose effects. In *Formation and characterization of particles. Report of the 1996 Health Effects Institute Workshop*, pages 1-5, 1997.

[23] N. O. M. Norma Oficial Mexicana NOM-020-SSA1-2014. *Diario Oficial de la Federación.* 19 August 2014. Mexico, 2014a. (In Spanish.)

[24] N. O. M . Norma Oficial Mexicana NOM-025-SSA1-2014. *Diario Oficial de la Federación.* 20 August 2014. Mexico, 2014b. (In Spanish.)

[25] N. O. M. Norma Oficial Mexicana NOM-020-SSA1-2021. *Diario Oficial de la Federación.* 28 October 2021. Mexico, 2021a. (In Spanish.)

[26] N. O. M . Norma Oficial Mexicana NOM-025-SSA1-2021. *Diario Oficial de la Federación.* 27 October 2021. Mexico, 2021b. (In Spanish.)

[27] E. R. Rodrigues, J. A. Achcar, and J. Jara-Ettinger. A Gibbs sampling algorithm to estimate the occurrence of ozone exceedances in Mexico City. In *Air Quality - Models and Applications*, D. Popovic (ed.), pages 131-150. InTech Open Access Publisher, Croatia, 2011.

[28] E. R. Rodrigues, G. Nicholls, M. H.Tarumoto, and G. Tzintzun. Using a non-homogeneous Poisson model with spatial anisotropy and change-points to study air pollution data. *Environmental and Ecological Statistics*, 26:53-184, 2019. https://doi.org/10.1007/s10651-019-00423-6

[29] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6:461-464, 1978.

[30] B. M. Súarez-Sierra, E. R. Rodrigues, and G. Tzintzun. Rate of ozone and PM_{10} exceedances: A case study comparing data from Mexico City and Bogota. *Communications in Statistics: Case Studies, Data Analysis and Applications*, 5:153-165, 2019. https://doi.org/10.1080/23737484.2019.1593066

[31] B. M. Súarez-Sierra, E. R. Rodrigues, and G Tzintzun. An application of a non-homogeneous Poisson model to study $PM_{2.5}$ exceedances in Mexico City and Bogota. *Journal of Applied Statistics*, 49:2430-2445, 2022. doi: 10.1080/02664763.2021.1897972

[32] G. D. Thurston. A critical review of PM_{10} mortality time series studies. *Journal of Exposure Analysis and Environmental Epidemiology*, 6:3–21, 1996.

[33] W. H. O. *Air Quality Guidelines-2005, Particulate Matter, Ozone, Nitrogen dioxide and Sulphur Dioxide*. European Union: World Health Organization Regional Office for Europe.

[34] Y.-F. Xing, Y.-H. Xu, M.-H. Shi, and Y.-X. Lian. The impact of $PM_{2.5}$ on the human respiratory system. *Thorac. Dis.*, 8:E69–E74, 2016.

[35] T. E. Yang and L. Kuo. Bayesian binary segmentation procedure for a Poisson process with multiple change-points. *Journal of Computational and Graphical Statistics*, 10:772-785 2001.

Appendix

In this appendix we present the tables with the results of the estimated values of the discrimination criteria and estimated parameters of the selected model for each dataset. We also give the plots related to the mean and rate functions of the selected models.

A. Tables

In this section we present the tables with the values of the BIC, ML and SAD for all models and datasets. We also give the table with the estimated distribution probability of the number of change-points, as well as the estimated mean obtained using the RJ-MCMC algorithm generated samples for all models and datasets. Finally, we present the table with the estimated parameters for the selected model when each dataset is considered.

BIC

	# chpt	0	1	2	3	4	5	6	7	8
ozone	$\lambda = 1$	–	17393.183	17416.943	17673.64	17772.99	17829.78	17791.54	17785.11	17923.25
	$\lambda = 1.5$	–	17383.92	17417.12	17473.531	17487.167	17498.677	17513.02	17500.94	17644.48
	$\lambda = 2$	–	17388.827	17416.996	17483.714	17613.43	174688.01	17770.97	17762.42	17776.49
PM_{10}	$\lambda = 0.9$	–	–	–	16200.346	16245.071	16266.493	16272.748	16231.302	16195.843
	$\lambda = 1$	–	–	–	16238.131	16270.708	16267.419	16271.339	16226.55	16242.217
	$\lambda = 1.2$	–	–	–	16264.76	16291.621	16301.804	16303.197	16334.874	16326.742
$PM_{2.5}$	$\lambda = 0.9$	–	–	–	303.013	335.025	350.677	348.923	332.114	300.662
	$\lambda = 1.2$	–	–	–	299.73	335.224	356.247	362.856	353.967	325.518
	$\lambda = 1.5$	–	–	–	299.004	334.572	355.722	362.586	352.154	328.976

ML

	# chpt	0	1	2	3	4	5	6	7	8
ozone	$\lambda = 1$	–	-8674.633	-8665.121	-10790.66	-10901.42	-11133.16	-10980.17	-11208.99	-10602.75
	$\lambda = 1.5$	–	-8672.769	-8665.373	-9521.734	-9773.483	-9849.739	-10113.93	-10756.78	-10007.65
	$\lambda = 2$	–	-8671.446	-8665.942	-9779.279	-10454.27	-10745.77	-11011.23	-10986.37	-10926.77
PM_{10}	$\lambda = 0.9$	–	–	–	-8078.125	-8094.161	-8110.535	-8121.468	-8079.054	-8068.361
	$\lambda = 1$	–	–	–	-8114.733	-8117.945	-8125.061	-8100.534	-8063.668	-7913.238
	$\lambda = 1.2$	–	–	–	-8102.293	-8122.191	-8130.786	-8124.669	-8151.185	-8138.848
$PM_{2.5}$	$\lambda = 0.9$	–	–	–	-95.145	-95.154	-95.133	-95.136	-95.19	-95.1
	$\lambda = 1.2$	–	–	–	-95.743	-95.704	-95.712	-95.807	-95.799	-95.64
	$\lambda = 1.5$	–	–	–	-95.305	-95.284	-95.279	-95.278	-95.261	-95.232

SAD

	# chpt	0	1	2	3	4	5	6	7	8
ozone	$\lambda = 1$	–	477853.314	116659.692	1991963.14	2688574.14	2697940.37	2499682.38	2315276.03	3123958.78
	$\lambda = 1.5$	–	291422.357	180443.688	161372.846	539478.189	342774.744	1504312.76	1719419.85	1985461.34
	$\lambda = 2$	–	321161.73	218057.004	622637.379	1504676.23	1761124.98	2793844.75	292248453	2154184.67
PM_{10}	$\lambda = 0.9$	–	–	–	717416.729	790202.274	788668.172	789823.963	768273.689	668077.692
	$\lambda = 1$	–	–	–	812272.138	819953.318	774984.935	756757.268	682515.924	917162.2
	$\lambda = 1.2$	–	–	–	808309.888	809742.14	805817.559	788002.613	831199.358	784061.76
$PM_{2.5}$	$\lambda = 0.9$	–	–	–	25918.111	24637.758	24499.217	25027.873	27011.622	23455.583
	$\lambda = 1.2$	–	–	–	26510.282	25049.377	24449.872	23738.459	25210.661	29402.703
	$\lambda = 1.5$	–	–	–	23769.717	24479.68	23267.907	23923.868	24756.449	20572.883

Table 1: BIC, ML and SAD values for each model and pollutant. We use "–" to indicate the fact that the model with the corresponding number of change-points either was not visited by the algorithm or was visited an insignificant number of times.

	chpt	0	1	2	3	4	5	6	7	8	Mean
Ozone	$\lambda = 1$	–	0.057	0.5288	0.1762	0.1562	0.0651	0.0141	0.0023	0.0003	3
	$\lambda = 1.5$	–	0.0334	0.5851	0.1332	0.1612	0.0671	0.0165	0.0029	0.0006	3
	$\lambda = 2$	–	0.0363	0.5423	0.1451	0.1334	0.0902	0.0392	0.0112	0.0023	3
PM_{10}	$\lambda = 0.9$	–	–	–	0.379	0.384	0.171	0.053	0.011	0.002	4
	$\lambda = 1$	–	–	–	0.282	0.382	0.222	0.082	0.026	0.006	4
	$\lambda = 1.2$	–	–	–	0.226	0.374	0.255	0.108	0.031	0.006	4
$PM_{2.5}$	$\lambda = 0.9$	–	–	–	0.328	0.415	0.196	0.05	0.01	0.001	4
	$\lambda = 1.2$	–	–	–	0.229	0.397	0.255	0.092	0.023	0.004	4
	$\lambda = 1.5$	–	–	–	0.172	0.364	0.289	0.129	0.037	0.009	4-5

Table 2: Distribution and mean of the number of change-points for each model and dataset estimated using the values generated by the RJ-MCMC algorithm. We use "–" to indicate that the corresponding value either was not part of that specific model or the value was of order smaller that 10^{-3}.

	Ozone		PM_{10}		$PM_{2.5}$	
	Mean (SD)	MC Error	Mean (SD)	MC Error	Mean (SD)	MC Error
α_1	0.976 (1.4E-2)	1.16E-3	0.97 (1.4E-2)	5.5E-3	0.821 (2.1E-2)	3.4E-3
α_2	0.9598 (1.8E-2)	1.57E-3	0.955 (1.5E-2)	2.4E-4	0.824 (2.6E-2)	2.7E-4
α_3	0.943 (1.6E-2)	1.43E-3	0.953 (1.5E-2)	1.8E-3	0.825 (2.6E-2)	2.7E-4
α_4	–	–	0.933 (1.4E-2)	3.4E-3	0.824 (2.5E-2)	8.9E-4
α_5	–	–	–	–	0.806 (2E-2)	1.2E-3
σ_1	0.9199 (0.101)	8.38E-3	0.999 (8.7E-2)	3.2E-2	1.236 (0.192)	3.2E-2
σ_2	0.9297 (0.164)	1.51E-2	1.076 (0.138)	2.1E-3	1.221 (0.275)	2.8E-3
σ_3	0.967 (0.166)	1.53E-2	1.133 (0.135)	1.7E-2	1.225 (0.273)	2.8E-3
σ_4	–	–	1.221 (0.124)	2.1E-2	1.224 (0.285)	1E-2
σ_5	–	–	–	–	1.335 (0.252)	1.8E-2
τ_1	3241 (315.7)	20.59	102 (1.534)	0.545	103 (2.64)	0.328
τ_2	5289 (241.9)	19.38	7439 (1127.42)	448.23	2880 (1258.87)	11.21
τ_3	–	–	7462 (1144.29)	458.08	5665 (59.41)	4.84
τ_4	–	–	–	–	5771 (50.07)	5.57

Table 3: Estimated means, standard deviations (indicated by SD) and MC Error of the parameters for each selected model and dataset. We use "–" to indicate that the corresponding parameter was not part of that specific model.

B. Figures

In the this appendix we present the plots of the accumulated observed and estimated means when all pollutants, hyperparameters λ and number of change-points present in each version of the non-homogeneous Poisson model are used. We also present the plots of their absolute differences which indicate the places of the best fit to the observed means of each estimated means. Additionally, the plots of the rate functions of the selected models are also given.

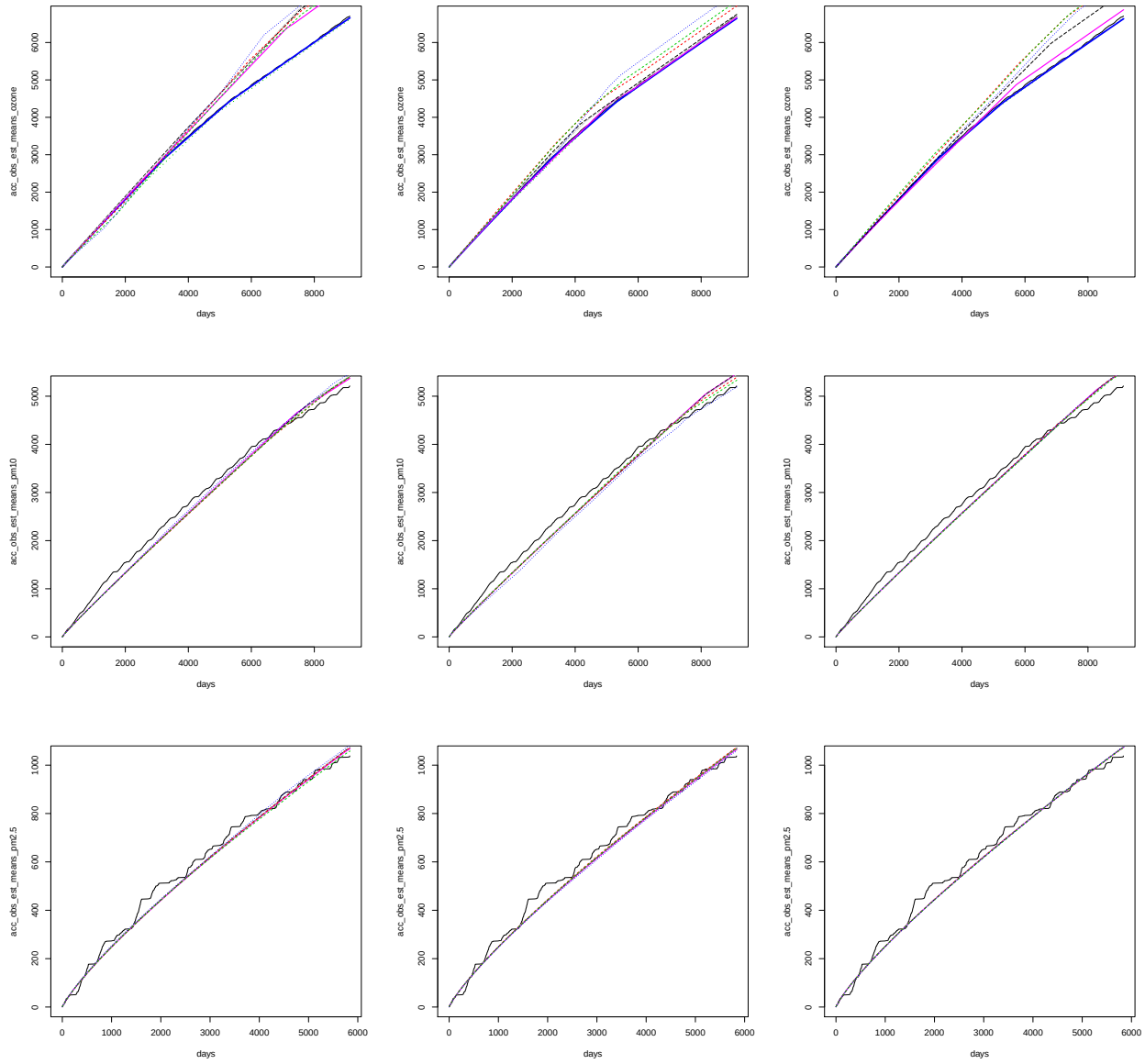

Figure 1: Accumulated observed and estimated means in the case of ozone with $\lambda = 1, 1.5, 2$ (from left to right, first row plots), PM_{10} with $\lambda = 0.9, 1, 1.2$ (from left to right, second row plots) and $PM_{2.5}$ with $\lambda = 0.9, 1.2, 1.5$ (from left to right, third row plots). Black continuous lines are the observed means, green dotted light are the cases where one change-point is allowed, blue continuous lines represent the cases with two change-points, pink continuous lines represent the cases where we have three change-points, black dashed lines are the cases where we have four change-points, dashed grey lines are the five change-points models, dashed red lines represent the cases with six change-points, dashed green lines are the cases where we have seven change-points and the dotted blue lines represent the cases with eight change-points.

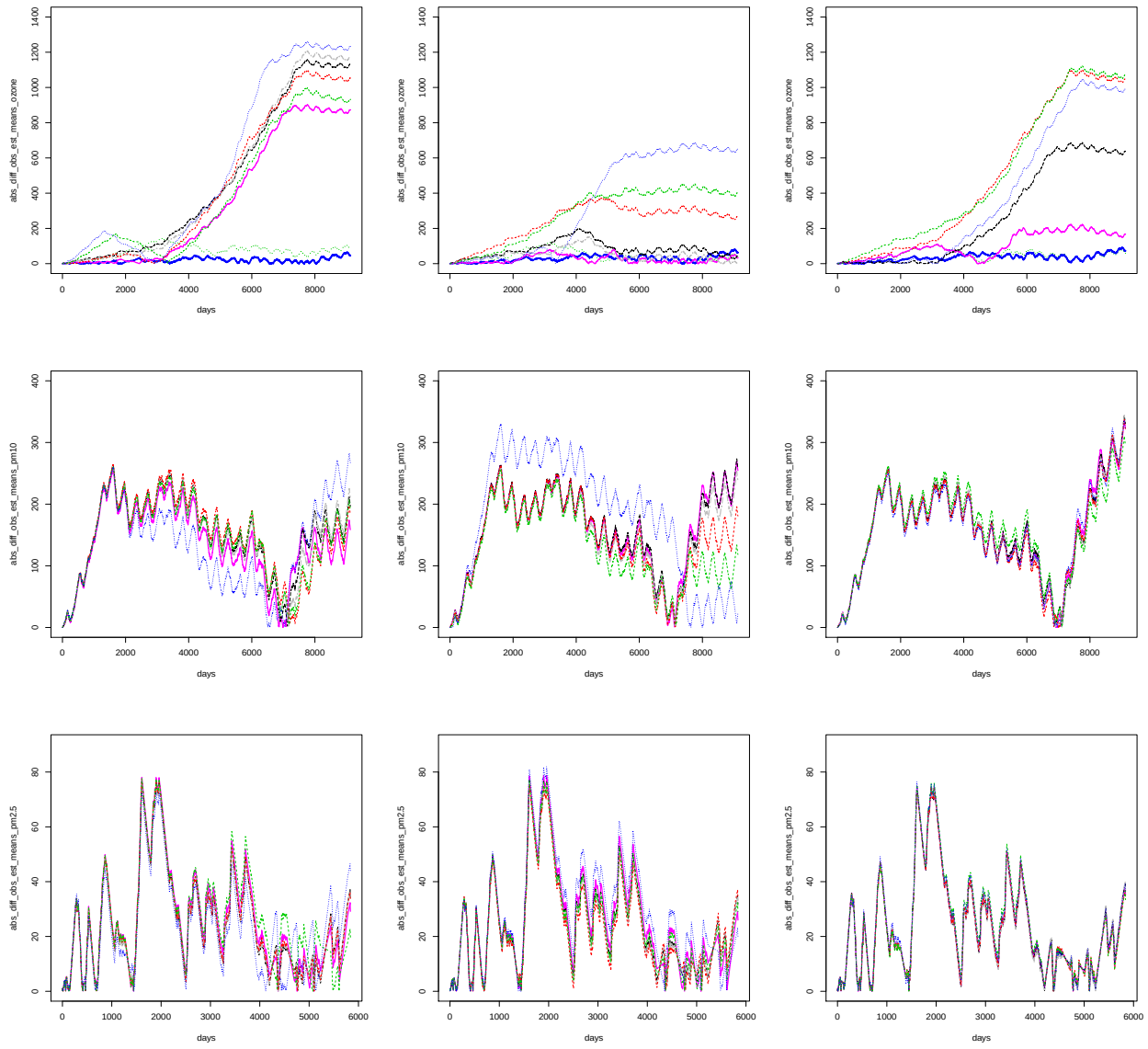

Figure 2: Absolute differences between the accumulated observed and estimated means in the case of ozone with $\lambda = 1, 1.5, 2$ (from left to right, first row plots), PM_{10} with $\lambda = 0.9, 1, 1.2$ (from left to right, second row plots) and $PM_{2.5}$ with $\lambda = 0.9, 1.2, 1.5$ (from left to right, third plots). Black continuous lines are the observed means, green dotted light are the cases where one change-point is allowed, blue continuous lines represent the cases with two change-points, pink continuous lines represent the cases where we have three change-points, black dashed lines are the cases where we have four change-points, dashed grey lines are the five change-points models, dashed red lines represent the cases with six change-points, dashed green lines are the cases where we have seven change-points and the dotted blue lines represent the cases with eight change-points.

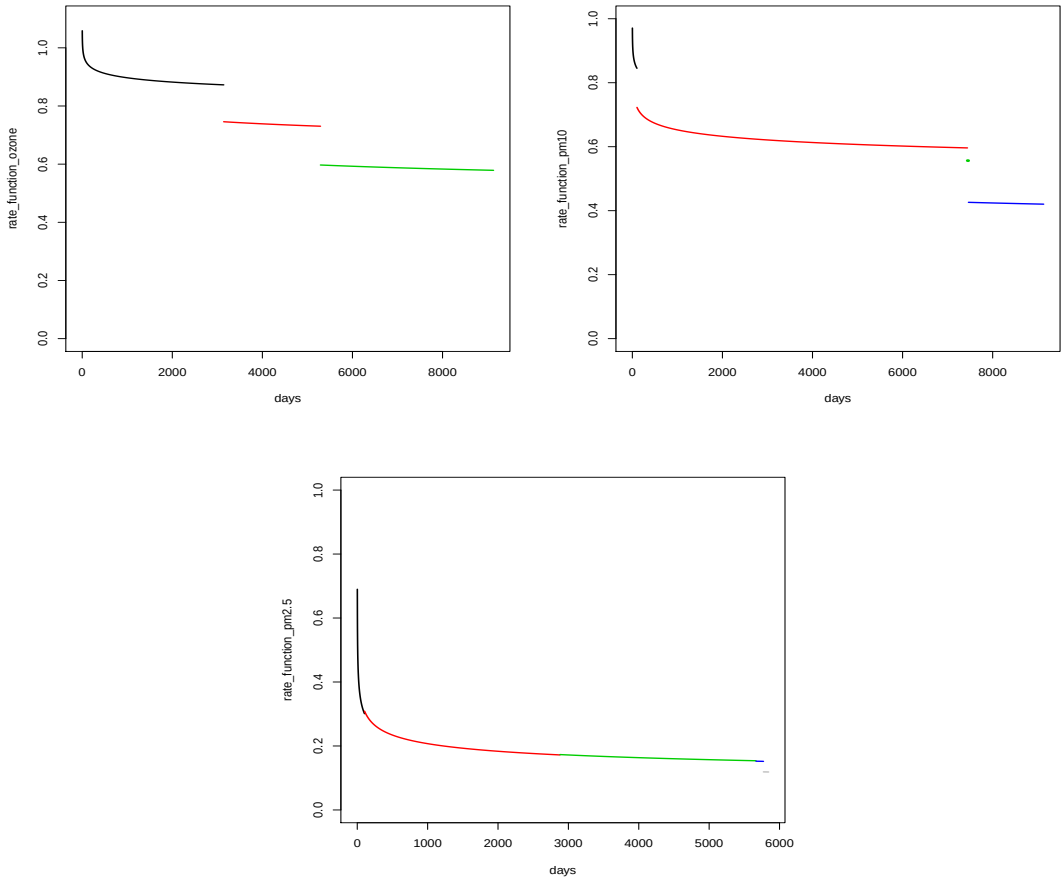

Figure 3: Rate functions of the selected model for each pollutant. Black, red, green, blue and grey lines correspond to the rate functions before de first change-point, between the first and second change-points, between the second and third change-points, between the third and fourth change-points and between the fourth and fifth change-points.