

Article

Detection of Changes in Ground-Level Ozone Concentrations via Entropy

Yuehua Wu ^{1,*}, Baisuo Jin ² and Elton Chan ³

¹ Department Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada

² Department Statistics and Finance, University of Science and Technology of China, Hefei, Anhui, China; E-Mail: jbs@ustc.edu.cn

³ Air Quality Research Division, Science and Technology Branch, Environment Canada, Toronto, Ontario, M3H 5T4, Canada; E-Mail: elton.chan@ec.gc.ca

* Author to whom correspondence should be addressed; E-Mail: wuyh@mathstat.yorku.ca;
Tel: +1-416-7365250; Fax: +1-416-7365757.

Academic Editor: Kevin H. Knuth

Received: 4 March 2015 / Accepted: 28 April 2015 / Published: 30 April 2015

Abstract: Ground-level ozone concentration is a key indicator of air quality. There may exist sudden changes in ozone concentration data over a long time horizon, which may be caused by the implementation of government regulations and policies, such as establishing exhaust emission limits for on-road vehicles. To monitor and assess the efficacy of these policies, we propose a methodology for detecting changes in ground-level ozone concentrations, which consists of three major steps: data transformation, simultaneous autoregressive modelling and change-point detection on the estimated entropy. To show the effectiveness of the proposed methodology, the methodology is applied to detect changes in ground-level ozone concentration data collected in the Toronto region of Canada between June and September for the years from 1988 to 2009. The proposed methodology is also applicable to other climate data.

Keywords: change-point detection; Box–Cox transformation; entropy; ozone concentration; spatial dependence; simultaneous autoregressive modelling

1. Introduction

Air quality has attracted more attention in the past 50 years. Climate change itself may have a direct impact on air quality. Air quality change may also be caused by human activities. The quantitative change of air quality includes its mean change, variance change, quantile change, correlation change, and so on.

Ground-level ozone concentration is a key indicator of air quality. Exposure to high levels of ozone can cause problems for people with respiratory and heart problems and agricultural crop loss. For this reason, specialists, in conjunction with public institutions, have been carrying out investigations in areas related to ozone and health [1]. Several statistical methodologies have been applied to model the ground-level ozone concentration data, which include multivariate models [2,3], quantile regression [4,5], non-linear time series [6–8] and hierarchical Bayesian kriging [9,10]. However, most of these approaches assume the temporal homogeneity of the stochastic processes involved, which may not hold over longer time horizons.

Ground-level ozone results from photochemical reactions between oxides of nitrogen and volatile organic compounds in the presence of sunlight. In many countries, the transportation sector is now the single largest source of ground-level ozone concentration. Regulations establishing limits for gaseous and particulate compounds emitted by on-road vehicles were promulgated by different countries. In order to monitor and assess the efficacy of these and future policies, it is important to develop adequate statistical methods to measure the impact of the regulations on the dynamics of various pollutants, especially with regard to the set standards [11].

To address this issue, some authors have modelled the exceedances of air pollution concentrations using non-homogeneous Poisson processes [11–13]. However a non-homogeneous Poisson process is only a point process, which does not include the spatial correlation between different areas. In contrast, entropy can be used to measure the various spatial uncertainties, which include the uncertainties in both spatial variance and spatial dependence. Thus, in this paper, we consider using entropy to investigate the spatial properties of the index of air quality. We remark that entropy has been used to predict ozone observations, e.g., [14], or to design national air pollution monitoring networks in Fuentes [9,15], among others.

It is noted that functional data analysis and control charts have been proposed to detect outliers in gas emissions in the literature (e.g., [16,17]). These methods can be used to monitor abnormal air quality due to a short-term climate change or unusual human activity. However, they are not appropriate for studying the long-term effect of air quality change caused by some policies and regulations of environmental agencies, because of their ability to find abnormalities.

The article is arranged as follows: Section 2 presents the methodology for the detection of changes in ground-level ozone concentrations via entropy. The proposed methodology is applied to a real data in Section 3. The discussion is given in Section 4.

2. The Methodology

Let $X_{i,t}$, $i = 1, \dots, N$; $t = 1, \dots, T$, be the ozone concentration data collected in T days from N monitoring stations. In general, $X_{i,t}$ are not normally distributed or even approximately normally

distributed. To tackle this problem, we can first transform the data by applying the Box–Cox power transformation, with the parameter λ :

$$Z_{i,t}(\lambda) = \begin{cases} \frac{X_{i,t}^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log(X_{i,t}), & \lambda = 0. \end{cases} \tag{1}$$

How to choose λ will be given later.

In order to account for the periodicity and temporal autocorrelation in $Z_{i,t}(\lambda)$, $t = 1, \dots, T$, for each fixed i , it is assumed that $Z_{i,t}(\lambda)$, $t = 1, \dots, T$, is an autoregressive time series with period $2L$. Thus, to model the data, we employ the Fourier series expansion to reflect its periodic properties, while using the autoregressive formulation to describe its autocorrelation structure as follows:

$$Z_{i,t}(\lambda) = a_{i,0}(\lambda) + \sum_{j=1}^p \left[a_{i,j}(\lambda) \cos\left(\frac{j\pi}{L}t\right) + b_{i,j}(\lambda) \sin\left(\frac{j\pi}{L}t\right) \right] + \sum_{k=1}^q c_{i,k}(\lambda) Z_{i,t-k}(\lambda) + \varepsilon_{i,t}(\lambda), \tag{2}$$

where $a_{i,0}(\lambda)$, $a_{i,j}(\lambda)$, $b_{i,j}(\lambda)$, $j = 1, \dots, p$, $c_{i,k}(\lambda)$, $k = 1, \dots, q$ are unknown regression coefficients, p is the order of the truncated Fourier series, q is the lag order of the autoregressive representation and $\varepsilon_{i,t}(\lambda)$, $t = 1, \dots, N$, are random errors.

The problem remains how to model $\{\varepsilon_{i,t}(\lambda)\}$, which should be allowed to vary in space and time. To tackle this problem, we can borrow the strength of a simultaneous autoregressive (SAR) model, which is often used in spatial statistics for modelling the spatial correlation of quantities of interest in a region and the regression relation between quantities of interest and explanatory variables. The parameter estimation for a SAR model can be given by employing the maximum likelihood method [18] or a Bayesian method [19]. Put $\varepsilon_{.t} = (\varepsilon_{1,t}, \dots, \varepsilon_{N,t})'$. We model $\{\varepsilon_{.t}\}$ by the following SAR model:

$$(I_N - \rho_t W)\varepsilon_{.t} = \epsilon_t, \tag{3}$$

where I_N is an $N \times N$ identity matrix, $\{\rho_t\}$ are spatial parameters, W is a weight matrix and $\epsilon_t = (\epsilon_{1,t}, \dots, \epsilon_{N,t})'$ are independently normally distributed random errors with zero means and diagonal covariance matrix $\sigma_t^2 I_N$. Thus, the density function of $\varepsilon_{.t}$ is:

$$f(\varepsilon_{.t}) = |2\pi\Sigma_t|^{-1/2} \exp\left\{-\frac{1}{2}\varepsilon_{.t}'\Sigma_t^{-1}\varepsilon_{.t}\right\},$$

where $\Sigma_t = \sigma_t^2 [(I_N - \rho_t W)^{-1}]' (I_N - \rho_t W)^{-1}$. Following Ahmed and Gokhale (1989) [20], the differential entropy of the multivariate normal distribution is:

$$h_t = -E\{\log[f(\varepsilon_{.t})]\} = \frac{1}{2} \log[(2\pi e)^n |\Sigma_t|]. \tag{4}$$

There may exist sudden changes in ozone concentration data over a long time horizon, which may be caused by the implementation of government regulations and policies, such as establishing exhaust emission limits for on-road vehicles. To monitor and assess the efficacy of these policies, there is a need to detect changes in ground-level ozone concentrations, which can be fulfilled by detecting sudden

changes in the time sequence $\{h_t\}$. Denote the number of sudden changes by g and denote these g change-points by k_1^*, \dots, k_g^* , such that $1 < k_1^* < k_2^* < \dots < k_g^* < T$. Thus, h_t can be expressed as:

$$h_t = \theta_0 + \sum_{l=1}^g \theta_l I_{(k_l^*, \infty)}(t), \tag{5}$$

where $I_A(t)$ is an indicator function of the set A , *i.e.*,

$$I_A(t) = \begin{cases} 1, & \text{if } t \in A, \\ 0, & \text{if } t \notin A, \end{cases}$$

and $\theta_l \neq 0$ for $l = 1, \dots, g$. The aim of this paper is to estimate g and k_1^*, \dots, k_g^* , which can be done by the method given in [21]. Let $m = \lfloor \sqrt{T} \rfloor$ and $p = \lfloor T/m \rfloor$, where $\lfloor c \rfloor$ denotes the largest integer less than or equal to c . Denote $\theta = (\theta_1, \dots, \theta_p)'$. By Jin, Shi and Wu (2013) [21], the estimate of θ is given by:

$$\hat{\theta} = \arg \min_{\theta} \left\{ \sum_{t=1}^T \left(h_t - \sum_{j=0}^{\lfloor t/m \rfloor} \theta_j \right)^2 / T + \sum_{j=0}^p p_{\lambda_T, \gamma_T}(|\theta_j|) \right\}, \tag{6}$$

where $\lambda_T > 0, \gamma_T > 0$ are chosen by the Bayesian information criterion (BIC), and the penalty function $p_{\lambda_T, \gamma_T}(|u|)$ satisfies the following assumption:

$$p_{\lambda, \gamma}(u) = \lambda u - \frac{u^2}{2\gamma} I_{[0, \gamma\lambda]}(u) + \frac{1}{2} \gamma \lambda^2 I_{(\gamma\lambda, \infty)}(u).$$

If $\hat{\theta}_j \neq 0$, we test if there is a change-point in $[T - (p - j + 2)m + 1, T - (p - j - 1)m]$ by the method of cumulative sum of squares. Let $\hat{k} = \arg \min_{T - (p - j + 2)m + 1 \leq k \leq T - (p - j - 1)m} Q_k$, where:

$$Q_k = \sum_{t=T - (p - j + 2)m + 1}^k \left(h_t - \frac{1}{k - (j - 1)m + 1} \sum_{i=T - (p - j + 2)m + 1}^k h_i \right)^2 + \sum_{t=k+1}^{T - (p - j - 1)m} \left(h_t - \frac{1}{(j + 2)m - k} \sum_{i=k+1}^{T - (p - j - 1)m} h_i \right)^2.$$

Let $b = (2 \log(\log(3m)) + \log(\log(\log(3m))))^2 / (2 \log(\log(3m)))$, $a = \sqrt{b / (2 \log(\log(3m)))}$ and $D = 3m(Q_{\hat{k}} - Q_{T - (p - j - 1)m}) / Q_{T - (p - j - 1)m}$. By Theorem 3.1.1 in [22], we have:

$$\lim_{T \rightarrow \infty} P((D - b)/a \leq x) = \exp(-2e^{-x/2}).$$

Thus, if $(D - b)/a \geq 2 \log(-2 / \log(0.95))$, it is claimed that there is a change-point located in $[T - (p - j + 2)m + 1, T - (p - j - 1)m]$, and \hat{k} is its estimate that is significant at the 5% level. Otherwise, there is no change-point in this interval.

The detailed implementation of the proposed methodology above consists of the following four steps.

- Step 1.** Select all of the stations, such that at least one pair of ozone concentration observations from any two of these stations is not missing.
- Step 2.** For the data from each station, do the following: Fit the temporal model (2) to the data. Since the data are not normally distributed, we transform the data by using the Box–Cox transformation given in (1). λ is chosen, such that the residuals obtained by fitting the temporal model are normally distributed. Test if the residuals are dependent.
- Step 3.** Compute the sample covariance of the residuals resulting from fitting two temporal models to the data from two stations. Find the relationship between the covariance and the distance between the two stations, and then, construct the spatial weights matrix W . For example, if the sample covariance is decreasing as the distance between the corresponding two stations is increasing, we can use the inverse of the distance as the corresponding off-diagonal element in the spatial weight matrix W .

Use the matrix W to establish the simultaneous autoregressive (SAR) model at each time. Estimate the parameters of the SAR model by using the residuals obtained by fitting N temporal models to the ozone concentration data.

- Step 4.** Estimate the entropy h_t of the SAR model at each time t and denote it by \hat{h}_t . Apply the change-point detection method given in [21] to the entropy time series $\{\hat{h}_t\}$ to detect multiple change-points.

3. Application to Real Ozone Concentration Data

In this section, we use the methodology proposed in the previous section to detect changes in ground-level ozone concentration data collected in the Toronto region of Canada between June and September for the years from 1988 to 2009. There are 19 monitoring stations in this region, and the rate of missing data at each station is below 50%. We primarily focus on the daily time scale in four consecutive summer months from June to September for the years ranging from 1988 to 2009. Thus, we have the original data $X_{i,t}$, $i = 1, \dots, 19$; $t = 1, \dots, 2684$, formed by 2684 (22 years \times 122 days) daily maximum eight-hour moving averages of hourly ozone concentration data recorded in micrograms per cubic meter from each of the 19 stations, which are displayed in Figure 1. Figure 2 displays the locations of the 19 stations and their indexes. The numbers of missing data at nine of the stations are under 200, while the numbers of missing data at the other five stations are between 400 and 800. The remaining five stations have a number of missing data close to 1000. Figure 3 presents the box-and-whisker plots of the data collected at each station. It is clear that the data at each station are not normally distributed. Thus, we apply the Box–Cox power transformation (1) to the data $\{X_{i,t}\}$ and obtain the transformed data $\{Z_{i,t}(\lambda)\}$, for each $\lambda \in \{0.3, 0.31, 0.32, \dots, 0.6\}$. The final value of λ will be decided later.

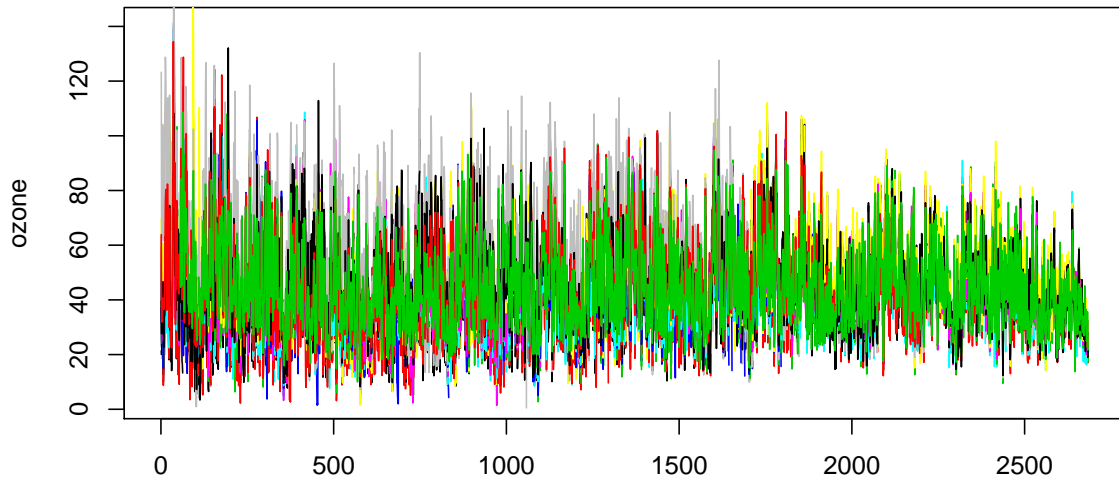


Figure 1. The ozone concentration data in 2,684 days and from 19 stations.

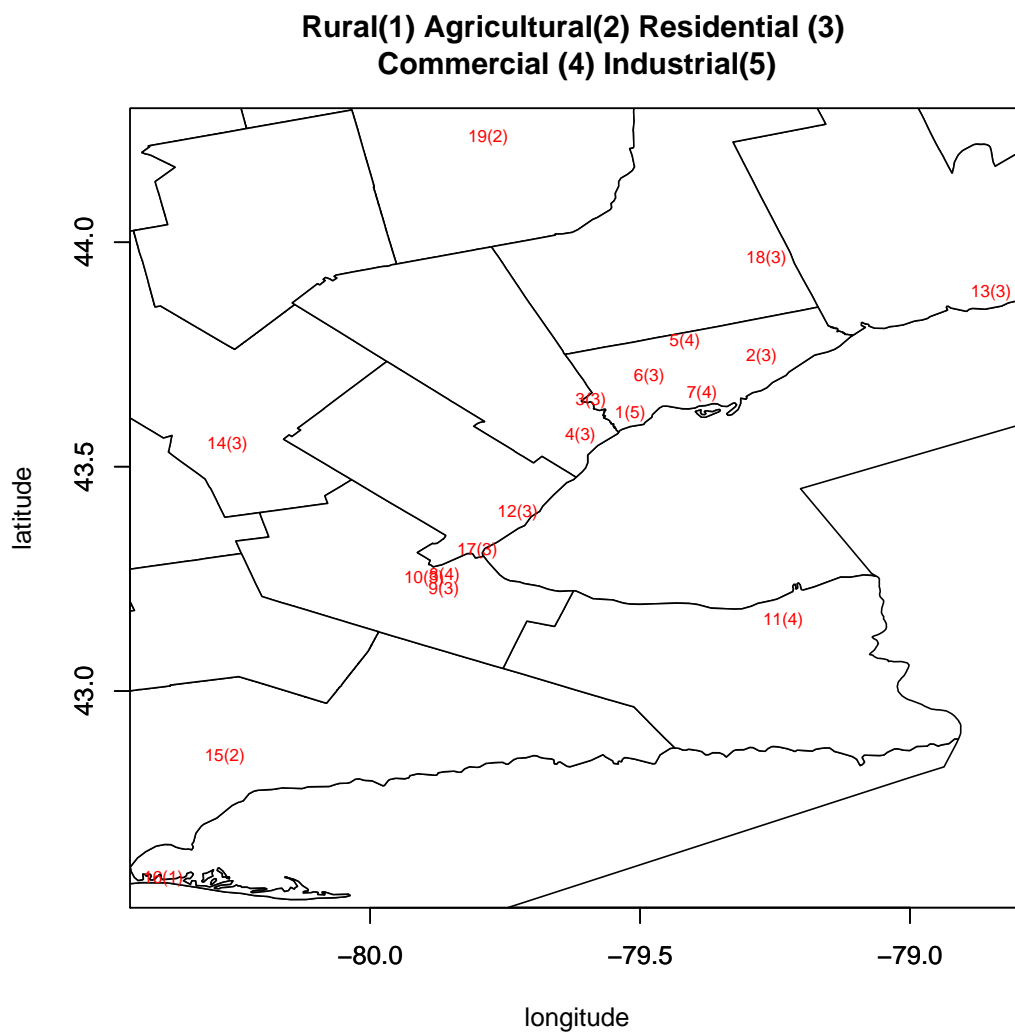


Figure 2. Locations of the 19 stations.

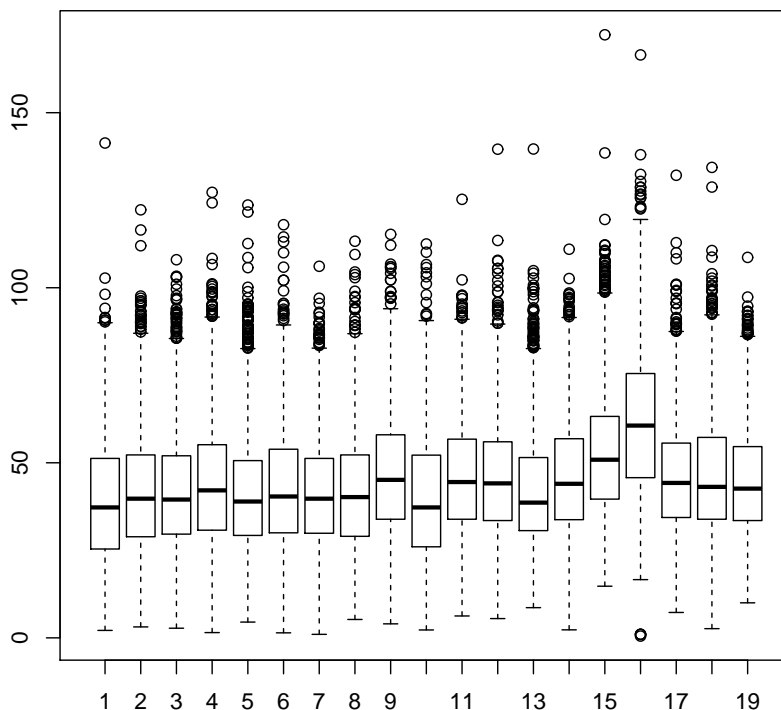


Figure 3. The respective box-and-whisker plots of the ozone concentration data from the 19 stations.

Since there are 122 days from 1 June to 31 September in each year, the time period is thus 122, so that L in the model (2) is 61. Preliminary data analysis shows that we may use the temporal model (2) with $p = 1$ and $q = 3$ to fit the data. We write the model (2) with $p = 1$ and $q = 3$ as follows:

$$Z_{i,t}(\lambda) = \beta_{0,i}(\lambda) + \beta_{1,i}(\lambda) \cos(t\pi/61) + \beta_{2,i}(\lambda) \sin(t\pi/61) + \beta_{3,i}(\lambda)Z_{i,t-1}(\lambda) + \beta_{4,i}(\lambda)Z_{i,t-2}(\lambda) + \beta_{5,i}(\lambda)Z_{i,t-3}(\lambda) + \varepsilon_{i,t}(\lambda). \tag{7}$$

Let $\lambda \in \{0.3, 0.31, 0.32, \dots, 0.6\}$. For each λ and a fixed i , we fit the model (7) to the data $\{Z_{i,t}(\lambda)\}$ by least squares and obtain the estimates $\hat{\beta}_{j,i}(\lambda)$, $j = 1, \dots, 5$, of the parameters $\beta_{j,i}(\lambda)$, $j = 1, \dots, 5$. We compute the residuals $\{\hat{\varepsilon}_{i,t}(\lambda)\}$ by:

$$\hat{\varepsilon}_{i,t}(\lambda) = Z_{i,t}(\lambda) - \left[\hat{\beta}_{0,i} + \hat{\beta}_{1,i} \cos(t\pi/61) + \hat{\beta}_{2,i} \sin(t\pi/61) + \hat{\beta}_{3,i}Z_{i,t-1}(\lambda) + \hat{\beta}_{4,i}Z_{i,t-2}(\lambda) + \hat{\beta}_{5,i}Z_{i,t-3}(\lambda) \right], \quad t = 1, \dots, T.$$

We remark that the purpose of applying the Box–Cox power transformation to the ozone concentration data is such that $\{\varepsilon_{i,t}(\lambda)\}$ are approximately normally distributed. Thus, we can choose λ in terms of p -values of a normality test on $\{\hat{\varepsilon}_{i,t}(\lambda)\}$ for each fixed pair of λ and i . In this application, the Pearson chi-squared test (R code: `pearson.test`) is employed. By applying this test to the residuals $\{\hat{\varepsilon}_{i,t}(\lambda)\}$ for fixed λ and i , we obtain the p -value $p_i(\lambda)$. Let $p(\lambda) = \text{Median}\{p_i(\lambda), i = 1, \dots, 19\}$ for each $\lambda \in \{0.3, 0.31, 0.32, \dots, 0.6\}$. $\lambda = \hat{\lambda}$ is chosen, such that $\hat{\lambda} = \arg \max_{\lambda} p(\lambda)$, which turns out to be 0.48. Hence, $\hat{\lambda} = 0.48$ is used in the Box–Cox power transformation (1) hereafter.

Let $Y_{i,t} = (X_{i,t}^{0.48} - 1)/0.48$. As discussed above, $\{Y_{i,t}\}$ for each fixed i are modelled as:

$$Y_{i,t} = \beta_{0,i} + \beta_{1,i} \cos(t\pi/61) + \beta_{2,i} \sin(t\pi/61) + \beta_{3,i}Y_{i,t-1} + \beta_{4,i}Y_{i,t-2} + \beta_{5,i}Y_{i,t-3} + \varepsilon_{i,t}, \tag{8}$$

where $t = 1, \dots, T$. As done previously, we estimate $\beta_{j,i}, j = 0, 1, \dots, 5$ by the least squares method. Denote these estimates by $\hat{\beta}_{j,i}, j = 0, 1, \dots, 5$. We can then compute the residuals $\hat{\varepsilon}_{i,t}$ for $t = 1, \dots, T$. $\{\hat{\beta}_{j,i}\}$ and $\{\hat{\varepsilon}_{i,t}\}$ are plotted respectively in Figures 4 and 5. To examine if the model has fitted the data from each station well, we compute R_i^2 (the coefficient of determination) obtained by fitting the model (8) to the data from each of 19 monitoring stations. $R_i^2, i = 1, \dots, 19$ are displayed in Table 1, which shows that the values of R_i^2 are all larger than 0.95. We also compute the p -value p_i obtained by performing Pearson chi-square test on $\{\hat{\varepsilon}_{i,t}, t = 1, \dots, T\}$ for $i = 1, \dots, 19$, which are also displayed in Table 1. From this table, it can be observed that only three p -values of the Pearson chi-square test are smaller than 0.01. Further, for each time series $\{\hat{\varepsilon}_{i,t}, t = 1, \dots, 2684\}$, we compute the Box–Pierce test statistic ([23]) for each of the two null hypotheses $H_0 : \rho(1) = \rho(2) = \rho(3) = \rho(4) = 0$ and $H_0 : \rho(1) = \rho(2) = \dots = \rho(7) = 0$, where $\rho(k)$ is the autocorrelation at lag k (R code: `Box.test`). The box-and-whisker plot of the p -values from the Box–Pierce test is displayed in Figure 6, which shows that both null hypotheses cannot be rejected, *i.e.*, the residuals can be considered as uncorrelated at Lags 1 to 7.

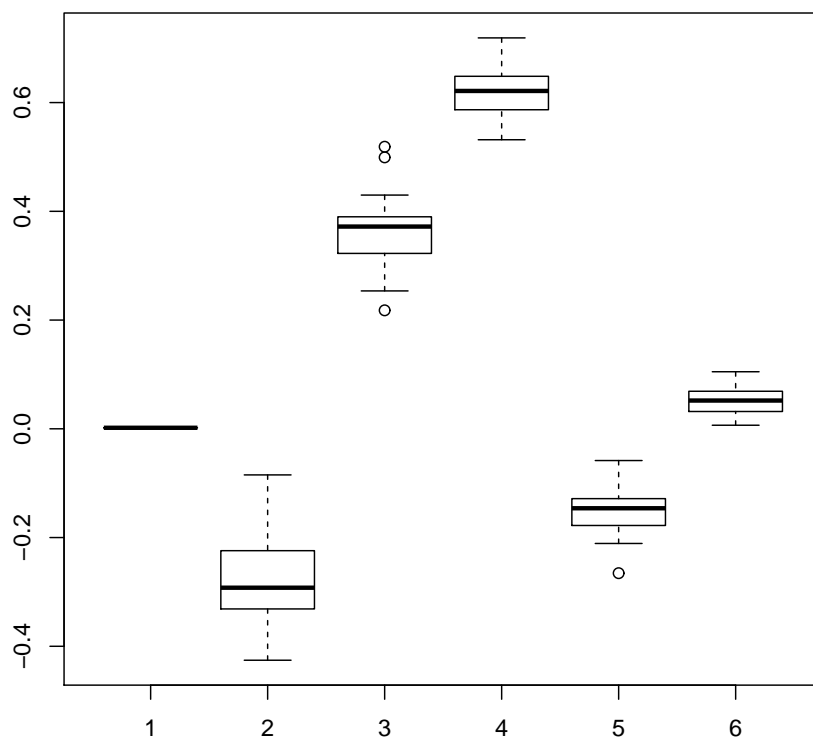


Figure 4. The respective box-and-whisker plots of $\hat{\beta}_{j,i}, j = 0, 1, \dots, 5$.

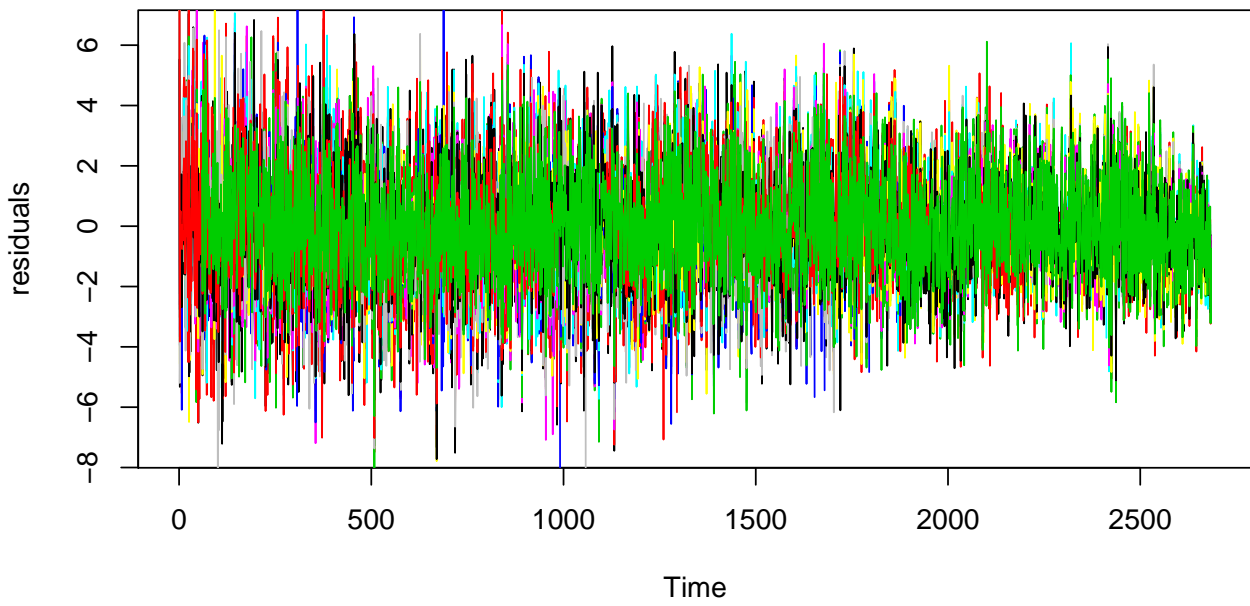


Figure 5. Plot of $\hat{\varepsilon}_{i,t}$, $i = 1, \dots, 19$.

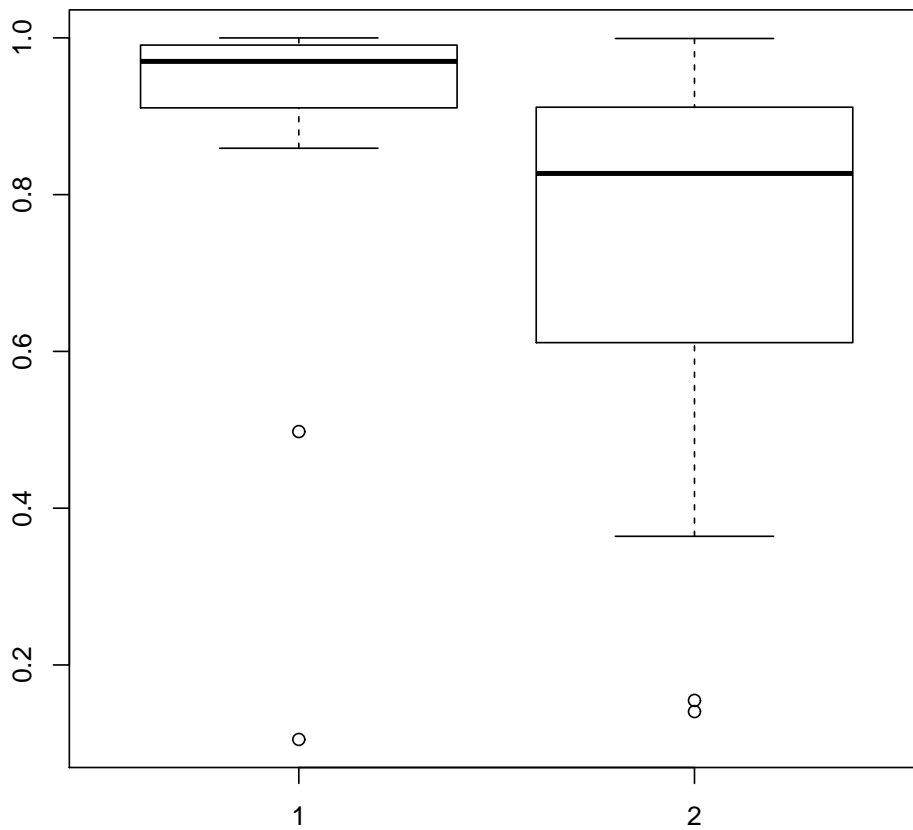


Figure 6. The respective box-and-whisker plots of p -values of the Box–Pierce test on the respective two null hypotheses $H_0 : \rho(1) = \rho(2) = \rho(3) = \rho(4) = 0$ and $H_0 : \rho(1) = \rho(2) = \dots = \rho(7) = 0$.

Table 1. The respective coefficient of determination, R_i^2 , and the p -value, p_i , for $i = 1, \dots, 19$.

	Station ID									
	1	2	3	4	5	6	7	8	9	10
R_i^2	0.9538	0.9662	0.9662	0.9647	0.96708	0.9641	0.9714	0.9657	0.9716	0.9493
p_i	0.3568	0.7291	0.0110	0.0547	0.06204	0.4119	0.5094	0.5559	0.3411	0.3453
	Station ID									
	11	12	13	14	15	16	17	18	19	
R_i^2	0.9742	0.9711	0.9700	0.9742	0.97906	0.9745	0.9754	0.9752	0.9785	
p_i	0.4131	0.7438	0.0140	0.4816	0.01905	0.0056	0.1779	0.0001	0.0001	

Let $\hat{\epsilon}_i = (\hat{\epsilon}_{i,1}, \dots, \hat{\epsilon}_{i,2684})'$. Figure 7 displays the sample covariance $C_{i,j} = \hat{\epsilon}_i' \hat{\epsilon}_j / 2684$, $i, j = 1, \dots, 19$ and $i \neq j$, against the distance $d_{i,j} = \sqrt{(s_{i,1} - s_{j,1})^2 + (s_{i,2} - s_{j,2})^2}$, where $(s_{i,1}, s_{i,2})$ is the rectangular coordinate of the location of the i -th station. It can be seen that the covariance decreases as the distance increases. Thus, we construct the spatial weight matrix $W = (w_{i,j})_{19 \times 19}$ in (3) by letting all of its diagonal elements $\{w_{i,i}\}$ be zeros and off-diagonal elements $\{w_{i,j}, i \neq j\}$ be the inverse distances between the stations i and j , i.e., $w_{i,j} = 1/d_{i,j}$.

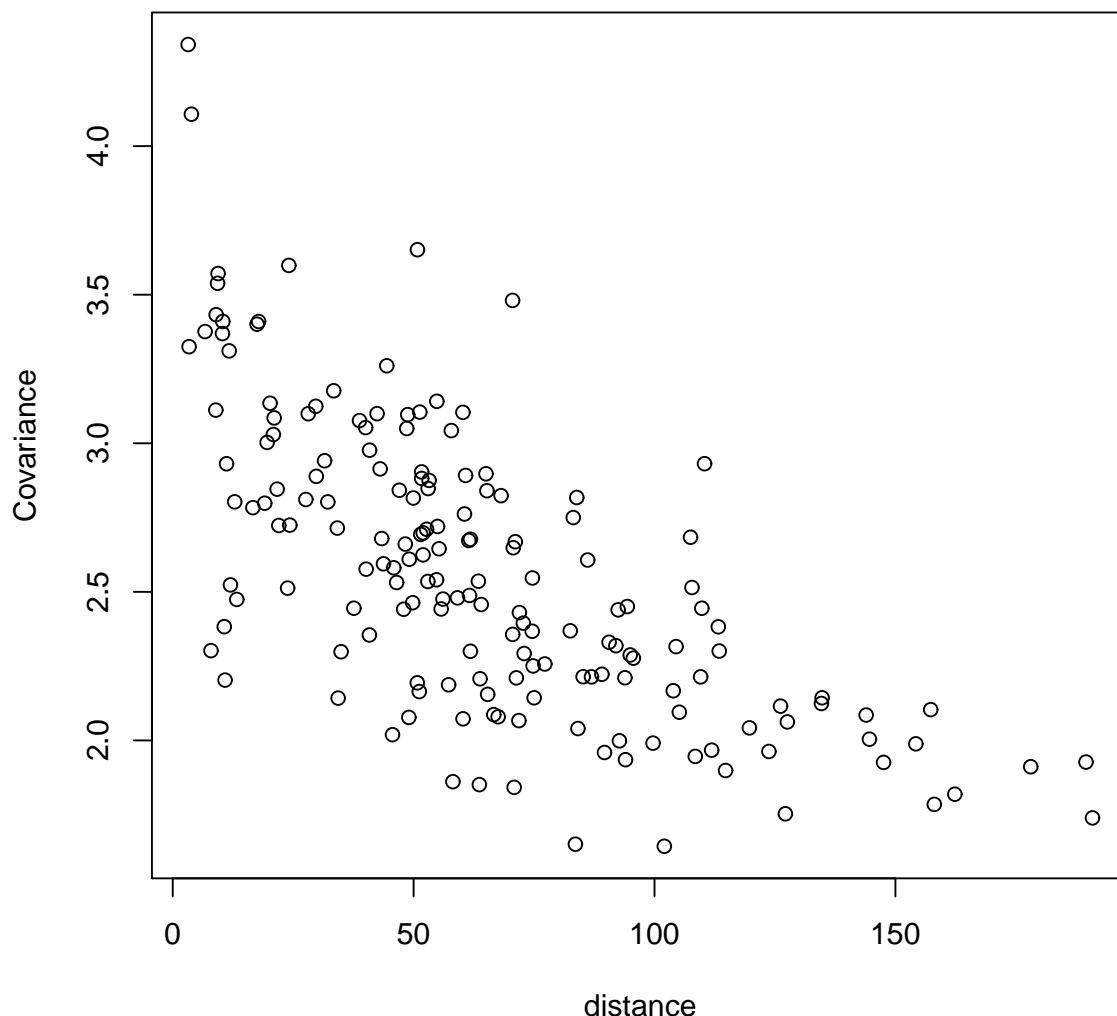


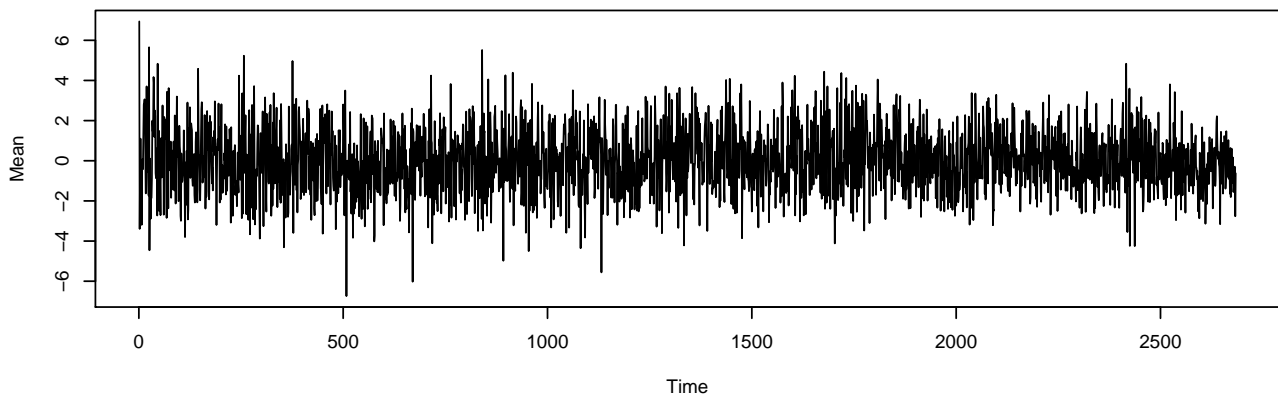
Figure 7. Plot of spatial covariance against the distance between two monitoring stations.

The data have been assumed to be spatially correlated. To confirm this, Moran’s I is used to test the dependence at each time, which is computed by:

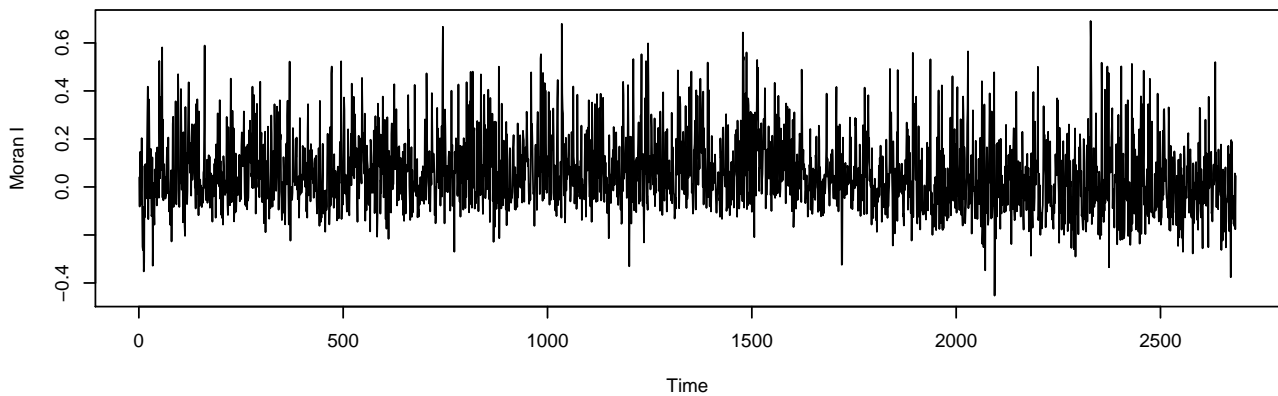
$$\mathcal{I}_t = \frac{19}{\sum_{i=1}^{19} \sum_{j=1}^{19} w_{i,j}} \times \frac{\sum_{i=1}^{19} \sum_{j=1}^{19} w_{i,j} (\hat{\epsilon}_{i,t} - \bar{\hat{\epsilon}}_t)(\hat{\epsilon}_{j,t} - \bar{\hat{\epsilon}}_t)}{\sum_{i=1}^{19} (\hat{\epsilon}_{i,t} - \bar{\hat{\epsilon}}_t)^2}$$

with $\bar{\hat{\epsilon}}_t = \sum_{i=1}^{19} \hat{\epsilon}_{i,t}/19$. More than 86% of the tests on the data at each time point are significant at the 0.05 level.

(a) mean



(b) Moran I



(c) Rho

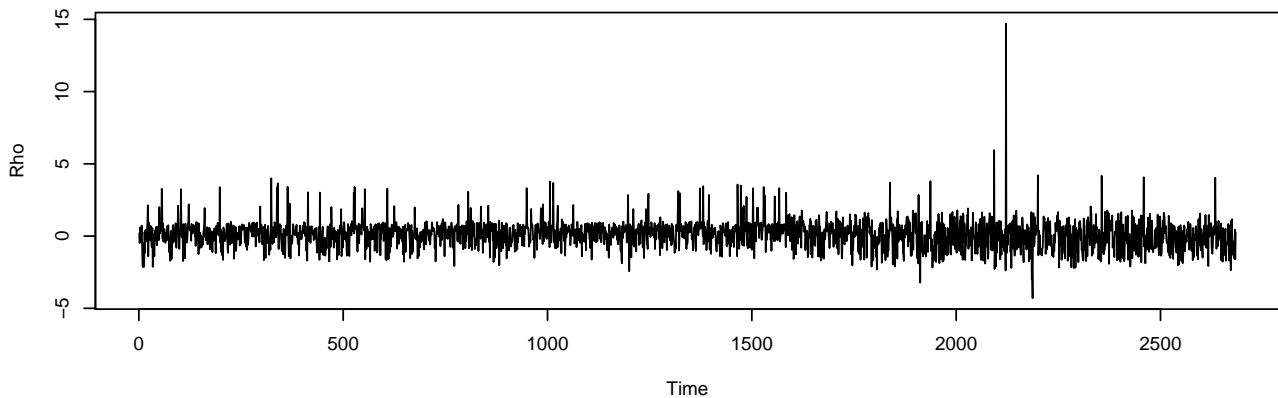


Figure 8. Respective plots of $\bar{\hat{\epsilon}}_t$, \mathcal{I}_t , and $\hat{\rho}_t$.

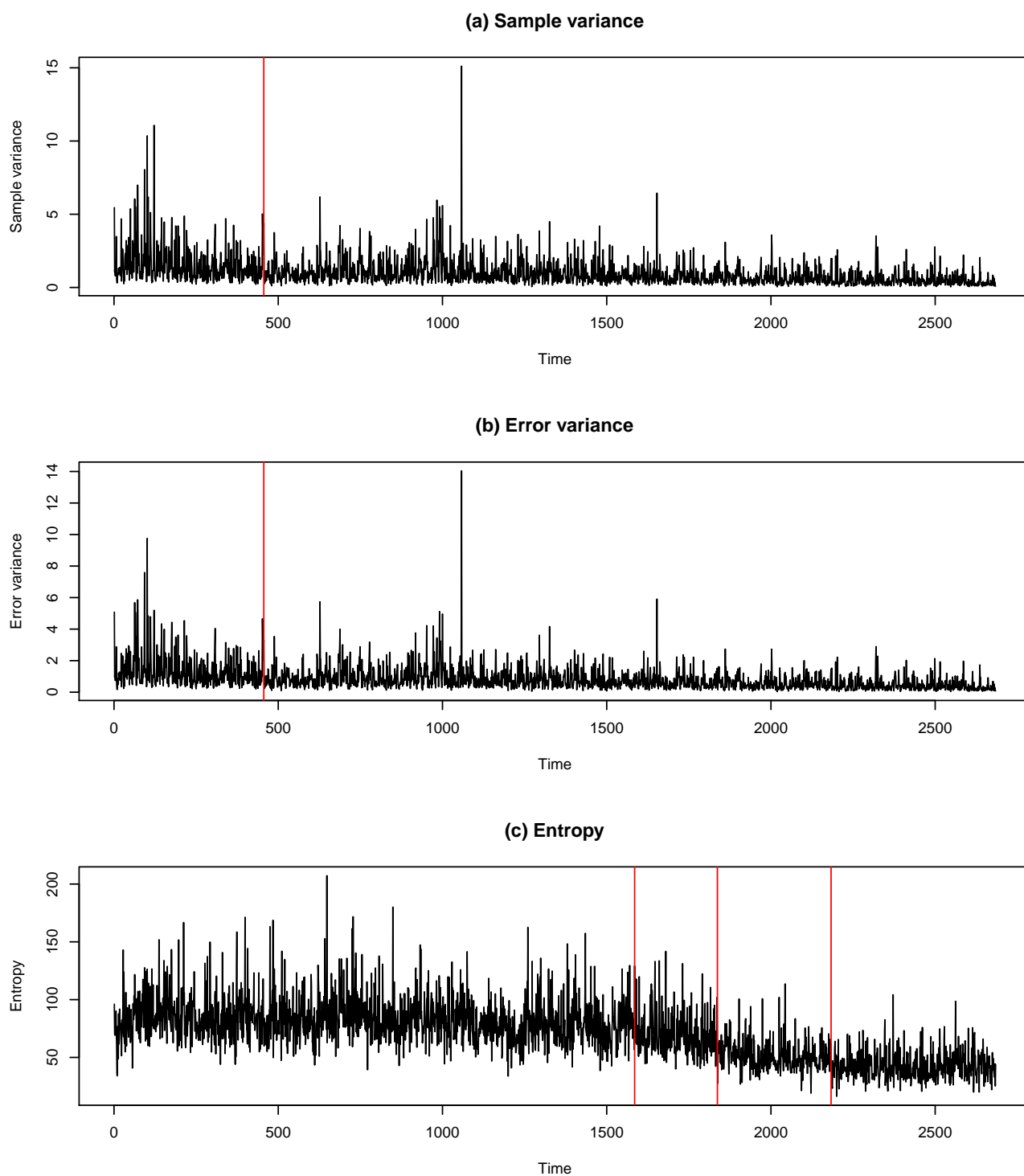


Figure 9. Respective plots of S_t^2 , $\hat{\sigma}_t^2$ and \hat{h}_t .

Replace ε_t by $((\hat{\varepsilon}_{1,t} - \bar{\varepsilon}_{.t}), \dots, (\hat{\varepsilon}_{19,t} - \bar{\varepsilon}_{.t}))'$ in Model (3). By Ord (1975) [18], we obtain the maximum likelihood estimates $\hat{\rho}_t$ and $\hat{\sigma}_t^2$ of (3), and then, we obtain the estimate of $\hat{\Sigma}_t = \hat{\sigma}_t^2(I_{19} - \hat{\rho}_t W)^{-2}$. Thus, we obtain $\hat{h}_t = \frac{1}{2} \log \left[(2\pi e)^n |\hat{\Sigma}_t| \right]$, an estimate of the differential entropy defined in (4).

As shown in Figure 6, $\{\hat{\varepsilon}_{.t}, t = 1, \dots, 2684\}$ can be considered to be independent distributed, and the same argument is also true for $\{\mathcal{I}_t\}$, $\{\hat{\sigma}_t^2\}$, $\{\hat{\rho}_t\}$ and $\{\hat{h}_t\}$. Let $S_t^2 = \sum_{i=1}^{19} (\hat{\varepsilon}_{i,t} - \bar{\varepsilon}_{.t})^2 / 18$ be the sample variance. The sample mean $\bar{\varepsilon}_{.t}$, Moran's I \mathcal{I}_t and $\hat{\rho}_t$ are respectively displayed in Figure 8. We

apply the change-point detection method given in [21] to each time series of $\bar{\varepsilon}_t$, \mathcal{I}_t and $\hat{\rho}_t$ and cannot find any change-point. Thus, if we only consider the time series $\bar{\varepsilon}_t$, \mathcal{I}_t and $\hat{\rho}_t$, we have to claim that there is no change in the ozone concentration in the Toronto region. In contrast, by applying the same method to both time series $\{S_t^2\}$ and $\{\hat{\sigma}_t^2\}$, we detect the same change-point at 456 (29 August 1991). If we also apply the same method to the time series $\{\hat{h}_t\}$, we find three change-points, 1585 (30 September 2000), 1837 (7 June 2003) and 2183 (17 September 2005). The sample variance S_t^2 , error variance $\hat{\sigma}_t^2$ and entropy \hat{h}_t are respectively displayed in Figure 9.

By Simmons (2002) [24], each year in Canada, 16,000 people die prematurely as a result of air pollution. Cars and light trucks are responsible for the majority of transportation emissions, but the heavy trucks in the trucking industry are also a major contributor, whose emissions have increased more rapidly than any other element of the Canadian transportation sector. Historically, Canada has taken a passive approach to the regulation of motor vehicle pollution. The estimated change-points, 1585 (30 September 2000), 1837 (7 June 2003) and 2183 (17 September 2005), are consistent with the following published regulations. By the 44th Working Party on Pollution and Energy (GRPE) of the United Nations [25], since 1988, Canadian on-road vehicle emission standards have been, through a combination of regulations and voluntary agreements, aligned with those of the U.S. EPA (Environmental Protection Agency). The Canadian Environmental Protection Act 1999 transferred the responsibility to the Department of the Environment. Environment Canada adopted the Sulphur in Gasoline Regulations in June, 1999, and proposed the Sulphur in Diesel Fuel Regulations in December, 2001. The Canadian Department of the Environment (Environment Canada) published proposed new on-road vehicle and engine emission regulations on 30 March 2002. Regulations for each of the five off-road groups were proposed later in 2002 and during 2003. Sulphur in gasoline was limited to on average 30 parts per million (ppm) in 2005, with an interim limit of 150 parts per million in 2002. It is noted that ground level ozone is not emitted directly into the air, but is created by chemical reactions between oxides of nitrogen and volatile organic compounds, which include sulphur content. Thus, limiting sulphur in gasoline can help to improve the air quality.

4. Conclusion

In this paper, we propose a methodology for detecting changes in ground-level ozone concentrations by using entropy. It is shown via a real data example that the entropy \hat{h}_t , a function of $\hat{\rho}_t$ and $\hat{\sigma}_t^2$, can be used for detecting changes in ground-level ozone concentration data. As demonstrated in Section 3, when the same change-point detection method is applied to each of the time series $\{\bar{\varepsilon}_t\}$, $\{\mathcal{I}_t\}$, $\{\hat{\rho}_t\}$, $\{S_t^2\}$, $\{\hat{\sigma}_t^2\}$ and $\{\hat{h}_t\}$, the time series that is the best for detection of multiple change-points is $\{\hat{h}_t\}$. This may be due to the fact that the entropy can be used to measure various spatial uncertainties, including both spatial variance and spatial dependence, and is able to extract more information from the data than some other statistics, e.g., $\hat{\rho}_t$ and $\hat{\sigma}_t^2$. The proposed methodology is also applicable to other climate data.

As shown in the data example, the changes in both the mean and spatial dependence of ozone concentrations may not be detectable statistically after the regulations of environmental agencies are proposed. In contrast, the changes in the spatial uncertainties of ground-level ozone concentrations, measured by entropy, may be detectable statistically. Thus, after a regulation is promulgated,

environmental agencies may be effective at monitoring the air quality change by employing the methodology presented in Section 2, which may help them to decide what is the next step for improving air quality.

Acknowledgements

The research was partially supported by York University and the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank the anonymous referees for the helpful comments and suggestions.

Author Contributions

All authors contributed equally to this paper. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Rodríguez, S.; Reyes, H.; Pérez, P.; Vaquera, H. Selection of a Subset of Meteorological Variables for Ozone Analysis: Case Study of Pedregal Station in Mexico City. *J. Environ. Sci. Eng. A* **2012**, *1*, 11–20.
2. Özbay, B.; Keskin, G.A.; Doğruparmak, Ş.Ç.; Ayberk, S. Multivariate Methods for Ground-level Ozone Modeling. *Atmos. Res.* **2011**, *102*, 57–65.
3. Shu, Y.; Lam, N.S.N. Spatial Disaggregation of Carbon Dioxide Emissions from Road Traffic Based on Multiple Linear Regression Model. *Atmos. Environ.* **2011**, *45*, 634–640, doi:10.1016/j.atmosenv.2010.10.037.
4. Baur, D.; Saisana, M.; Schulze, N. Modelling the Effects of Meteorological Variables on Ozone Concentration—a Quantile Regression Approach. *Atmos. Environ.* **2004**, *38*, 4689–4699.
5. Munir, S.; Chen, H.; Ropkins, K. Modelling the Impact of Road Traffic on Ground Level Ozone Concentration Using a Quantile Regression Approach. *Atmos. Environ.* **2012**, *60*, 283–291.
6. Chelani, A.B. Nonlinear dynamical analysis of ground level ozone concentrations at different temporal scales. *Atmos. Environ.* **2010**, *44*, 4318–4324, doi:10.1016/j.atmosenv.2010.07.028
7. Niu, X.F. Nonlinear Additive Models for Environmental Time Series, with Applications to Ground-Level Ozone Data Analysis. *J. Am. Stat. Assoc.* **1996**, *91*, 1310–1321.
8. Weng, Y.C.; Chang, N.B.; Lee, T.Y. Nonlinear Time Series Analysis of Ground-level Ozone Dynamics in Southern Taiwan. *J. Environ. Manag.* **2008**, *87*, 405–414.
9. Jin, B.; Chan, E.; Wu, Y. Hierarchical Bayesian Spatio-temporal Modelling of Regional Ozone Concentrations and Network Design. *J. Environ. Stat.* **2011**, *3*, 1–32.
10. Sahu, S.K.; Bakar, K.S. Hierarchical Bayesian Autoregressive Models for Large Space-time Data with Applications to Ozone Concentration Modelling. *Appl. Stoch. Models Bus. Ind.* **2012**, *28*, 395–415, doi: 10.1002/asmb.1951.

11. Gyarmati-Szabó, J.; Bogachev, L.V.; Chen, H. Modelling Threshold Exceedances of Air Pollution Concentrations via Non-homogeneous Poisson Process with Multiple Change-points. *Atmos. Environ.* **2011**, *45*, 5493–5503, doi:10.1016/j.atmosenv.2011.06.049.
12. Achcar, J.A.; Fernández-Bremauntz, A.A.; Rodrigues, E.R.; Tzintzun, G. Estimating the Number of Ozone Peaks in Mexico City Using a Non-homogeneous Poisson Model. *Environmetrics* **2008**, *19*, 469–485, doi: 10.1002/env.890.
13. Smith, R.L.; Shively, T.S. Point Process Approach to Modeling Trends in Tropospheric Ozone Based on Exceedances of a High Threshold. *Atmos. Environ.* **1995**, *29*, 3489–3499.
14. de Nazelle, A.; Arunachalam, S.; Serre, M.L. Bayesian Maximum Entropy Integration of Ozone Observations and Model Predictions: an Application for Attainment Demonstration in North Carolina. *Environ. Sci. Technol.* **2010**, *44*, 5707–5713, doi:10.1021/es100228w.
15. Fuentes, M.; Chaudhuri, A.; Holland, D.M. Bayesian Entropy for Spatial Sampling Design of Environmental Data. *Environ. Ecol. Stat.* **2007**, *14*, 323–340.
16. Martínez, J.; García, P.J.; Alejano, L.; Reyes, A. Detection of Outliers in Gas Emissions from Urban Areas Using Functional Data Analysis. *J. Hazard. Mater.* **2011**, *186*, 144–149.
17. Sancho, J.; Martínez, J.; Pastor, J.J.; Taboada, J.; Piñeiro, J.I.; García-Nieto, P.J. New Methodology to Determine Air Quality in Urban Areas Based on Runs Rules for Functional Data. *Atmos. Environ.* **2014**, *83*, 185–192.
18. Ord, K. Estimation Methods for Models of Spatial Interaction, *J. Am. Stat. Assoc.* **1975**, *70*, 120–126.
19. Oliveira, V.D.; Song, J.J. Bayesian Analysis of Simultaneous Autoregressive Models. *Sankhyā* **2008**, *70-B(2)*, 323–350.
20. Ahmed, N.A.; Gokhale, D.V. Entropy Expressions and Their Estimators for Multivariate Distributions. *IEEE Trans. Inf. Theory* **1989**, *35*, 688–692.
21. Jin, B.; Shi, X.; Wu, Y. A Novel and Fast Methodology for Simultaneous Multiple Structural Break Estimation and Variable Selection for Nonstationary Time Series Models. *Stat. Comput.* **2013**, *23*, 221–231, doi 10.1007/s11222-011-9304-6.
22. Csörgő, M.; Horvath, L. *Limit Theorems in Change-Point Analysis*; Wiley: Chichester, UK, 1997.
23. Box, G.E.P.; Pierce, D.A. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *J. Am. Stat. Assoc.* **1970**, *65*, 1509–1526.
24. Simmons, G. *Canadian Regulation of Air Pollution from Motor Vehicles*; Researched and prepared by Greg Simmons for Greenpeace and the Sierra Legal Defence Fund, 2002; Sierra Legal Defence Fund: Vancouver, BC, Canada, 2002.
25. United Nations. *Canadian On-road Vehicle and Engine Emission Regulations*; Informal document No.4 (44th GRPE, 10-14 June 2002, agenda item 10.); Available online: www.unece.org/fileadmin/DAM/trans/doc/2002/wp29grpe/TRANS-WP29-GRPE-44-inf04e.pdf (accessed on 29 April 2015).