INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **27**: 805–818 (2007)
Published online 18 October 2006 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/joc.1432

RMetS
ROYAL METEOROLOGICAL SOCIETY

Bayesian change-point analysis of heat spell occurrences in Montreal, Canada

M. N. Khaliq,[a,*] T. B. M. J. Ouarda,[b] A. St-Hilaire[b] and P. Gachon[c]

[a] *Ouranos, 550 Sherbrooke Street West, West Tower, 19th Floor, Montreal, Quebec H3A 1B9, Canada*
[b] *Canada Research Chair on the Estimation of Hydrological Variables, Hydro Quebec/NSERC Chair in Statistical Hydrology, INRS-ETE, University of Quebec, 490 de la Couronne, Quebec City, Quebec G1K 9A9, Canada*
[c] *Atmospheric Science and Technology Directorate, Adaptation and Impacts Research Division, Environment Canada at Ouranos, 550 Sherbrooke Street West, West Tower, 19th Floor, Montreal, Quebec H3A 1B9, Canada*

Abstract:

Positive/upward shifts in the rate of occurrence of heat spells can considerably impact socioeconomic sectors. Particularly, populous urban areas and centers of regional socioeconomic activities are more vulnerable to the enhanced activity of heat spells. In this study, 24 time series of annual counts of summer-season (June–August) heat spells are derived from homogenized records of daily minimum and maximum temperatures (i.e. $T_{\min}$ and $T_{\max}$) observed at McTavish station, located in the center of Montreal (Canada), over the period 1896–1991. Twelve of these time series, which fulfill the assumption of the Poisson process for heat spell occurrences, are examined for abrupt changes in the rate of occurrences using hierarchical Bayesian change-point approach. In these analyses, a heat spell is defined as an extreme climate event with $T_{\min}$ and $T_{\max}$ simultaneously above selected thresholds and a duration $\geq$1-day. The results of the Bayesian change-point analyses suggest structural inhomogeneities within the heat spell observations, i.e. the results do not support abrupt changes for all time series of annual counts of heat spells; this may not have been possible to detect by studying heat spells defined on the basis of just a single combination of $T_{\min}$ and $T_{\max}$ thresholds. Furthermore, the overall results of the Bayesian change-point analyses and those of commonly employed nonparametric trend detection and estimation techniques, when applied to change-point free smaller samples, suggest that there is inadequate evidence in favor of increased activity of heat spells in Montreal during the third last and second last decades (i.e. 1970s and 1980s) of the 20th century, which are the most recent decades of the observation period analyzed.

KEY WORDS bayesian inference; change-point analysis; climate variability; heat spells; MCMC; poisson process

Received 14 December 2005; Revised 29 May 2006; Accepted 28 August 2006

INTRODUCTION

Extreme weather and climate phenomena, such as heat spells, is an area of great research interest since positive/upward shifts in their occurrences can cause considerable damage to ecosystems and human society. An example is the European heat wave of 2003, which caused tremendous loss of life and also impacted various socioeconomic sectors (WHO, 2003; Beniston and Diaz, 2004; Beniston and Stephenson, 2004; Schär *et al.*, 2004). In general, heat spells are extreme in terms of intensity, duration and frequency, and are regional in scale and therefore, populous urban areas and centers of regional socioeconomic activities are more vulnerable to the enhanced activity of heat spells. Within southern Quebec, where most of the population of the province is concentrated, the greater Montreal area has occasionally experienced high daily maximum temperatures reaching above 30 °C, with varying durations, during the summer season (Khaliq *et al.*, 2005). Extended episodes of such extreme temperatures and the increase in their rate of occurrence can impact public health and regional economy. Thus, the main objective of this study is to investigate abrupt shifts (i.e. increases/decreases), as opposed to regular time-varying changes, in the rate of occurrence of heat spells in Montreal using Bayesian change-point analysis, assuming that the heat spell occurrences follow a Poisson point process. The Poisson point process has been successfully used for modeling climate data (Rodriguez-Iturbe *et al.*, 1987; Rootzén and Tajvidi, 1997; Katz, 2002) and the adequacy of this approach is also verified in this study. The use of Bayesian framework is intuitively appealing, particularly in situations where reliable data are scarce (which is often the case in climatology) to make statistical inference by incorporating information from other reliable sources, such as previous knowledge and expert opinion about the system being studied. The change-point analysis plays an important role in the analysis of climate data and can provide new insights into climate variability not available with other

* Correspondence to: M. N. Khaliq, Ouranos, 550 Sherbrooke Street West, West Tower, 19th Floor, Montreal, Quebec H3A 1B9, Canada. E-mail: khaliq.naveed@ouranos.ca

methods (Elsner *et al*., 2004). The change-point analysis (1) directly addresses the question of where the change in the mean value of the observations is likely to have occurred (Solow, 1987), (2) can serve to pinpoint potential inhomogeneities in records arising from improved observational technologies or changes in station location or modification in climatic conditions (Lund and Reeves, 2002; Elsner *et al*., 2004) and (3) can also be used for determining probabilistic effects of climate change (Cox *et al*., 2002).

Although the focus of this paper is on the Bayesian change-point analyses, we have also performed some nonparametric analyses to help develop a preliminary understanding of the temporal evolution of Montreal's heat spell climate. Nonparametric tools are considered to be more robust than parametric ones and are strongly recommended for climate data analyses (Lanzante, 1996). However, it should be noted that it is not the intention of the present study to investigate the relative weaknesses and/or strengths of Bayesian and nonparametric approaches but simply to analyze temporal changes in heat spell occurrences in Montreal.

Usually, a heat spell involves the occurrence of consecutive days of extreme high temperatures, which could be either daily maximum temperatures (henceforth denoted as $T_{\max}$) or daily minimum temperatures (henceforth denoted as $T_{\min}$) or the two together. These types of extreme climate events have been experienced in different parts of the world. For example, Kunkel *et al*. (1996) reported that the extreme heat event of 1995 in Chicago, Illinois, was characterized by high dewpoint and high minimum temperatures during the night and was classified as the most intense heat wave of the latter half of the 20th century in this region. The European heat wave of 2003 was characterized with extremely high mean daily temperatures (Schär *et al*., 2004). There is no uniform definition of a heat spell and its choice depends upon a number of factors, e.g. the type and length of the available data, the chosen sector impacted by the heat spell (i.e. public health, agriculture, wildlife, socioeconomics, etc.), time of the year, etc. Throughout this study, a heat spell is defined as an extreme climate event with both $T_{\min}$ and $T_{\max}$ simultaneously above suitably chosen reasonably high joint $T_{\min}$ and $T_{\max}$ thresholds, which are discussed in detail later in the paper, and having a duration $\geq$1-day. This type of spells are supposedly more severe from the viewpoint of impact on human health as compared to spells defined on the basis of either $T_{\max}$ or $T_{\min}$ alone, and hence the results of this study could be useful for the public health industry for planning emergency measures in and around Montreal. We have employed homogenized $T_{\min}$ and $T_{\max}$ observations to define heat spells and have not taken into account the effect of relative humidity, for which homogenized observations were not available. However, inclusion of the relative humidity factor in the definition of heat spells would be certainly useful since heat spells with increased atmospheric moisture are likely to be more damaging.

The outline of this paper is as follows. The data used for the study and the methods to obtain heat spells are discussed in the section '$T_{\min}$ and $T_{\max}$ Data and Their Thresholds', followed by a discussion of nonparametric analysis of long-term trends and split-sample comparisons in the section 'Nonparametric Trend Analyses and Split-sample C'. The section 'Poisson Point Process' deals with the suitability of the Poisson process as the primary model for heat spell occurrences and the section 'Bayesian Analyses' presents a hierarchical Bayesian change-point analysis approach. The section 'Results and Discussion' contains the detailed results and discussion of Bayesian analyses for heat spell climate of Montreal. A brief summary of the results and conclusions are given in the section 'Summary and Conclusions'.

$T_{\min}$ AND $T_{\max}$ DATA AND THEIR THRESHOLDS

Homogenized $T_{\min}$ and $T_{\max}$ data for the period 1895–2000 from McTavish station located in Montreal center were obtained from Environment Canada, which were developed by Vincent *et al*. (2000, 2002) for climate change studies in Canada. The use of homogenized data reduces the risk of artificially induced changes due to station relocation, equipment change, equipment drift, change in method of data collection, etc. Since we focus on heat spells, only hot summer-season (June–August) observations are considered for the current study. For the year 1895 and for the five years from 1992 to 1996, summer-season $T_{\min}$ and $T_{\max}$ observations are either not available or sparsely observed. Therefore, the analysis is carried out using data from the 1896–1991 period, which covers most of the 20th century and has the least missing values. The number of missing $T_{\min}$ and $T_{\max}$ observations was less than 0.06% of the total summer-season observations. Rather than filling in data for the few missing observations, termination of a heat spell at the missing value is considered throughout the analysis.

Total count of heat spells (CHS) as a function of selected $T_{\min}$ and $T_{\max}$ thresholds is displayed in Figure 1. The rare nature of extreme spells as both thresholds increase is obvious from this figure. Threshold values of 30 °C for $T_{\max}$ and 22 °C for $T_{\min}$ and a $\geq$3-day spell duration have been recommended by public health board (Drouin *et al*., 2005) to determine heat wave conditions in Montreal. It would have been interesting to analyze heat spells that satisfy these criteria (or exceed these criteria); however, only 113 (14), 53 (3) and 17 (0) spells having a $\geq$1-day duration ($\geq$3-day duration) and with $T_{\max}$ above 30 °C and $T_{\min}$ above 22, 23 and 24 °C, respectively, were found to have occurred (Figure 1). Therefore, in order to include a reasonable number of heat spells of duration $\geq$3 days in the present study, we selected four thresholds for the $T_{\min}$ (i.e. 18, 19, 20 and 21 °C) and six thresholds for the $T_{\max}$ (i.e. 27 to 32 °C, each at 1 °C interval) as compared with the single percentile-based threshold, both for the $T_{\min}$ and $T_{\max}$, used in Khaliq *et al*. (2006). Obviously, the choice of these thresholds is

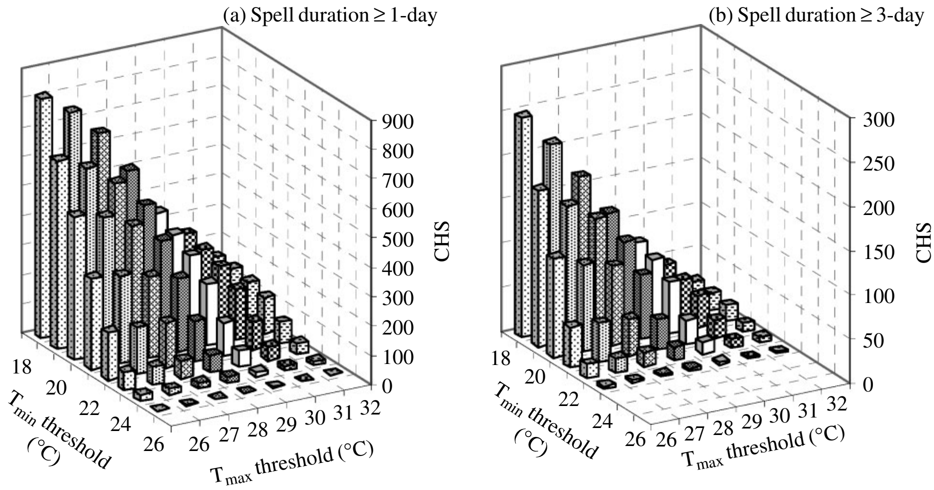

Figure 1. Total count of heat spells (CHS) having $T_{\min}$ and $T_{\max}$ values above the indicated thresholds and durations, (a) $\geq$1 day and (b) $\geq$3 days, observed during the June–August summer-season over the 1896–1991 period.

arbitrary, but it is reasonable in that it allows us to study a wide range of heat spells of various characteristics and at the same time helps maintain their extreme nature. For convenience of presentation, the time series of annual counts of heat spells for a particular combination of $T_{\min}$ threshold (u) and $T_{\max}$ threshold (v) is referred to as $Y^{u:v}$ throughout this paper.

NONPARAMETRIC TREND ANALYSES AND SPLIT-SAMPLE COMPARISONS

Nonparametric statistical tools are used to describe long-term temporal changes and to compare distributions of split samples of time series of annual counts of heat spells in order to have a preliminary understanding of their temporal evolution. These tools are generally considered to be more robust as compared to the parametric ones and are less affected by the presence of outliers or issues of nonnormality (Lanzante, 1996). The Mann–Kendall (Kendall, 1975) method is used to examine the statistical significance of temporal changes (i.e. time trends), and the Sen's method (Sen, 1968) and the three-group resistant line method (Hoaglin *et al.*, 1983) were used to obtain an estimate of the magnitude of the monotonic trend. To assess how the rate of occurrence of heat spells varies over the 96-year period (i.e. 1896–1991), three equal-sized split-samples (SS) are chosen to be examined: SS1 (1896–1927), SS2 (1928–59) and SS3 (1960–91). Probability distributions of the split samples are compared with each other using the distribution-free Wilcoxon rank-sum and robust rank-order tests (Siegel and Castellan, 1988). To assess the credibility of the null hypothesis, i.e. that there is no difference in the location of distributions, a 5% significance level is selected and used throughout this section.

Statistics of the annual counts of heat spells and their durations (i.e. mean number of heat spells $\hat{\mu}_Y$, variance of number of heat spells $\hat{\sigma}_Y^2$ and mean duration of heat spells $\hat{\mu}_D$), along with the proportion of spells having a duration $\geq$3 days ($\hat{p}_3$), and the estimates of time trends ($\hat{\delta}_Y$) and their significance are given in Table I. All the $Y^{u:v}$ time series are found independent by testing the first serial correlation coefficient at 5% significance level following the testing procedure described in Salas *et al.* (1980). Significant increasing (i.e. positive) time trends can be noticed for $Y^{18:27}$, $Y^{18:28}$, $Y^{18:29}$, $Y^{19:27}$, $Y^{19:28}$, $Y^{19:29}$, $Y^{20:27}$, $Y^{20:28}$, $Y^{20:29}$, $Y^{21:27}$ and $Y^{21:28}$ time series. In general, heat spells that have $T_{\min}$ above 18, 19, 20 and 21 °C and $T_{\max}$ above 30, 31 and 32 °C do not provide any clear evidence of increasing time trends. The proportion of spells having a $\geq$3-day duration ranges between 8 and 28%. Thus, majority of the spells have a duration <3 days. The mean duration $\hat{\mu}_D$ of spells ranges between 1.4 and 2.2 days for various combinations of selected $T_{\min}$ and $T_{\max}$ thresholds. Almost similar results were obtained using both Wilcoxon rank-sum and robust rank-order tests for split-sample analyses. Therefore, the results for the latter test are reported in Table II. The results of this table provide indications that (1) the mean rate of SS1 is significantly different from (or smaller than) that of SS3 for $Y^{18:27}$, $Y^{18:28}$, $Y^{18:29}$, $Y^{19:27}$, $Y^{19:28}$, $Y^{19:29}$, $Y^{20:27}$, $Y^{20:28}$, $Y^{21:27}$ and $Y^{21:28}$ time series, (2) the mean rate of SS1 is significantly smaller than that of SS2 for many $Y^{u:v}$ time series and (3) except for two time series (i.e. $Y^{18:30}$ and $Y^{19:30}$), no significant difference exists between the mean rates of SS2 and SS3. Thus, it can be concluded that there are indications that, at some time in the series, the mean rate of occurrences of heat spells shifts toward a higher rate and hence the credibility of statistically significant increasing time trends (discussed above and presented in Table I) is questionable, i.e. the presence of abrupt shifts might make the estimates of time trends invalid. However, it is difficult to generalize this conclusion on the basis of split-sample analyses because there are no clear indications of a shift for the $Y^{u:v}$ time series that have higher $T_{\max}$ thresholds (i.e. $v \geq 31$ °C). The time at which the rate of occurrence of heat spells shifts toward a higher rate is

DOI: 10.1002/joc

Table I. Statistics of the annual counts of heat spells and their durations, and the proportion of spells having a $\geq$3-day duration for the entire 1896–1991 period. Significant results of the Mann–Kendall test are indicated using '*' symbol. Slope estimates by Sen's method and the three-group resistant line method (indicated in parentheses) are in number of spells per 50 years. See text for additional explanation.

Statistic	T_{max} threshold v (°C)					
	27	28	29	30	31	32
T_{min} threshold u (°C): 18						
Mean, $\hat{\mu}_Y$	7.79	6.82	5.30	3.58	2.62	1.59
Variance, $\hat{\sigma}_Y^2$	7.16	7.58	6.49	4.60	3.67	2.43
Slope, $\hat{\delta}_Y$	1.56(1.81)*	1.52(1.92)*	1.05(1.56)*	0.0(0.0)	0.0(0.0)	0.0(0.0)
Proportion, $\hat{p}_3$	0.28	0.25	0.23	0.23	0.19	0.14
Mean, $\hat{\mu}_D$	2.19	2.08	1.93	1.86	1.74	1.54
T_{min} threshold u (°C): 19						
Mean, $\hat{\mu}_Y$	6.18	5.44	4.44	3.19	2.44	1.54
Variance, $\hat{\sigma}_Y^2$	6.48	6.92	5.53	4.05	3.60	2.48
Slope, $\hat{\delta}_Y$	1.48(2.34)*	1.48(1.56)*	0.86(1.17)*	0.0(0.39)	0.0(0.0)	0.0(0.0)
Proportion, $\hat{p}_3$	0.26	0.25	0.23	0.24	0.18	0.14
Mean, $\hat{\mu}_D$	2.08	2.03	1.91	1.84	1.72	1.53
T_{min} threshold u (°C): 20						
Mean, $\hat{\mu}_Y$	4.80	4.29	3.56	2.80	2.21	1.46
Variance, $\hat{\sigma}_Y^2$	6.01	5.70	4.48	3.59	3.35	2.36
Slope, $\hat{\delta}_Y$	1.04(0.83)*	0.74(0.78)*	0.0(0.39)*	0.0(0.0)	0.0(0.0)	0.0(0.0)
Proportion, $\hat{p}_3$	0.21	0.22	0.21	0.22	0.17	0.13
Mean, $\hat{\mu}_D$	1.87	1.86	1.82	1.78	1.69	1.52
T_{min} threshold u (°C): 21						
Mean, $\hat{\mu}_Y$	3.10	2.86	2.58	2.19	1.77	1.27
Variance, $\hat{\sigma}_Y^2$	5.04	4.35	4.06	3.21	2.70	2.18
Slope, $\hat{\delta}_Y$	0.89(0.91)*	0.60(0.80)*	0.0(0.78)	0.0(0.0)	0.0(0.39)	0.0(0.0)
Proportion, $\hat{p}_3$	0.15	0.15	0.14	0.13	0.12	0.08
Mean, $\hat{\mu}_D$	1.66	1.67	1.61	1.59	1.57	1.44

Table II. Split-sample (SS) comparisons of annual counts of heat spells having T_{min} and T_{max} above indicated thresholds. The p-values of the robust rank-order distribution test are shown and the significant cases are in bold. Split samples SS1, SS2 and SS3 consist of annual counts of heat spells observed during the periods 1896–1927, 1928–1959 and 1960–1991, respectively.

Split samples	T_{max} threshold v (°C)					
	27	28	29	30	31	32
T_{min} threshold u (°C): 18						
SS1:SS2	**<0.001**	**<0.001**	**<0.001**	**0.005**	0.086	0.260
SS1:SS3	**<0.001**	**<0.001**	**0.005**	0.512	0.955	0.892
SS2:SS3	0.841	0.191	0.138	**0.034**	0.144	0.175
T_{min} threshold u (°C): 19						
SS1:SS2	**<0.001**	**<0.001**	**<0.001**	**0.002**	0.096	0.198
SS1:SS3	**<0.001**	**<0.001**	**0.009**	0.400	0.737	0.856
SS2:SS3	0.879	0.665	0.368	**0.043**	0.286	0.136
T_{min} threshold u (°C): 20						
SS1:SS2	**0.023**	**0.040**	0.103	0.092	0.356	0.462
SS1:SS3	**<0.001**	**0.017**	0.206	0.670	0.883	0.799
SS2:SS3	0.612	0.995	0.766	0.259	0.533	0.300
T_{min} threshold u (°C): 21						
SS1:SS2	0.132	0.189	0.596	0.493	0.605	0.563
SS1:SS3	**0.001**	**0.020**	0.160	0.462	0.913	0.968
SS2:SS3	0.222	0.376	0.450	0.990	0.755	0.535

DOI: 10.1002/joc

studied in detail using the Bayesian change-point analysis technique, which is the main theme of the paper. Before presenting this technique, validation of the Poisson point process is provided in the section that follows.

POISSON POINT PROCESS

Poisson point process has been widely used for modeling hydro-meteorological phenomena. Typical applications can be found in the literature for modeling streamflow droughts (Zelenhasić and Salvai, 1987), rainfall occurrences (Rodriguez-Iturbe *et al.*, 1987), hurricane occurrences (Elsner and Bossak, 2001; Katz, 2002; Elsner *et al.*, 2004) and extreme wind storms (Rootzén and Tajvidi, 1997), in which it is assumed that the occurrence of events (e.g. wind storms or hurricanes) follows a homogeneous Poisson process. The number of events, Y_i, in any year, i, is independent of the number of events in any other year. Thus, the Poisson process is intended to represent events that occur at 'random' (Guttorp, 1995).

The random variable Y has a Poisson distribution with the mean rate parameter λ; that is, the probability of y events occurring in any year, i, can be expressed as

$$\Pr\{Y = y\} = e^{-\lambda}(\lambda)^y/y! \qquad y = 0, 1, 2, \ldots \text{and } \lambda > 0 \tag{1}$$

The mean and variance of Y are equal and are given by $\mu_Y = \sigma_Y^2 = \lambda$.

For Montreal heat spells, it is assumed that the occurrences of heat spells follow a Poisson point process. We chose to verify this assumption using the conventional frequentists' approach by using the fundamental property of equality of the mean and variance of the Poisson distribution and calculating the chi-squared test statistic given in Johnson *et al.* (1992), i.e.

$$\chi^2 = (m-1)\hat{\sigma}_Y^2/\hat{\mu}_Y \tag{2}$$

where $\hat{\sigma}_Y^2$ denotes the sample variance of the annual counts of heat spells (with divisor $m-1$) and m denotes the number of years. This statistic has an approximate chi-squared distribution, with $(m-1)$ degrees of freedom, under the null hypothesis of equality of the mean and variance. In a Poisson process of event occurrences, inter-event waiting times are exponentially distributed (Guttorp, 1995). Hence, in addition to Equation (2), this property of the Poisson process can also be used to test its adequacy by employing samples of inter-event waiting times obtained from recorded observations. Following the results of split-sample analyses presented in the section on 'Nonparametric Trend Analyses and Split-sample Comparisons', we consider 1932–1991 (60 years) and 1942–1991 (50 years) data for validating the suitability of the Poisson process. The calculated values of χ^2 and corresponding 95% confidence intervals are shown in Figure 2 for a number of combinations of $T_{\min}$ and $T_{\max}$ thresholds. It is obvious from this figure that the heat spell occurrences could be assumed to follow Poisson process only for a range of $T_{\min}$ and $T_{\max}$ thresholds as the calculated chi-squared values lie within the 95% confidence interval. It appears that the heat spells with $T_{\min}$ above 21 °C and $T_{\max}$ above the considered range of thresholds do not follow the Poisson process at 95% confidence level. Similar behavior can be noticed for heat spells with $T_{\max}$ above 31 and 32 °C and $T_{\min}$ above the considered range of thresholds. We also generated 5000 bootstrap samples for each of the time series and obtained 95% upper and lower confidence limits for the χ^2 test statistic given in Equation (2). Almost the same results were noted as presented and discussed above. Thus, it is reasonable to assume that the Poisson process adequately models the occurrences of Montreal heat spells for 12 out of 24 $Y^{u:v}$ time series. The validation of the Poisson process is a pre-requisite for the Bayesian change-point analysis technique presented in the next section.

BAYESIAN ANALYSES

Bayesian inference

Let θ be a population parameter and y the given data sample. The Bayesian inference consists of specifying three basic elements: (1) a prior probability distribution, $\pi(\theta)$, which summarizes our beliefs about the distribution of θ before we have the information from the data (its

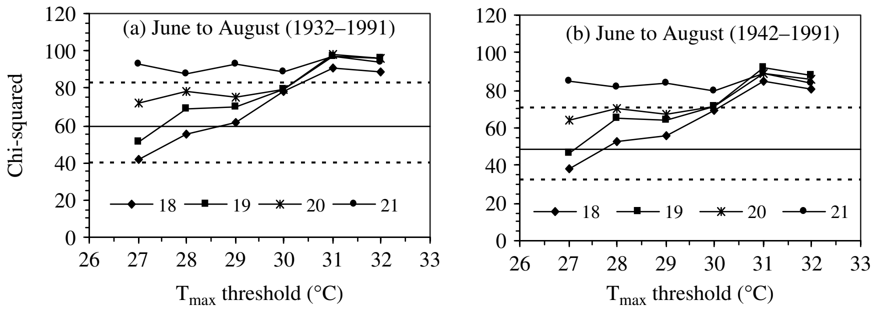

Figure 2. Calculated chi-squared values and their 95% confidence intervals (dotted lines) for the Poisson process model for selected combinations of $T_{\max}$ (shown on the x-axis) and $T_{\min}$ (shown in the legend of each panel) thresholds considering (a) 1932–1991 and (b) 1942–1991 summer-season (June–August) observations. Solid horizontal lines in both panels correspond to equality of mean and variance property of the Poisson distribution.

DOI: 10.1002/joc

choice is subjective and may vary from investigator to investigator), (2) the probability distribution of the data sample, y, given the parameter θ (i.e. the likelihood function) and (3) the posterior distribution, $p(\theta|y)$, of θ, which summarizes the information in the data, y, together with the information in the prior distribution, $\pi(\theta)$. The posterior distribution is obtained via the Bayes Theorem, i.e.

$$p(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta)\pi(\theta)d\theta} \tag{3}$$

Thus, we update the prior distribution to a posterior distribution after seeing the data via the Bayes Theorem. Any feature of the posterior distribution is legitimate for inference, including moments, quantiles, p-values, confidence intervals, etc. (Gelman *et al.*, 2003). The evaluation of the integral in the denominator (i.e. the normalizing constant) in Equation (3) is a source of practical difficulty. Analytic evaluation is usually impossible and numerical integration is difficult to apply if θ is a high-dimensional vector of parameters. The Markov chain Monte Carlo (MCMC) method is employed to solve this problem by exploiting relationships between conditional and marginal distributions to simplify the problem (Gelman *et al.*, 2003; Congdon, 2003). The MCMC procedure creates the opportunity to explore the use of Bayesian techniques in application areas that were previously thought of as impenetrable owing to the computations implied by Equation (3) (Coles, 2001). The term 'Monte Carlo' refers to drawing random numbers and the term 'Markov chain' refers to a series of values (a chain), of which any value depends only on the immediate prior one. A stationary (invariant) posterior distribution is guaranteed because the Monte Carlo sampling produces a Markov chain that does not depend on the starting value. A common MCMC algorithm is Gibbs sampling and the details of this can be found in Chen *et al.* (2000) and Gelman *et al.* (2003).

MCMC change-point analysis

A change-point occurs if, at some point in time k in the series, the annual counts of heat spells come from a Poisson distribution with a common rate up to that time, and from the same distribution but with a different rate thereafter. In other words, a change-point is a point in time that divides the Poisson rate process into independent epochs. The MCMC change-point algorithm consists of two steps. In the first step, we use the entire record to determine candidate change points on the basis of the expected value of the probability of a change as a function of time, which is presented later in this section. A large mean probability indicates a greater likelihood of a change-point. A plot of the mean probabilities as a function of time along with a minimum probability line (also presented later in this section) identifies the candidate change points against the hypothesis of no change points. The candidate change points have mean probabilities near or exceeding the minimum probability line. In the second step, we determine the posterior distributions of the Poisson rate parameter before and after the candidate change-point.

Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ be a vector of two parameters, where λ_1 is the mean Poisson rate before the change and λ_2 is the mean Poisson rate after the change, and we wish to simulate from the posterior $f(\boldsymbol{\lambda}|y)$. The hierarchical specification of the change-point model for the annual counts of heat spells (Y) is given as

$$Y_i \sim \text{Poisson}(\lambda_1) \text{ for } i = 1, \ldots, k \text{ and} \tag{4}$$

$$Y_i \sim \text{Poisson}(\lambda_2) \text{ for } i = k+1, \ldots, n = 96 \tag{5}$$

where $\lambda_1 \sim \text{gamma}(\alpha_1, \beta_1)$, $\lambda_2 \sim \text{gamma}(\alpha_2, \beta_2)$, k is the discrete uniform over $(1, \ldots, n)$, each independent, and $\beta_1 \sim \text{gamma}(\gamma_1, \varepsilon_1)$ and $\beta_2 \sim \text{gamma}(\gamma_2, \varepsilon_2)$ (Carlin *et al.*, 1992). The Poisson-gamma relationship is used because the annual counts could be assumed to follow a Poisson distribution for a range of combinations of $T_{\min}$ and $T_{\max}$ thresholds, as presented in the section 'Poisson Point Process', and the rate parameter of this distribution is usually taken as gamma distributed (Epstein, 1985). The specification is hierarchical because in stage one the annual counts are random values following a Poisson distribution with an unknown Poisson rate, and in stage two the Poisson rate follows a gamma distribution with two unknown parameters. In stage three the unknown scale parameter follows a gamma distribution, and in the fourth stage the parameters of the gamma distribution follow noninformative priors. The specification leads to the following conditional distributions used in the Gibbs sampling (Carlin *et al.*, 1992):

$$\lambda_1|Y, \lambda_2, \beta_1, \beta_2, k) \sim \text{gamma}\left(\alpha_1 + \sum_{i=1}^{k} Y_i, k + \beta_1\right) \tag{6}$$

$$\lambda_2|Y, \lambda_1, \beta_1, \beta_2, k) \sim \text{gamma}\left(\alpha_2 + \sum_{i=k+1}^{n} Y_i, n - k + \beta_2\right) \tag{7}$$

$$\beta_1|Y, \lambda_1, \lambda_2, \beta_2, k) \sim \text{gamma}(\alpha_1 + \gamma_1, \lambda_1 + \varepsilon_1) \tag{8}$$

$$\beta_2|Y, \lambda_1, \lambda_2, \beta_2, k) \sim \text{gamma}(\alpha_2 + \gamma_2, \lambda_2 + \varepsilon_2), \text{ and} \tag{9}$$

$$p(k|Y, \lambda_1, \lambda_2, \beta_1, \beta_2) = \frac{L(Y; k, \lambda_1, \lambda_2)}{\sum_{j=1}^{n} L(Y; j, \lambda_1, \lambda_2)} \tag{10}$$

where the likelihood function is given by

$$L(Y; k, \lambda_1, \lambda_2) = \exp\{k(\lambda_2 - \lambda_1)\}(\lambda_1/\lambda_2)^{\sum_{i=1}^{k} Y_i} \text{ and} \tag{11}$$

$p(k|Y, \lambda_1, \lambda_2, \beta_1, \beta_2)$ is the probability that year k is a change-point, given the data and the parameters. For the full derivation of the likelihood function, refer Carlin *et al.* (1992) and Elsner *et al.* (2004).

DOI: 10.1002/joc

Practical considerations for MCMC simulations

There are several points, e.g. the definition of the years relative to the change-point year, starting values of the model parameters, length of simulation (i.e. number of iterations involved in an MCMC simulation), burn-in period, etc., that need attention while implementing the MCMC change-point algorithm to determine change points and their statistical significance. In the model specification given above, it is stated that year k is the last year of the old epoch with $k+1$ being the first year of the new epoch. Throughout the analysis, the first year of the new epoch is referred to as the change-point year. Theoretically, the choice of initial values should not influence the final stationary (invariant) posterior distribution, if the chain is irreducible, meaning that any set of values results in a positive probability that the chain can reach any other set of values. Since the Poisson rate parameter λ is gamma(α, β) with mean α/β and variance α/β^2, the initial mean values for λ_1 and λ_2 are chosen equal to the climatological rate μ_Y, i.e. the average annual number of heat spells, $\gamma_1 = \gamma_2 = 0.1$, and $\varepsilon_1 = \varepsilon_2 = 1$ as the starting values, with the latter two being an educated guess. Thus, the mean values for β_1 and β_2 are $0.1/1 = 0.1$ and $\alpha_1 = \alpha_2 = \mu_Y/10$. It is useful to perform a number of simulations (i.e. independent runs of the change-point algorithm) with different starting values to check if the posterior distributions are sensitive to the choice of initial values. Results of a large number of simulations are presented and discussed in the section 'Results and Discussion' to verify this point. Unless indicated, throughout the analysis, each MCMC simulation (or run) of the change-point algorithm consists of 11 000 iterations. First 1000 iterations are discarded as burn-in and the last 10 000 iterations are used to estimate the posterior distributions. The choice of the length of burn-in is also discussed in the section 'Results and Discussion'.

The next point is related to the identification of candidate change points. If there are no change points in the observed counts of heat spells, the change-point probabilities should not be different from those based on counts arising from a Poisson distribution with a constant rate. One thousand random time series, each of length 96 (i.e. the length of record used in the present study), are generated from a Poisson distribution with a parameter equal to the climatological rate μ_Y. Ranking the posterior probabilities, the 995th (out of 1000) largest posterior probability is selected as the minimum probability necessary for identifying a candidate change-point year. This establishes a conservative baseline (i.e. an empirical confidence line) for identifying a candidate change-point when none exists. This approach can detect more than one candidate change-point. Once a change-point is established and qualified using formal statistical tests, the empirical confidence line is re-estimated for studying shorter records to identify additional change points. A similar approach has been adopted by Elsner *et al.* (2004) to identify candidate change points in time series of annual counts of hurricanes.

The last point is related to the posterior distributions of the mean Poisson rates before and after the change-point (i.e. λ_1 and λ_2) and their difference (i.e. $\lambda_2 - \lambda_1$), and the statistics obtained from them. The fraction of the posterior distribution of the difference in mean Poisson rates that is greater (or less) than 0 provides evidence of the direction of change and is termed p-value, and is often used in making statistical inferences. The p-value obtained this way is suitable for a one-sided alternate hypothesis if one had hypothesized in advance that the second epoch would have more (or less) events than the first epoch, and therefore a two-sided alternate hypothesis is more suitable for the present change-point problem because the change is observed after the data are collected. The mean rates λ_1 and λ_2 and p-value obtained from a single MCMC simulation are random variables, so additional simulations are required to obtain ensemble averaged values. Though the choice is subjective, one can employ any reasonable number of ensemble members to obtain ensemble averaged values. For the present problem of Montreal heat spells, we employ 100 ensemble members (i.e. additional runs of the change-point algorithm) to obtain ensemble averaged values. In addition to p-values, Bayesian Information Criterion (*BIC*, Schwarz, 1978) is used to confirm the presence of a change-point. The use of 'Bayes Factors' is a dominant method for Bayesian model assessment, but 'Bayes Factors' are often quite difficult to calculate and a suitable alternative is *BIC* (Gill, 2002). The *BIC* is defined as

$$BIC = -2\ln(L) + 2r\ln(n) \tag{12}$$

where L is the likelihood function, r is the number of model parameters and n is the record length. Smaller values of *BIC* are preferred in pair-wise nested modeling situations such as the change-point problem. If there is a change-point then two Poisson models are more suitable for the observed counts of heat spells, otherwise a single Poisson model is more likely. Thus, positive values of $\Delta_{BIC} = BIC_1 - BIC_2$ would favor the former situation. The subscript 1 and 2 on *BIC*, respectively, corresponds to one- and two-model situations.

RESULTS AND DISCUSSION

There is no single definition of what constitutes an extreme climate event. However, extreme climate events can be quantified from the viewpoint of their rarity, intensity, duration and impact on socioeconomic sectors (Beniston and Stephenson, 2004b). In the present work, we consider both $T_{\min}$ and $T_{\max}$ observations above reasonably high joint $T_{\min}$ and $T_{\max}$ thresholds in order to define heat spells as indicators of extreme climate. Other meteorological variables, such as relative humidity, could also be brought into the definition of a heat spell and that remains the subject of future studies. In the following, we present and discuss detailed results of the MCMC change-point algorithm applied to 12 $Y^{u:v}$ time series of annual

DOI: 10.1002/joc

counts of heat spells obtained from homogenized $T_{\min}$ and $T_{\max}$ observations. The use of homogenized observations rules out the possibility of artificially induced change-points because of logistic problems. Thus, any identified change-point, if it exists in recorded observations, could be considered to have at least an element of natural origin as a causative factor toward higher/lower activity of heat spells.

In a hierarchical Bayesian setting such as the present case, in which a large number of iterations of the MCMC algorithm are involved to obtain posterior distributions of model parameters, it is important to examine convergence of the algorithm, i.e. how quickly the chain converges to stationary distributions. To investigate this point, the $Y^{20:28}$ time series is selected and the analysis is run using the initial conditions prescribed in the section 'Bayesian Analyses'. Values of λ_1 and λ_2 are plotted in Figure 3 considering 1931 as the change-point year. It is evident from the plotted values that the chain quickly converges to the stationary distributions and there does not appear to be any trend or cycle in the generated values. As mentioned in section 'Bayesian Analyses', the first 1000 iterations are considered as burn-in to account for the effect of starting values, but the plots of Figure 3 show that even a smaller number (e.g. 100 or 200 iterations) is enough for this purpose. The use of 1000 iterations, as burn-in, guarantees that the starting point has no influence on the posterior distributions. These verification results help one to confidently derive conclusions from the posterior distributions of the model parameters. To further investigate the effect of burn-in, we ran the change-point algorithm for 15 000 iterations for three independent runs for the same $Y^{20:28}$ time series. The values of (λ_1, λ_2) obtained from the last 10 000 iterations, for each of the three independent runs, are (3.26, 4.88). These values of the Poisson rate parameters are not very different from the ones shown in Figure 3. This confirms that our choice of 1000 iterations as burn-in and 10 000 iterations for estimating posterior distributions is reasonable.

The results from the MCMC change-point algorithm applied to 12 $Y^{u:v}$ time series of annual counts of heat spells are shown in Figure 4, in which the expected value of the probability that the year k is a change-point, given the data and the parameters, i.e. $p(k|Y, \lambda_1, \lambda_2, \beta_1, \beta_2)$, is plotted as a function of year. The expected value represents the average posterior probability of the year k being the first year of a new epoch. Large probabilities indicate that a likely change occurred beginning with year k. The empirical confidence line (i.e. the dashed line, which is smoothed using a 5th degree polynomial smoother; the choice of smoother is not a critical issue in the present context) shows the minimum probability necessary for detecting a change-point based on the Poisson process model of constant rate parameter. Note that years near the beginning and end of the record require substantially larger posterior probabilities to surpass the empirical confidence line. The U-shaped confidence line indicates that there is larger variance on the posterior change-point probabilities near the beginning and end of the record. This is because of the small number of values used to estimate them near the end points. Years with the highest average posterior probabilities (APPs) above the empirical confidence line are indicated in Figure 4. The algorithm empirically chooses (i) 1929 for $Y^{18:27}$, $Y^{18:28}$ and $Y^{18:29}$ time series, (ii) 1931 for $Y^{19:27}$, $Y^{19:28}$, $Y^{19:29}$, $Y^{19:30}$, $Y^{20:27}$, $Y^{20:28}$, $Y^{20:29}$ and $Y^{20:30}$ time series and (iii) either 1906 or 1929 for $Y^{18:30}$ time series as the most likely year of the new epoch.

In the majority of the $Y^{u:v}$ time series, there are several years, near the most likely year of the new epoch, which are also candidate change points. This indicates that there is uncertainty associated with the most likely change-point year, i.e. which one to select as the change-point year. Only in two cases (i.e. $Y^{19:28}$ and $Y^{20:27}$) are the additional change-point years, which are located near the empirical confidence line, far away from the most likely change-point years noted in the above paragraph. One way to investigate the uncertainty associated with the most likely change-point year is to examine the influence of the choice of initial values on the posterior change-point probability for a cluster of candidate years. Indirectly, this will also reveal the robustness of the

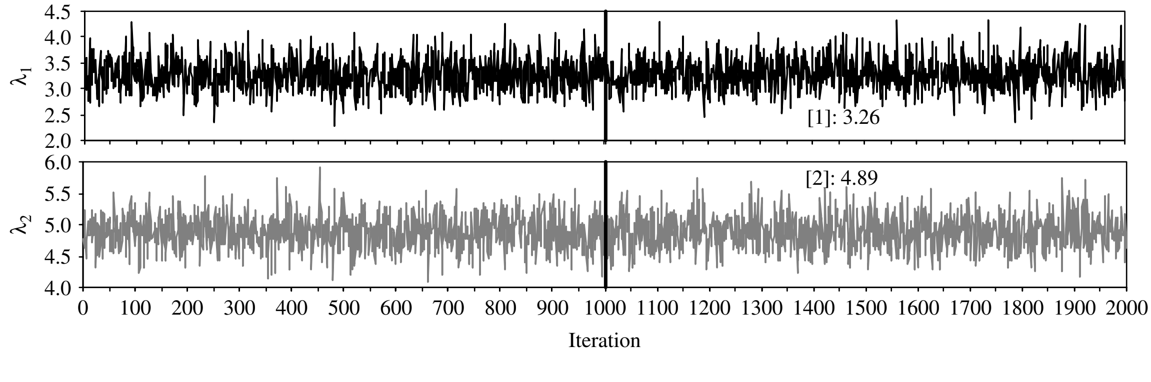

Figure 3. Values of the Poisson rates λ_1 (upper panel) and λ_2 (lower panel) for the first 2000 iterations obtained from a single run of the change-point analysis algorithm applied to $Y^{20:28}$ time series considering 1931 as the change-point year. Average values of the parameters (i.e. [1] and [2]) obtained from the last 10 000 iterations (out of 11 000) are shown in each panel. First 1000 iterations, left of the thick vertical line, are discarded as burn-in.

DOI: 10.1002/joc

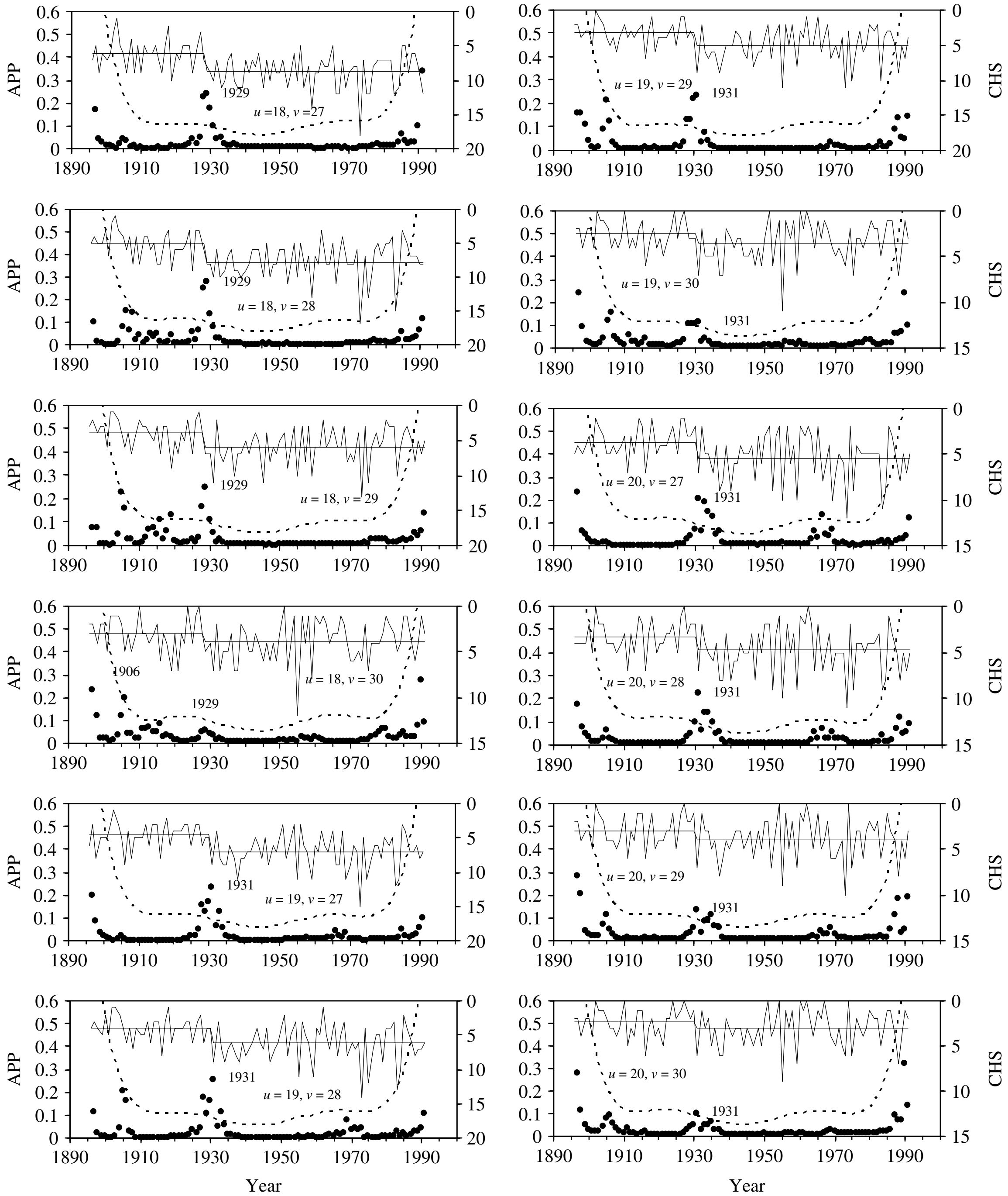

Figure 4. Average posterior probabilities (APPs) of each year being the first year of a new epoch (filled circles) for $Y^{u:v}$ time series of annual counts of heat spells (CHS; plotted against the right hand side y-axes) obtained from various combinations of $T_{\min}$ (u) and $T_{\max}$ (v) thresholds. The dashed line represents the empirical confidence line to identify candidate change points. Horizontal thin solid lines represent the change in level of the $Y^{u:v}$ time series.

change-point identification algorithm. For this purpose, $Y^{18:27}$ time series is selected and the Gibbs sampling is run 100 times for each value of the climatological rate μ_Y equal to 4, 6 and 8 numbers of spells. These values of μ_Y translate into alpha priors (i.e. α_1 and α_2) that are each equal to 0.4, 0.6 and 0.8. The APPs obtained from 100 independent runs of the change-point algorithm are plotted in Figure 5 for the four candidate change-point years (i.e. 1928, 1929, 1930 and 1931). In this figure, the range (minimum to maximum) of APPs is shown by plotting them against the right hand side y-axis for three out of four candidate change-point years because it is hard to see their variability when plotted against the left hand side y-axis. Three obvious conclusions can be drawn from Figure 5: (1) the average posterior probability of the most likely change-point year is always higher than its competing years, (2) the distribution of the average posterior probability has a very small spread (the variability of the average posterior probability could have been large for the candidate years located near

DOI: 10.1002/joc

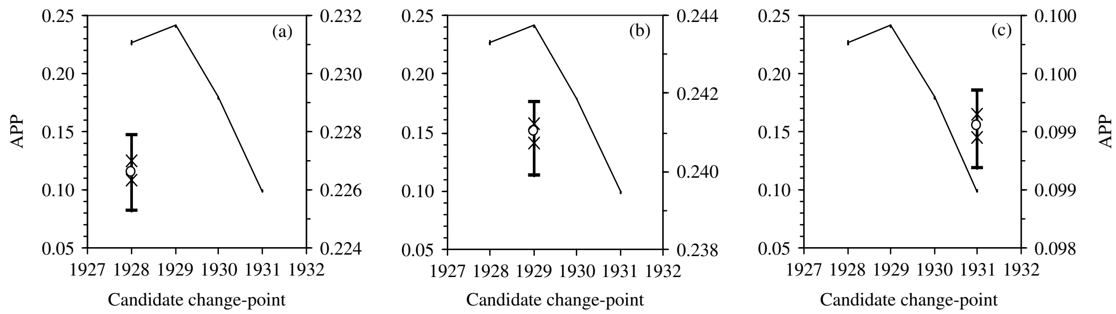

Figure 5. Average posterior probabilities (APPs) obtained from 100 independent runs (each run consisted of 10 000 iterations after the burn-in) of the change-point algorithm for a cluster of candidate change-point years (i.e. 1928, 1929, 1930 and 1931) for the $Y^{18:27}$ time series. For Gibbs sampling, starting values of the climatological mean rate μ_Y were taken equal to (a) 4, (b) 6 and (c) 8 numbers of spells. Starting values of the other parameters are indicated in the section 'Bayesian Analyses' of the paper. Thin lines are drawn through medians of APPs. Thick vertical lines correspond to the range (minimum to maximum) of APPs and the symbols correspond to 25th, 50th and 75th percentiles plotted against the right hand side y-axis (note the scale difference in each panel).

the beginning and end of the record because of the U-shaped nature of the empirical confidence line), and (3) the choice of prior values is not a critical factor in identifying the candidate change-point year as the chain quickly finds a stationary distribution regardless of where the chain is started. Thus, in the case of $Y^{18:27}$ time series, it can be stated that the change in the Poisson rate parameter occurred as early as 1928 and as late as 1931, with 1929 as the most likely year of change. Similar statements could be made in the cases of other time series where clusters of candidate change-point years exist. We remark here that when $Y^{18:27}$ data were analyzed using a graphical approach, i.e. using plots of cumulative departures from the mean (Dahmen and Hall, 1990), the most likely change-point year, 1929, which has the highest average posterior probability in the Bayesian approach, proved to be the first year of the new epoch and similar results were observed for most of the remaining $Y^{u:v}$ time series. Thus, it is safe to conclude that the year with the largest average posterior probability, being at or above the empirical confidence line, could be considered as the change-point year, i.e. the start of a new epoch of heat spells. These empirically observed change points need to be confirmed using additional confirmatory analyses given below.

From a probabilistic point of view, additional confirmatory analysis can be made by comparing the conditional posterior distributions of the Poisson rate parameters (i.e. λ_1 and λ_2) before and after the change-point and estimating the p-value from the distribution of the rate difference (i.e. $\lambda_2 - \lambda_1$). Posterior distribution estimates of λ_1 and λ_2 from Gibbs sampling, along with estimates of means of λ_1 and λ_2 and their 95% confidence intervals, are shown in Figure 6 for $Y^{19:29}$ and $Y^{20:30}$ time series. For $Y^{19:29}$ time series, the 95% confidence intervals of mean rates indicate no overlap and posterior distributions of the rate parameters show a little overlap implying a rate increase beginning with the year 1931. The ensemble averages of the mean Poisson rates before and after the change-point, i.e. $\langle\lambda_1\rangle$ and $\langle\lambda_2\rangle$, are, respectively, 3.29 and 5.10. The higher ensemble average of the mean rate after the change-point confirms the direction of change, i.e. an increased activity of heat spells. A small ensemble average of p-value on the rate increase (i.e. $\langle p\text{-value}\rangle = 0.003$) and positive change in *BIC* (i.e. $\langle\Delta^{a}_{BIC}\rangle = 8.11$ (Table III)) confirm that the change is significant. Figure 6(b) for the $Y^{20:30}$ time series indicates that the 95% confidence intervals of mean rates overlap and so do the posterior distributions of the rate parameters, implying a nonsignificant rate increase beginning with the year 1931. The $\langle p\text{-value}\rangle = 0.114$

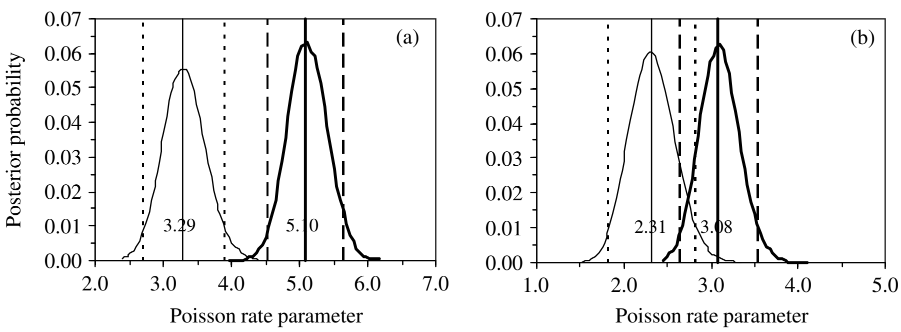

Figure 6. Posterior distribution estimates (smoothed versions of histograms) from the Bayesian change-point algorithm applied to (a) $Y^{19:29}$ and (b) $Y^{20:30}$ time series (1896–1991) considering 1931 as the change-point year. Distributions of Poisson rates λ_1 (thin solid line) before and λ_2 (thick solid line) after the change-point are shown. Mean values of λ_1 (thin vertical line) and λ_2 (thick vertical line) along with corresponding 95% confidence intervals (small dotted and large dotted vertical lines, respectively) are also shown.

DOI: 10.1002/joc

Table III. Ensemble averaged estimates ($\langle . \rangle$) of mean Poisson rates before and after the given change-point (CP), ensemble averaged two-sided p-values and ensemble averaged estimates of Δ_{BIC} for 12 $Y^{u:v}$ time series of annual counts of heat spells. Δ_{BIC}^{a} and Δ_{BIC}^{b} were respectively estimated from the average values of the posterior estimates of parameters and averages of the likelihood functions computed over 10 000 iterations of the Gibbs sampler. Change points for the cases in bold are likely to be insignificant.

$Y^{u:v}$	CP	$\langle \lambda_1 \rangle$	$\langle \lambda_2 \rangle$	$\langle p\text{-value} \rangle$	$\langle \Delta_{BIC}^{a} \rangle$	$\langle \Delta_{BIC}^{b} \rangle$	$Y^{u:v}$	CP	$\langle \lambda_1 \rangle$	$\langle \lambda_2 \rangle$	$\langle p\text{-value} \rangle$	$\langle \Delta_{BIC}^{a} \rangle$	$\langle \Delta_{BIC}^{b} \rangle$
$Y^{18:27}$	1929	6.00	8.73	<0.001	12.50	11.38	$Y^{19:29}$	1931	3.29	5.10	0.003	8.11	6.14
$Y^{18:28}$	1929	5.06	7.75	<0.001	14.92	13.80	$Y^{19:30}$	1931	2.54	3.56	**0.051**	**−1.70**	**−2.69**
$Y^{18:29}$	1929	3.91	6.03	<0.001	10.23	9.12	$Y^{20:27}$	1931	3.57	5.51	0.002	9.04	7.08
$Y^{18:30}$	1929	2.20	3.74	**0.093**	**−3.68**	**−4.67**	$Y^{20:28}$	1931	3.26	4.89	0.007	5.19	3.22
$Y^{19:27}$	1931	4.60	7.08	<0.001	14.06	12.95	$Y^{20:29}$	1931	2.86	3.97	**0.043**	**−1.18**	**−2.17**
$Y^{19:28}$	1931	3.94	6.30	<0.001	14.63	13.52	$Y^{20:30}$	1931	2.31	3.08	**0.114**	**−4.32**	**−5.30**

and $\langle \Delta_{BIC}^{a} \rangle = -4.32$ confirm that the rate increase is nonsignificant. For all the $Y^{u:v}$ time series, estimates of $\langle \lambda_1 \rangle$, $\langle \lambda_2 \rangle$, $\langle p\text{-value} \rangle$ and $\langle \Delta_{BIC} \rangle$ are given in Table III. The results of this table confirm a rate increase beginning with the most likely change-point year, which varies with respect to $Y^{u:v}$ time series (Figure 4), and is the very first point of significant change (for most of the $Y^{u:v}$ time series) during the history of recorded observations. Note that the confirmatory analyses of p-values do not approve the presence of change points for $Y^{18:30}$, $Y^{19:30}$ and $Y^{20:30}$ time series at a 5% significance level. In addition to these time series, the confirmatory analysis of Δ_{BIC} rejects the presence of a change-point for $Y^{20:29}$ time series as well. Thus, from the results presented and discussed so far, it is evident that there are structural inhomogeneities within the Montreal heat spells. The Bayesian change-point algorithm points out that, most likely, the change toward an increased activity of heat spells took place around 1930 for most of the $Y^{u:v}$ time series and there was no change in the occurrence of extreme heat spells. According to Environment Canada (Vincent, 2006; personal communications), both $T_{\min}$ and $T_{\max}$ temperatures were homogenized separately for known logistic problems and joint observations were not investigated. Because of exposure problems, the instruments were relocated after 1933 and this was taken care of when developing homogenized data. Thus, the identified change points are likely to be real and are not due to logistic problems.

We also calculated the monotonic time trends, discussed in the section 'Nonparametric Trend Analyses and Split-sample Comparisons', considering the observations of annual counts of heat spells after and including the above established change points (Table III). Sixty-seven percent $Y^{u:v}$ time series showed negative but insignificant trend and 33% showed (nearly zero) positive but insignificant trend. Thus, statistically significant increasing trends noted in the section 'Nonparametric Trend Analyses and Split-sample C' for 9 (out of 12) $Y^{u:v}$ time series did not show up when counts of heat spells after and including the identified change points were considered. If one wishes to estimate the time trends from the data that are free of change points, then this could be accomplished by either developing residual time series or by bringing the pre-change-point epoch at par with the post-change-point epoch by appropriately adjusting pre-change-point data. A similar technique has been suggested by Lund and Reeves (2002), which we did not attempt to implement here because the main focus of our study is on change points.

After having identified the above indicated change points (Table III), we applied the MCMC change-point algorithm to eight $Y^{u:v}$ time series (i.e. $Y^{18:27}$, $Y^{18:28}$, $Y^{18:29}$, $Y^{19:27}$, $Y^{19:28}$, $Y^{19:29}$, $Y^{20:27}$ and $Y^{20:28}$) considering the observations of annual counts of heat spells after and including the year of significant change. The same procedure as explained earlier was followed to identify the candidate change points. No potential candidate change points are found for any of these $Y^{u:v}$ time series.

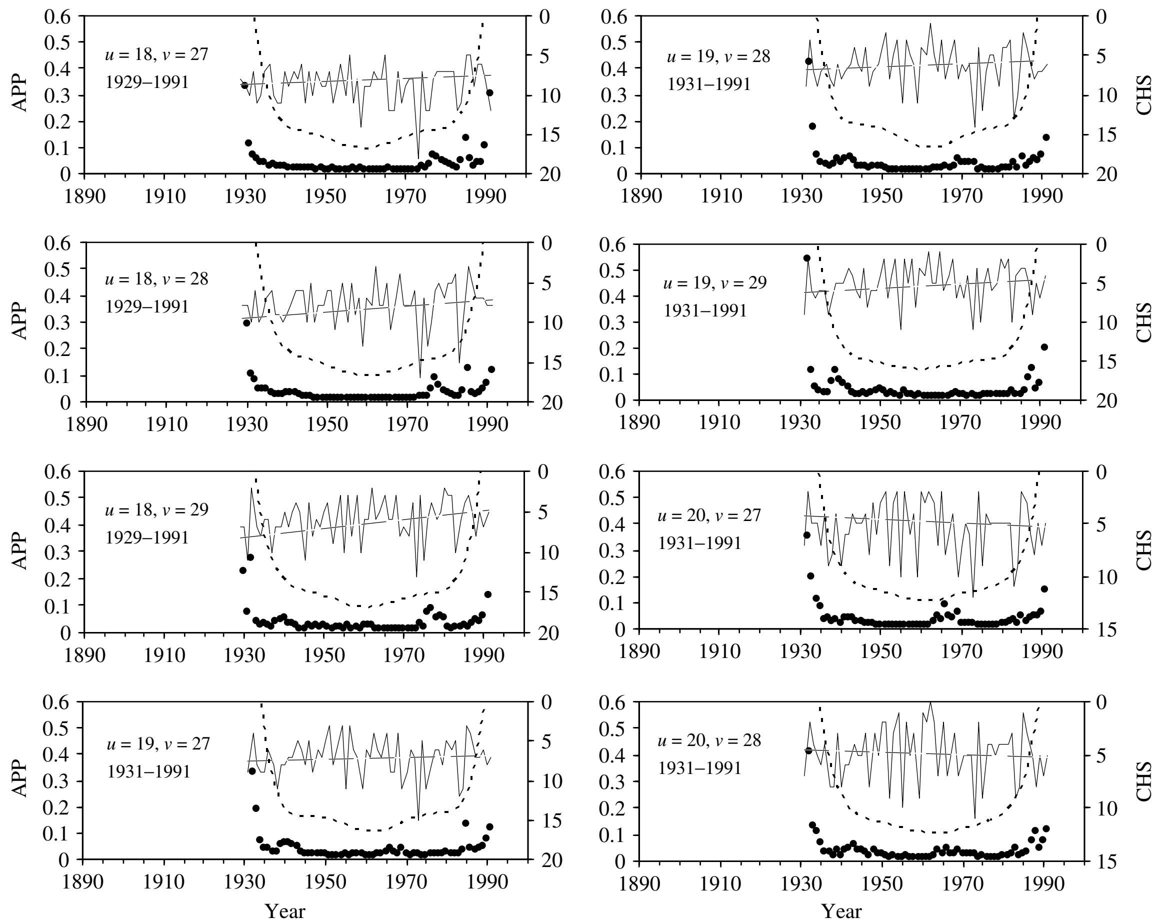

Figure 7. Same as in Figure 4 but for considering annual counts of heat spells after, and including, the most likely identified change-point indicated in Table III and Figure 4. Large dotted line represents the time trend plotted against the right hand side y-axis.

Detailed plots are shown in Figure 7. One of these time series (i.e. $Y^{20:27}$) is examined in detail for a possible change-point at year 1966, where a locally elevated jump in average posterior probability is visible. The distributions of λ_1 and λ_2 overlap each other as well as the 95% confidence intervals around the mean rates (figure not shown). The ensemble averaged p-value is 0.22 and $\langle \Delta^{a}_{BIC} \rangle$ is -5.21. This leads to the conclusion that 1966 is not really a change-point.

On the basis of the above results of the MCMC change-point algorithm applied to $Y^{u:v}$ time series considering observations of annual counts of heat spells for the years after, and including, the very first change-point year identified in the history of observations (Table III), it can be concluded that there is inadequate evidence of increased activity of heat spells at Montreal during the third and second last decades (i.e. 1970s and 1980s) of the 20th century. It would have been interesting to identify the tendency of heat spells during the last decade of the 20th century and the beginning of the first decade of the 21st century, but our analysis spanned the period 1896–1991 and we did not consider the latter observations as there were 5 years of missing/inadequate $T_{\min}$ and $T_{\max}$ observations during the 1990s. Future studies should address this point by appropriately filling in the missing observations.

SUMMARY AND CONCLUSIONS

In this study, annual counts of heat spells observed at Montreal are analyzed to identify points of abrupt change in the mean rate of their occurrence using a hierarchical Bayesian change-point approach based on the assumption that the occurrence of heat spells follows a Poisson point process. We are unaware of any other study on heat spells conducted in a manner as in the present work, in which a total of 24 time series of annual counts of heat spells are derived from a single station homogenized $T_{\min}$ and $T_{\max}$ observations by employing suitably chosen and reasonably high thresholds–four thresholds for the $T_{\min}$ (i.e. 18, 19, 20 and 21 °C) and six for the $T_{\max}$ (i.e. 27 to 32 °C). Twelve of these time series, which fulfill the assumption of the Poisson point process at 95% confidence level, are subjected to Bayesian change-point analysis. In addition to identifying the potential time point of change and its direction, the Bayesian change-point approach directly provides the estimates of uncertainty on the mean rate of occurrence of heat spells before and after the change-point. The following main conclusions and recommendations can be made from the analyses presented in this paper:

- The Bayesian change-point analysis approach proved to be a useful tool in identifying significant shifts in

DOI: 10.1002/joc

the mean rate of occurrence of heat spells at a priori unknown time points and is able to detect multiple change points. This confirms the findings of Elsner *et al.* (2004), who successfully applied the technique to identify temporal shifts in the rates of occurrence of North Atlantic and US hurricane activities, and Chu and Zhao (2004), who applied a similar Bayesian technique to study tropical cyclone activity in Central North Pacific.

- By employing various combinations of $T_{\min}$ and $T_{\max}$ thresholds to define heat spells, it has been demonstrated in this study that there are structural inhomogeneities within the heat spell observations, i.e. all time series of annual counts of heat spells, obtained from the same $T_{\min}$ and $T_{\max}$ observations by employing various combinations of their joint thresholds, do not show abrupt changes in the mean rate of occurrence. By using just a single threshold for the $T_{\min}$ and a single threshold for the $T_{\max}$ to define heat spells, identification of such inhomogeneities would not have been possible. Thus, the advantage of analyzing more than just one time series of annual counts of heat spells at a single location is obvious.
- The overall results of the change-point analyses suggest that there is inadequate evidence in favor of increased activity of heat spells in Montreal during the second last decade (i.e. 1980s) of the 20th century, which is the most recent observation period analyzed. Without implementing the change-point analysis approach, one could have inadvertently concluded that there is an increased activity (being statistically significant in 9 out of 12 time series of annual counts of heat spells) of heat spells in Montreal on the basis of commonly used nonparametric trend detection and estimation techniques, whereas the opposite is more likely because more than half the number of time series of annual counts of heat spells show statistically insignificant decreasing trends after the very first detected change-point in the chronology of recorded observations. This finding shows that the change-point analyses, whether Bayesian or non-Bayesian, should be used in conjunction with tests used to identify monotonic time trends in order to develop future climate change projections based on the analysis of recorded observations.
- Finally, the Bayesian change-point analyses presented in this paper can be extended to the most recent observation periods by inserting the missing $T_{\min}$ and $T_{\max}$ values for the years 1992–1996 using appropriate techniques, such as those discussed in Eischeid *et al.* (1995). Work in this direction is highly encouraged in order to establish the turning points, if there are any, in the rate of occurrence of heat spells in Montreal during the most recent years of the 21st century. Furthermore, it will be interesting to verify if the presence of structural inhomogeneities in the counts of heat spells is a characteristic of just Montreal data or is a regional behavior. This could be accomplished by analyzing the $T_{\min}$ and $T_{\max}$ observations from the surrounding stations having long observational records in a manner similar to the present study.

ACKNOWLEDGEMENTS

The financial support provided by Ouranos Consortium on Regional Climatology and Adaptation to Climate Change and NSERC is gratefully acknowledged. The access to Environment Canada homogenized temperature dataset is appreciated. Dr Laxmi Sushama of the Canadian Regional Climate Modeling group is acknowledged for reviewing an earlier version of the paper. We thank two anonymous referees for their valuable comments that contributed toward an improved version of the paper.

REFERENCES

Beniston M, Diaz HF. 2004. The 2003 heat wave as an example of summers in a greenhouse climate? Observations and climate model simulations for Basel, Switzerland. *Global and Planetary Change* **44**: 73–81.

Beniston M, Stephenson DB. 2004. Extreme climatic events and their evolution under changing climatic conditions. *Global and Planetary Change* **44**: 1–9.

Carlin BP, Gelfand AE, Smith AFM. 1992. Hierarchical Bayesian analysis of change-point problems. *Applied Statistics* **41**: 389–405.

Chen M-H, Shao Q-M, Ibrahim JG. 2000. *Monte Carlo Methods in Bayesian Computation*. Springer: New York.

Chu P-S, Zhao X. 2004. Bayesian change-point analysis of tropical cyclone activity: the Central North Pacific case. *Journal of Climate* **17**: 4993–4901.

Coles S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer: London.

Congdon P. 2003. *Applied Bayesian Modelling*. Wiley: New York.

Cox DR, Isham VS, Northrop PJ. 2002. Floods: some probabilistic and statistical approaches. *Philosophical Transactions of the Royal Society of London, Series A* **360**: 1389–1408.

Dahmen ER, Hall MJ. 1990. *Screening of Hydrological Data, ILRI Publication No. 49*. International Institute for Land Reclamation and Improvement (ILRI): Netherlands.

Drouin L, King N, Jacques L, Fortier I, Roy LA, Litvak E, Simard JO, Kosatsky T. 2005. The response of the Montreal Public Health Board to climate change: preventing access morbidity and mortality due to extreme summer temperatures in vulnerable human populations. In *The Proceedings of the Conference "Adapting to Climate Change in Canada 2005: Understanding Risks and Building Capacity"*, Montreal, Quebec; 4 to 7 May, 2005; 64.

Eischeid JK, Baker CB, Karl TR, Diaz HF. 1995. The quality control of long-term climatological data using objective data analysis. *Journal of Applied Meteorology* **34**: 2787–2795.

Elsner JB, Bossak BH. 2001. Bayesian analysis of US hurricane climate. *Journal of Climate* **14**: 4341–4350.

Elsner JB, Niu X, Bossak BH. 2004. Detecting shifts in hurricane rates using a Markov chain Monte Carlo approach. *Journal of Climate* **17**: 2652–2666.

Epstein ES. 1985. *Statistical Inference and Prediction in Climatology: A Bayesian Approach, Meteorological Monograph No. 42*. American Meteorological Society: Boston.

Gelman A, Carlin JB, Stern HS, Rubin DB (eds). 2003. *Bayesian Data Analysis*. Chapman and Hall/CRC: Boca Raton.

Gill J. 2002. *Bayesian Methods: A Social and Behavioral Sciences Approach*. Chapman and Hall/CRC: Boca Raton.

Guttorp P. 1995. *Stochastic Modeling of Scientific Data*. Chapman and Hall: London.

Hoaglin DC, Mosteller F, Tukey JW. 1983. *Understanding Robust and Exploratory Data Analysis*. Wiley: New York.

Johnson NL, Kotz S, Kemp AW. 1992. *Univariate Discrete Distributions*, 2nd edn. Wiley: New York.

Katz RW. 2002. Stochastic modeling of hurricane damage. *Journal of Applied Meteorology* **4**: 754–762.

Kendall MG. 1975. *Rank Correlation Methods*. Charless Griffin: London.

DOI: 10.1002/joc

Khaliq MN, St-Hilaire A, Ouarda TBMJ, Bobée B. 2005. Frequency analysis and temporal pattern of occurrences of Southern Quebec heatwaves. *International Journal of Climatology* **25**: 485–504.

Khaliq MN, Gachon P, St-Hilaire A, Ouarda TBMJ, Bobée B. 2006. Southern Quebec (Canada) summer-season heat spells over the 1941–2000 period: An assessment of observed changes. *Theoretical and Applied Climatology*. DOI 10.1007/s00704-006-0232-x.

Kunkel KE, Changnon SA, Reinke BC, Arritt RW. 1996. The July 1995 heat wave in the Midwest: a climate perspective and critical weather factors. *Bulletin of the American Meteorological Society* **77**: 1507–1518.

Lanzante JR. 1996. Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. *International Journal of Climatology* **16**: 1197–1226.

Lund R, Reeves J. 2002. Detection of undocumented change points: a revision of the two-phase regression model. *Journal of Climate* **15**: 2547–2554.

Rodriguez-Iturbe I, Cox DR, Isham V. 1987. Some models for rainfall based on stochastic point processes. *Proceedings of the Royal Society of London, Series A* **410**: 269–288.

Rootzén H, Tajvidi N. 1997. Extreme value statistics and wind storm losses: a case study. *Scandinavian Acturial Journal* **1997**(1): 70–94.

Salas JD, Delleur JW, Yevjevich VM, Lane WL. 1980. *Applied Modeling of Hydrologic Time Series*. Water Resources Publications: Littleton, CO; 484.

Schär C, Vidale PL, Luthi D, Frei C, Haberli C, Liniger MA, Appenzeller C. 2004. The role of increasing temperature variability in European summer heatwaves. *Nature* **427**: 332–335.

Schwarz G. 1978. Estimating the dimension of a model. *The Annals of Statistics* **6**: 461–464.

Sen PK. 1968. Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association* **63**: 1379–1389.

Siegel S, Castellan NJ Jr. 1988. *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill: New York.

Solow AR. 1987. Testing for climate change: An application of the two-phase regression model. *Journal of Applied Meteorology* **26**: 1401–1405.

Vincent LA, Zhang X, Bonsal BR, Hogg WD. 2000. Homogenized daily temperatures for trend analyses in extremes over Canada. In *Proceedings of the 12th Conference on Applied Climatology*, Ashville, NC, American Meteorological Society 86–89.

Vincent LA, Zhang X, Bonsal BR, Hogg WD. 2002. Homogenization of daily temperatures over Canada. *Journal of Climate* **15**: 1322–1334.

WHO. 2003. *The Health Impacts of 2003 Summer Heat-Waves*. Briefing note for the delegations of the fifty-third session of the WHO (World Health Organization) Regional Committee for Europe.

Zelenhasić E, Salvai A. 1987. A method of streamflow drought analysis. *Water Resources Research* **23**(1): 156–168.