Environmental Modeling & Assessment (2022) 27:385–398
https://doi.org/10.1007/s10666-021-09797-z

Climate Change: Use of Non-Homogeneous Poisson Processes for Climate Data in Presence of a Change-Point

Jorge Alberto Achcar[1] · Ricardo Puziol de Oliveira[2]

Received: 10 November 2020 / Accepted: 23 August 2021 / Published online: 15 September 2021

Abstract

In this study, non-homogeneous Poisson processes (NHPP) are used to analyze climate data. The data were collected over a certain period time and consist of the yearly average precipitation, yearly average temperature and yearly average maximum temperature for some regions of the world. Different existing parametric forms depending on time and on unknown parameters are assumed for the intensity/rate function $\lambda(t), t \geq 0$ of the NHPP. In the present context, the Poisson events of interest are the numbers of years that a climate variable measurement has exceeded a given threshold of interest. The threshold corresponds to the overall average measurements of each climate variable taking into account here. Two versions of the NHPP model are considered in the study, one version without including change points and one version including a change point. The parameters included in the model are estimated under a Bayesian approach using standard Markov chain Monte Carlo (MCMC) methods such as the Gibbs sampling and Metropolis–Hastings algorithms. The models are applied to climate data from Kazakhstan and Uzbekistan, in Central Asia and from the USA obtained over several years.

Keywords Temperature and precipitation data · Threshold exceedances · Bayesian inference · Change-point · MCMC methods

1 Introduction

Changes in climate either at global or at local level are monitored by following the behavior of climate variables such as precipitation volume and temperature as well as ocean levels, among others, paying special attention to the occurrence of events that deviate from the expected behavior (for instance, precipitation volume and/or temperature that are either higher or lower than the average value). The changes in climate have been observed throughout the world since the end of the nineteenth century (see, for instance, [1] and https://climate.nasa.gov/evidence/). It is possible to observe a significant increase in the average world temperature specially since the 1950s (see https://climate.nasa.gov/evidence/). For instance, nowadays we have an increase of 1.5 degrees Celsius in the average world temperature when compared to the pre-industrial era [1]. Additionally, climate events that deviate from the expected behavior are registered more and more frequently around the globe. One of the possible reasons for these changes could be the increase in the human made emissions into the atmosphere of carbon dioxide and other pollutants (see, for instance, https://www.ncdc.noaa.gov/monitoring-references/faq/indicators.php).

In addition to impacting the occurrence of climate events deviating from the average behavior, climate change has also a serious impact on ecosystems around the world, some of them very fragile. For instance, [2] presents several works studying the impact of climate change on forests in the USA. In [3], it is presented the effects of climate change on marine systems, and [4] introduces a review describing the impacts of the climate change in British Columbia diversity. In particular, changes in the temperature and/or precipitation patterns may have serious consequences in food production. Too little/much rain, high/low temperature and dry spells at the wrong time may affect crop production and meat production. Additionally, we may have economic and human consequences due to the possible forest fires which might reach households and cause economical losses and risk human lives.

✉ Ricardo Puziol de Oliveira
rpuziol.oliveira@gmail.com

Jorge Alberto Achcar
achcar@fmrp.usp.br

1 Medical School, University of São Paulo, São Paulo, Brazil

2 Department of the Environment, Maringá State University, Maringá, Brazil

Therefore, due to its serious consequences on the lives of the planet, it is very important to study climate change and the occurrence of climate events deviating from the expected behavior. Going in this direction, many statistical models are introduced in the literature dealing in some way with the problems related to this change. Among these models, [5] introduces a stochastic model, assuming different scenarios, which is used to estimate, for example, the probabilities of daily changes in precipitation. In another study, [6] considers stochastic differential models to study climate change problems and [7] presents a non-parametric renewal model for modeling daily precipitation data. A Markov model is also used by [8] to analyze precipitation data and [9] provides an overview of stochastic climate theory from the point of view of applied mathematics. Many other statistical models to infer the years that climatic anomalies occur are also presented in the literature [10–13].

Despite the existence of a very large number of papers using different statistical models introduced in the literature in recent years to study the climate changes, other statistical modeling alternatives may be of interest in the study of changes in precipitation and temperature patterns. Thus, as the main goal of this study, we consider statistical models in the presence of change-points assumed as unknown parameters that should be estimated from the climate data to accurately detect these changes. The presence of one or more change points is very common in time series data derived from many areas of application, for instance, epidemiology, finance, environment, medicine and many others, see, for instance, [14–18], among others. In general, we have change-points when either there is an intervention on the experiment whose results are being recorded or when there is an occurrence of a natural event, related to the data, that abruptly changes the behavior of the time series. This change in behavior may occur to any type of continuous or discrete data. In the case of discrete measurements a common type of change is in the counting of the occurrence of events, for instance, changes in the occurrence of threshold exceedances and number of hospitalizations, among others.

In addition to estimating specific change points, we may also be interested in the number of times a climate variable deviates from its average behavior over a fixed period of time. The change-points, as pointed out above, can be estimated by considering models in the presence of parameters denoting the change-points. If our interest is the number of times the average annual amount of precipitation and temperature exceeds their corresponding mean values over a given time interval, a natural choice for statistical modeling is given by counting processes and in particular non-homogeneous Poisson processes (NHPP). Even though in the NHPP formulation an assumption of independent inter-occurrence times is implicit, this type of model has provided good approximation for many problems, of similar nature as those considered here, that were studied in the past while allowing to calculate specific quantities of interest such as the probability of the number of occurrences in a given time interval.

NHPP have been used to study problems in several areas where problems of similar nature as those considered here are analyzed, see, for instance, in reliability theory [19–21], counting marine species [22], air pollution [14, 16], community noise pollution [23], in medicine [15], among many other areas. One common feature of those models is the assumption of several forms for the rate function associated with the NHPP. These rate functions may dependent not only on time, but also on some parameters that need to be estimated.

Different approaches have been considered in order to estimate the parameters of the proposed models including the location of the change-point. Many works, presented in the literature, consider this point especially under a Bayesian approach. This inference approach has some computational advantages when compared with standard maximum likelihood methods especially when dealing with non-homogeneous Poisson processes and, in particular, in the presence of change-points since the likelihood function may have a very complex form. Other advantage of the Bayesian approach is that we could incorporate prior opinions of experts leading to more accurate inference.

Bayesian inference for either homogeneous or non-homogeneous Poisson processes has been discussed by many authors such as [21, 24–27]. Those processes have also been used to obtain inference for change-point models [25, 28–30]. Raftery and Akman [31] consider a Bayesian analysis for homogeneous Poisson processes in the presence of a change-point. Ruggeri and Sivaganesan [32] and [33] consider a Bayesian analysis dealing with a random number of change-points. In the former work, NHPPs with the so-called power law processes (PLP) as rate function is considered and in the latter a stepwise constant rate function is used.

Within the Bayesian framework the estimation of the parameters may be performed using MCMC methods [34–36]. We follow this path when we develop a Bayesian analysis assuming different parametric structures for the rates in non-homogeneous Poisson processes either in the presence or in the absence of change-points for the climate time series considered here. The OpenBugs software [37] is used to simulate the MCMC samples.

The paper is organized as follows. Section 2 introduces the methodology where the NHPP models are presented in two situations: models with and without change-points. In Sect. 3 we give the results assuming different climate data sets. Section 4 presents some discussion on the obtained results and concluding remarks. Finally, an Appendix is included after the list of references where we give the

OpenBugs code used to generate the samples to estimate the parameters of interest as well as the data sets used in each of the applications.

2 Methodology

2.1 Non-Homogeneous Poisson Models

As mentioned before NHPP are used in many applications where, as in the present work, occurrences of events are of interest. Here, we use this type of model to estimate the probability that a climate variable (in the present case, either precipitation or temperature) exceeds a predetermined threshold a given number of times in a time interval of interest. The threshold, to which the climate variable measurement is compared, is the overall average measurement of the variable of interest taking into account the entire observed period indicated by $[0, T], T \geq 0$.

Hence, let $M_t \geq 0$ record the number of times the climate variable is above the threshold (represented by the overall average of the measurements) in the time interval $[0, t)$, $t \in [0, T]$. We assume that $M = \{M_t : t \geq 0\}$ follows a non-homogeneous Poisson process with rate and mean functions given, respectively, by $\lambda(t) > 0$ and,

$$m(t) = \int_0^t \lambda(s)\, ds, \quad t \geq 0. \tag{1}$$

Recall that the rate function $\lambda(t)$ dictates the behavior of the Poisson process.

Different parametric forms may be considered for the rate function. Due to the nature of the questions asked when using NHPP models, these functions may be borrowed from studies made in reliability theory. Hence, if θ is the vector of parameters present in the rate and mean functions, then in order to explicitly indicate this dependence, we will use, in the description of the models, $\lambda(t \mid \theta)$ and $m(t \mid \theta)$ to denote the rate and mean functions, respectively. Note that $m(t \mid \theta)$ denotes the expected number of events registered by M_t up to time t. The characterization of a non-homogeneous Poisson process of this type is specified by the functional form of $m(t \mid \theta)$, or equivalently, of its intensity function $\lambda(t \mid \theta)$, given by the first derivative of $m(t \mid \theta)$, that is, $\lambda(t \mid \theta) = dm(t \mid \theta)/dt$. For the analysis of climate data, it is interesting to have a rate function $\lambda(t \mid \theta), t \geq 0$ showing different behaviors as decreasing or increasing depending of time.

Different formulations of NHPP could be used in the climate data analysis as well as other types of analysis. One of these formulations, usually used in software reliability studies and denoted as NHPP-I, assumes that the mean value function is given by $m(t) = \alpha F(t)$ where $F(t)$ is the cumulative function of a specified probability distribution and α is a unknown parameter that should be estimated [21]. Another formulation, also used in software reliability studies and denoted as NHPP-II, is given by taking $m(t) = -\log(1 - F(t))$ where $F(t)$ is the cumulative function of a probability distribution also usually used in reliability and reliability software applications [38, 39].

For the climate data, we consider five parametric structures: the power law process (PLP) [40, 41]; the Musa–Okumoto process (MOP) [42]; the Goel–Okumoto process (GOP) [43]; a generalized form of Goel–Okumoto (GGOP); and the exponentiated-Weibull (GPLP) [21, 44, 45] which generalizes de PLP model. The PLP, MOP and GPLP models are defined as special cases of the mean function $m(t) = -\log(1 - F(t))$, that is, in the class NHPP-II, where $F(t)$ is the cumulative function of a Weibull distribution [46] given by $F(t) = \exp\{-(t/\sigma)^\alpha\}, t > 0$ for the PLP, the cumulative function of a Lomax or Pareto type II distribution [47, 48] given by $F(t) = 1 - (1 - t/\alpha)^{-\beta}, t > 0$ for the MOP and $F(t) = \{1 - \exp[-(t/\sigma)^\alpha]\}^\beta, t > 0$ the cumulative distribution of an exponentiated-Weibull distribution for the GPLP that generalizes de PLP process. The GOP and the GGOP are obtained from formulation of the mean value function given by $m(t) = \alpha F(t)$ where $F(t)$ is the cumulative function of an exponential distribution, that is, $F(t) = 1 - \exp(-\beta t)$ for the GOP model and $F(t) = 1 - \exp(-\beta t^\gamma)$ the cumulative distribution of a Weibull distribution for the GGOP model. Thus, the mean value functions, considered in the present work, are given by,

$$\begin{aligned} m_{PLP}(t \mid \theta) =& (t/\sigma)^\alpha, \text{ where } \theta = (\alpha, \sigma); \alpha, \sigma > 0, \\ m_{MOP}(t \mid \theta) =& \beta \log(1 + t/\alpha), \text{ where } \theta = (\alpha, \beta); \alpha, \beta > 0, \\ m_{GOP}(t \mid \theta) =& \alpha[1 - \exp(-\beta t)], \text{ where } \theta = (\alpha, \beta); \alpha, \beta > 0, \\ m_{GGOP}(t \mid \theta) =& \alpha[1 - \exp(-\beta t^\gamma)], \text{ where } \theta = (\alpha, \beta, \gamma); \alpha, \beta, \gamma > 0, \\ m_{GPLP}(t \mid \theta) =& -\log[1 - F_{EW}(t)], \text{ where } \theta = (\alpha, \beta, \sigma); \alpha, \beta, \sigma > 0, \end{aligned} \tag{2}$$

with $F_{EW}(t) = 1 - \exp[-(t/\sigma)^\alpha]^\beta$. The corresponding intensity/rate functions $\lambda(t \mid \theta) = dm(t \mid \theta)/dt$ for the mean functions (2) are given by,

$$\begin{aligned} \lambda_{PLP}(t \mid \theta) =& (\alpha/\sigma)(t/\sigma)^{\alpha-1}, \text{ where } \theta = (\alpha, \sigma); \alpha, \sigma > 0, \\ \lambda_{MOP}(t \mid \theta) =& \beta/(t + \alpha), \text{ where } \theta = (\alpha, \beta); \alpha, \beta > 0, \\ \lambda_{GOP}(t \mid \theta) =& \alpha\beta \exp(-\beta t), \text{ where } \theta = (\alpha, \beta); \alpha, \beta > 0, \\ \lambda_{GGOP}(t \mid \theta) =& \alpha\beta\gamma t^{\gamma-1} \exp(-\beta t^\gamma), \text{ where } \theta = (\alpha, \beta, \gamma); \alpha, \beta, \gamma > 0, \\ \lambda_{GPLP}(t \mid \theta) =& G(t)/[1 - F_{EW}(t)], \text{ where } \theta = (\alpha, \beta, \sigma); \alpha, \beta, \sigma > 0, \end{aligned} \tag{3}$$

where $G(t) = \alpha\beta\sigma^{-1}\{1 - \exp[-(t/\sigma)^\alpha]\}^{\beta-1}\exp[-(t/\sigma)^\alpha](t/\sigma)^{\alpha-1}$ and where $F_{EW}(t)$ is defined as in (2).

The intensity functions given by (3) define the hazard rates of the time between occurrence of events in the respective models. The several expressions for the rate functions cover a wide range of forms of behavior of the occurrences of the events of interest in function of time. For instance, the intensity function $\lambda_{PLP}(t \mid \theta)$ presents different behaviors depending on the value of α. These behaviors could be constant, decreasing or increasing depending on whether $\alpha = 1, \alpha < 1$ or $\alpha > 1$, respectively. The intensity functions $\lambda_{MOP}(t \mid \theta)$ and $\lambda_{GOP}(t \mid \theta)$ presents a decreasing behavior as functions of t and $\lambda_{GGOP}(t \mid \theta)$ describes the situation where the intensity increases slightly at the beginning and then begins to decrease with t. Moreover, for the rate $\lambda_{GPLP}(t \mid \theta)$ we observe that: if $\alpha \geq 1$ and $\alpha\beta \geq 1$, $\lambda(t)$ is an increasing function of t; if $\alpha \leq 1$ and $\alpha\beta \leq 1$, $\lambda(t)$ is a decreasing function of t; if $\alpha > 1$ and $\alpha\beta < 1$, $\lambda(t)$ has a bathtub form; if $\alpha < 1$ and $\alpha\beta > 1$, $\lambda(t)$ is unimodal.

Two versions of the models will be considered depending on the rate function used. In one version we assume that no change-points are present and in the other we assume the presence of a change-point. Since the Bayesian point of view will be used to estimate the parameters and the OpenBugs software will be used to program the MCMC algorithm, we only need to specify the likelihood function of the model as well as the prior distributions of the parameters involved. We start with the likelihood function.

In order to specify the likelihood function, we need to describe the information actually used in it. In the present case, this information consists of the years where exceedances of the corresponding threshold for each data set occurred. Hence, let n be the number of observed times where these events of interest occurred in the time interval $[0, T]$, $T \geq 0$ and let $0 < t_1 < t_2 < \ldots < t_n < T$ denote these times. Thus, the set of observed values is $D_T = \{n; t_1, \ldots, t_n; T\}$.

The likelihood functions of the two versions of the models considered here are given as follows.

2.2 Likelihood Function Without the Presence of Change-Points

Suppose that there are no change-points, then the likelihood function of the model is given by [49],

$$L(\theta \mid D_T) = \left[\prod_{i=1}^{n} \lambda(t_i \mid \theta)\right] \exp[-m(T \mid \theta)]. \tag{4}$$

2.3 Likelihood Function in the Presence of a Change-Point

When we have a single change-point τ making a transition between two NHPP models of the same type but with different parameters, the intensity function of the overall process is given by [14] ,

$$\lambda(t \mid \theta) = \begin{cases} \lambda(t \mid \theta_1) \text{ if } 0 \leq t \leq \tau, \\ \lambda(t \mid \theta_2) \text{ if } t \geq \tau, \end{cases} \tag{5}$$

where $\lambda(t \mid \theta_j), j = 1, 2$ are the intensity functions related to the intensity functions defined in (2) and $\theta_j, j = 1, 2$ are the parameters associated to the NHPP before and after the change-point. The corresponding mean value functions $m(t \mid \theta_j), j = 1, 2$, are given by,

$$m(t \mid \theta) = \begin{cases} m(t \mid \theta) = m(t \mid \theta_1) \text{ if } 0 \leq t \leq \tau, \\ m(\tau \mid \theta_1) + m(t \mid \theta_2) - m(\tau \mid \theta_2) \text{ if } t \geq \tau. \end{cases} \tag{6}$$

Hence, if n exceedances occurred in the time interval $[0, T]$ with the occurrence time given by D_T, then we may rewrite this set as $D_T = \{n; t_1, \ldots, t_{N_\tau}; t_{N_\tau+1}, \ldots, t_n; T\}$ where t_k, $k = 1, 2, \ldots, n$ is the time of occurrence of the kth event (in the present case is the kth exceedances of the climate threshold), τ is the change-point, and N_τ is the number of times the event occurred before the change-point. Therefore, when one change-point is allowed the likelihood function of the model is given by,

$$L(\theta \mid D_T) = \left[\prod_{i=1}^{N_\tau} \lambda(t_i \mid \theta_1)\right] \exp[-m(\tau \mid \theta_1)] \times \left[\prod_{i=N_\tau+1}^{n} \lambda(t_i \mid \theta_2)\right] \exp[-m(T \mid \theta_2) + m(\tau \mid \theta_2)]. \tag{7}$$

As a special case, for PLP model in the presence of a change-point, the intensity function (5) is given by,

$$\begin{aligned} \lambda(t \mid \theta) &= (\alpha_1/\sigma_1)(t/\sigma_1)^{\alpha_1 - 1} \text{ if } 0 \leq t \leq \tau, \\ \lambda(t \mid \theta) &= (\alpha_2/\sigma_2)(t/\sigma_2)^{\alpha_2 - 1} \text{ if } t \geq \tau, \end{aligned} \tag{8}$$

with the corresponding mean value function given by,

$$\begin{aligned} m(t \mid \theta) &= (t/\sigma_1)^{\alpha_1} \text{ if } 0 \leq t \leq \tau, \\ m(t \mid \theta) &= (\tau/\sigma_1)^{\alpha_1} + (t/\sigma_2)^{\alpha_2} - (\tau/\sigma_2)^{\alpha_2} \text{ if } t \geq \tau. \end{aligned} \tag{9}$$

In a similar way, we obtain the rate and mean functions for the MOP, GOP, GGOP and GPLP models in the presence of one change-point. Hence, we may consider a Bayesian approach entirely based on the marginal *posterior* densities. However, deriving analytical expressions for these densities is infeasible, mainly due to the complexity of the associated log-likelihood function. In our case, we used Markov chain Monte Carlo (MCMC) method based on Gibbs sampling algorithms to simulate samples for the joint posterior distributions of interest [34, 35]. A brief description of the Gibbs sampling algorithm is presented as follows:

– Suppose $\pi(\theta \mid y)$ a joint posterior distribution, where $\theta = (\theta_1, \ldots, \theta_k)$, on which we want to obtain inferences.

– For this, we simulate random quantities of complete conditional distributions $\pi(\theta_i \mid y, \boldsymbol{\theta}_{(i)})$, $\theta_{(i)} = (\theta_1, \dots, \theta_{(i-1)}, \theta_{(i+1)}, \dots, \theta_k)$.
– Consider the initial (arbitrary) values $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)})$.
– Generate $\theta_1^{(1)}$ from $\pi(\theta_1 \mid y, \theta_2^{(0)}, \dots, \theta_k^{(0)})$,
– Generate $\theta_2^{(1)}$ from $\pi(\theta_2 \mid y, \theta_1^{(1)}, \dots, \theta_k^{(0)})$,
– $\vdots$
– Generate $\theta_k^{(1)}$ from $\pi(\theta_k \mid y, \theta_1^{(1)}, \dots, \theta_{k-1}^{(1)})$.
– Replace the initial values by $\theta^{(1)} = (\theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_k^{(1)})$.
– The values $\theta_1^{(z)}, \theta_2^{(z)}, \dots, \theta_k^{(z)}$, for z sufficiently large, converge to a random quantity value with distribution $\pi(\theta \mid y)$.

For each generated sample, a chain with $N = 200,000$ values was generated for each component of the parameter vector of the model, considering a burn-in period of 5% of the chain's size. To obtain pseudo-independent samples from the *joint posterior* distribution, one out every 100 generated values was kept, resulting in chains of size 2,000 for each parameter. We assumed independent uniform $U(a, b)$ or $Gamma(c, d)$ prior distributions for the parameters of the proposed models, considering all data sets taken into account, where the hyperparameters a, b, c and d are known and Gamma(c, d) denotes a gamma distribution with mean c/d and variance c/d^2.

The discrimination of the models was made by comparing plots of the empirical accumulated number of climate violations with the estimated mean value functions versus time of occurrence. The Bayesian analysis for all models was made using the OpenBugs software [37]. Convergence of the Gibbs sampling algorithm was monitored by usual time series plots for the simulated samples. In the selection of the best model we also used the deviance information criterion (DIC) [50] which is an approximation for the Bayes factor (smaller values of DIC indicate better models). However, the results were very similar for all assumed models. That makes it difficult to choose the best model using this criterion.

3 Results

The data sets considered in the present work are the yearly averaged amount of rainfall (in mm) from 1879 to 2002 ($T = 124$) and the yearly average of the maximum temperatures (in degrees Celsius - °C) collected from 1915 to 2003 ($T = 88$) in a climate station in Almaty, Kazakhstan [51]; the yearly maximum temperature averages collected from 1894 to 2003 ($T = 110$) reported at a climate station in Tashkent, Uzbekistan; and the yearly average temperature (in degrees Fahrenheit - °F) collected from 1895 to 2019 ($T = 125$) in the USA (https://www.ncdc.noaa.gov/cag/).

The threshold considered in each case is the corresponding overall average measurement. For instance, the threshold for the rainfall data from Kazakhstan is the average of the 124 values belonging to this data set, and in the case of temperature, it is the average of the corresponding 88 measurements.

In the next three subsections we present the application of the NHPP model considered in the earlier sections to each of these data sets.

3.1 Kazakhstan Climate Data

In a first instance, we consider data from Almaty, Kazakhstan. Hence, we have the yearly precipitation averages (in mm) from 1879 to 2002 (given an observation period of $T = 124$ years) and the yearly average of the maximum temperatures (in °C) measured from 1915 to 2003 (given an observation period of $T = 88$ years) in a climate station in Almaty, Kazakhstan [51]. Figure 1 shows the plots of these two time series. The plot on the left corresponds to

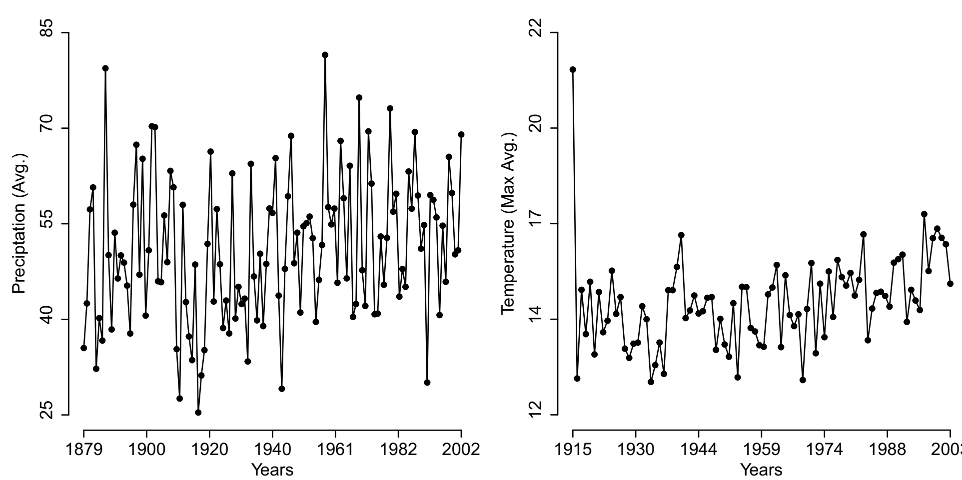

Fig. 1 Yearly average precipitation and averages of the maximum temperatures in Almaty, Kazakhstan

the yearly precipitation averages and that on the right corresponds to the yearly averages of maximum temperatures.

The overall precipitation average for the period of 124 years (1879 to 2002) is equal to 50.88 (data set in Table 5, in Appendix 1), and the overall average of averages of the maximum temperatures for the period of 88 years (1915 to 2003) is equal to 14.939 (data set in Table 6, in Appendix 1). Taking this into account, it is possible to observe an increase in trend, which is above average, in the precipitation as well as temperature data after the years 1920 and 1950, respectively. That could indicate a possible presence of a change-point in the time series since after those years a change in the behavior of the data may be observed.

3.1.1 Yearly Rainfall Averages

Since the overall precipitation average for the period of 124 years is equal to 50.88, we assume for this data analysis a threshold for the precipitation average equal to 51. That gives us $n = 57$ exceedances, that is, there were 57 years where precipitation averages were above 51 during the follow-up period of $T = 124$ years.

Under the NHPP setting, we assume that, in the case where no change-points are present, the prior distributions of the parameters appearing in the rate functions given by (2) are the following,

- PLP: $\alpha \sim U(0,5)$;$\sigma \sim U(0,10000)$.
- MO: $\alpha \sim$ Gamma $(0.01, 0.01)$;$\beta \sim$ Gamma $(0.01, 0.01)$.
- GO: $\alpha \sim$ Gamma $(0.01, 0.01)$;$\beta \sim$ Gamma $(0.01, 0.01)$.
- GGO: $\alpha \sim$ Gamma $(0.01, 0.01)$;$\beta \sim$ Gamma $(0.01, 0.01)$; $\gamma \sim U(0,5)$.
- GPLP: $\alpha \sim$ Gamma $(0.1, 0.1)$;$\sigma \sim$ Gamma $(0.1, 0.1)$; $\beta \sim U(0,100)$.

Observe that we are using non-informative prior distributions for the parameters of the proposed models. Estimation of the parameters was performed using a sample of size 1000 taken every 100th simulated Gibbs sample after a burn-in period of 11000 iterations.

Table 1 shows the posterior means, standard deviations and 95% credible intervals for the parameters of each model considering the yearly precipitation averages from Kazakhstan where the threshold 51 was used.

From Table 1, we have that assuming a squared error loss function, the Bayesian estimator is given by the posterior estimated mean which is equal to 42.53 (or year 43 which corresponds to the year 1921), but considering the posterior estimated median (32.63), this change-point occurred earlier, in year 33 corresponding to the year 1911.

In Fig. 2, leftmost figure, we have the plots of the accumulated average precipitation exceedances (values above the threshold 51) as well as the estimated mean value function assuming each one of the five proposed rate functions and when no change-points are present. It is possible to observe that the best fitted model is given by the PLP. Even though the GPLP model has three parameters and good convergence of the MCMC algorithm, it has not produced a better fit for this data set. Note that this model possibly has some identifiability problems, and for practical use it is necessary very informative prior distributions based on opinion of climate experts.

Table 1 Posterior summaries (Kazakhstan precipitation averages data)

Model	Par.	Post. Mean	Std. Dev.	95% Cred. Int.	
				Lower	Upper
PLP	α	1.376	0.173	1.069	1.718
	σ	6.795	2.542	2.832	12.35
GPLP	α	1.153	0.202	0.784	1.565
	β	2.465	3.411	0.263	11.19
	σ	5.862	3.044	1.305	12.79
MOP	α	201.6	91.03	73.71	417.6
	β	115.7	44.20	53.66	224.2
GOP	α	239.9	101.8	109.4	495.2
	β	0.003	0.001	0.001	0.005
GGOP	α	168.2	81.94	73.06	378.2
	β	0.002	0.001	0.001	0.002
	γ	1.470	0.193	1.116	1.844
PLP change-point	α_1	1.354	0.821	0.498	3.611
	α_2	1.307	0.208	0.799	1.604
	σ_1	5.082	2.217	1.293	9.396
	σ_2	6.012	2.653	0.593	9.861
	τ	42.63	36.01	1.714	120.0

Also in Fig. 2, middle figure, we have the plots of the empirical mean value function $m(t)$ and the estimated using the PLP model (the best fitted model when no change-points are present) when a change-point is allowed and when we assume the following prior distributions: $\alpha_j \sim U(1.3, 1)$; $\sigma_j \sim U(0, 10)$ and $\tau \sim U(1, 124), j = 1, 2$. Observe that when specifying the prior distributions in the case of one change-point, we use some prior information obtained from the Bayesian estimators assuming PLP model without the presence of change-points, especially for the parameters α_1 and α_2, i.e., we use a Bayesian empirical analysis [52]. Figure 2 also shows the histogram of the generated Gibbs samples for the change-point τ assuming the PLP model in the presence of a change-point (rightmost figure).

Table 1 also shows the posterior summaries of interest for the PLP model in the presence of a change-point. With the obtained Bayesian estimate for the change-point $\tau = 42.63$, that is, $\tau = 43$ which corresponds to the year 1921, from (8) we have $m(t) = (t/5.082)^{1.354}$ if $t \leq 43$, and

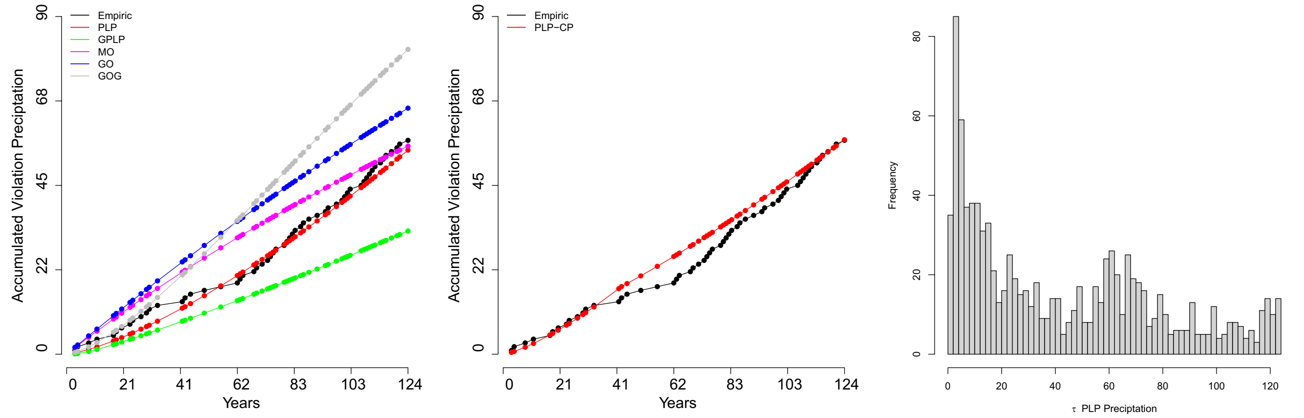

Fig. 2 Accumulated precipitation exceedances (empirical and fitted $m(t)$) and histogram of the generated Gibbs samples for τ in the case of the PLP model using the Kazakhstan precipitation data

$m(t) = (43/5.082)^{1.354} + (t/6.012)^{1.307} - (43/6.012)^{1.307}$ if $t > 43$.

3.1.2 Yearly Maximum Temperature Averages

When we consider the temperature data, we have seen that the overall average for the period of 88 years is 14.939. Hence, we take 15 as the threshold of the Kazakhstan temperature data. Using this threshold, there were $n = 43$ years where exceedances occurred, that is, there were 43 years with values above the threshold 15 in the follow-up period of $T = 88$ years.

Consider now the NHPP formulation with parametric forms for the intensity function given by (2). In the present case, we use the same prior distributions (9) assumed for the Kazakhstan's precipitation data as well as the same MCMC simulation procedure.

Table 2 shows the posterior means, standard deviations and 95% credible intervals for the parameters of each model taking into account the yearly averages of the maximum temperatures and the threshold 15.

In Fig. 3, leftmost figure, we have the plots of the accumulated exceedances of averages of the maximum temperatures (values above the threshold 15) and the estimated mean value functions assuming each one of the five proposed models and with no change-points present.

Looking at Fig. 3, it is observed that the best fitted model, when no change-points are allowed, is again given by the PLP model. Also in Fig. 3, middle figure, we have the plots of the empirical $m(t)$ and the fitted by a PLP in the presence of a change-point when we use the following prior distributions: $\alpha_j \sim U(1.3, 1); \sigma_j \sim U(0, 10)$ and $\tau \sim U(1, 88), j = 1, 2$.

The posterior summaries of interest for the PLP model in the presence of a change-point are also shown in Table 2. The estimated change-point $\tau = 33$ (approximation for the estimated value 32.56) corresponds to the year 1948. Hence, from (8), we have $m(t) = (t/3.884)^{1.09}$ if $t \geq 33$ and $m(t) = (33/3.884)^{1.09} + (t/6.24)^{1.407} - (33/6.24)^{1.307}$ if $t > 33$.

Figure 3 also shows the histogram of the simulated Gibbs samples for the change-point τ (rightmost plot) assuming the PLP model with one change-point. Moreover, from Table 2, we observe that assuming a squared error loss function the Bayesian estimator of the change-point, given by the posterior mean, is equal to 32.56, but considering the posterior median (23.505), this estimated change-point occurred early in the year 23 corresponding to the year 1938. Hence, the

Table 2 Posterior summaries (Kazakhstan maximum temperature averages data)

Model	Par.	Post. Mean	Std. Dev.	95% Cred. Int.	
				Lower	Upper
PLP	α	1.342	0.193	1.001	1.771
	σ	5.597	2.236	2.022	10.66
GPLP	α	1.129	0.267	0.671	1.711
	β	2.686	6.493	0.139	14.40
	σ	4.251	2.927	0.413	11.18
MOP	α	151.3	82.39	46.13	369.9
	β	90.90	40.74	37.74	199.8
GOP	α	191.9	91.55	78.85	419.2
	β	0.004	0.002	0.001	0.008
GGOP	α	160.0	103.2	60.00	453.9
	β	0.002	0.001	0.001	0.004
	γ	1.383	0.199	1.001	1.767
PLP change-point	α_1	1.090	0.386	0.455	2.038
	α_2	1.407	0.217	0.927	1.740
	σ_1	3.884	2.121	0.697	9.031
	σ_2	6.240	2.542	0.997	9.904
	τ	32.56	24.70	1.895	83.74

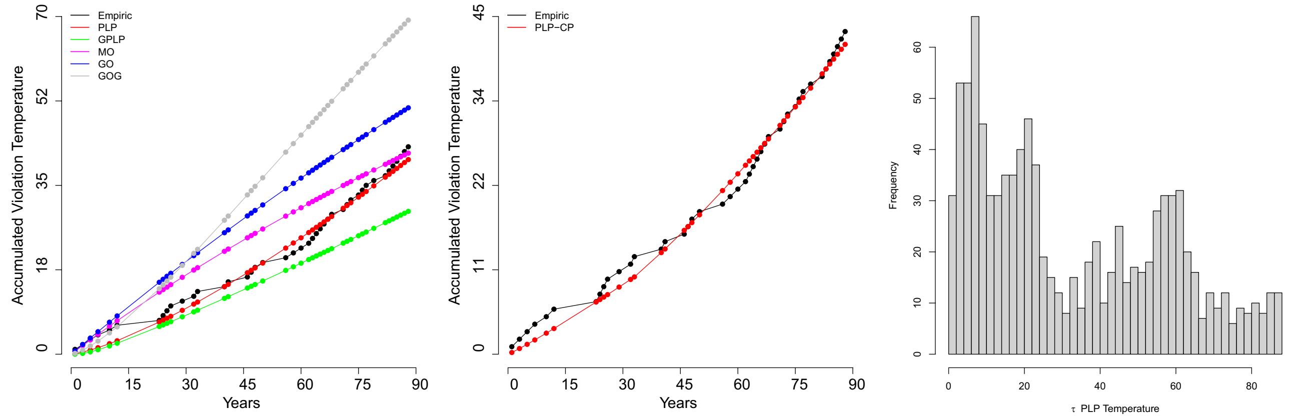

Fig. 3 Accumulated maximum temperature average exceedances (empirical and fitted $m(t)$) and histogram of the simulated Gibbs samples for τ in the case of the PLP model using and Kazakhstan temperature data

changes in the climate variable corresponding to temperature have been observed for quite a while.

3.2 Uzbekistan Maximum Temperature Data

We consider now the Uzbekistan's temperature (°C) data (data set in Table 7, in Appendix 1). Hence, we have the yearly averages of the maximum temperatures collected from 1894 to 2003 (giving an observed period of $T = 110$ years) in a climate station in Tashkent, Uzbekistan, another country in Central Asia [51]. Figure 4 shows the plot of this time series.

The overall average for the period of 110 years is 20.7395. Thus, we assume in the application of the model a threshold equal to 21. That gives us $n = 50$ exceedances, that is, there were 50 years in which the average of the maximum temperature averages is above 21 during the follow-up period of $T = 110$ years.

Hence, under the NHPP formulation we consider the parametric forms for the intensity function given by (2) and the prior distributions given in (9). Table 3 shows the posterior means, standard deviations and 95% credible intervals for the parameters of each model considering the yearly

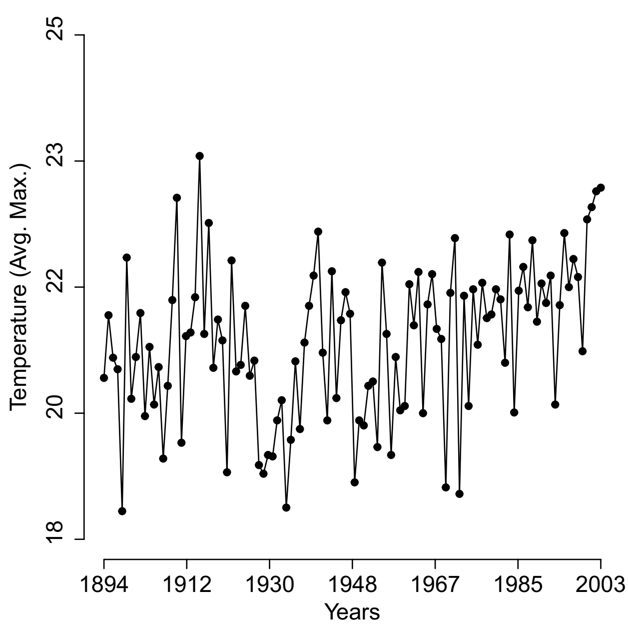

Fig. 4 Yearly average of the maximum temperature averages in Tashkent, Uzbekistan

Table 3 Posterior summaries (Uzbekistan maximum temperature averages data)

Model	Par.	Post. Mean	Std. Dev.	95% Cred. Int.	
				Lower	Upper
PLP	α	1.550	0.199	1.195	1.969
	σ	9.073	3.013	4.256	15.64
GPLP	α	1.354	0.268	0.903	2.002
	β	1.627	2.673	0.140	7.330
	σ	8.18	4.137	2.033	18.70
MOP	α	209.3	96.61	72.46	449.3
	β	114.7	45.80	51.62	230.3
GOP	α	247.1	113.3	107.6	555.2
	β	0.002	0.001	0.001	0.005
GGOP	α	146.0	66.91	66.72	315.4
	β	0.001	0.001	0.000	0.002
	γ	1.661	0.236	1.218	2.255
PLP change-point	α_1	1.264	0.412	0.571	2.289
	α_2	2.103	0.458	1.207	2.831
	σ_1	6.752	4.229	1.607	17.35
	σ_2	18.12	7.486	3.685	29.46
	τ	43.20	25.88	2.462	98.94

averages of the maximum temperature averages when the threshold equals to 21. The same MCMC scheme considered in Sect. 3.1 is used here as well.

Figure 5, leftmost figure, shows the plots of the accumulated maximum temperature averages exceedances (values above the threshold 21) and the estimated mean value functions assuming each of the five proposed forms for the rate function and when no change-points are present.

It is possible to observe, by looking at Fig. 5, that, in the case of no change-points, the best fit is again given by the PLP model. Figure 5 also shows, middle figure, the plots of the empirical and fitted $m(t)$ when the PLP model in the presence of a change-point is used and when the following prior distributions considered: $\alpha_j \sim U(1.5, 1); \sigma_j \sim U(0, 30)$ and $\tau \sim U(1, 110), j = 1, 2$.

From Table 3, we may observe that assuming a squared error loss function, the Bayesian estimator for the change-point, given by the posterior mean, is equal to 43.20 corresponding to a change-point $\tau = 43$. Using this change-point, from (8) we have that, $m(t) = (t/6.752)^{1.264}$ if $t \leq 43$ and $m(t) = (43/6.752)^{1.264} + (t/18.12)^{2.103} - (43/18.12)^{2.103}$ if $t > 43$.

3.3 US Climate Data

Now, consider the USA yearly average temperatures (°F) data from 1895 to 2019 giving an observed period of $T = 125$ years. The overall of the 125 consecutive years measurements is 52.16 (data set given in Table 8, in the Appendix 1). The data published in August 2020 (see https://www.ncdc.noaa.gov/cag/) are presented in Fig. 6.

Looking at that Fig. 6 we may also see that, apparently, from the year 1982 there is an increasing trend in the values of the average temperature until the year 2019.

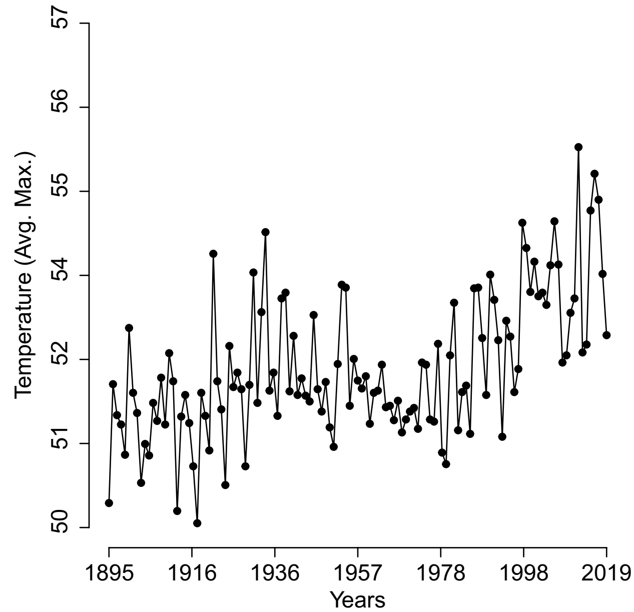

Fig. 6 Yearly average temperatures in USA from 1895 to 2019

Considering the five rate functions, described in Sect. 2, in the model without the presence of change-points, we assume the same prior distributions (9) for the parameters of the models, except for the parameter γ of the GGOP model where we assume a prior $\gamma \sim U(0, 100)$ instead of $\gamma \sim U(0, 5)$. The same MCMC scheme used for the climate data of Kazakhstan using the OpenBugs software is applied to the present data set. Since the mean temperature during the observed period is 52.16, for the present data set, we take 52 as the threshold value. That gives $n = 60$ exceedances, that is, there were 60 years in which the temperature averages were above 52 during the follow-up period of $T = 125$ years.

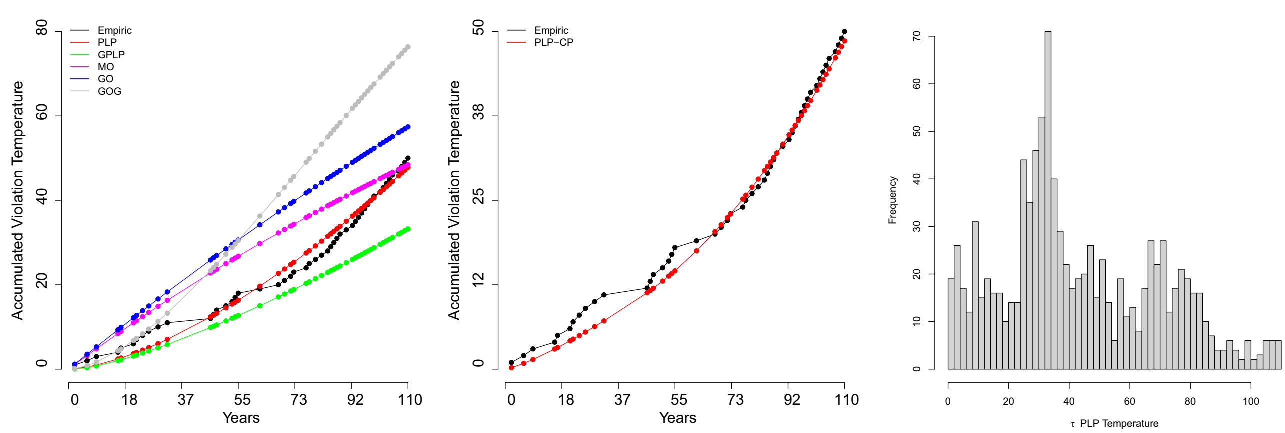

Fig. 5 Accumulated maximum temperature average exceedances (empirical and fitted $m(t)$) and the histogram of the generated Gibbs samples for τ (PLP in the presence of a change-point) for Uzbekistan's data

Table 4 Posterior summaries (USA yearly temperature averages)

Model	Par.	Post. Mean	Std. Dev.	95% Cred. Int.	
				Lower	Upper
PLP	α	1.730	0.219	1.327	2.190
	σ	11.95	3.581	5.494	19.84
GPLP	α	1.436	0.240	1.021	1.945
	β	3.114	4.316	0.319	12.76
	σ	10.46	4.257	3.723	19.79
MOP	α	241.5	96.82	102.4	476.3
	β	139.4	49.80	68.23	263.4
GOP	α	291.5	130.3	126.2	635.3
	β	0.002	0.001	0.001	0.004
GGOP	α	98.25	24.35	73.04	169.9
	β	0.001	0.001	0.000	0.002
	γ	2.043	0.191	1.639	2.414
PLP change-point	α_1	1.646	0.617	0.544	3.256
	α_2	1.978	0.417	1.199	2.709
	σ_1	11.50	5.318	3.592	25.56
	σ_2	16.51	7.514	3.243	29.20
	τ	54.67	36.36	2.477	116.5

Table 4 shows the posterior means, standard deviations and 95% credible intervals for the parameters of each model when we take into account the temperature averages in USA with a threshold equals to 52.

Figure 7, leftmost figure, shows the plots of the accumulated average precipitation exceedances (values above the threshold 52) and the estimated mean value functions assuming each one of the five proposed models when no change-points are allowed.

Looking at the leftmost plots in Fig. 7, we may see that the best fit is again given by the PLP. In the same Fig. 7, middle figure, we have the plots of the empirical $m(t)$ and fitted by a PLP model in the presence of a change-point where the following prior distributions were assumed: $\alpha_j \sim U(1.7, 1)$;$\sigma_j \sim U(0, 30)$ and $\tau \sim U(1, 124)$, $j = 1, 2$.

In Table 4 we also have the posterior summaries of interest for the PLP model in the presence of a change-point. The estimated change-point $\tau = 55$ (approximation to the Bayesian estimate $\tau = 54.67$) corresponds to the year 1949. Hence, from (8), we have that $m(t) = (t/11.5)^{1.646}$ if $t \leq 55$ and $m(t) = (55/11.5)^{1.646} + (t/16.51)^{1.978} - (55/16.51)^{1.978}$ if $t > 55$.

In Fig. 7, rightmost figure, we also have the histogram of the generated Gibbs samples for the change-point τ assuming the PLP model in the presence of a change-point. In this case, the posterior mean equal to 54.67 (or year 55) is close to the posterior median which is 56.38.

4 Discussion and Concluding Remarks

The use of the methodology based in NHPP in the presence of one change-point considered in this study could be applied to any climate data to detect the year in which a possible change in climate might have occurred. The proposed methodology also could be generalized for situations with more than one change-point. However, we do not take that into account. One reason for not doing so is that changes in climate are, in general, a very slow process and the occurrence of a comparatively new change-point detection may be considered a rare event to occur in the future, during the period of time considered in the present work. However, this might be a subject of a future study if an extended data set is obtained. The Bayesian inferences of interest (point estimators and credible intervals for the parameters of the models) are obtained using MCMC simulation methods where existing free software like the OpenBugs may be used with few

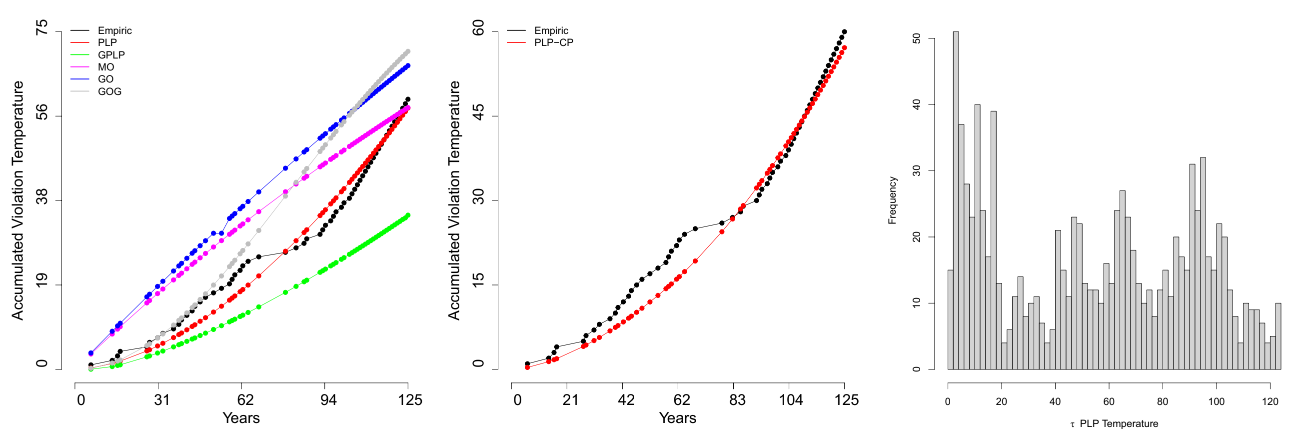

Fig. 7 Accumulated temperature average violations (empirical and fitted m(t)) for USA and histogram of the simulated Gibbs samples for τ (PLP in the presence of a change-point)

computational costs. Important results were obtained from the data analysis indicating significative climate changes in the applications considered in the study. They are listed as follows.

In the first application, i.e., where the rainfall data from Kazakhstan (1879 to 2002) are used, there are 11 years before the estimated change-point (year 43 corresponding to the year 1921) and 46 years after, with precipitation averages above the overall average. This might be considered as an indication that the average rain has been increasing after the year 1921. Moreover, in the second application, i.e., where the temperature data from Kazakhstan (1915 to 2003) are used we have that: there are 13 years before the estimated change-point (year 33 which corresponds to the year 1948) and 30 years after, with maximum temperature averages above the overall average, an indication that the average maximum temperature has been increasing after the year 1948.

In the third application, i.e., when the temperature data from Uzbekistan (1894 to 2003) are used, we have that: there are 11 years before the estimated change-point (year 43 which corresponds to the year 1936) and 99 years after, with maximum temperature averages above the overall average. That is also an indication that the average maximum temperature has been increasing after the year 1936.

Finally, when average temperature data from USA (1895 to 2019) are used, we have the following: there are 18 years before the estimated change-point (year 55 which corresponds to the year 1949) and 106 years after the estimated change-point with average temperatures above the overall average, also an indication that the average temperature has been increasing after the year 1949.

Another possible way of studying the type of problems considered here, is to use the so-called Hawkes process. In this case, the intensity function ruling the occurrences of events that we are trying to study is given by $\lambda^*(t) = f^*(t)/(1 - F^*(t)), t \geq 0$, where $f^*(\cdot)$ and $F^*(\cdot)$ are the density and distribution functions of the occurrence times conditioned to the past occurrences times. This type of models may be useful in the cases where exceedances may occur in small clusters. However, this paths is not followed here and is the subject of future studies. For a review regarding the Hawkes process see, for instance, [53].

Appendix 1

Table 5 A1. Average yearly precipitation in Kazakhstan (Almaty) from the year 1879 to 2002

35.5000	42.5000	57.2500	60.6667	32.2500	40.1667	36.6667	79.4417	50.0833
38.4167	53.5833	46.4167	50.0333	48.8833	45.3417	37.7750	57.9750	67.3917
47.0083	65.2250	40.5917	50.8333	70.3333	70.1583	45.9750	45.8250	56.3000
48.9750	63.3417	60.7167	35.3167	27.5750	57.9667	42.7083	37.3083	*
33.5625	48.6000	25.3750	31.2250	35.1750	51.8455	66.3250	42.8000	57.3250
48.6417	38.6333	42.9500	37.7917	62.9083	40.1333	45.1167	42.3583	43.2500
33.3917	64.3833	46.7250	39.8417	50.2917	38.9417	48.7167	57.4417	56.7250
65.2500	43.7333	29.1083	47.9250	59.3500	68.7750	48.7833	53.5500	41.0667
54.6333	55.0500	56.1167	52.7417	39.5583	46.2500	51.6750	81.5333	57.6000
54.9000	57.3750	45.7417	67.9500	59.0333	46.4250	64.1500	40.3833	42.4083
74.7750	47.7000	42.0833	69.4833	61.2750	40.7500	40.9000	52.9833	45.4333
52.7750	73.1333	56.8667	59.7000	43.5750	47.9417	45.1500	63.1917	57.3667
69.3500	59.4333	51.0833	54.7500	30.0833	59.5000	58.7500	56.0000	40.6667
54.7273	45.9167	65.5000	59.8333	50.1667	50.8333	69.0000		

Table 6 A2. Average yearly maximum temperature in Kazakhstan (Almaty) from the year 1915 to 2003

21.0250	12.9455	15.2750	14.1111	15.4800	13.5833	15.2083	14.1583	14.4583
15.7667	14.6417	15.0833	13.7333	13.4917	13.8583	13.8917	14.8417	14.5000
12.8583	13.3000	13.8917	13.0667	15.2583	15.2583	15.8667	16.7000	14.5333
14.7250	15.1167	14.6500	14.7083	15.0583	15.0833	13.7000	14.5083	13.8417
13.5167	14.9167	12.9833	15.3500	15.3417	14.2667	14.1833	13.8250	13.7750
15.1500	15.3250	15.9167	13.7667	15.6500	14.6083	14.3167	14.6333	12.9083
14.7667	15.9667	13.6083	15.4250	14.0333	15.7500	14.5583	16.0500	15.6000
15.3833	15.7083	15.1167	15.5333	16.7250	13.9500	14.7818	15.1917	15.2167
15.1083	14.8250	15.9750	16.0667	16.1917	14.4333	15.2750	14.9909	14.7417
17.2500	15.7583	16.6167	16.8750	16.6333	16.4583	15.4333		

Table 7 A3. Average yearly maximum temperature in Uzbekistan (Tashkent) from the year 1894 to 2003

20.24	21.11	20.52	20.36	18.39	21.91	19.95	20.53	21.14	19.71	20.67
19.87	20.39	19.12	20.13	21.32	22.74	19.34	20.82	20.87	21.36	23.32
20.85	22.39	20.38	21.05	20.76	18.93	21.87	20.33	20.42	21.24	20.27
20.48	19.03	18.91	19.17	19.15	19.65	19.93	18.44	19.38	20.47	19.53
20.73	21.24	21.66	22.27	20.59	19.65	21.72	19.96	21.04	21.43	21.13
18.79	19.65	19.58	20.13	20.19	19.28	21.84	20.85	19.17	20.53	19.79
19.85	21.54	20.97	21.71	19.75	21.26	21.68	20.92	20.78	18.72	21.42
22.18	18.63	21.38	19.85	21.47	20.70	21.56	21.07	21.12	21.47	21.33
20.45	22.23	19.76	21.45	21.78	21.22	22.15	21.02	21.55	21.28	21.66
19.87	21.25	22.25	21.50	21.89	21.64	20.61	22.44	22.61	22.83	22.88

Table 8 A4. Average yearly temperature in USA from the year 1895 to 2019

50.34	51.99	51.56	51.43	51.01	52.77	51.87	51.59	50.62	51.16	51.00
51.73	51.48	52.08	51.43	52.42	52.03	50.23	51.54	51.84	51.45	50.85
50.06	51.87	51.55	51.07	53.80	52.03	51.64	50.59	52.52	51.95	52.15
51.92	50.85	51.98	53.54	51.73	52.99	54.10	51.90	52.15	51.55	53.18
53.26	51.89	52.66	51.84	52.07	51.83	51.75	52.95	51.92	51.61	52.02
51.39	51.12	52.27	53.37	53.33	51.69	52.34	52.04	51.93	52.10	51.44
51.87	51.90	52.26	51.67	51.69	51.49	51.76	51.32	51.50	51.61	51.66
51.37	52.29	52.26	51.50	51.47	52.55	51.04	50.88	52.39	53.12	51.35
51.88	51.97	51.30	53.32	53.33	52.63	51.84	53.51	53.16	52.60	51.26
52.87	52.65	51.88	52.20	54.23	53.88	53.27	53.69	53.21	53.26	53.09
53.64	54.25	53.65	52.29	52.39	52.98	53.18	55.28	52.43	52.54	54.40

Appendix 2

OpenBugs code (PLP with a change point)

```
model
{
C <- 1000
   for (i in 1:N)
{
      zeros[i]<- 0
      phi[i]<- -log(L[i])+C
      zeros[i] ~ dpois(phi[i])
      log(lambda[i]) <- log(alpha[J[i]])-log(sigma[J[i]])
                              +(alpha[J[i]]-1)*(log(t[i]) -log(sigma[J[i]]))
      L[i]<-lambda[i]*m
      J[i] <- 1+step(t[i]-tau-0.5)
   }
   m <- exp(-(pow((tau/sigma[1]),alpha[1])+pow((T/sigma[2]),alpha[2])
          -pow((tau/sigma[2]),alpha[2]))/N)

# Prior distributions

tau ~ dunif(1,T)
sigma[2] ~ dunif(0,10)
alpha[1] ~ dgamma(1.3,1)
alpha[2] ~ dgamma(1.3,1)
sigma[1] ~ dunif(0,10)
}
```

Acknowledgements The authors are grateful to an anonymous reviewer for the comments and suggestions which helped to improve the presentation of the results and also for suggesting the work about Hawkes processes. The authors are also grateful to Dr Eliane R. Rodrigues for a review of the manuscript and important comments that led to the great improvement of the article.

Author Contributions Jorge Alberto Achcar and Ricardo Puziol de Oliveira contributed to revision, writing, analysis and methodology.

Funding Not applicable.

Data Availability See Appendix 1.

Code Availability See Appendix 2.

Declarations

Conflicts of Interest/Competing Interests There is no conflict of interest.

References

1. Intergovernmental Panel on Climate Change. (2019). IPCC, 2018: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. V. Masson-Delmotte, P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (Eds.).
2. Joyce, L. A., & Birdsey, R. (2000). The impact of climate change on America's forests: A technical document supporting the 2000 USDA

Forest Service RPA Assessment. In L.A. Joyce, R. Birdsey (Eds.) *Gen. Tech. Rep RMRS-GTR-59.* Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
3. Doney, S. C., Ruckelshaus, M., Duffy, J. E., et al. (2012). Climate change impacts on marine ecosystems. *Annual Review of Marine Sciences, 4,* 11–37.
4. Gayton, D. V. (2008). Impacts of climate change on British Columbia's biodiversity: A literature review. Extended Abstract. *BC Journal of Ecosystems and Management*, *9*, 26-30.
5. Katz, R. W. (1996). Use of conditional stochastic models to generate climate change scenarios. *Climate Change, 32,* 237–255.
6. Majda, A. J., Tomofeyev, I., & Eijnden, E. V. (2001). A mathematical framework for stochastic climate models. *Communications on Pure and Applied Mathematics LIV*, 891–974.
7. Rajagopalan, B., Lall, U., & Tarboton, D. G. (1984). A non parametric renewal model for modeling daily precipitation. In K.W. Hipel, A.I. McLeod, U.S. Panu, V.P. Sing (Eds.) *Stochastic and Statistical Methods in Hydrology and Environmental Engineering, 4*, 47-59.
8. Outayek, S. E., & Nguyen, V. -T. -V. (2019) Stochastic modeling of daily rainfall process in the context of climate change. *In: CSCE Annual Conference*.
9. Franzke, C. L. E., O'Kane, T. J., Bener, J., Williams, P. D., & Lucarini, V. (2015). Stochastic climate theory and modeling. *WIRES Climate Change, 6,* 63–78.
10. Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., et al. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres*, *111*(D5).
11. Foley, A. (2010). Uncertainty in regional climate modelling: A review. *Progress in Physical Geography, 34*(5), 647–670.
12. Powell, B. (2016). Chapter 22 - Statistical modelling of climate change. *In: Climate Change*. Elsevier, 341–354.
13. Richards, G. R. (1993). Change in global temperature: A statistical analysis. *Journal of Climate, 6*(3), 546–559.
14. Achcar, J. A., Rodrigues, E. R., Paulino, C. D., & Soares, P. (2010). Non-homogeneous Poisson models with a change-point: An application to ozone peaks in Mexico City. *Environmental and Ecological Statistics, 17*(4), 521–541.
15. Achcar, J. A., & Sicchieri, M. P. L. (2010). da Silva: Comparação de modelos para os excessos de internações hospitalares diárias por pneumonia na cidade de São Paulo. (In Portuguese) *Revista Brasileira de Biometria*, *28*, 87-103.
16. Achcar, J. A., Rodrigues, E. R., & Tzintzun, G. (2011). Using non-homogeneous Poisson models with multiple change-points to estimate the number of ozone exceedances in Mexico City. *Environmetrics, 22*(1), 1–12.
17. Coelho-Barros, E. A., Achcar, J. A., Martinez, E. Z., Davarzani, N., & Grabsh, H. I. (2019). Bayesian inference for segmented Weibull distribution. *Revista Colombiana de Estadística, 42,* 225–243.
18. Puziol de Oliveira, R., Chen, C., & Achcar, J. A. (2018). Bayesian estimation for change-points in counting time series: a unified approach using linear regression model and non-homogeneous Poisson processes. In: Workshop de Bioestatística, 05-07 December 2018, Maringá-PR, Brazil.
19. Achcar, J. A. (2001). Ch. 29. Bayesian analysis for software reliability data. *Handbook of Statistics*, *20*, 733–748.
20. Achcar, J. A., Dey, D .K., & Niverthi, M. (1998). A Bayesian approach using nonhomogeneous Poisson process for software reliability models. In Frontiers in Reliability, Basu, SK, Mukhopadhyay, S. (eds.); Series on Quality, Reliability and Engeneering Statistics 4; Calcutta University; India
21. Cid, J. E. R., & Achcar, J. A. (1999). Bayesian inference for nonhomogeneous poisson processes in software reliability models assuming nonmonotonic intensity functions. *Computational Statistics & Data Analysis, 32*(2), 147–159.
22. Wilson, S. P., & Costello, M. J. (2005). Predicting future discoveries of european marine species by using a non-homogeneous renewal process. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 54*(5), 897–918.
23. Guarnaccia, C., Quartieri, J., Tepedino, C., & Rodrigues, E. R. (2015). An analysis of airport noise data using a non-homogeneous Poisson model with a change-point. *Applied Acoustics*, *91*, 33-39.
24. Achcar, J., Martinez, E., Ruffino-Netto, A., Paulino, C., & Soares, P. (2008). A statistical model investigating the prevalence of tuberculosis in new york city using counting processes with two change-points. *Epidemiology & Infection, 136*(12), 1599–1605.
25. Achcar, J. A., & Loibel, S. (1998). Constant hazard function models with a change point: A Bayesian analysis using Markov chain Monte Carlo methods. *Biometrical Journal: Journal of Mathematical Methods in Biosciences, 40*(5), 543–555.
26. Kuo, L., & Yang, T. Y. (1996). Bayesian computation for nonhomogeneous Poisson processes in software reliability. *Journal of the American Statistical Association, 91*(434), 763–773.
27. Pievatolo, A., & Ruggeri, F. (2004). Bayesian reliability analysis of complex repairable systems. *Applied Stochastic Models in Business and Industry, 20*(3), 253–264.
28. Achcar, J. A., & Bolfarine, H. (1989). Constant hazard against a change-point alternative: A Bayesian approach with censored data. *Communications in Statistics-Theory and Methods, 18*(10), 3801–3819.
29. Carlin, B. P., Gelfand, A. E., & Smith, A. F. (1992). Hierarchical Bayesian analysis of changepoint problems. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 41*(2), 389–405.
30. Dey, D. K., & Purkayastha, S. (1997). Bayesian approach to change point problems. *Communications in Statistics-Theory and Methods, 26*(8), 2035–2047.
31. Raftery, A. E., & Akman, V. (1986). Bayesian analysis of a Poisson process with a change-point. *Biometrika*, 85–89.
32. Ruggeri, F., & Sivaganesan, S. (2005). On modeling change points in non-homogeneous Poisson processes. *Statistical Inference for Stochastic Processes, 8*(3), 311–329.
33. Gyarmati-Szabó, J., Bogachev, L. V., & Chen, H. (2011). Modelling threshold violations of air pollution concentrations using multiple change-points Poisson process. *Atmospheric Environment, 45,* 5493–5503.
34. Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician, 49*(4), 327–335.
35. Gelfand, A. E., & Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association, 85*(410), 398–409.
36. Smith, A. F., & Roberts, G. O. (1993). Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Methodological), 55*(1), 3–23.
37. Spiegelhalter, D. J., Thomas, A., Best, N., & Lunn, D. (2003). WinBUGS version 1.4 user manual. MRC Biostatistics Unit, Cambridge. URL: http://www.mrc-bsu.cam.ac.uk/bugs.
38. Oliveira, R. P., Achcar, J. A., Mazucheli, J., & Bertoli, W. (2021). A new class of bivariate Lindley distributions based on stress and shock models and some of their reliability properties. *Reliability Engineering & System Safety, 211,* 107528.
39. Yang, T. Y. (1994). Computational approaches to bayesian inference for software reliability. PhD Thesis, Department of Statistics, University of Connecticut, Storrs, USA.
40. Crow, L. H. (1974). Reliability analysis for complex repairable systems. In: Proschan, F., Serfling, R. J. (Eds.). *Reliability and Biometry*, 379–410.

41. Crow, L. H. (1982). Confidence interval procedures for the Weibull process with applications to reliability growth. *Technometrics, 24*(1), 67–72.
42. Musa, J. D., & Okumoto, K. (1984). A logarithmic Poisson execution time model for software reliability measurement. *In: Proceedings of the 7th international Conference on Software Engineering*. Citeseer, 230–238.
43. Goel, A., & Okumoto, K. (1978). An analysis of recurrent software failures in real-time control system. *In: Proceedings of ACM Conference*. Washington, 496–500.
44. Cancho, V., Bolfarine, H., & Achcar, J. (1999). A Bayesian analysis for the exponentiated-Weibull distribution. *Journal Applied Statistical Science, 8*(4), 227–42.
45. Mudholkar, G. S., Srivastava, D. K., & Freimer, M. (1995). The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. *Technometrics, 37*(4), 436–445.
46. Weibull, W., et al. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics, 18*(3), 293–297.
47. Kotz, S., Balakrishnan, N., & Johnson, N. L. (2004). Pareto distributions. Continuous multivariate distributions, Volume 1: Models and applications. John Wiley & Sons.
48. Lomax, K. S. (1954). Business failures: Another example of the analysis of failure data. *Journal of the American Statistical Association, 49*(268), 847–852.
49. Cox, D. R., & Lewis, P. A. (1996). *Statistical analysis of series of events*. U.K.: Methuem.
50. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583–639.
51. Williams, M., & Konovalov, V. (2008). Central Asia temperature and precipitation data, 1879–2003. Boulder, Colorado: USA National Snow and Ice Data Center. https://doi.org/10.7265/N5NK3BZ8. [Accessed in 10/20/2020].
52. Carlin, B. P., & Louis, T. A. (2000). Bayes and empirical Bayes methods for data analysis. Chapman & Hall/CRC.
53. Laub, P. J., Taimre, T., & Pollett, P. K. (2015). Hawkes processes. In arXiv:1507.02822v1.