SCIENTIFIC REPORTS

Nonstationary Temperature-Duration-Frequency curves

Taha B. M. J. Ouarda & Christian Charron

Persistent extreme heat events are of growing concern in a climate change context. An increase in the intensity, frequency and duration of heat waves is observed in several regions. Temperature extremes are also influenced by global-scale modes of climate variability. Temperature-Duration-Frequency (TDF) curves, which relate the intensity of heat events of different durations to their frequencies, can be useful tools for the analysis of heat extremes. To account for climate external forcings, we develop a nonstationary approach to the TDF curves by introducing indices that account for the temporal trend and teleconnections. Nonstationary TDF modeling can find applications in adaptive management in the fields of health care, public safety and energy production. We present a one-step method, based on the maximization of the composite likelihood of observed heat extremes, to build the nonstationary TDF curves. We show the importance of integrating the information concerning climate change and climate oscillations. In an application to the province of Quebec, Canada, the influence of Atlantic Multidecadal Oscillations (AMO) on heat events is shown to be more important than the temporal trend.

Extended periods of extreme temperature have significant adverse impacts on public health, infrastructure and natural ecosystems. It is well established that mean temperatures have increased globally since the middle of the 20th century[1]. This phenomenon has been attributed to the human influence on climate[2]. Increase in global mean temperature has resulted in an even larger increase in the probabilities of extreme temperature events[3–5]. Recently, a large number of heat waves occurred in different regions[6,7], causing considerable damage. It is expected that the frequency, duration and intensity of extreme heat events will increase in a future warmer climate[8–10]. Aside from climate change, global-scale modes of climate variability have also important influences on temperature extremes around the world. The links between temperature extremes and circulation patterns have been demonstrated in a number of studies[11–13].

Rainfall Intensity-Duration-Frequency (IDF) curves are widely used tools for the planning, design and operation of water resources infrastructure[14,15]. IDF curves relate rainfall intensities corresponding to different durations to a series of return periods. A similar concept applied to temperature extremes[16], referred to in this work as the Temperature-Duration-Frequency (TDF) curves approach, would be of interest for managers in a number of fields including healthcare, public security, agriculture and energy production.

In the only previous work on TDF curves, it was assumed that the characteristics of the probability distribution of extreme heat events are invariant through time[16]. However, temperature extreme characteristics change with time due to the influence of climate change and climate oscillation patterns. In this context, it is important to develop models for temperature extremes that account for external forcings. One approach to handle nonstationarity in temperature extremes is to make the parameters of the temperature extremes distribution dependent on covariates representing climate variability and change[17]. A covariate representing time can be used to account for the eventual temporal trend caused by climate change, and covariates representing climate indices can be introduced to take into consideration the influence of climate oscillation patterns.

The mathematical formulation of the non-stationary TDF curves presented here follows the one of the general IDF relationship (see Methods) but extended to the non-stationary case. In this formulation, a single model is defined in which the duration is a parameter. This model differs from other recent works introducing non-stationarity in IDF curves where an independent model is defined for each duration. Maximum temperatures are modeled with the Generalized Extreme Value (GEV) distribution, a 3-parameter distribution widely used to model climatic extremes.

Such a nonstationary tool would be particularly useful in adaptive management, a structure process of decision making in condition of uncertainty[18]. This concept has been used, for instance, in the fields of agriculture[19], health care[20] and water resources[21,22]. Adaptive management is expected to become an important tool for public

Canada Research Chair in Statistical Hydro-Climatology, INRS-ETE, 490 de la Couronne, Québec, QC, G1K 9A9, Canada. Correspondence and requests for materials should be addressed to T.B.M.J.O. (email: taha.ouarda@ete.inrs.ca)

Station	Latitude	Longitude	Period	μ	σ	κ	θ	η
Montréal Intl. Airport	45.47	−73.74	1941–2015	35.48	1.71	−0.22	1.94	0.09
Quebec Intl. Airport	46.80	−71.38	1943–2016	36.52	1.83	−0.23	2.77	0.11
Sherbrooke	45.44	−71.69	1962–2015	36.88	1.68	−0.35	3.54	0.12
Rimouski	48.45	−68.52	1953–2016	33.54	1.97	−0.30	1.05	0.11
Bagotville	48.33	−71.00	1943–2016	36.26	2.12	−0.31	1.35	0.12
Ste-Anne-de-la-Pérade	46.58	−72.23	1950–2016	35.03	1.76	−0.17	1.99	0.10

Table 1. Description of selected stations and parameters of the stationary TDF curves.

health policy makers where knowledge concerning trends in climate extremes can be used for mitigation and preparedness[23]. Non stationary TDF curves can also be used to integrate information concerning future temperature extreme characteristics in the design, operation, and safety analysis of thermal and nuclear power plants[24].

Results

Modelling of extreme temperature with nonstationary TDF curves.
The nonstationary TDF approach proposed here is applied to observed daily maximum temperatures at meteorological stations in the Province of Quebec (Canada). Six meteorological stations with long series and recent data were selected. Table 1 presents the selected stations with their coordinates and the parameters obtained for the stationary TDF approach. To build the TDF curves, the annual maximum of multi-day averages of daily maximum temperatures over durations of 1 to 7 and 10 days (a total of 8 durations) during the summer season (June-August) were extracted for each station. The slope of the annual maximum temperature time series was computed for each duration using the Theil-Sen method. The slopes corresponding to all durations have a positive sign at all stations except for the station of Quebec Intl. Airport where the slopes are negative for all durations (Supplementary Table 1). An analysis of the trends with the non-parametric Mann-Kendall statistical test reveals that, while trends are not generally significant, positive significant trends are observed for most of the durations for the stations of Sherbrooke and Rimouski.

The Atlantic Multidecadal Oscillation (AMO) is a large scale pattern of multidecadal variability related to variations of the sea surface temperature in the North Atlantic Ocean. Several examples of regional multidecadal climate variability have been related to the AMO including for the North American and European summer climate[25]. AMO is known to have the most important influence on summer temperature extremes in North America[26] and is selected here to be included as covariate in the nonstationary TDF approach. Annual AMO time series for the summer season are obtained by averaging the monthly values over the summer months (June-August) of each year. Correlations of summer AMO with the annual maximum temperatures for the different durations at the six stations are computed and the Student's t-test reveals that correlations are significant at a significance level of 5% for most of the durations (Supplementary Table 2).

In the nonstationary framework, we consider the location and scale parameters of the GEV distribution in the TDF relationship to be dependent upon the covariates. The variables "Time" representing the year, and the climate index AMO for the summer season, are used as covariates (see Methods). One nonstationary model uses Time as covariate (denoted TDF model "Time"), a second uses AMO (denoted TDF model "AMO") and a third model uses both covariates together (denoted TDF model "Time + AMO"). The parameters of the TDF relationships are estimated with the maximum composite likelihood method (see Methods for details). The criterion CL-AIC (see equation (14) in Methods), an analogue of the Akaike information criterion (AIC), is used for model comparison.

Table 2 presents the maximized independence log-likelihood, the CL-AIC statistic and the model parameters for each TDF model and each station. For a given station and a given selection of covariates, the table presents only the best model among the linear and quadratic relationships presented in equations (6–9) according to the CL-AIC statistic. Table 2 shows that the linear relationship is selected by CL-AIC in all cases except in the case of Montréal where a quadratic relationship is selected when the covariate is time. In all cases, with respect to the CL-AIC criterion, the goodness-of-fit is improved when a nonstationary model is used. For the majority of stations, the model with only AMO obtains the best CL-AIC values. This means, that for most stations, the influence of the climate oscillation pattern is more important than the temporal trend. This result is in agreement with the fact that correlations with the AMO index are more significant than temporal trends. For the Quebec and the Rimouski stations, the best goodness-of-fit is obtained with a combination of covariates Time and AMO. This indicates that in some cases the combination of the two covariates has significant impacts on extreme temperatures.

Stationary TDF curves and similar relationships (e.g. IDF curves) are generally represented on graphs with the intensity plotted against the duration where each curve represents a return period. Such representation is not possible with nonstationary TDF curves. With one covariate, TDF surfaces can be well defined and it is possible to represent them with 3D graphs of the intensity against the duration and the covariate. Two examples of such graphs are presented in Fig. 1a for Sherbrooke with the TDF model "Time" and in Fig. 1c for Montréal with the TDF model "AMO". Curves of the 3-day maximum temperatures against the covariate are also illustrated for the same examples in Fig. 1b–d. These curves represent cross sections of the TDF surfaces for fixed durations. A positive temporal trend with a decreasing variance can be observed for Sherbrooke and a linear positive relationship with AMO can be observed for Montréal. For Sherbrooke, the location and the scale parameters are related linearly to the covariate Time, and for Montréal, the location parameter is related linearly to the covariate AMO. With two covariates, it is possible to represent TDF surfaces with 3D graphs of the intensity against each covariate for separate durations or separate return periods. Examples of such illustrations are presented in Supplementary Fig. 10, where the 3-day maximum temperatures are plotted against Time and AMO for each station.

| Station	Model	ℓ_{ind}	CL-AIC	Model parameters
Montréal Intl. Airport	Stationary	−1084.2	2205.5	$\mu, \sigma, \kappa, \theta, \eta$
	Time	−1068.0	2200.7	$\mu_l = \mu_0 + \mu_1 Time + \mu_2 Time^2, \sigma, \kappa, \theta, \eta$
	AMO	−1064.7	**2179.7**	$\mu_l = \mu_0 + \mu_1 AMO, \sigma, \kappa, \theta, \eta$
	Time + AMO	**−1064.7**	2194.3	$\mu_l = \mu_0 + \mu_1 Time + \mu_2 AMO, \sigma, \kappa, \theta, \eta$
Quebec Intl. Airport	Stationary	−1083.7	2203.1	$\mu, \sigma, \kappa, \theta, \eta$
	Time	−1074.7	2197.2	$\mu, \sigma_l = \sigma_0 + \sigma_1 Time, \kappa, \theta, \eta$
	AMO	−1068.7	2187.8	$\mu_l = \mu_0 + \mu_1 AMO, \sigma, \kappa, \theta, \eta$
	Time + AMO	**−1046.3**	**2175.8**	$\mu_l = \mu_0 + \mu_1 Time + \mu_2 AMO,$ $\sigma_l = \sigma_0 + \sigma_1 Time + \sigma_2 AMO, \kappa, \theta, \eta$
Sherbrooke	Stationary	−704.9	1443.4	$\mu, \sigma, \kappa, \theta, \eta$
	Time	−686.2	1429.6	$\mu_l = \mu_0 + \mu_1 Time, \sigma_l = \sigma_0 + \sigma_1 Time, \kappa, \theta, \eta$
	AMO	−674.2	**1394.8**	$\mu_l = \mu_0 + \mu_1 AMO, \sigma, \kappa, \theta, \eta$
	Time + AMO	**−663.8**	1404.7	$\mu_l = \mu_0 + \mu_1 Time + \mu_2 AMO,$ $\sigma_l = \sigma_0 + \sigma_1 Time + \sigma_2 AMO, \kappa, \theta, \eta$
Rimouski	Stationary	−912.7	1856.0	$\mu, \sigma, \kappa, \theta, \eta$
	Time	−884.2	1815.6	$\mu_l = \mu_0 + \mu_1 Time, \sigma_l = \sigma_0 + \sigma_1 Time, \kappa, \theta, \eta$
	AMO	−893.4	1829.2	$\mu_l = \mu_0 + \mu_1 AMO, \sigma, \kappa, \theta, \eta$
	Time + AMO	**−872.2**	**1810.4**	$\mu_l = \mu_0 + \mu_1 Time + \mu_2 AMO,$ $\sigma_l = \sigma_0 + \sigma_1 Time + \sigma_2 AMO, \kappa, \theta, \eta$
Bagotville	Stationary	−1139.6	2309.1	$\mu, \sigma, \kappa, \theta, \eta$
	Time	−1139.5	2317.4	$\mu, \sigma_l = \sigma_0 + \sigma_1 Time, \kappa, \theta, \eta$
	AMO	−1123.8	**2289.6**	$\mu_l = \mu_0 + \mu_1 AMO, \sigma, \kappa, \theta, \eta$
	Time + AMO	**−1123.8**	2302.8	$\mu_l = \mu_0 + \mu_1 Time + \mu_2 AMO, \sigma, \kappa, \theta, \eta$
Ste-Anne-de-la-Pérade	Stationary	−988.6	2016.7	$\mu, \sigma, \kappa, \theta, \eta$
	Time	−984.6	2021.4	$\mu, \sigma_l = \sigma_0 + \sigma_1 Time, \kappa, \theta, \eta$
	AMO	−959.5	**1971.2**	$\mu_l = \mu_0 + \mu_1 AMO, \sigma, \kappa, \theta, \eta$
	Time + AMO	**−958.3**	1982.9	$\mu_l = \mu_0 + \mu_1 Time + \mu_2 AMO, \sigma, \kappa, \theta, \eta$

Table 2. Statistics and model parameters for the stationary and nonstationary TDF curves.

Impacts of considering nonstationarity. To illustrate the impacts of considering nonstationary TDF curves instead of the classical stationary model, Fig. 2 presents, for each station, graphs of the 10-year quantiles versus the duration for the stationary TDF model and the nonstationay TDF models for values of the covariates representing different scenarios. For the nonstationary TDF model with Time as covariate, the case for the last year of record is considered. This model represents the most probable TDF relationship for the last year of record when only the temporal trend is considered. For the nonstationary TDF models "AMO" and "Time + AMO", the years with the largest and lowest values of AMO during the period 1941–2016 are identified to illustrate extreme cases (1974 and 1998 for which AMO was respectively at the values −0.47 and 0.51). For the nonstationary TDF model "AMO", the two observed extreme values of AMO during these two years were considered. Finally, for the nonstationary TDF model "Time + AMO", these two years and their corresponding AMO values were considered.

Figure 2 illustrates the importance of considering climate variability when building TDF curves, as large discrepancies in quantiles are obtained for different years and for different nonstationary models. The mean difference between the quantiles of the two years selected for the TDF model "Time + AMO" is about 1.8 °C. In general, the climate index has much more impact than the temporal trend on the quantiles. For instance, quantiles obtained with the TDF model "Time" at the Sherbrooke, Rimouski and Bagotville stations are very similar to the ones obtained with the stationary TDF model. On the other hand, quantiles obtained with models including AMO are very different from the ones obtained with the stationary models. In general, the TDF model "Time + AMO" leads to very similar quantiles to the ones obtained with the TDF model "AMO". The only exceptions are for the Quebec and Rimouski stations where the quantiles for the years 1974 and 1998 are distinct for these two models. These two stations are also the ones for which the TDF model "Time + AMO" obtains better goodness-of-fit statistics than the TDF model "AMO". These last results can be explained by the fact that no important temporal trends are detected in the majority of time series while significant correlations are observed with AMO for the majority of durations for all stations.

Discussion

This study proposes a nonstationary approach to the development of TDF curves. With this approach, parameters of the TDF relationship are conditional upon time-dependent covariates representing climate change and climate oscillation patterns (Climate variability). We found that the goodness-of-fit is improved when using a nonstationary approach with covariates related to climate variability. For most stations in the study area, the impacts of considering the temporal trend on quantiles are lower than the impacts of considering the climate index AMO. This illustrates the importance of introducing climate indices in nonstationary models. Future work may focus on the introduction of additional relevant low frequency climate oscillation indices in the nonstationary TDF model. For the study area considered in the present work, few significant trends are detected in the annual maximum
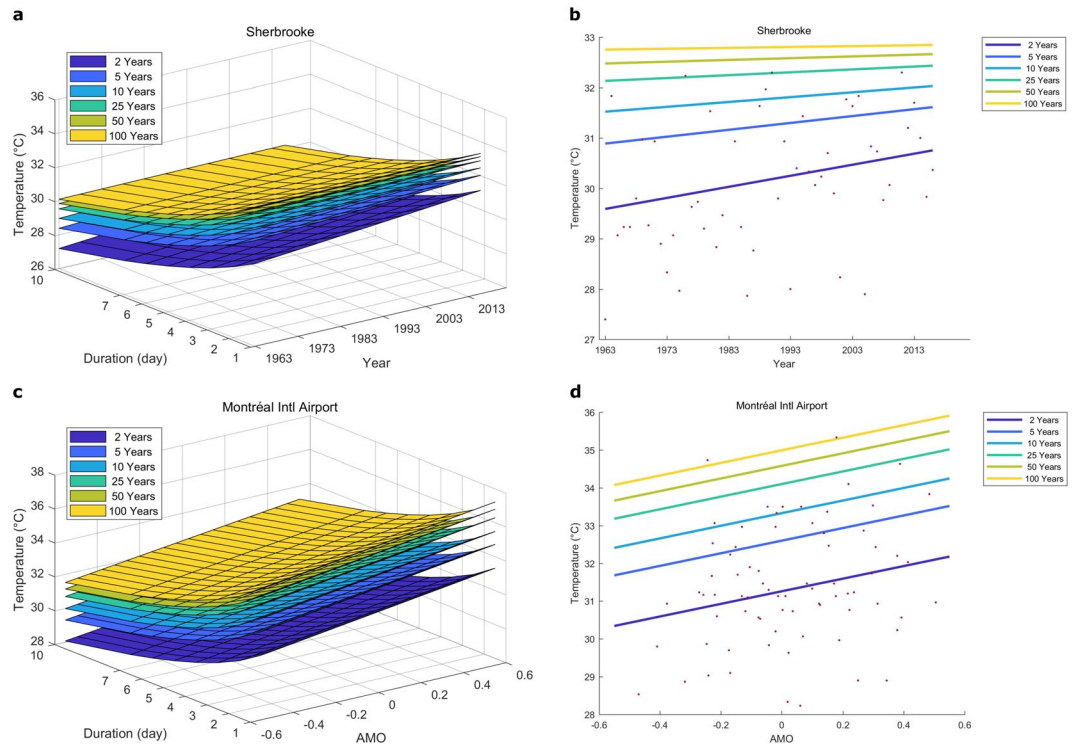
Figure 1. Nonstationary TDF surfaces and 3-day maximum temperatures against the covariate at Sherbrooke and Montréal stations. Nonstationary TDF surfaces are presented for the stations at Sherbrooke (**a**) and Montréal (**c**) with the covariates Time and AMO respectively. The 3-day maximum temperatures are represented for the stations of Sherbrooke (**b**) and Montréal (**d**) against the covariates Time and AMO respectively. Red dots represent observations. (**b,d**) are cross sections in (**a,c**) respectively.

temperature series. However, this may not be the case in other parts of the world where increasing heat wave frequency and intensity are reported[7,10,13].

Nonstationarity in TDF curves, IDF curves or similar tools have not been introduced until recently[14,27]. The general procedure presented in this work can be applied to other variables such as precipitations (IDF curves) and floods (QDF curves) for which the influence of climate variability is relatively well documented[28–30]. The nonstationary TDF model can be useful in adaptive management where predictions of the evolution of extreme values are needed. This approach will allow integrating information concerning future extreme temperature characteristics in a number of applications such as the analysis of the safety of thermal and nuclear power plants. Indeed, the efficiency and safety of thermal and nuclear power plants are heavily affected by heatwaves by increasing the scarcity and temperature of cooling water and by reducing thermal efficiency[24].

The approach proposed herein is useful if nonstationay models can be built for the near future. To achieve this, the covariate related to the temporal trend can be easily extrapolated. However, in the present approach, the climate index used is measured during the same period as the extreme temperature events (i.e. June-July-August). This is a limitation for prediction as the value of the climate index is unknown in the future. However, it is possible to predict climate indices[26,31] and use these predictions in nonstationary TDF curves.

Methods

Nonstationary TDF relationship formulation. To build TDF curves, the time series of maximum average temperatures are derived from the daily maximum observed temperatures. For each year, the maximum average temperature value is extracted for a number of D durations d_k, $k = 1, ..., D$. The maximum average temperature for year l, denoted by $T_l(d)$, $l = 1, ..., n$, is obtained using a moving average window during the summer season of June-August (JJA). In this formulation, n is the number of years with measurements and $d \in [d_1, d_D]$.

The return level of $T(d)$ for a return period of R years, denoted by $t_R(d)$, is given by:

$$t_R(d) = \frac{a(R)}{b(d)}. \tag{1}$$

In this formulation, $t_R(d)$ has a separate functional dependence on the return period R and the duration d. The function $a(R)$, determined by the distribution of $T(d)$, defines curves that are parallel for different return periods R. The function $b(d)$ models the shape of the TDF curves as a function of the duration d and can be expressed by:

$$b(d) = (d + \theta)^\eta, \tag{2}$$
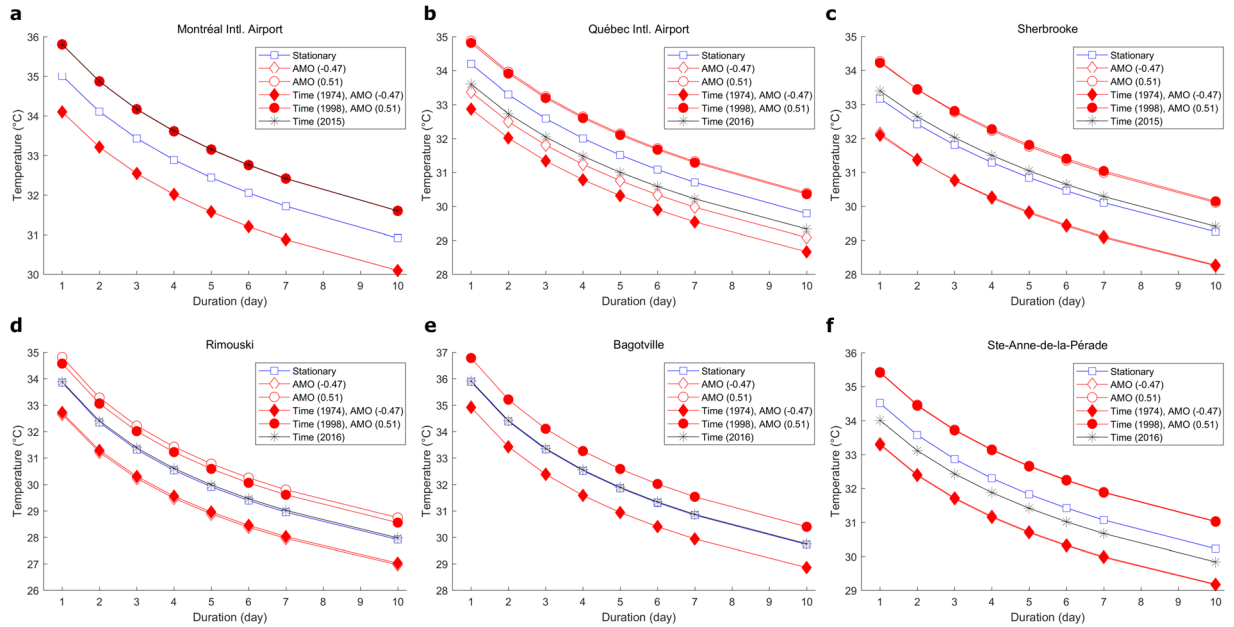
Figure 2. Comparison of the 10-year quantiles for the stationary TDF model and the nonstationary TDF models. 10-year quantiles for the stationary TDF model, the nonstationary TDF model Time for the case of the last year of record, the nonstationary TDF model AMO for the largest and lowest observed values of AMO (-0.47 and 0.51) and the nonstationary TDF model Time + AMO for the 2 years with the largest and lowest observed values of AMO (1974 and 1998) and their corresponding values.

where θ and η are the shape parameters subject to the inequality constrains that $\theta > 0$ and $0 < \eta < 1$.

If $F_{T(d)}(t; d)$ denotes the distribution function of $T(d)$, the scaled maximum average temperature $Y = T(d) \, b(d)$ will also be distributed as $F_{T(d)}(t; d)$ (i.e. $F_{T(d)}(t; d) = F_Y(y_R) = 1 - \frac{1}{R}$). Since $a(R)$ is the return level of $T(d)$, the following expression is obtained:

$$a(R) = F_Y^{-1}\left(1 - \frac{1}{R}\right).$$

(3)

The Gumbel and the GEV are the probability distributions that are the most widely used to model climate extremes. The GEV is the limiting distribution for block maxima and includes the Gumbel distribution as a subset[32]. The GEV is used here to model $T(d)$. The cumulative distribution function of the GEV is given by:

$$F(x) = \exp\left\{-\left[1 + \kappa\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\kappa}\right\},$$

(4)

where μ, σ and κ are the location, scale and shape parameters respectively. $F(x)$ is defined for $1 + \kappa(x - \mu)/\sigma > 0$ where $\sigma > 0$. The general TDF relationship assuming that $T(d)$ follows the GEV distribution is then given by the following expression:

$$t_R(d) = \frac{a(R)}{b(d)} = \frac{\mu - \frac{\sigma}{\kappa}\left\{1 - \left[-\log\left(1 - \frac{1}{R}\right)\right]^{-\kappa}\right\}}{(d + \theta)^{\eta}}.$$

(5)

In the nonstationary case, the distribution parameters are made dependent on covariates which can represent a climate oscillation index or time. The shape parameter of the GEV is usually kept constant in nonstationary analysis[33,34]. In this study, all shape parameters, including the parameter κ, θ and η, are also kept constant. Let us denote U_l and V_l, the values during the lth year of two time-dependent covariates U and V. For non-stationary models with one covariate, the location parameter can be stationary or can depend linearly or quadratically on U_l:

$$\mu_l = \begin{cases} \mu_0 \\ \mu_0 + \mu_1 U_l \\ \mu_0 + \mu_1 U_l + \mu_2 U_l^2 \end{cases}$$

(6)

and the scale parameter can be stationary or can depend linearly on U_l:

$$\sigma_l = \begin{cases} \sigma_0 \\ \sigma_0 + \sigma_1 U_l \end{cases}. \tag{7}$$

For non-stationary models with two covariates, the location parameter can be stationary or can depend linearly or quadratically on U_l and V_l:

$$\mu_l = \begin{cases} \mu_0 \\ \mu_0 + \mu_1 U_l + \mu_2 V_l \\ \mu_0 + \mu_1 U_l + \mu_2 U_l^2 + \mu_3 V_l \\ \mu_0 + \mu_1 U_l + \mu_2 V_l + \mu_3 V_l^2 \\ \mu_0 + \mu_1 U_l + \mu_2 U_l^2 + \mu_3 V_l + \mu_4 V_l^2 \end{cases} \tag{8}$$

and the scale parameter can be stationary or can depend linearly on U_l and V_l:

$$\sigma_l = \begin{cases} \sigma_0 \\ \sigma_0 + \sigma_1 U_l + \sigma_2 V_l \end{cases}. \tag{9}$$

The non-stationary models are built by considering any combinations of the models in equations (6) and (7) for one covariate and any combinations of the models in equations (8) and (9) for two covariates (the cases where both distribution parameters are constant are excluded since this is equivalent to the stationary model). For models with two covariates, it is not allowed for $\mu_l(\sigma_l)$ to depend only on one covariate and $\sigma_l(\mu_l)$ to depend only on the other covariate (see equations (8 and 9)). The relations of the distribution parameter σ_l with covariates are made constant or linear for the aim of obtaining simpler models considering that the number of parameters to be fitted increases rapidly with the complexity of the model, and that the size of the record may become a limiting factor. Parsimony considerations need to be taken into account in the selection of the optimum model to be adopted. The two covariates introduced in the nonstationary approach are the climate index AMO during summer and the year number denoted by Time. To obtain the AMO time series, the average of the AMO values over the months of the summer season (JJA) is computed. The covariate Time is defined by a series of integers incremented from 1 to the number of years of observed data.

Maximum composite likelihood method. The vectors of the distribution parameters $\psi = (\mu, \sigma, \theta, \eta)$ and $\psi = (\mu_0, \mu_1, \ldots, \sigma_0, \sigma_1, \ldots, \kappa, \theta, \eta)$ need to be estimated for stationary IDF curves and non-stationary IDF curves respectively. The method used here for the estimation of the parameters is the maximum composite likelihood. The characteristics and formulation of the nonstationary TDF model make it necessary to adopt this method instead of the classical maximum likelihood approach. Indeed, while observations from year to year are independent, the maximum temperatures over the different durations for the same year are dependent. Let us define $f(t; \psi, \alpha)$, the joint probability density of the random vector $T = \{T(d_1), \ldots, T(d_D)\}$ where α is a parameter vector that parameterizes the interdependence between temperatures corresponding to different durations and ψ parameterizes the marginal structure[35]. The full likelihood is then given by:

$$L(\psi; \ t) = \prod_{l=1}^{n} f(t_{1l}, \ \ldots, \ t_{Dl}; \ \psi, \ \alpha), \tag{10}$$

where t_{kl} denotes the maximum average temperature for the year l and for the duration group k. However, the joint density $f(t; \psi, \alpha)$ is unknown, making the estimation of the full likelihood difficult. To overcome this difficulty, a simplified likelihood function for TDF curves is obtained by assuming the independence of the temperatures over the different durations[36,37]. This function is referred to as the independence likelihood[35] and is given by:

$$L_{ind}(\psi; \ t) = \prod_{j=1}^{D} \prod_{l=1}^{n} f(t_{jl}; \ \psi), \tag{11}$$

where $f(t; \psi)$ represents the density function of $T(d)$. The independence likelihood can be considered as the simplest case of a composite likelihood defined as an inference function derived by multiplying a set of component likelihoods[38]. Composite likelihood is used in several applications as surrogate for the full likelihood when it is too cumbersome or impractical to compute[39]. If $T(d)$ is assumed to follow a GEV distribution, the probability function of $T(d)$ is given by[37]:

$$T(d) \sim \text{GEV}(\mu(d), \ \sigma(d), \ \kappa). \tag{12}$$

In the non-stationary case, the distribution parameters $\mu(d)$ and $\sigma(d)$ are expressed by:

$$\mu_l(d) = \frac{\mu_l}{(d + \theta)^\eta}, \ \sigma_l(d) = \frac{\sigma_l}{(d + \theta)^\eta}. \tag{13}$$

In practice, the log likelihood $\ell_{ind}(\psi; \ t) = \log L_{ind}(\psi; \ t)$ is maximized instead of $L_{ind}(\psi; t)$ with an optimization procedure. The optimization function *fmincon* in Matlab® is used to find $\hat{\psi}$, the estimate of ψ that maximizes

$\ell_{ind}(\psi; \ t)$. The algorithm used (interior-point) ensures that $\sigma_l > 0$: if $\sigma_l < 0$ for an iteration, the objective function returns a "not-a-number" value, the iterate is rejected and the next step is attempted.

For model comparison, information criteria such as the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) are frequently used. Such criteria account for the goodness-of-fit and penalize more complex models. Let $H(\psi) = -E[\nabla^2 \ell_{ind}(\psi; \ T)]$ be the sensitivity or Hessian matrix and $J(\psi) = \text{Var}\{\nabla \ell_{ind}(\psi; \ T)\}$ be the variability matrix. Because of the assumption of independence among the likelihood terms in the definition of the independence likelihood, composite likelihood can be seen as a misspecified likelihood. In that case, the second Bartlett identity fails (i.e. $H(\psi) \neq J(\psi)$) and classical criteria should be generalized. Analogous criteria for AIC and BIC based on composite likelihoods were introduced[39] and have the following forms[38]:

$$\text{CL-AIC} = -2\ell_{ind}(\hat{\psi}; \ t) + 2\text{dim}(\psi), \tag{14}$$

$$\text{CL-BIC} = -2\ell_{ind}(\hat{\psi}; \ t) + \text{dim}(\psi)\log(n), \tag{15}$$

where $\text{dim}(\psi)$ is the effective number of parameters estimated by $\text{tr}\{J(\psi)H(\psi)^{-1}\}$. The sample estimate of the sensitivity matrix H can be obtained by:

$$\hat{H}(\psi) = -\frac{1}{n}\sum_{l=1}^{n} \nabla u(\hat{\psi}; \ t_l), \tag{16}$$

where $u(\psi; \ t_l) = \nabla \ell_{ind}(\psi; \ t_l)$ and t_l denotes the vector of maximum average temperatures for the year l. However, the cumbersome computation of the Hessians in equation (16) can be avoided. Given that the second Bartlett identity is valid for each individual likelihood term[38], the matrix H can be obtained by the following sample estimate:

$$\hat{H}(\psi) = \frac{1}{n}\sum_{k=1}^{D}\sum_{l=1}^{n} u(\hat{\psi}; \ t_{kl})u(\hat{\psi}; \ t_{kl})^{\text{T}}. \tag{17}$$

The sample estimate of the variability matrix J can be obtained empirically by[38]:

$$\hat{J}(\psi) = \frac{1}{n}\sum_{l=1}^{n} u(\hat{\psi}; \ t_l)u(\hat{\psi}; \ t_l)^{\text{T}}. \tag{18}$$

Data Availability

The maximum daily temperature data for Canada is freely available online from the Government of Canada: http://climate.weather.gc.ca/historical_data/search_historic_data_e.html. The monthly values of the climate index AMO are freely available online from the Earth System Research Laboratory of the National Oceanic and Atmospheric Administration: https://www.esrl.noaa.gov/psd/data/timeseries/AMO/.

References

1. Hansen, J., Sato, M. & Ruedy, R. Perception of climate change. *Proc. Natl. Acad. Sci. USA* **109**, E2415–E2423 (2012).
2. Coumou, D. & Robinson, A. Historic and future increase in the global land area affected by monthly heat extremes. *Environ. Res. Lett.* **8**, 034018 (2013).
3. Mearns, L. O., Katz, R. W. & Schneider, S. H. Extreme High-Temperature Events: Changes in their probabilities with Changes in MeanTemperature. *J. Clim. Appl. Meteorol.* **23**, 1601–1613 (1984).
4. Katz, R. W. & Brown, B. G. Extreme events in a changing climate: Variability is more important than averages. *Clim. Change* **21**, 289–302 (1992).
5. Perkins, S. E., Alexander, L. V. & Nairn, J. R. Increasing frequency, intensity and duration of observed global heatwaves and warm spells. *Geophys. Res. Lett.* **39**, L20714 (2012).
6. Rohini, P., Rajeevan, M. & Srivastava, A. K. On the variability and increasing trends of heat waves over India. *Sci. Rep.* **6**, 26153 (2016).
7. Christidis, N., Jones, G. S. & Stott, P. A. Dramatically increasing chance of extremely hot summers since the 2003 European heatwave. *Nature Clim. Change* **5**, 46–50 (2015).
8. Coumou, D. & Rahmstorf, S. A decade of weather extremes. *Nature Clim. Change* **2**, 491–496 (2012).
9. IPCC. *Managing the risks of extreme events and disasters to advance climate change adaptation. A special report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC)* [Field, C. B. *et al.* (eds)]. 13 (Cambridge University Press, 2012).
10. Basha, G. *et al.* Historical and projected surface temperature over India during the 20th and 21st century. *Sci. Rep.* **7**, 2987 (2017).
11. Loikith, P. C. & Broccoli, A. J. The influence of recurrent modes of climate variability on the occurrence of winter and summer extreme temperatures over North America. *J. Clim.* **27**, 1600–1618 (2014).
12. Myoung, B., Kim, S. H., Kim, J. & Kafatos, M. C. On the relationship between the North Atlantic Oscillation and early warm season temperatures in the southwestern United States. *J. Clim.* **28**, 5683–5698 (2015).
13. Kumar, K. N., Molini, A., Ouarda, T. B. M. J. & Rajeevan, M. N. North Atlantic controls on wintertime warm extremes and aridification trends in the Middle East. *Sci. Rep.* **7**, 12301 (2017).
14. Yousef, L. A. & Ouarda, T. B. M. J. Adaptation of water resources management to changing climate: the role of Intensity-Duration-Frequency curves. *Int. J. Environ. Sci. Dev.* **6**, 478 (2015).
15. Madsen, H., Arnbjerg-Nielsen, K. & Mikkelsen, P. S. Update of regional intensity–duration–frequency curves in Denmark: Tendency towards increased storm intensities. *Atmos. Res.* **92**, 343–349 (2009).
16. Khaliq, M. N., St-Hilaire, A., Ouarda, T. B. M. J. & Bobée, B. Frequency analysis and temporal pattern of occurrences of southern Quebec heatwaves. *Int. J. Climatol.* **25**, 485–504 (2005).
17. Strupczewski, W. G., Singh, V. P. & Feluch, W. Non-stationary approach to at-site flood frequency modelling I. Maximum likelihood estimation. *J. Hydrol.* **248**, 123–142 (2001).

18. Ebi, K. Climate change and health risks: assessing and responding to them through 'adaptive management'. *Health Aff.* **30**, 924–930 (2011).
19. Fraisse, C. W. *et al*. AgClimate: A climate forecast information system for agricultural risk management in the southeastern USA. *Comput. Electron. Agric.* **53**, 13–27 (2006).
20. Ebi, K. L., Lewis, N. D. & Corvalan, C. Climate variability and change and their potential health effects in small island states: information for adaptation planning in the health sector. *Environ. Health Perspect.* **114**, 1957–1963 (2006).
21. Pulwarty, R. S. & Melis, T. S. Climate extremes and adaptive management on the Colorado River: Lessons from the 1997–1998 ENSO event. *J. Environ. Manage.* **63**, 307–324 (2001).
22. Ouarda, T. B. M. J. & Labadie, J. W. Chance-constrained optimal control for multireservoir system optimization and risk analysis. *Stoch. Environ. Res. Risk Assess.* **15**, 185–204 (2001).
23. Paterson, J., Berry, P., Ebi, K. & Varangu, L. Health care facilities resilient to climate change impacts. *Int. J. Environ. Res. Publ. Health* **11**, 13097–13116 (2014).
24. Linnerud, K., Mideksa, T. K. & Eskeland, G. S. The impact of climate change on nuclear power supply. *Energy J.* **32**, 149–168 (2011).
25. Knight, J. R., Folland, C. K. & Scaife, A. A. Climate impacts of the Atlantic Multidecadal Oscillation. *Geophys. Res. Lett.* **33**, L17706 (2006).
26. Sutton, R. T. & Hodson, D. L. R. Atlantic ocean forcing of North American and European summer climate. *Science* **309**, 115–118 (2005).
27. Cheng, L. & AghaKouchak, A. Nonstationary precipitation Intensity-Duration-Frequency curves for infrastructure design in a changing climate. *Sci. Rep.* **4**, 7093 (2014).
28. Ouachani, R., Bargaoui, Z. & Ouarda, T. B. M. J. Power of teleconnection patterns on precipitation and streamflow variability of upper Medjerda Basin. *Int. J. Climatol.* **33**, 58–76 (2013).
29. Thiombiano, A. N., St-Hilaire, A., El Adlouni, S.-E. & Ouarda, T. B. M. J. Nonlinear response of precipitation to climate indices using a non-stationary Poisson-generalized Pareto model: case study of southeastern Canada. *Int. J. Climatol* **38**, e875-e888 (2018).
30. Chandran, A., Basha, G. & Ouarda, T. B. M. J. Influence of climate oscillations on temperature and precipitation over the United Arab Emirates. *Int. J. Climatol.* **36**, 225–235 (2016).
31. Lee, T. & Ouarda, T. B. M. J. Prediction of climate nonstationary oscillation processes with empirical mode decomposition. *J. Geophys. Res. Atmos.* **116**, D06107 (2011).
32. Coles, S. *An introduction to statistical modeling of extreme values*. (Springer, 2001).
33. El Adlouni, S., Ouarda, T. B. J. M., Zhang, X., Roy, R. & Bobee, B. Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resour. Res.* **43**, W03410 (2007).
34. Katz, R. W., Parlange, M. B. & Naveau, P. Statistics of extremes in hydrology. *Adv. Water Resour.* **25**, 1287–1304 (2002).
35. Chandler, R. E. & Bate, S. Inference for clustered data using the independence loglikelihood. *Biometrika* **94**, 167–183 (2007).
36. Van de Vyver, H. Bayesian estimation of rainfall intensity–duration–frequency relationships. *J. Hydrol.* **529**, 1451–1463 (2015).
37. Muller, A., Bacro, J.-N. & Lang, M. Bayesian comparison of different rainfall depth–duration–frequency relationships. *Stoch. Environ. Res. Risk Assess.* **22**, 33–46 (2008).
38. Varin, C., Reid, N. & Firth, D. An overview of composite likelihood methods. *Stat. Sin.* **21**, 5–42 (2011).
39. Varin, C. & Vidoni, P. A note on composite likelihood inference and model selection. *Biometrika* **92**, 519–528 (2005).

Acknowledgements

Author Contributions

T.B.M.J.O. and C.C. performed the analysis and contributed to the writing of the paper.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-33974-y.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.