Title: Nonlinear response of precipitation to climate indices using a nonstationary Poisson-Generalized Pareto model: Case study of Southeastern Canada

Short title: Nonlinear and nonstationary Poisson-Generalized Pareto model

Alida N. Thiombiano *[(1)], André St-Hilaire [(1)], Salah-Eddine El Adlouni [(2)], Taha B.M.J. Ouarda [(1)]

[(1)] Centre Eau–Terre-Environnement, Institut National de la Recherche Scientifique, 490, de la Couronne Street, Québec, QC, G1K 9A9, Canada

[(2)] Université de Moncton, 18 Antonine-Maillet Avenue, Moncton, NB, Canada E1A 3EA

Corresponding author: (*)

E-mail address: alida_nadege.thiombiano@ete.inrs.ca; thiombianoalida@yahoo.fr

Mobile phone number: (403) 700-1373

Fax number: 418 654 2600

Abstract

Quantile estimates are generally interpreted in association with the return period concept in practical engineering. To do so with the peaks-over-threshold (POT) approach, combined Poisson-Generalized Pareto distributions (referred to as PD-GPD model) must be considered. In this paper, we evaluate the incorporation of nonstationarity in the Generalized Pareto distribution (GPD) and the Poisson distribution (PD) using respectively the smoothing-based B-spline functions and the logarithmic link function. Two models are proposed, a stationary PD combined to a nonstationary GPD (referred to as PD0-GPD1) and a combined nonstationary PD and GPD (referred to as PD1-GPD1). The teleconnections between hydro-climatological variables and a number of large scale climate patterns allow using these climate indices as covariates in the development of nonstationary extreme value models. The case study is made with daily precipitation amount time series from Southeastern Canada and two climatic covariates, the Arctic Oscillation (AO) and the Pacific North American (PNA) indices. A comparison of PD0-GPD1 and PD1-GPD1 models showed that the incorporation of nonstationarity in both POT models instead of solely in the GPD has an effect on the estimated quantiles. The use of the B-spline function as link function between the GPD parameters and the considered climatic covariates provided flexible nonstationary PD-GPD models. Indeed, linear and nonlinear conditional quantiles are observed at various stations in the case study, opening an interesting perspective for further research on the physical mechanism behind these simple and complex interactions.

Keywords: Nonlinearity; Nonstationarity; Poisson-Generalized Pareto model; AO index; PNA index; B-splines; Precipitation; Canada.

1 Introduction

Extreme value theory (EVT) provides a solid justification to the use of probabilistic distributions such as the generalized extreme value (GEV) and the generalized Pareto distribution (GPD) for extreme event frequency analysis purposes (Fisher and Tippett, 1928; Jenkinson, 1955; Pickands, 1975; Coles, 2001). Both distributions are widely used in hydrology to fit respectively the annual maxima (AM) and the peaks-over- threshold (POT) values. With the extreme value sampling using the POT approach, two variables can be characterized, the exceedance intensity (i.e. the exceedance value) and the exceedance frequency (i.e. the yearly number of exceedances) (Lang et al., 1999).

The interpretation in terms of return period of the quantiles obtained from the GPD model requires information concerning the yearly number of exceedances which is generally assumed to be Poisson distributed. The Point Process (PP) theory allows to represent the Poisson distribution (PD) and the GPD as a two-dimensional nonhomogeneous Poisson process (Katz et al., 2002). Thus, a correspondence can be established between the combined distributions PD-GPD and the GEV distribution through their parameters (Lang et al., 1999 and Silva et al., 2016). The result is a reformulation of GEV parameters as functions of those of the PD-GPD, which allows associating an annual return period to the estimated quantile. Even if this process is indirect, it has the advantage of using the POT approach, from which, more than one extreme event can be sampled each year and both processes (i.e., intensity and frequency) of the extreme events can be captured.

The classical formulation of the EVT models assumes that the observations are independent and identically distributed (iid). However, with the mounting evidence concerning climate change (IPCC, 2012), there is a growing interest in the development of nonstationary statistical methods (Khaliq et al., 2006; Katz, 2013; Dörte, 2013) which can lead to more reliable estimates of various quantiles compared to the stationary approach. A quantile is a design value associated with a given probability of non-exceedance. It is often expressed in terms of the return period concept often used in engineering applications.

Nonstationary modeling commonly uses dependence between the parameters of a given probabilistic distribution and covariates. The dependence could be expressed in the form of polynomial functions (El Adlouni et al., 2007), smoothing splines (Chavez-Demoulin and Davison, 2005) or smoothing based on B-spline functions (Nasri et al., 2013). The B-splines are semi-parametric functions and lead to more flexible nonstationary models than the polynomial parametric functions (Padoan and Wand, 2008). Various B-spline functions were evaluated with the GPD model in the work of Thiombiano et al. (2016) and allowed modeling linear as well as nonlinear interactions.

Climate indices are widely used as covariates for the modeling of hydrological variables in a nonstationary framework. Indeed, significant relationships were observed worldwide between large-scale atmospheric/oceanic climate indices and hydro-climatological variables thereby helping to understand the variability in such variables. The definition of climate indices from Lee et al. (2013), "*time series that allow quantifying the temporal evolution of climate process in a particular region*", can justify their use as explanatory variables for various hydro-climatological applications.

For example, Kenyon and Hegerl (2010) highlighted the worldwide effects of the El-Nino-Southern Oscillation (ENSO) on extreme precipitation events, and the significant response of the latter to the North Atlantic Oscillation (NAO) and the Arctic Oscillation (AO) respectively over the European continent and Northern Hemispheric midlatitude. Lee and Ouarda (2010) studied the future evolution of the regional scaled winter precipitation and the extreme hydro-climatological variables in Eastern Canada, by modeling and projecting a nonstationary oscillation process. The significant oscillation signal of the NAO winter index can be extended from the Empirical Mode Decomposition (EMD) process and used as a covariate. Zhang et al. (2010) showed that winter daily precipitation maxima over North America are significantly influenced by the ENSO and the Pacific Decadal Oscillation (PDO) indices. Stone et al. (2000) found that the Pacific North American (PNA) climatic pattern has a stronger influence on the frequency of extreme daily precipitation amounts than their intensity in Ontario and Southern Quebec during autumn and winter seasons. Coulibaly (2006) related the change around the year 1940 in the Canadian seasonal precipitation to AO, and also observed strong correlation between the former and the PNA index after 1970. Bonsal and Shabbar (2011) synthesized the spatial and seasonal effects of ENSO, PDO, PNA, NAO, AO and the Atlantic Multi-decadal Oscillation (AMO) on the Canadian climate. Thiombiano et al. (2016) identified the AO and PNA indices as the two dominant climatic patterns that influence the intensity of extreme daily precipitation amounts over Southeastern Canada using a rank-based correlation analysis. They also found an East-West correlation sign shift between the AO index and the studied extreme precipitation events.

Evidences in the literature about climatic teleconnections are dominated by a based-prior assumption of linear interactions between these low frequency climate indices and hydro-climatological variables. However, nonlinear dependences are also observed because of the physical structure (i.e. negative and positive phases) of the climate indices. For example, using composite and correlation analyses, Shabbar et al. (1997) showed that winter precipitation in the upper St. Lawrence valley is enhanced during the La Nina phase while no significant response occurring during the El Nino years. Quadratic response of precipitation to ENSO and AO indices were found over the Northern Hemisphere by Wu et al. (2005) and Hsieh et al. (2006) based on artificial neural network analysis. By applying an Akaïke Information Criterion-based polynomial selection approach, Fleming and Dahlke (2014) detected parabolic downward and upward interactions between annual total flow volume time series and climate indices ENSO and AO over the Northern Hemisphere. Chandran et al. (2015) found through linear correlation and wavelet analyses, that the negative phase of the Southern Oscillation Index (SOI) is significantly associated with the increase in precipitation over the United Arab Emirates. Canon (2015) compared linear and nonlinear GEV models and showed that El Nino is associated with a decreased likelihood of extreme precipitation over the Great Lakes (Ohio River valley) and Western Canada (Alaska). Silva et al. (2016) found that high flood frequency and magnitude are not monotonically increasing or decreasing with the Niño3.4 index in the Itajaí River basin located in southern Brazil.

Nonlinear relationships remain complex but need to be explored and used because they can improve the estimates of hydroclimatic variables. Hence, the present study aims at investigating such interactions and developing a flexible nonstationary PD-GPD model for the statistical modeling of hydro-climatological variables of interest. The methodology is described in Section 2 and the case study is presented in Section 3. The discussion and concluding remarks follow in Section 4.

2 Methods

2.1 PD-GPD model

The extreme value sample with the POT approach is generally constituted of all events that exceed a suitable high threshold. From the data sample, the exceedances over the threshold have the GPD as an asymptotic limiting model (Pickands, 1975) with cumulative distribution function (cdf) given by

$$G(x, \xi, \sigma, u) = \begin{cases} 1 - \left[1 - \xi\left(\frac{x-u}{\sigma}\right)^{\frac{1}{\xi}}\right] & (x-u) \geq 0;\ \xi \neq 0;\ \sigma > 0 \\ 1 - \exp\left[-\left(\frac{x-u}{\sigma}\right)\right] & (x-u) \geq 0;\ \xi = 0;\ \sigma > 0 \end{cases} \qquad (1)$$

where x and $x - u$ are respectively the exceedance intensity (i.e. the peak value) and the magnitude of exceedance values; ξ, σ *and* u are respectively the GPD shape, scale and threshold parameters.

From the POT sample, the yearly number of exceedances over a given period (i.e. the frequency of exceedances) can also be fitted by a PD whose probability mass function is given by

$$P(m = n) = \exp(-\lambda)\frac{\lambda^n}{n!}\ ,\ \ n \in \mathbb{N} \qquad (2)$$

where m is a random variable representing the number of exceedances per year and λ is the PD rate parameter which is the expected annual frequency of the exceedances.

The Quantile Q_p in the POT framework for a given non-exceedance probability p can be obtained directly by inverting the GPD cdf from equation (1) as follows

$$Q_p(p,\xi,\sigma) = \begin{cases} \left[\frac{\sigma}{\xi}\left[1-(1-p)^{\xi}\right]\right] + u & \xi \neq 0 \\ \left[-\sigma \ln(1-p)\right] + u & \xi = 0 \end{cases} \quad (3)$$

However, to estimate the return level of exceedances (Q_T) associated with a T-years return period, the PD rate parameter needs to be incorporated as follows

$$Q_T(T,\lambda,\xi,\sigma) = \begin{cases} \left[\frac{\sigma}{\xi}\left[1-\left(\frac{1}{\lambda T}\right)^{\xi}\right]\right] + u & \xi \neq 0 \\ \left[-\sigma \ln\left(\frac{1}{\lambda T}\right)\right] + u & \xi = 0 \end{cases} \quad (4)$$

In the case of a Poisson process, there is a correspondence between the POT and AM distributions (Lang et al., 1999) formulated as $F(y) = exp\left[-\lambda\left(1-G(x)\right)\right]$ where y represents the AM values; F(y) and G(x) are respectively the GPD and GEV cdf; λ is the expected annual frequency of exceedances.

A straight estimate of Q_T is obtained by inverting the GEV cdf (equation 5), but with its parameters expressed as function of the PD and GPD parameters.

$$F(y,\xi,\beta,\psi) = \begin{cases} \exp\left[-\left(1-\frac{\xi}{\beta}(y-\psi)\right)^{\frac{1}{\xi}}\right] & \xi \neq 0;\ \beta > 0 \\ \exp\left[-\exp\left(-\frac{(y-\psi)}{\beta}\right)\right] & \xi = 0;\ \beta > 0 \end{cases} \quad (5)$$

where β and ψ represent respectively the GEV scale and location parameters and are reparameterized as given in equation (6) (Silva et al., 2016); the shape parameter ξ is the same as in equation (1).

$$\begin{cases}\beta = \sigma/\lambda^{\xi} \\ \psi = u + \beta\left(\frac{\lambda^{\xi}-1}{\xi}\right)\end{cases} \tag{6}$$

Note that all the GEV parameters are deduced from the fit of those of the POT model and hence based on the sample of exceedances. Quantiles from the inverse of the GEV cdf are AM (Q_A) and are expressed in equation (7) with a non-exceedance probability p.

$$Q_A(p, \xi, \beta, \psi) = \begin{cases}\psi + \left(\frac{\beta}{\xi}\right)\left[1 - \left(-\ln(p)\right)^{\xi}\right] & \xi \neq 0 \\ \psi - \beta\ln(-\ln(p)) & \xi = 0\end{cases} \tag{7}$$

2.2 Assumptions of the POT models

The use of the classical POT model implies that observations are iid (Lang et al., 1999). Hence, before the statistical inference, it is important to check the independence of exceedances. Moreover, the threshold is traditionally fixed so that exceedances over such a value are Poisson distributed. This assumption for the yearly number of exceedances needs to be verified because sometimes, other distributions like the Binomial or Negative Binomial can be more suitable (Bezak et al., 2014).

The PD can be verified with the test based on the dispersion index I proposed by Cunnane (1979). It is defined as follows (Önöz and Bayazit, 2001)

$$I = \frac{1}{N-1}\sum_{i=1}^{N}\frac{(m_i-\bar{m})^2}{\bar{m}} \tag{8}$$

where N is the number of years, in the case of annual time series; m_i is the number of exceedances in year i and $\overline{m}$ is the mean of $(m_i)_{i=1,\ldots,N}$. The test statistic t corresponding to this index is given by

$$t = \sum_{i=1}^{N} \frac{(m_i - \overline{m})^2}{\overline{m}} = (N-1)I \quad (9)$$

The statistic d asymptotically follows a χ^2 distribution with $(N-1)$ degrees of freedom. As the PD has a dispersion index I=1, the poissonian hypothesis is not rejected if the computed t value is in the range $(\chi^2{}_{\alpha/2}, \chi^2{}_{1-\alpha/2})$ where α is the significance level. If $t < \chi^2{}_{\alpha/2}$, the Binomial distribution must be preferred, and if $t > \chi^2{}_{1-\alpha/2}$, the Negative Binomial distribution is more appropriate. The case of I >1 corresponds to what is called the overdispersion phenomenon and is indicated to be normally more realistic. However, it is possible to take it into account with a simple non-homogeneous Poisson process (Eastoe and Tawn, 2010).

The independence criterion can be validated by assessing suitable thresholds (e.g., high percentiles) with the mean excess plot and the GPD shape and scale parameters stability given an increasing sequence of threshold (Davison and Smith, 1990; Lang et al., 1999; Coles, 2001). To systematically meet this criterion, a declustering technique (Roth et al., 2012) is widely used and can be validated through the partial autocorrelation function and the Chi-square goodness-of-fit test for the GPD adequacy.

2.3 Nonstationary POT models

The nonstationarity is commonly incorporated separately in each component of the PD-GPD model and can thus be easily integrated indirectly into the reparameterized GEV distribution (Katz, 2013).

2.3.1 GPD B-spline model

To take into account nonstationarity in equation (1), the GPD threshold and shape parameters were kept constant, and only the scale parameter was allowed to vary as a function of the covariate Z, referred to as σ_Z. To assure positive values of the scale parameter, the logarithm is usually applied to σ_Z. The link function h, defined by $\sigma_Z = h(Z)$, was assumed to be a linear combination of suitable basis spline (B-spline) functions $B_{k,d}$ having k knots and used to form a piecewise polynomial function of degree d (Nasri et al., 2013) as follows

$$\sigma_Z = h(Z) = \beta_0 + \sum_{j=1}^{k} \beta_j B_{j,d}(Z) \quad (10)$$

where β_0 *and* β_j are the regression parameters.

For example, the B-spline functions with (k, d) = (1, 1) and (k, d) = (1, 2) are special cases of polynomial linear and quadratic functions respectively. There is flexibility in the use of the B-spline functions as they allow exploring linear as well as nonlinear linkages between the variables. Moreover, such spline smooth functions are robust for extreme value modeling as the impact of the outliers and non-local effects are limited (Chavez-Demoulin and Davison, 2012). Some applications with the cubic splines and B-spline functions for a

POT framework were provided respectively in Chavez-Demoulin and Davison (2005) and Thiombiano et al. (2016). The use of B-spline functions with the GEV distribution was also proposed in Padoan and Wand (2008) and Nasri et al. (2013). The nonstationary GPD is then obtained by replacing the stationary scale parameter in equation (1) by its expression from equation (10).

The Generalized Maximum Likelihood (GML) method is used to estimate the parameters of the proposed nonstationary GPD model because it can improve the estimation of the GEV or GPD shape parameter (El Adlouni and Ouarda, 2008; El Adlouni et al., 2007). It is a Bayesian estimator where a prior distribution, the Beta distribution (u=6; v=9) defined on the interval [-0.5; +0.5], is specified for the GPD shape parameter to avoid unfeasible estimates (Martins and Stedinger, 2001). The GML estimator of this parameter is the mode (or mean) of its empirical posterior distribution which is obtained using Markov Chain Monte Carlo (MCMC) computation methods.

However, since the prior information is brought only for the GPD shape parameter and is the same for all GPD models, the ratio of the posterior distributions is equivalent to the likelihood ratio. Equivalence can thus be established between the Bayes factors and the AIC or Bayesian Information Criterion (Kass and Wasserman 1995; Schwarz 1978) as described in Appendix 1. Therefore, instead of using a MCMC algorithm, we proceeded by constrained minimization of the negative log-likelihood of the reparameterized GEV or GPD probability density function, depending on the assessed model. To do so, a Newton-Raphson algorithm (Hosking and Wallis, 1987) is used and the constraint is applied on the shape parameter.

The selection of the best model among various candidates can be obtained through the deviance statistic or more generally the AIC, which penalizes the minimized negative log-likelihood (nll) depending on the number of parameters (r) of each assessed model (Katz, 2013). The AIC computation for each model is as follows

$$AIC\ (r) = 2nll\ (r) + 2r \quad (11)$$

The number of parameters in the case of a GPD B-spline model depends on the B-spline function's parameter (d and k) and is given by $[(d + k) + 1]$. The model with the lowest AIC can be considered as the best model.

An automatic selection tool based on the AIC value was thus implemented to select the best GPD-B-spline model (Thiombiano et al., 2016).

2.3.2 Non-homogeneous Poisson process

One of the main objectives of the present study is to incorporate the effect of nonstationarity on the distribution of the yearly number of exceedances. To this end, the PD parameter λ is also allowed to vary as a function of the covariate Z. The covariate Z can be the same or different from the one used for the GPD scale parameter. The resulting non-homogeneous Poisson process with rate parameter λ_Z is herein investigated and given by

$$\log\lambda_Z = \lambda_0 + \lambda_1 Z \quad (12)$$

The hyperparameters λ_0 and λ_1 can be estimated through a generalized linear model (GLM) for the PD with the logarithmic link function.

2.3.3 Nonstationary PD-GPD model

The set of stationary parameters (σ and λ) should now be replaced by (σ_Z and λ_Z) as formulated in equations (10) and (12), allowing subsequently to take into account the effect of the covariate in the reparameterized parameters of the GEV distribution expressed in equation (6). Two nonstationary PD-GPD models are evaluated in this paper: a stationary PD combined to a nonstationary GPD (referred to as PD0-GPD1) and a combined nonstationary PD and GPD (referred to as PD1-GPD1). These models can also be compared using the AIC.

2.3.4 Quantile estimation and uncertainty measure

The quantiles are estimated under the GPD model (i.e., Q_p from (3)) and the PD-GPD model (i.e., Q_A from (7)) with the non-exceedance probability p in the stationary (without covariate Z) and nonstationary (with covariate Z) frameworks. The following models are thus used to obtain the quantiles:

- Stationary quantile Q_p from the model named GPD0 expressed by (1).
- Stationary quantile Q_A from the model named PD0-GPD0 formulated by combining (5) and (6).
- Nonstationary quantile Q_p from model named GPD1 formulated by combining (1) and (10).
- Nonstationary quantile Q_A from model named PD0-GPD1 formulated by combining (5), (6) and (10).

- Nonstationary quantile Q_A from model named PD1-GPD1 formulated by combining (5), (6), (10) and (12).

The quantiles from models GPD0 and GPD1 are presented for comparison purpose to those estimated by the PD-GPD approach.

The measure of the uncertainty associated with estimated quantiles is provided as intervals of credibility (ICs) computed for the 95% probability of confidence. An adjusted asymptotic approach is used. This new approach was proposed by Ashkar and El Adlouni (2015) who showed that it allows improving the normality of GPD-based quantiles and leads to more accurate ICs for quantiles in the right-tail of the GPD than the ICs obtained from the traditional large-sample based theory (Ashkar and Ouarda, 1996).

2.4 Rank-based correlation and wavelet analyses

Linear correlation analysis is widely used to statistically test the relationships between hydrological variables and large scale climate patterns. For non-normal distributions like those of the hydro-climatological extremes, this approach can be less flexible (Yue et al., 2002; Chen et al., 2012). In this paper, the rank-based correlation using the Kendall's tau with 5% significance level is used to select the covariate Z for the development of the PD-GPD nonstationary model. The dependence between hydrological variables and potential covariates is measured for each station of a case study. Stations where significant correlation values are detected are of interest and the covariate with the highest number of stations associated with significant Kendall's tau is selected.

A complementary method, wavelet analysis, is considered given the fact that both linear and rank-based correlation analyses identify monotonic dependence. Wavelet analysis allows capturing the time-scale changes in and between time series through the continuous wavelet transform (CWT), the cross-wavelet transform (XWT) and the wavelet transform coherence (WTC) plots (Torrence and Compo, 1998; Grinsted et al., 2004).

3 Case study

3.1 Data

3.1.1 Precipitation

The hydro-climatological variable analyzed is the observed daily precipitation amounts (code 012) obtained from Environment Canada (http://climat.meteo.gc.ca/historical_data/search_historic_data_f.html). Based on the quality of time series and on the validation of a suitable threshold for the POT sample definition, 173 stations located in Southeastern Canada were considered. Each time series must have at least 30 full years of observations (further referred to as complete years), and the threshold allowing to have the maximum number of stations where independence between exceedances is validated, is chosen. This independence criterion was evaluated using the methods presented in Section 2.2. The 99th percentile was found to be the suitable threshold to define the sample of exceedances at these 173 stations. Figure 1 shows their geographical location in this region. The stations where the PD is statistically validated are distinguished based on the index dispersion assessment. The study region is composed by the provinces of Newfoundland (NF), Labrador (NFL), Prince Edward Island (PEI), New Brunswick (NB), Nova Scotia (NS), Quebec (QC) and Ontario (ON) in Canada.

[Figure 1]

From each sample of daily precipitation amount exceedances, the intensity and frequency variables (i.e., V1 and V2 respectively) were defined for each station. The V1 and V2 time series were further used for correlation and wavelet analysis with studied climate patterns. These climatic covariates describe large scale atmospheric and oceanic oscillations, while hydrological variables characterize local to regional observations (Rossi et al., 2009). Therefore, a time series smoothing is often employed to maximize the correlation results between variables like precipitation which shows an important variability, and climatic covariates (Assani et al., 2008). An annual average of V1 values was calculated for this purpose.

3.1.2 Climate indices

Thiombiano et al. (2016) showed that the AO and PNA indices are the two dominant climate patterns influencing the intensity of extreme daily precipitation amounts in Southeastern Canada. The former index is the first empirical orthogonal function of the Northern Hemisphere (20°-90°N) winter sea level pressure data, while the latter index is the dominant climatic pattern of low-frequency variability in the Northern Hemisphere extratropics (Rossi et al., 2011). In this paper, the same seven indices used in their study (i.e. AMO, AO, NAO, PDO, PNA, SOI and the Western Hemisphere Warm Pool (WHWP)) were considered to assess the interaction of each of these indices with the V2 time series. The monthly standardized time series of these indices were obtained from the Physical Sciences Division of the National Oceanic and Atmospheric Administration (NOAA) (http://www.esrl.noaa.gov/psd/data/climateindices/list/). With regard to the smoothing applied to the V1 series for correlation and wavelet analyses purposes, the indices time series are averaged from January to December (JD) on one hand, and on the other hand, a three-month moving average (i.e. December through February (DJF), January through March (JFM), etc.), are considered.

3.2 Results

3.2.1 Correlation analysis

The correlation analysis between the studied climate patterns and the V1 series on one hand and the V2 series on the other hand, showed that significant correlation between the V1 (V2) series and at least one of the 7 studied climate indices, were found at 136 (138) stations of the 173 stations originally analyzed. The AO and PNA indices showed the most significant influence on the variability of both variables (i.e. V1 and V2) than the other indices. Figures 2 and 3 summarize the results of the correlation analysis.

[Figure 2]

The identification of the AO and PNA indices as dominant covariates is based on the counting by province of the study region, of the number of stations where significant correlation was found between each of them or both indices and the V1 and/or V2 series.

[Figure 3]

The same approach was adopted to evaluate the influence of other studied indices in the study region in order to provide an overview of the co-influence of all studied indices on Southeastern Canada's extreme precipitations (Table I).

[Table I]

The highest level of the influence corresponds to the index with the largest number of correlated stations, while the absence of any correlation with stations is identified by "No" influence.

3.2.2 Nonstationary modeling

Among the stations where the V1 and V2 series showed significant correlations with the indices AO and/or PNA, 10 stations (15 stations) were considered for the combined nonstationary modeling of V1 and V2 using the same AO index (PNA index). The aim was to investigate the effect of AO (PNA) pattern on both the intensity and frequency of extreme precipitation in the study region.

For the GPD1 model, the GPD scale parameter was allowed to vary conditionally to the AO index or the PNA index given the station, using B-spline functions with (k, d) = (2, 1). This choice was based on the work of Thiombiano et al. (2016). For the PD0 model, the expected annual frequency of exceedances varied between 3 and 4 events per year. For the PD1 model, the GLM estimates of the associated rate parameter are used.

Three non-exceedance probabilities (p=0.5, 0.9 and 0.99) were used for quantile estimates. The fitting of the GPD1, PD0-GPD1, PD1-GPD1 models to the 25 stations (i.e.

10 stations and 15 stations associated with the AO and PNA indices respectively), showed four types of conditional quantiles (Figure 4).

[Figure 4]

The concave forms of conditional quantiles are associated with the AO index and are observed in the NS and NF provinces. Only one similar form is found in a more central province (Ontario). Concerning the convex forms of conditional quantiles, they are associated with the PNA index and are rather detected in central provinces (Quebec and Ontario). At these nonlinear responses of extreme precipitations to the AO or PNA index, common linear responses are obviously observed (i.e. monotonically downward and upward dependences). To illustrate the nonlinear responses, some analyses results are proposed with two stations highlighted in red in Figure 4.

3.2.3 Illustrations

The V1 and V2 time series from the Upper Stewiacke (ID 8206200) and "*Grandes Bergeronnes*" (ID 7042840) stations located respectively in NS and QC provinces are used herein. The analysis periods for these stations are 1951-2005 and 1951-2012 respectively, leading to 53 years (with 51 complete years) and 60 years (with 57 complete years) of daily observations. At the Upper Stewiacke (*Grandes Bergeronnes*) station, the covariate is the AO (PNA) index. The threshold values corresponding to the 99th percentile of the daily precipitation amounts dataset are 37 mm and 32 mm respectively at stations Upper Stewiacke and *Grandes Bergeronnes*, leading to a sample of one day declustered exceedances containing 210 and 207 independent events respectively. The average intensity value of exceedances is 49 mm and 46 mm respectively over the respective analyzed periods, while the expected annual frequency of exceedances is between 3 and 4. The dispersion index is 1.06 at the Upper Stewiacke station and 1.26 at the *Grandes Bergeronnes* station. The statistic t associated to these values validated the PD assumption, allowing the use of the PD-GPD model.

The Kendall tau measured with 5% significance level between 13 time series of the AO (PNA) index and variables V1 and V2 are presented in Table II for both stations with significant correlations highlighted in bold.

[Table II]

The average of the AO values from the months of September to November (SON) and October to December (OND) constituted the covariate AO data respectively for V1 and V2 variables for nonstationary modeling at the Upper Stewiacke station. In the case of the *Grandes Bergeronnes* station, DJF and JD time windows were retained for V1 and V2 respectively.

The assessment of the CWTs of the V1, V2, AO and PNA analyzed time series, showed the presence of significant features of variability predominately in the range of periods spanning 2-8 years. This detection of variability is more physically meaningful than the simple correlation analysis, thus sustaining the nonstationary frequency analysis framework. Common and coherent significant features were also clearly highlighted by the XWT and WTC plots between the explored variable-covariate datasets. These wavelet results can comfort or not the correlation results by comparing the direction of the interaction arrows as pointing right (or left) means a positive (negative) correlation. The wavelet analysis results for V1-AO (Upper Stewiacke station) and V2-PNA (*Grandes Bergeronnes* station) interactions are proposed in Figures 5 and 6.

[Figure 5]

In Figures 5 and 6, the thick black contour designates the 5% significance level against the red noise and, the cone of influence where the edge effects might distort the picture is shown in a lighter shade. The darker the red colour is in the enclose feature, the stronger the variability is. The phase relationship between the time series (see XWTs and WTCs plots) is represented as arrows with in-phase pointing right (positive correlation) and anti-phase pointing left (negative correlation).

Significant features are found from the Upper Stewiacke station V1 time series in the 2-3 year and the 5-8 year periods respectively around the decades 1960-1970 and 1980-1990 (Figure 5a). In its corresponding covariate AO data, a 3-5 year feature can also be observed around the year 1980 (Figure 5b). Common significant features effectively appear from 1960 to 1990 in the 2-7 year period between these two correlated datasets (Figure 5c), with a strong covariance between them in the 2-5 year period during the decade 1970-1980 (Figure 5d). Moreover, the XWT and WTC results highlight that these datasets are in-phase, confirming the significant positive correlation between them (Table II).

[Figure 6]

At the *Grandes Bergeronnes* station, significant features are also present in the different time series (Figure 6) with, however, less strong and synchronized enclosed features. Nevertheless, arrows in the XWT (Figure 6c) and WTC (Figure 6d) plots show that these datasets are also phase-locked, sustaining the positive significant correlation observed between them (Table II).

The choice of the best GPD1 model for Upper Stewiacke and *Grandes Bergeronnes* stations is highlighted in bold in Table III where comparative values of the AIC for nine GPD1 models are indicated given the B-spline function parameter (k, d) evaluated.

[Table III]

The combination (k, d) resulting in the smallest AIC value is highlighted in bold and is considered as the best GPD1 model to be used. At both stations, a B-spline with 2 knots and 1 degree (2, 1) is the best combination with AIC values of 360.22 and 433.76 respectively. These values are smaller than the AIC associated with the GPD0 model (364.22 and 434.11 respectively).

Figures 7 and 8 illustrate the quantiles associated with a non-exceedance probability p=0.9 at the Upper Stewiacke and *Grandes Bergeronnes* stations respectively. These quantiles are estimated using the models defined in Section 2.3.4 (GPD0, GPD1, PD0-GPD0, PD0-GPD1 and PD1-GPD1). The purpose of adding the AM observations on these Figures is to understand the difference between estimated quantiles resulting from a simple GPD versus a combined PD-GPD model. For the POT observations, they are independent exceedances of daily precipitation amounts. The same monthly value of the climatic covariate is thus used for exceedances that occurred in that same month, hence, the alignment of some observations.

The quantiles estimated using GPD0 and GPD1 are systematically inferior to the T-year return quantiles obtained from models PD0-GPD0, PD0-GPD1 and PD1-GPD1. The comparison between PD0-GPD0 and PD0-GPD1 then PD1-GPD1, shows a difference in values at around 40 mm and 70 mm respectively. In the nonstationary case, the estimates depend on the value of the covariate and show a clear nonlinear association between the climate index and precipitation extremes. In both concave (Figure 7) and convex (Figure 8) nonlinear structures, the AM quantiles obtained from the PD0-GPD1 model are superior (inferior) to those from the PD1-GPD1 model during the negative (positive) phase of the index. These results suggest that PD0-GPD1-based quantiles are not systematically above or below the PD1-GPD1-based quantiles for all covariate values.

[Figure 7]

At the Upper Stewiacke station (Figure 7), the quantile estimates increase with the increase in the absolute value of the AO index for both positive and negative index values. The conditional quantile curve has a concave form, highlighting clearly the nonlinear response of precipitation extreme events to AO index at this station. With regard to the comparison of models PD0-GPD0, PD0-GPD1 and PD1-GPD1 through AIC values, it is observed that the second model outperforms the third model which is better than the first model. This result suggests that incorporation of nonstationarity only in the estimation of the GPD scale parameter provided a better model.

[Figure 8]

At the *Grandes Bergeronnes* station (Figure 8), the nonlinear dependence between precipitation extremes and the PNA index takes a convex form in comparison to the concave form observed in Figure 7. Hence, the estimated quantiles decrease with the increase in the absolute value of the PNA index for both positive and negative values. However, the model PD1-GPD1 leads to the lowest AIC value. This model outperforms PD0-GPD1 which was found to be better than the PD0-GPD0 model. Thus, incorporation of nonstationarity both in the GPD scale and PD intensity parameters provides a better fit at this station.

4 Discussion and conclusion

In the present study, the PD-GPD model was suggested to estimate quantiles by statistically testing the nonstationarity hypothesis and modeling simple and complex interactions between large scale climate patterns and hydro-climatological variables. For this purpose, a nonstationary GPD model where the scale parameter is allowed to vary as a B-spline function of a climatic covariate is combined to a nonhomogeneous PD. In this latter process, the rate parameter is a logarithm function of the same covariate.

The use of the B-spline function instead of the classical polynomial function, allowed to automatically analyze various nonstationary GPD and PD-GPD models. This flexibility of the B-spline functions has attracted progressive interest in hydrology over the last 10 years (Padoan and Wand 2008; Nasri et al. 2013; Thiombiano et al. 2016). The use of such type of linkage must then be promoted for nonstationary statistical modeling in hydrology because of the physical structure of climate indices which are widely used as covariates, given the well-established interactions between large-scale atmospheric and oceanic variability and hydro-climatological variables. However, in addition to the statistical significance found in many studies as well as in the present paper, there are underlying physical mechanisms related to the flux dynamics of the air masses.

For example, Trenberth (1990) indicated that the shift in atmospheric circulation constitutes the principal cause of regional variability in observed wind, temperature, precipitation and other climatic variables. Thompson and Wallace (2001) also explained the association between the sea level pressure variability, wind direction and the warm-cool phases of AO, NAO and PNA. Wu et al. (2005) mentioned that anomalous northerlies from the Arctic area transferred colder air over Northeastern Canada, leading to negative temperature anomalies there, while, the anomalous alongshore flow along the West coast of North America brings the normal moist westerly flow farther North, generating negative precipitation anomalies from Oregon to Southern British Columbia. The identification of AO and PNA indices as potential dominant modes of daily precipitation amounts extremes variability in Southeastern Canada, must be sustained by similar physical explanations in future research. This additional exercise will result in the suggestion of these climate indices as covariates for precipitation quantiles prediction over (or in some part of) the Southeastern Canada.

Indeed, from the case study, the correlation results showed that AO and PNA indices have the highest influence on both the intensity and frequency of extreme precipitation time series. Moreover, the modeling results obtained from the 10 (15) stations where the AO (PNA) index showed significant correlations with both the V1 and V2 series, provided some insight about the risk of underestimation or overestimation of quantiles when assuming stationary distributions or linear dependences in the case of non-stationarity. It is thus important to go beyond the correlation results and understand the physical mechanism behind these teleconnections. Mainly, the concave and convex relationships found in this study must be confirmed with larger time series by using for example adjusted and homogenized hydro-climatological datasets. The correlation and wavelet analysis results can also be helpful to study the combined effect of more than one climatic covariate on modeling results.

Acknowledgments: The Authors are grateful to the International Center for Research and Development (ICRD) and to the Natural Sciences and Engineering Research Council of Canada (NSERC) for the financial support through the FACE (*Faire Face Aux Changements Ensemble*) project grant. The authors also wish to thank the editor Dr. Sergey Gulev as well as the two anonymous reviewers for the comments which helped improve considerably the quality of the paper.

References

Ashkar F, El Adlouni S. 2015. Adjusting for small-sample non-normality of design event estimators under a generalized Pareto distribution. *Journal of Hydrology* **530**:384-391.

Ashkar F, Ouarda TBMJ. 1996. On some methods of fitting the generalized Pareto distribution. *Journal of Hydrology* **177**: 117-141.

Assani AA, Lajoie F, Vadnais M-E, Beauchamp G. 2008. Analyse de l'influence de l'oscillation Arctique sur la variabilité interannuelle des précipitations dans le bassin versant de la rivière Saint-François (Québec, Canada) au moyen de la méthode des corrélations canoniques. *Revue des sciences de l'eau* **21**(1):19-33.

Bezak N, Brilly M, Sraj M. 2014. Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal* **59**:959-977.

Bonsal B, Shabbar A. 2011. Large-scale climate oscillations influencing Canada, 1900-2008. Canadian Biodiversity : Ecosystem Status and Trends 2010. *Technical Thematic Report No.4*. Canadian Councils of Resource Ministers. Ottawa, ON.

Cannon AJ. 2015. Revisiting the nonlinear relationship between ENSO and winter extreme station precipitation in North America. *International Journal of Climatology*. doi: 10.1002/joc.4263

Chandran A, Basha G, Ouarda T. 2015. Influence of climate oscillations on temperature and precipitation over the United Arab Emirates. *International Journal of Climatology*. doi:10.1002/joc.4339

Chavez-Demoulin V, Davison AC. 2012. Modeling time series extremes. *REVSTAT-Statistical Journal* **10**:109-133.

Chavez-Demoulin V, Davison AC. 2005. Generalized additive modelling of sample extremes. *Appl. Statist.* **54**:207-222.

Chen L, Singh VP, Guo S. 2012. Measure of correlation between River flows using the copula-entropy method. *J. Hydrol. Eng*. **18**:1591-1606.

Coles SG. 2001. An introduction to statistical modeling of extreme values. *Springer Series in Statistics*.

Coulibaly P. 2006. Spatial and temporal variability of Canadian season precipitation (1900-2000). *Advances in Water Resources* **29**:1846-1865. doi: 10.1016/j.advwatres.2005.12.013

Cunnane C. 1979. A note on the Poisson assumption in partial duration series models. *Water Resour Res*. **15**:489-493.

Davison AC, Smith RL. 1990. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society*. Series B (Methodological) **52**:393-442.

Dörte J. 2013. Nonstationarity in extremes and engineering design. In : AghaKouchak A et al. (eds.) Extremes in changing climate : detection, analysis and uncertainty. *Springer*, pp 363-417.

Eastoe EF, Tawn JA. 2010. Statistical models for overdispersion in the frequency of peaks over threshold data for a flow series. *Water Resources Research* **46**:1-12.

El Adlouni S, Ouarda TBMJ. 2008. Comparaison des méthodes d'estimation des paramètres du modèle GEV non stationnaire. *Revue des Sciences de l'Eau* **21**(1):35-50.

El Adlouni S, Ouarda TBMJ, Zhang X, Roy R, Bobée B. 2007. Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resources Research* **43**:1-13.

Fleming SW, Dahlke HE. 2014. Parabolic northern-hemisphere river flow teleconnections to El Nino-Southern Oscillation and the Arctic Oscillation. *Environmental Research Letters* **9**,104007. doi: 10.1088/1748-9326/9/10/104007

Grinsted A, Moore JC, Jevrejeva S. 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics* **11**:561-566.

Hosking JRM, Wallis TR. 1987. Parameter and quantile estimation for the Generalized Pareto distribution. *Technometrics* **29**:339-349.

Hsieh WW, Wu A, Shabbar A. 2006. Nonlinear atmospheric teleconnection. *Geophysical Research Letters* **33**, L07714. doi: 10.1029/2005GL025471

IPCC. 2012. Managing the risk of extreme events and disasters to advance climate change adaptation. Special Report of the Intergovernmental Panel on Climate Change (IPCC). *A special report of working groups I, II of the IPCC…. Cambridge University Press*, Cambridge, UK, and New York, NY, USA.

Jenkinson AF. 1955. The frequency distribution of the annual maximum (or minimum) of meteorological elements. *Quart J Roy Meteor Soc* **81**:158-171.

Kass RE, Wasserman L. 1995. A Reference Bayesian Test for Nested Hypotheses and Its Relationship to the Schwarz Criterion. *Journal of the American Statistical Association* **90**:928-934.

Katz RW. 2013. Statistical methods for nonstationary extremes. In : AghaKouchak A et al. (eds.) Extremes in changing climate : detection, analysis and uncertainty. *Springer*, pp 15-37.

Katz RW, Parlange MB, Naveau P. 2002. Statistics of extremes in hydrology. *Advances in Water Resources* **25**:1287-1304.

Kenyon J, Hegerl GC. 2010. Influence of modes of climate variability on global precipitation extremes. *Journal of Climate* **23**:6248-6262.

Khaliq MN, Ouarda TBMJ, Ondo J-C, Gachon P, Bobée B. 2006. Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observation: a review. *Journal of Hydrology* **329**:534-552.

Lang M, Ouarda TBMJ, Bobée B. 1999. Towards operational guidelines for over-threshold modeling. *Journal of Hydrology* **225**:103-117.

Lee T, Ouarda TBMJ, Li J. 2013. An orchestrated climate song from the Pacific and Atlantic oceans and its implication on climatological processes. *International Journal of Climatology* **33**:1015-1020.

Lee T-S, Ouarda TBMJ. 2010. Long-term prediction of precipitation and hydrologic extremes with nonstationary oscillation processes. *Journal of Geophysical Research* **115**, D13107. doi:10.1029/2009JD012801

Martins ES, Stedinger JR. 2001. Generalized maximum likelihood Pareto-Poisson estimators for partial duration series. *Water Resources Research* **37**:2551-2557.

Nasri B, El Adlouni S, Ouarda TBMJ. 2013. Bayesian estimation for GEV-B-Spline model. *Open Journal of Statistics* **3**:118-128.

Önöz B, Bayazit M. 2001. Effect of the occurrence process of the peaks over threshold on the flood estimated. *Journal of Hydrology* **244**:86-96.

Padoan SA, Wand MP. 2008. Mixed model-based additive models for sample extremes. *Stat Probab Lett* **78**:2850-2858.

Pickands J. 1975. Statistical inference using extreme order statistics. *Ann Stat* **3**:119-131.

Rossi A. Massei N. Laignel B. 2011. A synthesis of the time-scale variability of commonly used climate indices using continuous wavelet transform. *Global and Planetary Change* **78**:1-13.

Rossi A, Massei N, Laignel B, Sebag D, Copard Y. 2009. The response of the Mississippi River to climate fluctuations and reservoir construction as indicated by wavelet analysis of streamflow and suspended-sediment load, 1950-1975. *Journal of Hydrology* **377**:237-244.

Roth M, Buishand TA, Jongbloed G, Klein TAMG, van Zanten JH. 2012. A regional peaks-over-threshold model in a nonstationary climate. *Water resources research* **48**, W11533. doi: 10.1029/2012WR012214

Schwarz GE. 1978. Estimating the dimension of a model. *Annals of Statistics* **6**: 461–464. doi:10.1214/aos/1176344136

Shabbar A, Bonsal B, Khandekar M. 1997. Canadian precipitation patterns associated with the Southern oscillation. *Journal of Climate* **10**:3016-3027.

Silva AT, Naghettini M, Portela MM. 2016. On some aspects of peaks-over-threshold modeling of floods under nonstationarity using climate covariates. *Stoch Environ Res Risk Assess* **30**:207-224. doi: 10.1007/s00477-015-1072-y

Stone DA, Andrew JW, Zwiers FW. 2000. Trends in Canadian precipitation intensity. *Atmosphere-Ocean* **38**:321-347.

Thiombiano AN, El Adlouni S, St-Hilaire A, Ouarda TBMJ, El-Jabi N. 2016. Nonstationary frequency analysis of extreme daily precipitation amounts in Southeastern Canada using a peaks-over-threshold approach. *Theoretical and Applied Climatology* **124**(1-2). doi: 10.1007/s00704-016-1789-7

Thompson DWJ, Wallace JM. 2001. Regional climate impacts of the Northern hemisphere annular mode. *Science* **293**:85-89.

Torrence C, Compo GP. 1998. A pratical guide to wavelet analysis. *Bulletin of the American Meteorological Society* **79**:61-78.

Trenberth KE. 1990. Recent observed interdecadal climate changes in the Northern Hemisphere. *Bulletin of the American Meteorological Society* **71**:988-993

Zhang X, Wang J, Zwiers FW, Ya Groisman P. 2010. The influence of large-scale climate variability on winter maximum daily precipitation over North America. *Journal of Climate* **23**:2902-2915.

Wang C, Enfield DB. 2001. The tropical Western Hemisphere warm pool. *Geophysical Research Letters* **28**:1635-1638.

Wu A, Hsieh WW. 2005. The nonlinear patterns of North American winter temperature and precipitation associated with ENSO. *Journal of Climate* **18**:1736-1752.

Yue S, Pilon P, Cavadias G. 2002. Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of Hydrology* **259**:254-271.

Illustration of Tables

Table I. Level of influence of seven climate indices on daily precipitation amount extremes by province in Southeastern Canada

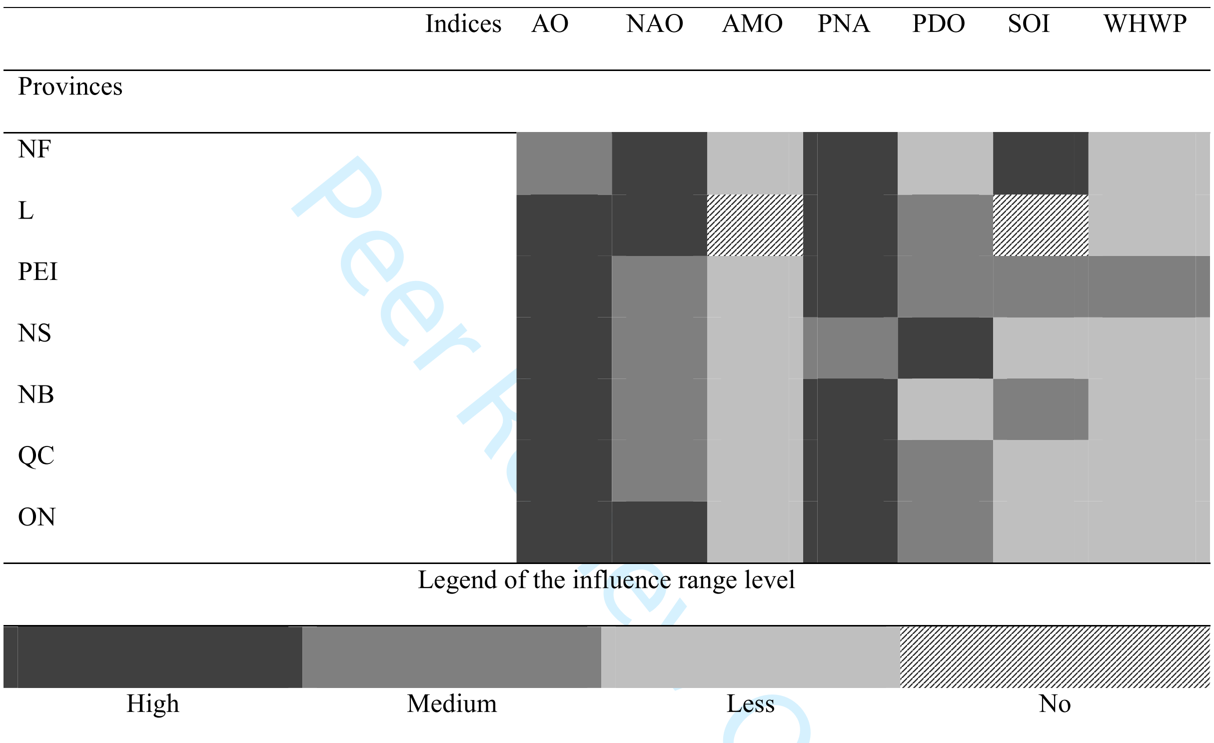


Table II. Kendall's Tau (in %) with significant correlation in bold

	Climate index windows												
	JD	DJF	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ
	Upper Stewiacke station												
AO&V1	**+20**	-3	+2	+1	+10	+17	+15	+12	+15	+18	**+25**	+16	+7
AO&V2	+14	+7	+6	+2	-7	-9	-4	+6	+13	+5	+15	**+23**	**+22**
	Grandes Bergeronnes station												
PNA&V1	+8	**+18**	+15	+5	-3	-4	+9	+3	+5	-1	+2	+1	+11
PNA&V2	**+26**	+6	+3	+14	**+20**	**+24**	+14	+12	+11	+16	**+20**	**+22**	+18

Table III. AIC values for different GPD B-spline models with the lowest AIC in bold

Knots k	2	3	4
Degree d	Upper Stewiacke station		
1	**360,22**	362,50	362,70
2	362,83	362,52	363,73
3	363,17	363,78	365,92
Degree d	*Grandes Bergeronnes* station		
1	**433,76**	436,33	437,78
2	436,56	437,65	439,58
3	437,10	439,51	440,65

Title of Figures

Figure 1. Southeastern Canada (a) and location of the 173 stations (b) with identification of stations where the PD, or BD or BND was validated.

Figure 2. Identification of stations where the AO and PNA indices were significantly correlated to the V1 series among the 136 stations where V1 series showed significant correlation with at least one of the seven studied climate indices.

Figure 3. Identification of stations where the AO and PNA indices were significantly correlated to the V2 series among the 138 stations where V2 series showed significant correlation with at least one of the seven studied climate indices.

Figure 4. Linear and nonlinear response of extreme precipitations to AO and PNA indices at 25 stations in Southeastern Canada. Highlighted stations in red are selected for illustrative purpose.

Figure 5. Wavelet analysis results with CWT illustration for V1 (a) and AO index: SON window (b). XWT (c) shows the common features between (a) & (b). WTC (d) highlights the covariance between (a) & (b).

Figure 6. Wavelet analysis results with CWT illustration for V2 (a) and PNA index: JD window (b). XWT (c) shows the common features between (a) & (b). WTC (d) highlights the covariance between (a) & (b).

Figure 7. Estimated quantiles from stationary (GPD0, PD0-GPD0) and nonstationary (GPD1, PD0-GPD1, PD1-GPD1) models, using the Arctic Oscillation as a covariate, associated with a non-exceedance probability of 0.9. Lower (BIC) and upper (BSC) 5% confidence intervals are also shown. Illustrated observations are independent daily peaks (POT) and annual maxima (AM) precipitation values for the Upper Stewiacke station.

Figure 8. Same description as in Figure 7, but using the Pacific North American index as a covariate for the *Grande Bergeronnes* station.

Appendix 1. Bayes factors and Laplace approximation

The selection of the most appropriate model is usually based on the Bayes factors when the inference is done in a Bayesian framework. Let *M* be a candidate model; the posterior distribution associated to model *M* is obtained through conditional probability formula and is given by

$$P(D \mid M) = \int_{\Omega} L(D \mid \theta, M)\, \pi(\theta \mid M) d\theta$$

Where $P(D|\theta, M)$ represents the likelihood of data D associated to model M and its vector of parameters θ; Ω is the dimension of parameters and $\pi(\theta|M)$ the prior distribution of the parameter.

To compare two models M1 and M2, the Bayes factors ratio B_{12} is

$$B_{12} = \frac{P(D \mid M_1)}{P(D \mid M_2)} . \frac{P(M_1)}{P(M_2)}$$

Where $P(M_i)$ is the prior probability associated to model M_i , $i = 1, 2$.

In absence of any prior discrimination between models M1 and M2, then $P(M_1) = P(M_2)$ and the Bayes factor is equivalent to the posterior distributions ratio.

The resulting integrals are often complex to assess, leading often to some approximations based on Laplace development (Kass and Wasserman, 1995).

When the prior distributions $\pi(\theta|M_1)$ and $\pi(\theta|M_2)$ of models M1 and M2 parameters are similar, the Laplace approximation is

$$B_{12} \approx \frac{L(D \mid \widehat{\theta_1}, M_1)}{L(D \mid \widehat{\theta_2}, M_2)} \cdot \left| \frac{H_1^{-1}\,(\widehat{\theta_1})}{H_2^{-1}\,(\widehat{\theta_2})} \right|^{1/2} \left(\frac{n}{2\pi}\right)^{(n2-n1)/2}$$

With $H_1^{-1}\,(\widehat{\theta_1})$ representing the Hessian matrix of the prior distribution; n the sample size; n_i the parameter dimension of model M_i , $i = 1, 2$.

In this study, the same prior distribution for the shape parameter is used for all models. Therefore, the Hessian matrix ratio equals 1. Consequently, the Bayes factors ratio is equivalent to the Schwartz criterion (Schwartz, 1978).

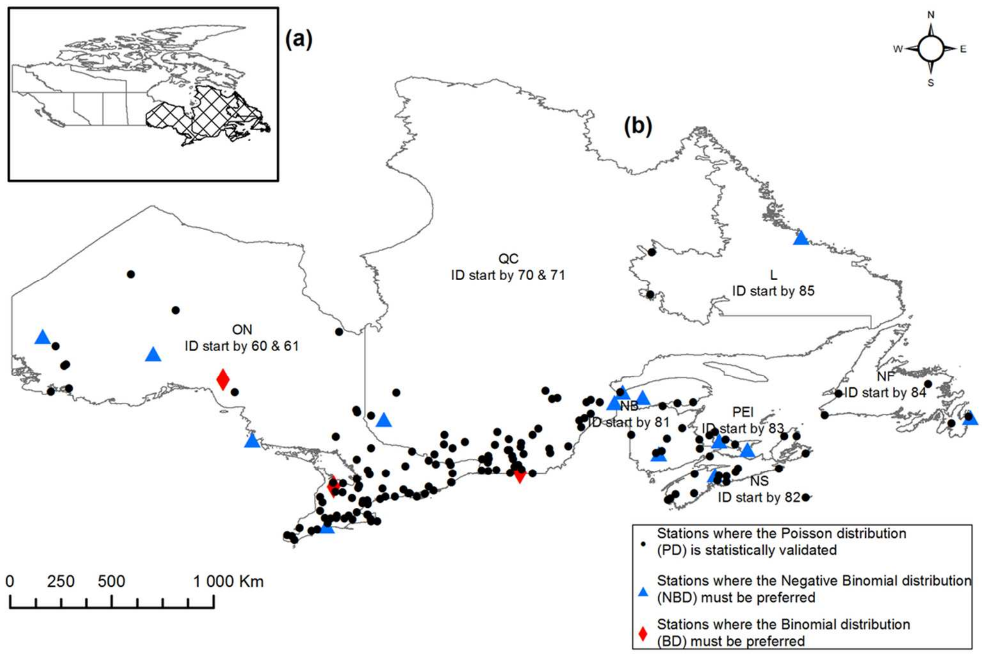


Figure 1. Southeastern Canada (a) and location of the 173 stations (b) with identification of stations where the PD, or BD or BND was validated.

101x67mm (300 x 300 DPI)

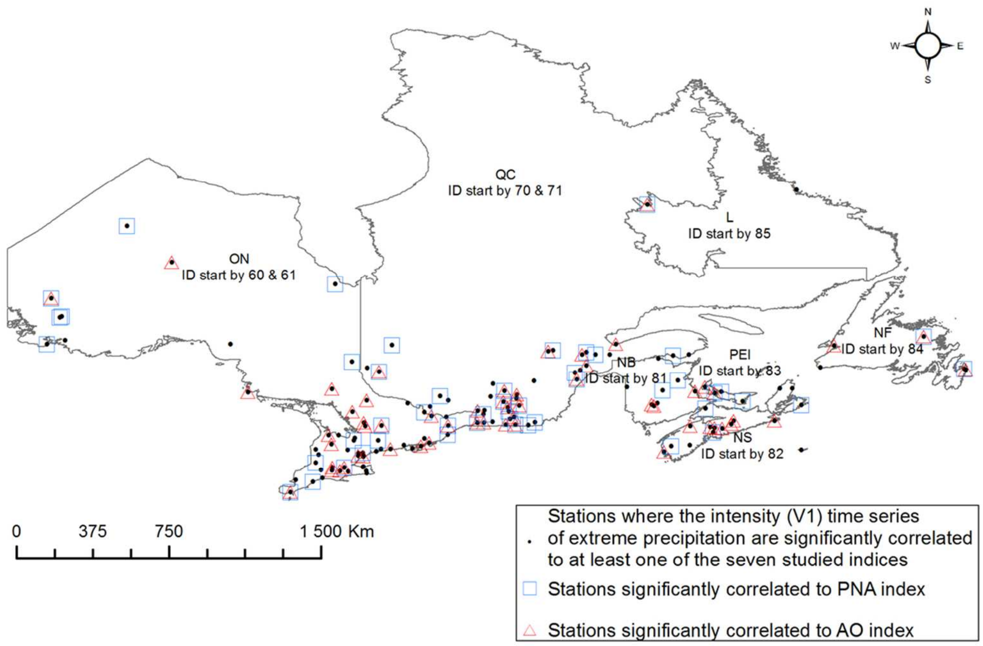


Figure 2. Identification of stations where the AO and PNA indices were significantly correlated to the V1 series among the 136 stations where V1 series showed significant correlation with at least one of the seven studied climate indices.

101x67mm (300 x 300 DPI)

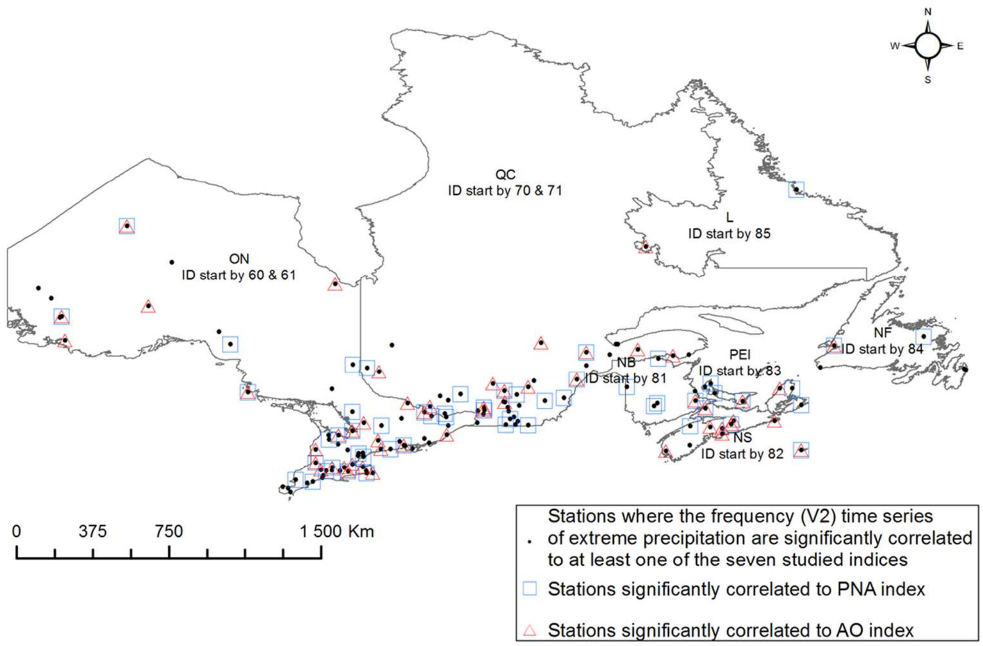


Figure 3. Identification of stations where the AO and PNA indices were significantly correlated to the V2 series among the 138 stations where V2 series showed significant correlation with at least one of the seven studied climate indices.

101x67mm (300 x 300 DPI)

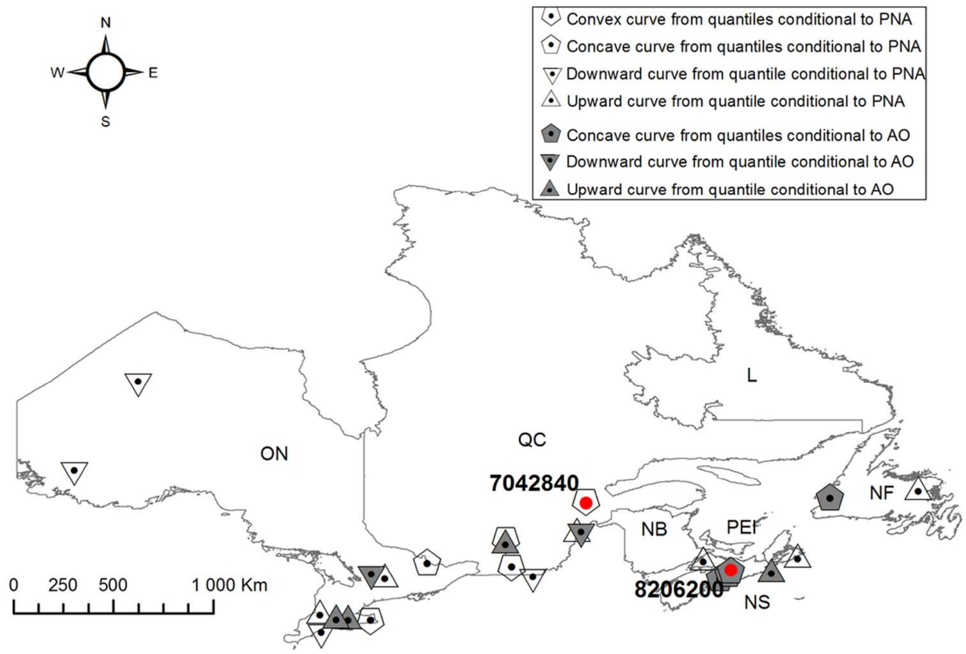


Figure 4. Linear and nonlinear response of extreme precipitations to AO and PNA indices at 25 stations in Southeastern Canada. Highlighted stations in red are selected for illustrative purpose.

101x67mm (300 x 300 DPI)

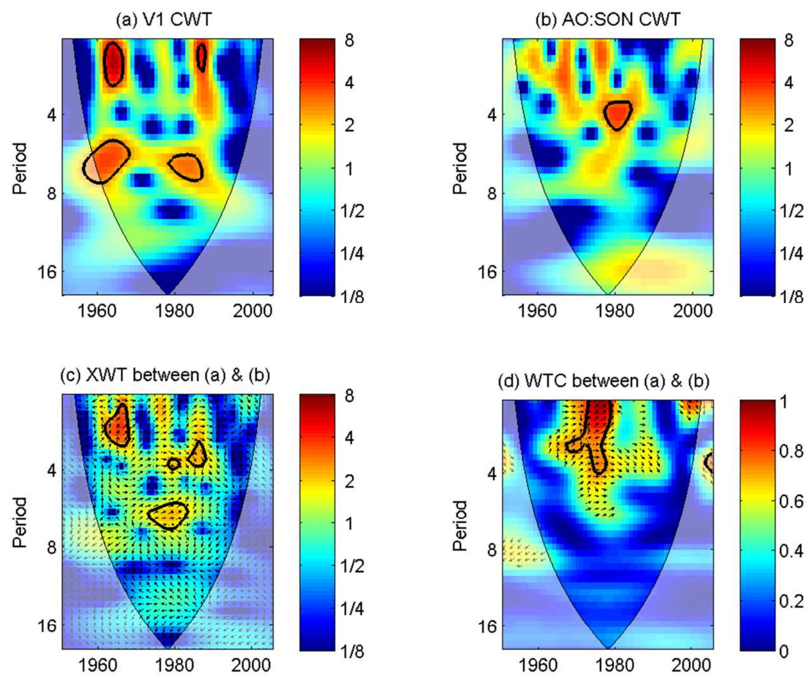


Figure 5. Wavelet analysis results with CWT illustration for V1 (a) and AO index: SON window (b). XWT (c) shows the common features between (a) & (b). WTC (d) highlights the covariance between (a) & (b).

203x152mm (150 x 150 DPI)

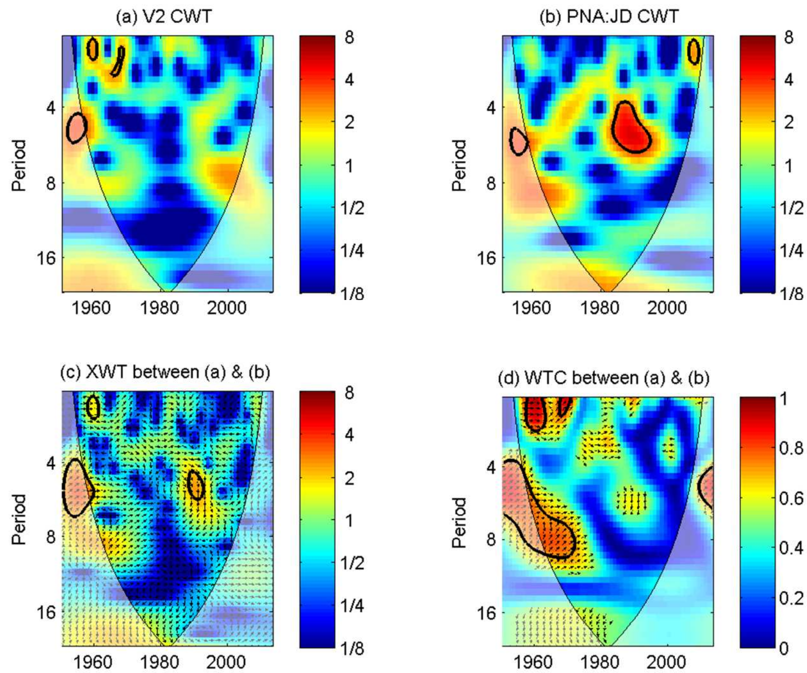


Figure 6. Wavelet analysis results with CWT illustration for V2 (a) and PNA index: JD window (b). XWT (c) shows the common features between (a) & (b). WTC (d) highlights the covariance between (a) & (b).

203x152mm (150 x 150 DPI)

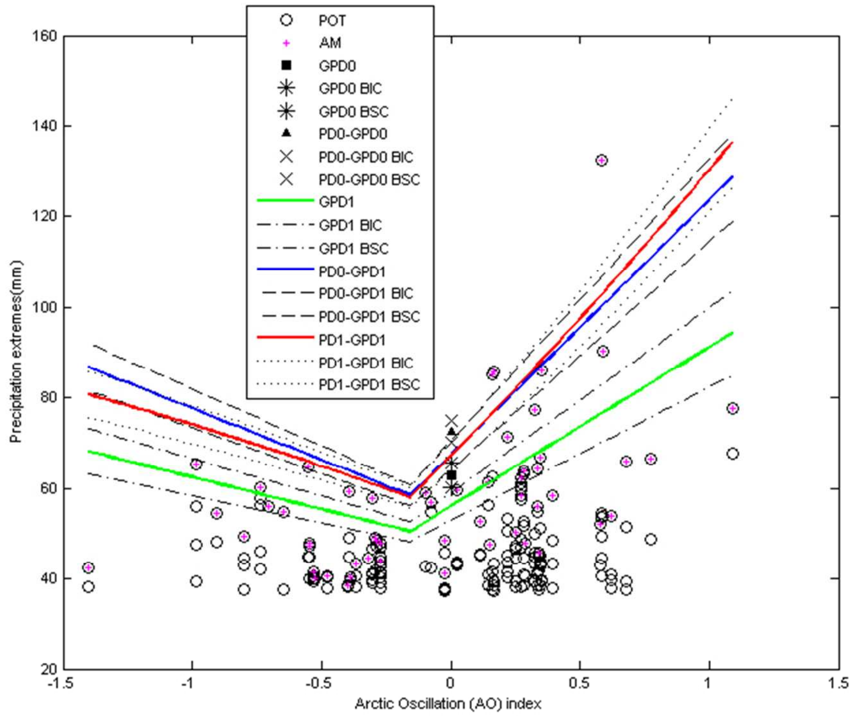


Figure 7. Estimated quantiles from stationary (GPD0, PD0-GPD0) and nonstationary (GPD1, PD0-GPD1, PD1-GPD1) models, using the Arctic Oscillation as a covariate, associated with a non-exceedance probability of 0.9. Lower (BIC) and upper (BSC) 5% confidence intervals are also shown. Illustrated observations are independent daily peaks (POT) and annual maxima (AM) precipitation values for the Upper Stewiacke station.

193x151mm (96 x 96 DPI)

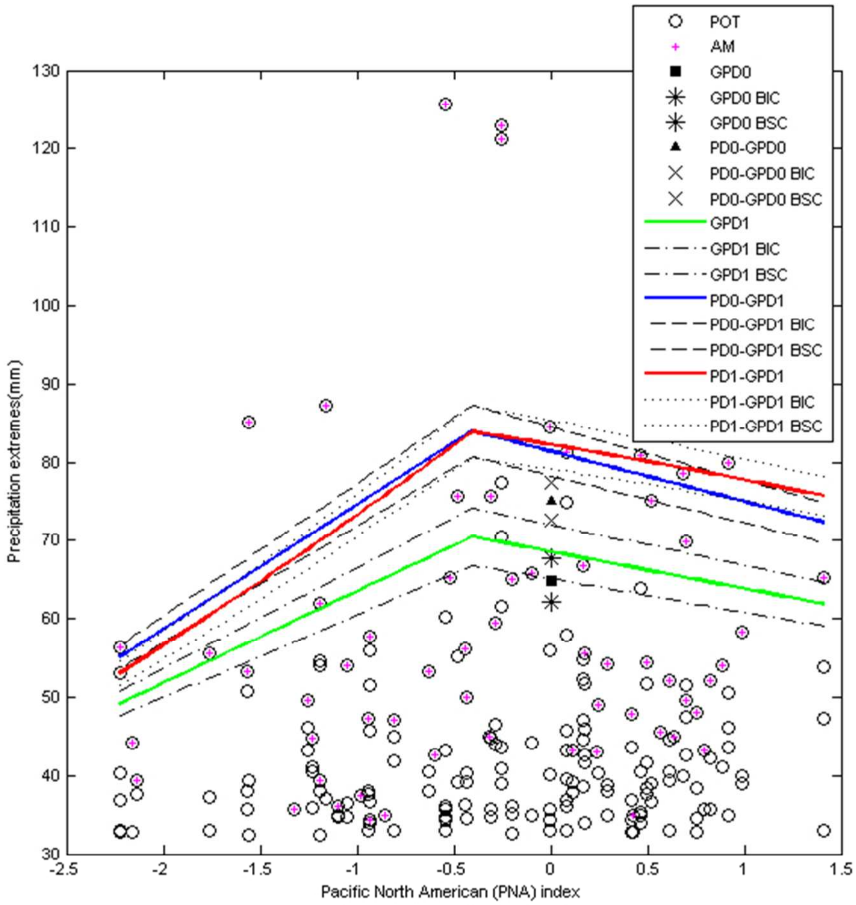


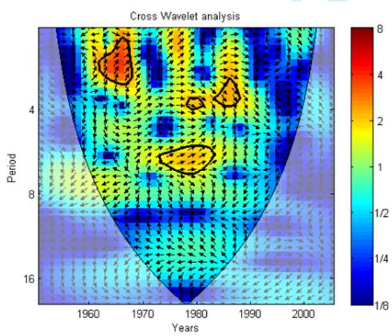
Figure 8. Same description as in Figure 7, but using the Pacific North American index as a covariate for the Grande Bergeronnes station.

174x168mm (96 x 96 DPI)

Graphical Table of Contents

Nonlinear response of precipitation to climate indices using a nonstationary Poisson-Generalized Pareto model: Case study of Southeastern Canada

Alida N. Thiombiano*, André St-Hilaire,·Salah-Eddine El Adlouni,·Taha B.M.J. Ouarda



Using statistical tools like the Cross Wavelet analysis illustrated in the above Figure, common features of variability are found between precipitation extreme events and the Artic Oscillation index at the Upper Stewiacke station located in Nova Scotia (Canada). Using this index as covariate, we developed nonstationary Poisson-Generalized Pareto models, which allow observing conditional quantiles with concave form. The proposed models are more flexible than classical extreme value nonstationary models which often used prior assumption of linear dependence.