Nonstationary and nonlinear Peaks-Over-Threshold model for the occurrence frequency and magnitude of extreme precipitation over Southeastern of Canada

Thiombiano A.N.[1], El Adlouni S.[2], St-Hilaire A.[1], Ouarda T.B.M.J.[1,3]

(1) INRS-ETE, (2) University of Moncton, (3) Masdar Institute of Science and Technology

INRS Université d'avant-garde

alida_nadege.thiombiano@ete.inrs.ca

STAHY 2016 Workshop
September 26-27, Quebec

Abstract

Significant interactions between hydroclimatic variables and climate indices time series have been found worldwide, indicating that low frequency oscillations indices may be useful in quantifying climate variability signals at longer temporal scales. With regard to climate change, such additional information can then be incorporated through the developing nonstationary statistical models for robust hydroclimatic extreme events frequency analysis.

Hence, extreme precipitation quantiles estimation is proposed in a Peaks-Over-Threshold (POT) framework, given a non-exceedance probability (p) and conditional on two identified dominant climate indices used as covariates over Southeastern of Canada: Arctic Oscillation (AO) on the one hand and Pacific North American (PNA) on the other hand.

The proposed methodology allowed investigating the performance of nonstationary models against stationary ones, and the advantages of taking into account additional information. It also showed its robustness for practical applications in hydrology, providing an additional statistical tool assessment for more realistic estimation of hydroclimatic extreme quantiles. Indeed, those constitute key information for hydraulic structure design as well as for water management.

Objectives

1. Determination of regional dominant climate indices significantly impacting hydroclimatic variables;

2. Development of extreme value theory-based models allowing to take into account the effect of such climate covariates on the intensity as well as on the frequency of extremes occurrence;

3. Estimation of quantiles for a given risk and special value of the covariate.

Methods

Objective 1 Tools

- Correlation analysis: nonlinear assumption based on literature (*Canon 2015; Shabbar et al. 1997; Thiombiano et al. 2016*) leading to Kendall tau computation at α=5%.
- Wavelet analysis: nonstationary hypothesis leading to temporal assessment of variability (*Chandran et al. 2015; Coulibaly 2006*).

Objectives 2&3 Tools

Let X a random hydroclimatic variable recorded over L-years and N the number of independent exceedances x_{i_k} i=1, 2, …, L and k=1, 2, …, N, over a threshold u. In stationary case, POT model (*Chavez-Demoulin and Davison 2005*), the number of exceedances per year m_i is assumed to follow a Poisson distribution (PD) with parameter λ so that $\hat{\lambda} = N/L$. The probability mass function of PD is:

$$\mathbf{P(m = k) = e^{-\lambda}\frac{\lambda^k}{k!}}$$

The magnitude of exceedances $y_{i_k} = x_{i_k} - u$ (so that $x_{i_k} > u$), is fitting by a Generalized Pareto distribution (GPD) whom cumulative distribution function (CDF) is:

$$\mathbf{G_{\xi,\sigma}(y) = 1 - \left(1 + \xi\frac{y}{\sigma}\right)^{-1/\xi} \qquad for \quad \xi \neq 0}$$

A complete POT model allows take into account the frequency of exceedances through the PD by combining PD and GPD into a Generalized Extreme Value (GEV) model (*Silva et al. 2016*). The inverse of below GEV CDF is considered for annual maxima (AM) quantiles estimation.

$$F_{\xi,\psi,\delta}(Z) = exp\left\{-\left[1 + \xi\left(\frac{Z - \psi}{\delta}\right)\right]^{-1/\xi}\right\} \qquad for\ \xi \neq 0$$

$$\text{where} \quad \delta = \frac{\sigma}{\lambda^{\xi}} \quad \text{and} \quad \psi = u + \delta\left(\frac{\lambda^{\xi} - 1}{\xi}\right)$$

In the nonstationary framework, σ and λ are allowed varying conditional to covariate A (σ_A and λ_A) using respectively a B-splines (*Nasri et al. 2011; Thiombiano et al. 2016*) and a PD link functions. The so-called GPD-B-splines model parameters are estimated with the generalized maximum likelihood method where a prior information is brought for GPD shape parameter (*El Adlouni and Ouarda 2008*), while the nonhomogeneous PD parameters are estimated through a Generalized Linear Model.

$$\mathbf{\sigma_A = h(A) = \beta_0 + \sum_{j=1}^{k} \beta_j B_{j,d}(A)}$$

$$\mathbf{\lambda_A = \log\lambda(A) = \lambda_0 + \lambda_1 A}$$

β_0, β_j are regression parameters, k is the number of control points, $B_{j,d}$ is a piecewise polynomial of degree d.

Case study

Preliminary analysis

- Enter variable= observed daily precipitation amounts over study region (Fig.1);
- Applied threshold u= 99th percentile with 1 day for exceedances declustering;
- Annual pooled data= V1 (V2) distribution of exceedances peaks (frequency);
- Climate covariates= AMO, AO, NAO, PDO, PNA, SOI, WHWP (*Bonsal and Shabbar 2011; Wang and Enfield 2001*).

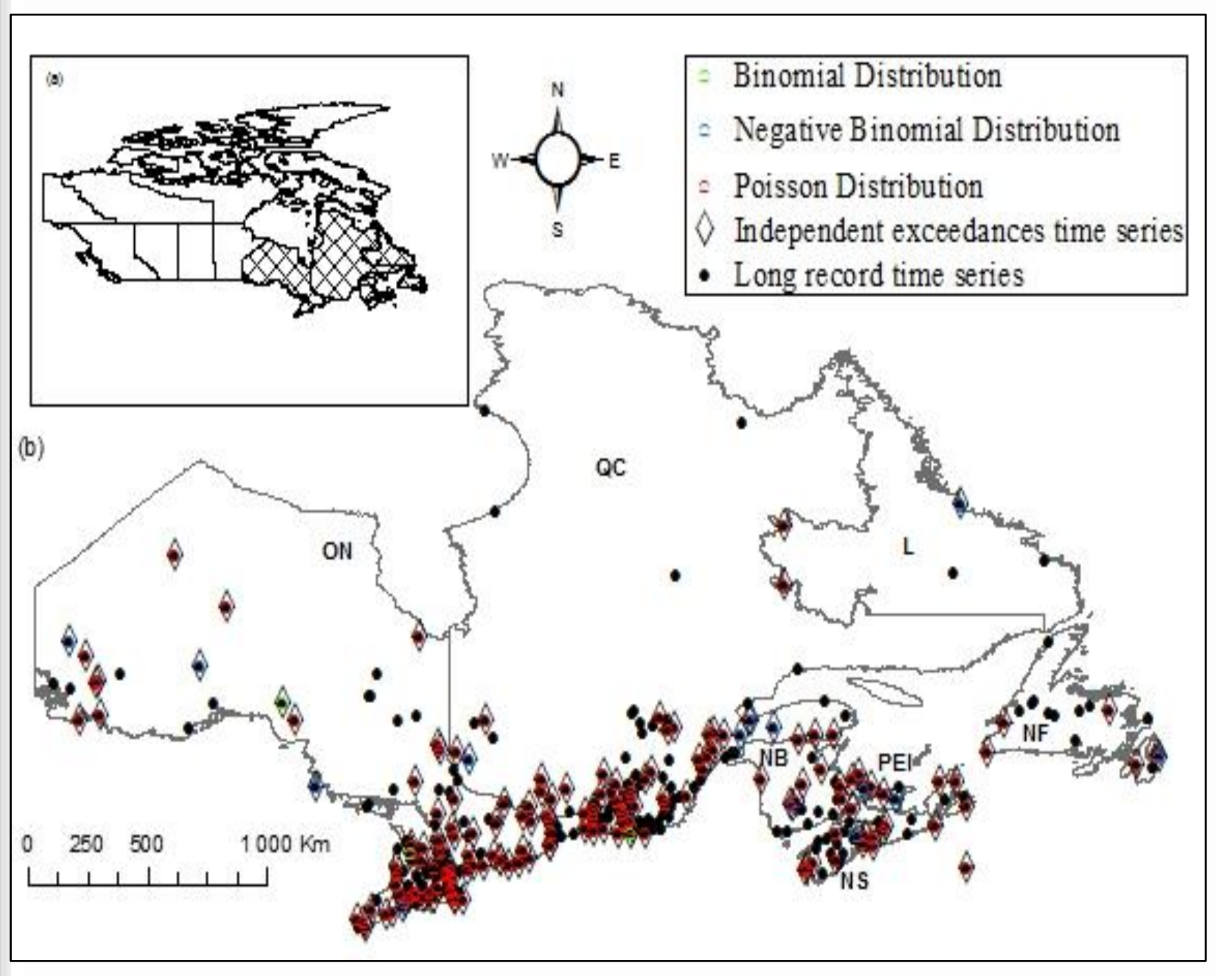


Fig1. Southeastern Canada (a) and stations location (b)

Table1. Stations selection process

Steps	PEI	NB	NFL	NS	ON	QC	Total
Full inventory	16	70	98	89	386	398	1057
Length >= 30 years	08	19	22	35	136	109	329
Independence criterion	06	09	08	18	78	54	173
Correlation V1 & >=1covariate	04	08	06	16	60	42	136
Correlation V2 & >=1covariate	04	06	07	14	66	41	138
Poisson	4	8	6	17	72	49	156
Binomial Negative	2	1	2	1	4	4	14
Binomial	0	0	0	0	2	1	3

Covariate, model, nonstationary quantile

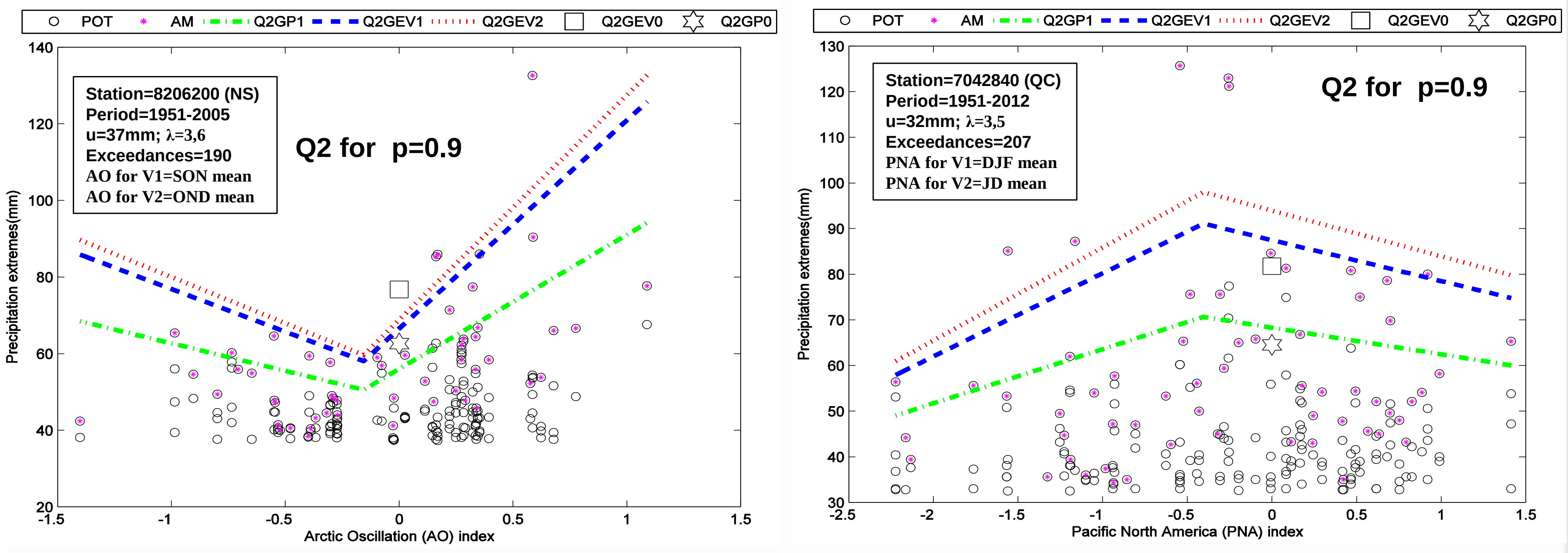
Table2. Nonlinear correlation analysis

Climate indices	AO	NAO	AMO	PNA	PDO	SOI	WHWP
Newfoundland (NF)	▓	░	▒	░	▒	░	▒
Labrador (L)	░	░	▨	░	▓	▨	▒
Prince Edward Island (PEI)	░	▓	▒	░	▒	▓	▓
Nova Scotia (NS)	░	▓	▒	▓	░	▒	▒
New Brunswick (NB)	░	▓	▒	░	▒	▓	▒
Quebec (QC)	░	▓	▒	░	▓	▒	▒
Ontario (ON)	░	░	▒	░	▓	▒	▒

░	▓	▒	▨
High correlation	Medium correlation	Less correlation	No correlation

- A significant and relatively strong correlation was found between PNA (AO) and V1&V2 at 15 (10) stations.
- GPD-B-splines with k=2 and d=1 was the best model (lowest AIC).
- 4 types of nonstationary quantiles identified: **Concave [**highest (lowest) quantiles are associated to indice extreme (moderate) values], **Convex, Upward and Downward** curves [quantiles decrease from indice negative to positive phase].

Illustrations



GP1= GPD-B-splines quantiles Q2; **GEV1**= AM quantiles from GPD-B-splines and λ); **GEV2**= AM quantiles from GPD-B-splines and λ_A); **GP0** and **GEV0** are stationary quantiles.

Fig2. Concave curve (left) & Convex curve (right) of quantiles

Conclusion

- Outperformance of nonstationary models given AIC.
- Risk of GP0 & GEV0 inaccuracy knowing covariate additional information.
- Risk of GP1 underestimation knowing λ and λ_A additional information.
- Enhancement of extremes during AO negative (positive) phase over QC&ON regions (over PEI, NF, NS & NB), and during PNA positive phase over study area.

Acknowledgments

Data Access Integration (DAI)
Environment Canada FACE portal

FACE
« Faire-face Aux Changements Ensemble »
Mieux s'adapter aux changements climatiques au Canada et en Afrique de l'Ouest dans le domaine des ressources en eau

http://www.esrl.noaa.gov/psd/data/climateindices/list/