



Space-time trends detection and dependence modelling approaches by functional Peaks-Over-Thresholds.

Application to precipitation in Burkina Faso

Béwentaoré Sawadogo, Diakarya Barro, Liliane Bel

► To cite this version:

Béwentaoré Sawadogo, Diakarya Barro, Liliane Bel. Space-time trends detection and dependence modelling approaches by functional Peaks-Over-Thresholds. Application to precipitation in Burkina Faso. JDS 2021 52èmes journées de statistique de la SFdS Nice, Jun 2021, Nice, France. hal-04489024

HAL Id: hal-04489024

<https://hal.science/hal-04489024>

Submitted on 4 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SPACE-TIME TRENDS DETECTION AND DEPENDENCE MODELLING APPROACHES BY FUNCTIONAL PEAKS-OVER-THRESHOLDS. APPLICATION TO PRECIPITATION IN BURKINA FASO

Béwentaoré Sawadogo^{1,2} & Liliane Bel¹ & Diakarya Barro^{2,3}

¹ *Université Paris Saclay, INRAe, AgroParisTech, UMR MIA-Paris, 75005, Paris, France, liliane.bel@agroparistech.fr*

² *LANIBIO, UJKZ, UFR-SEA, BP: 7021, Ouagadougou 03, Burkina Faso, sbewentaore@yahoo.fr*

³ *UFR-SEG, Université Thomas SANKARA, 12 BP: 417 Ouagadougou 12, Burkina Faso, dbarro2@gmail.com*

Abstract. In this paper, we propose a method for estimating trends in extreme spatio-temporal processes using both information from marginal distributions and dependence structure. We combine two statistical approaches of extreme value theory allowing on the one hand to handle temporal and spatial non-stationarities via a tail trend function with a spatio-temporal structure in the marginal distributions and by modelling on the other hand the dependence structure by a latent stationary process using generalized ℓ -Pareto processes. This methodology for trend analysis of extreme events is applied to precipitation data from Burkina Faso. We show that a significant increasing trend for the 50 and 100 years return levels in some parts of the country.

Keywords: Non-stationary POT, Generalized ℓ -Pareto process, Space-time Extremes, Dependence Modelling, Trends detection, Climate Change.

Résumé. Dans cet article, nous proposons une méthode pour estimer les tendances dans les extrêmes de processus spatio-temporels en utilisant à la fois des informations provenant des distributions marginales et de la structure de dépendance. Nous combinons deux approches statistiques de la théorie des valeurs extrêmes permettant d'une part de prendre en compte les non-stationnarités temporelles et spatiales via une fonction de tendance de queue à structure spatio-temporelle dans les distributions marginales et en modélisant d'autre part la dépendance spatiale à l'aide d'un processus stationnaire latent en utilisant les processus ℓ -Pareto généralisés. Cette méthodologie d'analyse des tendances des événements extrêmes est appliquée aux données de précipitations du Burkina Faso. Nous montrons que les niveaux de retour à 50 et 100 ans montrent une croissance significative dans certaines zones du pays.

Mots-clés: POT non stationnaire, processus ℓ -Pareto généralisé, extrêmes spatio-temporels, dépendance spatiale, détection de tendance, changement climatique.

1 Introduction

Climate extremes occur in a non stationary framework due to the action of climate change. To study extremes in this framework, the Extreme value Theory initially designed for independent or stationary processes must be extended. Several works focus on that topic, some of them present transform a non-stationary time series into a stationary series for which classical EVT can be applied ([1],[7]). A frequently alternative for addressing non-stationarity in EVT models is to include covariates, mostly space and time, to the model the parameters ([2],[11]). These approaches work marginally and do not take spatial dependence into account. Recently some forms of non-stationary dependence structures, have been studied by [8],[10], in the framework of spatial max-stable processes. Although attractive, these spatial models are very computationally expensive for large dimensions. And alternative for max-stable modelling is the modelling of threshold exceedances. The spatial Peak Overs Threshold (POT) approach introduced in a stationary framework by [6], and generalized by [5],[4] that have defined the family of generalized ℓ -Pareto process are good candidates for modelling the spatial dependence structure of extremes of spatio-temporal processes.

In this paper we propose a methodology coupling two statistical approaches of EVT, non-stationarity is handled in the marginal distributions via a trend function with spatio-temporal structure([7]) while the spatial dependence structure is modelled using a functional POT model([3, 4]) to obtain a flexible spatio-temporal method of threshold exceedances to assess presence of trends in the extremes. We aim at evaluating trend in extreme precipitation in sub-Saharan Africa by these means, and we illustrate theses evolution by representing non stationary high return levels over the area.

2 Methodology

2.1 Space-time trends modelling

Let $X = \{X_t(s), s \in S, t \in T\}$ be a continuous non-stationary space-time stochastic process with sample paths in the family of continuous functions $\mathcal{C}(S \times T)$ equipped with the uniform norm $\|\cdot\|_\infty$, where $S \times T \subset \mathbb{R}^d \times \mathbb{R}^+$ and $\mathcal{C}_+(S \times T)$ its restriction to non-negative functions deprived of the null function. In practice X is observed at each stations s_1, \dots, s_m and at dates $t = 1, \dots, n$. Let $F_{t,s}$ be the continuous univariate marginal distribution with a common right endpoint x_F and $Z = \{Z(s), s \in S\}$ an unobserved latent stationary stochastic process with sample paths in $\mathcal{C}(S \times T)$ satisfying the proportional tail condition such that

$$\lim_{x \rightarrow x_F} \frac{P(X_t(s) > x)}{P(Z(s) > x)} = c_\theta \left(\frac{t}{n}, s \right), \text{ with, } \frac{1}{m} \sum_{j=1}^m \int_0^1 c_\theta(u, s_j) du = 1, \quad u \in [0, 1], \quad (1)$$

where $c_\theta : [0, 1] \times S \rightarrow]0, \infty[$ is assumed to be a continuous and positive function depending on a parameter vector $\theta \in \Theta \subset \mathbb{R}$, called tail trend function or skedasis function ([6, 7]). The skedasis function describes the evolution of extreme events jointly in space and time. Moreover we assume that the continuous marginal distributions F_Z of the latent process has the same right endpoint x_F . F_Z is in the maximum domain of attraction condition for some constant $\gamma \in \mathbb{R}$ and appropriate real normalization constants $a_Z > 0$, $u_Z \in \mathbb{R}$. Thanks to equation(1) and the convergence of $\{Z(s), s \in S\}$ exceedances to a GPD distribution we deduce a pseudo-sample of $\{Z_t(s)\}$ from observations of $\{X_t(s), s \in S, t \in T\}$ in the following manner ([7]):

$$\hat{Z}_t(s_j) = \left\{ \hat{c}_\theta \left(\frac{t}{n}, s_j \right) \right\}^{-\hat{\gamma}} \left[X_t(s_j) - \frac{\{\hat{c}_\theta(\frac{t}{n}, s_j)\}^{\hat{\gamma}} - 1}{\hat{\gamma}} (\hat{a}_Z - \hat{\gamma} \hat{u}_Z) \right], \quad (2)$$

$$j = 1, \dots, m ; t = 1, \dots, n,$$

where $\hat{\gamma}$, \hat{a}_Z , \hat{u}_Z and \hat{c}_θ are respectively consistent estimators of γ , a_Z , u_Z and c_θ which we will discuss later in the section (2.2).

The modelling is then focused on the evaluation of the extreme spatial dependence structure in Z using functional POT([5], [3, 4]). In the multivariate and spatial framework a threshold exceedance for a random function $Z = \{Z(s), s \in S\}$ is defined by [5] to be an event of the form $\{\ell(Z) > u\}$ for some $u \geq 0$, where $\ell : \mathcal{C}(S) \rightarrow \mathbb{R}^+$ is a continuous and homogeneous non-negative risk function. The risk function ℓ can be for instance the maximum, minimum, average or value at a specific point $s_0 \in S$. As in ([6]), we assume that the stationary process Z is a general functional regular variation process. Under minimal assumptions on the risk function and for n large enough, the conditional distribution of ℓ -exceedance for some threshold $u \geq 0$ of the normalized process can be approximated by a generalized ℓ -Pareto process, i.e for some a_n and b_n functions ([4])

$$P \left\{ \left\lfloor \frac{Z - b_n}{\ell(a_n)} \right\rfloor \in A \mid \ell \left(\frac{Z - b_n}{\ell(a_n)} \right) \geq u \right\} \rightarrow P \{W_\ell \in A\}, n \rightarrow \infty, \quad (3)$$

where W_ℓ is a non-degenerate stochastic process over S and belongs to the family of generalized ℓ -Pareto processes with tail index γ .

2.2 Statistical inference

Marginal parameters of the process γ and a_n of the latent stationary process Z can be estimated by maximizing the independence log-likelihood taking $\hat{u}_Z = \ell(b_n) = q_{1-\alpha}\{\ell(Z)\}$ a high quantile of $\ell(Z)$. Assuming the dates of exceeding the marginal thresholds \hat{u}_Z form a Poisson process, its density is the trend function parameter c_θ which is estimated using maximum likelihood method, assuming a parametric log-linear model as in ([9]).

The inference on the spatial dependence structure of the latent stationary process, is driven by modelling the angular component of ℓ -Pareto process ([5], [3]) by a log-Gaussian

process with stationary increments within the framework of a Brown-Resnick model. The parameters of the isotropic semi-variogram are estimated using the gradient scoring rule method ([3]).

The non-stationary m -period return levels of the process X at each location $x_m(s)$ are evaluated inverting the equation(2) from return levels calculated from Z , $z_m(s)$.

$$x_m(s) = z_m(s)\hat{c}_\theta(t_m, s)^{\hat{\gamma}} + \frac{\hat{c}_\theta(t_m, s)^{\hat{\gamma}} - 1}{\hat{\gamma}} (\hat{a}_Z - \hat{\gamma}\hat{u}_Z), \quad s \in S, \quad (4)$$

$t_m = 1 + \frac{n_x m}{n}$ with n_x the number of days in the year and n the size of the sample observed. $z_m(s)$ are calculated at each site by extrapolating the marginal parameters on the whole area.

3 Main Results

This study uses time series of daily precipitation measurements from 1957 to 2016 provided by ten synoptic stations extracted from the Burkina Faso climatological database. These stations have been selected to ensure good spatial uniformity and representativeness of different climatic regimes and data quality. In order to limit the problems related to seasonal rainfall cycles, on each station we worked from the sub-series corresponding to rainy days that is the period from May to October. Figure (1) shows the estimated trend function c_θ by a log-linear model for the ten stations. The frequencies of extreme precipitation are increasing in areas such as Ouahigouya, Bogande, Boromo, Gaoua, and Po. On the other hand, at the Ouagadougou and Fada stations, extreme rainfall frequencies tend to decrease ($\theta < 0$). Figure (2) shows the non-stationary return levels $x_m(s)$

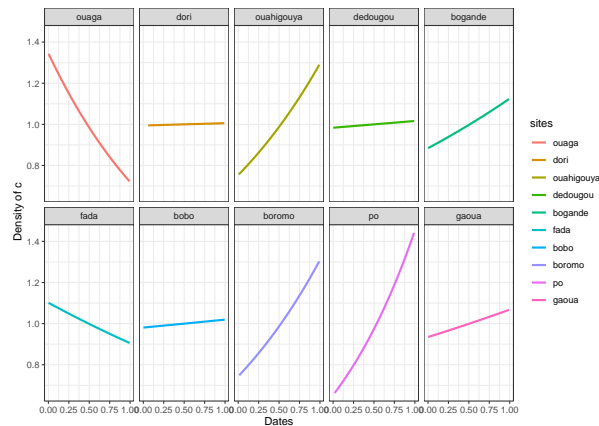


Figure 1: Local adjustment of evolution of the frequencies of extreme precipitation by a log-linear model trend c_θ .

Sudanian climate Sudano-Sahelian climate Sahelian climate

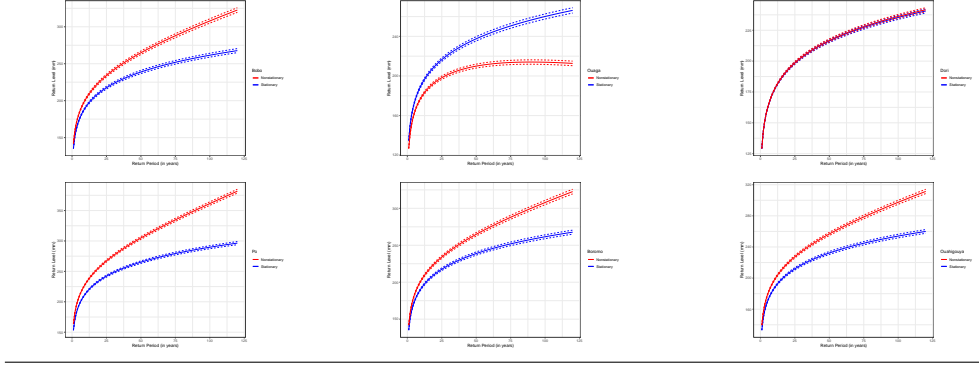


Figure 2: Trends of stationary and non-stationary return levels at reference stations calculated with a log-linear tail trend function.

together with the stationary return levels $z_m(s)$ for six reference stations.

The map of non-stationary return level for a return period $m = 50$ and 100 years are displayed in Figures (3). The extreme precipitation are likely to be observed on average at least once every 50 years (resp. 100 years), will be particularly intense in the Sudanian and Sudano-Sahelian zone and less intense in the Sahelian zone, with a potentially quite strong spatial dependence within a radius of 200 km. The southwest and eastern region of the country will be most affected by extreme precipitation.

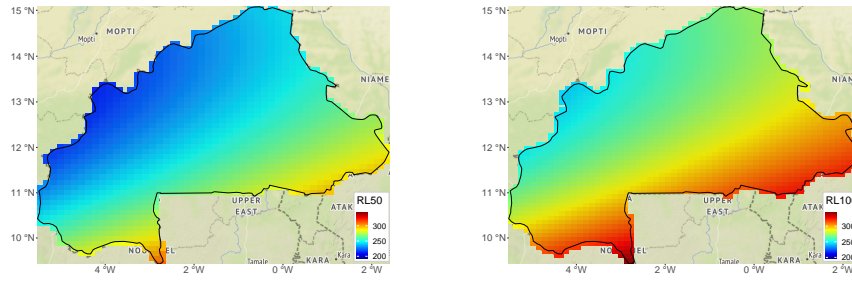


Figure 3: Maps of the non-stationary 50-years return levels(left) and 100-years return level (right) obtained by extrapolating the information from the log-linear tail trend function and the dependence structure.

4 Conclusion

In this study we proposed a new flexible methodology for trend detection in the extremes capable of capturing marginal non-stationarities and the dependence structure between margins using generalized ℓ -Pareto processes. We calculated the non-stationary return

levels of precipitation in Burkina Faso and exhibit some regions in which there may be a significant increasing or decreasing of extreme precipitation.

References

- [1] Didier Dacunha-Castelle, Thi Thu Huong Hoang, and Sylvie Parey. Modeling of air temperatures: preprocessing and trends, reduced stationary process, extremes, simulation. *Journal de la Société Française de Statistique*, 156(1):138–168, 2015.
- [2] Anthony C Davison and Richard L Smith. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52(3):393–425, 1990.
- [3] R. de Fondeville and A. C. Davison. High-dimensional peaks-over-threshold inference. *Biometrika*, 105(3):575–592, 2018.
- [4] R. de Fondeville and A.C. Davison. Functional peaks-over-threshold analysis, 2020. Preprint.
- [5] Clément Dombry and Mathieu Ribatet. Functional regular variations, pareto processes and peaks over threshold. *Statistics and its Interface*, 8(1):9–17, 2015.
- [6] Ana Ferreira and Laurens De Haan. The generalized pareto process; with a view towards application and simulation. *Bernoulli*, 20(4):1717–1737, 2014.
- [7] Ana Ferreira, Petra Friederichs, Laurens de Haan, Cláudia Neves, and Martin Schlather. Estimating space-time trend and dependence of heavy rainfall. *arXiv preprint arXiv:1707.04434*, 2017.
- [8] Raphaël Huser and Marc G Genton. Non-stationary dependence structures for spatial extremes. *Journal of agricultural, biological, and environmental statistics*, 21(3):470–491, 2016.
- [9] Aline Mefleh, Romain Biard, Clément Dombry, and Zaher Khraibani. Trend detection for heteroscedastic extremes. *Extremes*, 23(1):85–115, 2020.
- [10] Gilles Nicolet, Nicolas Eckert, Samuel Morin, and Juliette Blanchet. Assessing climate change impact on the spatial dependence of extreme snow depth maxima in the french alps. *Water Resources Research*, 54(10):7820–7840, 2018.
- [11] Sylvie Parey, Thi Thu Huong Hoang, and Didier Dacunha-Castelle. Different ways to compute temperature return levels in the climate change context. *Environmetrics*, 21(7-8):698–718, 2010.