Journal of Multivariate Analysis 159 (2017) 49–66

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Non-linear models for extremal dependence

Linda Mhalla [a,*], Valérie Chavez-Demoulin [b], Philippe Naveau [c]

[a] Faculté d'Économie et de Management (GSEM), Université de Genève, Genève, Suisse, Switzerland

[b] Faculté des Hautes Études Commerciales (HEC), Université de Lausanne, Lausanne, Suisse, Switzerland

[c] Laboratoire des Sciences du Climat et l'Environnement (LSCE), CNRS, Gif-sur-Yvette, France

ARTICLE INFO

Article history:
Received 12 October 2016
Available online 29 April 2017

AMS subject classifications:
62Fxx
62Gxx
62Hxx
62Jxx

Keywords:
Extreme value theory
Generalized additive models
Max-stable random vectors
Non-stationarity
Pickands function
Semi-parametric models
Temperature data

ABSTRACT

The dependence structure of max-stable random vectors can be characterized by their Pickands dependence function. In many applications, the extremal dependence measure varies with covariates. We develop a flexible, semi-parametric method for the estimation of non-stationary multivariate Pickands dependence functions. The proposed construction is based on an accurate max-projection allowing to pass from the multivariate to the univariate setting and to rely on the generalized additive modeling framework. In the bivariate case, the resulting estimator of the Pickands function is regularized using constrained median smoothing B-splines, and bootstrap variability bands are constructed. In higher dimensions, we tailor our approach to the estimation of the extremal coefficient. An extended simulation study suggests that our estimator performs well and is competitive with the standard estimators in the absence of covariates. We apply the new methodology to a temperature dataset in the US where the extremal dependence is linked to time and altitude.

1. Introduction

Extreme events are of major importance in different fields. Such events, defined as rare and severe, may cause considerable material damage, deaths, and economic losses. A proper statistical framework should then provide the tools for measuring extreme events through the quantification of their frequencies, as well as intensities, which can be formulated in terms of quantiles. Often, in practice, there is much more to gain by understanding simultaneous extreme events of several quantities. For example, in finance, people are interested in the values of several assets that constitute the portfolio and the risk that the assets collapse together. In environmental sciences, it is vital to measure the risk of simultaneous flooding at different points of a river [2] or at nearby rivers. In both cases, the quantity of interest is measured by the extremal dependence between different variables. There is a well-established literature on modeling multivariate extremes and thus, extremal dependence structures; some authors addressed this matter using Gaussian copulas [41] while others analyzed multivariate extreme value copulas [26]. When the interest is in modeling multivariate extremes, the notion of max-stable processes is very important. Max-stable processes arise as the limiting distribution of suitably normalized componentwise maxima of independent replications of some continuous stochastic process [17,18,45]. The dependence structure of a multivariate max-stable random vector can be characterized by its Pickands dependence function A (see

* Corresponding author.
E-mail addresses: linda.mhalla@unige.ch (L. Mhalla), valerie.chavez@unil.ch (V. Chavez-Demoulin), philippe.naveau@lsce.ipsl.fr (P. Naveau).

http://dx.doi.org/10.1016/j.jmva.2017.04.006
0047-259X/

Section 2.1) defined on the unit simplex $S_{d-1} = \left\{\boldsymbol{\omega} = (\omega_1, \dots, \omega_d) \in \mathbb{R}_+^d : \omega_1 + \cdots + \omega_d = 1\right\}$ through the Pickands representation theorem [39].

In this work, our main interest is modeling possible changes in the dependence structure among multivariate block maxima according to a given set of covariates $\mathbf{x}$. Moreover, the effect of some covariates does not need to have a linear form but may vary smoothly. In classical regression analysis, a popular and flexible set of models is the class of generalized additive models [24,28], which link the mean behavior of a random variable Y with a set of covariates $\mathbf{X} \in \mathbb{R}^q$ through

$$\mathrm{E}\,(Y \mid \mathbf{X} = \mathbf{x}) = g\left\{\mathbf{u}^\top \boldsymbol{\beta} + \sum_{k=1}^{K} h_k(t_k)\right\}, \tag{1}$$

where

- g is a link function,
- $(u_1, \dots, u_s)$ and $(t_1, \dots, t_K)$ are subsets of $\{x_1, \dots, x_q\}$,
- $\boldsymbol{\beta} \in \mathbb{R}^s$ is a vector of parameters, and
- $h_k : \mathbb{T}_k \to \mathbb{R}$ are smooth functions supported on closed $\mathbb{T}_k \subset \mathbb{R}$, for all k.

Chavez-Demoulin and Davison [6] proposed a first step to bridge extreme value theory and generalized additive models (GAM). They modeled the marginal behavior of extremes with a GAM, but they did not investigate the effect of covariates on the dependence structure. In this paper, we fill this gap in the context of block maxima analysis by developing a methodology for estimating conditional Pickands dependence functions $A(\cdot|\mathbf{x})$. The method relies on the max-projection of a d-dimensional max-stable random vector $\mathbf{Z}(\mathbf{x})$, where $\mathbf{Z}(\mathbf{x}) = (Z_1(\mathbf{x}), \dots, Z_d(\mathbf{x}))^\top$ given a set of covariates $\mathbf{x}$. We show that, conditionally on $\mathbf{x}$ and at any fixed value in the unit simplex, the univariate max-projection follows a classical two-parameter Beta distribution. The second parameter of the Beta distribution is equal to 1, and the first parameter is exactly the Pickands function evaluated at the specific value in the unit simplex. Under the resulting likelihood-based framework, we benefit from the well-established framework of the GAM to fit non-stationary Pickands functions evaluated at the specific value in the unit simplex. Repeating the max-projection procedure along with the GAM at several values in the unit simplex generates estimates of semi-parametric models for the Pickands function conditional on $\mathbf{x}$. In the bivariate case, for a fixed value of $\mathbf{x}$, the estimated Pickands function is regularized by constrained median smoothing B-splines. When $d \geq 3$, the focus is on the regularization of the non-stationary extremal coefficient estimate (see Section 3.4), as it allows for a simple interpretation of the extremal dependence in high dimensions.

To the best of our knowledge, only [21] addressed the dependence of the Pickands function on covariates. They considered a robust estimator, in the bivariate case, using local estimation with the minimum density power divergence criterion. Few works addressed specific non-stationarity in the multivariate and spatial contexts: Huser and Genton [31] developed non-stationary max-stable dependence structures in which covariates are incorporated, and their inference is based on pairwise likelihoods. de Carvalho and Davison [16] constructed a model for a family of spectral measures from several populations, to each of which a set of predictors is considered. They introduced the spectral density ratio model that they fitted using empirical likelihood methods and showed how to relate the resulting tilting parameters to covariates. Jonathan et al. [32] proposed a spline-based methodology to incorporate multiple covariates. Their approach used a Poisson process model for the rate of occurrence of the threshold excesses and a generalized Pareto model for the size of the threshold excesses. Our method provides a semi-parametric, flexible framework based on the classical setup of generalized additive models.

The paper is organized as follows: In Section 2, we summarize the theory of max-stable random vectors and then introduce the notion of max-projection from multivariate to univariate settings. In Section 3, we develop the generalized additive model for the Pickands dependence function in the multivariate context. We then propose a way to regularize the Pickands function in the bivariate case and the extremal coefficient in higher dimensions. We implement a bootstrapping procedure to compute the variability bands in Section 3.3. Finally, in the last two sections, we perform simulations and address the estimation of a conditional Pickands model for the monthly maxima of the maximum and minimum temperatures in the US.

2. Max-stable random vectors and max-projection

In this section, we first briefly recall some basics of max-stable random vectors and then propose a max-projection procedure that synthesizes the extremal dependence of multivariate max-stable random vectors within a univariate Beta random variable.

2.1. Max-stable random vectors

The usual strategy for modeling multivariate block maxima is to take advantage of the solid theoretical basis for modeling extreme events by completely characterizing the tail behavior of the underlying generating process. Multivariate extreme value theory (see [42, Chapter 5] and [18, Chapter 6]) addresses the limiting behavior of suitably normalized componentwise maxima $\mathbf{M}_n = (M_{n,1}, \dots, M_{n,d})^\top$ where $M_{n,j}$ represents the maximum, over a block of size n, of the jth component of a vector