Spatial Statistics 21 (2017) 187–208

Contents lists available at ScienceDirect

Spatial Statistics

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

Modeling extreme rainfall
A comparative study of spatial extreme value models

Quentin Sebille, Anne-Laure Fougères, Cécile Mercadier *

Univ. Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 43 blvd. du 11 novembre 1918, F-69622 Villeurbanne cedex, France

ARTICLE INFO

Article history:
Received 6 October 2016
Accepted 26 June 2017
Available online 3 July 2017

Keywords:
Spatial modeling of extreme events
Extreme value theory
Max-stable processes
Hierarchical models
Spatial prediction
Precipitation data

ABSTRACT

In this paper, focus is done on spatial models for extreme events and on their respective efficiency regarding the estimation of two risk measures: one extrapolating marginal distributions and one summarizing the spatial bivariate dependence of extremes. A wide comparison is performed on an innovative simulation plan that has been specifically designed from a daily precipitation dataset. The objective of this paper is twofold: firstly, pointing out the inherent properties of each model, and secondly, advising users on how to choose the model depending on the specific type of risk.

1. Introduction

Analyses of extreme values of environmental variables such as precipitation are of great importance since it involves human lives as well as considerable loss of money when a catastrophic event occurs. For instance, between May 28th and June 7th 2016, extreme precipitation events affected a part of Europe including France. This huge and sudden amount of rainfall caused nineteen deaths and, for France only, one billion Euro in damages. Accurate risk measures for such extreme phenomena are therefore needed to prevent from this type of scenario.

The risk estimation of these events is challenging because they involve values that are beyond the range of the observations. For this purpose, adapted tools come from extreme value theory. See for instance de Haan and Ferreira (2006), Beirlant et al. (2004), Finkenstädt and Rootzén (2004) and Coles (2001). Since precipitation phenomena have a spatial feature and data are generally observed at

* Corresponding author.
E-mail address: mercadier@math.univ-lyon1.fr (C. Mercadier).

http://dx.doi.org/10.1016/j.spasta.2017.06.009

several stations, the dedicated setting to handle this question is that of (conditionally) max-stable spatial models. Detailed and helpful reviews on these models are Cooley et al. (2012), Davison et al. (2012) or Ribatet (2013).

Five spatial models have been selected among the most popular and the most recent in the literature. This choice includes Bayesian and frequentist concepts, and goes from simple to more complex spatial dependence structure. Within this paper, the main goal is to answer: *Which of the five competing models yield the best spatial prediction for extreme behaviors of simulated processes mimicking precipitation data?* Addressing this question supposes in particular a careful consideration of the datasets involved, as well as a relevant choice of performance criteria.

Two comparative criteria adapted to extreme events prediction are evaluated. The estimation of rare events at a location where no data is available is handled first; this induces the capacity of spatial extrapolation of the extreme behavior when looking at marginal information only. A clear and well known way to summarize this marginal information into a concrete risk measure is the return level. Then a second and complementary criterion is the measure of extreme sets for the bivariate distribution at a pair of locations. This aims at capturing the spatial dependence structure of extremes. One could also look at indices involving more than two dimensions, but most of the dependent models for extremes have explicit formulae only in the bivariate case, so it would induce heavier numerical calculations.

Several options can be chosen to define the terms of comparison. One possible option could be to start from an expert point of view and compare each model with an a priori value of the previous criteria. To depart from a subjective choice, an intensive simulation study has been preferred. *How could one choose inside each of the five models a meaningful representative mimicking rainfall data?* Such a question has not a straight answer. A parametric bootstrap procedure has been set up here. More precisely:

- A real precipitation dataset is considered over a central-east region of France on which each of the five spatial models is fitted. These fitted models are then fixed to play the role of extreme rainfall generators.
- Both criteria (return level and bivariate extremal dependence) are evaluated on each generator so that every single generating process has its own true value of a criterion.
- The five spatial models are again considered and fitted on each generated sample drawn from one of the five generators. Both criteria are then evaluated on each fitted model and finally compared to the corresponding (known) true value.

To our knowledge, this way to proceed is original. We believe it offers a better insight into the comparative study than applying a given methodology to a single dataset, or estimating a given process inside its own class of model only.

The remainder of this article is organized as follows. The theoretical background of extreme value theory is addressed in Section 2, with an emphasis on the so-called block maxima approach and on the five (conditionally) max-stable spatial models that we consider. Section 3 describes the simulation study, the two criteria used to evaluate each case and the results. Conclusions are drawn in Section 4, where some recommendations are provided with respect to different purposes.

2. Notations and models

2.1. Definition of max-stable processes

Let $\mathcal{S}$ be a compact subset of $\mathbb{R}^d$ that represents the spatial region of interest, d being a positive integer. Consider a random process $Y(\cdot) = \{Y(s)\}_{s\in\mathcal{S}}$ defined over $\mathcal{S}$, with continuous sample paths. Write $Y_1(\cdot), \ldots, Y_T(\cdot)$ for independent copies of $Y(\cdot)$. The process Y is called *max-stable* if for each $T > 1$, there exist continuous functions $a_T(\cdot) > 0$ and $b_T(\cdot) \in \mathbb{R}$ such that:

$$\bigvee_{t=1}^{T} \frac{Y_t(\cdot) - b_T(\cdot)}{a_T(\cdot)} \stackrel{d}{=} Y(\cdot),$$