Journal of Statistical Planning and Inference 139 (2009) 1266 – 1275

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

An extended Gaussian max-stable process model for spatial extremes

Elizabeth L. Smith[a,*], Alec G. Stephenson[b]

[a]*Department of Statistics and Applied Probability, National University of Singapore, Singapore 11756*
[b]*Faculty of Life and Social Sciences, Hawthorn Campus, Swinburne University of Technology, VIC 3122, Australia*

ARTICLE INFO

Article history:
Received 31 May 2007
Received in revised form
5 August 2008
Accepted 8 August 2008
Available online 14 August 2008

Keywords:
Bayesian modelling
Extreme value theory
Max-stable processes
Multivariate extreme value theory
Spatial extremes

ABSTRACT

The extremes of environmental processes are often of interest due to the damage that can be caused by extreme levels of the processes. These processes are often spatial in nature and modelling the extremes jointly at many locations can be important. In this paper, an extension of the Gaussian max-stable process is developed, enabling data from a number of locations to be modelled under a more flexible framework than in previous applications. The model is applied to annual maximum rainfall data from five sites in South-West England. For estimation we employ a pairwise likelihood within a Bayesian analysis, incorporating informative prior information.

1. Introduction

Multivariate extreme value theory provides a way to quantify the joint extreme behaviour of several variables. Many environmental processes are spatial in nature and modelling multivariate data can be of great importance. Methods for modelling such data, based on a multivariate generalisation of the classical approach to univariate extremes, were proposed by Tawn (1988, 1990) and Smith et al. (1990). Multivariate extensions of the threshold exceedance approach were developed by Coles and Tawn (1991) and Joe et al. (1991).

Smith (1990) suggests a procedure using the theory of max-stable processes for modelling data which are collected on a grid of points in space. This approach can be considered as an infinite dimensional extension of multivariate extreme value theory. The extension has the advantage that it can be used to consider problems concerning aggregation of the process over the whole region, and interpolation to anywhere within the region. Another advantage of this approach is that models based on the resulting family of multivariate extreme value distributions are workable even for a large number of grid points. Other applications of max-stable processes are given in Coles (1993), Coles and Walshaw(1994) and Coles and Tawn (1996a). Coles (1993) gives a class of max-stable process models which can utilise all data exceeding pre-defined thresholds. Max-stable processes are used by Coles and Walshaw(1994) to model the directional dependence of extreme wind speeds, and by Coles and Tawn (1996a) to develop a model for spatially aggregated rainfall extremes.

Bayesian techniques can be used to incorporate information other than the data into the model in the form of prior distributions. These techniques have the potential to be very useful in improving estimation in extreme value problems, since extreme data are naturally scarce. Applications of Bayesian techniques to univariate problems have been considered by Coles and Powell (1996) and Coles and Tawn (1996b), and applications to multivariate problems have been considered by Smith and Walshaw (2003) and Smith (2005).

* Corresponding author at: School of Mathematics and Statistics, Newcastle University, Newcastle upon Tyne NE1 7RU, UK.
E-mail address: lizsmith6@msn.com (E.L. Smith).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2008.08.003

The Gaussian max-stable process, introduced by Smith (1990), suffers from a lack of flexibility, since the Gaussian storm profile that generates the process has a constant variance matrix across all points in the space of interest. In this paper the model is extended to a more flexible framework and is applied to annual maximum rainfall data from five locations in South-West England. For estimation, Smith (1990) suggests a collection of somewhat ad hoc techniques, due to the difficulty arising from the fact that the joint distribution of the process at more than two sites is unknown. The approach we take here is to estimate our extended model using a pairwise likelihood within a Bayesian analysis. The pairwise likelihood, which is the product of the likelihoods for all pairs of sites, comes from a general class of composite likelihoods introduced by Lindsay (1988). Hjort (1993) was the first to consider using a product of bivariate likelihoods. Hjort (1993) called the approach quasi-likelihood but since then the name pairwise likelihood has been adopted. In the Bayesian analysis we use informative priors for the marginal site parameters, based on expert prior information given in Coles and Tawn (1996a) . Prior distributions for parameters relating to the dependence structure of the process are specified to be non-informative.

2. Extreme value theory

Let $Y_1, Y_2, \ldots$ be an independent and identically distributed (IID) sequence of random variables. Classical univariate extreme value theory is concerned with the limiting behaviour of $M_n = \max\{Y_1, \ldots, Y_n\}$ as $n \to \infty$, after a linear normalisation of M_n. The key result of univariate extreme value theory states that the limiting distribution of M_n is the generalised extreme value (GEV) distribution:

$$G(z) = \exp\left[-\left\{1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right\}^{-1/\xi}\right], \tag{2.1}$$

for $\{z : 1 + \xi(z-\mu)/\sigma > 0\}$ and with $-\infty < \mu < \infty, \sigma > 0$ and $-\infty < \xi < \infty$.

For multivariate extremes, we focus on the limiting distribution, as $n \to \infty$, of the vector

$$\mathbf{M}_n = \left(\max_{i=1,\ldots,n} \{Y_{i,1}\}, \ldots, \max_{i=1,\ldots,n} \{Y_{i,p}\}\right)$$

after a linear normalisation of each component, where $\mathbf{Y}_1, \mathbf{Y}_2, \ldots$ is an IID sequence of random vectors on $\mathbb{R}^p$ and $\mathbf{Y}_i = (Y_{i,1}, \ldots, Y_{i,p})$. Each of the p marginal components can be considered separately as sequences of independent, univariate random variables. The margins of the limit distribution therefore have the form of a particular member of the GEV family in Eq. (2.1). All possible limit distributions of normalised $\mathbf{M}_n$ have the representation

$$G(\mathbf{x}) = G(x_1, \ldots, x_p) = \exp\left\{-\int_{S_p} \max_{j=1,\ldots,p} (\eta_j/y_j)\, \mathrm{d}H(\boldsymbol{\eta})\right\}, \tag{2.2}$$

where H is a positive measure on the simplex $S_p = \{\boldsymbol{\eta} = (\eta_1, \ldots \eta_p) \in \mathbb{R}_+^p : \sum \eta_j = 1\}$, subject to

$$\int_{S_p} \eta_j \, \mathrm{d}H(\boldsymbol{\eta}) = 1, \quad j = 1, \ldots, p,$$

and where

$$y_j = \left\{1 + \xi_j\left(\frac{x_j - \mu_j}{\sigma_j}\right)\right\}^{1/\xi_j}, \quad j = 1, \ldots, p. \tag{2.3}$$

The measure H characterises the dependence structure of the limiting distribution G, and (μ_j, σ_j, ξ_j) give the GEV parameters for the jth margin $G_j(x_j) = \exp(-1/y_j)$.

3. Max-stable processes

3.1. Theory

A stochastic process $\{Y_t, t \in T\}$, where T is an arbitrary index set, is a max-stable process (de Haan, 1984) if the following property holds:

> if $Y_t^{(1)}, \ldots, Y_t^{(n)}$ are n independent copies of the process, then there exists constants $a_{n_t} > 0$ and b_{n_t} such that $\{Y_t^*, t \in T\}$ is identical in law to $\{Y_t, t \in T\}$, where
>
> $$Y_t^* = \left(\max_{i=1,\ldots,n} Y_t^{(i)} - b_{n_t}\right) \Big/ a_{n_t}, \quad t \in T.$$