Smooth Spatial Maximum Likelihood GEV Fitting with gevreg

Harald Schellander

ZAMG Innsbruck

Abstract

The **gevreg** package provides functions for maximum likelihood estimation of smooth spatial regression models for the Generalized Extreme Value Distribution GEV. Models for the location, scale and shape parameter of the GEV can have different regressors. Suitable standard methods to compute predictions are provided as well. The model and its R implementation is introduced and illustrated by usinf snow depth data for Austria.

Keywords: regression, GEV, smooth spatial, R.

concordance:gevreg.tex:gevreg.Rnw:1 37 1 1 5 66 1

1. Introduction

A spatial representation of meteorological extremes such as snow depth is of crucial importance for numerous purposes such as the planning and construction of buildings, for avalanche simulation (Rudolf-Miklau and Sauermoser 2011) or in general risk assessment.

A more or less simple interpolation of GEV parameters to space has some disadvantages, as e.g. Blanchet and Lehning (2010) showed for extreme snow depths in Switzerland. As an improvement they suggested a direct estimation of a spatially smooth generalized extreme value (GEV) distribution, called *smooth spatial modeling*. With smooth modeling the GEV parameters are modeled as smooth functions of spatial covariates. Spatially varying marginal distributions are achieved by maximizing the sum of the log-likelihood function over all stations. Compared to several interpolation methods, smooth modeling for swiss snow depth led to more accurate marginal distributions, especially in data sparse regions. The key feature of smooth modeling, permitting to approximate the likelihood as a sum of GEV likelihoods at the stations, is the simplifying assumption that annual snow depth maxima are approximately independent in space and time. Nevertheless, smooth modeling does not provide any spatial dependence of extremes.

As a way to account for spatial dependence of extremes, *max-stable processes* as an extension of multivariate extreme value theory to infinite dimensions can be used (de Haan 1984). With max-stable processes, the margins and their spatial dependency can be modeled simultaneously but independently. The **SpatialExtremes** package provides functions for statistical modelling of spatial extremes using max-stable processes, copula or Bayesian hierarchical models. In addition, the **hkevp** package provides several procedures around the HKEVP model of Reich and Shaby (2012) and the Latent Variable Model of Davison, Padoan, and

Ribatet (2012). However, no package for easy use of the smooth modeling approach exists.

The **gevreg** package provides a function to fit a smooth spatial regression model to observation. It has a convenient interface to estimate the model with maximum likelihood and provides several methods for analysis and prediction. The outline of the paper is as follows. Section 2 describes the idea of smooth spatial (extreme value) modeling, and Section 3 presents its R implementation. Section 4 illustrates the package functions with Austrian snow depth data and finally Section 5 summarizes the paper.

2. Smooth Spatial Modeling of Extremes

The annual maximum of a variable can be interpreted as a time-space stochastic process $\{S_x^{(t)}\}$, where t denotes the corresponding year and $x \in \mathcal{A}$ the location ,e.g. in Austria. However, we assume that the distribution of $S_x^{(t)}$ does not depend on the time t and therefore for each of the GEV parameters, we consider a linear model, i. e. a model of the form

$$\eta(x) = \alpha_0 + \sum_{k=1}^{m} \alpha_k y_k(x) + f\left(y_{m+1}(x), \ldots, y_n(x)\right) \tag{1}$$

at location x, where η denotes one of the GEV parameters, $y_1, \ldots, y_n$ are the considered covariates as functions of the location, $\alpha_0, \ldots, \alpha_m \in \mathbb{R}$ are the coefficients of the linear part and f is a P-spline with 2 knots, evenly distributed across the spatial domain. For $k = 1, \ldots, K$ the k-th station is given by the location x_k and therefore, we have a realization $s_{x_k}^{(1)}, \ldots, s_{x_k}^{(N)}$ of the random sample $S_{x_k}^{(1)}, \ldots, S_{x_k}^{(N)}$ given as measurements. Note that $S_{x_k} \sim \text{GEV}(\mu_{x_k}, \sigma_{x_k}, \xi_{x_k})$ and $\mu(x_k), \sigma(x_k), \xi(x_k)$ are the GEV parameters given by the linear models in (1). By $\ell_k\left(\mu(x_k), \sigma(x_k), \xi(x_k)\right)$ we denote the log-likelihood function at the k-th station corresponding to (1). With the assumption of spatially independent stations, the log-likelihood function then reads as

$$l = \sum_{k=1}^{K} \ell_k\left(\mu(x_k), \sigma(x_k), \xi(x_k)\right),$$

where l only depends on the coefficients of the linear models for the GEV parameters, cf. (1). This approach was called *smooth modeling* by Blanchet and Lehning (2010).

The advantage of maximizing the sum of the log-likelihood functions at the stations compared to maximizing the log-likelihood function at each station lies in the following fact: A good fit at a single station leading to worse fits at several other stations will be penalized. As a consequence, the stations become intertwined in terms of the fitting. As the smooth model does not provide any spatial dependence, it is generally assumed to be less suited to spatially model extremes, compared to other approaches as fitting a max-stable process.

3. R implementation

4. Example

5. Summary

References

Blanchet J, Lehning M (2010). "Mapping snow depth return levels: smooth spatial modeling versus station interpolation." *Hydrol. Earth Syst. Sci*, **14**(12), 2527–2544. doi:10.5194/hess-14-2527-2010. URL http://www.hydrol-earth-syst-sci.net/14/2527/2010/.

Davison AC, Padoan SA, Ribatet M (2012). "Statistical Modeling of Spatial Extremes." *Statistical Science*, **27**(2), 161–186. doi:10.1214/11-STS376. URL http://dx.doi.org/10.1214/11-STS376.

de Haan L (1984). "A Spectral Representation for Max-stable Processes." *The Annals of Probability*, **12**(4), 1194–1204. ISSN 0091-1798. doi:10.1214/aop/1176993148. URL http://dx.doi.org/10.1214/aop/1176993148http://www.jstor.org/stable/2243357.

Reich BJ, Shaby BA (2012). "A hierarchical max-stable spatial model for extreme precipitation." *Annals of Applied Statistics*, **6**(4), 1430–1451. doi:10.1214/12-AOAS591. URL http://dx.doi.org/10.1214/12-AOAS591.

Rudolf-Miklau F, Sauermoser S (2011). *Handbuch Technischer Lawinenschutz*. Ernst & Sohn. A Wiley Company.

Affiliation:

Harald Schellander
ZAMG - Zentralanstalt für Meteorologie und Geodynamik
6020 Innsbruck, Austria
E-mail: Harald.Schellander@zamg.ac.at