Spatial Statistics 8 (2014) 86–103

Contents lists available at ScienceDirect

Spatial Statistics

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

Tapered composite likelihood for spatial max-stable models

Huiyan Sang [a,*], Marc G. Genton [b]

[a] Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA

[b] CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

ARTICLE INFO

Article history:
Received 6 February 2013
Accepted 18 July 2013
Available online 29 July 2013

Keywords:
Composite likelihood
Generalized extreme-value distribution
Max-stable process
Statistics of extremes
Weighted composite likelihood

ABSTRACT

Spatial extreme value analysis is useful to environmental studies, in which extreme value phenomena are of interest and meaningful spatial patterns can be discerned. Max-stable process models are able to describe such phenomena. This class of models is asymptotically justified to characterize the spatial dependence among extremes. However, likelihood inference is challenging for such models because their corresponding joint likelihood is unavailable and only bivariate or trivariate distributions are known. In this paper, we propose a tapered composite likelihood approach by utilizing lower dimensional marginal likelihoods for inference on parameters of various max-stable process models. We consider a weighting strategy based on a "taper range" to exclude distant pairs or triples. The "optimal taper range" is selected to maximize various measures of the Godambe information associated with the tapered composite likelihood function. This method substantially reduces the computational cost and improves the efficiency over equally weighted composite likelihood estimators. We illustrate its utility with simulation experiments and an analysis of rainfall data in Switzerland.

Published by Elsevier B.V.

1. Introduction

Statistical modeling of extreme values has recently drawn research attention. Many environmental problems involve extreme values observed over space, such as extreme precipitation, heavy snow, windstorms and high tides, to name a few. The primary interest in analyzing such data is to characterize recognizable and meaningful spatial patterns that are useful to understanding, predicting, and managing the risks of extreme environmental phenomena.

* Corresponding author. Tel.: +1 9798453156.
E-mail addresses: huiyan@stat.tamu.edu (H. Sang), marc.genton@kaust.edu.sa (M.G. Genton).

2211-6753/$ – see front matter. Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.spasta.2013.07.003

Recent developments of statistical models for spatial extreme values are mainly based on latent variables, on copulas and on spatial max-stable processes; see the review paper by Davison et al. (2012) and references therein. Among these models, max-stable stochastic processes have emerged as a fundamental class of models that are able to describe spatial extreme value phenomena. This is because they arise as the natural generalization of the univariate generalized extreme value (GEV) distribution in infinite dimensional continuous spaces, providing an asymptotically justified approach to modeling process extremes. Max-stable process models for spatial data were first constructed using the spectral representation of de Haan (1984). There have been several subsequent works on the construction of spatial max-stable process models (see, e.g., Smith, 1990; Schlather, 2002; Kabluchko et al., 2009; Smith and Stephenson, 2009; Davison and Gholamrezaee, 2012) and on their application (see, e.g., Coles, 1993; Buishand et al., 2008; Padoan et al., 2010).

Despite many attractive properties of max-stable process models, both classical and Bayesian inferences encounter difficulties because closed-form expressions of the corresponding joint likelihoods are typically not available except for some trivial cases. Taking advantage of the availability of low-dimensional marginal likelihoods, the composite likelihood method has been introduced for inference on the parameters of max-stable processes. Padoan et al. (2010) were the first to suggest maximum pairwise likelihood estimation for a particular class of max-stable process models, namely Gaussian extreme value processes (Smith, 1990). Genton et al. (2011) studied maximum composite likelihood estimators for the same model based on both pairs and triples. They demonstrated substantial gain in efficiency from $p = 2$ to $p = 3$ sites in $\mathbb{R}^2$ by means of a Monte Carlo simulation study. Pairwise and triplewise composite likelihoods have also been used for inference on other max-stable process models (see, e.g., Blanchet and Davison, 2011; Davison et al., 2012; Huser and Davison, 2013).

Several investigations have considered the choice of weights of the composite likelihoods of time series and spatial data to improve statistical efficiency or to reduce the computational burden associated with large datasets. In the context of Gaussian process models, Bevilacqua et al. (2012) suggested that down-weighting or excluding likelihood contributions from sites that are very far apart leads to efficiency gains over the full composite likelihood. In the context of time series, several works (see, e.g., Joe and Lee, 2009; Davis and Yau, 2011) have shown that including unnecessary pairs can cause some loss of estimation efficiency. Padoan et al. (2010) found in a simulation study that the composite likelihoods constructed only using neighboring sites can reduce the asymptotic variances of the model parameters from a Gaussian extreme value process referred to as the *Smith* model (Smith, 1990).

In this paper, we extend the results in Padoan et al. (2010) to investigate in detail the utility of a tapered composite likelihood approach to make inference on several formulations of max-stable process models, including the *Smith* model, the *Schlather* model (Schlather, 2002) and the *Brown–Resnick* model (Kabluchko et al., 2009). Two flexible and practical tapering strategies are proposed to improve the efficiency of the composite likelihood estimators. One is based on the trace of the estimated covariance matrix associated with parameter estimates, and the other is based on its determinant. The proposed tapering strategy is also useful to reducing the computational cost caused by the combinatorial explosion associated with the use of composite likelihood with large datasets. In particular, we extend our investigation beyond the tapered pairwise composite likelihood for the settings of isotropic max-stable processes to the context of anisotropic max-stable processes and triplewise composite likelihoods. Moreover, we discuss the connection between the choice of weights and the strength of the extremal dependence.

The paper is organized as follows. Section 2 reviews the theory of max-stable processes. The tapered composite likelihood approach is developed in Section 3, while Section 4 illustrates the performance of our method through a number of simulation studies. We conclude with an illustration of spatial extreme value analysis of precipitation data in Switzerland.

2. Max-stable processes

2.1. Spatial max-stable processes

Let $\{\tilde{Z}(\boldsymbol{s})^{(i)}\}$, $\boldsymbol{s} \in \mathcal{S} \subset \mathbb{R}^d$, $i = 1, \ldots, n$, be n independent replicates of a continuous spatial stochastic process, where $\mathcal{S}$ is an index set. A spatial stochastic process, $Z(\boldsymbol{s})$, is *max-stable* if there