Variograms for spatial max-stable random fields

Dan Cooley[1], Philippe Naveau[1,2], and Paul Poncet[3]

[1] Department of Applied Mathematics, University of Colorado at Boulder,USA
[2] Laboratoire des Sciences du Climat et de l'Environnement, IPSL-CNRS
philippe.naveau@colorado.edu
[3] Ecole des Mines de Paris, France

1 Introduction

For the univariate case in extreme value theory, it is well known (e.g. [EKM97]) that parametric models characterize the asymptotic behavior of extremes (e.g. sample maxima, exceedances above a high threshold) as the sample size increases. More precisely, a Generalized Extreme Value (GEV) distribution approximates the distribution of sample maxima and a Generalized Pareto Distribution fits asymptotically exceedances. These results can be extended to stationary observations under some general conditions, e.g. [LLR83]. In a multivariate setting and therefore, for spatial extremes, a general structure of the limiting behavior of component-wise maxima has also been proposed in terms of max-stable processes, e.g. [HR77], see Equation (3) for a definition of such processes. However, an important distinction between the univariate and the multivariate case is that no parametrized model can entirely represent max-stable processes. Hence, statistical inference has mostly focused on special bivariate cases, e.g. logistic or bilogistic model (see Chapter 8 in [Col01]). When dealing with spatial data, there are as many variables as locations and consequently, additional assumptions have to be made in order to work with manageable models. The multivariate framework has rarely been treated from a statistical point of view (e.g. [HT04]), but there has been a growing interest in the analysis of spatial extremes in recent years. For example, [HP05] proposed two specific stationary models for extreme values. These models depend on one parameter that varies as a function of the distance between two points. [DM04] proposed space-time processes for heavy tail distributions by linearly filtering i.i.d. sequences of random fields at each location. [Sch02, Sch03] simulated stationary max-stable random fields and studied the extremal coefficients for such fields. Bayesian or latent processes for modeling extremes in space has been also investigated by several authors. In this case, the spatial structure was modeled by assuming that the extreme value parameters were realizations of a smoothly varying process, typically a Gaussian process

with a spatial dependence [CC99]. In geostatistics, a classical approach called "Gaussian anamorphosis" consists of transforming a spatial field into a Gaussian one (e.g. Chapter 33 in [Wac03]). Indicator functions are then used to obtain information about the tails. Although the approach performs well for some specific cases, it is not based on extreme value theory and therefore, modeling extremes with a Gaussian anamorphosis does not take advantage of the theoretical foundation provided by extreme value theory.

In comparison with all these past developments, our research can be seen as a further step in the direction taken by [Sch03] and [Sch02]. We work with stationary max-stable fields and focus on capturing the spatial structure with extremal coefficients. The novelty is that we propose different estimators of extremal coefficients that are more clearly linked to the field of geostatistics. In contrast with the work of [HP05] in which a special structure was imposed (i.e. a parameter measured the dependence as a function of the distance between locations), our estimator can be used for any max-stable field. It is also worthwhile to emphasize the following two points. First, we will show that our estimators are closely related to a distance introduced in [DR93]. Second, although we do not pursue the Gaussian anomorphosis approach used in geostatistics, one should not forget that the field of geostatistics has a long history of modeling spatial data sets, and its development has been tremendous in terms of environmental, climatological and ecological studies. Hence, finding connections between geostatistics and extreme value theory is of primary interest for improving the analysis of complex fields of extreme values.

In geostatistics, it is classical to define the following second-order statistic

$$\gamma(h) = \frac{1}{2}\mathbb{E}|Z(x+h) - Z(x)|^2 \,, \tag{1}$$

where $\{Z(x)\}$ represents a spatial and stationary process with a well-defined covariance function and $x \in \mathbb{R}^2$. The function $\gamma(.)$ is called the (non-centered) (semi) variogram and it has been extensively used by the geostatistic community, see [Wac03], [CD99], [Ste99] and [Cre93]. With respect to extremes, this definition is not well adapted, because a second order statistic is difficult to interpret inside the framework of extreme value theory. Instead of taking the squared difference in $\gamma(.)$, we will show that working with the absolute difference $|Z(x+h) - Z(x)|$ is more appropriate when dealing with extreme values. The first-order moment of this difference leads to the definition of the madogram

$$\nu(h) = \frac{1}{2}\mathbb{E}|Z(x+h) - Z(x)| \,, \tag{2}$$

where the stationary process $\{Z(x)\}$ with an assumed finite mean represents extreme values at different locations x, say annual maximum precipitation at given weather stations. The basic properties of this first-order variogram have been studied by [Mat87] but he did not explore the connection between madograms and extreme value theory. Besides the basic properties $\nu(0) = 0$, $\nu(h) \geq 0$, $\nu(h) = \nu(-h)$ and $\nu(h+k) \leq \nu(h) + \nu(k)$, he showed that