Discussion of "Statistical Modeling of Spatial Extremes" by A. C. Davison, S. A. Padoan and M. Ribatet

D. Cooley and S. R. Sain

We congratulate the authors for their overview paper discussing modeling techniques for spatial extremes. There is great interest in spatial extreme data in the atmospheric science community, as the data is inherently spatial and it is recognized that extreme weather events often have the largest economic and human impacts. In order to adequately assess the risk of potential future extreme events, there is a need to know how the characteristics of phenomena such as precipitation or temperature could be altered due to climate change.

Because of the high interest level in the atmospheric science and (more broadly) the geoscience communities, it is imperative for the statistics community to develop methodologies which appropriately answer the questions associated with spatial extreme data. Davison, Padoan and Ribatet (2012) provide a comprehensive overview of existing techniques that can serve as a useful starting point for statisticians entering the field. That the paper is written as a case study helps to illustrate the advantages and disadvantages of the various methods. We hope that this Swiss rainfall data will serve as a test set by which future methodologies can be evaluated.

The authors analyze data which are annual maxima. This is natural from the classical extreme value theory point of view whose fundamental result establishes the limiting distribution of $\mathbf{Y} = (\bigvee_{i=1}^{n} X_{1i}, \ldots, \bigvee_{i=1}^{n} X_{Di})^T$ to be in the family of the multivariate max-stable distributions. In practice, modeling vectors of annual maxima seems less than ideal, and it is not clear how much dependence information is lost by discarding the coincident data. Scientists in other disciplines can be uncomfortable with the idea of constructing data vectors of events which most often occur on different days. We are aware that there is current work to extend spatial extremes work to deal with threshold exceedances, and we look forward to that work appearing in the literature.

Davison, Padoan and Ribatet (2012) divide the spatial approaches into three categories: latent variable models, copulas, and max-stable process models. In Section 7 they do a very nice job of detailing the strengths and weaknesses of the three approaches. However, it seems that the article does not make clear enough that the aim of the latent variable approach is fundamentally different than the aim of a copula or max-stable process model. As the authors state in Sections 2.2 and 2.3, current modeling of multivariate (or spatial) extremes requires two tasks: (1) the marginals must be estimated and transformed to something standard (e.g., unit Fréchet) so that (2) the tail dependence in the data can be modeled. The latent variable model is a method for characterizing how the marginal distribution varies over space, that is, task 1. In contrast, both copula models and existing max-stable process models explicitly model the tail dependence in the data once the marginals are known, that is, task 2. We refer to the dependence remaining after the marginals have been accounted for as "residual dependence," as Sang and Gelfand (2010) described the random variables after marginal transformation as "standardized residuals."

Davison, Padoan and Ribatet (2012) are correct to point out (Figure 4) that using a latent variable model is inappropriate for applications where the joint behavior of the random vector is required. However, there are applications which aim only to model the marginal behavior. There is a long history of producing return level maps such as those shown in Figure 3 of the manuscript. For instance, the recent effort to update the precipitation frequency atlases for the US (Bonnin et al., 2004a, 2004b) aimed only to characterize the marginal distribution's tail over the study region. Bonnin et al. (2004a, 2004b) employed a regional frequency analysis (Dalrymple, 1960; Hosking and

D. Cooley is Assistant Professor, Department of Statistics, Colorado State University, Fort Collins, Colorado 80523-1877, USA (e-mail: cooleyd@stat.colostate.edu). Stephan R. Sain is Scientist, Institute for Mathematics Applied to Geosciences, National Center for Atmospheric Research, Boulder, Colorado 80307-3000, USA.

Wallis, 1997), approach which, like the latent variable model approach, aims to borrow strength across sites when estimating marginal parameters. As Davison, Padoan and Ribatet (2012) clearly show, explicitly modeling residual dependence requires considerable effort, and when only the marginal effects need to be described, we feel it can be appropriate to ignore the residual dependence so long as one recognizes the limited scope of the questions that such an analysis can answer.

In situations where the joint behavior of multiple locations must be described, then one must explicitly model the residual dependence. As Davison, Padoan and Ribatet (2012) show, dependence models not specifically designed for extremes may be inadequate to capture tail dependence. However, models such as the extremal copulas or max-stable processes do not easily lend themselves to current atmospheric science applications with hundreds, thousands, or tens-of-thousands of locations. There are obvious avenues to explore toward adapting the pairwise likelihood methods for large spatial data sets, but, to date, pairwise likelihood methods have only been applied to applications similar to the one in Davison, Padoan and Ribatet (2012) with roughly 50 locations. We imagine scaling the methods to the size of current applications will be nontrivial, and perhaps new inference procedures or more computationally-feasible extremes dependence models will need to be developed. Until appropriate extremes techniques are available, people will continue to be tempted to apply high-dimensional models developed to describe nonextreme data (e.g., a Gaussian copula) to model tail dependence.

Most of the spatial extremes work to date has been primarily descriptive in nature. Such analyses are useful in assessing risk (i.e., the probability of an extreme event), but do not help to explain the underlying causes of extreme events. There is a desire in the atmospheric sciences to move beyond descriptive analyses and toward analyses which enhance understanding of the processes which lead to extreme events. For example, Sillmann et al. (2011) establish a link between extreme cold temperatures in Europe and a blocking phenomenon in the North Atlantic, Maraun, Osborn and Rust (2011) link extreme precipitation in Europe to large-scale airflow covariates, and Weller, Cooley and Sain (2012) link extreme precipitation on the Pacific coast of North America to surface pressure patterns.

Since it is generally believed that climate models are better at representing processes at large-scales, establishing links between extreme events and large-scale phenomena enable one to better conjecture how the nature of extreme events will change with the climate. While none of the analyses cited above involved extensive spatial modeling of extremes, it is foreseeable that science will move in this direction.

Finally, undertaking a pairwise likelihood fitting of a max-stable process model is challenging and would be beyond the capabilities of most geoscientists. The authors are to be commended for developing the SpatialExtremes (Ribatet, 2011) package in R which enables the general scientific community to utilize these methods.

REFERENCES

BONNIN, G., TODD, D., LIN, B., PARZYBOK, T., YEKTA, M. and RILEY, D. (2004a). *NOAA Atlas* 14, *Precipitation Frequency Atlas of the United States* **1**. U.S. Dept. Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, MD.

BONNIN, G., TODD, D., LIN, B., PARZYBOK, T., YEKTA, M. and RILEY, D. (2004b). *NOAA Atlas* 14, *Precipitation Frequency Atlas of the United States* **2**. U.S. Dept. Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, MD.

DALRYMPLE, T. (1960). Flood frequency analyses. Water Supply Paper 1543-a, U.S. Geological Survey, Reston, VA.

DAVISON, A., PADOAN, S. and RIBATET, M. (2012). Statistical modeling of spatial extremes. *Statist. Sci.* **27** 161–186.

HOSKING, J. R. M. and WALLIS, J. R. (1997). *Regional Frequency Analysis*: *An Approach Based on L-Moments*. Cambridge Univ. Press, Cambridge.

MARAUN, D., OSBORN, T. and RUST, H. (2011). The influence of synoptic airflow on UK daily precipitation extremes. Part I: Observed spatio-temporal relationships. *Climate Dynamics* **36** 261–275.

RIBATET, M. (2011). SpatialExtremes: Modelling spatial extremes. R package version 1.8-1.

SANG, H. and GELFAND, A. E. (2010). Continuous spatial process models for spatial extreme values. *J. Agric. Biol. Environ. Stat.* **15** 49–65. MR2755384

SILLMANN, J., CROCI-MASPOLI, M., KALLACHE, M. and KATZ, R. W. (2011). Extreme cold winter temperatures in Europe under the influence of North Atlantic atmospheric blocking. *Journal of Climate* **24** 5899–5913.

WELLER, G., COOLEY, D. and SAIN, S. (2012). An investigation of the pineapple express phenomenon via bivariate extreme value theory. *Environmetrics*. To appear. DOI:10.1002/env.2143.