Statistical Science
2012, Vol. 27, No. 2, 199–201
DOI: 10.1214/12-STS376REJ
Main article DOI: 10.1214/11-STS376
© Institute of Mathematical Statistics, 2012

Rejoinder

A. C. Davison, S. A. Padoan and M. Ribatet

1. INTRODUCTION

We are grateful to the discussants for their positive and interesting comments. In an area moving so rapidly it is to be expected that our review overlooks some work, and all the contributions helpfully supplement our paper. Our remarks focus on points of possible disagreement or where expansion seems useful.

2. COOLEY AND SAIN

Cooley and Sain bring to the discussion wide experience of statistical applications in atmospheric science, in addition to innovative methodological work. We entirely agree with them that the analysis of annual or seasonal maxima is often unsatisfactory from the statistical point of view: it fails to make full use of the available data, which typically comprise numerous simultaneous time series, and by reducing daily or even hourly data to annual maxima does not allow detailed modeling of the underlying process. In some cases it is useful to follow Stephenson and Tawn (2005) and to incorporate information on the occurrence times of annual maxima; Davison and Gholamrezaee (2012) show that this is quite feasible in the present context, and find some improvement in precision of estimation from doing so. There is a close relationship between models for annual maxima, as considered in our paper, and those for peaks over thresholds (Smith, 1989; Davison and Smith, 1990), and max-stable models of both types share the deficiencies mentioned at the end of Section 8 of our paper. Huser and Davison (2012) extend the ideas used for annual maxima in the present paper to a space-time treatment of extreme hourly rainfall data using the threshold approach. The use of pairwise likelihood poses some tricky issues in that context, however, because of the multiplicity of pairs, which can correspond to simultaneous events in different time series, events at different times in a single series, or at different times in different series. The application considered by Huser and Davison (2012) involves 10 hourly rainfall time series for 27 summers, around 580,000 observations giving 7 billion possible pairs, of which a subset of only around 30 million were used! Although heavy computational burdens arise also in other spatial modeling contexts, better approaches are clearly needed to deal with larger settings for spatial extremes, as Cooley and Sain remark. As an aside, the choice of subsets of observations that contribute to the composite likelihood can be more subtle than at first appears: Huser and Davison (2012) find that although one might think it best to include only strongly-dependent pairs, it can be preferable to include some for which observations are independent or nearly so, in order to get reasonable estimates of the ranges of extremal phenomena.

We entirely agree that the goals of analysis may differ, and that it may not be worthwhile to fit a spatial (or space-time) extremal model when a map of quantiles is the intended output. However, naive use of a latent variable model that ignores the correlations between the events may provide uncertainty measures that are overly precise, as pointed out in the discussion contribution by Gabda et al. Thus, building some form of spatial dependence between events, and not merely between model parameters, seems wise. A pragmatic way to do this may be the use of a Gaussian copula, as in Sang and Gelfand (2010).

Cooley and Sain's final comment concerns a crucial part of extremal modeling, namely, the incorporation of subject-matter knowledge. While the generalized extreme-value distribution, max-stable process and the like rest on elegant and mathematically compelling theory, the real world is a messy place to which the relevance of that theory may be unclear. Very often extremal data show much greater variation than a simplistic view of the theory might suggest, perhaps

Anthony Davison is Professor, Chair of Statistics, Institute of Mathematics, EPFL-FSB-IMA-STAT, Station 8, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (e-mail: Anthony.Davison@epfl.ch). Simone Padoan is a Senior Assistant Researcher, Department of Statistical Science, University of Padua, Via Cesare Battisti 241, 35121 Padova, Italy (e-mail: Simone.Padoan@stat.unipd.it). Mathieu Ribatet is a Maître de conference, I3M, UMR CNRS 5149, Universite Montpellier II, 4 place Eugene Bataillon, 34095 Montpellier, cedex 5, France (e-mail: mathieu.ribatet@math.univ-montp2.fr).