Journal of Statistical Planning and Inference 174 (2016) 11–19

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

A goodness-of-fit test for heavy tailed distributions with unknown parameters and its application to simulated precipitation extremes in the Euro-Mediterranean region

G. Jogesh Babu [a], Andrea Toreti [b,*]

[a] The Pennsylvania State University, United States

[b] European Commission, Joint Research Centre, Italy

ARTICLE INFO

Article history:
Received 9 January 2015
Received in revised form 16 October 2015
Accepted 3 February 2016
Available online 10 February 2016

Keywords:
Anderson–Darling
Extremes
Goodness-of-fit
Bootstrap

ABSTRACT

We establish a general bootstrap procedure combined with a modified Anderson–Darling statistic. This procedure is proved to be valid for heavy tailed generalized Pareto distributions that are commonly used to model excesses over a high threshold in extreme value theory. Then, the method is applied to daily precipitation excesses simulated over the Euro-Mediterranean region in autumn by four regional climate models from the EURO-CORDEX initiative.

1. Introduction

In many fields, e.g. climate sciences, there is an increasing need of modeling extreme values. The natural statistical framework to perform such task is the extreme value theory—EVT (de Haan and Ferreira, 2006; Reiss and Thomas, 2007) that is mainly based on the Fisher–Tippett theorem. Under some regularity conditions, this theorem states that the distribution of the maximum of m i.i.d. random variables converges to a distribution belonging to a specific parametric family: the generalized extreme value (GEV). Based on this result, a similar limiting theorem for excesses over a high threshold holds. In this case, under general regularity conditions, Balkema, de Haan and Pickands (Balkema and de Haan, 1974; Pickands, 1975) established that the limiting distribution belongs to the generalized Pareto (hereafter GP) family composed of three sub-families of distributions: Pareto, Exponential, Beta. A generic distribution belonging to the GP family, can be written as:

$$G_{\sigma,\xi}(x) = \begin{cases} 1-\left(1+\dfrac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}} & \xi \neq 0 \\ 1-\exp\left(-\dfrac{x}{\sigma}\right) & \xi = 0 \end{cases} \tag{1}$$

for $\sigma > 0$ and for $x > 0$ when $\xi \geq 0$ and $x \leq -\frac{\sigma}{\xi}$ when $\xi < 0$. Several methods have been developed and proposed to estimate the two parameters controlling the GP distribution, e.g.: maximum likelihood (Smith, 1985), generalized probability weighted moments (Diebolt et al., 2007). Nevertheless, the inference with small samples (especially of ξ) remains difficult as well as testing the convergence condition on which the model relies. Thus, assessing the goodness-of-fit of such

* Corresponding author.
E-mail address: andrea.toreti@jrc.ec.europa.eu (A. Toreti).

http://dx.doi.org/10.1016/j.jspi.2016.02.002

a model in applications to real data can be important. To address this issue, Choulakian and Stephens (2001) proposed tests based on the Cramér–von Mises and the Anderson–Darling statistics both for known and unknown parameters of the GP distribution. However, the former gives equal weight to all observations while the latter gives more weight to both tails. Therefore, when the interest is on heavy tailed distributions (i.e., GP with $\xi > 0$), a modification is needed. With this respect, a modified Anderson–Darling statistic (hereafter MADA) was proposed by Ahmad et al. (1988):

$$A_n = n \int_{-\infty}^{\infty} [F(x) - E_n(x)]^2 \cdot [1 - F(x)]^{-1} dx \tag{2}$$

where n denotes the sample size, F is the theoretical distribution and E_n is the empirical distribution function. However, when the parameters of F are not known and estimated, the asymptotic distribution of A_n (and the critical values for the goodness-of-fit test) is unknown too.

In this paper, we establish a valid general bootstrap procedure for goodness of fit for modified Anderson–Darling statistic under some general conditions on hazard function. The method is also valid for the heavy tailed GP family, as applied in previous studies (Toreti et al., 2013). Then, we apply the test to characterize daily precipitation extremes in autumn over the Euro-Mediterranean region simulated by a set of (recently released) regional climate models in the frame of the EURO-CORDEX initiative (Jacob et al., 2014). The achievement of a better understanding and characterization of precipitation extremes is very important due to the high impacts of these events on human and natural systems (IPCC, 2012), and this is especially true in a climate change context. Furthermore, a potential increase of vulnerability and exposure to climate extremes further enhances this importance. Concerning the Euro-Mediterranean region, its complexity in terms of topography, atmospheric processes, etc. (Lionello et al., 2012) is well reflected in the estimated and observed climate extremes over the region (Ulbrich et al., 2012; Toreti et al., 2010).

In the following section we establish a valid bootstrap procedure for goodness of fit for modified Anderson–Darling statistic under some general conditions on hazard function. The third section is focused on a simulation study, while the fourth one is devoted to the climate analysis and the last one on conclusions.

2. The bootstrap approach

The procedure (and the associated proof) to be combined with MADA builds on the work of Babu and Rao (2004). Let $\mathcal{F} = \{F(\cdot;\theta), \theta \in \Theta\}$ be a family of continuous distribution functions with Θ being an open region in a p-dimensional Euclidean space. For instance, the family of GP distributions with positive shape parameter, $\theta = (\sigma, \xi)$ and $\Theta = (0, \infty) \times (0, \infty)$. Then, let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables from a distribution F. The aim is to test $F = F(\cdot;\theta)$ for some $\theta = \theta_0 \in \Theta$ by using the MADA statistics, which is based on the empirical processes $Y_n(x;\theta) = \sqrt{n}\,[F(x) - E_n(x)]$. As soon as an estimator of θ is available (i.e., $\hat{\theta}_n$), n i.i.d. samples $X_1^\star, X_2^\star, \ldots, X_n^\star$ can be generated according to $F(\cdot;\hat{\theta}_n)$. Then, the same estimator of the first step can be used to get $\hat{\theta}_n^\star$ from $X_1^\star, X_2^\star, \ldots, X_n^\star$. Thus, this approach can be applied to obtain the critical levels of the statistic if we show that (under some specific conditions) $\int_{-\infty}^{\infty} Y_n^2(x;\hat{\theta}_n^\star)[1 - F(x;\hat{\theta}_n^\star)]^{-1}\, dF(x;\hat{\theta}_n^\star)$ with $Y_n(x;\hat{\theta}_n^\star) = \sqrt{n}\,[F(x;\hat{\theta}_n^\star) - E_n^\star(x)]$ converges for almost all sample sequences to the same limiting distribution of $\int_{-\infty}^{\infty} Y_n^2(x;\hat{\theta}_n)[1 - F(x;\hat{\theta}_n)]^{-1}\, dF(x;\hat{\theta}_n)$ with $Y_n(x;\hat{\theta}_n) = \sqrt{n}\,[F(x;\hat{\theta}_n) - E_n(x)]$.

To achieve this objective we need some technical results and the assumptions listed in the Appendix. Given $\theta_0 \in \Theta$ and $\Lambda \subset \Theta$ the closure of a given neighborhood of θ_0, suppose $\{\theta_n\}$ is a sequence in Λ converging to θ_0 as $n \to \infty$. Let $X_{1,n}, \ldots, X_{n,n}$ be i.i.d. random variables from the distribution $F(\cdot;\theta_n)$. Let $\mathbb{P}_{\theta_n}$ denote the probability measure induced by $X_{1,n}, \ldots, X_{n,n}$ and let E_n denote the empirical distribution of these random variables. Suppose $\hat{\theta}_n$ is an estimator of θ_n, we can just start by stating the following theorem of Babu and Rao (2004). See Appendix for assumptions.

Theorem 2.1 *(Theorem 4.1, in Babu and Rao, 2004). Suppose $\theta_n \to \theta_0$, assumption (A1) holds, and*

$$\hat{\theta}_n - \theta_n = \frac{1}{n}\sum_{i=1}^{n} \ell(X_{i,n};\theta_n) + \frac{1}{\sqrt{n}}\epsilon_n, \tag{3}$$

for a score function ℓ satisfying the assumptions (A2)–(A5), where $\epsilon_n \to 0$ in P_{θ_n}-probability. If $L(\theta_n) \to L(\theta_0)$, then the process Y_n given by

$$Y_n(x;\hat{\theta}_n) = \sqrt{n}\big(E_n(x) - F(x;\hat{\theta}_n)\big)$$

converges weakly to a centered, $\mathbb{E}\{Y(x)\} = 0$, Gaussian process Y, where $L(\theta)$ is defined in the Appendix (see A3).

From this theorem and assuming conditions (E) and (P) of Appendix to be valid, it follows that for almost all sample sequences the processes $Y(\cdot, \hat{\theta}_n^*)$ and $Y(\cdot, \hat{\theta}_n)$ converge weakly to the same limiting centered Gaussian process Y. Now, let $\lambda(\cdot;\theta)$ denote the hazard function of $F(\cdot;\theta)$ i.e.,

$$\lambda(x;\theta) = \frac{f(x;\theta)}{1 - F(x;\theta)},$$

where $f(\cdot;\theta)$ denotes the density function of $F(\cdot;\theta)$.

We need the following lemmas, whose proofs are given in the Appendix.

Lemma 2.1. *Suppose*

$$\sup_x \left|\left(\lambda(x;\hat{\theta}_n)/\lambda(x;\theta_n)\right) - 1\right| \to 0 \quad \text{in probability} \tag{4}$$

and $\sqrt{n}\,\|\hat{\theta}_n - \theta_n\|$ *is bounded in probability. Let* $g(x;\theta) = \nabla_\theta F(x,\theta)$. *Suppose that for some* $\epsilon > 0$ *and for all* $x > 0$,

$$\|g(x,\theta)\|^2\,\lambda(x,\theta') = O(x^{-1-\epsilon}) \tag{5}$$

holds uniformly for all θ *in its domain and for* θ' *such that* $\|\theta - \theta'\| < \delta$ *(for a fixed* $\delta > 0$*).*

Then we have $\hat{I}_n(A) = O_p(A^{-\epsilon})$, *and* $I_n(A) = O_p(A^{-\epsilon})$ *for all sufficiently large* $A > 0$, *where*

$$\hat{I}_n(A) = \int_A^\infty n\frac{[F(x;\theta_n) - F(x;\hat{\theta}_n)]^2}{1 - F(x;\hat{\theta}_n)}dF(x;\hat{\theta}_n) \tag{6}$$

$$I_n(A) = \int_A^\infty n\frac{[F(x;\theta_n) - F(x;\hat{\theta}_n)]^2}{1 - F(x;\theta_n)}dF(x;\theta_n). \tag{7}$$

Lemma 2.2. *Let* $X_{1n}, X_{2n}, \ldots, X_{nn}$ *be i.i.d. random variables from a continuous distribution function* F_n. *For any* $\eta > 0$, *there exists* A *large such that for all large* n

$$\mathbb{P}(S_n(A) > \eta) < \eta, \quad \text{where } S_n(A) = \int_A^\infty \frac{n[E_n(x) - F_n(x)]^2}{1 - F_n(x)}\,dF_n(x) \tag{8}$$

and E_n *denotes the empirical distribution function of* $X_{1n}, X_{2n}, \ldots, X_{nn}$.

Lemma 2.3. *Suppose* $\mathcal{F}$ *be the 2-parameter family of generalized Pareto distributions with positive shape parameter with* $\theta = (\sigma, \xi) \in \Theta = (0,\infty) \times (0,\infty)$. *If* $\min(\sigma_n, \hat{\sigma}_n, \xi_n, \hat{\xi}_n) \geq \delta$ *for some* $\delta > 0$, *then (4) and (5) hold.*

We can finally state the theorem with the main result:

Theorem 2.2. *With the same assumptions of Theorem 2.1 and Lemma 2.1 for* $\hat{\theta}_n$ *and* $\hat{\theta}_n^\star$ *and supposing conditions* (E) *and* (P) *of Appendix to be valid, the following two processes have the same limiting distributions for almost all sample sequences*

$$\int_{-\infty}^\infty \frac{Y_n^2(x,\hat{\theta}_n^\star)}{1 - F(x,\hat{\theta}_n^\star)}\,dF(x,\hat{\theta}_n^\star) \quad \text{and} \quad \int_{-\infty}^\infty \frac{Y_n^2(x,\hat{\theta}_n)}{1 - F(x,\hat{\theta}_n)}\,dF(x,\hat{\theta}_n). \tag{9}$$

Proof. It follows directly by applying Theorem 2.1, Lemmas 2.1, 2.2 and noticing that $\sqrt{n}\|\hat{\theta}_n - \hat{\theta}_n^\star\|$ is bounded in probability by (E) and (P) of Appendix.

This result together with Lemma 2.3 implies that the bootstrap approach can be applied with MADA for GP distributions with positive shape parameter.

3. Simulation

In order to provide an assessment of the proposed approach, a set of simulations is performed by using different distributions (i.e., Gamma, lognormal) and mixtures of GP distributions. 10^4 samples with size equal to 150 are simulated for each case study. As shown in Table 1, MADA performs very well when the gamma and the lognormal are used to generate the samples. While, the performance decreases when GP-mixtures are used, especially for values of the shape parameter close to zero (see Table 1). This intrinsic difficulty of distinguishing extreme value distributions when the shape gets closer to zero has been also noticed and pointed out by Toreti and Naveau (2015) and Naveau et al. (2013). Although this simulation study is far from being complete, it provides a good overview of the power of MADA.

4. Precipitation extremes

The previously described approach is here applied to investigate precipitation extremes simulated by recently released regional climate models' runs. Daily precipitation values have been retrieved from four ERA-Interim (Dee et al., 2011) driven

Table 1
Simulation study based on 10^4 samples of size 150. $M(a, b)$ denotes an equiweighted mixture of GP distributions with shape parameter a and b, respectively. Values are expressed in percentage.

Distribution	Power
$Gamma(2, 1)$	92
$Lognorm$	93
$M(-0.4, 0.4)$	79
$M(-0.2, 0.4)$	64
$M(-0.3, 0.2)$	63
$M(-0.1, 0.4)$	59
$M(-0.05, 0.2)$	51

Table 2
List of the EURO-CORDEX models used in this study.

Model	Institution
CLM	*CLM Community with contributions by BTU, DWD, ETHZ, UCD, WEGC*
DMI	*Danish Meteorological Institute*
KNMI	*Royal Netherlands Meteorological Institute*
SMHI	*Swedish Meteorological and Hydrological Institute*

runs of the EURO-CORDEX initiative (Table 2). These runs cover the period 1989–2009 and have a horizontal resolution of about 50 km. Since they are driven by reanalysis, they can be interpreted as a *plausible* representation of what have happened in the past. Their use in the context of precipitation extremes is particularly worthy at the regional European scale, because the availability of observations from weather stations is limited and does not reach an equivalent spatial resolution in many areas of the region. Moreover, gridded data sets based on observations at similar or higher spatial resolution (e.g. E-OBS; Haylock et al., 2008) are affected by several issues with respect to precipitation extremes, especially in areas with a not too high station density.

Daily precipitation extremes are investigated for autumn (September to November). Excesses over a high threshold (here, 90th percentile) are extracted. In EVT, the choice of the threshold represents a trade-off between the need of data for the inference and the need of being in the domain of attraction of an extreme distribution (de Haan and Ferreira, 2006). Few objective approaches have been proposed (Süveges and Davison, 2010; Toreti et al., 2010); however, in most of the studies *a priori* choices, e.g. the 90th percentile, are taken.

In this exercise, the proposed approach (with a 5% test on each grid point) is applied to the aforementioned excesses in combination with a generalized probability weighted moments estimator (Diebolt et al., 2007) to infer the parameters of the GP distribution, see eq. (1). This estimator is based on the moments $\eta_\omega = \mathbb{E}\left\{X\omega(1 - G_{\sigma,\xi}(X))\right\}$, with ω being a continuous function null and with right derivative at 0, and the associated estimator $\hat{\eta}_{\omega,n} = \int_0^\infty W(1 - E_n(x))\,dx$ where n is the number of excesses and W is the primitive of ω. Here, $\omega(x) = x^r$ with $r = 1, 1.5$ is chosen, as two moments are needed to estimate the two GP-parameters. As soon as an estimation of the two GP-parameters is available, return levels can be derived as well:

$$\hat{z}_R = u + \hat{\sigma}\hat{\xi}^{-1}[(R\zeta_u)^{\hat{\xi}} - 1] \tag{10}$$

where u is the threshold, R the return period (here, 5 years) and ζ_u is the Poisson process for the occurrence of an event above the threshold.

Here, the inference is applied to each grid point independently, while advantages could be taken by modeling the spatial (and spatio-temporal) dependence (e.g. Blanchet and Davison, 2011; Davison et al., 2012). However, as reported by Davison et al. (2012) the extension of the available spatial methods to threshold excesses has been only recently explored (Turkman et al., 2010; Huser and Davison, 2014; Thibaud et al., 2013) and the application to large spatial domain (such as the Euro-Mediterranean region) is still challenging.

Figs. 1–3 show, respectively, the estimated parameters ($\hat{\xi}$ and $\hat{\sigma}$) and the associated 5-year return levels $\hat{z}_5$. All four models are characterized by heavy tailed distributions in the majority of grid points, although some limited areas having bounded tails (i.e., $\xi < 0$) can be identified in Fig. 1. Furthermore, all four models generally agree on heavier tails in the southern part of the domain, although there are remarkable spatial differences. For instance, KNMI shows heavier tails over northern Africa (e.g., over the Atlas mountains); while for the CLM run, the Mediterranean basin shows heavier tails and a more spatial homogeneous behavior. Concerning the application of MADA, the run of the DMI model is the only one having a large area, in the southern part of the domain, where the goodness-of-fit test is not passed. As shown in Fig. 4, p-values lower than 0.01 are associated with the test applied to grid points in the southern part of the region. In terms of $\hat{\sigma}$, overall the investigated models' run show a good spatial agreement and some evident differences over the mountain areas, e.g. the Alps and the Pyrenees. Concerning the estimated $\hat{z}_5$, values range from less than 10 mm to more than 60 mm. All four models agree on having the highest values over the Alps, the Gulf of Lion and the Balkans, but the spatial extension of these *hot-spots*

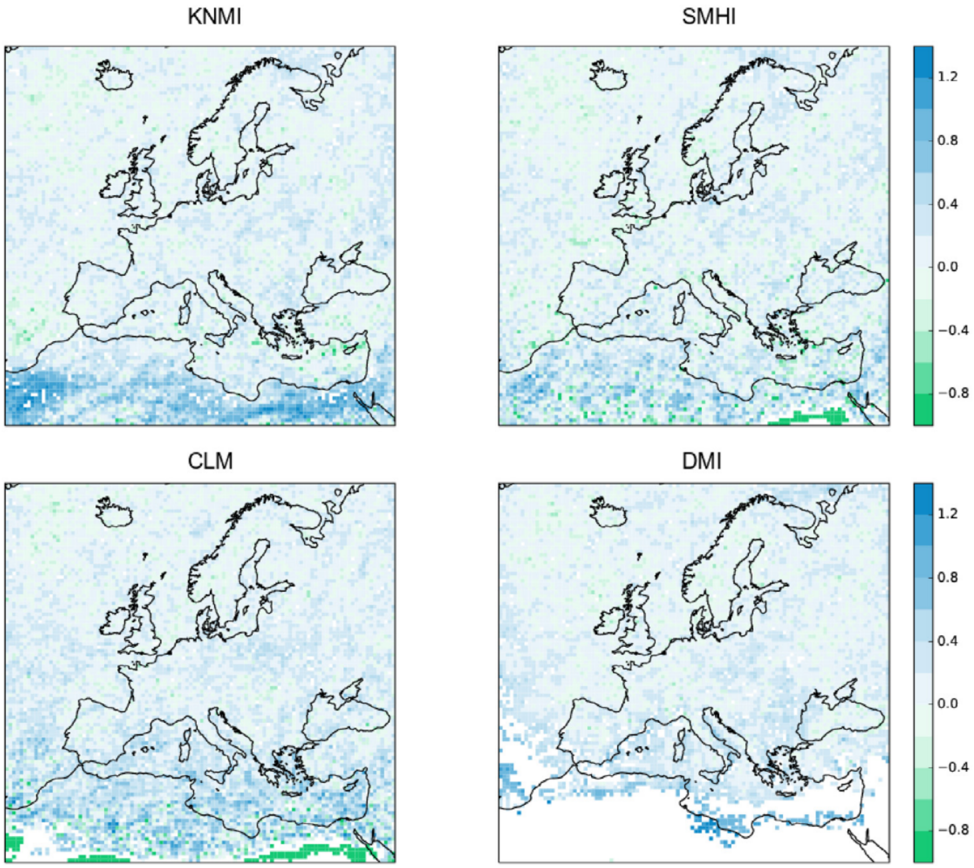

Fig. 1. Estimated shape parameter, $\hat{\xi}$, in autumn for each grid point. White areas are associated with cases where the goodness-of-fit test is not passed.

is not the same. Some interesting differences can be also observed over the Mediterranean coast of Turkey, where the KNMI model run shows higher return levels.

5. Conclusions

Classical models developed in the extreme value theory can be successfully applied to characterize extremes. However, a measure of reliability is often needed. Here, a general approach based on the combination of a modified Anderson–Darling statistic with a bootstrap procedure has been proved to work. This result could be extended to the broad family of ϕ-divergences (Jager and Wellner, 2007), although more efforts are surely needed as such extension does not appear to be straightforward.

The analysis of the simulated precipitation extremes over the Euro-Mediterranean area (in the period 1989–2009) done in the framework of the EURO-CORDEX initiative highlights the applicability and the potentialities of the procedure that can also be used in combination with other estimators. Although the four runs agree on the main spatial pattern, remarkable inter-model spatial differences are evident. Only one model shows a rejection of the goodness-of-fit over a large area, i.e., the southern part of the domain.

Acknowledgments

We acknowledge the World Climate Research Programme's Working Group on Regional Climate, and the Working Group on Coupled Modeling, former coordinating body of CORDEX and responsible panel for CMIP5. We also thank the climate modeling groups (listed in Table 2 of this paper) for producing and making available their model output. We also acknowledge the Earth System Grid Federation infrastructure an international effort led by the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison, the European Network for Earth System Modeling and other partners in the Global Organization for Earth System Science Portals (GO-ESSP). The work of G.J. Babu is supported in part by NSF grant AST-1047586. The analyses have been performed by using R while maps have been produced with the Iris python package. We thank three anonymous referees for constructive suggestions which helped in improving the paper.

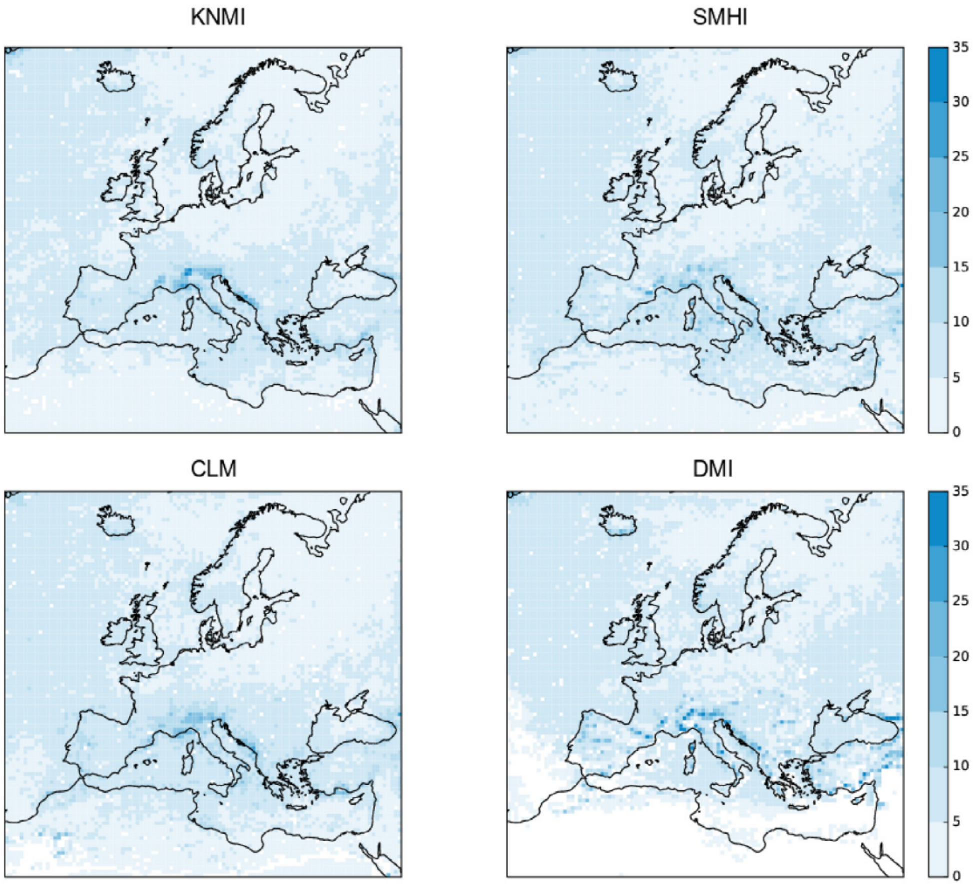

Fig. 2. As Fig. 1 but for $\hat{\sigma}$.

Appendix

Let $\Lambda \subset \Theta$ be the closure of a given neighborhood of a point $\theta_0 \in \Theta$. We use some of the assumptions of Babu and Rao (2004) listed below on the estimators $\hat{\theta}_n$ and $\hat{\theta}_n^*$, where $\ell(.; \theta), \theta \in \Lambda$ is a measurable p-dimensional row vector valued function:

(E) For some $\epsilon_n = \epsilon_n(X_1, \ldots, X_n) \to 0$ in probability,

$$\hat{\theta}_n - \theta_0 = \frac{1}{n} \sum_{i=1}^{n} \ell(X_i; \theta_0) + \frac{1}{\sqrt{n}} \epsilon_n.$$

(P) For some $\epsilon_n^* \to 0$ in probability under the bootstrap measure,

$$\hat{\theta}_n^* - \hat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} \ell(X_i^*; \hat{\theta}_n) + \frac{1}{\sqrt{n}} \epsilon_n^*.$$

We now list an additional set of assumptions on ℓ and F used in the main results.

(A1) The row vector $g(x; \theta) = \nabla_\theta F(x; \theta)$ is uniformly continuous in x and $\theta \in \Lambda$.
(A2) For $\theta \in \Lambda$, $\int \ell(x; \theta)\, dF(x; \theta) = 0$.
(A3) For $\theta \in \Lambda$, $L(\theta) = \int \ell'(x; \theta)\ell(x; \theta)\, dF(x; \theta)$ is a finite non-negative definite matrix.
(A4) As $\gamma \to \infty$,

$$\sup_{\theta \in \Lambda} \int_{\{\|\ell(x;\theta)\| > \gamma\}} \|\ell(x; \theta)\|^2\, dF(x; \theta) \to 0.$$

(A5) For all x, the function $h(x; .)$ defined by

$$h(x; \theta) = \int_{-\infty}^{x} \ell(t; \theta)\, dF(t; \theta)$$

is continuous at θ_0.

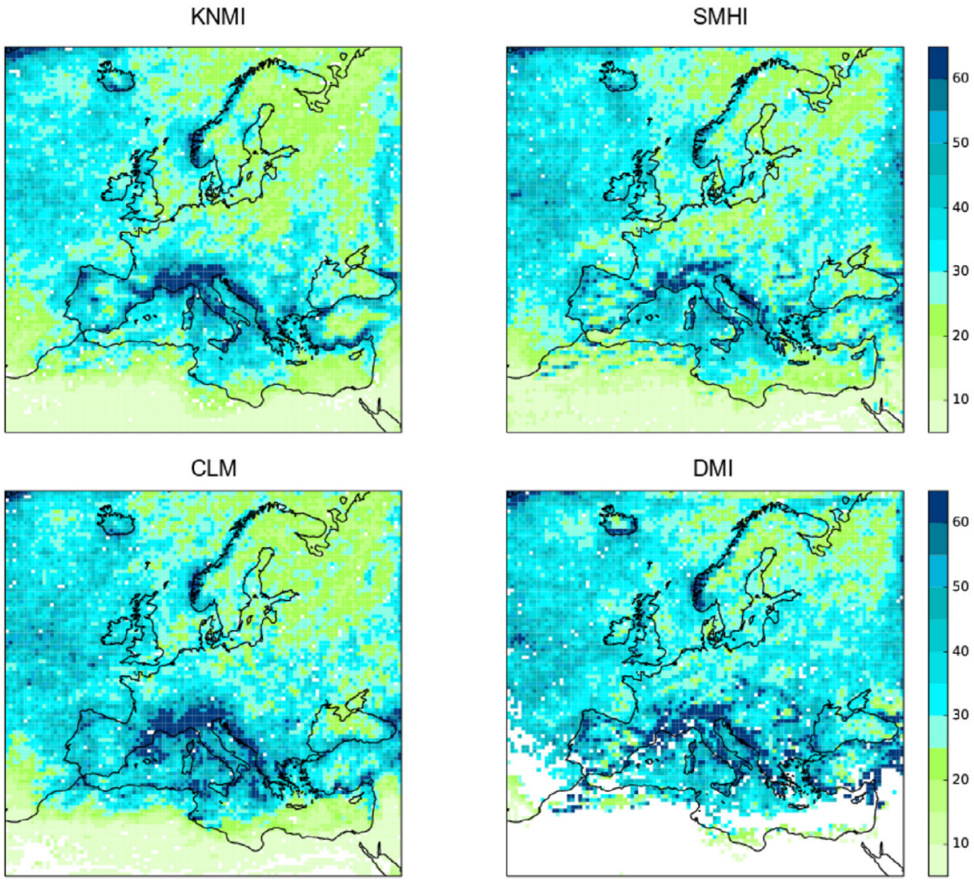

Fig. 3. As Fig. 1 but for 5-year return levels. Values in mm.

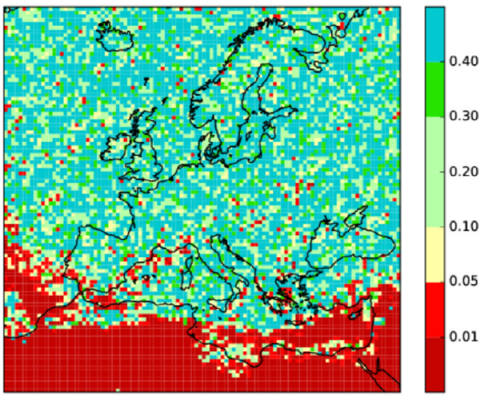

Fig. 4. p-values of the MADA goodness-of-fit test applied to each grid point of the DMI run in autumn.

Proof of Lemma 2.1. By the mean value theorem,

$$I_n(A) \leq \int_A^\infty n \frac{[(\theta_n - \hat{\theta}_n) \cdot g(x, \beta)]^2}{1 - F(x; \theta_n)} f(x; \theta_n)\, dx \quad \text{for some } \beta \text{ lying on the line joining } \theta_n \text{ and } \hat{\theta}_n.$$

So, $I_n(A) = O(n\|\theta_n - \hat{\theta}_n\|^2 \int_A^\infty \|g(x, \beta)\|^2 \lambda(x, \theta_n)\, dx)$ and (5) gives the result for $I_n(A)$. The same holds for $\hat{I}_n(A)$ by just noticing that (4) implies $\hat{I}_n(A) = O(I_n(A))$.

Proof of Lemma 2.2. Let $U_n(t)$ be the empirical process defined by $U_n(t) = n^{-1/2}\left[\sum_{i=1}^n \left(I_{(F_n(X_{in}) \leq t)} - t\right)\right]$, where I denotes the indicator function.

Let $0 < \gamma < 1/4$. Since $\|U_n(t)\|$ is bounded in probability by the Dvoretzky–Kiefer–Wolfowitz inequality (Shorack and Wellner, 1986), the following holds:

$$\int_0^\gamma \frac{U_n(t)^2}{t(1-t)\,dt} \le \|U_n(t)\| \left\| \frac{U_n(t)}{t^{1/4}(1-t)^{1/4}} \right\|_0^\gamma \int_0^\gamma \frac{dt}{t^{3/4}(1-t)^{3/4}}$$
$$= O_p(1)\, O\left(\left\| \frac{U_n(t)}{t^{1/4}(1-t)^{1/4}} \right\|_0^\gamma \right) O(\gamma^{1/4}).$$

Then, by applying the inequality of Corollary 1 in Wellner (1977) (Pyke–Shorack inequality type)

$$\mathbb{P}\left(\left\| \frac{U_n(t)}{t^{1/4}(1-t)^{1/4}} \right\|_0^\gamma > \eta \right) \le \frac{16}{\eta^2} \int_0^\gamma \frac{dt}{\sqrt{t(1-t)}} < \eta \quad \text{if } \gamma^{1/2} < d\,\eta^3, \text{ for some constant } d.$$

This completes the proof by noticing that if distribution V is uniform then $1-V$ is uniform too on $(0, 1)$ and that the following holds

$$\int_{1-\gamma}^1 \frac{U_n(t)^2}{1-t}\, dt \le \int_{1-\gamma}^1 \frac{U_n(t)^2}{t(1-t)}\, dt$$
$$S_n(A) \le \int_{F_n(A)}^1 \frac{U_n(t)^2}{t(1-t)}\, dt.$$

Proof of Lemma 2.3. (4) is valid by noticing that

$$\frac{\lambda(x, \hat{\theta}_n)}{\lambda(x, \theta_n)} = \frac{\sigma_n + \xi_n x}{\hat{\sigma}_n + \hat{\xi}_n x} \to 1 \quad \text{uniformly in } (0, \infty).$$

To prove (5), we note that $\theta = (\sigma, \xi), g = (g_1, g_2) = (\partial F/\partial\sigma, \partial F/\partial\xi)$ and

$$g_1(x, \sigma, \xi) = -f(x; \theta)\frac{x}{\sigma}$$
$$g_2(x, \sigma, \xi) = f(x; \theta)\,\frac{x}{\xi} - f(x; \theta)\,\frac{\sigma + \xi x}{\sigma\xi^2}\, \log\left(1 + \xi\,\frac{x}{\sigma}\right).$$

Thus, $\|g\| = O\big(|x|^{1+\alpha} f(x; \theta)\big)$ for any $\alpha > 0$. Since $f(x; \theta) = O(|x|^{-1-1/\xi})$, (5) holds.

References

Ahmad, M., Sinclair, C., Spurr, B., 1988. Assessment of flood frequency models using empirical distribution function statistics. Water Resour. Res. 24, 1323–1328.

Babu, G., Rao, C., 2004. Goodness-of-fit tests when parameters are estimated. Sankhya 66, 63–74.

Balkema, A., de Haan, L., 1974. Residual life time at great age. Ann. Probab. 2, 792–804.

Blanchet, J., Davison, A., 2011. Spatial modelling of extreme snow depth. Ann. Appl. Stat. 5, 1699–1725.

Choulakian, V., Stephens, M., 2001. Goodness-of-fit tests for the generalized pareto distribution. Technometrics 43, 478–484.

Davison, A., Padoan, S., Ribatet, M., 2012. Statistical modeling of spatial extremes. Statist. Sci. 27, 161–186.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Hólm, E.V., Isaksen, L., Kallberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 137, 553–597.

de Haan, L., Ferreira, A., 2006. Extreme Value Theory: An Introduction. Springer, New York.

Diebolt, J., Guillou, A., Rached, I., 2007. Approximation of the distribution of excesses through a generalized probability-weighted moments method. J. Statist. Plann. Inference 137, 841–857.

Haylock, M., Hofstra, N., Tank, A.K., Klok, E., Jones, P., New, M., 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. J. Geophys. Res. 113, D20119.

Huser, R., Davison, A., 2014. Space–time modelling of exteme events. J. R. Stat. Soc. Ser. B 76, 439–461.

IPCC Field, C.B., barros, V., Stocker, T.F., Qin, D., Dokken, D.J., Ebi, K.L., Mastrandrea, M.D., Mach, K.J., Plattner, G.-K., Allen, S.K., Tignor, M., Midgley, P.M. (Eds.), 2012. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O., Bouwer, L., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J., Teichmann, C., Valentini, R., Vautard, R., Weber, B., Yiou, P., 2014. EURO-CORDEX: new high-resolution climate change projections for european impact research. Reg. Environ. Change 14, 563–578.

Jager, L., Wellner, J., 2007. Goodness-of-fit tests via phi-divergences. Ann. Statist. 35, 2018–2053.

Lionello, P., Abrantes, F., Congedi, L., Dulac, F., Gacic, M., Gomis, D., Goodess, C., Hoff, H., Kutiel, H., Luterbacher, J., Planton, S., Reale, M., Schröder, K., Struglia, M., Toreti, A., Tsimplis, M., Ulbrich, U., Xoplaki, E., 2012. Mediterranean climate—background information. In: The Climate of the Mediterranean Region—From the Past to the Future. Elsevier.

Naveau, P., Guillou, A., Rietsch, T., 2013. A non-parametric entropy-based approach to detect changes in climate extremes. J. R. Stat. Soc. Ser. B 76, 861–864.

Pickands III, J., 1975. Statistical inference using extreme order statistics. Ann. Statist. 3, 119–131.

Reiss, R., Thomas, M., 2007. Statistical Analysis of Extreme Values: With Applications to Insurance, Finance, Hydrology and Other Fields. Birkhäuser-Verlag, Basel.

Shorack, G., Wellner, J., 1986. Empirical Processes with Applications to Statistics. John Wiley & Sons, New York.
Smith, R., 1985. Maximum likelihood estimation in a class of nonregular cases. Biometrika 72, 69–90.
Süveges, M., Davison, A., 2010. Model misspecification in peaks over threshold analysis. Ann. Appl. Stat. 4, 203–221.
Thibaud, E., Mutzner, R., Davison, A., 2013. Threshold modeling of exteme spatial rainfall. Wat. Resour. Res. 49, 4633–4644.
Toreti, A., Naveau, P., 2015. On the evaluation of climate model simulated precipitation extremes. Environ. Res. Lett. 10, 014012.
Toreti, A., Naveau, P., Zampieri, M., Schindler, A., Scoccimarro, E., Xoplaki, E., Dijkstra, H., Gualdi, S., Luterbacher, J., 2013. Projections of global changes in precipitation extremes from coupled model intercomparison project phase 5 models. Geophys. Res. Lett. 40, 4887–4892.
Toreti, A., Xoplaki, E., Maraun, D., Kuglitsch, F., Wanner, H., Luterbacher, J., 2010. Characterisation of extreme winter precipitation in the Mediterranean coastal sites and associated anomalous atmospheric circulation patters. Nat. Hazards Earth. Syst. Sci. 10, 1037–1050.
Turkman, K., Turkman, M., Pereira, J., 2010. Asymptotic models and inference for extremes of spatio-temporal data. Extremes 13, 375–397.
Ulbrich, U., Lionello, P., Belušić, D., Jacobeit, J., Knippertz, P., Kuglitsch, F., Leckebusch, G., Luterbacher, J., Maugeri, M., Maheras, P., Nissen, K., Pavan, V., Pinto, J., Saaroni, H., Seubert, S., Toreti, A., Xoplaki, E., Ziv, B., 2012. Climate of the Mediterranean: synoptic patterns, temperature, precipitation, winds and their extremes. In: The Climate of the Mediterranean Region—From the Past to the Future. Elsevier.
Wellner, J., 1977. A martingale inequality for the empirical process. Ann. Probab. 5, 303–308.