Nonlinear Processes
in Geophysics

Open Access

On the identification of Dragon Kings among extreme-valued outliers

M. Riva[1,2]**, S. P. Neuman**[2]**, and A. Guadagnini**[1,2]

[1]Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Piazza L. Da Vinci 32, 20133 Milan, Italy
[2]Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721, USA

Correspondence to: M. Riva (monica.riva@polimi.it)

Abstract. Extreme values of earth, environmental, ecological, physical, biological, financial and other variables often form outliers to heavy tails of empirical frequency distributions. Quite commonly such tails are approximated by stretched exponential, log-normal or power functions. Recently there has been an interest in distinguishing between extreme-valued outliers that belong to the parent population of most data in a sample and those that do not. The first type, called Gray Swans by Nassim Nicholas Taleb (often confused in the literature with Taleb's totally unknowable Black Swans), is drawn from a known distribution of the tails which can thus be extrapolated beyond the range of sampled values. However, the magnitudes and/or space–time locations of unsampled Gray Swans cannot be foretold. The second type of extreme-valued outliers, termed Dragon Kings by Didier Sornette, may in his view be sometimes predicted based on how other data in the sample behave. This intriguing prospect has recently motivated some authors to propose statistical tests capable of identifying Dragon Kings in a given random sample. Here we apply three such tests to log air permeability data measured on the faces of a Berea sandstone block and to synthetic data generated in a manner statistically consistent with these measurements. We interpret the measurements to be, and generate synthetic data that are, samples from α-stable sub-Gaussian random fields subordinated to truncated fractional Gaussian noise (tfGn). All these data have frequency distributions characterized by power-law tails with extreme-valued outliers about the tail edges.

1 Introduction

The statistics of extremes has been of central interest in many fields (see examples below) including hydrology (Katz et al., 2002) where the subject continues to be the focus of active research (Benson et al., 2007; Clarke, 2007; Bara et al., 2009; El Adlouni and Ouarda, 2009; Laio et al., 2009; Langousis et al., 2009; Veneziano et al., 2009; Ebtehaj and Foufoula-Georgiou, 2010; Fernandes et al., 2010; Katz, 2010; Lee and Ouarda, 2010; Li et al., 2010; Towler et al., 2010; Verhoest et al., 2010; Ailliot et al., 2011; De Michele et al., 2011; Ribereau et al., 2011; Rojas et al., 2011; Zhang et al., 2011; Lee et al., 2012; Viglione et al., 2012). Extreme values commonly cluster around heavy tails of data frequency distributions which are often approximated by stretched exponential, log-normal or power functions. Discrimination between these functional representations is difficult due to the rarity of extreme events (Clauset et al., 2009; Golosovsky and Solomon, 2012) and the ubiquitous presence of runaways, or outliers, among them (e.g. Barnett and Lewis, 1994; Johansen and Sornette, 1998; Sornette, 2003). Recently there has been an interest in distinguishing between extreme-valued outliers that belong to the parent population of most data in a sample and those that do not (Sornette, 2009). The first type, termed Gray Swans by Taleb (2007), is inherently unpredictable (often confused in the literature with Taleb's totally unknowable Black Swans). Eliminating the second type, called Dragon Kings by Sornette (2009), may help minimize ambiguity in defining (and extrapolating) the tails. More intriguing is Sornette's suggestion that Dragon Kings may sometimes be predicted on the basis of precursors lurking in the data (Sornette and Ouillon, 2012). The prospect has motivated some authors to

examine closely, in a special issue of the European Physical Journal, the nature of extreme-valued outliers associated with a wide range of phenomena including rainfall events (Peters et al., 2012); snow avalanches on mountain slopes (Ancey, 2012); rupture events associated with the propagation of cracks or sliding along faults in brittle materials including rock failure, landslides and earthquakes (Amitrano, 2012; Lei, 2012; Main and Naylor, 2012) as well as volcanic eruptions, landslides, wildfires and floods (Sachs et al., 2012; Schoenberg and Patel, 2012; Süveges and Davison, 2012); demographic and financial crises (Akaev et al., 2012; Janczura and Weron, 2012); neuronal avalanches and coherence potentials in the mammalian cerebral cortex (de Arcangelis, 2012; Plenz, 2012); citations of scientific papers (Golosovsky and Solomon, 2012); distributions of city sizes (Pisarenko and Sornette, 2012); and others. The prospect of predicting Dragon Kings has further motivated Janczura and Weron (2012) and Pisarenko and Sornette (2012) to propose statistical tests capable of identifying (not predicting) such outliers in a given data set.

In this paper we apply the latter three tests to log air permeability data measured on the faces of a Berea sandstone block and to synthetic data generated in a manner statistically consistent with these measurements. We focus on identifying Dragon Kings in a given sample, if and after they have occurred, as opposed to predicting the future occurrence of Dragon Kings as proposed by Sornette and Ouillon (2012). We interpret the measurements to be, and generate synthetic data that are, samples of α-stable sub-Gaussian random fields (Samorodnitsky and Taqqu, 1994) subordinated to truncated fractional Gaussian noise (tfGn). All these data have frequency distributions characterized by power-law tails with extreme-valued outliers about the tail edges.

2 Statistical tests to diagnose Dragon Kings

As noted in our introduction, there is no unique way to diagnose Dragon Kings in a sample. Statistical tests proposed in the literature for this purpose have been devised with independent identically distributed (iid) random variables in mind. Below we present versions of these tests tailored to stable distributions of the kind we deal with in this paper.

2.1 Rank-ordering plots

The most common way of identifying Dragon Kings (DKs) has been a visual examination of rank-ordering plots (Sornette, 2009). In these, events are ordered by magnitude from highest to lowest and plotted versus rank-order on semi-log or log-log scale. Highest ranked events that deviate markedly from a model fitted to the majority of extreme values, typically (but not exclusively) a power model, are designated DKs. Rank ordering plots differ from empirical cumulative distributions or survival functions merely by a transposition of their axes.

2.2 Confidence interval test

In the confidence interval test proposed by Janczura and Weron (2012), one considers the empirical cumulative distribution function, ecdf(x), of a random variable X, defined as

$$\text{ecdf}(x) = \frac{1}{n+1} \sum_{i=1}^{n} I_{\{x_i \leq x\}}, \tag{1}$$

where n is the number of observations, x_i, and I is an indicator function. Denoting the cumulative distribution function of X by cdf(x) and making use of the central limit theorem yields the asymptotic ($n \to \infty$) relationship

$$\sqrt{n} \frac{\text{ecdf}(x) - \text{cdf}(x)}{\sqrt{\text{cdf}(x)(1 - \text{cdf}(x))}} \xrightarrow{d} N[0, 1], \tag{2}$$

where $N[0, 1]$ is the standard normal distribution. From Eq. (2) it follows that

$$P\left(z_{\omega/2} < \sqrt{n} \frac{\text{ecdf}(x) - \text{cdf}(x)}{\sqrt{\text{cdf}(x)(1 - \text{cdf}(x))}} < z_{1-\omega/2} \right) = 1 - \omega$$

$$n \to \infty \tag{3}$$

where $z_{\omega/2}$ and $z_{1-\omega/2}$ are the $\omega/2$ and $1 - \omega/2$ quantiles of $N[0, 1]$, respectively. Since $z_{1-\omega/2} = -z_{\omega/2}$ Eq. (3) leads to

$$P\left(\text{cdf}(x) + \sqrt{\frac{\text{cdf}(x)(1 - \text{cdf}(x))}{n}} z_{\omega/2} < \text{ecdf}(x) < \text{cdf}(x) \right.$$

$$\left. - \sqrt{\frac{\text{cdf}(x)(1 - \text{cdf}(x))}{n}} z_{\omega/2} \right) = 1 - \omega \quad n \to \infty. \tag{4}$$

Equation (4) provides, with probability $1 - \omega$, the confidence intervals within which each point of the empirical cumulative distribution function lies for any given cdf model fitted to the data. Janczura and Weron (2012) consider Eq. (4) to be accurate for sample sizes, n, on the order of several hundred.

In this paper, we consider α-stable probability distributions characterized by four parameters: stability index $\alpha \in (0, 2]$, skewness $\beta \in [-1, 1]$, scale $\sigma > 0$ and shift μ. The distribution is defined by its characteristic function which, following Samorodnitsky and Taqqu (1994), we write as

$$\ln\left\langle e^{i\phi X} \right\rangle = i\mu\phi - \sigma^{\alpha} |\phi|^{\alpha} \left[1 + i\beta \text{sign}(\phi) \omega(\phi, \alpha) \right];$$

$$\omega(\phi, \alpha) = \begin{cases} -\tan\frac{\pi\alpha}{2} & \text{if } \alpha \neq 1 \\ \frac{2}{\pi} \ln|\phi| & \text{if } \alpha = 1 \end{cases} \tag{5}$$

where $\langle \rangle$ represents expected value (ensemble mean), ϕ is a real-valued parameter and sign$(\phi) = 1, 0, -1$ if $\phi > 0, = 0, < 0$, respectively. When $\alpha = 2$, the distribution becomes Gaussian. When $\alpha < 2$, one cannot compute confidence intervals explicitly on the basis of Eq. (4). For these cases one can solve Eq. (4) numerically. As we are concerned mainly

with power-law tails of the stable distribution in cases characterized by $\alpha > 1$ and skewness parameter $\beta \neq 1$, we take advantage of its asymptotic power-law limits

$$\lim_{x \to \infty} P(X > x) = \lim_{x \to \infty} 1 - \mathrm{cdf}(x) \to \sigma^{\alpha} C_{\alpha} (1 + \beta) x^{-\alpha}, \quad (6)$$

$$\lim_{x \to -\infty} P(X < x) = \lim_{x \to -\infty} \mathrm{cdf}(x) \to \sigma^{\alpha} C_{\alpha} (1 - \beta) |x|^{-\alpha}, \quad (7)$$

where $C_{\alpha} = \Gamma(\alpha) \sin(\pi\alpha/2)/\pi$, $\Gamma(\alpha)$ being the gamma function. Substituting Eqs. (6) and (7) into Eq. (4) yields

$$P\left(1 + \sqrt{\frac{1 - \sigma^{\alpha} C_{\alpha} (1 + \beta) x^{-\alpha}}{n \sigma^{\alpha} C_{\alpha} (1 + \beta) x^{-\alpha}}} z_{\omega/2} < \frac{1 - \mathrm{ecdf}(x)}{\sigma^{\alpha} C_{\alpha} (1 + \beta) x^{-\alpha}} \right.$$
$$\left. < 1 - \sqrt{\frac{1 - \sigma^{\alpha} C_{\alpha} (1 + \beta) x^{-\alpha}}{n \sigma^{\alpha} C_{\alpha} (1 + \beta) x}} z_{\omega/2} \right) = 1 - \omega \quad n \to \infty \quad (8)$$

$$P\left(1 + \sqrt{\frac{1 - \sigma^{\alpha} C_{\alpha} (1 - \beta) |x|^{-\alpha}}{n \sigma^{\alpha} C_{\alpha} (1 - \beta) |x|^{-\alpha}}} z_{\omega/2} < \frac{\mathrm{ecdf}(x)}{\sigma^{\alpha} C_{\alpha} (1 - \beta) |x|^{-\alpha}} \right.$$
$$\left. < 1 - \sqrt{\frac{1 - \sigma^{\alpha} C_{\alpha} (1 - \beta) |x|^{-\alpha}}{n \sigma^{\alpha} C_{\alpha} (1 - \beta) |x|^{-\alpha}}} z_{\omega/2} \right) = 1 - \omega \quad n \to \infty \quad (9)$$

for the right and left tails of the stable distribution, respectively.

2.3 DK test

For the right tail of a stable distribution, one has, from Eq. (6),

$$\lim_{x \to \infty} \mathrm{cdf}(x) \to 1 - \sigma^{\alpha} C_{\alpha} (1 + \beta) x^{-\alpha} \quad (10)$$

which is the same as a Pareto distribution with threshold, h, equal to $\sigma [C_{\alpha} (1 + \beta)]^{1/\alpha}$. The corresponding DK test (Pisarenko and Sornette, 2012) would aim to verify the null hypothesis that all large values of a sample are generated by this distribution. Suppose that n^* sampled values, x, of X exceed h. Then the normalized random variable $V = \ln(X/h)$ has an exponential distribution

$$\lim_{v \to \infty} \mathrm{cdf}(v) \to 1 - \exp(-\alpha v). \quad (11)$$

One arranges the corresponding sample values, v, of V in decreasing order, $v_1 > v_2 > \ldots > v_{n_*}$, and evaluates corresponding sample values, z, of a random variable Z according to

$$z_k = k(v_k - v_{k+1}) \quad k = 1 \ldots (n^* - 1)$$

$$z_{n*} = k v_{n*}. \quad (12)$$

Since V is exponentially distributed, Z has a χ^2 distribution

with two degrees of freedom. Associate the r ($< n^*$) largest values of z_i with a random variable T defined as

$$T = \frac{\frac{1}{r} \sum_{k=1}^{r} z_k}{\frac{1}{n^*-r} \sum_{k=r+1}^{n^*} z_k}. \quad (13)$$

As the numerator in Eq. (13) is the sum of r random variables, each of which has a χ^2 distribution with two degrees of freedom, the numerator has a χ^2 distribution with $2r$ degrees of freedom. Correspondingly, the denominator has a χ^2 distribution with $2(n^* - r)$ degrees of freedom. It follows that T has an f distribution with $[2r, 2(n^* - r)]$ degrees of freedom. The p value of the null hypothesis is thus

$$p = 1 - F[T, 2r, 2(n^* - r)], \quad (14)$$

where $F[T, 2r, 2(n^* - r)]$ is the cdf of the f distribution. The null hypothesis is rejected (i.e. the DK test diagnoses the existence of r positive DKs) when p is less than a given significance level. Pisarenko and Sornette suggest testing at a significance level of 0.1. Note that each but the largest DK is preceded in rank order by another DK. Pisarenko and Sornette observe that the outcomes of the DK test should not depend on h. They suggest conducting the test for a number of n^* values, starting with $n^* = r + 1$ but not exceeding 20–30, to avoid small values of $(v_k - v_{k+1})$ and large values of z_k. A similar test applies to the left power-law tail of the stable distribution.

2.4 U test

The U test, proposed by Pisarenko and Sornette (2012), makes use of the observation that if a continuous random variable is inserted into its own distribution, then the resulting random variable is uniformly distributed in the interval [0, 1]. Hence if the variable V defined earlier is introduced into Eq. (11), the resulting variable $U = 1 - \exp(-\alpha V)$ is uniformly distributed in the interval [0, 1]. Consequently the m-th order term, U_m, in an ordered set $U_1 < U_2 < \ldots < U_{n^*}$ of U values has a beta distribution with parameters m and $n^* - m + 1$ according to

$$\mathrm{pdf}(u_m) = \frac{1}{B(m, n^* - m + 1)} u_m^{m-1} (1 - u_m)^{n^* - m}, \quad (15)$$

where $B(m, n^* - m + 1)$ is the beta function

$$B(m, n^* - m + 1) = \frac{(m-1)!(n^* - m)!}{n^*!}. \quad (16)$$

The cumulative distribution of U_m is

$$\mathrm{cdf}(u_m) = \frac{B_{u_m}(m, n^* - m + 1)}{B(k, n^* - m + 1)} = I_{u_m}(m, n^*, m + 1), \quad (17)$$

where $B_{u_m}(m, n^* - m + 1)$ is the incomplete beta function,

$$B_{u_m}(m, n^* - m + 1) = \int_0^{u_m} x^{m-1}(1-x)^{n-m}\, dx, \qquad (18)$$

and $I_{u_m}(m, n^* - m + 1)$ is a regularized incomplete beta function. The p value of U_m (i.e. the probability of exceeding u_m) is

$$P(U_m > u_m) = p_{u_m} = 1 - \text{cdf}(u_m) = 1 - I_{u_m}(m, n^* - m + 1). \quad (19)$$

We find it convenient to rephrase Eq. (19) in terms of ranking order k in a sequence $u_1 > u_2 > \ldots u_{n^*}$ according to

$$P(U_k > u_k) = p_{u_k} = 1 - \text{cdf}(u_k) = 1 - I_{u_k}(n^* - k + 1, k), \quad (20)$$

which represents the probability of exceeding the observed value u_k under the null hypothesis. If p_{u_k} is below a given significance level, one concludes at confidence level $1 - p_{u_k}$ that u_k (and thus v_k) is a (positive) DK. Pisarenko and Sornette suggest conducting the test at significance level 0.1. In contrast to the DK test, the U test does not require that a DK be preceded in ranking order by a larger DK.

3   Analysis of synthetic signals

3.1   Stable random fields subordinated to truncated fractional Gaussian noise

Even though all three DK tests in Sect. 2 are theoretically limited to iid random variables, their stated purpose is to identify DKs in real data, which are seldom such. This motivates our application of the tests to Berea sandstone data in Sect. 4 and to synthetic data, statistically consistent with the latter, in this section. We start with the synthetic data because, by providing full control of parameters describing the underlying population, they allow us to examine the DK tests under various conditions, including iid. Whereas the CI test has been applied by its authors to both synthetic and real data, we are not aware of published attempts to apply the U and DK tests to synthetic signals.

Increments of the Berea sandstone data as well as of several other earth variables have been shown by us elsewhere (Guadagnini et al., 2012; Siena et al., 2012; Riva et al., 2013a, b) to be consistent with α-stable random fields subordinated to truncated fractional Gaussian noise (tfGn). We express these fields as (e.g. Neuman et al., 2013)

$$\Delta Y(x, s; \lambda_1, \lambda_u) = W^{1/2}\left[G'(x; \lambda_1, \lambda_u) - G'(x + s; \lambda_1, \lambda_u)\right]$$

$$= W^{1/2}\Delta G(x, s; \lambda_1, \lambda_u), \qquad (21)$$

where x is distance in a given direction, s the lag in the same direction, and W an $\alpha/2$-stable random variable totally skewed to the right of zero with scale parameter $\sigma_W = \left(\cos\frac{\pi\alpha}{4}\right)^{2/\alpha}$, unit skewness $\beta = 1$ and zero shift, independent

of $G'(x; \lambda_1, \lambda_u)$. The latter represents a zero-mean Gaussian random field (or process) that renders ΔY sub-Gaussian (Samorodnitsky and Taqqu, 1994). The field G' possesses a truncated power variogram (TPV; Di Federico and Neuman, 1997)

$$\gamma_i^2(s; \lambda_1, \lambda_u) = \gamma_i^2(s; \lambda_u) - \gamma_i^2(s; \lambda_1), \qquad (22)$$

where $i = 1$ or 2 and, for $m = l$ or u,

$$\gamma_i^2(s; \lambda_m) = \sigma_G^2(\lambda_m)\,\rho_i(s/\lambda_m), \qquad (23)$$

$$\sigma_G^2(\lambda_m) = A\lambda_m^{2H}/2H, \qquad (24)$$

$$\rho_1(s/\lambda_m) = \left[1 - \exp(-s/\lambda_m) + (s/\lambda_m)^{2H}\right.$$
$$\left. \Gamma(1 - 2H, s/\lambda_m)\right] \quad 0 < H < 0.5, \qquad (25)$$

$$\rho_2(s/\lambda_m) = \left[1 - \exp\left(-\pi(s/\lambda_m)^2/4\right) + \left(\pi(s/\lambda_m)^2/4\right)^H\right.$$
$$\left. \Gamma\left(1 - H, \pi(s/\lambda_m)^2/4\right)\right] \quad 0 < H < 1. \qquad (26)$$

A is a constant, H a Hurst scaling exponent, $\Gamma(\cdot, \cdot)$ the incomplete gamma function, λ_1 a lower cut-off proportional to the scale of data measurement or resolution, and λ_u an upper cut-off proportional to the scale of the sampling domain. For $\lambda_u < \infty$, the increments $\Delta Y(s; \lambda_1, \lambda_u)$ are stationary with zero-mean symmetric stable distribution characterized by a lag-dependent scale parameter

$$\sigma^\alpha(s; \lambda_1, \lambda_u) = \left[\gamma_i^2(s; \lambda_1, \lambda_u)\right]^{\alpha/2}. \qquad (27)$$

In the limits $\lambda_1 \to 0$ and $\lambda_u \to \infty$, the TPV $\gamma_i^2(s; \lambda_1, \lambda_u)$ converges to a power variogram (PV) $\gamma_i^2(s) = A_i s^{2H}$ where

$$A_1 = A\Gamma(1 - 2H)/2H$$

and

$$A_2 = A(\pi/4)^{2H/2}\Gamma(1 - 2H/2)/2H.$$

3.2   Generation of synthetic signals

We generate realizations $\Delta y(x, s; \lambda_1, \lambda_u)$ of $\Delta Y(x, s; \lambda_1, \lambda_u)$, subordinated to truncated fractional Gaussian noise $\Delta G(x, s; \lambda_1, \lambda_u)$ according to Eq. (21), using parameters similar to those we infer with this model from Berea sandstone data in Sect. 4. The generating algorithm is described by Riva et al. (2013a). Here we start by creating a realization $\Delta g(x, s; \lambda_1, \lambda_u)$ of $\Delta G(x, s; \lambda_1, \lambda_u)$ at 5×10^4 discrete points on an interval of length $L_u = 5$ along the x axis, spaced a distance $\delta = 10^{-4}$ (measured in arbitrary consistent length units) apart, using a TPV of exponential modes defined in

Eqs. (23)–(25) with $A = 1$, $H = 0.25$, $\lambda_l = 10^{-4}$ and $\lambda_u = 1$ or 10. Varying λ_u in this manner affects the component $\sigma_G^2(\lambda_u)$ of the variance, σ_G^2, of $G'(x; \lambda_l, \lambda_u)$ in Eq. (24). The variance, obtained upon setting $s = 0$ in Eqs. (23)–(24), is given by

$$\sigma_G^2(\lambda_l, \lambda_u) = \frac{A}{2H}\left(\lambda_u^{2H} - \lambda_l^{2H}\right). \tag{28}$$

With the above parameters, $\sigma_G^2 = 1.98$ for $\lambda_u = 1$ and $\sigma_G^2 = 6.30$ for $\lambda_u = 10$. We then multiply each discrete $\Delta g(x, s; \lambda_l, \lambda_u)$ value by a random value, w, of W to yield a discrete signal $\Delta y(x, s; \lambda_l, \lambda_u)$ characterized by stability indices $\alpha = 1.2$ or 1.5. Figure 1 illustrates signals corresponding to four combinations of the parameters λ_u and α, and to $\beta = \mu = 0$, at lag $s = 10^{-2}$. The signals are seen to be highly intermittent, punctuated by numerous (positive and negative) spikes about a near-zero sample mean. The intermittency is qualitatively similar to that exhibited by many earth, environmental, ecological, biological, physical, social and financial variables including several discussed by Sornette (2009) and Janczura and Weron (2012). Since we have generated all signals with a given model, we know that its spikes represent the same population as the rest of the synthetic data and are thus not DKs. We now apply the above statistical tests of diagnosing DKs to these signals with the hope that they would confirm this prior knowledge.

3.3 Diagnostic tests of synthetic signals

Table 1 lists maximum likelihood (ML) estimates $\hat{\boldsymbol{\theta}} = \left(\hat{\alpha}, \hat{\beta}, \hat{\sigma}, \hat{\mu}\right)^T$ of the parameter vector $\boldsymbol{\theta} = (\alpha, \beta, \sigma, \mu)^T$ for 12 signals we have generated with $\beta = \mu = 0$ and all possible combinations of $\lambda_u = 1$ and 10, $\alpha = 1.2$ and 1.5, $s = 10^{-3}$, 10^{-2}, and 10^{-1} including those in Fig. 1. ML estimation was performed with code STABLE of Nolan (2001) by maximizing an equivalent form of the likelihood

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(Y_i | \boldsymbol{\theta}), \tag{29}$$

with respect to $\boldsymbol{\theta}$ where n is sample size and $f(Y|\boldsymbol{\theta})$ the univariate conditional α-stable probability density function (pdf) of Y. Also listed are corresponding 95 % confidence intervals of the estimates, confirming that the univariate parameter estimates are reasonably close to their generating (true) multivariate counterparts. The quality of the ML fits is illustrated graphically for $s = 10^{-2}$ in Figs. 2–5, which plot cumulative distributions of $|\Delta y|$ for $\Delta y < 0$ and survival functions of Δy for $\Delta y > 0$ together with ML estimates (based on all generated values) and confidence intervals (CIs) at three confidence levels. The main part of each figure emphasizes behaviour at the tails, indicating a slight breakdown in power-law distribution of the data near the edges. Will the DK tests in Sect. 2 recognize these outliers to be members of
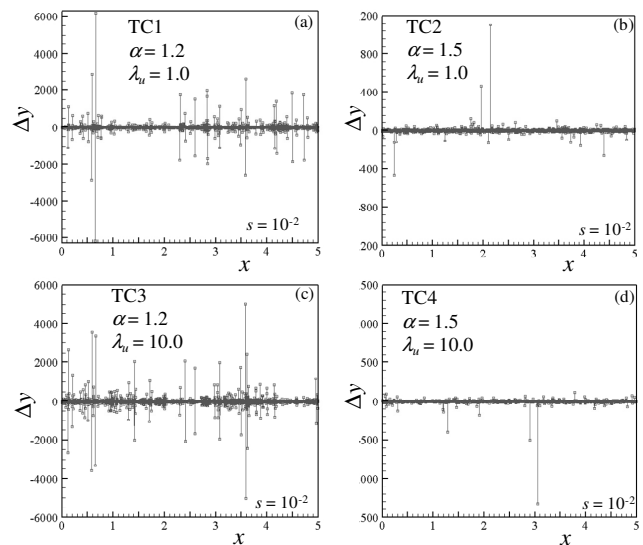

Fig. 1. Increments $\Delta y(x, s; \lambda_l, \lambda_u)$ evaluated at $s = 10^{-2}$ for two values of α and λ_u.

the modelled population even if visually some may appear to be DKs?

Before addressing this question, it is important to clarify that just because a sample is inferred by ML to fit an α-stable distribution, this does not necessarily imply that it might not fit other heavy-tailed distributions equally well or better. Our aim is not to identify the best fitting distribution for a set of data but to verify that (a) the data are consistent with an α-stable distribution having power-law tails and that (b) the DK tests identify the correct nature of outliers falling above or below this tail.

We start addressing the above question by applying the confidence interval test to the 12 test cases listed in Table 1. Table 2 lists the number (N) of points in each case that lie outside 90 %, 95 % and 99 % confidence intervals in Figs. 2–5 and their equivalents (not shown). There is not one case in which a point (or several points) would not fall outside one (or more) of these confidence intervals. Table 2 lists the smallest ($k_{\min}$) and largest ($k_{\max}$) rank orders of all N outliers and the ratio between $k_{\max}$ and sample size in each case. The largest of these ratios being 0.3 % qualifies all outliers as rare large events; when $k_{\min} = 1$, the largest value in the sample forms an outlier.

Next we try the U test with thresholds $h = 10$ and 20, beyond which the fitted stable distributions in all 12 test cases are well represented by their asymptotic power-law tails. We find the outcomes to depend only in a minor way on the choice of h and to agree quite closely with those of the confidence interval test. In way of illustration we plot in Fig. 6a p values for the left tail of the distribution associated with test case TC1 at $s = 10^{-2}$ versus rank order k for both thresholds. Following the suggestion of Pisarenko and Sornette (2012) to test for DKs at a significance level of 0.1, we identify the

Table 1. True and ML-estimated parameters of tfGn signals. True values of β and μ are equal to zero in all test cases.

True values		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$	$\hat{\mu}$
TC1	$s = 10^{-3}: \alpha = 1.20, \sigma = 0.30$	1.19 ± 0.01	0.00 ± 0.02	0.30 ± 0.00	0.00 ± 0.00
	$s = 10^{-2}: \alpha = 1.20, \sigma = 0.56$	1.21 ± 0.01	0.01 ± 0.02	0.59 ± 0.01	0.02 ± 0.01
	$s = 10^{-1}: \alpha = 1.20, \sigma = 0.95$	1.20 ± 0.01	0.00 ± 0.02	1.00 ± 0.01	0.03 ± 0.01
TC2	$s = 10^{-3}: \alpha = 1.50, \sigma = 0.30$	1.48 ± 0.01	-0.01 ± 0.03	0.30 ± 0.00	0.00 ± 0.00
	$s = 10^{-2}: \alpha = 1.50, \sigma = 0.56$	1.51 ± 0.01	-0.01 ± 0.03	0.59 ± 0.01	-0.01 ± 0.01
	$s = 10^{-1}: \alpha = 1.50, \sigma = 0.95$	1.50 ± 0.01	-0.02 ± 0.03	1.01 ± 0.01	-0.02 ± 0.02
TC3	$s = 10^{-3}: \alpha = 1.20, \sigma = 0.30$	1.20 ± 0.01	0.00 ± 0.02	0.30 ± 0.00	0.00 ± 0.00
	$s = 10^{-2}: \alpha = 1.20, \sigma = 0.57$	1.19 ± 0.01	0.02 ± 0.02	0.59 ± 0.01	0.05 ± 0.01
	$s = 10^{-1}: \alpha = 1.20, \sigma = 1.02$	1.21 ± 0.01	0.09 ± 0.02	1.01 ± 0.01	0.33 ± 0.01
TC4	$s = 10^{-3}: \alpha = 1.50, \sigma = 0.30$	1.50 ± 0.01	0.03 ± 0.03	0.30 ± 0.00	0.01 ± 0.00
	$s = 10^{-2}: \alpha = 1.50, \sigma = 0.57$	1.49 ± 0.01	0.03 ± 0.03	0.59 ± 0.01	0.02 ± 0.01
	$s = 10^{-1}: \alpha = 1.50, \sigma = 1.02$	1.52 ± 0.01	0.10 ± 0.03	1.01 ± 0.01	0.16 ± 0.02

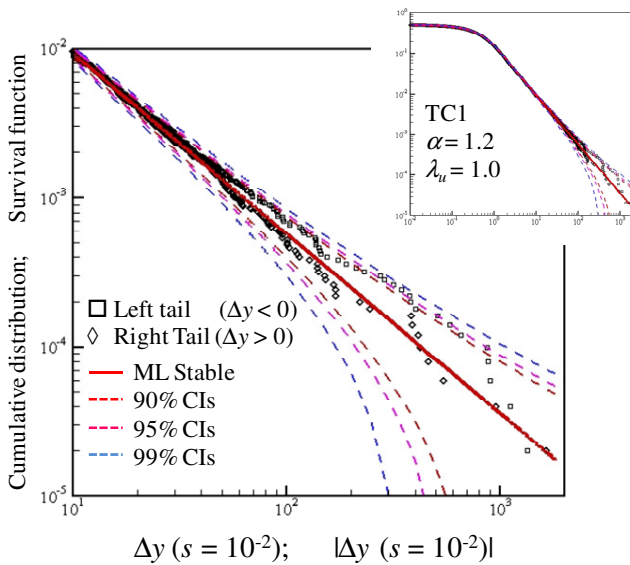**Fig. 2.** Cumulative distribution versus $|\Delta y|$ for $\Delta y < 0$ and survival function versus Δy for $\Delta y > 0$ of TC1 at $s = 10^{-2}$. Main figure emphasizes behaviour at tails. Solid curves represent ML fitted stable distribution; dashed curves indicate confidence intervals, CIs.
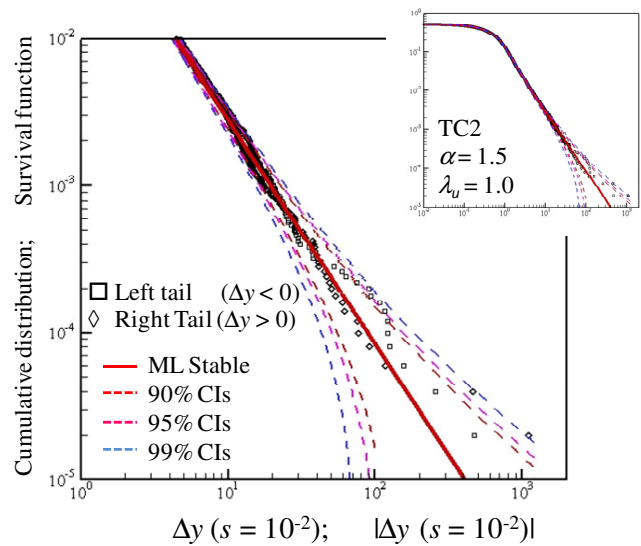

Fig. 3. Cumulative distribution versus $|\Delta y|$ for $\Delta y < 0$ and survival function versus Δy for $\Delta y > 0$ of TC2 at $s = 10^{-2}$. Main figure emphasizes behaviour at tails. Solid curves represent ML fitted stable distribution; dashed curves indicate confidence intervals, CIs.

fifth ($k = 5$) largest value as DK; so did the confidence interval test (Table 2 and Fig. 2). Strangely, the U test suggests the presence of DKs at relatively large k values (> 100). The same test does not identify any DKs on the right tail; nor did the confidence interval test (Fig. 2). A similar plot at $s = 10^{-1}$ in Fig. 6b does not identify the highest ranked ($k = 1$) outlier as a DK even though it was diagnosed as such by the confidence interval test; this outlier would, however, be diagnosed as a DK by the U test at just a slightly higher significance level of 0.15. A plot of p versus k for the right tail of the distribution associated with test case TC2

at $s = 10^{-2}$, in Fig. 7a, identifies the first two ($k = 1$ and 2) highest ranking outliers as DKs; so did the confidence interval test (Table 2 and Fig. 3). A similar plot for the left tail at $s = 10^{-1}$, in Fig. 7b, diagnoses the outlier at $k = 1$ as a DK; the same applies for the confidence interval test (Table 2). Plots of p versus k for the left tail associated with TC4 at $s = 10^{-2}$ and $s = 10^{-1}$, respectively, in Fig. 8a and b, identify the first three ($k = 1-3$) and first ($k = 1$) highest ranking outliers as DKs, in accord with the confidence interval test (Table 2).

Table 2. Confidence interval test. N is number of outliers, k_{min} and k_{max} being rank orders of first and last outlier, respectively. Ratios between k_{max} and sample size are in parentheses.

		CI	N	k_{min}	k_{max}
TC1	$s = 10^{-3}$	90 %	51	5	57 (0.23 %)
		95 %	43	9	55 (0.22 %)
		99 %	23	10	33 (0.13 %)
	$s = 10^{-2}$	90 %	19	5	45 (0.18 %)
		95 %	8	5	15 (0.06 %)
		99 %	0	–	–
	$s = 10^{-1}$	90 %	12	1	30 (0.13 %)
		95 %	5	1	16 (0.07 %)
		99 %	0	–	–
TC2	$s = 10^{-3}$	90 %	8	17	31 (0.12 %)
		95 %	2	18	22 (0.09 %)
		99 %	0	–	–
	$s = 10^{-2}$	90 %	7	1, 2	11 (0.04 %)
		95 %	4	1, 2	10 (0.04 %)
		99 %	1	1	–
	$s = 10^{-1}$	90 %	6	1	27 (0.11 %)
		95 %	1	1	–
		99 %	1	1	–
TC3	$s = 10^{-3}$	90 %	21	5	21 (0.08 %)
		95 %	15	8	21 (0.08 %)
		99 %	1	14	14 (0.06 %)
	$s = 10^{-2}$	90 %	20	3	40 (0.16 %)
		95 %	8	3	11 (0.04 %)
		99 %	1	10	10 (0.04 %)
	$s = 10^{-1}$	90 %	62	2	77 (0.30 %)
		95 %	18	2	77 (0.30 %)
		99 %	0	–	–
TC4	$s = 10^{-3}$	90 %	3	4	8 (0.03 %)
		95 %	0	–	–
	$s = 10^{-2}$	90 %	3	1, 2	3 (0.01 %)
		95 %	3	1, 2	3 (0.01 %)
		99 %	2	1	3 (0.01 %)
	$s = 10^{-1}$	90 %	5	1	22 (0.09 %)
		95 %	1	1	
		99 %	1	1	

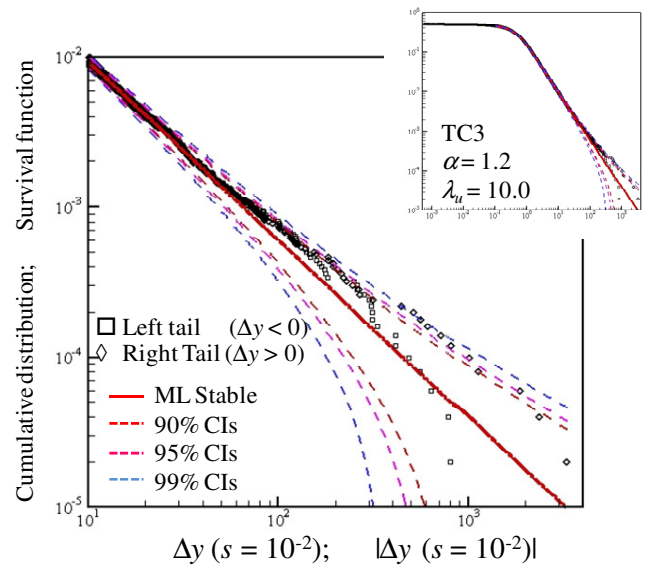**Fig. 4.** Cumulative distribution versus $|\Delta y|$ for $\Delta y < 0$ and survival function versus Δy for $\Delta y > 0$ of TC3 at $s = 10^{-2}$. Main figure emphasizes behaviour at tails. Solid curves represent ML fitted stable distribution; dashed curves indicate confidence intervals, CIs.
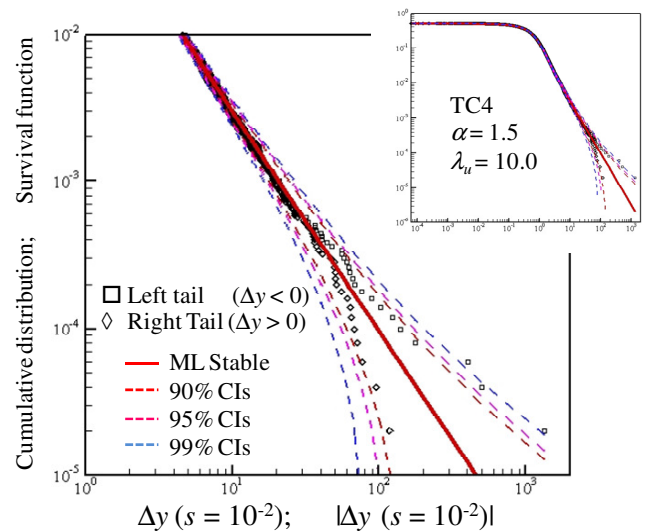

Fig. 5. Cumulative distribution versus $|\Delta y|$ for $\Delta y < 0$ and survival function versus Δy for $\Delta y > 0$ of TC4 at $s = 10^{-2}$. Main figure emphasizes behaviour at tails. Solid curves represent ML fitted stable distribution; dashed curves indicate confidence intervals, CIs.

We end this analysis with an application of the DK test to our 12 synthetic test cases. As this test requires that each DK (except the highest ranked) be preceded (in rank order) by another DK, we apply it only to cases in which the highest ranked outlier was previously diagnosed to be a DK. Since outcomes obtained with $h = 10$ are identical to those obtained with $h = 20$, we present below only the former. In test case TC1 at $s = 10^{-1}$, the highest ranked value on the left tail of the distribution was diagnosed by the confidence interval test as a DK; the U test would have indicated the same at a

significance level of 0.15 but not at 0.1. Accordingly we set $r = 1$ and plot in Fig. 9 p values obtained with the DK test for various numbers n^* of events on the left tail exceeding h. The plot confirms the diagnosis of the U test, the p values being slightly larger than 0.1. In TC2 at $s = 10^{-2}$ the two highest ranked events on the right tail were identified as DKs by both the confidence interval and the U tests. Correspondingly we set $r = 2$ and plot in Fig. 10a p values obtained with the DK
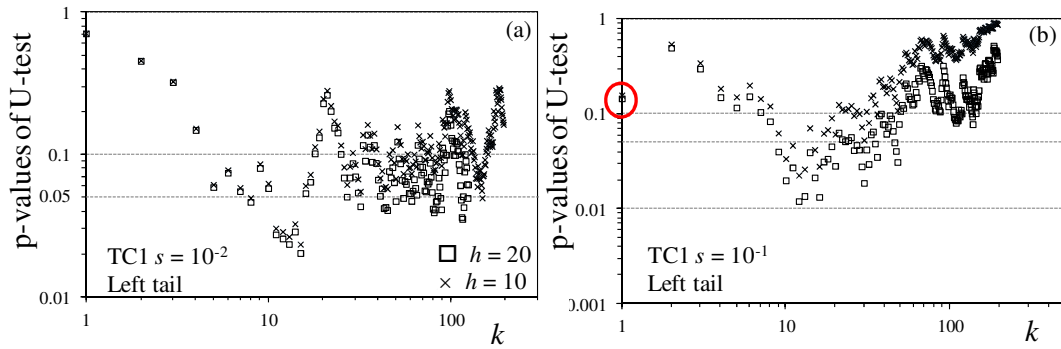
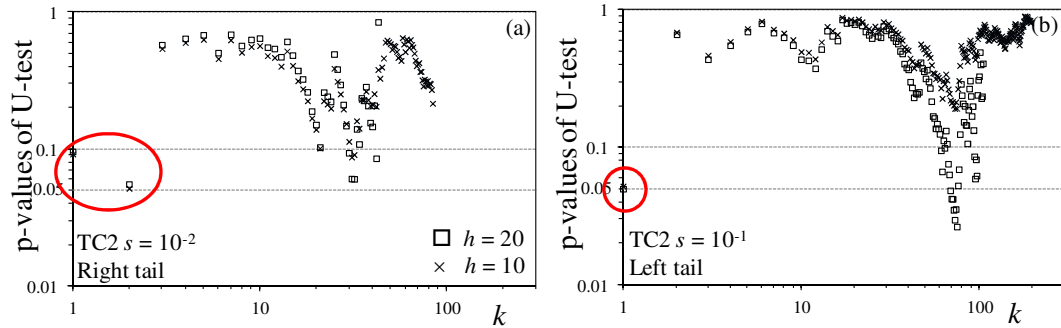**Fig. 6.** U test p values for TC1 at **(a)** $s = 10^{-2}$ and **(b)** $s = 10^{-1}$.

Fig. 7. U test p values for TC2 at **(a)** $s = 10^{-2}$ and **(b)** $s = 10^{-1}$.
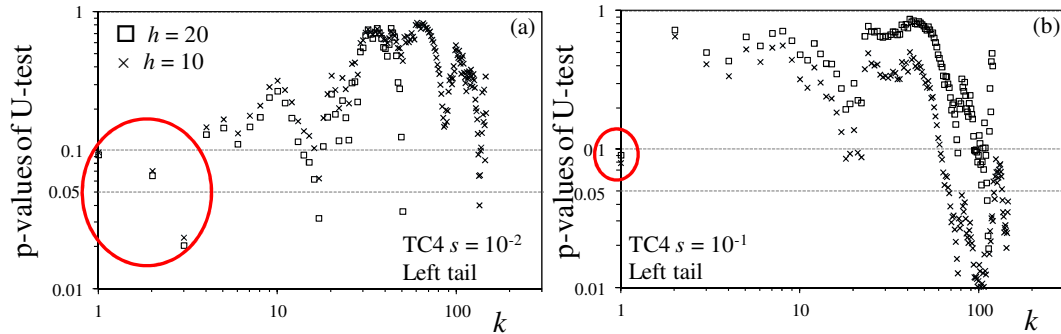

Fig. 8. U test p values for TC4 at **(a)** $s = 10^{-2}$ and **(b)** $s = 10^{-1}$.
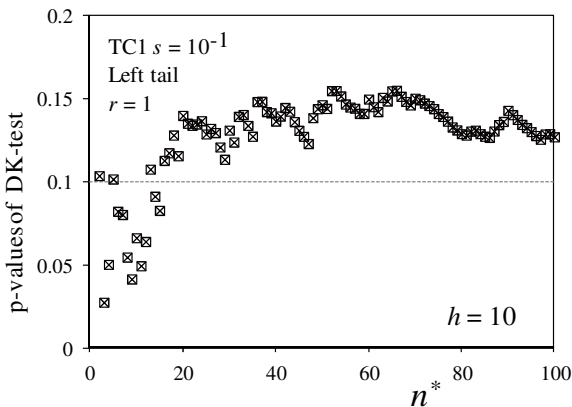

Fig. 9. DK test p values for TC1 and $r = 1$ at $s = 10^{-1}$.

test on the right tail for various numbers n^* of events exceeding h. The plot clearly confirms the previous findings. In TC2 at $s = 10^{-1}$ the highest ranked event on the left tail was identified by both previous tests as a DK, in agreement with DK test results in Fig. 10b. In TC4 at $s = 10^{-2}$ the three highest ranked events on the left tail were identified as DKs by both previous tests; setting $r = 3$ in the DK test yields a plot of p values versus n^* (Fig. 11a) that are generally larger than 0.1, in contrast to these previous tests. In TC4 at $s = 10^{-1}$ the highest ranked value on the left tail of the distribution was diagnosed by the confidence interval and U tests as a DK, in agreement with the DK test results in Fig. 11b.

We conclude that the three statistical tests (confidence interval, U and DK) yield similar outcomes in most but not all cases. All three tests identify Dragon Kings where none exists, casting doubt on their robustness. We found the same to
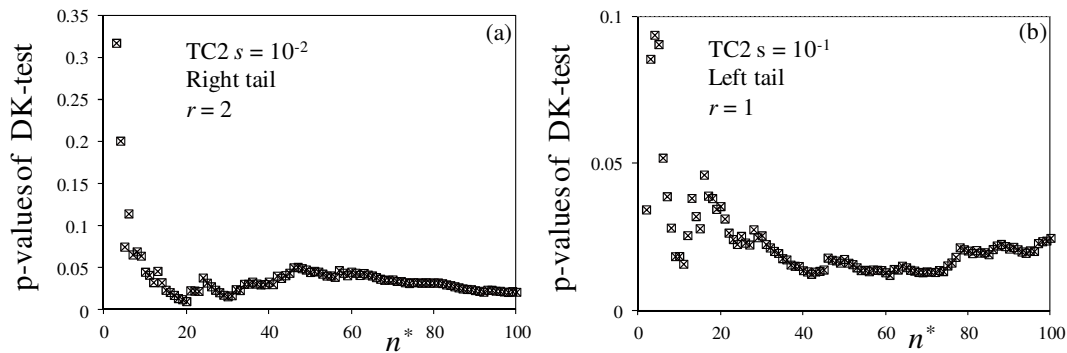
Fig. 10. DK test p values for TC2 and **(a)** $r = 2$ at $s = 10^{-2}$ and **(b)** $r = 1$ at $s = 10^{-1}$.
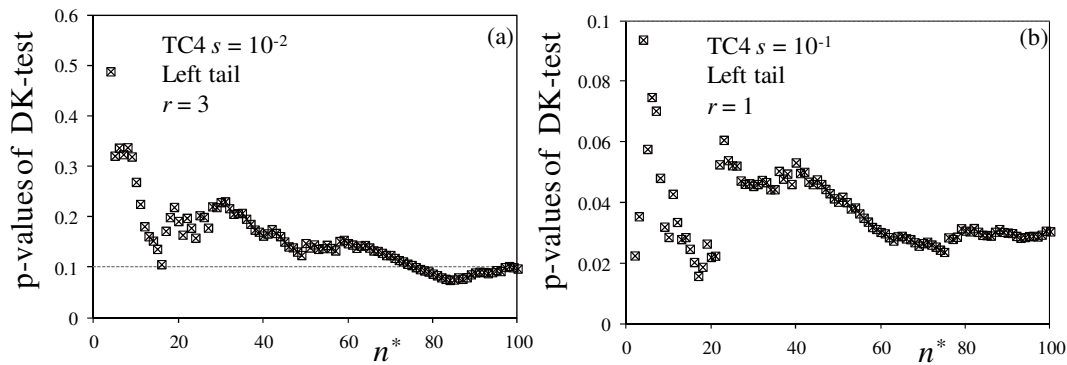

Fig. 11. DK test p values for TC4 and **(a)** $r = 2$ at $s = 10^{-2}$ and **(b)** $r = 1$ at $s = 10^{-1}$.

be true for synthetic samples of iid α-stable random variables characterized by the same values of parameters α, β, σ, and μ (not shown).

4 Analysis of Berea sandstone data

4.1 Statistical analysis of Berea sandstone data

Our air permeability (K) data consist of measurements conducted with a multi-support permeameter (MSP) by Tidwell and Wilson (1997, 1999) on six faces of a Berea sandstone block extending 30 cm on each side. Measurements were taken at intervals of $\Delta = 0.85$ cm on a grid of $24 \times 24 = 576$ points along each face covering a surface of 20×20 cm^2. We (Riva et al., 2013b) analysed natural $\log K$ values measured with the smallest inner radius of the MSP equal to 0.15 cm. Upon fitting a stable distribution to $\log K$ increments parallel to stratification by the maximum likelihood method, we found the shift and skewness parameters to be virtually zero, the stability (Lévy) index α to vary little with separation distance or lag, and scale parameter σ to increase systematically with lag. We concluded that the increments are consistent with α-stable fields subordinated to tfGn, similar to those generated in the previous section.

Figure 12a plots all 1344 available $\log K$ increments, Δy, computed along direction x parallel to stratification on four faces of the block (two parallel and two normal to stratification) at lag $s = 10\Delta$. Figure 12b depicts their sample cumulative distribution together with corresponding maximum likelihood (ML) estimates, obtained by fitting α-stable (Nolan, 2001) and Gaussian models to the corresponding sample frequency distribution. The sample distribution is evidently non-Gaussian, being represented closely by a near-symmetric α-stable distribution with stability index $\alpha = 1.39 \pm 0.08$, skewness $\beta = 0.09 \pm 0.14$, scale $\sigma = 0.26 \pm 0.02$ and shift $\mu = 0.05 \pm 0.02$. Figure 12c and d compare the sample frequency distribution of the increments with ML estimates of α-stable and Gaussian probability density functions in natural and semi-log coordinates, respectively. The semi-log plot accentuates oscillations near the tails of the distribution, which are difficult to discern in natural coordinates.

Figure 13 shows in log-log coordinates how the sample cumulative distribution and its ML equivalent vary with absolute values, $|\Delta y|$, of negative increments $\Delta y < 0$ along the left tail and how the sample survival function and its ML equivalent vary with positive increments $\Delta y > 0$ along the right tail. In both cases the data fall below the straight line, or power law, portion of each tail in a way that resembles negative Dragon Kings as defined by Pisarenko and Sornette (2012). While negative Dragon Kings are not
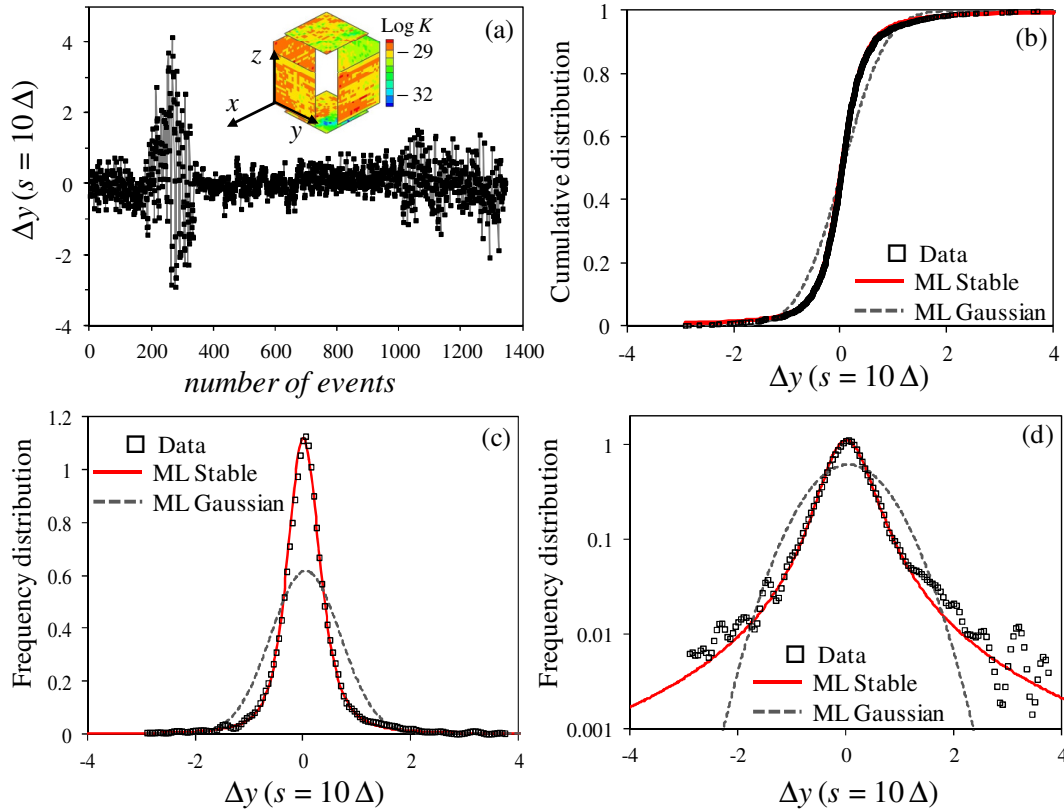
Fig. 12. Log K increments, Δy, along x direction parallel to stratification, computed on Berea sandstone faces at lag $s = 10\Delta$; inset shows log K measured on some block faces **(a)**. Cumulative **(b)** and frequency distributions in natural **(c)** and semi-log **(d)** coordinates compared with stable and Gaussian maximum likelihood fits.
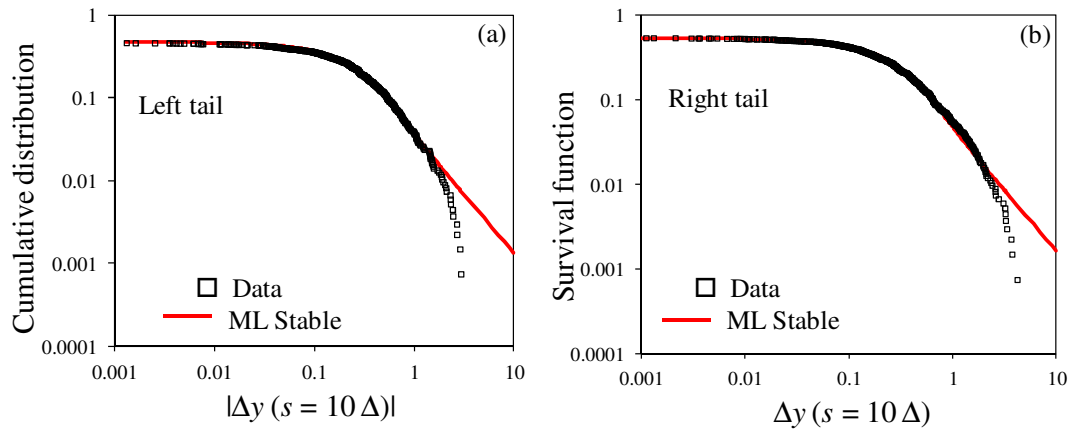

Fig. 13. (a) Empirical cumulative distribution and **(b)** survival function of increments Δy along x direction parallel to stratification computed on Berea sandstone faces at lag $s = 10\Delta$ compared with ML fits of stable distribution.

commonly discussed in the literature, they are of potential interest to hydrologists concerned with the possible effects of low-permeability geologic materials on resource extraction from, and contaminant migration through, the subsurface. The number of points falling below the power-law tails

of the stable distribution in Fig. 13 constitutes less than 2 % of all incremental data, justifying our ML fit of a stable distribution to all the data, including the former. Whether or not these outliers form Dragon Kings will be addressed next.
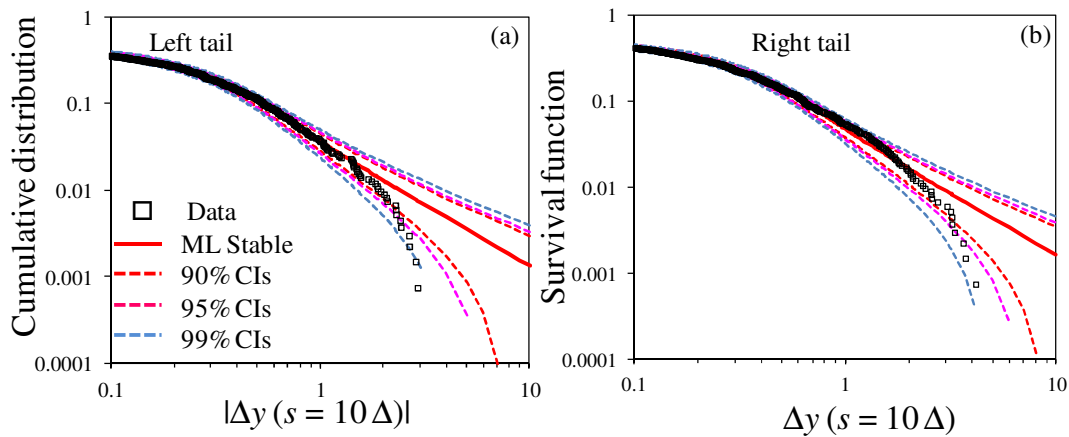
Fig. 14. Tails of **(a)** cumulative distribution versus $|\Delta y|$ for $\Delta y < 0$ and **(b)** survival function versus Δy for $\Delta y > 0$ of increments evaluated along x direction on Berea sandstone block at $s = 10\Delta$. Solid curves represent ML fitted stable distribution; dashed curves indicate CIs.
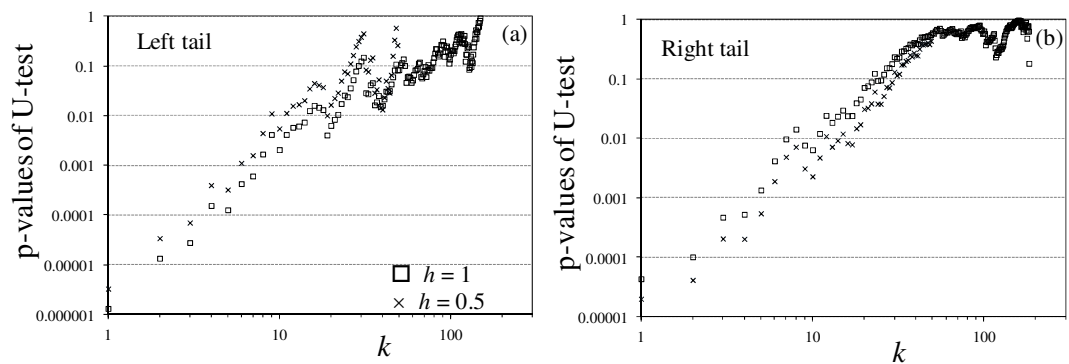

Fig. 15. U test p values for **(a)** left and **(b)** right tails of increments along x direction on the Berea sandstone block at $s = 10\Delta$.
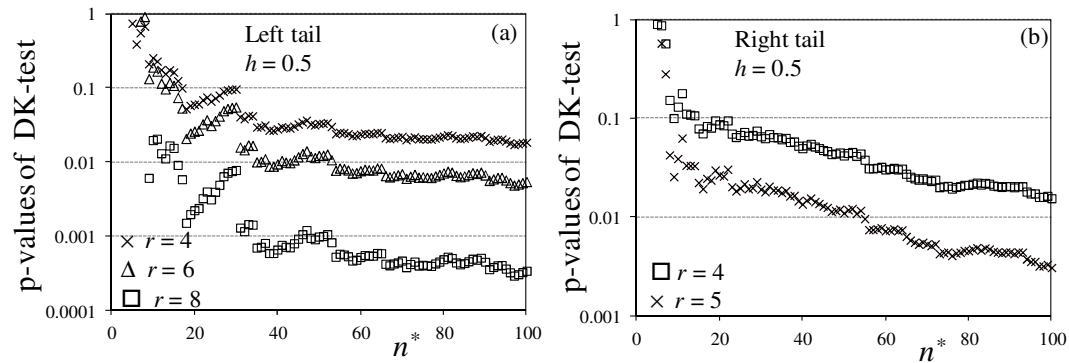

Fig. 16. DK test p values of increments along x direction on Berea sandstone block at $s = 10\Delta$: **(a)** left tail and $r = 4, 6, 8$, **(b)** right tail and $r = 4, 5$.

4.2 Diagnostic tests of Berea sandstone data

We end by applying the DK tests to our Berea sandstone data. As seen in Fig. 14, the confidence interval test diagnoses 13 DKs (8 on the left tail, 5 on the right tail, totaling fewer than 1 % of the sample) at a significance level of 0.1, 10 DKs (6 on the left tail, 4 on the right tail) at a significance level of 0.05, and 2 DKs (on the left tail) at significance level

of 0.01. This is in close agreement with results of the U test in Fig. 15 and those of the DK test (applied to r extreme values identified by the confidence interval test) in Fig. 16. As here our tests concern negative DKs, the p values for the U test and for the DK test are

$$p = F\left[T, 2r, 2\left(n^* - r\right)\right] \quad \text{and} \quad p_{u_k} = I_{u_k}\left(n^* - k + 1, k\right) \quad (30)$$

respectively.

In other words, all statistical tests infer the existence of negative DKs among the Berea sandstone incremental log permeability data. Our previous analysis of synthetic signals generated with a probabilistic model that appears to capture quite closely observed statistical scaling properties of the Berea sandstone data, and the lack of any visual evidence for aberrant geologic structure within the block, suggests to us that the inferred DKs are not real but rather artefacts of sampling.

5 Conclusions

We have tested the ability of three recently proposed statistical tests to diagnose the occurrence of Dragon Kings (DKs) in a random signal. Our tests relied on (a) synthetically generated signals extracted from α-stable sub-Gaussian random fields subordinated to truncated fractional Gaussian noise (tfGn) and on (b) natural log air permeabilities measured on the faces of a $30 \times 30 \times 30$ cm block of Berea sandstone. Though we know extreme-valued outliers in the synthetic signals to be part of the parent population, all three tests identify some of them (erroneously) as Dragon Kings (members of a different population). The same happens when the generated signals are independent and identically distributed (iid) as required formally by the tests. Application of the same statistical tests to the Berea data suggests the presence of negative DKs, for which there appears to be no independent hydrogeologic evidence. We conclude that existing statistical tests are not robust enough to reliably diagnose Dragon Kings in a sample.

Edited by: V. Lucarini
Reviewed by: R. T. Clarke and one anonymous referee

References

Ailliot, P., Thompson, C., and Thomson, P.: Mixed methods for fitting the GEV distribution, Water Resour. Res., 47, W05551, doi:10.1029/2010WR009417, 2011.

Akaev, A., Sadovnichy, V., and Korotayev, A.: On the dynamics of the world demographic transition and financial-economic crises forecast, Eur. Phys. J.-Spec. Top., 205, 355–373, doi:10.1140/epjst/e2012-01578-2, 2012.

Amitrano, D.: Variability in the power-law distributions of rupture events, how and why does b-values change, Eur. Phys. J.-Spec. Top., 205, 199–215, doi:10.1140/epjst/e2012-01571-9, 2012.

Ancey, C.: Are there "dragon-kings" events (i.e. genuine outliers) among extreme avalanches?, Eur. Phys. J.-Spec. Top., 205, 117–129, doi:10.1140/epjst/e2012-01565-7, 2012.

Bara, M., Kohnová, S., Gaál, L., Szolgay, J., and Hlavĉová, K.: Estimation of IDF curves of extreme rainfall by simple scaling in Slovakia, Contributions to Geophysics and Geodesy, 39, 187–206, 2009.

Barnett, V. and Lewis, T.: Outliers in statistical data, 3rd Edn., John Wiley & Sons, Chichester, 584 pp., 1994.

Benson, D. A., Schumer, R., and Meerschaert, M. M.: Recurrence of extreme events with power-law interarrival times, Geophys. Res. Lett., 34, L16404, doi:10.1029/2007GL030767, 2007.

Clarke, R. T.: Consistency and normality of estimates of hydrological extremes derived from a model for fragmented data, J. Hydrol., 346, 159–168, doi:10.1016/j.jhydrol.2007.09.010, 2007.

Clauset, A., Shalizi, C. R., and Newman, M. E. J.: Power law distributions in empirical data, SIAM Rev., 51, 661–703, doi:10.1137/070710111, 2009.

de Arcangelis, L.: Are dragon-king neuronal avalanches dungeons for self-organized brain activity?, Eur. Phys. J.-Spec. Top., 205, 243–257, doi:10.1140/epjst/e2012-01574-7, 2012.

De Michele, C., Zenoni, E., Pecora, S., and Rosso R.: Analytical derivation of rain intensity–duration–area–frequency relationships from event maxima, J. Hydrol., 399, 385–393, doi:10.1016/j.jhydrol.2011.01.018, 2011.

Di Federico, V. and Neuman, S. P.: Scaling of random fields by means of truncated power variograms and associated spectra, Water Resour. Res., 33, 1075–1085, 1997.

Ebtehaj, M. and Foufoula-Georgiou, E.: Orographic signature on multiscale statistics of extreme rainfall: A storm-scale study, J. Geophys. Res., 115, D23112, doi:10.1029/2010JD014093, 2010.

El Adlouni, S. and Ouarda, T. B. M. J.: Joint Bayesian model selection and parameter estimation of the generalized extreme value model with covariates using birth-death Markov chain Monte Carlo, Water Resour. Res., 45, W06403, doi:10.1029/2007WR006427, 2009.

Fernandes, W., Naghettini, M., and Loschi, R.: A Bayesian approach for estimating extreme flood probabilities with upper-bounded distribution functions, Stoch. Env. Res. Risk A., 24, 1127–1143, doi:10.1007/s00477-010-0365-4, 2010.

Golosovsky, M. and Solomon, S.: Runaway events dominated the heavy tail of citation distributions, Eur. Phys. J.-Spec. Top., 205, 303–311, doi:10.1140/epjst/e2012-01576-4, 2012.

Guadagnini, A., Neuman, S. P., and Riva, M.: Numerical Investigation of Apparent Multifractality of Samples from Processes Subordinated to Truncated fBm, Hydrol. Process., 26, 2894–2908, doi:10.1002/hyp.8358, 2012.

Janczura, J. and Weron, R.: Black swans or dragon-kings? A simple test for deviations from the power law, Eur. Phys. J.-Spec. Top., 205, 79–93, doi:10.1140/epjst/e2012-01563-9, 2012.

Johansen, A. and Sornette, D.: Stock market crashes are outliers, Eur. Phys. J. B, 1, 141–143, doi:10.1007/s100510050163, 1998.

Katz, R. W.: Statistics of extremes in climate change, Climatic Change, 100, 71–76, doi:10.1007/s10584-010-9834-5, 2010.

Katz, R. A, Parlange, M. B., and Naveau, P.: Statistics of extremes in hydrology, Adv. Water Resour., 25, 1287–1304, doi:10.1016/S0309-1708(02)00056-8, 2002.

Laio, F., Di Baldassarre, G., and Montanari, A.: Model selection techniques for the frequency analysis of hydrological extremes,

Water Resour. Res., 45, W07416, doi:10.1029/2007WR006666, 2009.

Langousis, A., Veneziano, D., Furcolo, P., and Lepore, C.: Multifractal rainfall extremes: Theoretical analysis and practical estimation, Chaos Soliton. Fract., 39, 1182–1194, doi:10.1016/j.chaos.2007.06.004, 2009.

Lee, T. and Ouarda, T. B. M. J.: Long-term prediction of precipitation and hydrologic extremes with nonstationary oscillation processes, J. Geophys. Res., 115, D13107, doi:10.1029/2009JD012801, 2010.

Lee, T., Ouarda, T. B. M. J., and Jeong, C.: Nonparametric multivariate weather generator and an extreme value theory for bandwidth selection, J. Hydrol., 452, 161–171, doi:10.1016/j.jhydrol.2012.05.047, 2012.

Lei, X.: Dragon-Kings in rock fracturing: Insights gained from rock fracture tests in the laboratory, Eur. Phys. J.-Spec. Top., 205, 217–230, doi:10.1140/epjst/e2012-01572-8, 2012.

Li, W., Zhang, C., Dey, D. K., and Wang, S.: Estimating threshold-exceeding probability maps of environmental variables with Markov chain random fields, Stoch. Environ. Res. Risk. A., 1113–1126, doi:10.1007/s00477-010-0389-9, 2010.

Main, I. and Naylor, M.: Extreme events and predictability of catastrophic failure in composite materials and in the Earth, Eur. Phys. J.-Spec. Top., 205, 183–197, doi:10.1140/epjst/e2012-01570-x, 2012.

Neuman, S. P., Guadagnini, A., Riva, M., and Siena, M.: Recent Advances in Statistical and scaling analysis of earth and environmental variables, in: Recent Advances in Hydrogeology, edited by: Mishra, P. K. and Kuhlman, K. L., Springer, New York, 1–25, doi:10.1007/978-1-4614-6479-2_1, 2013.

Nolan, J. P.: Maximum likelihood estimation of stable parameters, in: Levy processes: theory and applications, edited by: Barndorff-Nielsen, O. E., Mikosch, T., and Resnick, S. I., Birkhäuser, Boston, 379—400, 2001.

Peters, O., Christensen, K., and Neelin, J. D.: Rainfall and Dragon-Kings, Eur. Phys. J.-Spec. Top., 205, 147–158, doi:10.1140/epjst/e2012-01567-5, 2012.

Pisarenko, V. F. and Sornette, D.: Robust statistical tests of Dragon-Kings beyond power law distributions, Eur. Phys. J.-Spec. Top., 205, 95–115, doi:10.1140/epjst/e2012-01564-8,

Plenz, D.: Neuronal avalanches and coherence potentials, Eur. Phys. J.-Spec. Top., 205, 259–301, doi:10.1140/epjst/e2012-01575-5, 2012.

Ribereau, P., Naveau, P., and Guillou, A.: A note of caution when interpreting parameters of the distribution of excesses, Adv. Water Resour., 34, 1215–1221, doi:10.1016/j.advwatres.2011.05.003, 2011.

Riva, M., Neuman, S. P., and Guadagnini, A.: Sub-Gaussian model of processes with heavy tailed distributions applied to permeabilities of fractured tuff, Stoch. Env. Res. Risk A., 27, 195–207, doi:10.1007/s00477-012-0576-y, 2013a.

Riva, M., Neuman, S. P., Guadagnini, A., and Siena, M.: Anisotropic scaling of Berea Sandstone log air permeability statistics, Vadose Zone J., doi:10.2136/vzj2012.0153, in press, 2013b.

Rojas, R., Feyen, L., Dosio, A., and Bavera, D.: Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations, Hydrol. Earth Syst. Sci., 15, 2599–2620, doi:10.5194/hess-15-2599-2011, 2011.

Sachs, M. K., Yoder, M. R., Turcotte, D. L., Rundle, J. B., and Malamud, B. D.: Black swans, power laws, and dragon-kings: Earthquakes, volcanic eruptions, landslides, wildfires, floods, and SOC models, Eur. Phys. J.-Spec. Top., 205, 167–182, doi:10.1140/epjst/e2012-01569-3, 2012.

Samorodnitsky, G., and Taqqu, M.: Stable Non-Gaussian Random Processes, New York: Chapman and Hall, 1994.

Schoenberg, F. P. and Patel, R. D.: Comparison of Pareto and tapered Pareto distributions for environmental phenomena, Eur. Phys. J.-Spec. Top., 205, 159–166, doi:10.1140/epjst/e2012-01568-4, 2012.

Siena, M., Guadagnini, A., Riva, M., and Neuman, S. P.: Extended power-law scaling of air permeabilities measured on a block of tuff, Hydrol. Earth Syst. Sci., 16, 29–42, doi:10.5194/hess-16-29-2012, 2012.

Sornette, D.: Why Stock Markets Crash, Critical Events in Complex Financial Systems, Princeton University Press, 2003.

Sornette, D.: Dragon-kings, black swans, and the prediction of crises, International Journal of Terraspace Science and Engineering, 2, 1–8, 2009.

Sornette, D. and Ouillon, G.: Dragon-kings: Mechanisms, statistical methods and empirical evidence, Eur. Phys. J.-Spec. Top., 205, 10–26, doi:10.1140/epjst/e2012-01559-5, 2012.

Süveges, M. and Davison, A. C.: A case study of a "Dragon-King": The 1999 Venezuelan catastrophe, Eur. Phys. J.-Spec. Top., 205, 131–146, doi:10.1140/epjst/e2012-01566-6, 2012.

Taleb, N. N.: The Black Swan: The Impact of the Highly Improbable, Random House, 2007.

Tidwell, V. C. and Wilson, J. L.: Laboratory method for investigating permeability upscaling, Water Resour. Res., 33, 1607–1616, doi:10.1029/97WR00804, 1997.

Tidwell, V. C. and Wilson, J. L.: Permeability upscaling measured on a block of Berea Sandstone: Results and interpretation, Math. Geol., 31, 749–769, doi:10.1023/A:1007568632217, 1999.

Towler, E., Rajagopalan, B., Gilleland, E., Summers, R. S., Yates, D., and Katz, R. W.: Modeling hydrologic and water quality extremes in a changing climate: A statistical approach based on extreme value theory, Water Resour. Res., 46, W11504, doi:10.1029/2009WR008876, 2010.

Veneziano, D., Langousis, A., and Lepore, C.: New asymptotic and preasymptotic results on rainfall maxima from multifractal theory, Water Resour. Res., 45, W11421, doi:10.1029/2009WR008257, 2009.

Verhoest, N. E. C., Vandenberghe, S., Cabus, P., Onof, C., Meca-Figueras, T., and Jameleddine, S.: Are stochastic point rainfall models able to preserve extreme flood statistics?, Hydrol. Process., 24, 3439–3445, doi:10.1002/hyp.7867, 2010.

Viglione, A., Castellarin, A., Rogger, M., Merz, R., and Blöschl, G.: Extreme rainstorms: Comparing regional envelope curves to stochastically generated events, Water Resour. Res., 48, W01509, doi:10.1029/2011WR010515, 2012.

Zhang, Q., Sun, P., Chen, X., and Jiang, T.: Hydrological extremes in the Poyang Lake basin, China: changing properties, causes and impacts, Hydrol. Process., 25, 3121–3130, doi:10.1002/hyp.8031, 2011.