Eur. Phys. J. B **12**, 457–469 (1999)

THE EUROPEAN PHYSICAL JOURNAL B

EDP Sciences

Rare events in a log-Weibull scenario – Application to earthquake magnitude data

T. Huillet[1,a] and H.-F. Raynaud[2]

[1] LIMHP-CNRS, Université Paris XIII, Institut Galilée, 93430 Villetaneuse, France
[2] L2TI, Université Paris XIII, Institut Galilée, 93430 Villetaneuse, France

Received 15 March 1999 and Received in final form 20 May 1999

Abstract. We discuss the pertinency of the log-Weibull model in the statistical understanding of energy release for earthquake magnitude data. This model has many interesting features, the most remarkable of which being: depending on the value $\alpha > 0$ of the deformation index of the source, it may present tails ranging from moderately heavy ($\alpha < 1$) to very heavy (with tail index zero as $\alpha > 1$), through hyperbolic (power law) for the critical value $\alpha = 1$. Under this model (for which a precise tail study is supplied), the occurrence of power laws appears as a critical phenomenon: this reinforces the current trend predicting that some departure from the ideal (strictly scaling fractal) model may be ubiquitous. Having applied an affine transformation in the logarithmic scale, quantile estimation and the Kolmogorov-Smirnov statistics are used to fit the log-Weibull distribution to a realization of an *iid* sample. This enables to decide whether the upper tail of the phenomenon under study is light/heavy/very heavy. A comparative study of recorded French and Japanese earthquake magnitudes suggests that they exhibit comparable tail behaviour, albeit with different centrality and dispersion parameters.

PACS. 02.50.-r Probability theory, stochastic processes, and statistics – 05.20.-y Statistical mechanics – 89.90.+n Other areas of general interest to physicists

1 Introduction

The so-called “heavy-tailed” power-law distributions are defined through a *cpdf* (cumulative probability distribution function) in the form

$$F_E(z) = 1 - (z/z_0)^{-a} \qquad a > 0. \tag{1}$$

Such distributions have been used to model a wide range of natural or social phenomena (see, *e.g.*, [1–3]). A simple distinctive feature of this class of distributions is that the log-log plot of the complementary cumulative distribution function $1 - F_E(z)$ is a straight line with negative slope $-a$. However, empirical cumulative distribution functions have been shown to exhibit at most a limited quasi-linear regime followed by significant curvature [4]. This is the hint that, in some sense to be made precise in the sequel, the power-law class should be “*unstable*”.

In [4], the authors argued that such departures from the power-law description should not necessarily be explained by the finite size of the data, but could result from a deeper departure from the power-law hypothesis. Using rank-ordering statistics to back up their claim, they suggested that occurrences of numerous phenomena, ranging from earthquake death tolls and energies [5,6] to radio light emissions in galaxies, apparently fit the so-called “stretched exponential” (*Weibull*) subexponential distribution

$$F_E(z) = 1 - \exp\left[-(z/z_0)^{1/\alpha}\right] \qquad \alpha > 1. \tag{2}$$

However, in a recent paper [7], it has been underlined that the stretched exponential model could *not* account for the *skewness* of the empirical distribution function of earthquake energy E in France; rather, the related heavy-tailed *Fréchet* model has been shown to present nice fit properties in this respect. It turns out that both models can be generated by deforming the exponentially distributed driving source S through

$$E = (S/s_0)^{\alpha} \tag{3}$$

Here, s_0 is a scale parameter, and α is a deformation index of the energy source governing the tail behavior of the distribution.

However, in the *Weibull* model the deformation index α is positive, whereas in the *Fréchet* model it is negative: the first model amplifies (expands) the triggering noisy source whereas the second attenuates (damps) its effect as it can be easily checked from the graph of the deformation function $x \to x^{\alpha}$. This “damping” property of the *Fréchet* distribution is questionable and puzzling, since high (low) energies are produced by small (high) amplitudes of the source.

[a] e-mail: huillet@limhp.univ-paris13.fr

We then ask the following question: is there a simple stochastic model for E which is source-*expanding* and can account for both low and high energy tails?

In this article, we therefore advocate a simple (source-expanding) class of distribution functions, the so-called *log-Weibull* model class, presenting many interesting features in this respect: depending on the value $\alpha > 0$ of its deformation index, it may present tails ranging from moderately heavy ($\alpha < 1$) to very heavy (with tail index zero as $\alpha > 1$), through hyperbolic (power law) for the *critical* value $\alpha = 1$. Thus, power-laws appear as a critical phenomenon separating *subcriticality* ($\alpha < 1$) from *supercriticality* ($\alpha > 1$). In addition, the stretched exponential model may itself be recovered from the *log-Weibull* model zooming in the far ends of its characteristic scales.

Physically, this model arises from the observation that the ignition energy of the source should first be amplified (as in the *Weibull* setup), *then* exponentiated to produce the desired distribution for energy release: some additional cascade phenomenon (an explosion of the type of a chain reaction) is assumed to take place. As a result, a new type of behavior for the tails of the energy release is forecasted in the supercritical case: distributions with tail index zero whose tails are fatter than the ones of any power law with positive tail index. These phenomena exhibit quite uncanny statistical properties, especially their tail properties (with its various implications), which are studied in some detail in Section 2.

Identifying the value of the model parameters which fits the data in the best way according to some criterion would enable to decide whether the phenomenon under study exhibits light (moderately heavy), heavy or very heavy extremal behavior from the tails. Limit theorems for rank-ordering statistics, including the *Kolmogorov-Smirnov* test, can then be used to test whether the identified model is compatible with the empirical cumulative distribution as a whole or with its upper tail only.

It turns out that solving both the parameter identification problem and the compatibility tests is made a lot easier by using an affine transformation of the variable $\log E$, which therefore emerges as the "natural" choice of coordinates for this class of problems: the logarithmic deformation smooths out the tails of the original energy variable. The range of the observable variable X is now the real line.

Crossing of the value $\alpha = 1$ from below, the most distinctive feature is that the upper-tails of X (for positive values) become fatter than the exponential lower-tails (for negative values). Superexponential upper-tails, lighter than exponential lower-tails, therefore stand for a *signature* of the event $\alpha < 1$ whereas subexponential upper-tails, heavier than exponential lower-tails, indeed is a signature of the event $\alpha > 1$ for which energy is of tail index zero: crossing $\alpha = 1$ chiefly translates into a skewness inversion of the *df* for X.

2 The log-Weibull energy model

2.1 The Weibull (stretched exponential) model

Consider the class of random variables defined as

$$E = (S/s_0)^{\alpha} \qquad \alpha > 0 \qquad s_0 > 0 \tag{4}$$

where S is an exponentially distributed random variable with mean unity, *i.e.* with *cpdf*

$$F_S(s) = 1 - \exp(-s). \tag{5}$$

The variable E can be seen as the output of some deterministic "machine", with parameters (α, s_0), triggered by the stochastic source of disorder S [8,7].

While s_0 is simply a scaling factor, the parameter α defines, roughly speaking, the way in which the disorder generated by the source S is amplified through the transformation (4) over the positive real axis. For positive z, the density function (*df*) and *cpdf* of E are obtained easily by combining (4) with (5), yielding the *Weibull* distribution [9]:

$$f_E(z) = \frac{s_0}{\alpha} z^{1/\alpha - 1} \exp\left(-s_0 z^{1/\alpha}\right) \tag{6}$$

$$F_E(z) = 1 - \exp\left(-s_0 z^{1/\alpha}\right) \text{ with } \alpha > 0. \tag{7}$$

E is a special case of the so-called Von Mises variables [10], whose complementary cumulative distributions can be written in the form

$$\overline{F}_E(z) := 1 - F_E(z) = \overline{F}_E(z_0) \exp - \int_{z_0}^{z} h_E(z)\, \mathrm{d}z \tag{8}$$

where the *hazard energy density* h_E defined by this formula verifies

$$\lim_{z \uparrow +\infty} z h_E(z) = +\infty. \tag{9}$$

The cumulative distribution of a *Von Mises* variable decreases towards zero faster than hyperbolically, so that these distributions are light-tailed (or rapidly varying). As a consequence, these variables have moments of arbitrary positive order. If in addition the function h_E verifies

$$\lim_{z \uparrow +\infty} h_E(z) = 0 \tag{10}$$

the variable E is said to be subexponential (with moderately heavy tails); otherwise, it is superexponential (with thin tails). In the stretched exponential case, we get

$$h_E(z) = \frac{s_0}{\alpha} z^{1/\alpha - 1}. \tag{11}$$

Thus, when $\alpha > 1$ the *Weibull* variable E is subexponential, whereas for $0 < \alpha < 1$ it is superexponential.

For subexponential (superexponential) distributions, the tails of the *pdf* decrease towards zero at exponential rate or slower (faster) at $+\infty$.

Note that $\alpha = 1$ yields the critical exponential model and that the larger α is, the fatter the tails are for E, while remaining exponential.

With symbol $\stackrel{d}{\to}$ standing for convergence in distribution, a remarkable attribute of the *Weibull* distribution is the min-stable property:

$$n^{\alpha} E_{1:n} \stackrel{d}{\to} E \text{ as } n \uparrow \infty. \tag{12}$$

Here $E_{1:n} := \min(E_1, .., E_n)$ and $E_1^n := (E_1, .., E_n)$ is an *iid* sample drawn from a distribution satisfying

$$\overline{F}_E(z) \underset{z \uparrow 0^+}{\sim} \exp\left(-z^{1/\alpha}\right). \tag{13}$$

The $n^{-\alpha}$ scaling of $E_{1:n}$ only shows the common behavior at $z = 0$ of the constitutive elements of the minimum, discarding all other details.

Note also that if $E_1^n := (E_1, .., E_n)$ is an *iid* sample drawn from the distribution of E, $(E_{1:n})_{n \geq 1}$ is a *Markov self-similar* sequence (with similarity parameter $-\alpha$) in the sense that for any integer $a \geq 1$

$$a^{\alpha} E_{1:an} \stackrel{d}{=} E_{1:n}\ . \tag{14}$$

2.2 The chain reaction model

We shall next assume that this model for energy release is not complete in the following sense: the random energy produced by the triggering of some random source S actually proceeds in two steps; the first one is the one just mentioned: $S \to S^{\alpha}$ which informs us on the way the source was initially deformed: in this respect, α may be called the primary deformation index of the source. The second one is $S^{\alpha} \to \exp S^{\alpha}$ which one may relate to some "avalanche" effect: the first step energy release enters into a chain reaction (explosion) which snowballs into an exponential term to produce the final state variable.

In more precise terms, we shall assume that the right model for energy release is

$$E = z_0 \left(\exp\left(S/s_0\right)^{\alpha} - 1\right) > 0 \tag{15}$$

where $z_0 > 0$ is the (unknown) characteristic scale for E, just like $s_0 > 0$ was for S, and α is thus the primary deformation index.

As the variable $(S/s_0)^{\alpha}$ has just been shown to be distributed according to *Weibull* distribution, (15) now defines a variate which is *log-Weibull* distributed in the sense that $\log(1 + E/z_0)$ simply is *Weibull* distributed. This model will be shown to present a certain number of interesting new features.

For positive z, the *df* and *cpdf* of E are now obtained easily by combining (15) with (1), yielding the *log-Weibull* distribution:

$$f_E(z) = \frac{s_0}{\alpha z_0}\left(\log\left(1 + \frac{z}{z_0}\right)\right)^{1/\alpha - 1} \times \left(1 + \frac{z}{z_0}\right)^{-1} \overline{F}_E(z) \tag{16}$$

$$F_E(z) = 1 - \exp -s_0\left(\log\left(1 + \frac{z}{z_0}\right)\right)^{1/\alpha} \text{ with } \alpha > 0. \tag{17}$$

Note from this expression and criterion arising from (10) that the energy E is subexponential whatever the value of parameter α, so that the tails are at least moderately heavy. Let us discuss the tails shape in more details, depending on the value of parameter α. We shall distinguish three cases:

– Subcritical energy release: $\alpha < 1$.

When $\alpha < 1$, it can easily be checked that, with the *hazard energy density* h_E defined from (8) applied to (17): $\lim_{z \uparrow +\infty} h_E(z) = 0$. Thus model (15) for energy release is *Von Mises'*, with moderately heavy tails in the subcritical region for α. This model can account for situations showing significant departure from the power-law description of nature.

In the subcritical range for the parameter α, this model may be compared to the *log-normal* model [11] which also belongs to the *Von Mises*, with moderately heavy tails' class. For this model, it can easily be shown that

$$\overline{F}_E(z) \underset{z \uparrow \infty}{\sim} \exp\left[-(\log z)^2\right] \tag{18}$$

Thus, the *log-Weibull* distribution has heavier tails than the *log-normal* distribution as soon as $\alpha > 1/2$. The *log-normal* model has found numerous applications in various fields such as nanosciences (where E is the grain size of the nanoparticle) [12], internet traffic (network resource demands) [13], finance with the geometrical *Brownian* motion [10], earth sciences especially in the evaluation of the world oil and hydrocarbon reserve sizes [4,14,15] and turbulence [16].

– Critical energy release: $\alpha = 1$.

When $\alpha = 1$, the *cpdf* of E is easily obtained from (17), it is:

$$F_E(z) = 1 - \left(1 + \frac{z}{z_0}\right)^{-s_0}. \tag{19}$$

Let us recall [17] that a distribution is said to be *heavy-tailed* (or *slowly varying*), with tail index $a > 0$, if there exists some finite strictly positive constant a such that

$$\overline{F}_E(z) \underset{z \uparrow +\infty}{\sim} z^{-a} L(z) \tag{20}$$

where L is some function with regular variation, *i.e.* such that for all strictly positive t:

$$\lim_{z \uparrow +\infty} \frac{L(tz)}{L(z)} = 1. \tag{21}$$

Such distributions have only moments of order less than a, and the smaller a is, the fatter the tails are for the distribution of E.

This condition applied to (19) shows that the critical model (15) for energy release is heavy-tailed, with tail index $s_0 > 0$: in other words, it follows a *Pareto-Zipf* power law [1,18] with exponent $s_0 > 0$. Note that crossing the value $s_0 = 1$ in the critical model ($\alpha = 1$) is again critical in the (weaker) sense that the energy E has infinite mean when $s_0 \leq 1$, finite when $s_0 > 1$. When $s_0 > 1$, the conditional excess mean value of $E - z_c$ given $E > z_c$ is $z_c / (s_0 - 1)$ for large cutoff positive z_c: scaling supercritical power laws have no intrinsic characteristic scale but the one of the (tail) observer.

– Supercritical energy release: $\alpha > 1$.

This is the most interesting feature of the model described in this article.

When $\alpha > 1$, the *cpdf* of E satisfies the following interesting property, as a result of (17):

For any strictly positive constant a

$$\frac{\overline{F}_E(z)}{z^{-a}} \underset{z \uparrow +\infty}{\rightarrow} +\infty \tag{22}$$

Thus the tails of the distribution of E are fatter than any power-law with exponent $a > 0$: they are said to be heavy tailed with tail index zero. As a result, such distributions have no moment of any arbitrary positive order! Thus, model (15) for energy release in the supercritical regime is heavy-tailed, with tail index zero; we shall call such models "very heavy tailed": fitting this distribution to a natural phenomenon would mean that there exists in nature extreme situations with very heavy tails (and hence very special properties); this fact, to the authors' knowledge, has not been underlined and discussed, at least in the physics' literature. It has been argued that distributions with infinite mean are unrealistic in some physical applications, so that distributions with tail index zero should be ruled out even more vigorously. However, the practical relevance of such random variables remains on our opinion an open question: after all, they can be considered "realistic" in the more limited sense that they take finite values with probability one.

Let us make three remarks on this model, underlining its importance in our physical context:

Remark 1 As conventional wisdom suggests, the larger α, the heavier the tails for E, ranging from moderately heavy tails ($\alpha < 1$) to very heavy tails ($\alpha > 1$), through heavy tails (power law) with exponent $s_0 > 0$ in the critical situation when $\alpha = 1$. In this model, the power laws appear as a critical phenomenon. This may explain why (although they are appealing due to the underlying self-similarity property [2,19]), it is hard to observe such distributions in nature: it separates two regimes very distinct from the tail behavior point of view.

Remark 2 As noted above, for any value of $\alpha > 0$, the energy variable E is subexponential, *i.e.* exhibits at least moderately heavy tails.

The maximum $E_{n:n} := \max(E_1, .., E_n)$ of an n-sample $(E_1, .., E_n)$ is tail equivalent to the sum $\overline{E}_n := \sum_{m=1}^{n} E_m$ [10], in the sense that

$$\frac{P(E_{n:n} > x)}{P(\overline{E}_n > x)} \underset{x \uparrow +\infty}{\rightarrow} 1 \tag{23}$$

The tail of the maximum determines the tail of the sum. If $\alpha > 1$, hence with distributions with tail index zero, two stronger results actually hold. They are

$$\frac{E_{n:n}}{\overline{E}_n} \rightarrow 1 \text{ (in probability).} \tag{24}$$

In addition, if and only if: $\alpha > 2$ ([20,21]),

$$\frac{E_{n:n}}{\overline{E}_n} \rightarrow 1 \text{ (almost surely).} \tag{25}$$

In this case, a single event explains (in probability or even almost surely) a cumulative event.

Remark 3 The *Weibull* (stretched exponential) model itself may be recovered from the following interesting scaling property of the *log-Weibull* model (15):

$$E \overset{d}{\rightarrow} (S/s)^{\alpha} \tag{26}$$

as $z_0 \uparrow \infty$, $s_0 \uparrow \infty$ while $s_0/(z_0^{1/\alpha}) = s > 0$. As a result, when the characteristic scales for E and S both tend to infinity while the ratio $s_0/z_0^{1/\alpha}$ is held constant, one recovers the stretched exponential model for energy release as a limit case. This may explain why the stretched exponential distribution may be identified with a phenomenon with no characteristic scale in itself.

3 Collecting data

3.1 The logarithmic model

In many physical situations, the random variable E is not directly observed. Rather, the observed variable is the real-valued variable

$$X := \mu + \sigma \log(E/z_0). \tag{27}$$

In this equation, $\sigma > 0$ is now the characteristic scale (dispersion parameter) for the observable X, whereas μ is a location (centrality) parameter.

The distinctive feature of the logarithmic scale is that it measures the distance between two values through their ratio rather than their difference. Thus, the intensity of noise, as perceived by the human ear, is usually measured in decibels, *i.e.* using a logarithmic scale. Similarly, earthquake magnitude is determined from the logarithm of the amplitude of waves recorded by seismographs, properly adjusted to compensate for the variation in the distance between the various seismographs and the epicenter of the earthquake.

Another motivation for working with X rather than with E is that the logarithmic transformation has a regularizing effect on the distribution's tail: logarithms are notorious for contracting data. Most notably, as noted above,

E may have no mean value (in both critical heavy-tailed range $\alpha = 1$, when $0 < s_0 < 1$ and in the whole supercritical range $\alpha > 1$ where no positive moment at all exist), whereas, as we shall see, X always has positive moments of any order – a fact which suggests that when dealing with heavy tailed distributions for energies one should rather consider their geometrical empirical mean rather than the arithmetical one: *multiplicative* micro-pulses generate *log-Weibull* energies.

Elementary calculations yield the *df* and *pdf* for the variable X

$$f_X(x) = \frac{s_0}{\alpha\sigma}\left(\log\left(1+\mathrm{e}^{(x-\mu)/\sigma}\right)\right)^{1/\alpha-1} \times \left(1+\mathrm{e}^{(\mu-x)/\sigma}\right)^{-1}\overline{F}_X(x) \tag{28}$$

$$F_X(x) = 1 - \exp - s_0\left(\log\left(1+\mathrm{e}^{(x-\mu)/\sigma}\right)\right)^{1/\alpha}. \tag{29}$$

For all choice of the parameter vector $(\alpha, s_0, \sigma, \mu)$, the variable X is *Von Mises*', as

$$\lim_{x\uparrow+\infty} x h_X(x) = +\infty \tag{30}$$

with *hazard energy density* given by

$$h_X(x) = \frac{s_0}{\alpha\sigma}\left(\log\left(1+\mathrm{e}^{(x-\mu)/\sigma}\right)\right)^{1/\alpha-1} \times \left(1+\mathrm{e}^{(\mu-x)/\sigma}\right)^{-1}. \tag{31}$$

More precisely,

$$\overline{F}_X(x) \underset{x\uparrow+\infty}{\sim} \exp - s_0 (x/\sigma)^{1/\alpha} \tag{32}$$

so that

$$h_X(x) \underset{x\uparrow+\infty}{\sim} \frac{s_0}{\alpha\sigma}(x/\sigma)^{1/\alpha-1}. \tag{33}$$

Consequently, when $\alpha < 1$ (respectively $\alpha > 1$) it is superexponential (respectively subexponential), which means that the tails of its *pdf* decrease towards zero at rate faster (slower) than exponential at $x = +\infty$. In this case, the distribution is tail-equivalent to a *Weibull* model.

At $x = -\infty$, we have

$$F_X(x) \underset{x\uparrow-\infty}{\sim} s_0 \exp \alpha (x/\sigma). \tag{34}$$

Thus, X always exhibits exponential tails at $x = -\infty$.

Hence, in all cases, the distribution of the observable is *Von Mises$'$* (and thus with integral moments of arbitrary order). It is very asymmetric except in the critical case $\alpha = 1$, with strictly exponential tails at both ends of the support. When $\alpha = 1$, the upper distribution tail is exponential, with parameter s_0/σ, whereas the lower tail is exponential, with parameter $1/\sigma$. Perfect symmetry of the distribution is obtained only when $\alpha = s_0 = 1$.

Therefore, the most distinctive feature of the fact of crossing the value $\alpha = 1$ from below is that the upper-tails (for positive x) become fatter than the exponential lower-tails (for negative x). Superexponential upper-tails, lighter than exponential lower-tails, therefore stand out as a *signature* of the event $\alpha < 1$. Conversely, subexponential upper-tails, heavier than exponential lower-tails, indeed is a *signature* of the event $\alpha > 1$ for which energy is of tail index zero. Thus, the distribution's asymmetry gets inverted upon crossing the threshold $\alpha = 1$.

3.2 Large deviation from the mean

From the law of large numbers, the empirical mean $\frac{1}{n}\overline{X}_n$ converges almost surely towards the theoretical mean, say $\overline{m}_X := \mathbb{E}X$ of X, which is known to exist. Large deviation theory is concerned with the evaluation of the (small) probability [22]

$$P\left(\frac{1}{n}\overline{X}_n > x\right) \tag{35}$$

when x exceeds the mean. More precisely, it is concerned with the rate at which this probability tends to zero, as a function of the sample size n. We shall distinguish two cases, depending on the tail behavior of X:

- $\alpha \leq 1$. In this case the Laplace transform of $f_X(x)$, say

  $$Z(\beta) := \int_{\mathbb{R}} \mathrm{e}^{-\beta x} f_X(x)\,\mathrm{d}x \tag{36}$$

  (with real β) is defined in a open neighborhood of $\beta = 0$. As a result, it is well-known that, with $x > \overline{m}_X$

  $$n^{-1}\log P\left(\frac{1}{n}\overline{X}_n > x\right) \underset{n\uparrow\infty}{\to} s((x-\mu)/\sigma) < 0. \tag{37}$$

  Here $s((x-\mu)/\sigma)$ the concave *Cramér-Chernoff* transform of the "free energy" $-\log Z(\beta)$. The probability that the empirical mean deviates from the mean tends to zero exponentially fast as n grows, with rate $s((x-\mu)/\sigma)$.
- $\alpha > 1$. The function $Z(\beta)$ is no longer defined in a neighborhood of $\beta = 0$, so that the previous result does *not* hold in this (subexponential) case. However, from the tail equivalence of the maximum and sum for subexponential distributions, one gets

  $$P\left(\frac{1}{n}\overline{X}_n > x\right) \underset{n\uparrow\infty}{\sim} P\left(\frac{1}{n}X_{n:n} > x\right) = 1 - (F_X(nx))^n. \tag{38}$$

  From (29), with $x > \overline{m}_X$

  $$1 - (F_X(nx))^n \underset{n\uparrow\infty}{\sim} n \exp - s_0\left(\log\left(1+\mathrm{e}^{n(x-\mu)/\sigma}\right)\right)^{1/\alpha} \underset{n\uparrow\infty}{\sim} \exp - s_0 (n(x-\mu)/\sigma)^{1/\alpha} \tag{39}$$

showing that (37) should be replaced by

$$n^{-1/\alpha} \log P\left(\frac{1}{n}\overline{X}_n > x\right) \underset{n\uparrow\infty}{\rightarrow} -s_0\left(\left(x-\mu\right)/\sigma\right)^{1/\alpha} < 0. \quad (40)$$

This formula exhibits a slower decay of $P\left(\frac{1}{n}\overline{X}_n > x\right)$ towards zero, due to the presence of moderately heavy tails as $\alpha > 1$.

Note that the arithmetic mean for the sequence X_1^n, $\frac{1}{n}\sum_{m=1}^{n} X_m$, corresponds to a geometric mean of the energy records E_1^n: $\prod_{m=1}^{n} E_m^{1/n}$. This observation and the large deviation results just mentioned show that there are large deviation results for the energy sequence itself but not of its empirical arithmetic mean (it simply could not converge) but rather of its empirical geometrical mean. In explicit form, (37,40) read in terms of the geometrical mean of the energies

$$n^{-1} \log P\left(\prod_{m=1}^{n} E_m^{1/n} > z\right) \underset{n\uparrow\infty}{\rightarrow} s\left(\log\left(z/z_0\right)\right) < 0 \quad (41)$$

as $\alpha \leq 1$.

$$n^{-1/\alpha} \log P\left(\prod_{m=1}^{n} E_m^{1/n} > z\right) \underset{n\uparrow\infty}{\rightarrow} -s_0 \log^{1/\alpha}\left(z/z_0\right) < 0. \quad (42)$$

as $\alpha > 1$.

3.3 Asymptotic behavior for the max(min)imum of the observable

When dealing with extreme events, it may be useful to understand the way the max(min)imal event behaves as the sample size grows [23,24].

First observe the obvious fact that $X_{n:n} \overset{a.s.}{\rightarrow} \infty$, as $n \uparrow \infty$. This observation does not enclose too much information and one would like a deeper insight on how the order of magnitude of the maximum evolves, as $n \uparrow \infty$. This is the purpose of what follows.

Define the increasing quantile sequence $\left(x_n^*\right)_{n\geq 1}$ by

$$n\overline{F}_X\left(x_n^*\right) = 1. \quad (43)$$

Then, the *Fisher-Tippett* theorem [9,10] for *Von Mises'* variable yields the following convergence in distribution for the maximum $X_{n:n} := \max\left(X_1, ..., X_n\right)$,

$$h_X\left(x_n^*\right)\left(X_{n:n} - x_n^*\right) \overset{d}{\rightarrow} G \text{ as } n \uparrow \infty. \quad (44)$$

where G is a *Gumbel* random variable, *i.e.* with *cpdf:* $P\left(G \leq t\right) = \mathrm{e}^{-\mathrm{e}^{-t}}$.

In our case, with $\sigma > 0$, one may check that

$$x_n^* \underset{n\uparrow+\infty}{\sim} \sigma\left[\frac{1}{s_0}\log n\right]^{\alpha} \quad (45)$$

and that

$$h_X\left(x_n^*\right) \underset{n\uparrow+\infty}{\sim} \frac{s_0}{\alpha\sigma}\left[\frac{1}{s_0}\log n\right]^{1-\alpha}. \quad (46)$$

As a result, the maximum typically grows like $\left(\log n\right)^{\alpha}$. Remark also from (44) that the quantity $\left(h_X\left(x\right)\right)^{-1}$ can be interpreted as the *absolute* standard fluctuation for the maximum around $x = x_n^*$: from the previous expression of $h_X\left(x_n^*\right)$, this *absolute* fluctuation tends to $+\infty$ (zero) as n grows depending on $\alpha > 1$ ($\alpha < 1$) and thus showing (again) two distinct regimes. However, in any case, the *relative* fluctuation for the maximum: $1/\left(x_n^* h_X\left(x_n^*\right)\right)$ tends to zero as $n \uparrow \infty$.

Concerning the minimum of n *iid* observed random variables, say $X_{1:n} := \min\left(X_1^n\right)$, one has

$$X_{1:n} = -\max\left(-X_1^n\right) \quad (47)$$

so that it suffices to reason in a similar way as for the maximum, but working with $-X$.

Proceeding this way, one easily shows that the minimum typically behaves like

$$x_n^* \underset{n\uparrow+\infty}{\sim} -\frac{\sigma}{s_0}\log\left(ns_0\right) \quad (48)$$

with constant *absolute* standard fluctuation σ/α.

These constructions allow for an approximation of the width of the confidence interval of the max(min)imum around their "typical" value x_n^*, as the sample size n grows. To do this, one should fix a small real number, say $\delta = 0.05$, and compute the radius of the ball centered at x_n^* which is likely (with probability $1-\delta$) to include the maximum value in the data set. Using the asymptotic behavior (44), this confidence interval is therefore obtained in the form $\left[x_n^* - \varepsilon\left(\delta\right); x_n^* + \varepsilon\left(\delta\right)\right]$, where

$$P\left(\left|G\right| > \varepsilon\left(\delta\right) h_X\left(x_n^*\right)\right) = \delta. \quad (49)$$

Thus, for $\delta = 0.05$, we get

$$\varepsilon\left(\delta\right) \simeq \frac{3}{h_X\left(x_n^*\right)} \quad (50)$$

and the test for adequacy of the tail distribution is as follows: accept the hypothesis if $x_{n:n} \in \left[x_n^* - \varepsilon\left(\delta\right); x_n^* + \varepsilon\left(\delta\right)\right]$, reject it otherwise.

4 Parameter identification

We now address the problem of fitting this model to a particular data set. The problem is to decide whether, and for what choice of the parameter set $\left(\alpha, s_0, \sigma, \mu\right)$, the distribution function F_X is a good statistical model for a particular data set $\left(x_1, x_2, \ldots, x_n\right)$; or equivalently, whether F_E is a good model for the data set $\left(\mathrm{e}^{x_1}, \mathrm{e}^{x_2}, \ldots, \mathrm{e}^{x_n}\right)$ under the hypothesis that $\left(x_1, x_2, \ldots, x_n\right)$ is a realization of an *iid* sequence $X_1^n := \left(X_1, X_2, \ldots, X_n\right)$ with *pdf* F_X.

The solution advocated here will be a procedure based on the *Kolmogorov-Smirnov* test.

The *Kolmogorov-Smirnov* test allows one to decide whether or not an *iid* sample has been drawn according to some completely known probability distribution function, using a rescaled measure of the distance between the empirical and theoretical *pdfs* (the *Kolmogorov-Smirnov* statistics). However, if the distribution function depends on unknown parameters and if the sample has been used to estimate those parameters, the *Kolmogorov-Smirnov* statistics should be corrected to take into account the dependency resulting from parameter identification. This is a difficult problem which, to the authors' knowledge, has so far been solved only for very special cases, such as normal or exponential distributions [25,26]. A simple, albeit suboptimal, alternative is to cut the data set in two subsamples; the first one will be used to identify the distribution's parameters, and the second one to test the adequacy of the identified model.

We therefore propose the following three-steps parameter identification procedure:

– The first step, discussed in Section 4.1, is to use quantile estimation techniques and the first subsample to obtain an initial value for the parameter set $(\alpha, s_0, \sigma, \mu)$. The purpose of this initial identification is to determine the order of magnitude for the model parameters.

– The second step, discussed in Section 4.2, is to search the value of the parameter quadruple which minimizes the *Kolmogorov-Smirnov* distance for the first subsample, using the initial value obtained in step 1 as a starting point for a local non-linear optimization algorithm.

– The third step, discussed in Section 4.3, is to test the adequacy of this estimated distribution with the empirical distribution of the second subsample, using a standard *Kolmogorov-Smirnov* test.

In the sequel, the two subsamples used for identification and testing will be denoted respectively

$$\mathbf{x}^{(1)} := \left(x_1^{(1)}, x_2^{(1)}, \ldots, x_{n_1}^{(1)}\right) \tag{51}$$

$$\mathbf{x}^{(2)} := \left(x_1^{(2)}, x_2^{(2)}, \ldots, x_{n_2}^{(2)}\right) \tag{52}$$

with $n_1 + n_2 = n$.

4.1 Step one: quantile estimation

For any $p \in (0,1)$, the theoretical p-quantile $x(p)$ is defined by: $F_X(x(p)) = p$. From (29), this is also

$$\left(-\frac{1}{s_0}\log(1-p)\right)^{\alpha} = \log\left(1 + \exp\frac{x(p)-\mu}{\sigma}\right) > 0. \tag{53}$$

Taking the logarithm of this equality yields

$$-\alpha\log s_0 + \alpha\log\log(1/(1-p)) = \log\log\left(1+\exp\frac{x(p)-\mu}{\sigma}\right). \tag{54}$$

This formula shows that, if the model fits the data for some parameter set $(\alpha, s_0, \sigma, \mu)$, the quantities

$$y(p) := \log\log(1/(1-p)) \tag{55}$$

$$z(p) := \log\log\left(1+\exp\frac{x(p)-\mu}{\sigma}\right) \tag{56}$$

should be aligned, as p varies in $(0,1)$, with positive slope α and abscissa at the origin $-\alpha\log s_0$.

Now, choose four distinct probability values (p_1, p_2, p_3, p_4) for parameter p – for example $p = (0.05, 0.25, 0.75, 0.95)$. Denote as $(x_{1:n_1}, .., x_{n_1:n_1})$ the ordered version of subsample $\mathbf{x}^{(1)}$, which means

$$x_{1:n_1} < .. < x_{m:n_1} < .. < x_{n_1:n_1} \tag{57}$$

The empirical p-quantiles are respectively

$$x_{n_1}(p) = x_{[n_1p]:n_1} + \left(x_{[n_1p]+1:n_1} - x_{[n_1p]:n_1}\right)(n_1p - [n_1p]) \tag{58}$$

for $p = p_i$, $i = 1, \ldots, 4$, where $[x]$ stands for the integer part of x.

Next, the quantile estimation algorithm works as follows:

– for each $i = 1, \ldots, 4$, compute the abscissa $y(p_i)$ and the associated empirical ordinate

$$z(p_i) := \log\log\left(1+\exp\frac{x_{n_1}(p_i)-\mu}{\sigma}\right) \tag{59}$$

– Find the value of (μ, σ), say $(\overline{\mu}, \overline{\sigma})$, which aligns the four points $(y(p_i), z(p_i))$ for $i = 1, \ldots, 4$. Once this alignment is performed, the slope of the straight line, say $\overline{\alpha}$, is the estimated value of parameter α, while from the reading a of the abscissa at the origin, an estimator $\overline{s}_0$ for the characteristic scale s_0 of the source is

$$\overline{s}_0 = \exp(-a/\overline{\alpha}) \tag{60}$$

4.2 Step two: minimizing the Kolmogorov-Smirnov distance

Let $(X_{1:n_1}, .., X_{n_1:n_1})$ be the ordered version of the *iid* subsample $\mathbf{X}^{(1)}$ with *pdf* F_X, and let F_X^- be the quantile distribution function (*qdf*) of X :

$$F_X^-(p) := \inf(x : F_X(x) > p) \tag{61}$$

This *qdf* is easily computed from (17-19):

$$F_X^-(p) = \mu + \sigma\log\left(\exp\left[-\frac{1}{s_0}\log(1-p)\right]^{\alpha} - 1\right). \tag{62}$$

Denote as F_{n_1} the empirical *pdf* of subsample one, *i.e.*

$$F_{n_1}(x) := \frac{1}{n_1}\sum_{m=1}^{n_1}\mathbf{1}(x_{m:n_1} \le x). \tag{63}$$

The *Kolmogorov-Smirnov* distance between the empirical and theoretical *pdfs* is defined by

$$\sup_x |F_{n_1}(x) - F_X(x)|. \quad (64)$$

This distance is used as an optimization criterion for identifying an optimal value of the parameter set. Starting from the initial value $(\overline{\alpha}, \overline{s}_0, \overline{\sigma}, \overline{\mu})$, we search for

$$(\widehat{\alpha}, \widehat{s}_0, \widehat{\sigma}, \widehat{\mu}) := \underset{(\alpha, s_0, \sigma, \mu)}{\operatorname{argmin}} \left(\sup_x |F_{n_1}(x) - F_X(x)| \right) \quad (65)$$

using a local minimization algorithm.

4.3 Step three: Kolmogorov-Smirnov test

Finally, we test whether the empirical distribution of subsample two has been drawn from the theoretical *pdf* (29), say $\widehat{F}_X$, corresponding to $(\alpha, s_0, \sigma, \mu) = (\widehat{\alpha}, \widehat{s}_0, \widehat{\sigma}, \widehat{\mu})$. Since these parameters have been estimated from subsample one, which by assumption is independent from subsample two, we can now take $\widehat{F}_X$ as a completely known *pdf* for subsample two. We are therefore in a position to apply the *Kolmogorov-Smirnov* test in its standard form.

Let F_{n_2} be the empirical *pdf* of subsample two. The relevant *Kolmogorov-Smirnov* distance is now

$$\sup_x \left| F_{n_2}(x) - \widehat{F}_X(x) \right|. \quad (66)$$

Using the transformation $x = \widehat{F}_X^{-}(p)$, this is also

$$\sup_{p \in [0,1]} \left| F_{n_2}^U(p) - p \right| \quad (67)$$

where $F_{n_2}^U$ is the empirical *pdf* of an *iid* uniform sequence of size n_2 on the interval $(0, 1)$. Using this notation, it is shown [27] that

$$\sqrt{n_2} \sup_{p \in [0,1]} \left| F_{n_2}^U(p) - p \right| \underset{n_2 \uparrow \infty}{\overset{d}{\to}} M \quad (68)$$

where the variable M is the absolute supremum of a *Brownian* bridge, whose *pdf* is

$$F_M(z) = 1 - 2 \sum_{k \geq 1} (-1)^{k-1} \exp\left(-2k^2 z^2\right). \quad (69)$$

A first consequence of (68) is that, as the sample size n_2 increases, the graph of the ordered sample *versus* the quantile function should converge towards a straight line with slope one. To test whether this graph is close enough to this limit, one should search for the level value $\gamma_{n_2}(\delta)$ such that

$$P\left\{ \sup_{p \in [0,1]} \left| F_{n_2}^U(p) - p \right| > \gamma_{n_2}(\delta) \right\} = \delta \quad (70)$$

for small δ (say $\delta = 0.05$). For large n_2, we get from (68,69):

$$\gamma_{n_2}(\delta) \simeq \frac{1}{\sqrt{n_2}} \left[\frac{\log(2/\delta)}{2} \right]^{1/2}. \quad (71)$$

Thus, the *Kolmogorov-Smirnov* test works as follows: a) transform the subsample two into

$$U_1^{n_2} := \left(\widehat{F}_X\left(X_1^{(2)}\right), .., \widehat{F}_X\left(X_{n_2}^{(2)}\right) \right) \quad (72)$$

b) compute

$$\max_{m=1,..,n_2} |m/n_2 - U_{m:n_2}| = \sup_{p \in [0,1]} \left| F_{n_2}^U(p) - p \right| \quad (73)$$

where $(U_{1:n_2}, \ldots, U_{m:n_2})$ is the ordered version of $U_1^{n_2}$; c) if this number exceeds $\gamma_{n_2}(\delta)$, reject the hypothesis that the sample has been generated with the theoretical *pdf* $\widehat{F}_X$, otherwise accept it. δ is the probability to decide that the sample is not a realization of the distribution $\widehat{F}_X$ when it really is.

Note that another possible way to test the model adequacy is to use the fact that for any sequence of integers $1 \leq m_1 < m_2 < \cdots < m_k \leq n_2$, the difference between $\mathbf{X}_{m_k:n_2} := (X_{m_1:n_2}, \ldots, X_{m_k:n_2})$ and the multidimensional *qdf*

$$\widehat{\mathbf{F}}_X^{-}(\mathbf{m}_k/n) := \left(\widehat{F}_X^{-}(m_1/n_2), \ldots, \widehat{F}_X^{-}(m_k/n_2) \right) \quad (74)$$

when properly normalized, converges towards a *Gaussian* distribution as the sample size n_2 increases [7].

5 Application to Japanese and French recorded earthquake magnitudes

5.1 Modelling issues for earthquake magnitude data

In the standard *Gutenberg-Richter* model for earthquake magnitude [28], the logarithm of the probability for the magnitude to be greater than x is given by a relation in the form

$$\log_{10}(P(X > x)) = a - bx \quad (75)$$

The magnitude X is related to the seismic moment E (the amount of energy released by the earthquake) through

$$X = \frac{1}{\beta} [\log_{10}(E) - 9]. \quad (76)$$

From these two relations, one deduces that E follows a power-law with exponent $\mu = b/\beta$. For small and intermediate magnitude earthquakes, *e.g.* $0 < X < 7$ or $0 < X < 8$, statistical evidence shows that $b \simeq 1$ and $\beta \simeq 3/2$, so that $\mu \simeq 2/3 < 1$ [29]. Should this distribution be extended to large and very large magnitudes, it would require E to have infinite mean, which is considered by seismologists as physically abhorrent. One way to circumvent this drawback is to assume that the power-law model with $\mu \simeq 2/3$ holds only below some crossover energy E_c, and that a different power-law distribution with exponent $\mu' > 1$ holds for $E > E_c$ [6]. A different approach is to introduce a distribution which is close to a power-law with exponent $\mu < 1$ for intermediate values of E, while showing a significant departure from the power-law for high energies. Such distributions include the stretched exponential model [4], the parabolic fractal model [14] (which in addition has compact support), and the so-called "soft" magnitude cut-off model based on a Gamma-like distribution, first introduced by Kagan [29–32] and Main [33]. In the latter model, it is assumed that there exists a cut-off value E_c under which the power-law holds, and that above this threshold the density is in the form $Cz^{-(1+\mu)}e^{-\beta z}$, *i.e.* the power-law multiplied by an exponential roll-off at large moments. The corresponding *cpdf* decreases towards zero like $z^{-\mu}e^{-\beta z}$ for large z, so that this model enables to account for a significant departure from the pure power-law distribution. In addition, it has been shown that this distribution is the best model, in the sense of a Kullback "distance", under the two hypothesis that the Gutenberg-Richter power law holds in the absence of any condition and that one additional constraint (distribution's moment release rate) is imposed [34].

The properties of the *log-Weibull* model discussed in Section 2 can be related to these considerations by underlining that when $\alpha < 1$, E is *Von Mises* with moderately heavy tail, so that the average energy is finite, and that when $\alpha = 1$ the power-law model is recovered as a critical case, with tail index s_0. When α is less than yet close to one, the distribution is close to a power-law with exponent s_0, albeit with a *cpdf* which decreases towards zero faster than $Cz^{-(1+s_0)}$. However, the model does not incorporate an explicit cut-off value.

An important feature of earthquake catalogs is that data on small earthquakes are strongly deficient, due to incomplete, bad registration of low magnitude events. One way to address this problem would have been to consider the data as a realization of a truncated distribution; in effect, this amounts to the assumption that there exists some fixed detection threshold above which all earthquakes are recorded. Such a threshold could have been either estimated from the data or deduced from physical considerations, at the price of discarding part of the data. However, the probability for an earthquake to be detected is influenced by many factors, including purely human ones, so that the very notion of a cutoff threshold is ill-defined. A possible approach is to introduce a detection probability as an additional gaussian random variable [35]. This complicates the estimation problem by adding two parameters to the stochastic model.

Alternatively, one may simply seek to identify the distribution of recorded earthquakes, while being aware that this phenomenon results from the combination of a natural phenomenon and observational factors. This is justified by the fact that bad registration only concerns low magnitudes, while the main purpose of the modelling effort is to identify the upper tail of the distribution.

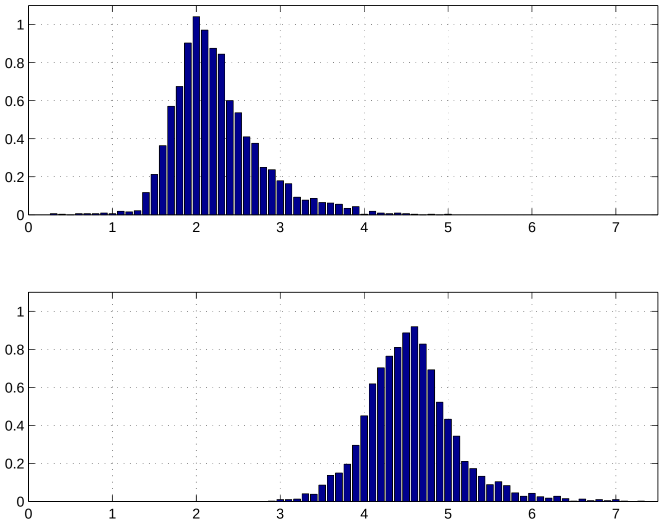

Fig. 1. Normalized histograms of magnitude data (Richter scale) for France (above) and Japan.

However, this requires that the distribution model has sufficient flexibility (in practice, a sufficient number of parameters) to take into account different combinations of lower and upper tail behaviors. In Section 2, it has been shown that the *log-Weibull* model does indeed possess such flexibility.

5.2 Fitting the log-Weibull model

The identification procedure described in Section 4 was tested on two distinct earthquake magnitude records obtained from the Northern California Earthquake Data Center (NCEDC). The first data set comprises all earthquakes recorded from November 1995 to October 1998 in a polygon corresponding roughly to the boundaries of metropolitan France, excluding Corsica (November 1995 was chosen as a starting point because prior to this date, this catalog apparently assigned a magnitude of one to all recorded small earthquakes). The second data set includes all recorded earthquakes in an area enclosing the Japanese archipelago from 1962 to February 1999. Records of earthquake magnitude are well suited to our purposes because of the logarithmic basis of the scale.

On the Richter scale, magnitude is expressed in whole numbers and decimal fractions. This round-up effect makes the raw data inconsistent with any continuous model of the probability distribution function. To overcome this minor obstacle, we regularized the data by adding to each recorded value an *iid* random noise uniformly distributed in the range $(-0.05; 0.05)$.

These two samples are of comparable sizes, containing respectively $n = 3245$ and $n = 3927$ events ranging from 0.3 to 5.0 (France) and from 2.9 to 7.4 (Japan) on the Richter scale. The corresponding normalized histograms are presented in Figure 1. Both empirical distributions

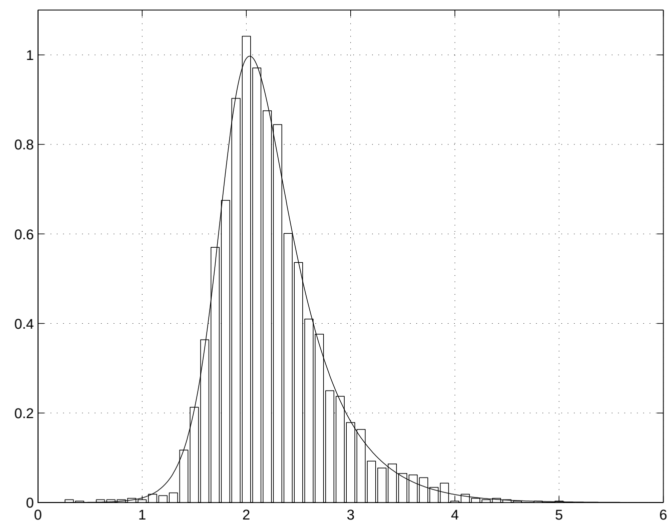

Fig. 2. Estimated *df* and normalized empirical histogram for France.

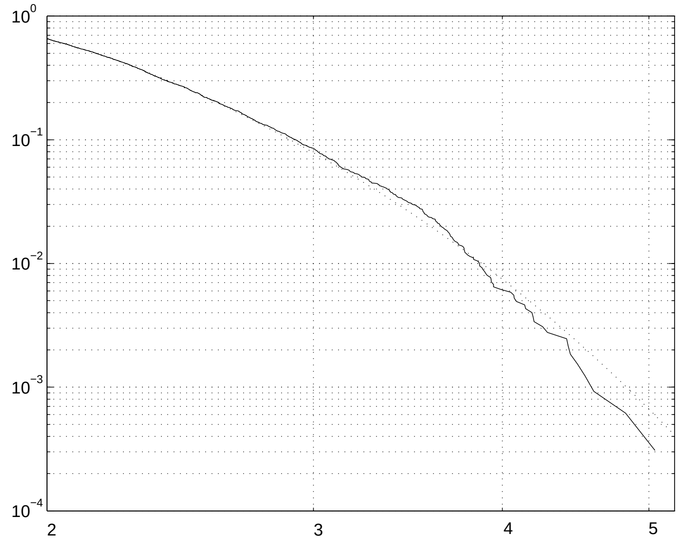

Fig. 3. Complementary probability distributions as a function of magnitude for France, log-log scale: empirical (solid line) and estimated (dotted line).

are asymmetrical. In both cases, the subsample used for parameter identification (steps one and two in Sect. 4) was obtained by taking one out of two successive events in chronological order.

Applying the first two steps of the estimation procedure in Section 4, with the choice of quantile vector $p = (0.05, 0.25, 0.75, 0.95)$ and with $n_1 = 1623$, we get for France

$$(\overline{\alpha}, \overline{s}_0, \overline{\sigma}, \overline{\mu}) = (0.848, 1.237, 0.164, 1.758) \quad (77)$$
$$(\widehat{\alpha}, \widehat{s}_0, \widehat{\sigma}, \widehat{\mu}) = (0.949, 0.331, 0.167, 1.840). \quad (78)$$

Thus, the estimated value for α is less than one, which means that the distribution tail for the underlying energy variable $E = z_0 \exp((X - \mu)/\sigma)$ is subcritical. The *Kolmogorov-Smirnov* distance for subsample two and the theoretical *pdf* $\widehat{F}_X$ obtained from subsample one is

$$\sup_x \left| F_{n_2}(x) - \widehat{F}_X(x) \right| = 9.3 \times 10^{-3} \quad (79)$$

to be compared, for the risk $\delta = 0.05$ and $n_2 = 1622$, with the level value

$$\gamma_{n_2}(\delta) \simeq \frac{1}{\sqrt{n_2}} \left[\frac{\log(2/\delta)}{2} \right]^{1/2} \simeq 3.37 \times 10^{-2}. \quad (80)$$

Therefore, the hypothesis that subsample two has been generated with the theoretical *pdf* $\widehat{F}_X$ with parameters $(\widehat{\alpha}, \widehat{s}_0, \widehat{\sigma}, \widehat{\mu})$ should be accepted. The corresponding *df* *versus* the histogram for the complete sample is presented in Figure 2. In order to assess the quality of the fit for the upper distribution tail, Figure 3 shows the estimated and empirical *cpdfs*, in log-log scale, for large magnitudes. In addition, Figure 4 shows the estimated quantile function $\widehat{F}_X^-$ *versus* the ordered complete sample.

In both Figure 3 and Figure 4, the discrepancy between estimated and empirical cumulative distributions

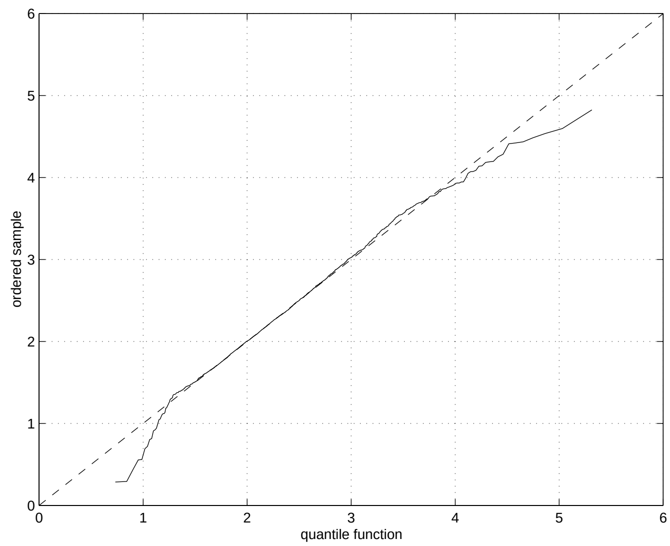

Fig. 4. Ordered sample *versus* quantile function for France.

increases as the magnitude approaches the empirical maximum, where data are sparse. One needs to test whether this discrepancy is compatible with the estimated distribution; in other words, is it likely that an *iid* sample drawn according to the estimated distribution exhibit the same fluctuations? One way to address this problem is to test whether the maximum value in subsample two is compatible with the estimated distribution, using the results in Section 3.3. In this case, the confidence interval with level $\delta = 0.05$ is

$$\left[x^*_{n_2} - \varepsilon(\delta) ; x^*_{n_2} + \varepsilon(\delta) \right] = [3.74; 6.32] \quad (81)$$

and contains the empirical maximum $x_{n_2:n_2} = 5$. This proves that the empirical maximum is statistically compatible with the estimated distribution. This procedure

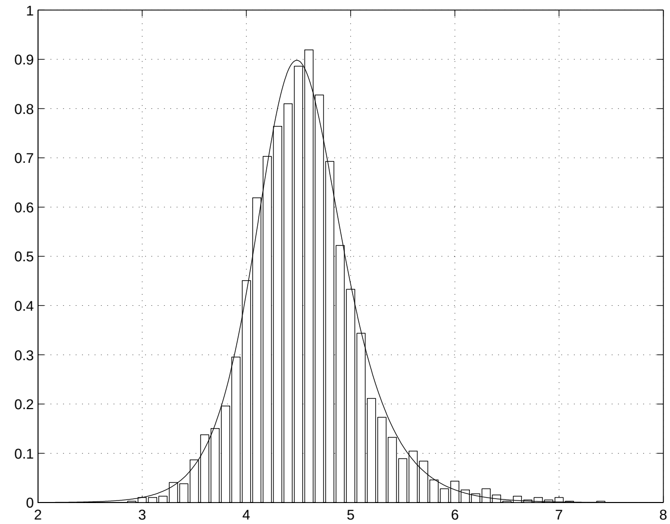

Fig. 5. Estimated *df* and normalized empirical histogram for Japan.

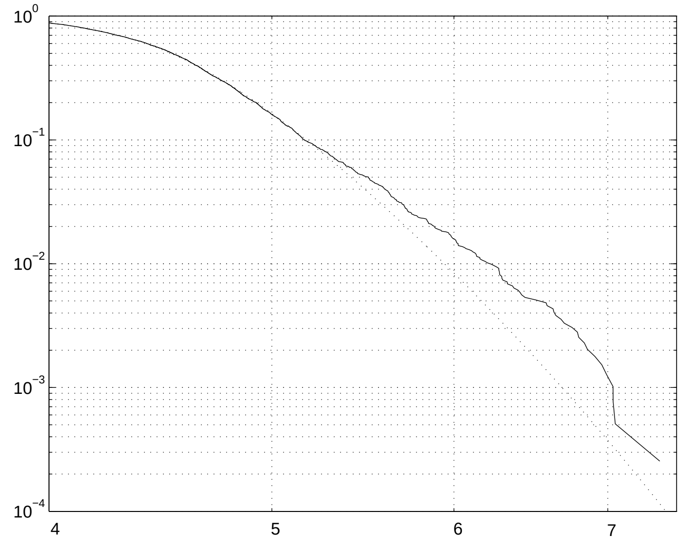

Fig. 6. Complementary probability distributions as a function of magnitude for Japan, log-log scale: empirical (solid line) and estimated (dotted line).

can be extended to test the compatibility of the $n-p$' largest value in the sample, with $n-p$ "close" to n [7].

Since the estimated value for α is close to one, one should wonder whether this conclusion really constitutes solid statistical evidence of subcriticality. While based on a very different model class, these results are consistent with the findings in [7], using the same data set.

For Japan, using the same quantile vector p and $n_1 = 1964$, we get

$$(\overline{\alpha}, \overline{s}_0, \overline{\sigma}, \overline{\mu}) = (1.088, 0.312, 0.229, 4.414) \qquad (82)$$

$$(\widehat{\alpha}, \widehat{s}_0, \widehat{\sigma}, \widehat{\mu}) = (0.962, 0.708, 0.257, 4.388). \qquad (83)$$

Here too, the estimated value for α is less than one, and close to the estimated value for France. The value of the *Kolmogorov-Smirnov* distance for subsample two, with $n_2 = 1963$, is 1.47×10^{-2}, to be compared for the risk $\delta = 0.05$ with the level value 3.07×10^{-2}. Therefore, the hypothesis that the subsample two has been generated with the theoretical *pdf* $\widehat{F}_X$ obtained from subsample one should be accepted. The corresponding *df*, estimated and empirical *cpdfs* in log-log scale and quantile function *versus* ordered sample graph are presented in Figures 5, 6 and 7. The confidence interval for the maximum also includes the maximum value in subsample two, $x_{n_2:n_2} = 7$:

$$\left[x^*_{n_2} - \varepsilon(\delta); x^*_{n_2} + \varepsilon(\delta)\right] = [5.94; 7.86]. \qquad (84)$$

As in the case of France, there is no solid statistical evidence of subcriticality. Also, since α is close to one and s_0 is close to $2/3$, the identified distribution is close to a power-law model with exponent $\mu = 2/3$.

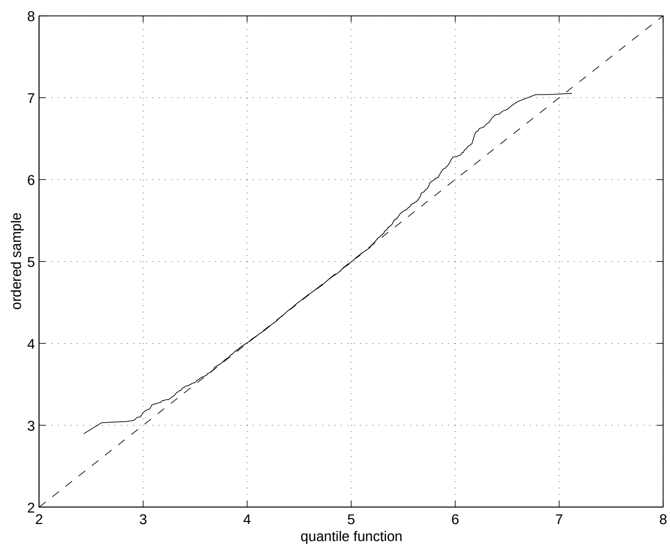

Fig. 7. Ordered sample *versus* quantile function for Japan.

5.3 Expected time until next earthquake

An obvious question of interest is: how long should one wait before the next earthquake of magnitude greater than a given level? This problem can be addressed through a detailed statistical study of the time separating two consecutive events [36]. A more pedestrian approach is to compute the "mean time between failure", which is defined as

$$N_x := \inf(n : X_n > x) \qquad (85)$$

for some magnitude level x. It is easily shown that N_x follows a geometrical distribution with parameter $F_X(x)$. Assuming that the number of earthquakes per year, say θ, is constant, equal to the total number of recorded earthquakes divided by the total collection period, the number of years one would have to wait before the next earthquake of magnitude greater than x is therefore equal to N_x/θ,

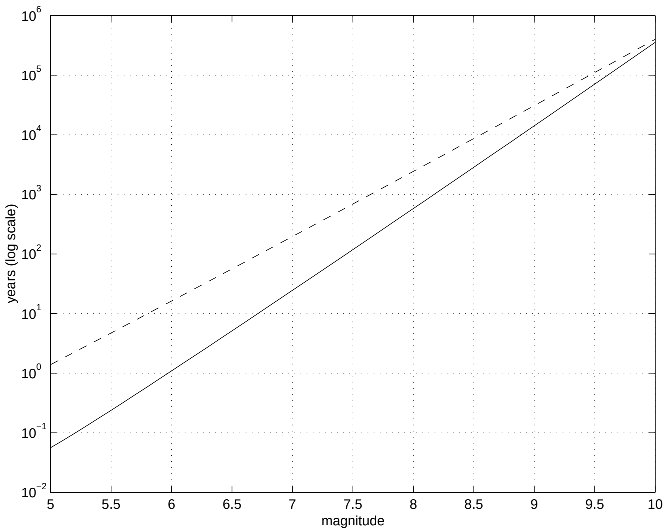

Fig. 8. Estimated average time before next earthquake of magnitude greater than x for France (dotted line) and Japan.

and can be estimated by

$$\mathbb{E}\{N_x/\theta\} = \frac{1}{\theta\left[1 - F_X(x)\right]} \cdot \qquad (86)$$

Figure 8 presents this estimated average time as a function of the magnitude in the cases of France and Japan. In order to interpret such plots, one should recall that magnitudes exceeding $x = 8$ correspond to very large and rare events (on average, one earthquake of such size occurs somewhere in the world each year). However, such forecasts should be taken *cum grano saltis*, because $F_X(x) \simeq 1$ for large x, so that the standard deviation of N_x/θ is of the same order of magnitude than its mean:

$$\sigma\{N_x/\theta\} = \frac{F_X(x)^{1/2}}{\theta\left[1 - F_X(x)\right]} \simeq \frac{1}{\theta\left[1 - F_X(x)\right]} \cdot \qquad (87)$$

The assumption that the number of earthquakes per year is constant may seem rather restrictive. One may replace this assumption by the following one, which is more realistic: the (random) number of events occurring on time interval $[0;t)$, say N_t, is a *Poisson* variable with rate $\overline{\theta}$. When t is measured in years, $\overline{\theta}$ is the average number of events per year. Under this assumption, the exact distribution function for the random variable $\max(0, X_1, \ldots, X_{N_t})$ is given for $x > 0$ by

$$\begin{aligned} & P\left(\max\left(0, X_1, \ldots, X_{N_t}\right) \le x\right) = \\ & \sum_{n \ge 0} P\left(\max\left(0, X_1, \ldots, X_n\right) \le x\right) P\left(N_t = n\right) = \\ & \sum_{n \ge 0} \left[F_X(x)\right]^n \frac{\exp\left(-\overline{\theta}t\right)\left(\overline{\theta}t\right)^n}{n!} = \\ & \exp\left[-\overline{\theta}t\left(1 - F_X(x)\right)\right]. \end{aligned} \qquad (88)$$

The event "$\max(0, X_1, \ldots, X_{N_t}) \le x$" coincides with the event "$T_x > t$", where T_x is the first time when some X exceeds the level x. From (88), T_x is exponentially distributed, with mean and standard deviation equal to

$$\mathbb{E}\{T_x\} = \sigma\{T_x\} = \frac{1}{\overline{\theta}\left[1 - F_X(x)\right]} \cdot \qquad (89)$$

Thus, if $\overline{\theta}$ is estimated by the average number of events per year in the sample, this result is consistent with (86,87).

The authors are grateful to the Northern California Earthquake Data Center (NCEDC) and to the Members Networks of the Council of the National Seismic System (CNSS) for supplying the raw data on earthquakes magnitudes through their internet `http` site

`http : //quake.geo.berkeley.edu/cnss/cnss.catalog.html`

They would also like to thank the referees for their competent scrutiny of the initial version of this article, their constructive remarks, and for providing additional useful references.

References

1. V. Pareto, *Cours d'Économie politique* (1896), reprinted as a volume of *Oeuvres complètes* (Droz, Geneva, 1965).
2. B.B. Mandelbrot, *The Fractal geometry of nature*, edited by W.H. Freeman (New York, 1983).
3. *Fractals in physics*, edited by A. Aharony, J. Feder, Physica D, **38**, 1-3 (1989).
4. J. Laherrère, D. Sornette, Eur. Phys. J. B, **2**, 525-539 (1998).
5. L. Knopoff, D. Sornette, J. Phys I France **5**, 1681 (1995).
6. D. Sornette, L. Knopoff, Y.Y. Kagan, C. Vanneste, J. Geophys. Res. **101** (B6), 13883 (1996).
7. T. Huillet, H.-F. Raynaud, J. Phys. A: Math. Gen. **32**, 1099 (1999).
8. J.-P. Bouchaud, M. Mézard, J. Phys. A: Math. Gen. **30**, 7997 (1997).
9. J. Galambos, *The Asymptotic theory of Extreme Order Statistics* (John Wiley and Sons, 1978).
10. P. Embrechts, C. Klüppelberg, T. Mikosh, *Modelling extremal events* (Springer-Verlag, Applications of mathematics **33,** 1997).
11. J. Aitchinson, J.A.C. Brown, *The log-normal distribution* (Cambridge University Press, 1957).
12. J. Soderlund, L.B. Kiss, G.A. Niklasson, C.G. Granqvist, Phys. Rev. Lett. (to be published).
13. P. Burchard, *Statistical properties of WWW*, `http://www.concentric.net/ ~Aleph0/www/stats` (1999).
14. J. Laherrère, *Distributions de type fractal parabolique dans la Nature* (*CRAS*, Série II a Paris), **322**, 535 (1996).
15. C. Barton, *A new approach to estimating hydrocarbon resources*, U.S. Geological Survey, http://energy.usgs.gov/factsheets/HydroRes/estimat.htm, 1995.

16. A. Arneodo, S. Manneville, J.-F. Muzy, Eur. Phys. J. B **1**, 129-140 (1998).
17. W. Feller, *An introduction to probability theory and its applications* **2** (Wiley, New York, 1971).
18. G.K. Zipf, *Human behaviour and the Principle of Least Effort* (Addison-Wesley, 1949).
19. B.B. Mandelbrot, *Fractales, hasard et finance* (Flammarion, Paris 1997).
20. W.E. Pruitt, Ann. of Probability **15**, 885 (1987).
21. R.A. Maller, S.I. Resnick, Proc. London Math. Soc. **49**, 385 (1984).
22. U. Frisch, D. Sornette, J. Phys. I France **7**, 1155 (1997).
23. H.A. David, *Order statistics,* 2nd edn (Wiley, New York, 1981).
24. E.J. Gumbel, *Statistics of Extremes* (Columbia University Press, Columbia 1958).
25. M.A. Stephens, Jour. of Am. Stat. Assoc. **69-347**, Theory and Methods Section, 730 (1974).
26. M.A. Stephens, *Introduction to Kolmogorov (1933) "On the empirical determination of a distribution"*, in *Breakthroughs in statistics* **2**, edited by S. Kotz, N.L. Johnson (Springer-Verlag, New York, 1992) pp. 93-105.
27. P. Billingsley, *Convergence of probability measures,* (John Wiley and Sons, 1968).
28. B. Gutenberg, C.F. Richter, Bull. Seism. Soc. Am. **34**, 185 (1944).
29. Y.Y. Kagan, Physica D **77**, 160 (1994).
30. Y.Y. Kagan, Bull. Seism. Soc. Am. **83**, 7 (1993).
31. Y.Y. Kagan, Bull. Seism. Soc. Am. **86**, 274 (1996).
32. Y.Y. Kagan, J. Geophys. Res. **102**, 2835 (1997).
33. I. Main, Rev. of Geophysics **34**, 433 (1996).
34. S. Sornette, A. Sornette (to appear in Bull. Seism. Soc. Am.).
35. Y. Ogata, K. Katsura, Geophys. J. Int. **113**, 727 (1993).
36. S. Sornette, L. Knopoff, Bull. Seism. Soc. Am. **87**, 789-798, (1997).