Multiscaling comparative analysis of time series and geophysical phenomena

Nicola Scafetta[1] and Bruce J. West[1,2]

[1] *Physics Department, Duke University, Durham, NC 27708 and*

[2] *Mathematics Division, Army Research Office,*

Research Triangle Park, NC 27709.

(Dated: April 19, 2022)

Abstract

Different methods are used to determine the scaling exponents associated with a time series describing a complex dynamical process, such as those observed in geophysical systems. Many of these methods are based on the numerical evaluation of the variance of a diffusion process whose step increments are generated by the data. An alternative method focuses on the direct evaluation of the scaling coefficient of the Shannon entropy of the same diffusion distribution. The combined use of these methods can efficiently distinguish between fractal Gaussian and Lévy-walk time series and help to discern between alternative underling complex dynamics.

The evaluation of the scaling exponents is of fundamental importance to describe a number of complex systems [1, 2, 3]. The mathematical definition of scaling is as follows [4]. The function $\Phi(r_1, r_2, \ldots)$ is termed scaling invariant, if it fulfills the property:

$$\Phi(r_1, r_2, \ldots) = \gamma^a \ \Phi(\gamma^b r_1, \gamma^c r_2, \ldots) \ . \tag{1}$$

Thus, if we scale all coordinates $\{r_i\}$ by means of an appropriate choice of the exponents $a, b, c....$, then we always recover the same function. This scaling invariance is the basic property that characterizes fractal functions [1, 2, 3]. The theoretical and experimental search for the correct scaling exponents is intimately related to the discovery of deviations from ordinary statistical mechanics. Fractal time series are particularly important in geophysics, as well as in several other field of research including biophysics and econophysics, where the phenomena of interest may present specific self-similarity patterns on different time scales.

Two methods of analysis of time series commonly used to determine scaling properties are autocorrelation analysis and power spectral analysis [5]. The autocorrelation function of the fractal noise $\{\xi_i\}$ results in the relation

$$C(r) = \frac{< \xi_i \ \xi_{i+r} >}{< \xi_i^2 >} \propto r^{2H-2} \ . \tag{2}$$

The power spectral representation of the same scaling property reads:

$$S(f) = \int_{-\infty}^{\infty} C(r) \ e^{-i2\pi fr} dr \propto f^{1-2H} \ . \tag{3}$$

It is easy to recognize the self-similarity or scaling property of the above two equations in their power-law form. The scaling exponent H was called the *Hurst exponent* by Mandelbrot [1] in honor of the civil engineer Hurst who first understood the importance of scaling laws to describe the long-range memory in time series. In particular, Hurst was interested in evaluating the strength of the persistence of the annual level of the floods of the Nile river and, for such a scope, developed a time series analysis method to determine the scaling parameter H [6].

A value $H = 1$ corresponds to 1/f-noise or *pink* noise. The adoption of a color name "pink" derives from the fact that a light source characterized by a 1/f spectrum looks pink. These type of noises are particular important because they represent a kind of perfect balance between randomness and order, or between unpredictability and predictability. In fact, for pink noises the autocorrelation function between two events separated by a time interval

$\Delta \tau = r$ is independent on r, ($C(r) \approx const$). Pink noises, $H \approx 1$, are found in countless natural phenomena from heart-beat intervals to music [7]. A value $0 < H < 0.5$ corresponds to antipersistent noise, $H = 0.5$ corresponding to uncorrelated or random noise, also known as *white* noise, and $0.5 < H < 1$ corresponds to correlated or persistent noise. It is possible to extend the definition of H for values larger than 1. So, a value $H = 1.5$ corresponds to Brownian motion, which, as it is well known, describes the erratic motion of a particle, such as a pollen grain, in suspension on a fluid; this erratic motion is cause by random collisions between the particle and the molecules of the fluid [8]. A value $H = 2$ corresponds to *brown* noise and a value $H > 2$ is known as *black* noise. These noises are characterized by a very smooth shape and may be adopted, for example, to generate artificial mountain landscapes [7].

It is important to point out that there are two common alternative deviations from ordinary statistical mechanics: anomalous Gaussian statistics and Lévy statistics [9]. These two different statistics are indicative, in particular, of two different kind of complex noises: the monofractal Gaussian noise [1] and the Lévy-walk noise [10]. These two types of noises present similar long-range correlation patterns, but are generated by quite different complex dynamics. The monofractal Gaussian noise, in its persistent form, presents long range memory in the sense that future events are strongly related to the frequency of occurrence of past events and the waiting time distribution between events has finite variance. The Lévy-walk intermittent noise, instead, presents long-range correlation patterns which are generated by random waiting time intervals between events obeying to an inverse power law distribution with exponents that yield infinite variance, and there is no real correlation between events [11]. Figs. 1 show examples of these noises.

Herein, we briefly describe two alternative time series scaling analysis methods, whose combined adoption can be used to distinguish the two above alternative noises. The diffusion entropy analysis (DEA) and standard deviation analysis (SDA) [12]. Both techniques are based on the prescription that a time series $\{\xi_i\}$ of N elements are the fluctuations of a diffusion trajectory [12]. Note that there exist several other scaling analysis methods such as the detrended fluctuation analysis [13] and several wavelet based methods [14, 15, 16], which are variance based methods and are theoretically equivalent to the SDA.

According to the prescription of Scafetta and Grigolini [12], we shift our attention from the time series $\{\xi_i\}$ to the probability distribution function (pdf) $p(x, t)$ of the corresponding
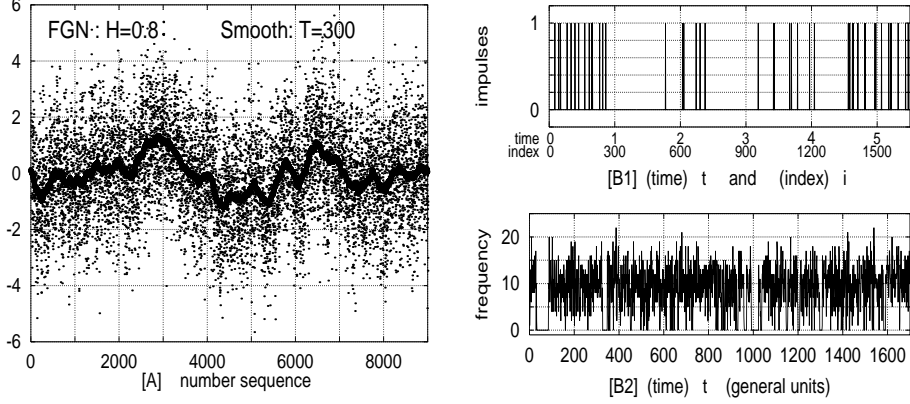
FIG. 1: [A] Fractional Gaussian noise with $H = 0.8$; [B] Two forms of *Lévy-walk intermittent noise* with $\psi(\tau) \propto \tau^{-\mu}$ and $\mu = 2.5$. B2 gives the frequency of impulses every 300 units of B1.

diffusion process, that is, the pdf of the diffusion process, $p(x,t)$, is evaluated by means of the $N - t$ sub-trajectories

$$x_n(t) = \sum_{i=0}^{t} \xi_{i+n} \qquad (4)$$

with $n = 0, 1, \ldots$ Therefore, x denotes the variable collecting the fluctuations and is referred to as the diffusion variable. The scaling property of the diffusion process, if it exists, takes the form

$$p(x,t) = \frac{1}{t^\delta} \, F\left(\frac{x}{t^\delta}\right) \;, \qquad (5)$$

where δ is the scaling exponent. The DEA [12] is based on the evaluation of the Shannon entropy $S(t)$ using the pdf (5). If the scaling condition of Eq. (5) holds true, it is easy to prove that

$$S(t) = -\int p(x,t) \ln[p(x,t)]dx = A + \delta \, \ln(t) \;, \qquad (6)$$

where, A is a constant. Numerically, we first evaluate the pdf with histogram of size-bin equal to the standard deviation of the data, and then use a discrete form of Eq. (6).

The SDA [12] is based on the evaluation of the standard deviation $D(t)$ of the same variable x, and yields

$$D(t) = \sqrt{\frac{\sum_{n=0}^{N-t} [x_n(t) - \langle x;t \rangle]^2}{N - t - 1}} \propto t^H, \qquad (7)$$

where $\langle x;t \rangle = \frac{1}{N-t} \sum_{n=0}^{N-t} x_n(t)$ is the mean value of $\{x_n(t)\}$, and H is the Hurst exponent.

If the data are fractal Gaussian noise the two scaling exponents are related to each other via the fractal Gaussian relation
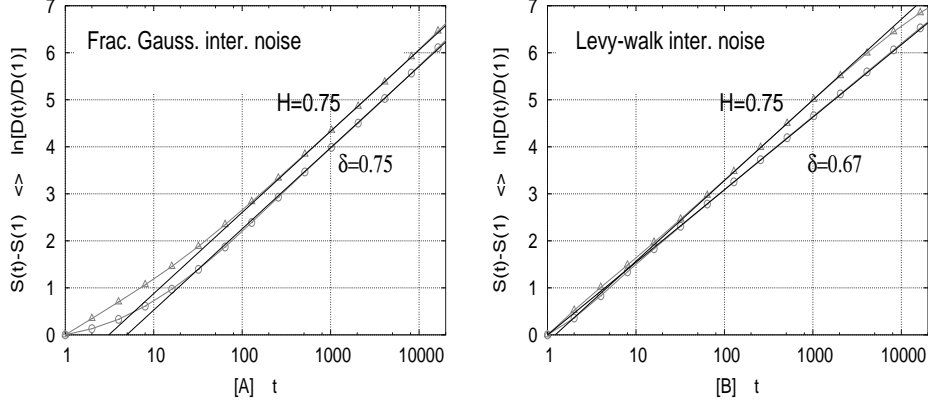
$$H = \delta \;. \qquad (8)$$

4

FIG. 2: DEA end SDA of: [A] a *fractal Gaussian intermittent noise* with $\psi(\tau) \propto \exp(-\tau/\gamma)$ with $\gamma = 25$ and $H = \delta = 0.75$; the fractal Gaussian relation (8) of equal exponents is fulfilled; [B] a *Lévy-walk intermittent noise* with $\psi(\tau) \propto \tau^{-\mu}$ and $\mu = 2.5$; note the bifurcation between $H = 0.75$ and $\delta = 0.67$ in accordance with the Lévy-walk relation (10).

If the data are generated by a Lévy-walk they are characterized by an inverse power law waiting time distribution of the type

$$\psi(\tau) \propto \frac{1}{(1+\tau)^\mu} \ , \tag{9}$$

where $2 < \mu < 3$, which ensures that although the first moment of τ is finite, the second moment diverges. The scaling exponents are related to each other via the Lévy-walk relation (LWR) [11]

$$0.5 < \delta = \frac{1}{3-2H} = \frac{1}{\mu-1} < H < 1. \tag{10}$$

There are several complex ways to generate a Lévy-walk sequence, see Ref. [11, 12, 17, 19]. Some of these noises involve mixed Lévy-Gaussian properties. The simplest Lévy-walk sequence is a dichotomous signal made by a series of zeros and ones, where $\xi = 1$ represents the occurrence of an event and $\xi = 0$ represents no event. The time intervals $\{\tau_i\}$ obeying to Eq. (9) give the intervals between events.

Thus, by evaluating δ and H and using Eq. (8) and (10) it is possible to distinguish the two kinds of time series [12], while the adoption of only one of the two techniques can lead to a misinterpretation of the characteristics of a phenomenon. Figs. 2A and 2B show DEA and SDA applied to a fractal Gaussian intermittent noise with $\delta = H = 0.75$ and to a Lévy walk intermittent noise with $\mu = 2.5$, which correspond to $\delta = 0.67$ and $H = 0.75$, respectively [11, 12].
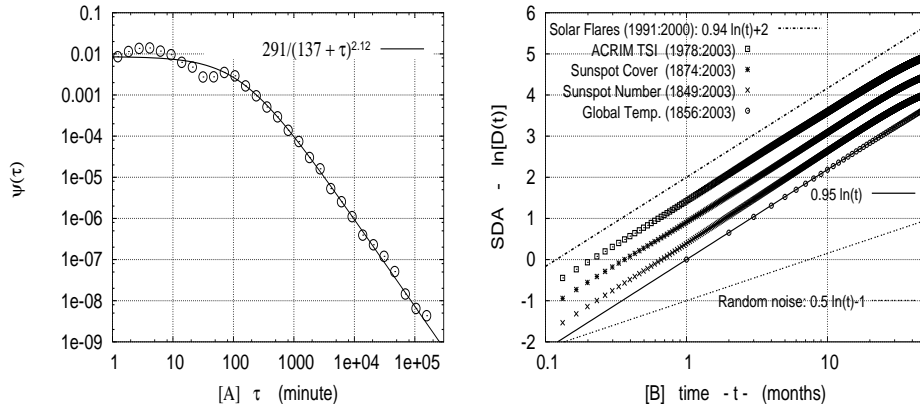
FIG. 3: [A] Waiting time distribution of flares. The distribution is fit with an inverse power law Eq. (9) with exponent $\mu = 2.12 \pm 0.05$ [Grigolini *et al.* 2002]. [B] SDA applied to the ACRIM composite TSI time series, sunspot number and cover sequence and global temperature anomalies. The uppermost line represents the theoretical Lévy-walk scaling of the solar flare intermittency, $H_T = 0.94 \pm 0.02$, obtained via Eq. (10) with $\mu = 2.12 \pm 0.05$. The bottom line shows the scaling for random noise, $H = 0.5$, for comparison.

Particularly interesting applications of the above scaling analysis techniques are found in geophysical phenomena such as earthquakes, solar flares and global temperature patterns, where long time series of data are available [11, 17, 18]. Herein, we briefly summarize some of our findings.

In Ref. [17] it was shown that the waiting time interval distribution $\psi(\tau)$ between solar flares [20] is an inverse power law of the type (9) with exponent $\mu = 2.12 \pm 0.05$. According to the LWR (10) this would induce a Levy-walk with theoretical exponents $H = 0.94 \pm 0.04$ and $\delta = 0.89 \pm 0.04$. Fig. 3A shows the waiting time distribution of flares. Fig. 3B shows the SDA applied to several solar data such as ACRIM TSI [21], sunspot cover [23] , sunspot number [22], global surface temperature anomalies (1856-2003) [24] and the theoretical prediction derived from the solar flare intermittency. The curves are quite parallel and suggest that a Lévy like process regulate the dynamics of the solar activity and that the Earth climate seems to contain the same statistics. The latter statement seems further confirmed by Figs. 4A and 4B that show DEA and SDA applied to the ACRIM TSI and the global temperature record showing the typical LWR bifurcation. We observe that if these findings are not accidental, they might imply the existence of a non-negligible complex Sun-Climate nonlinear coupling on a short time-scale as some studies seem to confirm [25, 26].
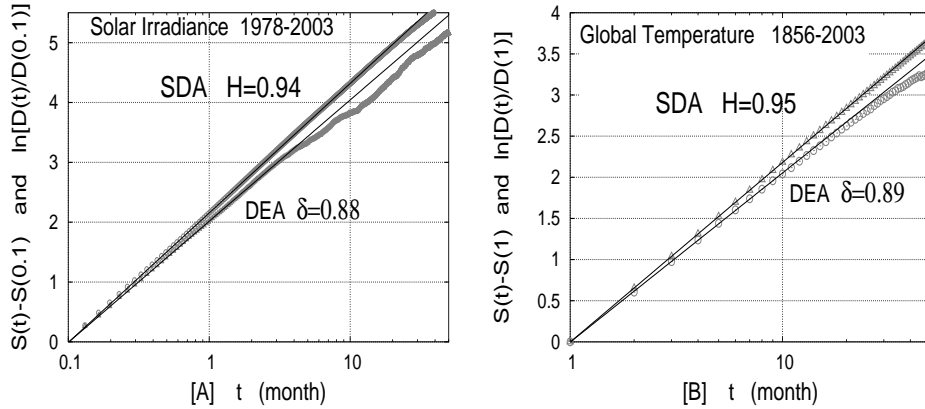
FIG. 4: [A] DEA and SDA applied to the ACRIM composite TSI time series. The two straight lines correspond to the scaling coefficients $\delta = 0.88 \pm 0.02$ and $H = 0.94 \pm 0.02$. [B] DEA and SDA applied to the global temperature anomalies (1856-2003) time series. The two straight lines correspond to the scaling coefficients $\delta = 0.89 \pm 0.02$ and $H = 0.95 \pm 0.02$.

The above techniques can be applied also to earthquake concurrence [11]. An issue about seismic phenomena is whether: (1) they obey a statistics according to which the waiting times between Omori's earthquake clusters [29] are uncorrelated from one another, as the traditional Generalized Poisson model [27, 28], the "ETAS" model [30]; (2) the Omori's earthquake clusters obey to some knind of Lévy-walk statistics [28]; (3) or whether the data may also be characterized by intercluster 1/f long-range correlations between Omori's clusters that may disclose the *earthquake conversations* recently suggested by Stein [27]. Understanding the nature of the long-range correlation is fundamental for building reliable earthquake models.

Fig. 5A shows the waiting time PDFs between earthquakes in California [31] using four earthquake magnitude thresholds $M_t = 1, 2, 3$ and 4. The PDFs show an initial Omori's law [29] ($P(\tau) \propto 1/\tau$), but the pdf tails present a large inverse power law exponent $\mu > 4$ and may even approach an exponential (or Poisson) distribution asymptotically. The Omori's law is determined by the short-range correlated aftershocks [29] and lasts for a time that increases with the magnitude threshold. Fig. 5B shows the DEA and SDA applied to the intermittent time signal $\xi(t)$, where t is the physical time, obtained by assigning a value equal to 1 at the occurrence of an event, and a value equal to 0 when no event occurred. The latter figure suggests that the data fulfill FGR (8). Thus, beyond the Omori's law, the earthquake clusters might be uncorrelated if the observed super-diffusion $\delta = H = 0.94$ is
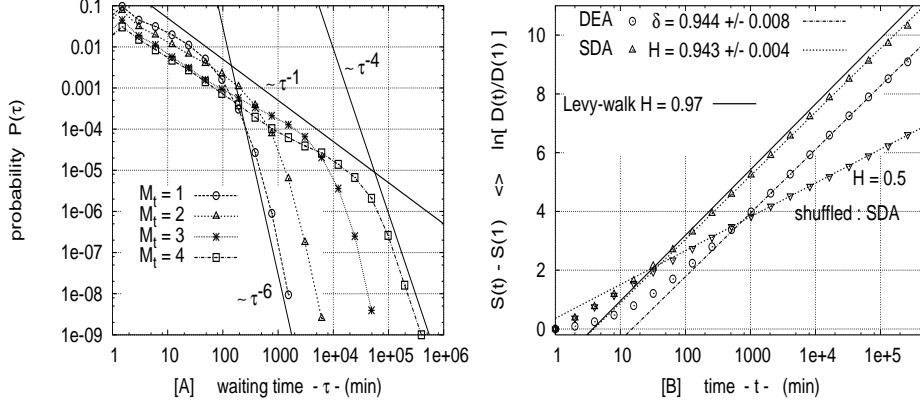
FIG. 5: [A] pdf of the waiting times τ_i of earthquakes with a magnitude $M \geq M_t = 1, 2, 3$ and 4. The initial $P(\tau) \propto 1/\tau$ is the Omori's law [29]. [B] DEA and SDA of the intermittent time signal $\xi(t)$ for magnitude $M \geq M_t = 1$. The data are fitted with scaling exponents $\delta = 0.944 \pm 0.008$ and $H = 0.943 \pm 0.004$. The uppermost solid line with $H = 0.97$ corresponds to the expectation of H if the Levy-walk condition (10) holds true.

generated by a long-tailed Omori's law involving multiple clusters, or there might be the possibility that the clusters are correlated as a 1/f Gaussian noise [11]. In the latter case, traditional earthquake models such as the Generalized Poisson model [27, 28] or the "ETAS" model [30] should be improved by adding additional correlations between clusters.

In conclusion, we have discussed some properties of complex time series analysis showing two different types of anomalous statistics: fractal Gaussian noise and Lévy-walk noise. We have shown how the multiscaling comparative analysis of time series can be used to distinguish the two types of noises and applied it to study some complex patterns of geophysical phenomena. Thus, we conclude that there are some difficulties in interpreting intermittent sequences. Models with alternative statistics can reproduce some patterns of a time series equally well. This fact suggests the need of an analysis involving complementary tests for addressing complex systems.

[1] B.B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, New York, (1983).

[2] Feders J., *Fractals*, Plenum Publishers, New York, (1988).

[3] Peitgen H.-O., J. Hartmut, D. Saupe, *Chaos and Fractals, new frontiers of science*, sec. edition, Springer, New York, (2004).

[4] Goldenfeld N., *Lectures on Phase Transitions and the Renormalization Group* (Perseus Book, Reading, Massachusetts,1985).

[5] R. Badii and A. Politi, *Complexity, Hierarchical structures and scaling in physics*, Cambrige University Press, UK, (1997).

[6] Hurst H.E., R. P. Black, Y.M. Simaika, *LongTerm Storage: An Experimental Study*, Constable, London, (1965).

[7] Schroeder M.,*Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*, W.H. Freeman & Company,(1992).

[8] Gardiner C.W., *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 2nd edition, Springer-Verlag, New York, New York, (1997).

[9] A.I. Khinchin, *Mathematical Foundations of Statistical Mechanics*, Dover Publications, Inc. New York, (1949).

[10] Shlesinger M.F., B.J. West, J. Klafter, "Lévy dynamics of enhanced diffusion: Application to turbulence," Phys. Rev. Lett. **58**, 1100 (1987).

[11] Scafetta N., and B.J. West, "Multi-scaling comparative analysis of time series and a discussion on earthquake conversations in California," Phys. Rev. Lett. **92** 138501 (2004).

[12] Scafetta N., and P. Grigolini, "Scaling detection in time series: diffusion entropy analysis," Phys. Rev. E **66**, 036130, (2002).

[13] Peng C.-K., S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, and A.L. Goldberger, "Mosaic organization of DNA nucleotides," Phys. Rev, E **49**, 1685, (1994).

[14] Mallat S.G., *A Wavelet Tour of Signal Processing* (2nd edition), Academic Press, Cambridge (1999).

[15] Muzy J.F., E. Bacry, A. Arneodo, "The Multifractal Formalism Revisited with Wavelets," Int. J. Bifurc. Chaos **4**, No. 2, 245-302 (1994).

[16] Percival D.B., and A.T. Walden, *Wavelet Methods for Time Series Analysis*, Cambrige University Press, Cambrige (2000).

[17] Grigolini P., D. Leddon, N. Scafetta, The Diffusion entropy and waiting time statistics of hard x-ray solar flares, Phys. Rev. E **65**, 046203 (2002).

[18] Scafetta N., and B.J. West, Solar Flare Intermittency and the Earth's Temperature Anomalies,

Phys. Rev. Lett. **90**, 248701 (2003).

[19] Scafetta N., V. Latora and P. Grigolini, "Scaling without detrending: the diffusion entropy method applied to the DNA sequences," Phys. Rev. E **66**, 031906 (2002).

[20] Solar Flare Catalog, (2003). *http://umbra.nascom.nasa.gov*

[21] Willson R.C., and A.V. Mordvinov , "Secular total solar irradiance trend during solar cycles 21-23," Geophys. Res. Lett., **30**, 1199, doi: 10.1029/2002GL016038 (2003). *http://www.acrim.com*

[22] Solar Influences Data analysis Center (2003). *http://sidc.oma.be/index.php3*

[23] Royal Greenwich Observatory/USAF/NOAA Sunspot Record 1874-2003 (2003). *http://science.msfc.nasa.gov/ssl/pad/solar/greenwch.htm*

[24] Climatic Research Unit, UK, (2003) *http://www.cru.uea.ac.uk.*

[25] D'Aleo Joe, "Solar influence may be evident in three of the last four seasons," (2002). *http://www.intellicast.com/DrDewpoint/Library/1339*

[26] Jackman C.H, and R.D. McPeters, "The effect of solar proton events on ozone and other constituents," Chapter in *Solar Variability and its Effects on Climate* by Pap J.M, and P. Fox, Geophysical Monograph Series Volume 141 (2004).

[27] Stein R.S., "Earthquake Conversation," Scientific American, **288** Jan 72 (2003).

[28] Mega M.S., P. Allegrini, P. Grigolini, V. Latora, L. Palatella, A. Rapisarda and S. Vinciguerra, "Power-Law Time Distribution of Large Earthuakes," Phys. Rev. Lett. **90** 188501 (2003).

[29] Omori F., "On the aftershocks of earthquakes," J. College Sci. Imper. Univ. Tokyo **7**, 111 (1895). P. Bak, K. Christensen, L. Danon and T. Scanlon, "Unified Scaling Law for Earthquakes," Phys. Rev. Lett. **88** 178501 (2002).

[30] Kagan Y.Y., and L. Knopoff, "Statistical short-term earthquake prediction," Science **236**, 1563 (1987).

[31] Southern California Earthquake Data Center (2003). http://www.scecdc.scec.org/ftp/catalogs/SCNS/