Statistics of Extreme Values in Time Series with Intermediate-Term Correlations

Cecilia Pennetta
Dipartimento di Ingegneria dell’Innovazione, Università del Salento and CNISM,
Via Arnesano, Edificio "Stecca", I-73100 Lecce, Italy

It will be discussed the statistics of the extreme values in time series characterized by finite-term correlations with non-exponential decay. Precisely, it will be considered the results of numerical analyses concerning the return intervals of extreme values of the fluctuations of resistance and defect-fraction displayed by a resistor with granular structure in a nonequilibrium stationary state. The resistance and defect-fraction are calculated as a function of time by Monte Carlo simulations using a resistor network approach. It will be shown that when the auto-correlation function of the fluctuations displays a non-exponential and non-power-law decay, the distribution of the return intervals of extreme values is a stretched exponential, with exponent largely independent of the threshold. Recently, a stretched exponential distribution of the return intervals of extreme values has been identified in long-term correlated time series by Bunde et al. (2003) and Altmann and Kantz (2005). Thus, the present results show that the stretched exponential distribution of the return intervals is not an exclusive feature of long-term correlated time series.

PACS numbers: 05.40.-a, 05.45.Tp, 02.50.-r

Keywords: Extreme values, Fluctuation phenomena, stochastic processes, Time series analysis

I. INTRODUCTION

The puzzling questions posed by the observations of climate changes have been the subject of further motivations to the investigation of extreme events and dynamics. In particular, the statistics of the return times (RTs) of extreme values is one of the subjects that have recently attracted the attentions of many authors. Return times (or intervals) $r_q$, of the extreme values of a fluctuating quantity $x(t)$, associated with the overcoming of a given threshold $q$, can be defined as the time intervals between two consecutive occurrences of the condition $x(t) > q$. In the case of uncorrelated events it is easy to show that the RTs are exponentially distributed. Exponential distributions describe the statistics of return intervals of prices in financial markets, wind speed data or daily precipitations in a given place for the same time windows.

On the other hand, in recent years it has been realized that several other important examples of time series display long-term correlations. This is the case of physiological data (heartbeats and neuron spikes), hydro-meteorological records (daily temperatures, geophysical or astrophysical data (occurrence of earthquakes or solar flares), internet traffic and stock market volatility records. Long-term correlated time series are characterized by an auto-correlation function, $C_x(s)$, decaying as a power-law:

$$C_x(s) = \langle x_i x_{i+s} \rangle = \frac{1}{N-s} \sum_{i=1}^{N-s} x_i x_{i+s} \sim s^{-\gamma_x}$$

where the correlation exponent $\gamma_x$ ranges between 0 and 1 (normalized series $\{x_i\}_{i=1,...,N}$ with zero average and unit variance are considered). Therefore, in these conditions, the mean correlation time $\tau$ diverges ($\tau$ is defined as the integral over $s$ of $C_x$). The effect of long-term correlations on the RT statistics has been recently studied by Bunde et al. and by Altmann and Kantz. Both these studies were performed on the ground of numerical calculations on Gaussian and long-term correlated time series generated by the Fourier transform technique and by imposing a power-law decay of the power spectrum. The main conclusions of Bunde et al. were the following: i) The mean return interval $R_q$ is unchanged by the presence of long-term correlations. ii) The distribution of the RTs becomes stretched exponential:

$$P_q(r) = a/R_q \exp[-(b/rR_q)\gamma]$$

where the exponents $\gamma$ and $\gamma_x$ were found to be equal. iii) The RTs themselves are long-term correlated with a correlation exponent $\gamma' \approx \gamma_x$. Actually, as noted by Altmann and Kantz, the statement i) is true in general and can be identified with Kac’s lemma introduced in the context of dynamical systems while the statement iii) is true only in very special conditions. Moreover, it must be noted that the results ii) and iii) only apply to linear time series (i.e. to series whose properties are completely defined by the power spectrum and by the probability density). However, apart from this restriction, the stretched exponential distribution of the RTs can be considered a general feature in presence of long-term correlations in a time series. This feature has important consequences on the observation of extreme events: indeed it implies a strong enhancement of the probability of having return intervals well below and well above $R_q$, in comparison with the occurrence of extreme events in an uncorrelated time series. Furthermore, it must be noted that the distribution in Eq. (2) only depends on the parameter $\gamma$, being $a$ and $b$ function of $\gamma$. 
Here it will be considered the effect on the distribution of the RTs of the presence of finite-term correlations with non-exponential decay, a situation which can occur in systems which are approaching criticality, where intermediate behaviors with non-exponential and non-power-law decay of correlations can emerge \cite{11}. Thus, in the next section it will be analyzed the RT statistics of extreme values of the resistance and defect-fraction fluctuations displayed by a resistor with granular structure in a nonequilibrium stationary state \cite{30}.

II. METHOD AND RESULTS

The time series analyzed consist in the fluctuations of resistance and defect-fraction of a thin resistor with granular structure, biased by an external current $I$ and in contact with a thermal bath at temperature $T$. The resistance values are calculated by using the multi-species resistor network (MSN) model \cite{30}. The MSN model describes a conducting film with granular structure as a resistor network made by different species of elementary resistors. Precisely, the resistor species are distinguished by their resistances and by their energies associated with thermally activated stochastic processes of breaking and recovery of the elementary resistors. The broken elementary resistors, named defects, correspond to resistors of resistance about $10^9$ higher than that of normal resistors. As a result of the competition between pairs of breaking and recovery processes for each species, the network can reach a stationary state whose properties will be dependent on the external conditions. The resistance and the defect fraction of the network are calculated by Monte Carlo simulations as a function of the time for different temperatures \cite{30}. All the details about the MSN model and its results can be found in Ref. \cite{30}. It should be noted, however, that this model provides a correlation time of the resistance fluctuations dependent on the temperature: $\tau \sim (T - T^*)^{-\theta}$ with $\theta = 2.7$, thus diverging for $T \to T^*$. As a result, the power spectral density of the resistance fluctuations shows a $1/f^\alpha$ dependence of frequency, with $\alpha \approx 1$ for $T \leq T^*$. It should be also noted that the MSN model represents an extension to systems characterized by $1/f$ noise of a previous model existing in the literature, the stationary and biased resistor network (SBRN) model \cite{31,32,33}. In any case, it must be underlined that, apart from the specific system described by the MSN model, the method used here for generating the time series can be also viewed as a pure numerical algorithm for generating numerical series with different and tunable correlation properties.

Long $R(t)$ and $p(t)$ time series (typically made of $1 \pm 2 \times 10^6$ records) have been generated and analyzed for different temperatures. Here $R$ indicates the resistance of the network expressed in Ohm and $p$ the defect-fraction, while the time $t$ is expressed in iteration steps. The analysis has been performed by considering normalized series: $x(t) = (R(t) - <R>)/\sigma$, where $<R>$ is the average value of the network resistance and $\sigma$ the root-mean-square deviation from the average (a similar procedure has been adopted for the $p(t)$ series). The following quantities have been considered: the auto-correlation function, the PDF of the $x$ records, the return intervals $r_q$ of the extreme values for different threshold $q$ and their distribution $P_q(r)$ ($q$ is expressed in units of $\sigma$). All the results shown here are obtained by considering a network of size $75 \times 75$, biased by a current $I = 1.0A$ and made by 15 species of resistors. The details concerning the network can be found in Ref. \cite{31}.

Figure 1 displays a typical trend of $x(t)$. Only a small portion of the total number of records, $N = 2 \times 10^6$, is
shown. In this case $x(t)$ represents the normalized resistance fluctuations calculated at a temperature of 400 K. The probability density of several $x(t)$ series has been also analyzed. It has been found that the PDF of the resistance fluctuations presents a significant non-Gaussianity that progressively increases at higher temperatures (going from 300 to 500 K) [30, 32, 35], while the PDF of the defect-fraction fluctuations is Gaussian at all the temperatures.

Figure 2 reports the auto-correlation functions, $C_x(t)$, of three resistance fluctuation time series obtained for three different temperatures. Precisely, the three black curves correspond to a temperature of 300, 400 and 500 K respectively going from the upper to the lower one. The grey dashed curve shows the best-fit to the $C_x$ data at 300 K with a power-law of exponent $\gamma = 0.22$, while the grey solid curves represent the best-fit to the data at 400 and 500 K with the function:

$$C_x(s) = C_0 t^{-h} \exp(-t/u)$$ (3)

The fitting parameters are $C_0 = 1.23$, $h = 0.30$ and $u = 1.42 \times 10^4$ for the data at 400 K and $C'_0 = 0.98$, $h' = 0.36$ and $u' = 1.51 \times 10^3$ for the data at 500 K. Many other functions have been also considered for the best-fit of the $C_x$ data. However, it has been found that the function in Eq. 3 optimizes the best-fit procedure with the minimum numbers of fitting parameters. By using Eq. 3, it is possible to calculate the correlation times characterizing the resistance fluctuations: $\tau = 1051$ at 400 K and $\tau = 148$ at 500 K (see Ref. 30 for further details). Figure 3 shows the auto-correlation functions of the defect-fraction times series obtained at 300, 400 and 500 K (black curves, again going from the upper to the lower one) and the best-fit to the $C_x$ data with Eq. 3 (grey solide curves). The fitting parameters are $C_0 = 1.12$, $h = 0.14$ and $u = 5.96 \times 10^3$ for the data at 300 K, $C'_0 = 1.19$, $h' = 0.19$ and $u' = 8.17 \times 10^3$ for the data at 400 K and $C''_0 = 1.08$, $h'' = 0.23$ and $u'' = 1.29 \times 10^3$ for the data at 500 K. The correlation times are: $\tau = 1.40 \times 10^4$ at 300 K, $\tau = 1653$ at 400 K and $\tau = 306$ at 500 K.

Let us start the analysis of the return intervals by considering first the RT distribution for extreme values of the resistance fluctuations at 300 K. As pointed out by Fig. 2, in this case $C_x$ scales as a power-law over more than four orders of magnitude, indicating the existence of long-term correlations. Figure 4 reports the distribu-

![Figure 3: Auto-correlation functions of the defect-fraction fluctuations calculated at 300 K (up circles), 400 K (squares) and 500 K (crosses). The grey solid curves represent the best-fit with the function given in Eq. 3.](image1)

![Figure 4: Distribution of the return intervals for values above the threshold $q = 2.5$ (dashed black curve at the border of the grey area) obtained from the resistance fluctuations calculated at 300 K (long-term correlated records). For comparison, the distribution of the return intervals for the same threshold value and obtained after shuffling the data is shown by circles.](image2)

![Figure 5: Double-logarithmic plot of the normalized probability density of the return intervals for different threshold values $q$. The $x(t)$ records are the same of Fig. 4, i.e. they present long-term correlations. The slope of the dashed line is 0.3.](image3)
that the distribution of the return intervals of extreme values of these series is well described by a stretched exponential and that the value of the exponent \( \gamma \) is independent of the threshold \( q \) in a large range of \( q \)-values. This occurs even in absence of long-term correlations and in presence of a finite correlation time.

### III. CONCLUSIONS

The distribution of return intervals of extreme values has been studied in several time series with different correlation properties: long-term and finite-term correlations. Precisely, it has been analyzed the return interval distribution of the fluctuations of resistance and defect-fraction displayed by a resistor with granular structure in nonequilibrium stationary states at different temperatures \( T \). The resistance fluctuations were calculated by using the MSN model which is based on a resistor network approach \([30]\). It has been found that when the auto-correlation function displays a non-exponential and a non-power-law decay, the distribution of the return intervals is well described by a stretched exponential with exponent \( \gamma \) largely independent of the threshold \( q \). This result shows that the stretched exponential distribution describes the distribution of the return intervals of extreme values not only when long-term correlations are present in the time series \([13, 16, 17]\), but even when finite-term correlations with non-exponential decay exist among the records, a situation typical of many systems which are approaching criticality.
Acknowledgments

Support from MIUR cofin-05 project "Strumentazione elettronica integrata per lo studio di variazioni conformazionali di proteine tramite misure elettriche" is acknowledged. The author thanks S. Ruffo (University of Florence, Italy), P. Olla (ISAC-CNR, Cagliari, Italy), G. Salvadori (University of Lecce, Italy) and E. G. Altman (Max Planck Inst. for Phys. of Complex Systems, Dresden, Germany) for helpful discussions.

[1] J. H. Christensen, O. B. Christensen, “Severe summertime flooding in Europe”, Nature, 421, pp. 805-806, 2003.
[2] M. Mudelsee, M. Böhringer and G. Tetzlaff, “No upward trends in the occurrence of extreme floods in central Europe”, Nature, 425, pp. 166-168, 2003.
[3] G. A. Mechi and C. Tebaldi, “More intense, more frequent and longer lasting heat waves in the 21st century”, Science, 305, pp. 994-997, 2004.
[4] A. L. Perry, P. J. Low, J. R. Ellis and J. D. Reynolds, “Climate change and distribution shift in marine fishes”, Science, 308, pp. 1912-1915, 2005.
[5] B. N. Goswami, V. Venugopal, D. Sengupta, M. S. Madhusoodanan and P. K. Xavier, “Increasing trend of extreme rain events over India in a warming environment”, Science, 314, pp. 1442-1445, 2006.
[6] J. T. Overpeck, B. L. Otto-Bliesner, G. H. Miller, D. R. Muhs, R. B. Alley and J. T. Kiehl, “Paleoclimatic evidence for future ice-sheet instability and rapid seal-level rise”, Science, 311, pp. 1747-1750, 2006.
[7] G. Ekström, M. Nettles and V. C. Tsai, “Seasonality and increasing frequency of Greenland glacial earthquakes”, Science, 311, pp. 1756-1758, 2006.
[8] E. J. Gumbel, Statistics of Extremes, Columbia University Press, New York, 1958.
[9] H. von Storch and F. W. Zwiers, Statistical Analysis in Climate Research, Cambridge University Press, Cambridge, 2001.
[10] S. Kotz and S. Nadarajah, Extreme Value Distributions, Theory and Applications, Imperial College Press, London, 2002.
[11] D. Sornette, Critical Phenomena in Natural Sciences, Chaos, Fractals, Selforganization and Disorder: Concepts and Tools, Springer, Berlin, 2004.
[12] T. Antal, M. Droz, G. Győrgyi, Z. Rácz, “1/f noise and extreme value statistics”, Phys. Rev. Lett., 87, pp. 24061-1-4, 2001.
[13] G. Salvadori and C. De Michele “Statistical characterization of temporal structure of storms”, Adv. Water Res., 29, pp. 827-842, 2006.
[14] C. Schmitt and C. Nicolis, “Scaling of return times for a high-resolution rainfall time series”, Fractals, 10, pp. 285-290, 2002.
[15] A. Bunde, J. F. Eichner, S. Havlin and J. W. Kantelhardt, “The effect of long-term correlations on the return periods of rare events”, Physica A, 330, pp. 1-7, 2003.
[16] A. Bunde, J. F. Eichner, J. W. Kantelhardt and S. Havlin, “Long-term memory: a natural mechanism for the clustering of extreme events and anomalous residual times in climate records”, Phys. Rev. Lett., 94, pp. 045701-1-4, 2005.
[17] E. G. Altman and H. Kantz, “Recurrence time analysis, long-term correlations and extreme events”, Phys. Rev. E, 71, pp. 056106-1-9, 2005.
[18] C. Pennetta and E. Allinoto, “Distribution of return periods of rare events in correlated time series”, in Unsolved Problems of Noise and Fluctuations, L. Reggiani, C. Pennetta, V. Akimov, E. Allinoto and M. Rosini, eds., AIP Conf. Proc. 800, pp. 546-552, 2005.
[19] C. Pennetta, “Distribution of return intervals of extreme events”, Eur. Phys. J. B, 50, pp. 95-98, 2006.
[20] A. Bunde, S. Havlin, J. W. Kantelhardt, T. Penzel, J. H. Peter and K. Voigt, “Correlated and uncorrelated regions in heart-rate fluctuations during sleep”, Phys. Rev. Lett., 85, pp. 3736-3739, 2000.
[21] Y. Ashkenazy, P. C. Ivanov, S. Havlin, C. K. Peng, A. L. Goldberger and H. E. Stanley, “Magnitude and sign correlations in heartbeat fluctuations”, Phys. Rev. Lett., 86, pp. 1900-1903, 2001.
[22] J. Davidsen and H. G. Schuster, “Simple model for 1/fα noise”, Phys. Rev. E, 65, pp. 026120-1-4, 2002.
[23] E. Koscielny-Bunde, A. Bunde, S. Havlin, H. E. Roman, Y. Goldreich and H. J. Schellnhuber, “Indication of a universal persistence law governing atmospheric variability”, Phys. Rev. Lett., 81, pp. 729-732, 1998.
[24] P. Bak, K. Christensen, L. Danon and T. Scanlon, “Unified scaling law for earthquakes”, Phys. Rev. Lett., 88, pp. 178501-1-4, 2002.
[25] A. Corral, “Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes”, Phys. Rev. Lett., 92, pp. 108501-1-4, 2004.
[26] G. Boffetta, V. Carboni, P. Giuliani, P. Velti and A. Vulpiani, “Power-laws in solar flares: self-organized criticality or turbulence?”, Phys. Rev. Lett., 83, pp. 4662-4665, 1999.
[27] Y. Liu, P. Cizeau, M. Meyer, C. K. Peng, H. E. Stanley, “Correlations in economic time series”, Physica A, 245, pp. 437-440, 1997.
[28] H. A. Makse, S. Havlin, M. Schwartz and H. E. Stanley, “Method for generating long-range correlations for large systems”, Phys. Rev. E, 53, pp. 5445-5449, 1996.
[29] M. Kac, Bull. of the Am. Math. Soc., 53, pp. 1002, 1947.
[30] C. Pennetta, E. Allinoto and L. Reggiani, “Long-Term correlations and 1/fα noise in the steady states of multispecies resistor networks”, cond-mat/0701712v1.
[31] C. Pennetta, L. Reggiani, G. Treffán and E. Allinoto, “Resistance and resistance fluctuations in random resistor networks under biased percolation”, Phys. Rev. E, 65, pp. 066119-1-10, 2002, and C. Pennetta, “Resistance noise near to electrical breakdown: steady state of random networks as a function of the bias”, Fluct. and Noise Lett., 2, pp. R29-49, 2002.
[32] C. Pennetta, E. Allinoto, L. Reggiani and S. Ruffo, “Non-Gaussian resistance noise near breakdown in granular material”, Physica A, 340, pp. 380-387, 2004.
[33] C. Pennetta, E. Allinoto, L. Reggiani, F. Fantini, I. De Munari and A. Scorzon, “Biased resistor network model for electromigration failure and related phenomena in metallic lines”, Phys. Rev. B, 70, pp. 174305-1-15, 2004.
[34] C. Pennetta, L. Reggiani and G. Trefán, “Scaling law of resistance fluctuations in stationary random resistor networks”, *Phys. Rev. Lett.*, **85**, pp. 5238-5241, 2000.
[35] S.T. Bramwell, P. C. W. Holdsworth and J. F. Pinton, “Universality of rare fluctuations in turbulence and critical phenomena”, *Nature*, **396**, pp. 552-554, 1998.