A review of power laws in real life phenomena

Carla M.A. Pinto, A. Mendes Lopes, J.A. Tenreiro Machado

a b s t r a c t

Power law distributions, also known as heavy tail distributions, model distinct real life phenomena in the areas of biology, demography, computer science, economics, informa- tion theory, language, and astronomy, amongst others. In this paper, it is presented a review of the literature having in mind applications and possible explanations for the use of power laws in real phenomena. We also unravel some controversies around power laws.

Keywords
Power law behavior, Parento law, Zipf law, Heavy tail distributions, Applications

1. Introduction

Power law (PL) distributions, also known as heavy tail distributions, Pareto-like laws, or Zipf-like laws, have been largely reported in the modeling of distinct real phenomena. Nevertheless, this is still a controversial issue since there are authors that state that PLs are simply statistical phenomena, that is, spurious facts. Others suggest that the debate around Pareto or Zipf laws should be about how well or how poorly these distributions fit the data and not about if they can be rejected or not. Some authors focus on the lack of strong proves of the validity of the PL models. In this paper, we summarize results found in the extended literature on PLs, trying to enumerate the most interesting and important issues. PLs are, in our understanding, a very interesting tool to analyze data, that should be supported by strong analytical validation and control.

In 1896 [79], Pareto introduced a distribution to describe income, that was known as Pareto law. It was demonstrated that the relative number of individuals with an annual income larger than a certain value x was proportional to a power of x. Since then considerable amount of work has been done in the study of income and wealth distribution, and expenditure [69,77,52,4,8,24,36,58,45] in distinct economic communities. Estoup [42] and Zipf [110,111] applied PLs to words frequen- cies. Their studies were followed by other researchers that considered other cases, such as the frequency of occurrence of personal names in most cultures [109]. In 1913, Auerbach [9] found that the product of the population size of a city by its rank in the distribution appeared to be roughly constant for a given territory. This study brought a new interest to the study of city size distribution [111,89,80,91,62,38,47,15,78,92]. Other studies on informetrics can be found in [35,106,39]. Scientific production by chemists and physicists [66] is also believed to follow a PL, as well as the number of citations re- ceived by papers [82,83]. PLs of various forms also appear in computer science [32,33,54,11,3], namely in studies of the num- ber of in- and out-degrees in models of dynamically growing web graphs [29,61,2], and the number of hits in web pages [1]. Other examples are the application of PLs to distribution of biological species [107,31], to earthquakes [53,21], to rainfall [81], moon craters [74], solar flares [67], war and terrorists attacks [53,85,21,88,37,22,16], books sales [19,75,55,44], music recordings and other commodities [30], to the fraction of mass metal or metal oxide powders produced by electronic means, in the case of limited quantitative observations of their particle size [43], to marine ecology [95], and non-coding DNA se- quences [59]. Interesting reviews on PL behavior and applications can be found in [64,100,71].

Detecting or proving the existence of a PL behavior in natural or human-made systems can be a very difficult task. PLs tell us that the size of an event is inversely proportional to its frequency. Thus, usually, researchers represent a log–log plot of these quantities and verify if a straight line is obtained. This is, however, a limited scientific way to proceed, and this issue will be discussed later in this paper.

Let X be a non-negative discrete random variable following a PL distribution. Then, its complementary cumulative distribution function is of the form $F(x) = P(X \geq x) = Cx^{-(\alpha-1)}$, where $\alpha > 0; C > 0$. This means that the tail falls asymptotically according to the value of α. This fact translates in heavy tails, comparatively to other distributions, such as exponential distributions. The Pareto distribution [79] is an example of a PL. It satisfies

$$P(X \geq x) = \frac{C}{\tilde{\alpha}} x^{-\tilde{\alpha}} \tag{1}$$

here $\tilde{\alpha} = \alpha - 1$ and $\tilde{C} = \frac{C}{\tilde{\alpha}}$.

Usually, in a PL distribution, $0 < \tilde{\alpha} \leqslant 2$. For these values of $\tilde{\alpha}$, X has infinite variance. Additionally, if $\tilde{\alpha} \leqslant 1$ then X has infinite mean. The probability function of a discrete random variable following Pareto distribution is given by

$$P(X = x) = Cx^{-\alpha} \tag{2}$$

Zipf's law [110,111] (aka rank-size rule) is a special case of the Pareto law. Zipf [111] proposed that the distribution of city sizes was described by a Pareto distribution with $\alpha = 2$. Consider that cities are ordered by population size, with the one with more population being ranked as 1. The rank-frequency chart is the plot of the occurrence frequency f_r versus the rank r, in logarithmic scales. In log–log scales, the Zipf distribution gives a straight line with slope equal to $-\alpha = -1$. More generally, for a random variable following a Pareto distribution the rank-frequency plot, in log–log scales, is asymptotically a straight line. For the specific case of a Pareto distribution, the behavior is exactly linear, as

$$\ln(P[X \geqslant x]) = \ln C - \ln \tilde{\alpha} - \tilde{\alpha} \ln x \tag{3}$$

Pareto and Zipf's distributions are also known as skew distributions. Other examples are the lognormal distribution. A continuous random variable Y is said to follow a lognormal distribution if the random variable $Z = \ln Y$ follows a normal distribution. The probability density function of Y is given by

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma y} e^{-(\ln y - \mu)^2/2\sigma^2} \tag{4}$$

where μ and σ are the mean and standard-deviation of variable Y, respectively. The complementary cumulative distribution function for variable Y is of the form:

$$P(Y \geqslant y) = \int_{z=y}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma z} e^{-(\ln z - \mu)^2/2\sigma^2} dz \tag{5}$$

Table 1
PL distributions seen in the literature, where α is the PL coefficient, $\tilde{\alpha} = \alpha - 1$, and $C > 0$. In Gibrat's law, $R = x_2/x_1$ is the growth rate, for fixed x_1, where x_1 is the income at time t and x_2 is the income at a later time. For the DPLD, α and β are the Pareto coefficients for the upper and lower tails, espectively, μ_0 and σ_0 are the lognormal parameters, U is the normal cumulative density function (cdf), and $U^c = 1 - U$ represents the complementary cdf.

Power law	Probability (density) function
Pareto	$P(X = x) = Cx^{-\alpha}$
Zipf	$P(X = x) = Cx^{-2}$
Gibrat	$P(R\|x_1) = P(R), R = x_2/x_1$
Lognormal	$f(y) = \frac{1}{\sqrt{2\pi}\sigma y} e^{-(\ln y - \mu)^2/2\sigma^2}$
Discrete Gaussian	$P(X = k) = \frac{A(\mu,\sigma)}{k} \exp\left[-\frac{(\ln(k)-\mu)^2}{2\sigma^2}\right]$,
Exponential	$A(\mu,\sigma) = \left\{\sum_{k=1}^{\infty} \frac{1}{k} \exp\left[-\frac{(\ln(k)-\mu)^2}{2\sigma^2}\right]\right\}^{-1}$
Pareto-positive	$f(x) = \frac{\lambda\nu[\log(x/\sigma)]^{\nu-1}}{x} \exp\left[-\lambda(\log(x/\sigma))^{\nu}\right], \quad x \geqslant \sigma$
Distribution	$f(x) = 0, \quad x < \sigma$
Double Pareto (DPLD)	$f(x) = \frac{\alpha\beta}{\alpha+\beta}\left[x^{-\alpha-1} \exp\left(\alpha\mu_0 + \frac{\alpha^2\sigma_0^2}{2}\right) \Phi^c\left(\frac{\log(x)-\mu_0-\alpha\sigma_0^2}{\sigma_0}\right) + \right.$
Lognormal distribution	$\left. + x^{\beta-1} \exp\left(-\beta\mu_0 + \frac{\beta^2\sigma_0^2}{2}\right) \Phi^c\left(\frac{\log x-\mu_0+\beta\sigma_0^2}{\sigma_0}\right)\right]$

A lognormal distribution results when the variable Y is the product of a large number of independent, identically-distributed variables. The lognormal probability plot shows the relation between the inverse of the cumulative distribution function of a variable Y, following a lognormal distribution, vs the natural logarithm of n realizations of the variable Y. Analogously to what is seen in the log–log Pareto or Zipf representations, this graph is a straight line. The lognormal distribution has been competing with Pareto's and Zipf's laws in the modeling a variety of data, from economics to city sizes.

In Table 1, we review some of the probability functions of PL distributions seen in the literature.

2. Power laws for cities

Cities are complex systems, that differ in many distinct ways, such as size, shape, and scale. These issues have been a major research theme in a variety of scientific areas, from geography up to economy [12]. The aspect that got more attention was the city size distribution. It is consensual that there are two major contributions for the distribution of city sizes: the Pareto distribution (which includes the usual Zipf distribution, as a particular case) and the lognormal distribution. Nevertheless, other distributions have appeared to model city size distributions, such as the Pareto-positive stable, discrete Gaussian exponential, and others. In this section, we summarize some of the literature concerning city size distributions.

Auerbach [9] (followed by Zipf [111]) proposed that the city size satisfied a Pareto distribution. In 1980, Rosen and Resnick [89] did a cross-country investigation of city sizes in 44 countries and found that Pareto exponent was in the interval $\alpha \in [0{:}81; 1{:}96]$, with sample mean of $\alpha = 1{:}14$. These values indicated that population was more evenly distributed than expected by Zipf's law, since the higher the value of the exponent, more even the city sizes became. In the limit, if $\alpha \to \infty$, all cities would have identical sizes. Moreover, Rosen and Resnick [89] found that the mean value of α was $\alpha = 1{:}136$. Only in 12 of the 44 countries studied, obtained $\alpha > 1$. They also tried to explain variations in the Pareto exponent, and showed that it is sensitive to city definition and city sample size. The Pareto coefficient was lower when considering population of metropolitan areas rather than "city proper", suggesting that results might depend on city definition. As to the sample size, authors compared the values of the Pareto coefficient for 6 countries using two methods. The first was to consider only the 50 largest cities and the other was to consider cities with more than 100,000 inhabitants. Here, again, changes in the value of the Pareto exponent were observed. These results were explained by the significance of the non-linear terms included, by the authors, in the logarithmic version of the Pareto distribution. Finally, these authors also tried to explain variations in the Pareto exponent using economic, demographic, and geographic factors, such as total population, area, GNP per capita, and railway mileage density. They conclude that cities of wealthier and more populous countries, with a better railway system, grew more evenly, i.e., they verified that Pareto coefficient was bigger than for poorest countries.

Krugman [62], in 1991, studied 135 USA metropolitan areas and calculated values of α close to one. Using the same data set, Gabaix [48,47] derived a statistical explanation of Zipf's law for cities. He showed that if cities followed similar growth processes, i.e., if they evolved randomly with the same growth mean and variance, for a certain range of (normalized) sizes, then, in the steady state, the distribution of city sizes would follow a Zipf law. Similar growth processes were often referred as Gibrat's law [50] (aka, Gibrat's law of proportional effects). Thus, Zipf law was considered a steady state distribution arising from Gibrat's law, that is to say, Zipf's law was the limit of a stochastic process. Simon [96] showed that proportional growth could explain several different lognormal, Pareto and Yule distributions. The Yule distribution results from inserting the birth-and-death process in Gibrat's law, considering a constant rate of birth [96], or serial correlation between periodic growth rates [56].

Zanette and Manrubia [108] developed an intermittency model to large-scale city size distributions. The model predicted a PL distribution with a coefficient α that was close to one, the coefficient for large city sizes known in the literature.

Brakman et al. [18] used historical data from The Netherlands to prove that the Pareto distribution is a good model for city sizes over a wide range of time, from pre-industrialization, passing by industrialization, up to post-industrialization.

Overman and Ioannides [78], implemented a test to check validity of Zipf's law for cities. Their results confirmed the validity of Zipf's law for the US cities, though it was observed a variation in the estimates of the Zipf coefficient across city sizes. They also proved that Gibrat's law hold generically for city growth processes. Urzúa [103] also proposed a test for Zipf's law. He defined the Lagrange multiplier test, where he tested the null hypothesis $H_0 : r = l;\ \alpha = 1$, corresponding to the PL distribution of city sizes of US metropolitan areas. His results showed that Zipf's law only held for cities of a certain sample size (greater than 250,000 inhabitants).

Davis and Weinstein [37], tried to unfold the problem of dissimilar economic activity across regions. They studied data on 39 Japanese regions for 8,000 years, from the Stone Age up to the modern era (till 1998). They also examined 303 Japanese cities (with more than 30,000 inhabitants in 1925) in order to study the implications, of the Allied bombing of Japanese cities during World War II (WWII), to the relative city sizes. They found that variation in Japanese regional population density, as well as the distribution of city sizes, obeyed a Pareto distribution, at all points in time. Nevertheless, they defended a hybrid theory in which spatial patterns of relative regional densities were explained by locational fundamentals, whereas the degree of spatial differentiation were due to increasing returns. Increasing returns tried to explain how spatial differentiation might emerge even in a framework where all locations were identical in terms of economic primitives. The locational fundamentals theory stated that permanent features of specific locations (i.e., closeness to the sea, access to rivers) made them valuable places for economic activity. A specific distribution of these features supported Pareto distribution [48]. The simple random growth theory [96,48] did not apply since it predicted that city sizes would grow randomly, and this was shown to be wrong by the effect on city sizes after the bombings during WWII.

Soo [99] tested the validity of Zipf's law for cities of 73 countries using two estimation methods, namely, the ordinary least squares (OLS) and the Hill estimator. He found that Zipf's law was rejected in 73% of the countries using OLS estimator, and 41% of the countries using Hill estimator. Soo updated α values for the interval $[0.73; 1.72]$. Soo also obtained a Pareto value for urban agglomerations less than 1, contradicting Rosen and Resnick [89], that stated that Zipf's law hold for urban agglomerations. Soo also tried to explain variations in the Pareto exponent. Unlike Rosen and Resnick [89], he argued that political economy might be the main factor influencing city size distribution.

Nitsch [76] analyzed 515 estimates of the Zipf exponent α, from 29 studies of city size distribution, for a wide range of different territories and time periods. He used meta-analysis to check the validity of Zipf's law as distribution to city sizes. His results showed that the estimated Zipf exponent was close to $\alpha = 1.1$, agreeing with Rosen and Resnick [89] and Soo [99], pointing Zipf's law not to be the most adequate and suggesting the Pareto distribution as a better alternative. Nitsch also enumerated some factors influencing the variation of α exponents, such as the base population (metropolitan areas or inner cities), the country (USA, Japan, China), and the small number of observations.

Moura and Ribeiro [73] studied Brazilian cities with more than 30,000 inhabitants. They showed that Pareto distribution was not valid for smaller cities. For these cases, the city size cumulative distribution function did not follow a PL behavior. The coefficient α values were calculated with three methods: maximum likelihood estimator, least squares fitting and average parameter estimator, fitting the interval [1.3, 1.4].

Bosker et al. [17], studied the evolution of city size distribution of 62 of the largest West-German cities, between 1925 and 1999. They found that large scale events, such as the WWII and the splitting and reunification of Germany, were main contributors towards evolution. Before WWII, city size distribution followed a Zipf law (Pareto coefficient equal to 1). However, the Pareto coefficient increased from $\alpha = 1.09$ in 1939 up to about $\alpha = 1.16$ right after the WWII, while in 1999 reached $\alpha = 1.20$. These outcomes revealed that big events influence remarkably the city size distribution. Moreover, Bosker et al. [17] findings supported the theory of increasing returns to scale for city growth rather than Gibrat's law. This was in part in agreement with Davis and Weinstein [37].

Rozenfeld et al. [90]) proposed the City Cluster Algorithm (CCA) to define a metropolitan area. This issue was quite important since results on urban distributions might be related to this definition. Authors used CCA to compute clusters in data sets of city sizes of different countries and continents, such as Great Britain, USA, and Africa. Results suggested the existence of scale-invariant growth mechanism, but deviations from Gibrat's law were also observed. The deviations were shown to arise from the existence of long-range spatial correlations in population growth.

The lognormal distribution [34] applied as a model for city size distribution was due to Gibrat [50]. The lognormal distribution appears when cities grow randomly and proportionately. In 2005, Anderson and Ge [6] compared US and Chinese city size distributions, using Pearson goodness-of-fit tests, and found significant differences between them. These differences revealed that lognormal distribution was well suited for Chinese cities, whereas Pareto led to a better fit for US cities. These differences were explained by the major economic reforms that took place in China, after 1980, forcing a structural change in Chinese urban system. The Chinese city size distribution remained stable before the reforms, but exhibited a convergent growth pattern after the reforms.

Gan et al. [49] used Monte-Carlo simulations to show that Zipf law for city size distribution was a statistical phenomena, a spurious fact. In the simulations, authors studied the rank-size relationships generated from various probability distributions (normal, lognormal, Pareto, and Gamma), by running regressions on random numbers (and their ranks). Results demonstrated that the requirement of an economic theory determining city-size distributions was erroneous.

Warren [104] showed that lognormal and Pareto distributions were insufficient to model city sizes. He argued that there should be two distinct Pareto distributions, one for modeling the largest city sizes and another for modeling smaller cities.

Meanwhile, other distributions to model city sizes were proposed in the literature. In 2001, Bi et al. [14], proposed the discrete Gaussian exponential (DGX) distribution for mining, massive skewed data. A random variable with DGX distribution is a lognormal random variable digitized to the nearest integer. If the lognormal random variable becomes zero after the rounding, then it is omitted. The Pareto law was included as a special case of the DGX. Authors applied DGX to distinct data sets, a text from the Bible, clickstream data, sales data and telecommunication data. In all cases, the DGX fitted well the data. Sarabia and Prieto [92] introduced the Pareto-positive stable (PPS) distribution as a new model for city size distribution. Pareto and Zipf distributions were included as particular cases. PPS distribution could be obtained by a monotonic transformation of the classical Weibull distribution, or by mixing the shape parameter a of the classical Pareto distribution with a positive stable law. Authors compared their results on data from population of Spanish cities from 1998 up to 2007, with Pareto, lognormal and Tsallis [102] distributions, and PPS provided better fits than the three previous distributions.

Benguigui and Blumenfeld–Lieberthal [13] proposed a dynamical growth model for cities. The model simulated the evolution, over time, of measurable entities, such as city size. They computed both size and number for each entity. Authors validated data from the model with real city size data from the Israeli system of cities, from 1961 to 2006. Israel was an example of a country in which cities grew in size and in number, over time. They concluded that there was a good correspondence between simulated and real data. Changes in the values of the Pareto coefficient were observed and were justified by the variation of the number of entities with time.

Giesen et al. [51] introduced the Double Pareto lognormal distribution (DPLN) to model city sizes. They used data from eight countries and their results showed that DPLN was a good fit for cities of all sizes. It was compatible with Zipf's law among large cities and followed a lognormal shape for other cases.

3. Economy and power laws

In 1896, Pareto [79] introduced the Pareto distribution to describe income distribution. It was the first PL to appear in the literature. Nowadays, in this highly globalized society, precise databases of market prices, company's and people's wealth, are known for smaller and smaller sampling periods. These databases are crucial in the understanding of the world's economic activity.

The study of the Pareto's law validity for describing income or wealth distribution is very important, since, though validity is restricted to a very small amount of income or wealth distribution, the fact is that this small amount can influence greatly the global economy. Many theories have been proposed to explain this Pareto fat tail behavior, using, for example, Markov chains to model evolution of person's income [23] or Lévy distributions [69]. The value of the Pareto exponent lies between 1 and 2 [63], in most studies, varying from economy to economy.

The wealth or income of the majority of the population and companies, that is, of the poor and the low-income group of the population, does not follow a PL and some say it obeys a Gibbs distribution [40].

Many new results concerning the income distribution constantly appear in the literature. Nevertheless, there is still no consensus of what should be the most suitable distribution to model the whole income distribution of countries.

Mantegna and Stanley [70] studied 1,447,514 records of the Standard & Poor's 500 cash index, between 1984 and 1989. They measured both price differences and returns. Results showed that the central part of the price differences distribution was modeled by a Lévy stable process, whereas the tails were measured by an "approximately exponential" distribution. Lévy's exponent value was close to 1.4. The log–log plot of the probability of return of index variations versus time was a straight line, with slope $\alpha \approx 1{:}4$. Lévy and Soloman [63] analyzed the wealth of the 400 richest people in US in 1996 (Forbes), as a function of their rank. They obtained a PL fit with $\alpha \approx 1{:}36$. This value of α was in good agreement with the work of Mantegna and Stanley [70], where stock market fluctuation was analyzed. Lévy and Solomon [63] conclude that this observation supports the relation between wealth distribution and stock market fluctuations.

In 2003, Reed [84] developed a stochastic model for income distribution that predicted a PL behavior in both tails. The author derived a four-parameter distribution, the double Pareto-lognormal distribution, that provided a very good fit to observed data on income and earnings, ranging over a number of different countries, incomes, and times (USA, 1997; Canada, 1996; Sri Lankan, 1981; Bohemia, 1933). The double Pareto-lognormal distribution was derived under the assumption of a workforce or population growing at a constant rate. It took advantage of the fact that income distribution over a population depended on the age profile of the population, that in turn, relied upon the demographic dynamics. In this sense, there would be more younger workers than older ones in the workforce (with an average lower income) for a growing population and conversely for a declining population.

Chatterjee et al. [24] introduced an ideal-gas model for markets' trading. Each gas molecule was considered an economic agent, and collisions between two molecules where identified with the trading. Saving propensity of the agents was introduced in the model and inserted some randomness. Results showed that Gibbs distribution fitted well trading with no savings and Pareto distribution modeled trading with (quenched) random saving propensity.

In 2004, Fujiwara et al. [46] studied data sets of European firm sizes in order to explain firm size distribution and growth. They considered large size firms, where Pareto behavior was observed in the upper tail and checked validity of Gibrat's law in that regime. In the larger part of the distribution, concerning lower firm sizes, lognormal behavior was predominant. Results

showed that Pareto law holded for a firm size larger than a given threshold, Gibrat's law was valid, supporting the fact that the growth rate of each firm is independent of initial size, and the law of detailed balance could be applied to data, stating that the transition between two states has the same frequency as its reverse. These authors also analyzed data from 16,526 Japanese bankrupted firms in the year of 1997. Their goal was to shed some light in the phenomenon of death of a large number of firms. They argued that results might help to understand the firm size growth. Results obtained showed that PL distribution, in particular, the Zipf distribution, fitted well debt distribution of bankrupted firms. The bankrupted firms size and debts were related by a multiplicative factor, independent of the debt. Life-time distribution of bankrupted firms was exponential, correlating with the rate of appearance of new firms.

Aoyama et al. [8] developed a framework, using the law of detailed-quasi balance, to find the relations between the Pareto's law and the Gibrat's law. The later could be observed in the distribution of quantities such as personal or company's income, firm size, firm number of employees, amongst others.

Andersson et al. [7] proposed a new model for geographic distribution of land values, that combined together transportation costs and multiplicative growth (Simon, Gibrat). They compared model's results with data of geographical maps of land values from Sweden, in year 2000. These authors found that transportation costs did not affect the PL distribution of land values per unit area, nonetheless, they could force new PLs on higher levels of aggregation. Results showed that these relations could be consequences of a growth model for geographic economic interdependencies. That is to say, they verified that the model replicated the random multiplicative growth model and opened space for microeconomic arguments and geography.

Das and Yarlagadda [36] studied the processes that led to wealth distribution in society. They considered tiny and gross trades. The first case consisted in trading between two people, while the later involved trading between one person and the gross system. Tiny trade was associated with poor people, that traded only with other poor people. Gross trade was due to wealthier people that traded with governments, nations' markets, and so on. At a steady state, a PL behavior was observed for the interaction of wealthy entities distribution, whereas a Boltzmann–Gibbs distribution was seen in the tiny trade interaction. The Boltzmann–Gibb's law was obtained in models of random trading between two people, in the absence of any savings [24].

In 2005, Ishikawa [57] verified Pareto's law in profits data from Japanese companies for two years, 2002 and 2003. He attested the derivation of Pareto's law and Pareto's coefficient from the law of detailed balance and Gibrat's law. He divided the companies in two categories, small-scale and large-scale job classes. The Pareto exponent was larger in the small-scale category, revealing that these might grow more than large-scale companies. Pareto coefficient varied closely to $\alpha \approx 1$.

Ishikawa [58] studied data concerning the values of lands in Japan, from 1983 till 2005. He was interested in the annual variation of Pareto's exponents and applied the detailed quasi-balance and Gibrat's law to non-equilibrium systems. He obtained a relation between the variation in the Pareto exponent and the symmetry in the law of detailed quasi-balance. The model fitted well the data except for two abrupt jumps in the Pareto's exponents, in the periods 1985–1986 and 2001–2002.

In 2006, Chebotarev [25] proposed a hierarchical model for the income distribution of agents. Interaction between agents was asymmetric and the price was invariant. Asymmetric trading was defined by trading where commodities moved from older to younger traders and money did the reverse. The price-invariance property meant that the ratio of incomes influenced the pairwise interaction, that was independent of the price or absolute income. This type of income distribution followed a double-Pareto function, with upper and lower Pareto tails. Chebotarev computed Pareto exponents for upper income to be 1 for gross trading and 2 for net trading, for linear demanding. For unlimited slowly varying demand, Pareto's coefficient was 1, for both types of trading. The lower income distribution had a Pareto exponent of 3. Chebotarev used simple balance equation, multiplicative Markov chains and exponential moments of random geometric progressions in his study and argued that his model could fit well certain phenomena in macroeconomics.

In 2008, Mizuno et al. [72] studied the modeling of the expenditure of a person in convenience stores. He analyzed 100 million receipts, between January and March 2007, from costumers of a Japanese store chain. Results showed that the density function of the expenditure had a fat tail following a PL, with Pareto coefficient equal to 2. They computed the Lorenz curve and estimated the value of Gini coefficient to be 0.7. The Lorenz curve was one of the measures of social inequality in economics. The Gini coefficient was used as a measure of inequality of wealth distribution and was estimated by twice the value of the area between the line $y = x$ and the observed Lorenz curve. A value of 0.7 suggested that there was inequality in the stores sales, and that the loyal customers contributed significantly to sales.

Figueira et al. [45] developed a model of income distribution that combined the Gompertz curve and the Pareto law. The PL modeled the income of the wealthier part of the population, whereas the Gompertz curve fitted the poorer part of the population. They used data from Brazil between 1981 and 2007. The Gompertz curve was commonly used in biology to model growth processes, population dynamics and mortality. The model introduced in this study required only three parameters to be linearly fitted from the data. Results revealed that the combination of Gompertz and Pareto distribution was concordant with the data. They computed the equation for the Lorenz curve, the Gini coefficient and the percentage share of the Gompertz part of the distribution relative to the total income.

4. City and forest fires

The study of city and forest-fire distributions is extremely important. Conclusions from this study can be used to take measures beforehand in view of possible hazards, thus saving natural resources and animal and human lives.

Let the size of a fire, either a city or a forest fire, be measured by the burned area, then the frequency-size distribution of fires obeys a PL. Several models have been proposed in the literature to explain this behavior.

In 1992, Drossel and Schwbal [41] proposed a stochastic cellular automaton to model forest-fire distribution. The authors have improved a model introduced by Bak et al. [10]. The forest was denoted by a d-dimensional lattice, with size L, and Ld sites. The growth of trees, ignition and spread of fires were simulated in a random way in the model. A burning tree ignited the non burnt neighboring trees. Thus, single events were characterized by a complete burn out of a forest cluster. This forest-fire model reached a steady state, defined by frequency-size distribution of fires following a PL. Mean field theory was used to explain this behavior at the critical state.

Malamud et al. [68] examined data from USA and Australian forest fires. They found that the graph of the frequency of occurrence vs the burned area followed a PL behavior, over distinct orders of magnitude, and was invariant with time. The same behavior was observed in a simple forest fire model. As a result, one practical implication was that the risk of large fires could be assessed by the frequency-area distribution of small and medium fires.

Ricotta et al. [87] studied 9164 wildfire records of a northern region of Italy (Liguria), from 1986 to 1993. They found that the relation between the frequency of occurrence and the size of burned area of those fires followed a PL, with coefficient $a \frac{1}{4} 0{:}723$. They also concluded that, due to the fact that 90% of the wildfires in Liguria had human origin, the comprehension and modeling of of the ignition of wildfires was very complicated.

Song et al. [97], improved the model of Drossel and Schwbal [41], by including the effects of rain and human fire fighting efforts, such as firebreaks. The authors simulated the new model and compared their results with Chinese forest-fire records, from 1950 to 1989. They concluded that there was a good match between simulated and real data. Their findings supported the claim that small effects, like firebreaks, could preserve trees and prevent large, hazardous forest fires.

Satoh et al. [93] developed a system to predict forest fire danger. Important variables in the model were the weather conditions and the available data on fire occurrence. They have simulated the model including also population density data. It was concluded that there was a good agreement between the predicted fire danger and the locations of real fire.

In 2006, Weiguo et al. [105] explored three distinct PL behaviors seen in frequency-size, frequency-interval, and frequency-density-of-population distributions of forest fires. They have used data from 518 fires, from 1950 to 1989, in China, 4284 fires, from 1986 to 1995, in USA, and 30,498 fires, from 1989 to 2000, and 5493 fires, from years 1999 and 2000, in Japan. These researchers concluded that the frequency-size distribution of forest-fires in these countries followed a PL behavior, with exponents $a \frac{1}{4} 1{:}31$, $a \frac{1}{4} 1{:}30$, and $a \frac{1}{4} 1{:}75$, respectively. They considered the environmental differences between these countries (the weather, tree species, environment and fire fighting systems, etc.) and human-related variables the main factors contributing to the distinct exponent values. Namely, most Japanese forest fires were caused by human activities, such as smoking, farmers' and camping fires, whereas, in USA, forest-fires were mainly due to lightings [93]. The frequency-interval distribution, where an interval is considered to be the interval of time between two consecutive ignitions, was also found to follow a PL. This behavior was observed for over two decades, depending on the magnitudes of fire intervals. The overlapping of 1999 and 2000 frequency-interval distributions, suggested that the distribution was time invariant. The exponent value was $a \frac{1}{4} 1{:}43$. The third PL distribution observed by the authors was related to the relation between population and fire probability. The PL approximating this relation had a positive value of the exponent. Finding a regular relation was considered important in rating fire danger.

In 2003, Song et al. [98] studied fire distribution in Chinese and Swiss cities. The city fires were complex since they aggregated industrial fires, building fires, amongst others, and had greater human interaction. The city-fire system was thus more intricate than the forest-fire system. These researchers computed the frequency loss plot and the rank-size plot and verified validity of a PL to these plots. The frequency loss was the frequency of fires with loss L. The rank was computed by sorting city fires from large to small, and considering the largest with rank 1. Results showed that large city fires follow a PL distribution for China and Switzerland. It was found that this distribution was invariant for scale and time, meaning that fire distribution is common for different places and times. Small fires did not follow a PL, while for large fires, Pareto coefficient for rank-size plots were in the interval $a 2 \frac{1}{4} 0{:}81; 0{:}88]$.

5. Online sales

Online sales help e-commerce businessman to make profit and constitute nowadays a valid alternative to traditional selling methods. The total number of sales of products such as books, music CDs, films, that are less popular is comparable to the number of sales of most sold items. This is due to the long tail behavior [75].

Chevalier and Goolsbee [26] used publicly available data on the sales ranks of 18,000 books to derive quantity proxies at two leading online booksellers: Amazon and Barnes and Noble. They have modeled the distribution of the number of sales by a Pareto distribution, with an estimated Pareto coefficient of $a \frac{1}{4} 1{:}2$, for both booksellers. Also in 2003, Brynjolfsson et al. [19], fitted sales of obscure titles, in total sales at Amazon.com, and sales rank data by a Pareto-like distribution. The Pareto coefficient obtained was $a \frac{1}{4} 1{:}15$. Based on the latter estimate of the power-law exponent and assuming, that the most popular 100,000 titles are stocked in Brick-and-mortar stores, these researchers concluded that about 40% of Amazon.com's sales were titles that could not be bought in those stores.

Newman [75], in 2005, used data of the 663 bestselling books in America, from 1895 up to 1965, to show that the rank-frequency plot of the number of copies sold followed a PL behavior. He estimated the proportion of tail sales by regression

techniques. Newman also showed cumulative distributions of twelve different quantities, of physical, biological, technological and social systems, such as word frequency, telephone calls, magnitude of earthquakes, citation of scientific papers, amongst others. The log–log plot also revealed a PL behavior. The corresponding Pareto coefficients were also computed, and were in the interval [0.94,2.5].

Anderson [5] made a critical review, based on other researchers' work and their own, of analytical frameworks for estimating Amazon's sales.

Iba et al. [55] studied books sales' data of over 2000 bookstores, in Japan, from April 2005 up to March 2006. They found that the logarithmic plot of sales volume and sales rank followed a PL behavior, both in monthly and annual sales. Researchers also explored, in particular, the titles that constituted the top 1.5% of sales, concluding that top sales tend to be always at the top, translating in increasing index and market share. Authors emphasized that understanding consumers' behavior was easier from then on, since results known from other phenomena, following PL behaviors, could be applied. The PL behavior used in this study was called "emergent order", due to the fact that it emerged from interactions at market level, that could not be reduced to the individual level. That is, the exposure of the individuals of top sales affected their other consumers' choice.

In 2010, Fenner et al. [44] introduced a new technique, based on a generative model, to explain the PL behavior observed in books sales. The model simulated a simplified sales process that asymptotically followed a PL. The two parameters in the model were a function of the rates at which new products were introduced and existing ones were discontinued. Researchers' identified strengths of the model as being: (i) prediction of tail sales, based on the number of available products and the total volume of sales, (ii) estimation of sales volumes from the tail, (iii) reduction of the error margin when computing the PL coefficient. Computing the PL coefficient using regression techniques was more susceptible to error. Moreover, simulation results from the model suggested that the tail sales of Amazon.com was close to 20%, the proportion estimated by its CEO, rather than 40% as argued in [19].

6. Wars and terrorists attacks

Patterns seen in wars and terrorists attacks have been at close attention by many researchers [53,85,86,21,88,37,22,60,16] and many attentive explanations have arisen in the literature. Nevertheless, political, ideological, cultural, historical and geographical influences, make the understanding of these patterns a complex goal to achieve.

The nature of war has been changing during the last few decades. Large scale conflicts, such as WWI and WWII, seem to be receding and a new kind of war is becoming prevalent. The later is characterized by lopsided affairs, where guerilla forces, insurgent groups, and terrorists oppose incumbent governments. These events seem to be random attacks to vulnerable targets of opportunity, and appear to be helpless in predicting future events. Nevertheless, considering the number of casualties and frequency of these natural and human-made disasters, there seems to exist a common underneath behavior. It is observed that large casualties are associated with low frequency phenomena, that is, very severe wars, producing huge numbers of casualties, are lesser frequent than other wars, not so harmful in terms of preserving human lives [85,86].

This behavior is also seen in earthquakes. Moreover, when thinking in the frequency of occurrence of earthquakes, there is a large number of earthquakes with few casualties, fewer that cause a large number of casualties, and a very small amount of huge disasters, where the number of casualties is extremely high [53,22].

Richardson [85], in 1948, explored domestic and international cases of violence, between 1820 and 1945. He divided those into five logarithmic categories, from 3 up to 7, that corresponded to casualties measured in powers of 10. In [86], the same author recorded 188, 63, 24, 5, and 2 events that matched each category listed above, respectively, the latter two being the two world wars. His results showed that as events increased by powers of 10 in severity, their frequency decreased by a factor close to three.

Cederman [22] followed Richardson's work [85,86]. He introduced an agent-based model of war and showed the power-law behavior. Its model allowed conflict to spread and diffuse, potentially over long periods of time, due to the quasi-parallel execution. This turned to be a new development concerning previous models, since it allowed modeling the territorial entities with dynamically fluctuating borders. This aspect was extremely important since great power wars coincided with massive rewriting of the geopolitical map. Europe and the two WWs were an exemplificative.

Johnson et al. [60] suggested a microscopic theory to explain why patterns of violence, such as war and global terrorism, evolved dynamically in a similar way, regardless of their underlying ideologies, motivations and the terrain in which they operated. Researchers have used wars in Iraq and Afghanistan, and long-term guerrilla war in Colombia, as examples. On global terrorism, attacks to London, Madrid, and New York (September 11) were main choices. Their theory considered the insurgent force as a generic, self-organizing system, which evolved dynamically through the continual coalescence and fragmentation of its constituent groups. The PL behavior was in excellent agreement with the data from Iraq, Colombia and non-G7 terrorism, and also with data obtained from the war in Afghanistan. Their findings indicated that the continual coalescence and fragmentation of attack units were the mechanism driving modern insurgent wars. Previous studies [85,86,75,22] had shown that the distribution obtained from wars in the period 1816 up to 1980, exhibited a PL with $\alpha \approx 1{:}80$. Note that, each data-point represented the total casualty figure for one particular war, in previous studies. In this study, authors analyzed casualties' patterns arising within a given war. In terms of global terrorist events, since 1968, it was known that they also obeyed a PL with $\alpha \approx 2{:}51$ for non-G7 countries [27] and $\alpha \approx 1{:}713$ for G7 countries. Results obtained

by Johnson et al. [60], showed a PL behavior for Iraq, Colombia and Afghanistan, with coefficient value (close to) $\alpha \approx 2.5$. This value of the coefficient equalized the coefficient value characterizing non-G7 terrorism. This result suggested that PL patterns would emerge within any modern asymmetric war, fought by loosely-organized insurgent groups.

Clauset et al. [28], focused in the frequency and severity of terrorist attacks, since 1968. They found that the log–log plot of the frequency vs the severity of the attacks follows a PL distribution. Severity was measured as: (i) number of deaths, (ii) injuries, and (iii) sum of the later. Economic development, the type of weapon used in the attack, or short time scales, did not affect the results. Researchers also conjectured that there was periodicity in the global terrorism, with period close to 13 years, since the autocorrelation function of the average log-severity, revealed a strong periodicity in the breadth of the distribution at roughly 13 years. Researchers introduced a toy model to fit the frequency of severe terrorist attacks and showed that existing models failed to reproduce the heavy-tail behavior observed in the log–log plot.

Bohorquez et al. [16] studied the quantitative relation between human insurgency, global terrorism and ecology. They considered a unified model of human insurgency to explain universal patterns occurring across wars. The main goals were to explain the size distribution of casualties or the timing of within-conflict events. Their model was based in ecology, considering the insurgent populations as self-organized groups that dynamically evolved through decision-making processes. Researchers explained the model as being consistent with work on human group dynamics in everyday environments, and a new perception of modern insurgencies, as fragmented, transient and evolving. Additionally, the decision-making process about when to attack was based on competition for media attention. Authors computed the Pareto coefficient to a value close to $\alpha \approx 2.5$, agreeing with previous work on Iraq and Colombia wars, and with the other insurgent wars studied in this paper (Peru, Sierra Leone, Afghanistan, amongst others). On the other hand, results from the model on data of the Spanish and American Civil Wars suggested that if a PL distribution was fitted then the coefficient value would be much smaller (around $\alpha \approx 1.7$). Moreover, for the later, a lognormal distribution might be applied to fit the data. As a consequence of these findings, researchers claimed that insurgent wars were qualitatively distinct from traditional wars. A coefficient value of $\alpha \approx 2.5$ was in concordance with the coefficient value of $\alpha \approx 2.48 \pm 0.07$ obtained by Clauset et al. [28] on global terrorism.

Researchers also argued that their model also suggested a remarkable link between violent and non-violent human actions, due to its similarity to financial market models.

7. Other examples of power law in sciences

This section is devoted to other examples of PL application.

Schuster et al. [94], in 1994, explored ribonucleic acid (RNA) folding of sequences into secondary structures. The shapes/structures were obtained by folding random sequences of fixed chain length (30 and 100). Researchers computed the frequencies of occurrence for individual shapes and showed that they followed a PL behavior. Ranks were assigned according to decreasing frequency values. The expression of the distribution was given by $f(x) = a x^{-c}$, where x was the rank of a shape and $f(x)$ its frequency. Parameter values of the best fit were $a \approx 1.25$; $b \approx 71.2$; $c \approx 1.73$. The frequency distribution of all secondary structures was invariant to chain lengths.

Li and Yang [65] analyzed the fitting of a PL to micro-array data (DNA chip). Their goal was to infer the importance of PL behavior in the selection process, namely, in the classification of important and unimportant genes. Researchers have used the maximum likelihood, normalized by the sample size, to rank the genes. The likelihood quantity measured the success of a gene towards the classification task. A high value of the likelihood indicated that the gene was successful in distinguishing distinct phenotypic changes, such as, cancer. On the contrary, a low likelihood value translated in an inactive/non important gene in that particular phenotypic change. PL behavior was seen in most of the ranked likelihood plots, and appeared also to be a good approximation for permuted data. Moreover, the PL behavior prevented a cutoff point between essential and unimportant genes. The PL coefficients computed in the examples shown in this study were smaller than $\alpha \approx 0.1$.

In 2006, Stroud [101] simulated epidemic dynamics of avian-related influenza in synthetic populations representing three US metropolitan areas (Los Angeles, Chicago, and Portland). They found that the number of new infections per day per infectious person scaled as a power (greater than one) of the fraction of the population that is susceptible. They computed the Pareto coefficient to be in the interval $\alpha \in [1.7; 2.0]$, depending on strength, duration, and number of contacts per person, that were distinct from city to city. This PL mixing formulation was different from the traditional disease modeling, where the expected number of new infections per day per infectious person was considered (linearly) proportional to the fraction of the population that was susceptible. In the new PL framework, the effects of social contact on epidemic dynamics were taken into account and the resulting simulated epidemic dynamics were significantly different from the traditional scaling. The framework was implemented in epidemic models, such as the S(usceptible) I(nfectious) R(ecovered), S(usceptible) E(xposed) I(nfectious) R(ecovered), and histogram-based models. The PL scaling produced better fittings with simulated data than the linear scaling.

In 2007, Caron et al. [20] focused in models of semantic extraction of different regions of interest (ROI) in images. A ROI could be defined, generally, by visually and structurally distinctive features than the rest of the image. This issue was relevant in machine image processing, since machines, unlike humans, failed to have cultural references and knowledge to identify the whole panorama. Researchers have used Zipf and inverse Zipf law to model the frequency of appearance of the patterns contained in images. They considered image patterns defined as blocks of 3×3 adjacent pixels. The number of patterns was minimized, as well as the distortion image, by defining nine pattern classes, corresponding to a different grey scale.

Each pixel was assigned a value of a class. Zipf and inverse Zipf laws were applied to an example image. First the image was scanned with a 3 ×3 mask, then patterns were encoded and frequencies of each pattern were computed. Researchers plotted a log–log chart of the decreasing frequencies of the patterns vs their rank, and concluded Zipf law was a good fit, despite the fact that two linear zones appeared in the graph. Homogeneous regions in the image were modeled by one PL and small regions were fitted by another independent PL. In order to check if the inverse Zipf law could be applied in the analysis of a ROI, researchers used a different approach, they have counted the number of distinct patterns having a given frequency, and represented the results in a double-logarithmic plot. They obtained a linear trace, concluding that this distribution fitted the number of distinct patterns, with respect to their frequencies. Caron et al. [20] concluded that both Zipf and the inverse Zipf laws could be used to detect ROI in image, since the differences observed in the log–log graphs of these distributions helped to understand better the details in an image.

8. Practical examples of PLs

This section investigates the cumulative distributions of several random variables of different kinds. All have been suggested in the literature to follow PLs, at least over some part of their range. Bearing this idea in mind, we consider one variable in the area of demographics (the size of populations of a set of cities around the world), two variables are in the area of economics (the wealth of the 400 richest people in the USA and the annual revenue of the largest American private companies), two variables in social sciences (the number of unique visitors of the 100 most-visited web sites in a set of countries and the severity of terrorist attacks worldwide), and four variables address natural phenomena (the sizes of earthquakes and forest fires in California, the size of craters on a set of planets of the solar system and, finally, the severity of tornadoes in the USA).

8.1. Populations of cities

In this study the concept of administratively defined cities is adopted. The data was collected from the web site Thomas Brinkhoff: City Population, http://www.citypopulation.de. The site advertises that the data is mostly based on official censuses and estimations, but some might be of "suspect accuracy". Anyway, this data has been used by several authors that checked their reliability when compared to data acquired from other sources. The web site supplies data about city populations for all countries and, for each country there is data corresponding to several censuses that have been taken place in the last 40 years.

A sample of 48 countries, quite different with respect to wealth, size and geography, was considered. The sample includes twelve Western European, five Eastern European, nine African, seven American and fourteen Asian countries. The used data corresponds to the last available census.

For each country a rank/frequency log–log plot was built. Firstly, the data was collected, sorted and ranked and then, a normalization of the values was carried out. That is, the data (x-axis) was divided by the value corresponding to the population of the largest city, and the rank (y-axis) was divided by the rank of the smallest city, yielding intervals in the range [0, 1]. Finally, a PL was adjusted using a least squares algorithm.

Fig. 1 depicts the result obtained for Turkish cities. It is clear that a PL having parameters ð Æ; ±Þ ¼ ð0:0022; 1:021Þ fits very well to the available data, yielding a squared correlation coefficient of R^2¼ 0:9955.

We computed the values of the PL parameters ðÆ; ±Þ for the set of 48 countries. The obtained values for the coefficient, Æ, are in the range 0:0009 6 Æ 6 0:0332, where the extreme values correspond to UK and to Bolivia, respectively. With respect to the exponent ±, it is comprised in the interval 0:661 6 ± 6 1:648. The minimum and maximum values correspond to

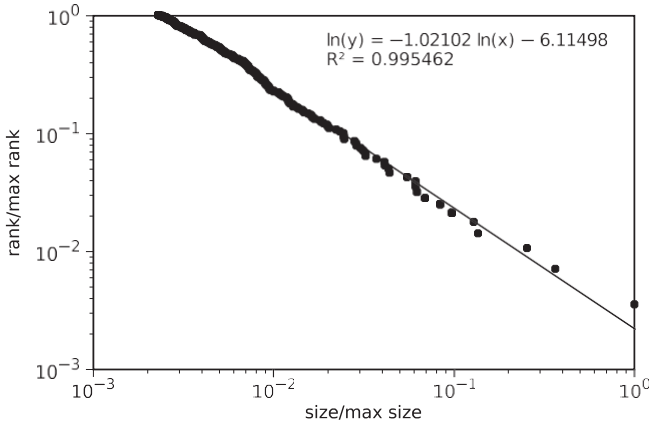

Fig. 1. Rank/frequency log–log plot of the size of populations of the Turkish cities.

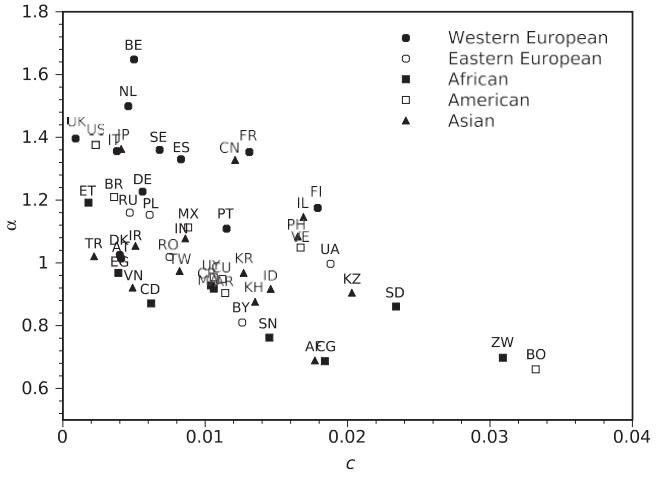

Fig. 2. Locus of (c, α) points for the sample of 48 analyzed countries.

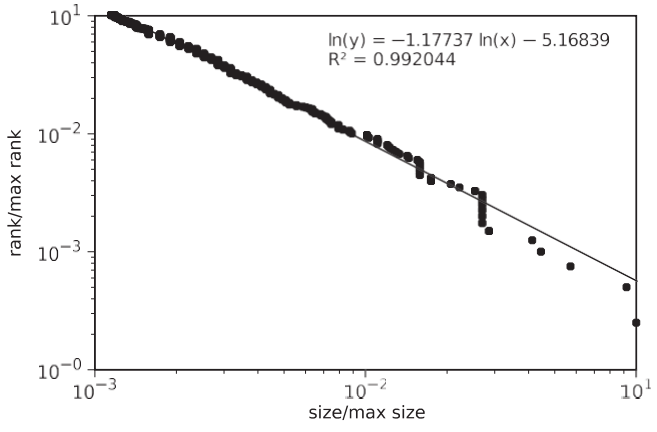

Fig. 3. Rank/frequency log–log plot of the wealth of the 400 richest Americans in 2000.

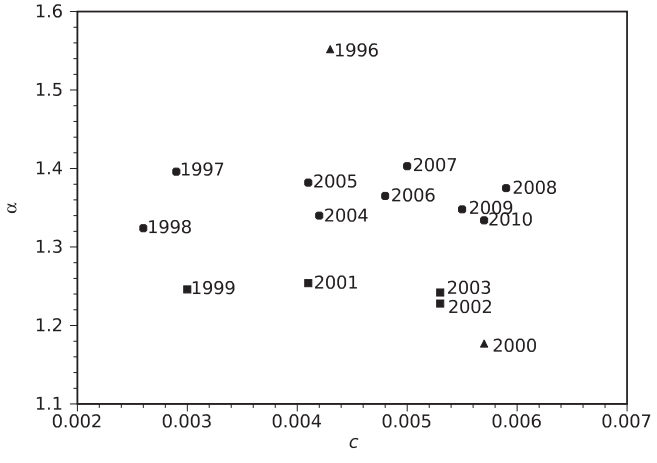

Fig. 4. Locus of (c, α) points for the distribution of the wealth of the 400 richest Americans.

Bolivia and to Belgium, respectively. Fig. 2 shows the locus of exponents α versus coefficients c. It can be observed that the countries presenting higher wealth levels, as is the case of Western European countries, Japan and USA, reveal higher values

of the exponent α and smaller to moderate values of coefficient c. Conversely, most African countries unveil smaller values of α and higher values of c. This leads to clusters in the left upper and right lower parts of the locus for the wealthy and not-so-wealthy countries, respectively.

8.2. Forbes 400 richest Americans

This subsection analyses the cumulative distribution of the wealth of the 400 richest individuals in the United States, measured in billions of dollars. The data was collected from the American Forbes (http://www.forbes.com/) that publishes the Forbes magazine, an important periodical in the business area. The Forbes magazine is well-known for its lists, including the Forbes 400, which lists the 400 richest Americans every year.

The last fifteen years of the Forbes 400 list, from 1996 up to 2010 were analyzed. For each year the collected data was sorted and ranked. After that, normalization took place, in a similar way to the described previously for the populations of cities. The normalized data was represented in a log–log plot and a PL adjusted using the least squares algorithm. For example, Fig. 3 depicts the result obtained for the year 2000. In that year the richest person was Bill Gates, with a fortune of 63 billions of dollars. At the bottom of the list were Christel Dehaan, Malcolm Glazer and Richard Haworth, with a fortune of 0.725 billions of dollars. It can be seen that the distribution of wealth unveils a clear statistical regularity in a range of almost two orders of magnitude.

The PLs coefficients were calculated for each year. The resulting values are in the ranges $0{:}0026 \leq c \leq 0{:}0059$ and $1{:}177 \leq \alpha \leq 1{:}552$. Fig. 4 shows the locus of the exponents, α, versus the coefficients, c. In terms of α, two periods can be highlighted: the first one corresponds to the periods 1997–1998 and 2004–2010, and the second corresponds to the years 1999 and 2001–2003. For both periods, although the coefficient c varies moderately, the exponent α reveals small changes. Moreover, the transitions from years 1996 to 1997 and from 1999 to 2000 indicate discontinuities on the plane (c, α).

8.3. Largest private American companies

We analysed the Forbes largest private American companies http://www.forbes.com/, with respect to their annual revenue, measured in billions of dollars, from 1996 up to 2010 (exception for 2005, because the corresponding data was not available).

Following a procedure similar to the one depicted in the previous sections, the results depicted in Figs. 5 and 6 were obtained. As an example, the frequency/rank plot, corresponding to year 2010, is shown in Fig. 5. This result unveils a remarkable statistical regularity over the distribution of companies' annual revenue, in a range of almost two orders of magnitude. It is worth mentioning that in 2010 the largest private company in the USA was Cargill, in the business area of farm products and annual revenue of 109.84 billions of dollars. At the end of the list, in the 223rd place, was ShopKo Stores, with annual revenue of 2 billions of dollars.

The PL parameters $(c; \alpha)$ were calculated for each year over the considered period. The resulting values of c and α are in the ranges $0{:}0020 \leq c \leq 0{:}0050$ and $1{:}296 \leq \alpha \leq 1{:}490$, respectively. Fig. 6 shows the locus of $(c; \alpha)$ points. It can be seen that there are four clusters on the plane $(c; \alpha)$. One corresponds to the years 1996–1998 plus 2007–2008, another one corresponds to the years 1999–2000, 2006 and 2010, a the third one comprises the period 2001–2004, and, finally, year 2009 appears quite apart from the others.

8.4. Forest fires

In this subsection, the analysis of the cumulative distribution of the size of forest fires in California, USA, over the ten year period 2001–2010, is presented. The adopted measure to quantify the size of a forest fire is the total burned area, usually

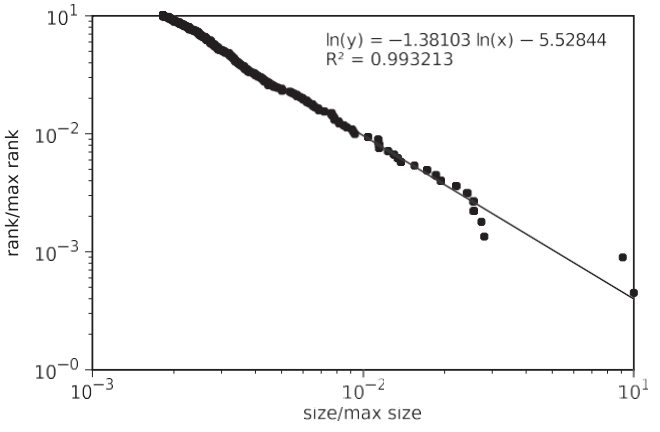

Fig. 5. Rank/frequency log–log plot of the largest private American companies, year 2010.

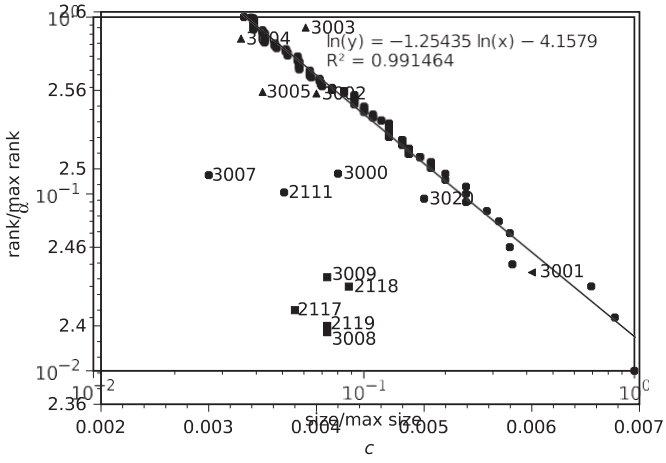

Fig. 6. Locus of (c, α) points for the distribution of the size of the largest private American companies.

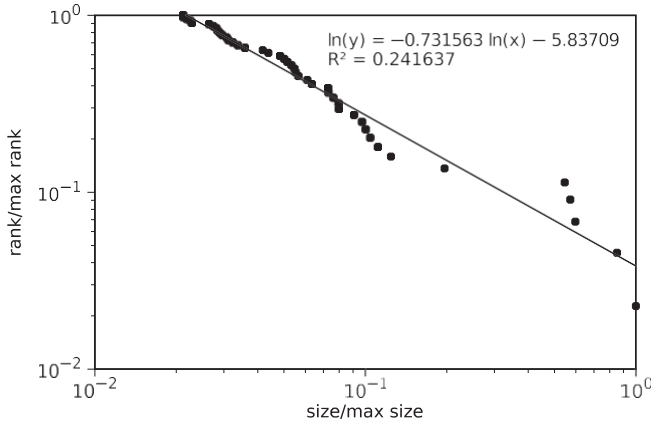

Fig. 7. Rank/frequency log–log plot of the size of Californian forest fires, year 2010.

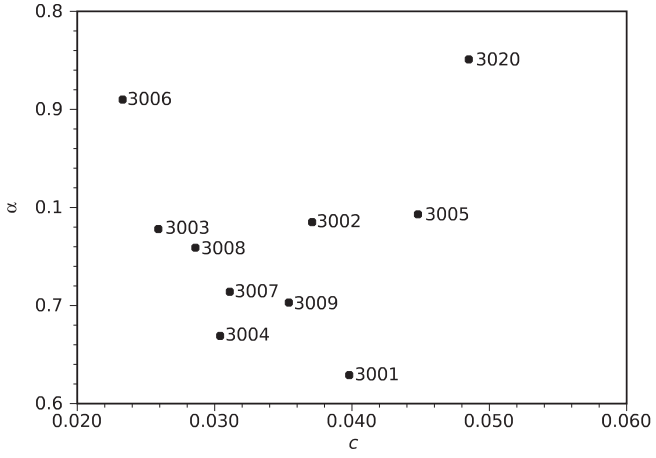

Fig. 8. Locus of (c, α) points for the distribution of the size of forest fires in California.

expressed in Acres. The used data is available on the California Department of Forestry & Fire Protection web site http://www.fire.ca.gov/ and considers all fires greater than 300 acres in total burned area.

Fig. 9. Rank/frequency log–log plot of the number of unique visitors of the 100 American most-visited web sites, January 2011.

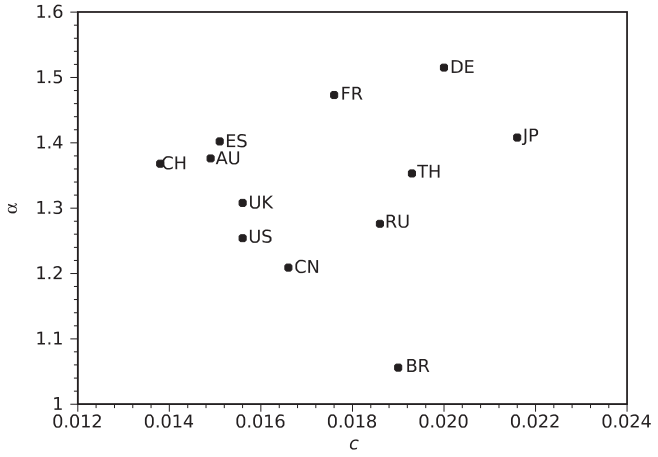

Fig. 10. Locus of ð€; ªÞ points for the number of unique visitors of the 100 most-visited web sites in the sample of analyzed countries.

Fig. 7 represents the rank/frequency log–log plot of the data corresponding to the year 2010. In this example, it can be seen that the data fits to a PL with parameters ð€; ªÞ ¼ ð0:0385; 0:851Þ. In the analyzed ten years period, the computed values vary are in the range 0:0133 6 € 6 0:0385 and 0:529 6 ª 6 0:851. From Fig. 8, depicting the locus of ð€; ªÞ points, it can be observed a moderate variability both in the exponent ª and the coefficient C€ However, there is a significant randomness in the transitions between consecutive years.

8.5. Most visited web sites

The cumulative distribution of the number of unique visitors of the 100 most-visited web sites in January 2011 is analysed in this subsection. In this context, unique visitors mean the estimated, unduplicated number of people who visit a site over a specific period of time. The adopted data can be collected from the Google's DoubleClick Ad Planner web site http://www.google.com/adplanner/static/top100countries/index.html, a free media planning tool that supplies this information for all web sites except Google's itself.

The rank/frequency log–log plot of the data corresponding to the USA is shown in Fig. 9. The most visited web site was facebook.com, with 130 million visitors and, at the bottom was lot.com, receiving 4.7 million visitors. It can be seen that the data fits very well to a PL with parameters ð€; ªÞ ¼ ð0:0156; 1:254Þ. Moreover, the computed values of € and ª, for a set of twelve analyzed countries, appeared in the ranges 0:0138 6 € 6 0:0216 and 1:056 6 ª 6 1:515. Fig. 10 shows the locus of ð€; ªÞ points. It can be observed a fairly low variability in both PL parameters.

8.6. Terrorist attacks

In this subsection it is analyzed the severity of terrorist attacks, in the period 1999–2009, according to the RAND Corporation's concept of terrorism http://www.rand.org. The severity of each event is quantified by the number of people killed

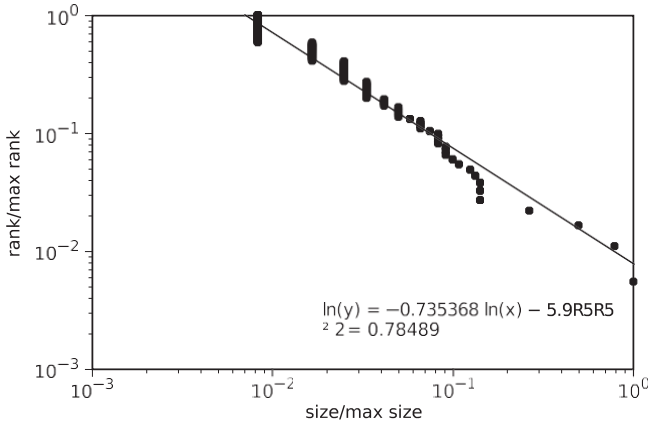

Fig. 11. Rank/frequency log–log plot of the severity of terrorist attacks worldwide, year 1999.

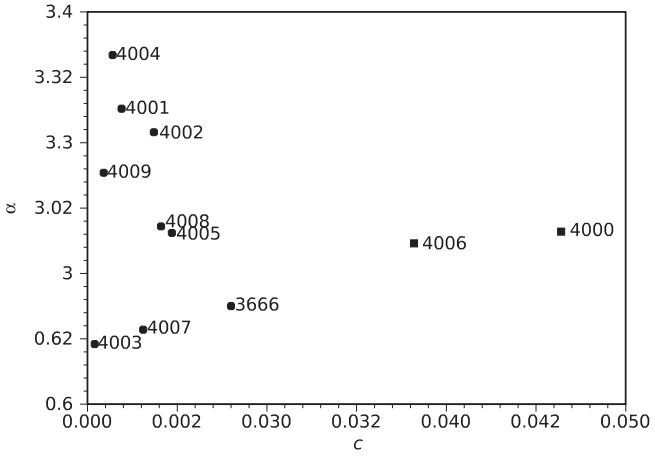

Fig. 12. Locus of ðC; ⍺Þ points for the distribution of the severity of terrorist attacks worldwide.

in it. The used data is available in the RAND Database of Worldwide Terrorism Incidents (RDWIT) http://smapp.rand.org/rwtid that holds a compilation of data related to incidents dated from 1972 up to 2009.

As an example, in 1999 people resulted killed in 180 incidents. Additionally, the records show that in the severest event 121 people died and there were 75 incidents resulting in one dead. The rank/frequency log–log plot of the data corresponding to year 1999 is shown in Fig. 11. An identical behavior can be observed for the whole decade. A PL with parameters ðℰ; ⍺Þ ¼ ð0:008; 0:975Þ fits well to the data. For the analyzed period, the PL parameters vary in the intervals 0:0004 ⩽ ℰ ⩽ 0:0264 and 0:946 ⩽ ⍺ ⩽ 1:167.

Fig. 12 shows the locus of the exponent ⍺ versus the parameter ℰ. In this case, although ⍺ varies slightly, the parameter ℰ unveils strong variability. According to the values of ℰ , two clusters can be identified: one contains years 2000 and 2009 and the other includes all other years. Moreover, the transitions to and from year 2000, from 2001 to 2002 and from 2008 to 2009 indicate large discontinuities on the plane ðℰ; ⍺Þ.

8.7. Tornadoes

A tornado is a rotating column of air, either pendant from a cumuliform cloud, or underneath a cumuliform cloud that is in contact with the ground. Very often it appears like a funnel cloud.

This subsection analyzes the severity of tornadoes, in the USA, from 2000 up to 2010. The total number of human victims (deaths and injured) directly related to a given event is used to quantify the severity of a tornado. The used data is available at the United States Department of Commerce's National Oceanic and Atmospheric Administration http://www.noaa.gov/, National Weather Service, Storm Prediction Center web site http://www.spc.noaa.gov/.

The rank/frequency log–log plot corresponding to year 2000 is depicted in Fig. 13 and a PL with parameters ðℰ; ⍺Þ ¼ ð0:022; 0:697Þ is adjusted to the data. In that year, the severest event caused 186 victims, while 28 events were re-

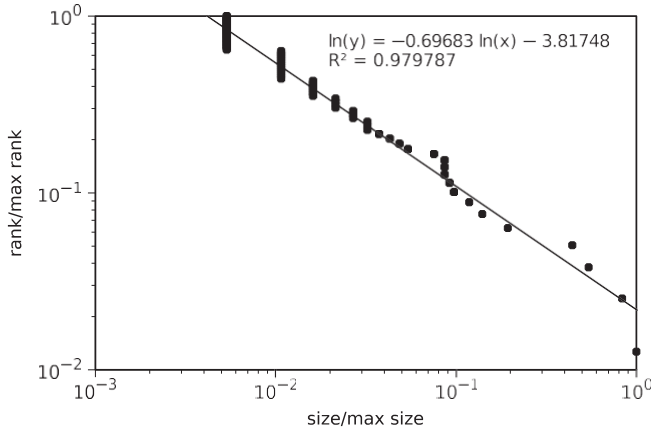

Fig. 13. Rank/frequency log–log plot of the severity of tornadoes in the USA, year 2000.

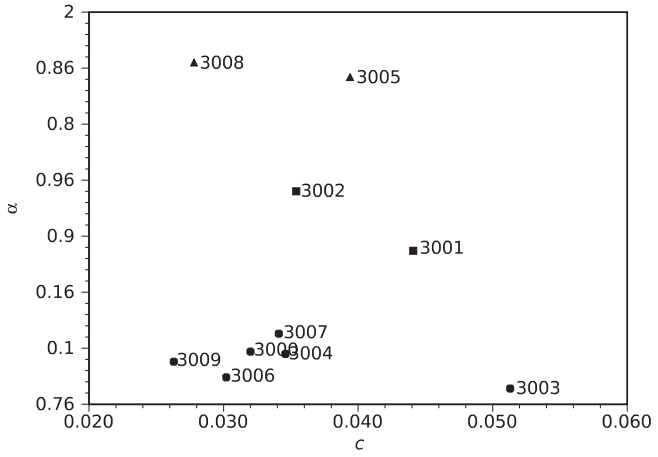

Fig. 14. Locus of ð€; ª Þ points for the distribution of the severity of tornadoes in the USA.

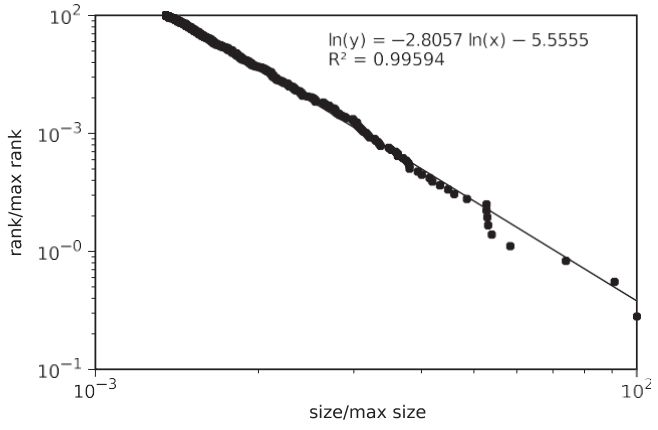

Fig. 15. Rank/frequency log–log plot of the size of craters on the Moon.

corded with one victim. This kind of behavior can be observed for every year in the analyzed period. The values of the parameters of the PLs vary in the interval 0:0163 6 € 6 0:0413 and 0:664 6 ª 6 0:955. Fig. 14 represents the locus of exponents ª versus parameters €. With respect to the values of ª, three clusters can be identified: one includes years 2000, 2002–2003,

2005–2006 and 2008, another one includes years 2001 and 2007, finally, a third cluster includes years 2004 and 2009. On the other hand, abrupt transitions on the plane ðC; αÞ can be observed between consecutive years.

8.8. Craters on planets

Another interesting natural phenomenon that follows a PL is the cumulative distribution of the size of craters, measured by their diameter, on planets and moons belonging to the solar system. In this subsection we analyzed the rank/frequency log–log graphs of three planets, namely Mercury, Venus and Mars, and three moons, namely Moon, Calisto and Ganymede. The data is available on Wikipedia http://en.wikipedia.org/wiki/Impact_crater.

As an example, Fig. 15 represents the rank/frequency plot for the data corresponding to the Moon. A PL fits very well to the data for values of the measured variable greater than 80 km. A similar behavior can observed for the six analyzed planets. The parameters of the PLs vary in the intervals $0{:}0039 \leq \varepsilon \leq 0{:}0455$ and $2{:}600 \leq \alpha \leq 3{:}009$. Statistical regularity is revealed in a range of almost one order of magnitude. Fig. 16 depicts the locus of ðε; αÞ points for the six planets. This natural phenomenon reveals values of α considerable greater than those found for the other phenomena. On the other hand, parameter ε varies significantly.

8.9. Earthquakes

An earthquake is a consequence of a sudden release of energy in the crust of the Earth. The propagated seismic waves may cause serious damage over large areas, depending on the earthquake magnitude and depth. The most popular scale to measure the magnitude of an earthquake is the Richter scale, also known as Local Magnitude scale ðMW Þ. In fact, it is a quantitative logarithmic scale that was introduced in 1930, to measure earthquakes size in Southern California. This scale presents several limitations because it saturates as higher values of earthquake magnitudes. Therefore, the Moment Magnitude scale ðMW Þ was established in 1977, preserving consistency with the Richter scale for moderate amplitude earthquakes. This scale gives most reliable estimates for moderate and large amplitude earthquakes and, nowadays, is the most used scale. The energy, E, released in an earthquake can be computed by $E \propto 10^{1{:}5 M_W + 4{:}8}$, which means that an earthquake having moment magnitude one point larger than another, will release 32.6 times more energy. In this subsection we study the cumulative distribution of the size of earthquakes occurred in California between 1999 and 2010 (except for 2007, because there is no available data), as recorded in the Southern California Earthquake Data Center http://www.data.scec.org/index.html. The data includes all earthquakes having moment magnitude $M_W \geq 3{:}5$, latitude and longitude from 32°N up to 37°N and 114°W up to 120°W, respectively, and depth smaller than 1000 km.

The rank/frequency log–log plot of the data corresponding to year 2009 is shown in Fig. 17. The statistical behavior observed for the measured variable is identical for all years in the analyzed period. A PL with parameters $(\varepsilon; \alpha) = (0{:}0086; 0{:}612)$ fits quite well to the data. For the entire period, the PL parameters vary in the intervals $0{:}0016 \leq \varepsilon \leq 0{:}0543$ and $0{:}439 \leq \alpha \leq 0{:}778$. Fig. 18 depicts the locus of (c, α) points, revealing low variability in both PL parameters, with the exception of the value of c for year 2008.

8.10. A global perspective

In this subsection we compare the previous results. Fig. 19 shows the temporal behavior of the exponents α, for the set of phenomena that were studied in an annual basis. It can be seen that, accordingly to α, the phenomena are grouped in three

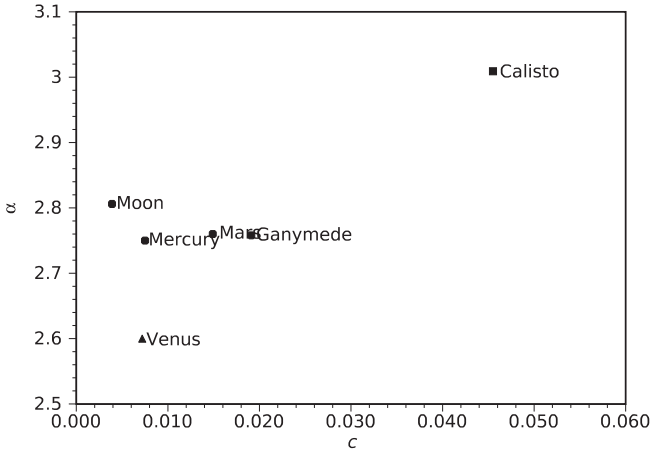

Fig. 16. Locus of ðε; αÞ points for the distribution of the size of craters.

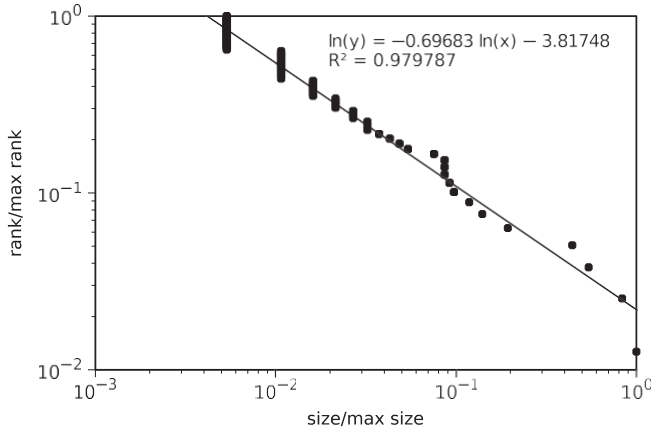

Fig. 17. Rank/frequency log–log plot of the size of earthquakes in California, year 2009.

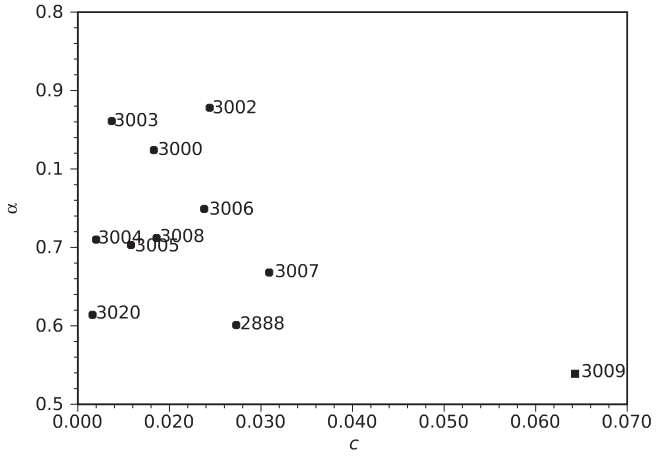

Fig. 18. Locus of $\{\epsilon, \alpha\}$ points for the distribution of the size of earthquakes in California.

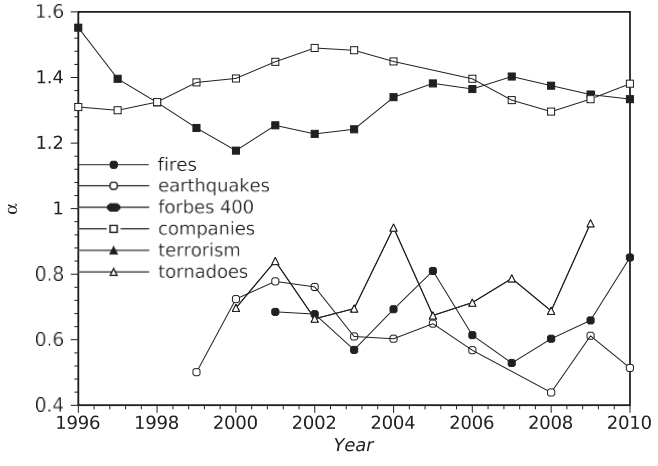

Fig. 19. Temporal evolution of the parameter α for the phenomena studied in an annual basis.

clusters. The distribution of the wealth of the 400 richest individuals in the United States and the largest private American companies constitute one cluster. It can be noticed that the corresponding curves, in Fig. 19, display an interesting behavior,

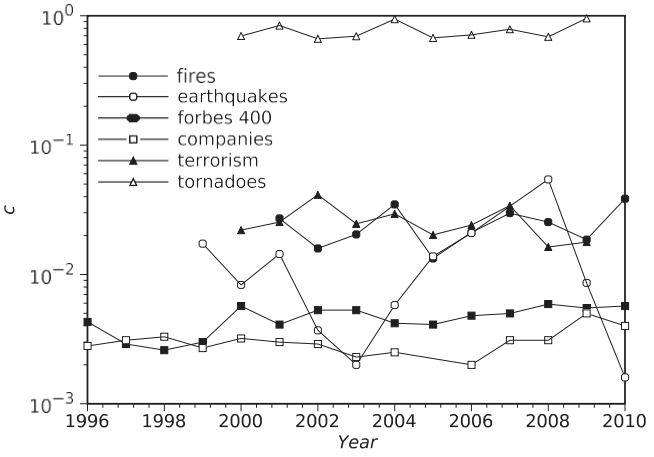

Fig. 20. Temporal evolution of the parameter $\hat{c}$ for the phenomena studied in an annual basis.

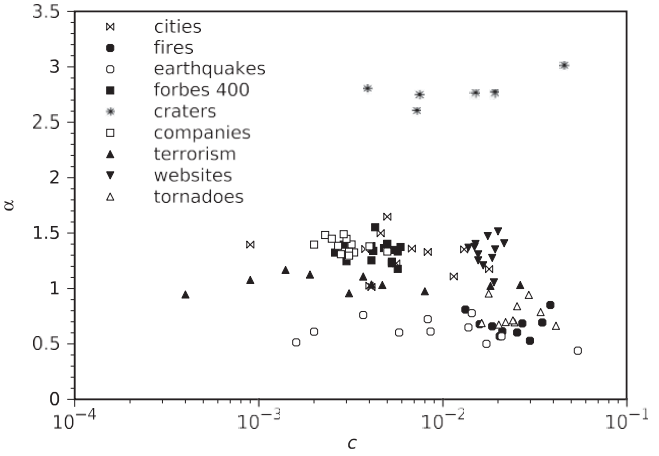

Fig. 21. Locus of $(\hat{c}, \alpha)$ points for the whole set of nine phenomena.

being in opposition of phase. The size of forest fires, the severity of tornadoes and the size of earthquakes can be grouped into another cluster. Finally, the severity of terrorist attacks is by itself one cluster.

Parameter $\hat{c}$ reveals a more complex temporal evolution (Fig. 20). Nevertheless, with respect to the wealth of the 400 richest Americans and the largest private American companies, the behavior of the parameters $\hat{c}$ and α is similar. Moreover, in the severity of tornadoes and size of forest fires, the parameter $\hat{c}$ varies only slightly. In the contrary, the severity of terrorist attacks and size of earthquakes have the associated parameters $\hat{c}$ varying strongly.

Fig. 21 shows the locus of $(\hat{c}, \alpha)$ points for the whole set of nine phenomena. In order to improve the readability of the figure, a logarithmic scale was adopted for the $\hat{c}$-axis and just cities of Western European countries were represented. It can be seen that the size of craters, the size of Western European cities, the size of earthquakes and the severity of terrorist attacks, spread significantly along the $\hat{c}$ axis. The wealth of the 400 richest individuals in the United States and the largest private American companies are very close in the $(\hat{c}, \alpha)$ plane. Besides, they present values for α close to those showed by the 100 most-visited web sites and Western European cities. A similar behavior is unveiled by the size of forest fires and the severity of tornadoes, which also present values of α close to those associated to the sizes of earthquakes. On the other hand, the size of craters has values of α clearly different from all the other phenomena. In conclusion, the locus of $(\hat{c}, \alpha)$ points revealed global characteristics of phenomena described by PLs and useful consequences will be drawn by adding new data and phenomena.

9. Conclusion

In this paper we reviewed interesting and important results on PL distributions and their applications to the modeling of distinct real life phenomena. We also unraveled some strong controversies around power laws, such as the consideration

that PLs are just statistical phenomena, or the lack of strong proves of the validity of the PL behavior. Researchers are acknowledging the need to make PL results scientifically strong and more directly useful. This is critical since papers invalidating previous PL modelings of some phenomena keep appearing and assumptions made for certain PL models are way far from reasonable. In one point we all agree, PL research has contributed to put scientists, of what seemed, completely distinct areas of knowledge to work together and to exchange experience and expertise.

Acknowledgments

Research funded by the European Regional Development Fund through the programme COMPETE and by the Portuguese Government through the FCT – Fundação para a Ciência e a Tecnologia under the project PEst-C/MAT/UI0144/2011.

References

[1] Adamic LA, Huberman BA. The nature of markets in the World Wide Web. Q J Electron Comm 2000;1:512.
[2] Adamic LA, Huberman BA. Zipfs law and the Internet. Glottometrics 2002;3:143–50.
[3] Aiello W, Chung F, Lu L. A random graph model for massive graphs. In: Proceedings of the 32nd annual ACM symposium on theory of computing, 2000. p. 171–80.
[4] Anazawa M, Ishikawa A, Suzuki T, Tomoyose M. Physica A 2004;335:616. Available from: cond-mat/0307116.
[5] Anderson C. A methodology for estimating Amazon's long tail sales. <http://longtail.typepad.com/the long tail/2005/08/a methodology f.html>; August 2005.
[6] Anderson G, Ge Y. The size distribution of Chinese cities. Reg Sci Urban Econ 2005;35:756–76.
[7] Andersson C, Hellervik A, Lindgren K. A spatial network explanation for a hierarchy of urban power laws. Physica A 2005;345:227–44.
[8] Ayoama H, Fujiwara Y, Souma W. Kinematics and dynamics of Pareto–Zipf's law and Gibrat's law. Physica A 2004;344:117–21.
[9] Auerbach F. Das Gesetz der Belvolkerungskoncentration. Petermann Geogr Mitt 1913;59:74–6.
[10] Bak P, Chen K, Tang C. A forest-fire model and some thoughts on turbulence. Phys Lett A 1990;147(5–6):297–300.
[11] Barford P, Bestavros A, Bradley A, Crovella M. Changes in Web client access patterns: characteristics and caching implications. World Wide Web 1999;2:15–28.
[12] Batty M. Cities and complexity: understanding cities through cellular automata, agent-based models, and fractals. Cambridge, MA: MIT Press; 2005.
[13] Benguigui L, Blumenfeld-Lieberthal E. The temporal evolution of the city size distribution. Physica A 2009;388(7):1187–95.
[14] Bi Z, Faloutsos C, Korn F, The DGX distribution for mining massive, skewed data. In: Proceedings of the seventh ACMSIGKDD international conference on knowledge discovery and data mining, San Francisco, California, 2001.
[15] Blank A, Solomon S. Power laws in cities population, financial markets and internet sites (scaling in systems with a variable number of components). Physica A 2000;287:279–88.
[16] Bohorquez JC, Gourley S, Dixon AR, Spagat M, Johnson NF. Common ecology quantifies human insurgency. Nature 2009;462(7275):911–4.
[17] Bosker M, Brakman S, Garretsen H, Schramm M. A century of shocks: the evolution of the German city size distribution 1925–1999. Reg Sci Urban Econ 2008;38(4):330–47.
[18] Brakman S, Garretsen H, van Marrewikj C, van de Berg M. The return of Zipf: towards a further understanding of the rank-size distribution. J Reg Sci 1999;39(1):182–213.
[19] Brynjolfsson E, Hu YJ, Smith MD. Consumer surplus in the digital economy: estimating the value of increased product variety at online booksellers. MIT sloan working paper No. 4305-03. Available at: SSRN: <http://ssrn.com/abstract=400940> or doi:10.2139/ssrn.400940; June 2003.
[20] Caron Y, Makris P, Vincent N. Use of power law models in detecting region of interest. Pattern Recognit 2007;40:2521–9.
[21] Carson M, Langer JS. Mechanical model of an earthquake fault. Phys Rev A 1989;40:6470–84.
[22] Cederman LE. Modeling the size of wars: from billiard balls to sandpiles. Am Polit Sci Rev 2003;97:135–50.
[23] Champernowne DG. A model of income distribution. Econ J 1953;63:318–51.
[24] Chatterjee A, Chakrabarti BK, Manna SS. Pareto law in a kinetic model of market with random saving propensity. Physica A 2004;335:155–63.
[25] Chebotarev AM. On stable Pareto laws in a hierarchical model of economy. Physica A 2007;373:541–59.
[26] Chevalier J, Goolsbee A. Measuring prices and price competition online: Amazon.com and BarnesandNoble.com. Quant Mark Econ 2003;1:203–22.
[27] Clauset A, Young M. Scale invariance in Global Terrorism. e-print physics/0502014 at <http://xxx.lanl.gov>.
[28] Clauset A, Young M, Gleditsch KS. On the frequency of severe terrorist events. J Conflict Resolut 2007;51(1):58–87.
[29] Cooper C, Frieze A. On a general model of undirected Web graphs. In: Proceedings of the 9th annual European symposium on algorithms, 2001. p. 500–11.
[30] Cox RAK, Felton JM, Chung KC. The concentration of commercial success in popular music: an analysis of the distribution of gold records. J Cult Econ 1995;19:333–40.
[31] Crawley MJ, Harral JE. Scale dependence in plant biodiversity. Science 2001;291:864–8.
[32] Crovella ME, Bestavros A. Self-similarity in World Wide Web traffic: evidence and possible causes. In: Gaither BE, Reed DA, editors. Proceedings of the 1996 ACM SIGMETRICS conference on measurement and modeling of computer systems. New York: Association of Computing Machinery; 1996. p. 148–59.
[33] Crovella M, Taqqu MS, Bestavros A. Heavy-tailed probability distributions in the world wide web Chapter 1. In: Adler RJ, Feldman RE, Taqqu MS, editors. A pratical guide to heavy tails. Chapman and Hall; 1998. p. 3–26.
[34] Crow EL, Shimizu K. Lognormal distributions: theory and application. New York: Dekker; 1988.
[35] Dahui W, Menghui L, Zengru D. True reason for Zipf's law in language. Physica A 2005;358:545–50.
[36] Das A, Yarlagadda S. An analytical treatment of the Gibbs–Pareto behavior in wealth distribution. Physica A 2005;353:529–38.
[37] Davis DR, Weinstein DE. Bones, bombs, and break points: the geography of economic activity. Am Econ Rev 2002;92(5):1269–89.
[38] Dobkins L, Ioannides Y. Dynamic evolution of the US citysize distribution. In: Huriot J-M, Tisse J-F, editors. The economics of cities. New York, NY: Cambridge University Press; 1998.
[39] Dorogovtsev SN, Mendes JFF, Oliveira JG. Frequency of occurrence of numbers in the World Wide Web. Physica A 2006;360:548–52.
[40] Dragulescu AA, Yakovenko VM. Statistical mechanics of money, income, and wealth: a short survey. Eur Phys J B 2001;20:585–9.
[41] Drossel B, Schwbal F. Self-organized critical forest-fire model. Phys Rev Lett 1992;69(11):1629–32.
[42] Estoup JB. Gammes stenographiques. Institut de France; 1916.
[43] Fahidy TZ. Applying Pareto distribution theory to electrolytic powder production. Electrochem Commun 2011.
[44] Fenner T, Levene M, Loizou G. Predicting the long tail of book sales: unearthing the power-law exponent. Physica A 2010;389(12):2416–21.
[45] Figueira FC, Moura NJ, Ribeiro MB. The Gompertz–Pareto income distribution. Physica A 2011;390:689–98.
[46] Fujiwaraa Y, Aoyamac H, Guilmib CD, Souma W, Gallegati M. Gibrat and ParetoZipf revisited with European firms. Physica A 2004;344:112–6.
[47] Gabaix X. Zipf's law and the growth of cities. In: American economic review papers and proceedings, vol. LXXXIX, 1999. p. 129–32.
[48] Gabaix X. Zipf's law for cities: an explanation. Q J Econ 1999;114:739–67.
[49] Gan L, Li Dong, Song S. Is the Zipf law spurious in explaining city-size distributions? Econ Lett 2006;92:256–62.

[50] Gibrat R. Les inégalités économiques. Paris, France: Librairie du Recueil Sirey; 1931.
[51] Giesen K, Zimmermann A, Suedekum J. The size distribution across all cities Double Pareto lognormal strikes. J Urban Econ 2010;68:129–37.
[52] Gopikrishnan P, Plerou V, Liu Y, Amaral LAN, Gabaix X, Stanley HE. Scaling and correlation in financial time series. Physica A 2000;287(3–4):362–73.
[53] Gutenberg B, Richter RF. Frequency of earthquakes in California. Bull Seismol Soc Am 1944;34:185–8.
[54] Huberman BA, Adamic LA. Growth dynamics of the World-Wide Web. Nature 1999;401:131.
[55] Iba T, Yoshida M, Fukami Y, Saitoh M, Power-law distribution in Japanese book sales market, in: Fourth joint Japan–North America mathematical sociology conference, 2008.
[56] Ijiri Y, Simon HA. Skew distributions and the sizes of business firms. Amsterdam: North Holland; 1977.
[57] Ishikawa A. Pareto law and Pareto index in the income distribution of Japanese companies. Physica A 2005;349:597–608.
[58] Ishikawa A. Annual change of Pareto index dynamically deduced from the law of detailed quasi-balance. Physica A 2006;336:367–76.
[59] Israeloff NE, Kagalenko M, Chan K. Can Zipf distinguish language from noise in noncoding DNA? Phys Rev Lett 1995;76:1976–9.
[60] Johnson NF, Spagat M, Restrepo JA, Becerra O, Bohorquez JC, Suarez N, Restrepo EM, Zarama R. Universal patterns underlying ongoing wars and terrorism. <arXiv:physics/0605035v1>; 2006.
[61] Krapivsky PL, Redner S. Organization of growing random networks. Phys Rev E 2001;63:066123.
[62] Krugman P. The self-organizing economy. Cambridge, MA: Blackwell; 1996.
[63] Lévy M, Solomon S. New evidence for the power-law distribution of wealth. Physica A 1997;242:90.
[64] Li W. References on Zipf's law. <http://www.nslij-genetics.org/wli/zipf/>.
[65] Lin W, Yang Y. Zipf's law in importance of genes for cancer classification using microarray data. J Theor Biol 2002;219:539–51.
[66] Lotka AJ. The frequency distribution of scientific productivity. J Washington Acad Sci 1926;16:317–24.
[67] Lu ET, Hamilton RJ. Avalanches of the distribution of solar flares. Astrophys J 1991;380:89–92.
[68] Malamud BD, Morein G, Turcotte DL. Forest fires: an example of self-organized critical behavior. Science 1998;281(5384):1840–2.
[69] Mandelbrot B. The Pareto–Levy law and the distribution of income. Int Econ Rev I 1960:79–106.
[70] Mantegna RN, Stanley HE. The scaling behavior of an economic index. Nature 1995;376:46–9.
[71] Mitzenmacher M. A brief history of generative models for power law and lognormal distributions. Int Math 2004;1:226–51.
[72] Mizunoa T, Toriyama M, Terano T, Takayasu M. Pareto law of the expenditure of a person in convenience stores. Physica A 2008;387:3931–5.
[73] Moura Jr NJ, Ribeiro MB. Zipf law for Brazilian cities. Physica A 2006;367:441–8.
[74] Neukum G, Ivanov BA. Crater size distributions and impact probabilities on Earth from lunar, terrestial-planet, and asteroid cratering data. In: Tucson AZ, Gehrels T, editors. Hazards due to comets and asteroids. University of Arizona Press; 1994. p. 359–416.
[75] Newman MEJ. Power laws, Pareto distributions and Zipf's law. Contem Phys 2005;46:323–51.
[76] Nitsch V. Zipf zipped. J Urban Econ 2005;57:123–46.
[77] Okuyama K, Takayasu M, Takayasu H. Zipfs law in income distribution of companies. Physica A 1999;269:125–31.
[78] Overman HG, Ioannides Y. Zipf law for cities: an empirical examination. London: Centre for Economic Performance, London School of Economics and Political Science; 2000.
[79] Pareto V. Cours d'economie politique. Geneva, Switzerland: Droz; 1896.
[80] Parr J. Note on the size distribution of cities over time. J Urban Econ 1985;XVIII:199–212.
[81] Pinho STR, Andrade RFS. An abelian model for rainfall. Physica A 1998;255:483–95.
[82] Price DJ de S. Networks of scientific papers. Science 1965;149:510–5.
[83] Redner S. How popular is your paper? An empirical study of the citation distribution. Eur Phys J B 1998;4:131–4.
[84] Reed WJ. The Pareto law ofincomes – an explanation and an extension. Physica A 2003;319:469–86.
[85] Richardson LF. Variation of the frequency of fatal quarrels with magnitude. J Am Stat Assoc 1948;43:523–46.
[86] Richardson LF. Statistics of deadly quarrels. Chicago: Quadrangle Books; 1960.
[87] Ricotta C, Avena G, Marchetti M. The flaming sandpile: self-organized criticality and wildfires. Ecol Model 1999;119(1):73–7.
[88] Roberts DC, Turcotte DL. Fractality and selforganized criticality of wars. Fractals 1998;6:351–7.
[89] Rosen K, Resnick M. The size distribution of cities: an examination of the Pareto law and primacy. J Urban Econ 1980;8:165–86.
[90] Rozenfeld HD, Rybski D, Andrade Jr JS, Batty M, Stanley HE, Makse HA. Laws of population growth. Proc Natl Acad Sci USA 2008;105:18702–7.
[91] Rozman G. East Asian urbanization in the nineteenth century. In: van der Woude AD, Hayami Akira, Vries Jan de, editors. Urbanization in history: a process of dynamic interactions. Oxford: Clarendon Press; 1990. p. 61–73.
[92] Sarabia JM, Prieto F. The Pareto-positive stable distribution: a new descriptive model for city size data. Physica A 2009;388:4179–91.
[93] Satoh K, Weiguo S, Yang KT. A study of forest fire danger prediction system in Japan. In: 15th international workshop on database and expert systems applications (DEXA'04), 2004. p. 598–602.
[94] Schuster P, Fontana W, Stadler PF, Hofacker IL. From sequences to shapes and back: a case study in RNA secondaru structures. Proc R Soc Lond B 1994;255:279–84.
[95] Seuront L, Mitchell JG. Towards a seascape typology. I. Zipf versus Pareto laws. J Marine Syst 2008;69:310–27.
[96] Simon HA. On a class of skew distribution functions. Biometrika 1955;42:425–40.
[97] Song WG, Fan W, Wang B. Self-organized criticality of forest fires in China. Chin Sci Bull 2001;46(13):1134–7.
[98] Song WG, Zhang HP, Chen T, Fan WC. Power-law distribution of city fires. Fire Safety J 2003;38:453–65.
[99] Soo KT. Zipfs law for cities: a cross-country investigation. Reg Sci Urban Econ 2005;35(3):239–63.
[100] Sornette D. Critical phenomena in natural sciences. 2nd ed. Heidelberg: Springer; 2003. Chapter 14.
[101] Stroud PD, Sydoriak SJ, Riese JM, Smith JP, Mniszewski SM, Romero PR. Semi-empirical power-law scaling of new infection rate to model epidemic dynamics with inhomogeneous mixing. Math Biosci 2006;203:301–18.
[102] Tsallis C. Possible generalization of Boltzmann–Gibbs statistics. J Stat Phys 1988;52:479–87.
[103] Urzúa CM. A simple and efficient test for Zipf's law. Econ Lett 2000;66:257–60.
[104] Warren D. Examining city size distributions using urban areas. In: 45th Annual meeting of the western regional science association, Santa Fe, New Mexico, 2006.
[105] Weiguoa S, Jiana W, Kohyub S, Weicheng F. Three types of power-law distribution of forest fires in Japan. Ecol Model 2006;196:527–32.
[106] Wichmann S. On the power-law distributon of language family sizes. J Linguist 2005;41:117–31.
[107] Willis JC, Yule GU. Some statistics of evolution and geographical distribution in plants and animals, and their significance. Nature 1922;109:177–9.
[108] Zanette DH, Manrubia SC. Role of intermittency in urban development: a model of large-scale city formation. Phys Rev Lett 1997:523–6.
[109] Zanette DH, Manrubia SC. Vertical transmission of culture and the distribution of family names. Physica A 2001;295:1–8.
[110] Zipf G. Selective studies and the principle of relative frequency in language. Cambridge, MA: Harvard University Press; 1932.
[111] Zipf G. Human behavior and the priciple of least effort. Cambridge, MA: Addison-Wesley; 1949.