Acta Geophysica (2021) 69:395–404
https://doi.org/10.1007/s11600-021-00550-4

RESEARCH ARTICLE - SOLID EARTH SCIENCES

Entropy of earthquakes: application to Vrancea earthquakes

B. F. Apostol[1] · L. C. Cune[2]

Received: 28 October 2020 / Accepted: 2 February 2021 / Published online: 18 February 2021

Abstract
The entropy of earthquakes is derived by using the Gutenberg–Richter statistical distributions. Both canonical and microcanonical earthquake distributions are given, and Einstein's fluctuation formula is deduced for earthquakes. The seismic activity of Vrancea in the period 1980–2019 is analyzed, for earthquakes with magnitude greater than two, and the results are compared with the theoretical results. It is shown that the parameter of the magnitude distribution exhibits a tendency of increasing with time, due to the accumulation of small-magnitude earthquakes, interrupted from time to time by ruptures towards smaller values, caused by earthquakes with greater magnitudes. These variations do not obey the normal distribution of the fluctuations. The (small) time variations of the distribution parameter provide a measure of the departure of the seismic activity from an equilibrium process. For Vrancea, these deviations are very small (up to 1% per year).

Keywords Earthquake entropy · Fluctuation distribution · Vrancea earthquakes

Introduction

As it is well known, the seismic activity is viewed as an open-system non-equilibrium statistical process, in the sense that the energy released by earthquakes is not transformed into tectonic energy which would produce other earthquakes (at least not directly, or integrally). Consequently, it is important to estimate the departure of this process from equilibrium, in order to have more information about the rate of the seismic activity in a given region and a given period of time. The most convenient tools of characterizing a statistical process are the entropy and the probability density. We derive here the entropy of earthquakes, in terms of the parameter of the Gutenberg–Richter magnitude distribution, and use it as a reference point for the seismic activity in Vrancea in the period 1980–2019, for 8455 earthquakes with magnitude $M \geq 2$. We find that the rate of increase in the seismic activity in this space–time window is extremely small, and the fluctuations are, practically, absent (or, at least, they do not obey the usual Gaussian distribution).

The entropy of earthquakes has been introduced in seismology from Statistical Physics, as a measure of the disorder produced by the seismic activity in a given region and a given time interval (Berrill and Davis 1980; Shen and Mansinha 1983; De Santis et al. 2011). Also, the entropy is viewed as a measure of the information content (Main and Burton 1984; Masinha and Shen 1987). The maximum entropy principle has been employed in earthquake recurrence relationships (Dong et al. 1984) and earthquake hypocentre distributions (Nicholson et al. 2000). The statistical entropy characterizes a statistical ensemble. In equilibrium the distribution function does not change in time. If the time dependence is very slow, we say that the process is in quasi-equilibrium. Also, a (quasi-) equilibrium statistical process may exhibit fluctuations, which obey a Gaussian distribution. Deviations of the entropy from its (quasi-) equilibrium value indicate a non-equilibrium state. In this connection, the time dependence of the entropy of earthquakes is of primary interest, it being related more to Physical Kinetics approaches, or the information theory (Shannon 1948; Wiener 1948). Abrupt variations in the time dependence of the entropy (singularities of its derivative) are associated with a criticality regime, occurring near large-magnitude earthquakes (e.g. De Santis et al. 2019). We show here, by analyzing a large set of earthquakes of Vrancea region, that

Communicated by Ramon Zuñiga, Ph.D. (CO-EDITOR-IN-CHIEF).

✉ B. F. Apostol
afelix@theory.nipne.ro

[1] Institute of Earth's Physics, PO Box MG-6, Magurele-Bucharest, Romania

[2] Institute of Physics and Nuclear Engineering, PO Box MG-35, Magurele-Bucharest, Romania

the fluctuations are, practically, absent (or, at least, they do not obey the usual Gaussian distribution), the data indicating a steady departure from equilibrium, although very small. The steady decrease in the entropy is due to the occurrence of many small earthquakes, interrupted from time to time by rather large ruptures towards higher values, due to larger earthquakes.

The statistical analysis is based on a few general hypotheses. First, we view the seismic activity as a sequence of independent earthquakes with various magnitudes. The statistical ensemble should be as large as possible, with a large set of data recorded in a long period of time. We assume that the earthquake occurrence can be described by a statistical distribution in magnitude. We leave aside any possible correlations between earthquakes, like triggering or blocking effects, as well as possible correlations between foreshocks, aftershocks and main shocks. As discussed below, the correlations affect mainly the small-magnitude earthquakes, which are left aside in our analysis. The earthquake distribution may exhibit fluctuations. It is claimed that such fluctuations have been identified in yearly variations of global earthquake populations (Main and Al-Kindy 2002), looking as quasi-periodic variations (these results are obtained for rather small-sized samples). The existence of a time-independent statistical distribution implies a statistical equilibrium. The statistical equilibrium is characterized by (quasi-) stationarity, i.e. the statistical distribution may exhibit only slow variations, at most, in long periods of time. The parameter of the fluctuation distribution changes correspondingly in such a long-time seismic activity. The slow variations of the equilibrium may give the possibility to identify the variation tendency by relatively short-time sampling of the data. This short-time seismic activity may be affected by fluctuations, so we are led to analyze the fluctuation distribution. The analysis of a large number of earthquakes with magnitude $M \geq 2$ (8455 earthquakes), which occurred in Vrancea during the time interval (years) 1980–2019 indicates a departure from equilibrium, as expected. In particular, the variations of the statistical distribution do not obey the fluctuations normal distribution. The reason for this behaviour originates in the fact that the seismic activity is fed continuously by the tectonic energy, such that the statistical ensemble is not in (quasi-) equilibrium (at least over the analyzed periods of time), though the deviation from equilibrium is very small.

The technical means of analyzing a statistical process and its fluctuations is the entropy. We derive in this paper the Gutenberg–Richter statistical distribution in magnitude for a canonical ensemble by using the standard method of maximizing the entropy under the constraint of a fixed mean magnitude. This constraint characterizes the tectonic loading and the energy release by seismic activity. Similarly, by using the standard method of a microcanonical statistical ensemble, we derive the fluctuations normal distribution (known as Einstein's fluctuation formula). The parameter of this formula is precisely the parameter of the Gutenberg–Richter distribution. The rate of change in time of the parameter of this distribution provides a measure of the departure of the seismic activity from equilibrium. For Vrancea, in the analyzed time periods, we find a very small deviation, and, practically, the absence of fluctuations.

Gutenberg–Richter statistical distributions

The Gutenberg–Richter magnitude probability

$$P(M) = \beta \mathrm{e}^{-\beta M} \tag{1}$$

is well known in Seismology; in Eq. (1) M denotes the earthquake magnitude and β is a parameter. The statistical variable M takes all possible values, i.e. from 0 to ∞. The validity of this formula is checked by fitting the empirical data, over a limited range of magnitudes. As it is well known, the lack of completeness of the seismological catalogs introduces a cutoff for small magnitudes; for large magnitudes the probability density given above falls off rapidly to zero. Originally, this law was derived from empirical observations. Indeed, it is well documented by statistical analysis (Gutenberg and Richter 1944, 1956; Richter 1958; Bullen 1963) that the number of earthquakes $N(M)$ with magnitude greater than M is given by

$$\ln N(M) = \ln N(0) - \beta M \tag{2}$$

(cumulative, or exceedance, distribution); hence, $\mathrm{d}N/N(M) = -\beta \mathrm{d}M$, and the density of magnitude probability (Ranalli 1969) is

$$\frac{\Delta N}{N(0)\Delta M} = \beta \mathrm{e}^{-\beta M}. \tag{3}$$

If the total, large, number of earthquakes $N(0)$ occurs in a long time T, we can define the seismicity rate $1/t_0$, where $N(0) = T/t_0$, and Eq. (3) becomes

$$\frac{t_0 \Delta N}{T \Delta M} = \beta \mathrm{e}^{-\beta M}, \quad \frac{\Delta N}{T} = \frac{\beta \Delta M}{t_0} \mathrm{e}^{-\beta M}, \tag{4}$$

or

$$\ln\left(\frac{\Delta N}{T}\right) = \ln\left(\frac{\beta \Delta M}{t_0}\right) - \beta M; \tag{5}$$

hence, we may define a mean recurrence time

$$t_r = \frac{T \beta \Delta M}{\Delta N} = t_0 \mathrm{e}^{\beta M}. \tag{6}$$

for the earthquakes with magnitude M.

Equations (1), (3) and (5) define statistical distributions. They may be called Gutenberg–Richter statistical distributions. The deviation from the mean value of the statistical variable (magnitude M) is the standard deviation

$$\delta M = \left(\overline{M^2} - \overline{M}^2\right)^{1/2} = \left(-\frac{\partial \overline{M}}{\partial \beta}\right)^{1/2} = \frac{1}{\beta} = \overline{M}; \tag{7}$$

it follows that the deviation of the mean recurrence time given by Eq. (6) is $\delta t_r / t_r = 1$, which implies an error of the order $\left[\left(\overline{M^2}\right)^{1/2} - \overline{M}\right] / \overline{M} = \sqrt{2} - 1$, at least.

At the same time, the magnitude was introduced in Seismology as a logarithmic measure of the energy released by an earthquake. More precisely, it is assumed that the energy E of an earthquake with magnitude M is given by $E/E_0 = e^{bM}$, or

$$\ln E = \ln E_0 + bM, \tag{8}$$

where E_0 is an energy cutoff and $b = \frac{3}{2} \ln 10 = 3.45$, by convention ($\ln 10 \simeq 2.3$) (Utsu and Seki 1955; Utsu 1969). Later, the energy was related to the magnitude of the seismic moment, and the earthquake magnitude entering Eq. (8) was called moment magnitude (Kanamori 1977; Hanks and Kanamori 1979). Equations of the type (8) may be called Hanks–Kanamori (or Gutenberg–Richter) law. As we can see, they have a definition character.

The statistical Gutenberg–Richter distributions in time, energy and magnitude have been derived from a geometrical-growth model of accumulation of energy in the focal region (Apostol 2006a, b; the exact relationship between energy and the magnitude of the seismic moment was established in Apostol 2019a). According to this model the relation between the parameters β and b is $\beta = br$, where the parameter r is related to the number of effective dimensions of the focal region and the rate of energy accumulation. For a uniform pointlike focal region $r = 1/3$, for a two-dimensional focal region $r = 1/2$, while for a one-dimensional region r tends to unity. Very likely, the parameter r varies in the range $1/3 < r < 1$, which entails a variation $\frac{1}{2} \ln 10 < \beta < \frac{3}{2} \ln 10$ i.e. $1.15 < \beta < 3.45$). According to the theory of energy accumulation in the focal region, the relationship between the accumulation time t and the accumulated energy E is

$$1 + t/t_0 = \left(1 + E/E_0\right)^r. \tag{9}$$

This relationship leads to a frequency of events $1/(1 + t/t_0)$, a time probability

$$P(t)\mathrm{d}t = \frac{1}{\left(1 + t/t_0\right)^2} \frac{\mathrm{d}t}{t_0}, \quad 0 < t < \infty \tag{10}$$

and energy and magnitude probabilities

$$\begin{aligned} P(E)\mathrm{d}E &= \frac{r}{\left(1+E/E_0\right)^{1+r}} \frac{\mathrm{d}E}{E_0}, \quad 0 < E < \infty \\ P(M)\mathrm{d}M &= 2^r \frac{\beta e^{bM}}{\left(1+e^{bM}\right)^{1+r}} \mathrm{d}M, \quad 0 < M < \infty. \end{aligned} \tag{11}$$

Since in the law $E/E_0 = e^{bM}$ the energy is measured from E_0, we may omit the unity in Eq. (9), which becomes

$$t/t_0 \simeq \left(E/E_0\right)^r = e^{\beta M}; \tag{12}$$

this equation leads to the Gutenberg–Richter statistical distributions $P(M) = \beta e^{-\beta M}$ (Eqs. 1, 11).

Fitting the data

The Gutenberg–Richter magnitude distribution given by Eqs. (1), (2) and (5) is widely used to fit data. By such a fitting we derive the parameters β (and, therefore, r) and t_0 (seismicity rate). An important problem in such fitting procedures is the choice of the data. The data and the fitting parameters depend on the seismic region, the time period and the cutoff parameters.

In the region of small magnitudes, the data may exhibit a smaller slope, i.e. a smaller parameter β. This roll-off effect in the Gutenberg–Richter distribution is usually assigned to the insufficiency in the determination of the small-magnitude data. The problem of the small-magnitude cutoff (completeness of earthquake catalogs) enjoys many discussions, especially in connection with the aftershocks and foreshocks sequences (Lombardi 2002; Marzocchi and Sandri 2003; Console et al. 2003). It was shown recently that the roll-off effect may arise, at least partially, from earthquake dynamical correlations (Apostol 2020). We note that the Gutenberg–Richter distribution $\sim e^{-\beta M}$ corresponds to independent events, in the sense that the distribution for the sum $M = M_1 + M_2$ of two magnitudes is the product of the distributions of the individual magnitudes, $e^{-\beta M} = e^{-\beta(M_1+M_2)} = e^{-\beta M_1} e^{-\beta M_2}$. This property may indicate that the correlations are ineffective. This may be the case for moderate and large magnitudes, but for small magnitudes there is a departure from the Gutenberg–Richter law, and the correlations are effective. Indeed, the small-magnitude earthquakes are more affected by correlations.

On the other hand, for large magnitudes, the data distribution is uncertain. It is difficult to include high-magnitude earthquakes in a statistical analysis, because they are rare events and may not belong to a statistical ensemble. However, their minor weight in the statistical ensemble does not affect the results too much.

An analysis of a large set of global earthquakes with $5.8 < M < 7.3$ ($\Delta M = 0.1$) indicates $\beta = 1.38$ (and

$1/t_0 = 10^{5.5}$ per year), corresponding to $r = 0.4$, a value which suggests an intermediate two/three-dimensional focal mechanism (Bullen 1963). Equations (1), (2) and (5) have been fitted to a set of 1999 earthquakes with magnitude $M \geq 3$ ($\Delta M = 0.1$), which occurred in Vrancea between 1974 and 2004 (31 years) (Apostol 2006a, b). The mean values of the fitting parameters are $-\ln t_0 = 9.68$ and $\beta = 1.89$ ($r = 0.54$). The same fit has been done for a set of 3640 earthquakes with magnitude $M \geq 3$ which occurred in Vrancea during 1981–2018 (38 years). The fitting parameters for this set are $-\ln t_0 = 11.32$ and $\beta = 2.26$ ($r = 0.65$) (Apostol 2019b). The fitting values given above have an error of approximately 15%. The data for Vrancea have been taken from Romanian Earthquake Catalog ROMPLUS (2018, updated). The analyzed earthquakes occurred in Vrancea within 45°–46° latitude and 26°–27° longitude. The magnitude of completeness of the ROMPLUS Catalog is assumed to be $M = 2$, starting 1980. The magnitude errors in this catalog are $\Delta M = 0.1$. It is easy to see that the error in the parameter β is $\Delta\beta/\beta = 2\Delta M/M_{max}$, where M_{max} is the maximum magnitude included in analysis; for $\Delta M = 0.1$ and $M_{max} = 6$ we get $\Delta\beta/\beta = 0.03$.

In decimal logarithms the parameter β reads $\beta = br = \frac{3}{2}r$. Since $1/3 < r < 1$, this parameter varies in the range $1/2 < \beta < 3/2$. Usually, the average value $\beta = \frac{3}{2}r = 1$ ($\beta = 2.3$ in natural logarithms) is currently used as a reference value, corresponding to $r = 2/3$ (Lay and Wallace 1995; Udias 1999; Stein and Wysession 2003). We note that this value is close to $\beta = 2.26$ ($r = 0.65$) given above for Vrancea.

Entropy of earthquakes

Let us assume that a region is loaded with seismic energy. This energy is released in time, by a sequence of earthquakes with various magnitudes M. We may associate a statistical distribution $\rho(M)$ to this seismic activity. The probability density $\rho(M)$ should be normalized,

$$\int_0^\infty \mathrm{d}M \rho(M) = 1, \tag{13}$$

and the mean magnitude should be a constant,

$$\int_0^\infty \mathrm{d}M \cdot M\rho(M) = \overline{M}; \tag{14}$$

this condition corresponds to the original load of energy, which is well determined. We may view a continuous loading, and a continuous energy release, because the statistical distributions imply a large set of data, which may be viewed as being continuously distributed. Since we deal here with a large set of data, we may use a continuum model of magnitude distribution. Also, we note that we use the magnitude as statistical variable, because, it being a dimensionless number, it is suitable to label the states. We may assume that the continuous regime of releasing energy through earthquakes is well determined and characterized by the mean magnitude $\overline{M}$. If the region is "free" , or "isolated", i.e. it is not subject to external influences, and the earthquakes are independent events, it is reasonable to assume that the release of the seismic energy will be completed in a sufficiently long period of time, i.e. in the conditions given above, the amount of released energy is the same (and, practically, equal to the load) for all the realizations of the statistical ensemble. Consequently, we seek a functional of ρ which attains its extremum value for a certain function ρ; this ρ corresponds to the seismic activity in that region. Let us introduce the functional

$$S = -\int_0^\infty \mathrm{d}M \cdot \rho \ln \rho - \alpha\left[\int_0^\infty dM \cdot \rho(M) - 1\right] - \beta\left[\int_0^\infty dM \cdot M\rho(M) - \overline{M}\right], \tag{15}$$

and look for its extremum value under the conditions (13) and (14). The parameters α and β are determined from Eqs. (13) and (14). From the first-order variation

$$\int_0^\infty \mathrm{d}M(-\ln\rho - 1 - \alpha - \beta M)\delta\rho(M) = 0 \tag{16}$$

we get immediately

$$\rho(M) = \beta \mathrm{e}^{-\beta M} \tag{17}$$

and, by Eq. (14),

$$\overline{M} = \frac{1}{\beta} \tag{18}$$

(and $-1 - \alpha = \ln\beta$). We can see (Eq. 17) that we recover the Gutenberg–Richter magnitude distribution given by Eq. (1). The quantity S is called the entropy of the earthquakes. It characterizes the seismic activity of a seismic region in a given interval of time. This seismic activity proceeds in such a manner as to maximize the entropy; indeed, the second-order variation of Eq. (15) is

$$\int_0^\infty \mathrm{d}M \cdot (-1/\rho)[\delta\rho(M)]^2 < 0, \tag{19}$$

which shows that S is maximal with respect to ρ, for ρ given by Eq. (17). We may say, according to Statistical Physics, that this maximal S corresponds to the statistical seismic activity in that region and in that period of time. We note that the entropy defined in this way is a measure of the disorder produced by independent statistical events. In Statistical Physics the distribution given by Eq. (17) is called canonical distribution, or Gibbs distribution (Landau and Lifshitz 1980).

Making use of the distribution given by Eq. (17), we get from Eq. (15) the entropy

$$S = \beta \overline{M} - \ln \beta \tag{20}$$

(or $S = 1 - \ln \beta$ for $\beta \overline{M} = 1$). We can see that

$$\left(\frac{\partial S}{\partial \beta} \right)_{\overline{M}} = 0, \tag{21}$$

according to Eq. (18). Equation (20) provides a relation between entropy and the parameters of the earthquake distribution (see also De Santis et al. 2011). The procedure used here to derive the entropy of earthquakes is very useful in deriving also the fluctuation distribution, which is another relevant characteristics of a statistical ensemble.

In Eq. (20) we can view β and $\overline{M}$ as independent variables, such that $\left(\partial S / \partial \overline{M} \right)_\beta = \beta$ and $\mathrm{d}S = \beta \mathrm{d}\overline{M}$. This latter relation (which defines an "equilibrium" transformation, where $\beta \overline{M} = 1$) indicates the physical meaning of the entropy: its changes are proportional to the changes in the mean magnitude. We note that in equilibrium $\beta, S, \overline{M} = \text{const}$. In practice, it is convenient to employ finite-difference variations for an equilibrium transformation, written as $\Delta S = \beta \Delta \overline{M} = -\frac{1}{\beta} \Delta \beta$; if $\Delta \beta = 0$, then $\Delta S = 0$ and we have, over that variation interval, an equilibrium (reversible) process. If the variation in time of the parameter β is very slow, the transformation is called adiabatic (and the process is a quasi-equilibrium process). We use the finite-difference variation of the entropy in the end section, in the discussion concerning Fig. 2.

Shannon entropy

The entropy given by Eq. (15), which implies a probability density, is called the Boltzmann–Gibbs entropy (Landau and Lifshitz 1980). From Eqs. (15) and (18) we get

$$S = -\int_0^\infty \mathrm{d}M \cdot \rho \ln \rho = -\int_0^\infty \mathrm{d}M \cdot \beta \mathrm{e}^{-\beta M} (\ln \beta - \beta M) = 1 - \ln \beta \tag{22}$$

(for $\beta \overline{M} = 1$), where $\rho = \beta \mathrm{e}^{-\beta M}$. If we view the entropy S as a measure of the disorder degree, then it should be positive, a condition which entails the inequality $\beta < e = 2.72$ (we note that this condition is satisfied for Vrancea, where $\beta = 2.26$). Such a condition is well known in Statistical Physics (it is called the classical limit of the ideal gas). If the density of small-magnitude earthquakes is excessively large, this inequality may be violated; in this case, the events corresponding to $M \longrightarrow 0$ should correspond to what is called "ordered" events in Statistical Physics.

The above description offers the opportunity to extend the analogy between the seismic activity and the behaviour of a thermodynamic ensemble. From equation $\mathrm{d}S = \beta \mathrm{d}\overline{M}$, derived above, we can view the parameter β as the inverse of a temperature $T = 1/\beta$ and the mean magnitude $\overline{M}$ as the thermodynamic energy. The equation becomes $\mathrm{d}\overline{M} = T\mathrm{d}S$, where $T\mathrm{d}S$ is the "heat" of the seismic activity; this equation leads to a Helmholtz free energy $F = \overline{M} - TS = -T \ln T = \frac{1}{\beta} \ln \beta$. This thermodynamic behaviour is preserved down to the low temperature $T_c = 1/e$ ($\beta_c = e$). For lower temperatures, the thermodynamic state must be reconstructed. We may assume a zero entropy $S = S_0 = 0$ for $T < T_c$ ($\beta > \beta_c = e$) and a non-vanishing entropy $S = 1 - \ln \beta = 1 + \ln T$ for $T > T_c$; then, the derivative $\partial S / \partial T$ has a discontinuity at $T = T_c$, similar with a phase transition.

In empirical studies the Shannon entropy is often used, which implies a discrete summation over point-process probabilities p_n. The Shannon entropy is defined as

$$\Sigma = -\sum_{n=0}^{\infty} p_n \ln p_n; \tag{23}$$

for a discretization with the step δ the probabilities are given by

$$p_n = \left(1 - \mathrm{e}^{-\beta\delta}\right) \mathrm{e}^{-n\beta\delta}, \tag{24}$$

where $M_n = n\delta$. The Shannon entropy is always positive (or zero). Since we deal here with a large set of data, it is more convenient to use a continuum model of magnitude distribution. Therefore, it is important to establish the relation between the Boltzmann–Gibbs entropy and the Shannon entropy in the limit $\delta \longrightarrow 0$. Also, we note that we do not compare directly the entropy to the empirical data; instead, we derive the parameter β of the Gutenberg–Richter distribution from fitting the data, and discuss its implications upon the entropy. This is the most direct approach of extracting information from data, because it gives access to the central concept of probability density.

The direct calculation of the summation in Eq. (23) leads to

$$\begin{aligned} \Sigma &= -\ln\left(1 - \mathrm{e}^{-\beta\delta}\right) + \frac{\beta\delta}{\mathrm{e}^{\beta\delta} - 1} \\ &= 1 - \ln(\beta\delta) + \frac{3}{8}(\beta\delta)^2 + \cdots, \end{aligned} \tag{25}$$

the series expansion being valid in the limit $\beta\delta \longrightarrow 0$. We can see that

$$\Sigma = S - \ln \delta + \cdots, \tag{26}$$

where S is given by Eq. (22) (De Santis et al. 2019). In the limit $\delta \longrightarrow 0$ the Shannon entropy increases indefinitely, which makes it inconvenient for a continuum model (leaving aside the ambiguity in defining the step δ). On the other hand, since we are interested only in variations of the entropy, we may leave aside the constant (and indefinite) contribution $-\ln\delta$. In this sense, the Shannon entropy is equivalent with the Boltzmann–Gibbs entropy, the latter being sometimes called the differential form of the former. For $S \ll 0$ ($\beta \gg \beta_c = e$) the probability density is $\rho(M=0)=\beta$ and $\rho(M>0)\simeq 0$. This indicates an accumulation of earthquakes on the state defined by $M=0$. In this case, according to Eq. (26) the parameter δ should go to zero. We may assume that all $p_n = 0$ for $n>0$ and $p_0 = \beta\delta = 1$. Then, the entropy is $\Sigma_0 = -\beta\delta\ln(\beta\delta) = 0 = S_0$. However, in the region $M \longrightarrow 0$ the correlations are effective, the events are not independent anymore, and the above formulae of the entropy are not valid.

The connection given by Eq. (26), between a discrete summation and its integral, is a general phenomenon known as the Euler–MacLaurin corrections (Abramowitz and Stegun 1964). Indeed, the transition from a summation to the integral of a function $f(x)$ is done according to the formula

$$\int_0^\infty \mathrm{d}x f(x) = \delta\sum_{n=0}^{\infty} f(\delta n) - \frac{1}{2}\delta f0) + \frac{1}{12}\delta^2 f'(0) + \cdots. \quad (27)$$

If we apply this formula to the function $f(M) = -\rho\ln\rho = -\beta e^{-\beta M}(\ln\beta - \beta M)$, we get

$$S = \Sigma(\beta\delta \longrightarrow 0) + \ln\delta + \frac{1}{2}\beta\delta\ln(\beta\delta) + \cdots, \quad (28)$$

which, in the limit $\beta\delta \longrightarrow 0$, is the relation given in Eq. (26). A similar relationship exists between the continuum and discrete normalization conditions,

$$\int_0^\infty \mathrm{d}M \cdot \beta e^{-\beta M} = \beta\delta\sum_{n=0}^{\infty} e^{-\beta\delta n} - \frac{1}{2}\beta\delta - \frac{1}{12}(\beta\delta)^2 + \cdots = 1. \quad (29)$$

Fluctuations

The definition (15) may lead to another viewpoint as regards the entropy. We may view the seismic activity as consisting of a sequence $M_i, i = 1, 2, \ldots, N$, of magnitudes, each realized by N_i random, independent processes (in Statistical Physics this is Boltzmann's hypothesis of the so-called molecular chaos); the probability ρ_i of each of these processes is $\rho_i = 1/N_i$. We may define the entropy $s_i = -\ln\rho_i = \ln N_i$ and the average entropy $\overline{s}_i = -N_i\rho_i\ln\rho_i = \ln N_i = s_i$. It follows that the probability ρ_i is given by $\rho_i = e^{-s_i} = e^{-\overline{s}_i}$. In Statistical Physics this distribution is called the microcanonical distribution, and $s_i = \overline{s}_i$ is called microcanonical entropy (Gibbs 1902). Obviously, this entropy is maximal under the condition of normalized probabilities ρ_i: ρ_i is a constant for each M_i. We may use $\rho_i = \beta e^{-\beta M_i}$, and get the microcanonical entropy

$$s_i = \overline{s}_i = \beta M_i - \ln\beta. \quad (30)$$

Since the processes are independent and random, irrespective of their magnitude, their probability ρ_m is given by

$$\rho_m^N = \prod_i \rho_i = e^{-\sum_i s_i}, \quad (31)$$

whence the entropy

$$S = -\ln\rho_m = \frac{1}{N}\sum_i s_i = \beta\overline{M} - \ln\beta, \quad (32)$$

where N is the number of the i-processes. We can see that S, given by Eq. (32) is the canonical entropy given by Eq. (20). The microcanonical probability

$$\rho_m = e^{-S} = \beta e^{-\beta\overline{M}}, \quad (33)$$

is the value of the canonical probability for the mean magnitude.

In equilibrium the entropy is maximal ($(\partial S/\partial\beta)_{\overline{M}} = 0$, Eq. 21); therefore, we may have small variations $\Delta S = \frac{1}{2}(\partial^2 S/\partial\beta^2)_{\overline{M}}(\Delta\beta)^2 = (\Delta\beta)^2/2\beta^2$. This variation is positive. It indicates the tendency of reaching equilibrium. Introducing this variation in entropy in Eq. (33), the probability ρ_m gives the fluctuation probability

$$\rho_f = \frac{1}{\sqrt{2\pi}\beta} e^{-\frac{(\Delta\beta)^2}{2\beta^2}} \quad (34)$$

(properly normalized). We can see that the statistical ensemble may have fluctuations, whose measure is the standard deviation

$$\delta\beta = \left[\overline{(\Delta\beta)^2}\right]^{1/2} = \beta \quad (35)$$

given by the normal distribution in Eq. (34). This is Einstein's fluctuation formula (Einstein 1909). If, in a distribution $\overline{M}e^{-\beta\overline{M}}$, we may view the parameter β as a statistical variable, then we get the standard deviation $\delta\beta = 1/\overline{M}$, which coincides with Eq. (35) for $\beta\overline{M} = 1$. A similar analysis can be made for the canonical distribution $\beta e^{-\beta M}$, leading to fluctuations $\delta M = 1/\beta = \overline{M}$. Various realizations of the statistical ensemble (in the same conditions) exhibit fluctuations, i.e. for another sampling, i.e. for another region, or another period of time, with the same mean magnitude, i.e. for another realization of the statistical ensemble, the entropy may fluctuate. We note that the Gutenberg–Richter distribution is a distribution in magnitude (Eq. 17), while the normal law given by Eq. (34) is a distribution in the parameter β of the Gutenberg–Richter distribution. Also, we note that the fluctuations are variations of the mean values. Standard deviations as large as the mean value indicate a serious limitation of the information we may get from statistical analysis of earthquake distributions. In addition, the "same conditions" requirement of the fluctuation formula (known as the "null hypothesis") may not be fulfilled; for instance, over a similar period of time, the geological conditions of the seismic region may change, or the accuracy of the measured magnitudes may differ.

We note that the parameter β may depend on the time in these formulae. Then, the variation of the entropy is given by $\Delta S = -\frac{1}{\beta}\Delta\beta$, over corresponding intervals of time variations. The fluctuations distribution remains valid in this case (since $\beta\overline{M} = 1$), and an analysis of this distribution may give the parameter β and, consequently, its variation. If the process would be a (quasi-) equilibrium process, we may expect an opposite variation of β in the next period of time, such that $\Delta\beta = 0$ and $\Delta S = 0$.

Analysis of Vrancea seismic activity

Let us assume that we fitted the Gutenberg–Richter distribution (e.g. the cumulative distribution, Eq. 2) to data gathered over a long period of time t_0 for a given region. At the moment of time t_0 we have the fitting parameter β_0. For a sufficiently long period of time t_0 we may assume that this seismic activity is statistically well defined. Let us take (at random) the next moments of time t_i, $i = 0, 1, 2, \dots, N$ and update the fitting to get the parameters β_i. If the time intervals $t_{i+1} - t_i$ are sufficiently small (but still as large as to have a measurable seismic activity in each) we may expect that the variations $\Delta\beta_i = \beta_{i+1} - \beta_i$ are fluctuations. For a sufficiently large N we may fit their distribution with the normal law given by Eq. (34). Thus, we get the fitting parameter β. If $\beta = \beta_0$ (within the fitting errors), the seismic equilibrium has not changed. If $\beta \neq \beta_0$ the equilibrium has changed over the period $t_N - t_0$. Consequently (in the absence of an external agent), we may expect a tendency to recover the equilibrium over the next period of time $t_{2N+1} - t_{N+1}$. Such an information might be regarded as a short-time prediction. For instance, if $\beta < \beta_0$, we may expect in the next time interval a decrease in the mean magnitude, i.e. the number of earthquakes with low magnitudes will increase, and high-magnitude earthquakes are not likely. On the contrary, if $\beta > \beta_0$, then we may expect an increase in the number of earthquakes with higher magnitude. We note that an increase (decrease) in β amounts to a decrease (increase) in equilibrium entropy $S = 1 - \ln\beta$ (Eq. 20).

It may happen that the distribution of the parameter changes $\Delta\beta_i$ is not a normal distribution. Then, the ensemble is not in equilibrium in the time period $t_N - t_0$. Under the equilibrium hypothesis, we may expect an evolution towards equilibrium in the next period. For instance, if the distribution of the variations $\Delta\beta_i$ is shifted towards higher values, i.e. if the parameters β_i show a tendency to increase, we may expect a decrease in these parameters in the next period, i.e. an increase in the mean magnitude.

It is worth noting that the discussion given above is valid under the main assumption of independent seismic events. If correlations exist, the entropy formulae derived above do not apply. A special case in this connection is the short-term foreshock (and aftershock) activity. The accompanying seismic activity obeys, approximately, the Gutenberg–Richter magnitude distribution (Kisslinger 1996), and a decrease in the parameter β, observed for the foreshock activity, was interpreted as an increase in entropy (De Santis et al. 2011). Moreover, recently it was shown that a real-time discrimination between foreshocks and aftershocks might be attained by monitoring the variations in the parameter β (Gulia and Wiemer 2019).

If the correlations are included, we expect a change in the distribution. In this case, the formulae given above for the entropy are not valid anymore. It was shown that the change caused by correlations in the Gutenberg–Richter distribution affects mainly the small-magnitude region (Apostol 2020). For moderate and large earthquakes the distribution preserves its independent-event form $\sim e^{-\beta M}$, which ensures the validity of the entropy formulae used here. Small-magnitude earthquakes ($M < 2$) are excluded from our analysis.

Also, we note that the practical application of the procedure described above depends on the choice of the (long) time period t_0, the (short) time intervals $t_{i+1} - t_i$ and the (large) number N of these intervals. This choice can only be made in close connection with the particular character of the seismic activity in the given region and in the given (long) time period.

The description given above is not supported by data, at least for Vrancea region, in the analyzed time periods. The parameter β is quasi-uniformly increasing in time, at a slow rate, due to the accumulation of small-magnitude earthquakes. This quasi-uniform tendency is interrupted from time to time by large-magnitude earthquakes, which decrease suddenly the parameter β (there is no evidence for foreshocks and aftershocks). Two examples of short-time variations of beta are given in Fig. 1 for Vrancea seismic activity. The time t_0 is from January 1*st* 1980 to December 31st 2009, with 5391 earthquakes with magnitude greater than 2. The Gutenberg–Richter fit to these data gives $\beta_0 = 2.125$ (error 15%). We have updated the parameter β for each week of the year 2010 (Fig. 1, panel a). We can see that this parameter increases continuously over this whole year (due to the accumulation of small-magnitude earthquakes). A similar procedure was used for each of the next years up

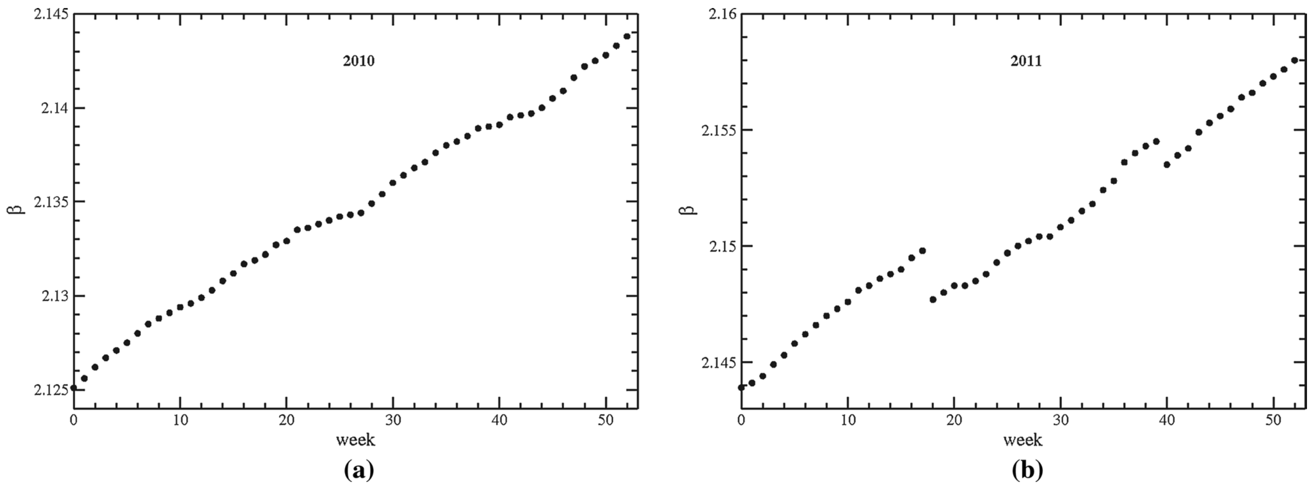

Fig. 1 Short-time (weeks) continuous (panel a, year 2010) and discontinuous (ruptures, panel b, year 2011) variations of the parameter β for Vrancea (Romanian Earthquake Catalog ROMPLUS (2018, updated)

to 2019 (8455 earthquakes with $M \geq 2$ in the whole period 1980–2019, taken from Romanian Earthquake Catalog ROMPLUS (2018, updated). In some years the continuous increase in the parameter β is disrupted by the occurrence of greater-magnitude earthquakes, like in the year 2011 (Fig. 1, panel b). Such variations of the parameter β cannot be fitted by a normal distributions, and, therefore, they cannot be viewed as fluctuations. We can only say, very imprecisely and qualitatively, that after a period of small-magnitude seismic activity it is likely to follow a few earthquakes with greater magnitude, and vice versa, which is a useless, common-sense expectation. Rigorously speaking, the seismic activity is not in (quasi-) equilibrium, because the tectonic energy source feeds it continuously. We expect this behaviour to have a general character. We note that such abrupt variations in the parameter β (and an abrupt increase in the entropy) have been reported by De Santis et al. (2011) for the accompanying seismic activity of the L'Aquilla earthquake (magnitude 6.3, 6 April 2009) and the Colfiorito earthquake (magnitude 6, 26 September 1997).

The overall variation of the parameter β is a slow increase in time, which may suggest a quasi-equilibrium adiabatic process. In the neighbourhood of a greater earthquake the parameter β suffers an abrupt variation. A qualitative sketch of the typical variations of the parameter β for a time interval (t_1, t_2) which includes a main shock (MS) at the moment t_{MS} is given in Fig. 2. After a (slight) increase the parameter β suffers an abrupt decrease $\Delta\beta_1$, possibly in a foreshock (f) region, down to the main shock, followed by an abrupt increase $\Delta\beta_2$ which may include, possibly, an aftershock region (a). We can see that the total variation $\Delta\beta = \Delta\beta_1 + \Delta\beta_2 = 0$, such that we may say that over this region there exists an equilibrium process ($\Delta S = 0$). Such a qualitative behaviour is shown in Fig. 1 for Vrancea and, also, is reported by De Santis et al. (2011) and Gulia and Wiemer (2019). The latter reference suggests to use the precursory decrease in the foreshock region, distinct from an increase in the aftershock region, as a real-time prediction of a main shock.

There exists another interpretation of the sudden variation of the parameter β in the vicinity of a large-magnitude earthquake, where the continuous variable is the time t (Fig. 1, panel b). Indeed, the sudden jump in the parameter β indicates a time derivative $\frac{\partial\beta}{\partial t} = \Delta\beta_1\delta(t - t_{MS})$, with the notations in Fig. 2. This abrupt variation induces a similar variation in the time derivative of the entropy $\frac{\partial S}{\partial t} = -\frac{\Delta\beta_1}{\beta_1}\delta(t - t_{MS})$. The singularity indicated by the function $\delta(t - t_{MS})$ is associated with a phase-transition singularity, which would correspond to a critical regime (see,

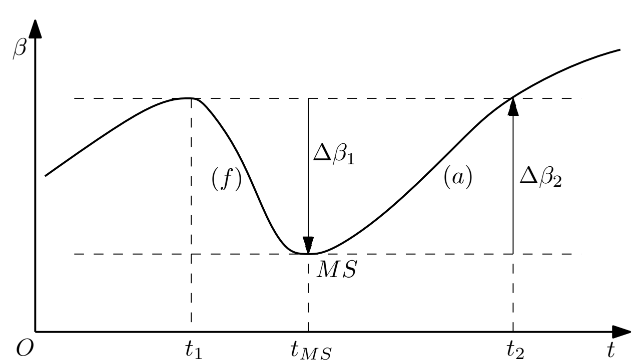

Fig. 2 A qualitative sketch of the variation of the parameter β versus time t around the moment t_{MS} of the occurrence of a main shock (MS), possibly including a foreshock region (f) and an aftershock region (a). The total variation of the parameter β in the time intervals (t_1, t_{MS}) and (t_{MS}, t_2) is zero ($\Delta\beta_1 + \Delta\beta_2 = 0$), indicating an equilibrium process (constant entropy, $\Delta S = 0$)

e.g. De Santis et al. 2019). However, this is not a thermodynamic interpretation of the singularity, because the continuous parameter is the time, not the temperature $T = 1/\beta$, which has also a jump (it is discontinuous) at t_{MS}.

The absence of fluctuations in the seismic activity analyzed here raises an interesting question. In Statistical Physics the fluctuations are analyzed for an ensemble consisting of a large number $N \gg 1$ of sub-ensembles (subsystems). The standard deviation $\delta\beta = \beta$ computed in Eq. (35) corresponds to one sub-ensemble. The sub-ensemble method is convenient for an extensive ensemble (like a gas of N particles). The seismic activity lacks this extensive property. However, we may view the successive updates of the parameter β described above as a series of N distinct, random, independent realizations of our ensemble, such that the average value of the parameter β is given by

$$B = \frac{1}{N}\sum_{i=1}^{N}\beta_i, \tag{36}$$

with the mean value $\overline{B} = \overline{\beta}$. We note that these assumptions imply the statistical equilibrium. The mean square deviation of B is

$$\overline{\Delta B^2} = \frac{1}{N^2}\sum_{i,j=1}^{N}\overline{\Delta\beta_i\Delta\beta_j} = \frac{1}{N^2}\sum_{i=1}^{N}\overline{\Delta\beta_i^2}, \tag{37}$$

which can be written as $\overline{\Delta B^2} = \frac{1}{N}\overline{\Delta\beta^2}$; hence, we find $\delta B/\overline{B} = \sqrt{\overline{\Delta B^2}}/\overline{B} = \frac{1}{\sqrt{N}}\left(\delta\beta/\overline{\beta}\right)$, where we may take $\delta\beta = \beta$ and $\overline{\beta} = \beta$. We can see that the relative fluctuation of B is vanishing for large N and the dispersion of the variables β_i is $\delta\beta = \sqrt{\overline{\Delta\beta^2}} = \beta$. This corresponds to the normal (Gaussian) distribution given by Eq. (34) (this result is also known as the central limit theorem). If the normal distribution changes in time, under the assumption of (quasi-) statistical equilibrium, we would have the possibility to do a prediction, as discussed above. From the discussion given above we see that this picture is not supported by data. The values β_i of the parameter β exhibit a slight, uniform increase in time, interrupted by disparate, apparently random abrupt decreases. This behaviour suggests a non-equilibrium process.

The time variation of the parameter β of the Gutenberg–Richter distribution may be used as a quantitative measure of the departure from equilibrium of a given seismic activity. For example, from Fig. 1 we can estimate a variation $\simeq 1\%$ for β during the year 2010 and $\simeq 0.5\%$ for year 2011. The latter is smaller, due to the occurrence of two large earthquakes in 2011, which lowered the parameter β (1 May 2011, magnitude 4.9 and 4 October 2011, magnitude 4.8). Similar values are obtained for other years, which may indicate that the seismic activity in Vrancea has a rather steady character, with a constant rate of change in time of the magnitude distribution, over a long period. Although these figures are very small, and we might be tempted to assign a quasi-equilibrium character to the seismic activity, such a conclusion is not supported by the lack of fluctuations; the existence of the fluctuations is a necessary element for the statistical (quasi-) equilibrium. However, if we view the small-earthquake increase in the parameter β together with the decrease brought about by larger earthquakes as long-period quasi-oscillations over a long period of time (including a large set of data), then we may assume that these quasi-oscillations are fluctuations. Such quasi-oscillations have been identified previously on the data corresponding to 3640 earthquakes with magnitude $M \geq 3$ which occurred in Vrancea during 1981−2018 (Apostol 2019b). Also, it seems that such quasi-oscillations have been seen by Main and Al-Kindy (2002), although on small-sized samples. From this perspective, long-period quasi-oscillations in the parameter β, corroborated with a small, overall increase, might lead to assuming that the seismic activity may be approximated by a quasi-equilibrium process over such very long periods of time. However, we note that the steady increase in the parameter β and its disparate ruptures are not easily distributed on a normal Gaussian.

Concluding remarks

The Gutenberg–Richter statistical distribution in magnitude is derived for a canonical ensemble by the standard procedure of maximizing the entropy. By assuming the seismic activity as consisting of a sequence of random, independent earthquakes with various magnitudes, the corresponding microcanonical ensemble is examined by standard methods and the fluctuations normal distribution is derived for earthquakes (Einstein's fluctuation formula). These results are tested by an analysis of 8455 earthquakes with magnitude $M \geq 2$, which occurred in Vrancea during the time interval (years) 1980−2019. The analysis indicates a departure from equilibrium, with a slight increase in the parameter β (0.5%−1% per year) and the absence of the fluctuations (at least in the usual Gaussian form of their distribution). The seismic activity is fed continuously by the tectonic energy and, consequently, it cannot be viewed, rigorously, as a (quasi-) equilibrium statistical ensemble. The Gutenberg–Richter distribution is shifted continuously towards small-magnitude earthquakes, with random re-arrangements caused by higher-magnitude earthquakes.

Acknowledgements The authors are indebted to the colleagues in the Institute of Earth's Physics, Magurele-Bucharest, for many enlightening discussions, and to the anonymous Reviewers for thoughtful suggestions. This work was partially carried out within the Program Nucleu 2019, funded by Romanian Ministry of Research and Innovation, Research Grant #PN19-08-01-02/2019.

References

Abramowitz M, Stegun IA (1964) Handbook of mathematical functions, with formulas, graphs and mathematical tables. National Bureau of Standards, Applied Mathematics Series #55, Washington, DC, p 806

Apostol BF (2020) Bath's law, correlations and magnitude distributions. arXiv:2006.07591v1 [physics.geo-ph], 13 June

Apostol BF (2006a) A model of seismic focus and related statistical distributions of earthquakes. Rom Reps Phys 58:583–600

Apostol BF (2006b) Model of seismic focus and related statistical distributions of earthquakes. Phys Lett A 357:462–466

Apostol BF (2019a) An inverse problem in seismology: derivation of the seismic source parameters from P an S seismic waves. J Seismol 23:1017–1030

Apostol BF (2019b) Statistical seismology, internal report. Institute of Earth's Physics, Magurele

Berrill JB, Davis RO (1980) Maximum entropy and the magnitude distribution. Bull Seismol Soc Am 70:1823–1831

Bullen KE (1963) An introduction to the theory of seismology. Cambridge University Press, London

Console R, Lombardi AM, Murru M, Rhoades D (2003) Bath's law and the self-similarity of earthquakes. J Geophys Res 108:2128. https://doi.org/10.1029/2001JB001651

De Santis A, Cianchini G, Favali P, Beranzoli L, Boschi E (2011) The Gutenberg–Richter law and entropy of earthquakes: two case studies in Central Italy. Bull Sesimol Soc Am 101:1386–1395

De Santis A, Abbattista C, Alfonsi L, Amoruso L, Campuzano SA, Carbone M, Cesaroni C, Cianchini G, De Franceschi G, De Santis A, Di Giovambattista R, Marchetti D, Martino L, Perrone L, Piscini A, Rainone ML, Soldani M, Spogli L, Santoro F (2019) Geosystemics view of earthquakes. Entropy 21:412. https://doi.org/10.3390/e21040412

Dong WM, Bao AB, Shah HC (1984) Use of maximum entropy principle in earthquake recurrence relationships. Bull Seismol Soc Am 74:725–737

Einstein A (1909) Zum gegenwaertigen Stand des Strahlungsproblem. Phys Z 10:185–193

Gibbs JW (1902) Elementary principles in statistical mechanics. Scribner's sons, New York

Gulia L, Wiemer S (2019) Real-time discrimination of earthquake foreshocks and aftershocks. Nature 574:193–199

Gutenberg B, Richter C (1944) Frequency of earthquakes in California. Bull Seismol Soc Am 34:185–188

Gutenberg B, Richter C (1956) Magnitude and energy of earthquakes. Ann Geofis 9: 1–15 ((2010) Ann Geophys 53:7–12)

Hanks TC, Kanamori H (1979) A moment magnitude scale. J Geophys Res 84:2348–2350

Kanamori H (1977) The energy release in earthquakes. J Geophys Res 82:2981–2987

Kisslinger C (1996) Aftershocks and fault-zone properties. Adv Geophys 38:1–36

Landau L, Lifshitz E (1980) Statistical physics, course of theoretical physics, vol 5. Elsevier, Oxford

Lay T, Wallace TC (1995) Modern global seismology. Academic Press, San Diego

Lombardi AM (2002) Probability interpretation of "Bath's law". Ann Geophys 45:455–472

Main I, Al-Kindy F (2002) Entropy, energy and proximity to criticality in global earthquake populations. Geophys Res Lett. https://doi.org/10.1029/2001GL014078

Main I, Burton PW (1984) Information theory and the earthquake frequency-magnitude distribution. Bull Seismol Soc Am 74:1409–1426

Marzocchi W, Sandri L (2003) A review and new insights on the estimation of the *b*-value and its uncertainty. Ann Geophys 46:1271–1282

Masinha L, Shen PY (1987) On the magnitude entropy of earthquakes. Tectonophysics 138:115–119

Nicholson T, Sambridge M, Gudmundsson O (2000) On entropy and clustering in earthquake hypocentre distributions. Geophys J Int 142:37–51

Ranalli G (1969) A statistical study of aftershock sequences. Ann Geofis 22:359–397

Richter CF (1958) Elementary seismology. Freeman, San Francisco

Romanian Earthquake Catalogue (ROMPLUS Catalog), National Institute for Earth Physics, Romania (2018) (updated)

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27(379–423):623–666

Shen PY, Mansinha L (1983) On the principle of maximum entropy and the earthquake frequency-magnitude relation. Geophys J R Astr Soc 74:777–785

Stein S, Wysession M (2003) An introduction to seismology, earthquakes, and earth structure. Blackwell, New York

Udias A (1999) Principles of seismology. Cambridge University Press, New York

Utsu T (1969) Aftershocks and earthquake statistics (I, II): Source parameters which characterize an aftershock sequence and their interrelations. J Fac Sci Hokkaido Univ Ser VII 3:129–195

Utsu T, Seki A (1955) A relation between the area of aftershock region and the energy of the mainshock. J Seismol Soc Jpn 7:233 (in Japanese)

Wiener N (1948) Cybernetics. MIT Press, Cambridge