Seismic Magnitudes Entropy and b-value

Fidencio Alejandro Nava-Pichardo[1]

[1]CICESE

April 12, 2024

Abstract

How seismic magnitudes are distributed is important for estimating stress levels in seismic hazard studies, and two methods of characterizing the magnitude distribution are through the Gutenberg-Richter b-value, or equivalently through , and through the information entropy. A closed relationship between the b-value and the entropy (applicable to any exponential distribution and its entropy) is presented and is checked through numerical evaluation of the entropy using exact probabilities derived directly from the magnitude exponential distribution. Since the numerical evaluation of the entropy is done over a finite magnitude range, it is possible to assess the possible contribution to the entropy of real or hypothetical very large magnitudes, and these contributions are found to be quite small. The relationship is also compared with entropies calculated from synthetic data, and Monte Carlo simulations are used to explore the behavior of entropy determinations as a function of sample size. Finally, it is considered how, for the usual case of having data from a single realization, in spite of the relation between them, because entropy and Aki-Utsu b-value are measured in different ways, both measures are not redundant and may be complementary and useful in determining when a sample is large enough to give reliable results.

Hosted file

`AGU_Manuscript_Word_Mar_23.docx` available at https://authorea.com/users/766355/articles/813863-seismic-magnitudes-entropy-and-b-value

Seismic Magnitudes Entropy and *b*-value

F. A. Nava[1]

[1]Centro de Investigación Científica Y Educación Superior de Ensenada, Baja California (CICESE).

Corresponding author: F. Alejandro Nava (fnava@cicese.mx)

Key Points: Information entropy, Gutenberg-Richter *b*-value, seismic magnitudes

Abstract

How seismic magnitudes are distributed is important for estimating stress levels in seismic hazard studies, and two methods of characterizing the magnitude distribution are through the Gutenberg-Richter *b*-value, or equivalently through $\beta = b \ln 10$, and through the information entropy. A closed relationship between the *b*-value and the entropy (applicable to any exponential distribution and its entropy) is presented and is checked through numerical evaluation of the entropy using exact probabilities derived directly from the magnitude exponential distribution. Since the numerical evaluation of the entropy is done over a finite magnitude range, it is possible to assess the possible contribution to the entropy of real or hypothetical very large magnitudes, and these contributions are found to be quite small. The relationship is also compared with entropies calculated from synthetic data, and Monte Carlo simulations are used to explore the behavior of entropy determinations as a function of sample size. Finally, it is considered how, for the usual case of having data from a single realization, in spite of the relation between them, because entropy and Aki-Utsu *b*-value are measured in different ways, both measures are not redundant and may be complementary and useful in determining when a sample is large enough to give reliable results.

Plain Language Summary

Two important measures for seismic hazard, that describe the relative abundance of small to medium and large earthquakes, are the slope of the logarithmic cumulative magnitude histogram, known as *b*-value, and the information entropy, *S*, of the probabilities of the magnitudes. We find an analytic relationship between both measures and check it numerically. The effects of a studying a finite magnitude range are explored. It is proposed that since *b*-value and *S* are measured in different ways, their estimates can be complementary for interpreting results from a single set of data.

1 Introduction

The Shannon, or information, entropy and the *b*-value of the Gutenberg-Richter distribution, both discussed in detail below, have become useful and widely-used tools in the study of seismicity, because both seem to quantify behaviors of seismicity related to the levels of stress. Here, a relationship between *b* and the entropy of the seismic magnitudes will be presented, some of its features will be discussed, and ways in which these measures can complement each other will be proposed.

In what follows unrounded magnitudes will be denoted by m and magnitudes rounded to ΔM by M (usually $\Delta M = 0.1$).

1.1 The G-R *b*-value

Ishimoto and Ida (1939) and Gutenberg and Richter (1944) observed that seismic magnitudes are distributed as

$$\log_{10} N(M) = a - b\,(M - M_c); \quad M \geq M_c \tag{1}$$

where $N(M)$ is the number of magnitudes $\geq M$ and b describes the proportion of large magnitudes to small ones (Richter, 1958). The magnitude origin has been shifted by M_c, the completeness magnitude below which $\log_{10} N(M)$ ceases to behave linearly due to insufficient coverage (e.g. Wiemer and Wyss, 2000). Although the physical meaning of this relation, and of related distributions of seismic energy and moment are still subject to discussion (e.g. El-Isa and Eaton, 2014; Wyss, 1973) the b-value has been widely used to characterize seismicity in different regions in the world (e. g. Kagan, 1999; Utsu, 2002), and it has been proposed that b is related to the fractal dimension (Aki, 1981; Hirata, 1989; Wyss et al., 2004). There are many studies that relate b to the level of stress and observe changes in its value before large earthquakes (DeSalvio & Rudolph, 2021; El-Isa & Eaton, 2014; Enescu & Ito, 2001; Frohlich & Davis, 1993; Godano et al., 2024; Hu et al, 2024; Li & Chen, 2021; Nanjo et al., 2012; Scholtz, 2015; Schorlemmer et al, 2005; Utsu, 2002; Wang, 2016; Wyss, 1973; Wyss et al., 2004; and many others), which gives b an important role in earthquake hazard estimation and forecasting.

b-values can be estimated directly from the slope of the linear range on the G-R histogram (e.g. Guttorp, 1987; Monterroso & Kulhanek, 2003), but frequently b-values are estimated from the mean magnitude (Aki, 1965; Marzocchi & Sandri, 2003; Tinti & Mulargia, 1987; Utsu, 1965), and most studies use the Aki-Utsu maximum likelihood estimate

$$b = \frac{\log_{10}(\mathrm{e})}{\overline{M} - m_c}, \tag{2}$$

where $\overline{M}$ is the observed mean of the data (Aki, 1965; Utsu, 1965). This estimate shares with the entropy determinations the problem of determining m_c, but otherwise it is based on the mean magnitude that, in a way, incorporates the information from all magnitudes. This measure is extremely easy to obtain but, unfortunately, many people use (2) as a magic formula without considering that the estimate will be good only if the observed $\bar{M}$ is close to μ (compare (2) with (7)), which requires having a sample large enough to be representative (Geffers et al., 2022; Marzocchi et al. 2020; Nava et al., 2017; Ogata and Yamashina, 1986; Shi and Bolt, 1982).

1.2 The Information Entropy

Another important statistical-probabilistical concept is Shannon's definition of the information entropy, S, of a process characterized by K states or classes of events, each having probability P_i, with

$$\sum_{i=1}^{K} P_i = 1\,, \tag{3}$$

as

$$S = -\sum_{i=1}^{K} P_i \log_2 P_i \equiv \sum_{i=1}^{K} s_i \tag{4}$$

(Shannon, 1948), where the logarithm can have any base; we will use base 2 because it is the one most commonly used for information purposes and yields an entropy expressed in bits, easy to interpret. Capital letters have been used for the probabilities to emphasize that they are not densities, and in this definition it is implicitly assumed that $0 \leq P_i \leq 1$, so that $\log_2 P_i \leq 0 \ \forall i$. Each term in the first sum in (3) is the contribution to the total entropy S of the probability of each rounded magnitude class, called *entropy score* by Harte and Vere-Jones (2005), and will be denoted by s_i, where i is the index of the class, or generally as s.

Remembering that the self-information of an event with probability P_i is

$$I_i = -\log_2 P_i \tag{5}$$

(Fano, 1961), the entropy (4) can be recognized as the expected self-information of the process. Although the self-information ranges from zero to infinity, the contribution to the entropy from any probability ranges from zero, for both $s(0)$ and $s(1)$, to the maximum $s(\mathrm{e}^{-1}) = 0.530738$ bit, as shown in Figure 1. This point will be retaken later.

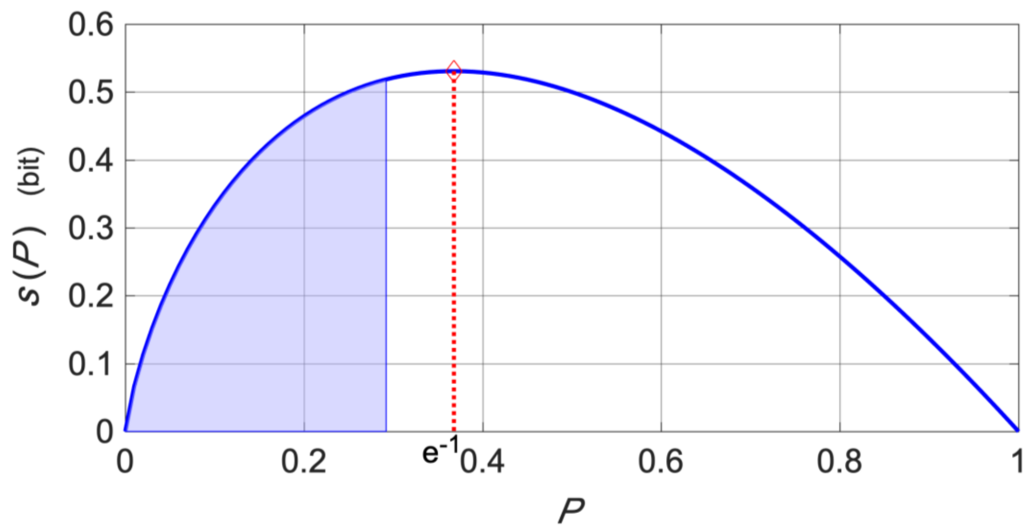


Fig. 1 Contribution of each particular probability value to the entropy. The dotted line indicates the position of the maximum for $P = \mathrm{e}^{-1}$. The shaded area indicates the entropy for the range of probabilities corresponding to magnitude distributions with $b \leq 1.5$ and $\Delta M = 0.1$ (discussed below).

The concept of entropy has been widely used in seismology, particularly through the Principle of Maximum Entropy (PME), to study distributions, recurrence relationships, model stress fields, estimate seismic hazard, etc. (Berrill & Davis, 1980; Bookstein, 2021; De Santis et al., 2011; Dong et al., 1984; Feng and Luo, 2009; Janes, 1957; Mansinha & Shen, 1987; Main & Naylor, 2008; Shen & Mansinha, 1983; Telesca et al 2004). Other studies use entropy as an indicator of proximity to criticality (Main & Al-Kindy, 2002; Vogel et al., 2020), some using so-called natural time (Ramírez-Rojas et al., 2018; Rundle et al. 2019; Sarlis et al, 2018; Varotsos et al., 2004; Varotsos et al., 2022; Varotsos et al., 2023) many using other definitions of entropy, and some for seismic electric signals (Varotsos et al., 2006). Entropy has also been used to study the spatial distribution of seismic sources (e.g. Bressan et al., 2017; Goltz, 1966; Goltz and Böse, 2002; Nava et al., 2021; Nicholson et al., 2000; Ohsawa, 2018) and to study noise (e. g. Lyubushin, 2021).

It is because of the possible usefulness of both the *b*-parameter and the magnitude entropy, that it is important to explore the relationship between these two observables.

2. The entropy of seismic magnitude distributions

Now, let the process considered in the information entropy be the seismic magnitudes and the classes be the classes of a magnitude histogram, and let us see what can the entropy be expected to be like by assuming that the magnitudes obey a G-R distribution.

The G-R relation (1) is a reverse cumulative histogram corresponding to an exponential magnitude probability density function,

$$p(m) = \beta\, \mathrm{e}^{-\beta\,(m-m_c)} \tag{6}$$

where

$$\beta = b\, \ln(10) = 1/\,(\mu - m_c)\,, \tag{7}$$

μ is the mean of the exponential distribution, and, since it needs to include all magnitudes that contribute to the rounded ones, is defined for unrounded magnitudes $m \geq m_c$, where $m_c = M_c - \Delta M/2$.

Let the classes considered in (3) correspond to the magnitudes rounded to $\Delta M = 0.1$, and the probability P_i of the class of a given rounded magnitude M_i, where $M_1 = M_c$, is determined from the pdf (6).

Commonly, P_i is approximated from (6) as

$$P_i \approx p(M_i)\, \Delta M \tag{8}$$

(e.g. Rundle et al., 2019); a better procedure will be proposed below, but for now let us digress to discuss some reported results based on this approximation.

2.1 The Entropy of a Continuous Distribution

Substitution of (8) in (4) yields

$$S = -\sum_{i=1}^{K} p(M_i)\, \Delta M\, \log_2[p(M_i)\, \Delta M]\,, \tag{9}$$

which can be written as

$$S = -\sum_{i=1}^{K} p(M_i)\, \log_2 p(M_i)\, \Delta M - \log_2 \Delta M\,, \tag{10}$$

On letting $\Delta M \to 0$ the first term on the right side of (10) becomes what Shannon (1948) defined as the *entropy of a continuous distribution for a process having probability density distribution* $p(m)$:

$$S^c = -\int_{-\infty}^{\infty} p(m)\log_2 p(m)\,\mathrm{d}m \tag{11}$$

which we will denote by S^c to differentiate it from what would be the limit of the entropy in 10). Formula (11), without the minus sign, corresponds to what Wiener (1948) defined as *the amount of information of* $p(M)$, not as entropy. Shannon (1948) states that "The entropies of continuous distributions have most (but not all) of the properties of the discrete case.", and it is clear they differ in this case, because the second term on the right-hand side of (10) has not been included in the limit and this term grows as ΔM decreases and tends to infinity as $\Delta M \to 0$ (Mansinha & Shen, 1987). Goldman (1953) is aware of the $-\log_2 \Delta M$ term, but states that it always cancels out, which is certainly not the case for the problem at hand.

A second problem is that the meaning of $-\log_2 p(m)$ is not clear, because the definitions of self-information and information entropy refer to mass probabilities, not to densities. For exponential distributions, unless $\beta < 1$, i. e., $b < 1/\ln 10 \approx 0.43429448$, which is an unrealistic value, the integral in (11) will include a range with $p(m) > 1$ that would imply negative information and result in negative entropy.

Equation (11) has been used in several studies (e.g. De Santis et al., 2011; Main and Burton, 1964; Posadas et al., 2002; Posadas et al., 2021 Shen & Mansinha, 1983;) with varying results, some of them unfortunate. For example, De Santis et al. (2011) obtained $b_{max} = \mathrm{e}\log_{10}\mathrm{e} {\approx} 1.1805$ as the upper limit for b-values, which illustrates the perils of using (11). Here, we will keep to the original definition of entropy (4)

2.2 Entropy of exponential distributions

Although this paper is oriented towards seismic magnitude distributions, what follows is applicable to any exponential probability distribution.

Coming back to equation (4), instead of using the approximation (8), the exact probability corresponding to the class of a rounded magnitude M_i can be calculated exactly as

$$P_i \equiv P(M_i) = \int_{M_i-\Delta M/2}^{M_i+\Delta M/2} \beta\,\mathrm{e}^{-\beta\,(m-m_c)}\,\mathrm{d}m\,, \tag{12}$$

which results in

$$\begin{aligned} P_i &= \mathrm{e}^{-\beta(M_i-m_c)}\left(\mathrm{e}^{\beta\frac{\Delta M}{2}} - \mathrm{e}^{-\beta\frac{\Delta M}{2}}\right) = \\ &= \mathrm{e}^{-\beta(M_i-M_c)}\left(1-\mathrm{e}^{-\beta\Delta M}\right) \equiv \mathrm{e}^{-\beta(M_i-M_c)}\,\Delta M_P\,. \end{aligned} \tag{13}$$

To show how this probability estimation compares with the approximation shown before, (8) can be written as

$$\beta\, \mathrm{e}^{-\beta(M_i - m_c)} \Delta M = \mathrm{e}^{-\beta(M_i - M_c)} \beta\, \mathrm{e}^{-\beta\, \Delta M/2}\, \Delta M \tag{14}$$

so both (13) and (14), consist of the same exponential multiplied by different factors that are shown in Figure 2 for various values of b. Both factors differ by very little for small b-values, but for large b-values ΔM_p is appreciably larger than the factor in (14), which shows that it is worthwhile to use the exact probability from (13).

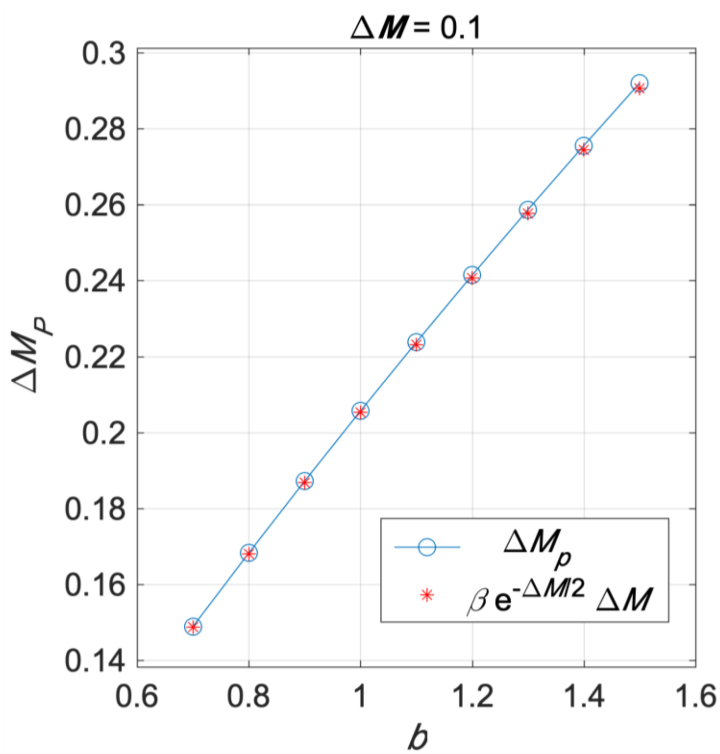


Figure 2. Comparison of the factors that multiply an exponential to evaluate or estimate probabilities.

Substituting probability (13) in (4), yields

$$S = -\sum_{i=1}^{K} \mathrm{e}^{-\beta(M_i - M_c)}\, \Delta M_P \quad \log_2\!\left[\mathrm{e}^{-\beta(M_i - M_c)}\, \Delta M_P\right], \tag{15}$$

which is the expression for the entropy that will be used to calculate the theoretical entropy corresponding to a given magnitude distribution, to illustrate how the elements of the magnitude distribution contribute to the entropy, and to estimate through Monte Carlo simulation, what can be expected from data samples of different sizes.

To obtain an estimate for the theoretical value of S, let $K \to \infty$ in (15) because the theoretical G-R distribution does not have an upper limit; this limit will be discussed below. Equation (15) can be written as:

$$S = -\sum_{i=1}^{\infty} \mathrm{e}^{-\beta(M_i - M_c)}\,\Delta M_P\ \left[-\beta(M_i - M_c)\log_2 \mathrm{e} + \log_2\ \Delta M_P\right],$$

or

$$S = \ \Delta M_P \log_2 \mathrm{e} \sum_{i=1}^{\infty} -\beta(M_i - M_c)\,\mathrm{e}^{-\beta(M_i - M_c)} - \log_2 \Delta M_P \sum_{i=1}^{\infty} \mathrm{e}^{-\beta(M_i - M_c)}\,\Delta M_P\,. \qquad (16)$$

The sum in the second right-hand term of (16) is the total probability equal to unity. In the first right-hand term, the factor $(M_i - M_c)$ takes values $0\Delta M, 1\Delta M, 2\Delta M, 3\Delta M,$ ..., so the sum written explicitly as:

$$0 - 1\Delta M\beta \mathrm{e}^{-1\Delta M\beta} - 2\Delta M\beta \mathrm{e}^{-2\Delta M\beta} - 3\Delta M\beta \mathrm{e}^{-3\Delta M\beta} - \cdots, \qquad (17)$$

can be recognized as the series representation of

$$\Delta M\ \beta\ \frac{d}{dx}(1-\mathrm{e}^{x})^{-1} = \Delta M\beta\,\frac{\mathrm{e}^{x}}{(1-\mathrm{e}^{x})^{\,2}}, \qquad (18)$$

for $x = -\Delta M\,\beta$. Hence, the total entropy of an exponential distribution with parameter β expressed in bits is

$$S = \beta\,\Delta M\ \frac{\mathrm{e}^{-\beta\,\Delta M}}{1-\mathrm{e}^{-\beta\,\Delta M}}\ \log_2 \mathrm{e} - \log_2\!\left(1-\mathrm{e}^{-\beta\,\Delta M}\right). \qquad (19)$$

Equation (19) is a closed, analytic expression for the entropy of an exponential distribution with parameter β and class width ΔM. For a magnitude distribution, since $\Delta M = 0.01$ can be considered to be a set, constant value, (19) can be considered a direct relation between S and β (or $b = \beta \log_{10} \mathrm{e}$). Although β has been used in the derivation of (19), results will be expressed in terms of b, because it is a more familiar parameter and its global average value, a good reference, is conveniently very close to 1.0 (e.g. El-Isa & Eaton, 2014).

The direct, closed, relationship (19) between the b-value and the magnitude entropy is shown in Figure 3. Figure 3 also shows the range of entropies for reasonable b-values: from

$S = 2.98$ bit for $b = 1.5$ to $S = 4.08$ bit for $b = 0.7$; a range of ≈ 1.1 bit for a b range of 0.8. This range has been chosen to illustrate the results because, although b-values in the $0.3 \leq b \leq 2.5$ range have been reported (El-Isa & Eaton, 2014), for estimates based on magnitudes scaling as M_W (Kanamori, 1983; Hanks & Kanamori, 1979) $b = 0.7$ is an adequate lower limit for global b-values (Frohlich & Davis, 1993) and an upper limit of $b = 1.5$ has proposed on physical grounds by Olsson (1999).

Figure 3 shows that S increases as b decreases, so that entropy appears to be directly related to the state of stress in the medium; indeed, since low b corresponds to probabilities being less concentrated around m_c, the significant probabilities are spread over a larger magnitude range, so the medium can be considered as being less ordered.

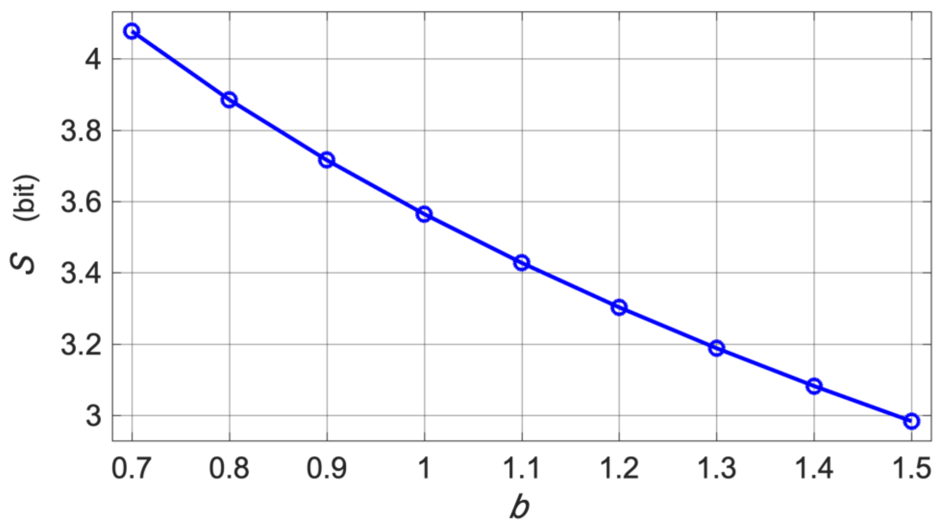


Figure 3 Relationship between b and S.

2.3 Numerical entropies over a finite magnitude range

Now the results of (19) will be checked against numerical results from (13) and (4) to see how results for finite K differ from those for $K \to \infty$. Although the G-R relation does not contemplate an upper limit for M, there are physical limits to how large a magnitude can be, so it is important to consider how results from a finite magnitude range differ from those of an infinite one. It is also important to consider the role of large magnitudes in the entropy determinations.

The very interesting problem of a maximum possible magnitude has been widely addressed (e.g. Beirlant et al., 2019; Chinnery and North, 1975; Kijko, 2004; Kijko & Singh, 2011; Smith, 1976; Sornette, 2009) and manners of dealing with modified G-R distributions or using other distributions have been proposed (e.g. Cornell & Vanmarke, 1969; Cosentino et al., 1977; Holschneider et al., 2011; Lomnitz-Adler & Lomnitz, 1979; Main, 1996; Main & Burton, 1984). The problem of a maximum magnitude is outside the scope of this paper, but it will be seen that the effects of very large magnitudes on entropy estimates are quite low and the possible existence, or not, of very large earthquakes does not affect the results shown here.

To check the results of (19), the entropy of the magnitude distribution will be computed by evaluating exactly from (13) the probabilities for rounded magnitudes in a finite magnitude range. The $2.0 \leq M_i \leq 9.0$ range has been chosen to illustrate the probablities, because M 2.0 is not an uncommon M_c and because M 9.0 is sufficiently rare as to be a practical upper limit because magnitudes much larger than 9.0 (including infinite ones) are not realistic. Since probabilities from a finite range will be considered, they have to be normalized by dividing by a factor

$$f_N = \int_{m_c}^{m_x} \beta\, \mathrm{e}^{-\beta(m-m_c)}\, \mathrm{d}m \;=\; 1 - e^{-\beta(m_x-m_c)}\,. \tag{20}$$

For the proposed magnitude range, this factor differs from 1 by 2×10^{-6} for $b = 0.8$, and by 3×10^{-9} for $b = 1.2$, so corrections are very small and do not affect significantly the results.

Next, these exact theoretical probabilities will be used to calculate each term s_i in the sum (4). and finally $S = \sum_{i=1}^{K} s_i$ will be computed and compared with the analytic total entropy values.

Figure 4 shows in (A) the theoretical probability mass distribution for three representative b-values; (B) shows the $s(M_i)$ corresponding to the probabilities shown in (A), and (C) shows the entropies computed using (4). Strictly speaking, the entropies correspond to the (larger) markers at the end of each curve, but the cumulative s leading to the total entropies is also shown, to illustrate its different behaviors for different b-values. The dotted hrizontal lines in (C) correspond to the analytical entropies.

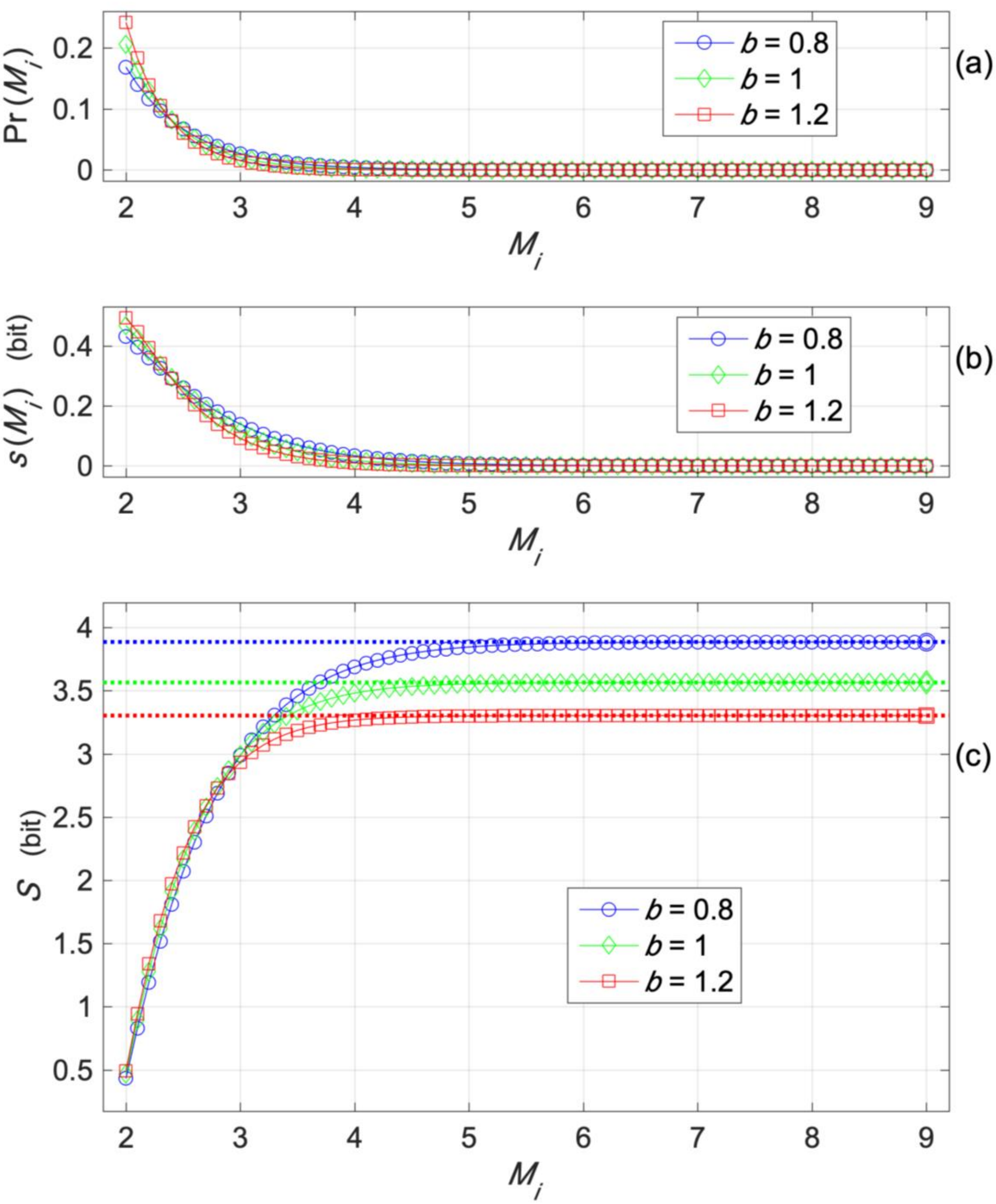


Fig. 4 Exact numerical probabilities (a), corresponding information scores (b) and entropies (c), for three representative *b*-values and a finite magnitude range. Panel (c) shows the numerical entropy values as large symbols over the largest magnitude, and the analytical entropies as dotted lines; also shown are the cumulative *s* values.

For the smallest magnitudes s is largest for the higher b, but about one magnitude unit above M_c the roles are reversed and the entropies for smaller b-values grow faster and soon the entropy for the smallest b is the largest of all. All entropies tend asymptotically to their theoretical values, with the largest b-values approaching it earlier. The magnitudes that make more difference are those in the $3.0 \leq M \leq 5.0$ range.

As shown in Figure 4 (A), large magnitudes have very small probabilities which are close to the left end of the shaded area in in Figure 1 and contribute very little to the total entropy, as shown in (B) and (C). Hence, the presence of magnitudes above 6.5 or 7.0 is not necessary for obtaining good, approximate estimates of S.

The numerical values for the total entropies differ from the analytic ones by only 4.2×10^{-5} for $b = 0.8$, 2.0×10^{-6} for $b = 1.0$, and 9.0×10^{-8} for $b = 1.2$, differences too small to be of practical concern. As would be expected from the properties of the exponential distribution, shifting the magnitude range while conserving the same width, to $1.5 \leq M_i \leq 8.5$, say, results in exactly the same entropy estimates.

Estimates do change if the range is enlarged, for example considering the $1.5 \leq M_i \leq 9.0$ range (five classes wider) reduces the differences between numerical and analytical to 1.7×10^{-5} for $b = 0.8$, 6.7×10^{-7} for $b = 1.0$, and 2.4×10^{-8} for $b = 1.2$, because of the contributions from the extra five terms in (15).

For reference, the entropy of a uniform distribution with K classes is

$$S_U = -\sum_{i=1}^{K} \frac{1}{K} \log_2 \frac{1}{K} = \log_2 K \,, \qquad (22)$$

so for the example, with range $2.0 \leq M_i \leq 9.0$ and $K = 71$, the entropy of the uniform distribution, i.e., the largest possible entropy, would be $S_U = 6.15$ bit, some 2.26 bit larger than the entropy for $b = 0.8$.

The total entropies are distintictly larger for the smaller b-values, which means that measuring entropies can be a good method for identifying regions of low or large b, that is, of large or low stress.

2.4 Numerical entropy from samples

Next, it will be seen how entropy measured from samples compares to the entropy computed from exact probabilities, and how it depends on the sample size; the samples will be synthetics from random simulations, for the same magnitude range and the three representative *b*-values used above.

For each *b*-value, *N* exponentially distributed random magnitudes are generated as

$$m = m_c - \ln(1 - r * \rho) / \beta , \tag{23}$$

where *r* is a uniformly distributed pseudo-random number in the zero to one range, and

$$\rho = 1 - \mathrm{e}^{\beta\,(m_c - m_x)} \tag{24}$$

maps this range onto the range that results in probabilities $m_c \leq m \leq m_x$.

With these magnitudes a histogram with classes ΔM wide, corresponding to the rounded magnitudes, is constructed and the number of events in each class $n(M_i)$ is counted. The probabilities are estimated as

$$P_i = n(M_i) / N \tag{25}$$

(c.f. Feng & Luo, 2009) and used to calculate the s_i values and thence *S*.

Figure 5 shows simulations for three *b*-values, each having $N = 5000$ magnitudes, a reasonably good-sized sample. The magnitude histograms $n(M_i)$ are shown in (A), and the contributions $s(M_i)$ are shown in (B); the cumulatives for *s* and the entropies are shown in (C), together with the theoretical entropies.

A comparison of panels (C) of Figures 4 and 5 shows very good agreement between entropies from theoretical and simulated magnitudes, both converging nicely to the analytic entropies from (19).

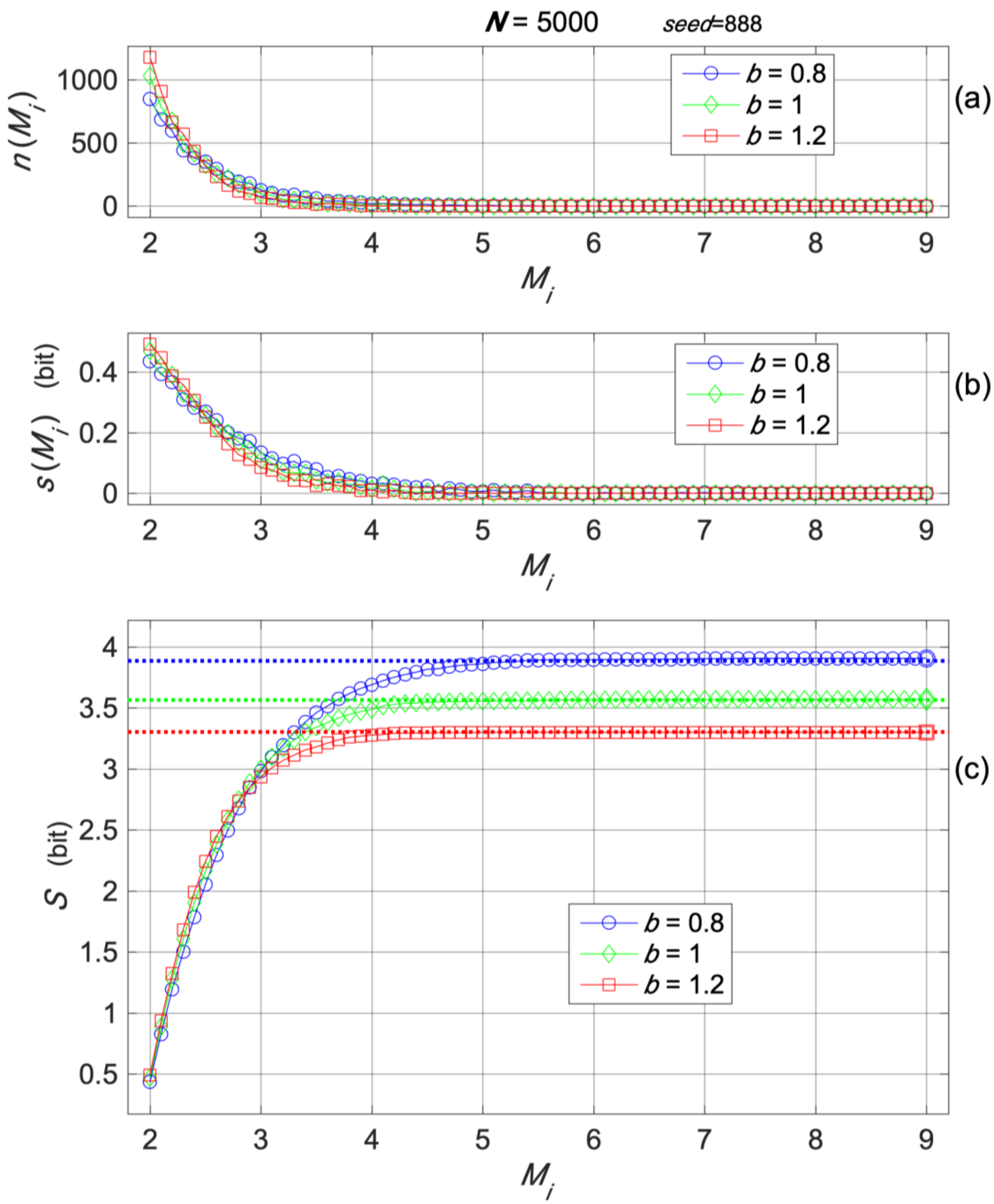
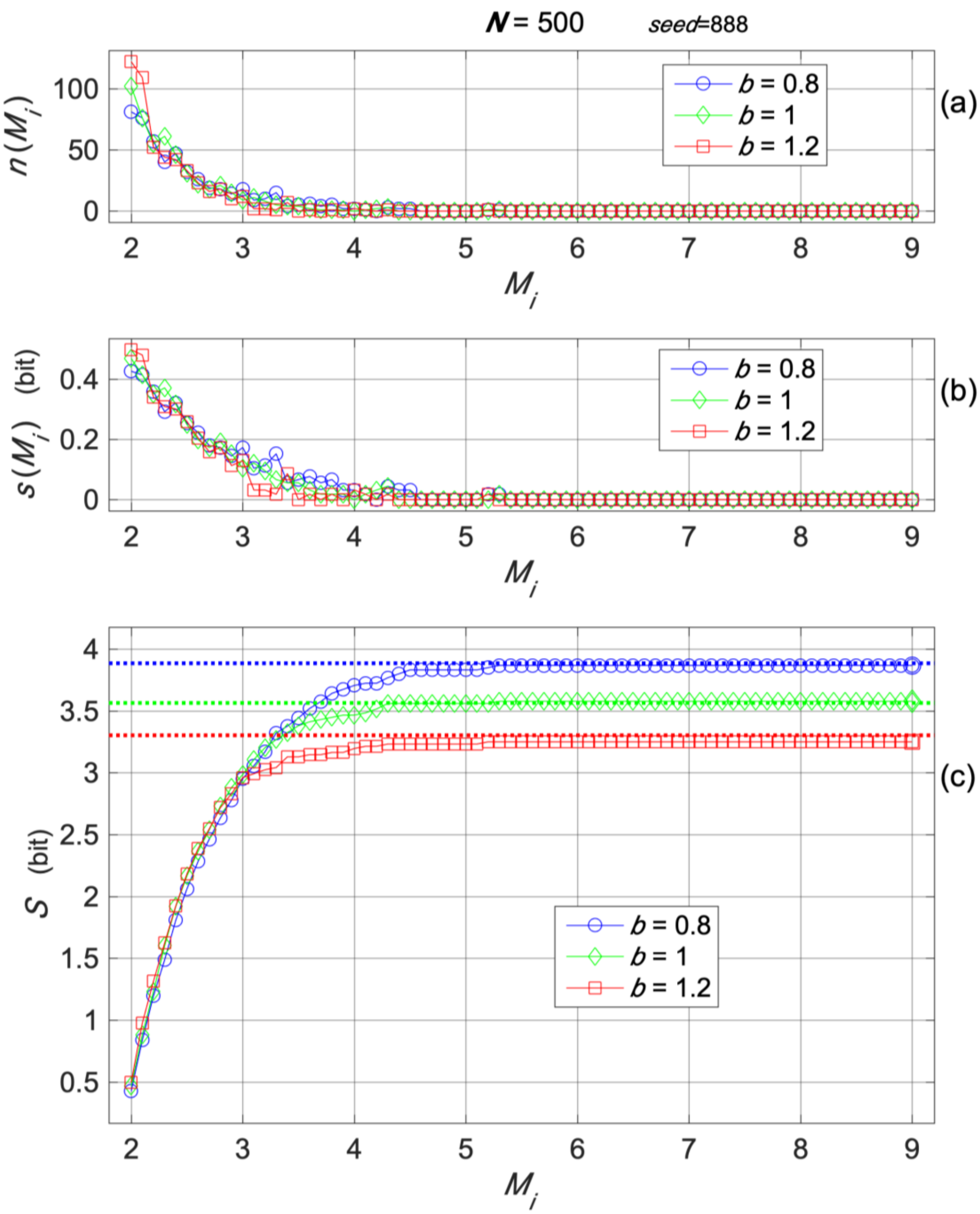


Fig. 5 Numerical probabilities from a synthetic sample of 5,000 magnitudes (a), corresponding information scores (b) and entropies c), for three representative *b*-values and a finite magnitude range. Panel (c) shows the numerical entropy values as large symbols over the largest magnitude, and the analytical entropies as dotted lines; also shown are the cumulative *s* values.

The simulations shown in Figure 6 are like those of Figure 5, but for a much smaller sample of $N = 500$ magnitudes. The histograms in the (A) and (B) panels show clear differences from the respective graphs in Figure 5; differences are less apparent between panels (C), but there is a noticeable difference for the entropy corresponding to the largest $b = 1.2$, which is well below the analytic entropy.



342 Fig. 6 Numerical probabilities from a synthetic sample of 5,000 magnitudes (a), corresponding
343 information scores (b) and entropies (c), for three representative *b*-values and a finite magnitude
344 range. Panel (c) shows the numerical entropy values as large symbols over the largest magnitude,
345 and the analytical entropies as dotted lines; also shown are the cumulative *s* values.
346

2.5 Monte Carlo simulations and sample size

Monte Carlo simulations are used to characterize how numerical entropies depend on sample size, each simulation consisting of $N_r = 5000$ realizations, like those shown in the previous section, of magnitude samples of different sizes, from $N = 250$ to $N = 5000$. The means and standard deviations of the N_r entropies calculated for each combination of b and N, are shown in Figure 7.

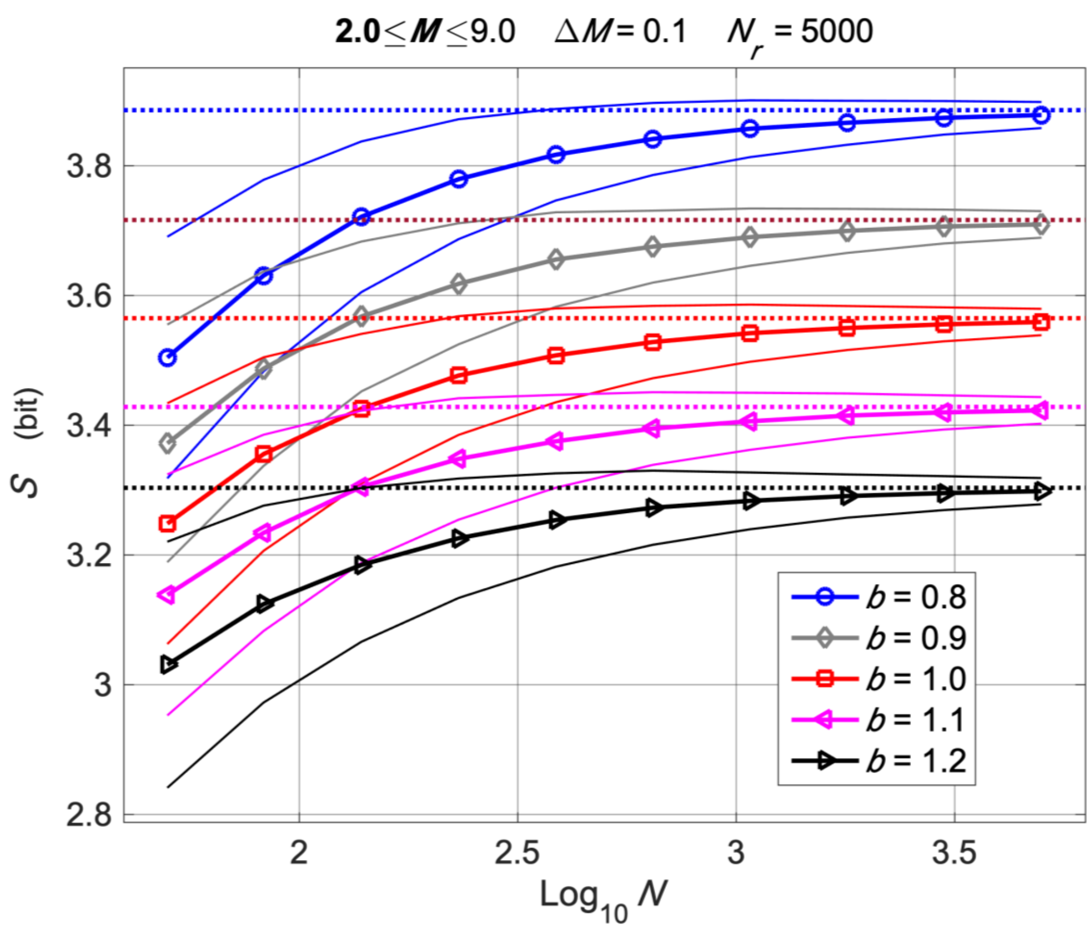


Fig. 7 Monte Carlo analysis of entropies S determined from synthetic samples for different sample sizes N. Thick lines with different colors and symbols, corresponding to representative b-values are the means of 5,000 realizations for each combination of b and N. The thin lines show the means plus/minus one standard deviation, and the horizontal dotted lines indicate the analytical entropies.

Figure 7 shows the mean calculated S for each b-value as a thick line with a particular color and symbol, shows the mean plus/minus one standard deviation as thin lines and the true analytical value as a dotted line in the corresponding color. In order to interpret correctly the information in the standard deviations it is necessary to determine how the entropy value determinations are distributed, and Figure 8 shows an example of these distributions for $b = 1.0$ and $N = 5000$, which tells that the values can be considered to be normally distributed around the mean.

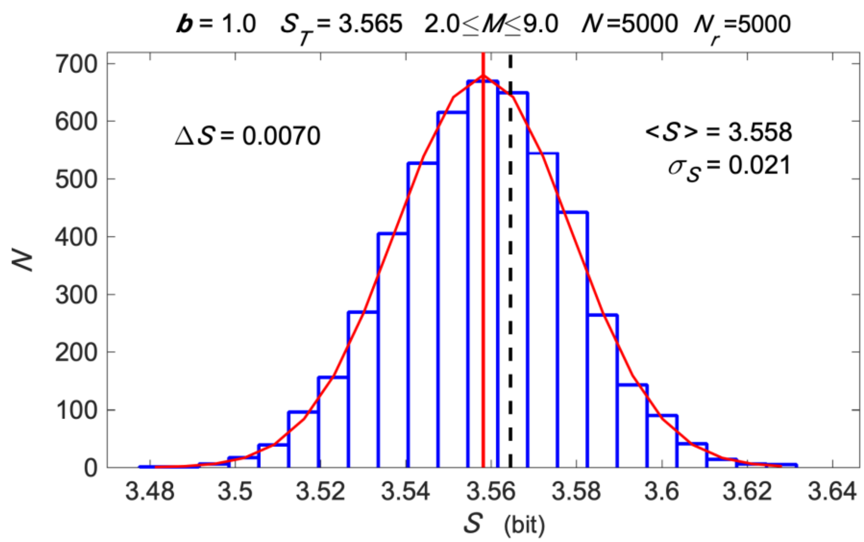


Fig.8. Histogram of $N_r = 5{,}000$ Monte Carlo entropy determinations for $b = 1.0$ and magnitudes in the $2.0 \leq M \leq 9.0$ range (blue line); te vertical red line shows the mean value, and the vertical dashed line is the analytical S value. The thin line is the normal distribution for the observed standard deviation σ_S multiplied by N_r.

Figure 7 shows that the entropy estimated from samples smaller than ~200 will almost certainly be undervalued, particularly for low b. Entropies corresponding to b-values differing by as much as 0.1 cannot be distinguished with 0.7 confidence for samples smaller than about 350 for low b and about 550 for high b; distinguishing them with 0.95 confidence requires ~1,500 and ~3,00 samples, respectively.

For samples ~2000 to ~2500, mean values underestimate the analytical entropy by ~0.01 bit, and for samples of 5000 the underestimations go from 0.0067 bit for $b = 0.8$ to 0.0051 bit for $b = 1.2$, with standard deviations ~0.02 bit. For the larger samples the means tend to the analytical entropies very slowly, and including larger magnitudes does not help very much because their number is very small and, as shown in Figures 1, 5, and 6, their contribution to the total entropy is almost insignificant.

Standard deviations diminish slowly, and even for large samples ~5,000 the standard deviation corresponding to $b = 0.8$, $\sigma_S = 0.0201$, is ~0.005 of the mean value $\bar{S} = 3.8778$, while for $b = 1.2$, $\sigma_S = 0.0202$ is ~0.0061 of the mean value $\bar{S} = 3.2978$. These normalized standard deviations are smaller than the corresponding ones for *b*-values estimated by the Aki-Utsu method for the same synthetic samples used to evaluate the entropies.

Figure 7 shows that, although entropies evaluated over a finite magnitude range should be smaller than the analytical ones, the entropies measured from samples could be overvalued and thus be slightly larger.

2.6 Measured entropies and *b*-values for single trials

It has been discussed how entropies are measured from data, and Figure (7) shows how the measurements can expected to agree with the real values, but in practice the real values are not known nor are there thousands of realizations; usually the data correspond to a single realization and there is no way of knowing how well it conforms to the behavior of the means shown above.

Since there is an explicit relation between S and b, it would seem that their measures would be redundant but this is not exactly the case because they are measured in different ways. b-value measurements (2) depend only on $\bar{M}$, while entropy estimations depend on the values of all entropy scores s_i.

In order to illustrate how single realizations agree with, or differ from, the means of many realizations and from the true values, let us look at four examples of single sample realizations, and see how each single realization depends on sample size. All realizations share exactly the same parameters and differ only in the number used as seed for the pseudo-random number generator. Each realization was a set of $N_T = 5{,}000$ magnitudes, and we obtained estimates of S, using (25) and (4), and b, using (2), for subsets of $N = 500, 600, 700, \ldots, 5{,}000$, and from each b, we calculated the entropy using (19).

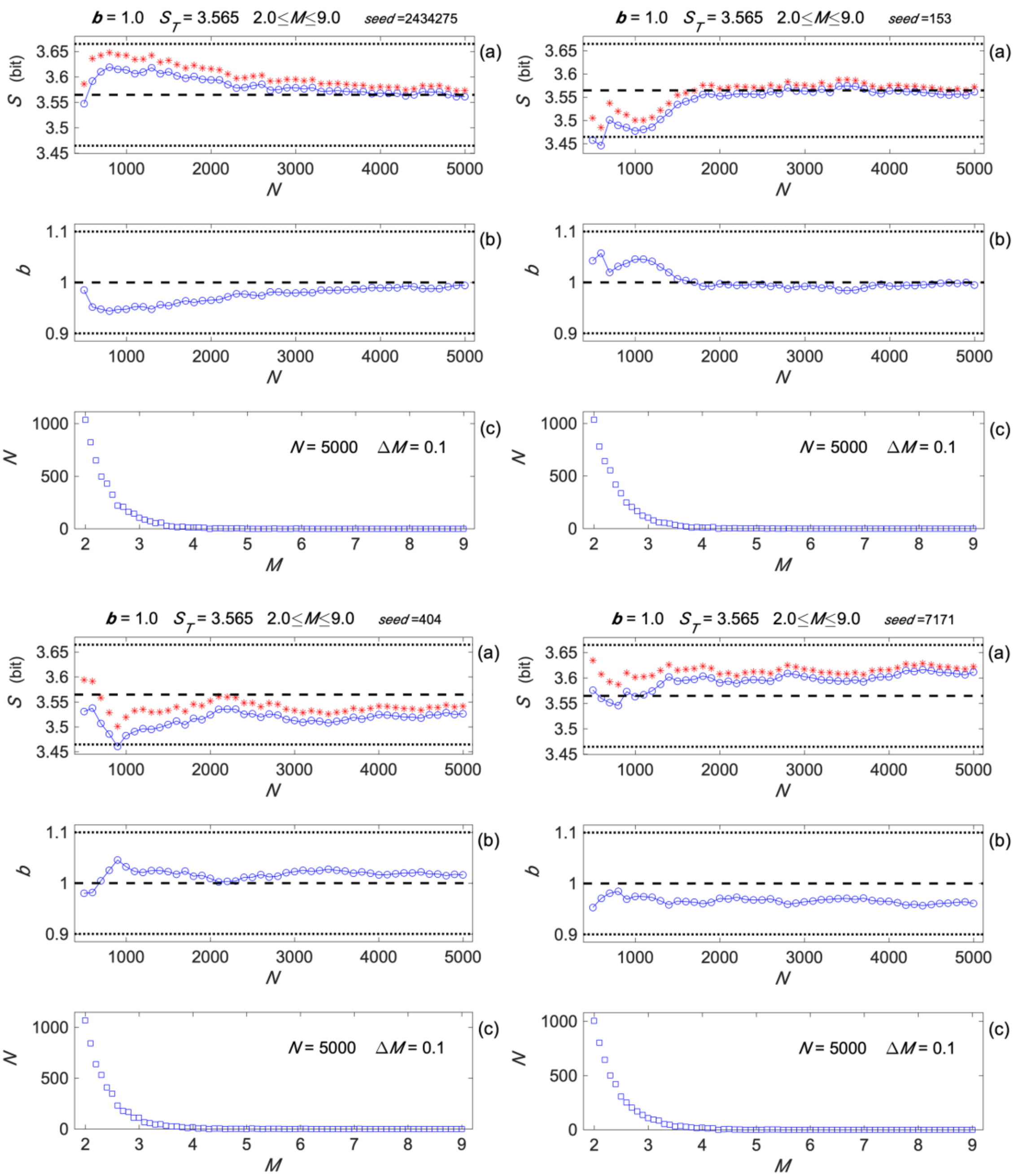


Figure 9 Four examples of entropy and b-value determinations from single realizations of $N_T = 5{,}000$ synthetic exponentially distributed magnitudes, taken N elements at a time. The (c) panels show the magnitude histograms for the total N_t data, the (b) panels show the b-values estimated using (2) with $\bar{M}$ deternined from N data (blue circles) and the true b-value (dashed red line). The (a) panels show as blue circles the entropies determined for each N data, and as asterisks the entropies estimated from the measured b-values in (b), using (19); the analytical entropy corresponding to the true b is shown as a dashed red line.

The examples are shown in Figure 9, where panels (c) plot the histograms of the total N_T magnitudes to show that the synthetic magnitudes are indeed exponentially distributed. Panels (b) show the estimated b-values and, for reference, the true b-value, while panels (a) show the estimated entropies as blue circles, the analytic entropy corresponding to the true b, and show as asterisks the entropies computed from the estimated b-values.

As mentioned above, the realizations in Fig. 9 differ only in the random number seed, and illustrate how a realization corresponding to some set of real data can vary randomly while being a product of a given conditions on a given seismic system. The two upper examples show “expected” behaviors, with values varying considerably for short samples and gradually converging to a value close to the true one, albeit one (upper left) from above and the other (upper right) from below. The example at lower left does converge but does not reach the true value, and the example at lower right does not converge to the true value at all. It should be said that most realizations behave more like the good examples, so that many different seeds were tried before the ugly example at lower right was obtained.

All the examples show that for small data sets the measured entropies and those eestimated from the b-values differ very much for small samples, but run almost parallel for large samples. Entropies from b estimates are larger than measured ones, but that is to be expected because of the finite magnitude range. Thus, it is proposed that, although related and calculated from the same data, entropy and b measurements are not just scaled versions of each other, because they are calculated in different ways that are sensitive to different kinds of errors, and when both measurements are correct they should agree within the limitations. Hence, the differences between directly estimated entropies and those estimates from b-values can help us determine when samples are adequate and results are trustworthy.

As an example of how the entropy and Aki-Utsu b-value measurements are not equivalent, consider the contribution of a very large magnitude. Let a data set have $N-1$ elements, and let the entropy determined from the sample be

$$S_{[N-1]} = \sum_{i=1}^{N-1} p_i \log_2 p_i\,, \tag{26}$$

and the b-value be

$$b_{[N-1]} = \frac{\ln 10}{\bar{M}_{[N-1]} - m_c}\,, \tag{27}$$

where $\bar{M}_{[N-1]} = \frac{1}{N-1}\sum_{i=1}^{N-1} M_i$. Now, let the next magnitude M_N be large enough so so that it stands alone in a class, then, because there is only one event in the class, its probability will be 1/N, so

$$S_{[N]} = \frac{N-1}{N} S_{[N-1]} + \frac{N-1}{N} \log_2 \frac{N-1}{N} + \frac{1}{N} \log_2 \frac{1}{N} \qquad (28)$$

and the change of entropy does not depend on the value of M_N, as long as it is large enough to stand alone in a class. On the other hand,

$$b_{[N]} = \frac{\ln 10}{\bar{M}_{[N-1]} - m_c + \frac{1}{N}\left(M_N - \bar{M}_{[N-1]}\right)} \qquad (29)$$

does depend on the actual value of M_N. Hence, unless N is very large the effect of a large magnitude is different for entropies and for *b*-values.

5 Discussion and Conclusions

An analytical relationship between the *b*-value, or β, that characterize the magnitude G-R distribution, or any other exponential distribution, and the information entropy of the distribution has been found and checked by means of the numerical evaluation of the entropy computed using the exact probabilities derived from the distribution.

Since neither the G-R distribution nor the associated exponential distributions contemplate a maximum magnitude it was possible to evaluate the effect of working with a finite magnitude range on the entropy, and it was found that, because very small probabilities contribute very little to the entropy, the difference between the analytical and the finite range entropies is quite small.

Next, the results of the relationship were compared with entropies estimated from synthetic sets of exponentially distributed random data, and very good agreement was found.

Using Monte Carlo simulations, the accuracy and precision for entropy evaluations as a function of sample size were explored. The evaluations were found to be distributed normally around their means, which allows setting familiar confidence limits to the capacity of discriminating between different values of the entropy.

Although *b*-values and entropies are formally related, they are evaluated from the data by different methods and so are affected differently by different characteristics of the data, particularly for small data sets. Hence, it is proposed that entropy and G-R *b*-value measurements can be complementary and help to estimate when a sample is large enough for results to be reliable.

Acknowledgments

Statement

No real of perceived finantial conflicts for the author.

No other affiliations.

No external funds.

Open Research

No data were used

References

Aki, K. (1965). Maximum likelihood estimate of b in the formula log_N_= a - bM and its confidence limits. *Bull. Earthq. Res. Inst. Tokyo Univ., 43*, 237–239.

Aki, K. (1981). A probabilistic synthesis of precursory phenomena, in *Earthquake Prediction: An International Review*, Maurice Ewing Set., vol. 4, edited by D. W. Simpson and P. G. Richards, pp. 566-574, AGU, Washington, D.C.

Beirlant, J., Kijko, A., Reynkens, T., & Einmahl, J. (2018). Estimating the maximum possible earthquake magnitude using extreme value methodology: the Groningen case. *Natural Hazards, 98*, 1091-1113.

Berrill, J., & Davis, R. (1980). Maximum entropy and the magnitude distribution, *Bull. Seismol. Soc. Am., 70*, 1823–1831.

Bookstein, F. (2021). Estimating earthquake probabilities by Jaynes's method of maximum entropy. *Bull. Seismol. Soc. Am, 111*(5), 2846-2861. DOI: 10.1785/0120200298

Bressan. G., Barnaba, C., Gentili, S., & Rossi, G. (2017). Information entropy of earthquake populations in northeastern Italy and western Slovenia. *Pure Appl. Geophys., 271,* 29–46. https://doi.org/10.1016/j.pepi.2017.08.001

Chinnery, M., & North, R. (1975). The frequency of very large earthquakes. *Science, 190* (4220), 1197-1198.

Cornell, C., & Vanmarcke, E. (1969). The major influences in seismic risk. *Proceeds 4th World Conf. Earthquake Eng. Santiago, Chile*, 69-83.

Cosentino, P., Ficarra, V., & Luzio, D. (1977). Truncated exponential frequency-magnitude relationship in earthquake statistics. *Bull. Seism. Soc. Am., 67*, 1615-1623.

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. Seismol. Soc. Am., 101*, 1386-1395. doi: 10.1785/0120090390

DeSalvio, N., & Rudolph, M. (2021). A Retrospective Analysis of b-Value Changes Preceding Strong Earthquakes. *Seismol. Res. Lett., 93*, 364–375, doi: 10.1785/0220210149.

Dong, W., Bao, A., & Shah, H. (1984). Use of maximum entropy principle in earthquake recurrence relationships. *Bull. Seismol. Soc. Am., 74*, 2, 725–737.

El-Isa, Z., & Eaton, D. (2014). Spatiotemporal variations in the b-value of earthquake magnitude–frequency distributions: Classification and causes. *Tectonophysics, 615*, 1-11.

Enescu, B., & Ito, K. (2001). Some premonitory phenomena of the 1995 Hyogo-Ken Nanbu (Kobe) earthquake: seismicity, *b*-value and fractal dimension. *Tectonophysics, 338*(3-4), 297-314.

Fano, R. (1961). *Transmission of information*. MIT Press-J.Wiley & Sons, 389 pp.

Feng, L., & Luo, G. (2009). The relationship between seismic frequency and magnitude as based on the maximum entropy principle. *Soft. Comp., 13*, 979–983.

Frohlich, C., & Davis, S. (1993). Teleseismic *b* values; or, much ado about 1.0. *J. Geophys. Res., 98* (B1), 631–644, doi: 10.1029/ 92JB01891.

Geffers, G., Main, I., & Naylor, M. (2022). Biases in estimating b-values from small earthquake catalogues: how high are high b-values? *Geophjys. J. Int., 229*, 1840-1855. https://doi.org/10.1093/gji/ggac028

Godano, C., Tramelli, A., Petrillo, G., & Convertito, V. (2024). Testing the predictive power of b value for Italian seismicity. *Seismica,* 3.1. DOI: 10.26443/seismica.v3i1.1084.

Goldman, S. (1953). *Information theory*. Prentice Hall, Inc., 385pp.

Goltz, C. (1996). Multifractal and entropic properties of landslides in *Japan. Geol Rundsch, 85*, 71–84.

Goltz, C., & Böse, M. (2002). Configurational entropy of critical earthquake populations. *Geophys. Res. Lett., 29*(20):51-1–51-4. https://doi.org/10.1029/2002GL015540

Gutenberg, B., & Richter, C. (1944). Frequency of earthquakes in California. *Bull. Seismol. Soc. Am., 34*, 185–188. [D]

Guttorp, P. (1987). On least-squares estimation of b values. *Bull. Seism. Soc. Am., 77*, 2115-2124.

Hanks, T., & Kanamori, H. (1979). A moment magnitude scale, *J. Geophys. Res., 84* (B5), 2348-2350.

Harte, D., & Vere-Jones, D. (2005). The entropy score and its uses in earthquake forecasting. *Pure and Applied Geophysics, 162*, 1229-1253.

Hirata, T. (1989). A correlation between the b value and the fractal dimension of earthquakes. *Journal of Geophysical Research: Solid Earth, 94*(B6), 7507-7514. [D Contradicts Aki $D = 2.3 - 0.73b$.

Holschneider, M., Zöller, G., & Hainzl, S. (2011). Estimation of the Maximum Possible Magnitude in the Framework of a Doubly Truncated Gutenberg–Richter Model. *Bull. Seism. Soc. Am., 101*, 1649-1659. DOI: 10.1785/0120100289

Hu, N., Han, P., Wang, R., Shi, F., Chen, L., & Li, H. (2024). Spatial Heterogeneity of *b* Values in Northeastern Tibetan Plateau and Its Interpretation. *Entropy, 26*, 182. https://doi.org/10.3390/e26030182

Ishimoto, M., & Iida, K. (1939). Observations sur les seisms enregistres par le microseismograph construite derniement (I). *Bull. Earthq. Res. Inst., 17,* 443-478.

Jaynes, E. (1957). Information theory and statistical mechanics. *Phys. Rev., 106*(4), 620–630.

Kagan, Y. (1999). The universality of the frequency-magnitude relationship. *Pure Appl. Geophys., 155*, 537–574.

Kanamori, H. (1983). Magnitude scale and quantification of earthquakes. *Tectonophysics*, *93* (3-4), 185–199.

Kijko, A. (2004). Estimation of maximum earthquake magnitude, m_{max}. *Pure Appl. Geophys., 161*, 1655-1681. DOI 10.1007/s00024-004-2531-4

Kijko A, & Singh M (2011). Statistical tools for maximum possible earthquake estimation. *Acta Geophys., 59*(4), 674–700.

Li, Y., & Chen, X. (2021). Variations in apparent stress and b value preceding the 2010 MW8.8 Bio-Bío, Chile earthquake. *Pure Appl. Geophys. 178*, 4797-4813. https://doi.org/10.1007/s00024-020-02637-3

Lomnitz-Adler, J., & Lomnitz, C. (1979). A modified form of the Gutenberg-Richter magnitude-frequency relation. *Bull. Seism. Soc. Am., 69*, 1209-1214.

Lyubushin, A. (2021). Seismic noise wavelet-based entropy in Southern California. *Journal of Seismology*, *25*(1), 25-39.

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

Main, I. (1996). Statistical physics, seismogenesis and seismic hazard. *Rev. Geophys., 34*(4):433–462. ttps://doi.org/10.1029/96 RG02808

Main, I., & Al-Kindy, F. (2002). Entropy, energy, and proximity to criticality in global earthquake populations, *Geophys. Res. Lett., 29* (7). doi 10.1029/2001GL014078.

Main, I., & Naylor, M. (2008). Maximum entropy production and earthquake dynamics. *Geophys. Res. Lett.*, *35*(19), L19311, doi:10.1029/2008GL035590

Mansinha, L., & Shen, P. (1987). On the magnitude entropy of earthquakes. *Tectonophysics*, *138*(1), 115-119.

Marzocchi, W., & Sandri, L. (2003). A review and new insights on the estimation of the b-value and its uncertainty. *Ann. Geophys., 46*(6), 1271–1282.

Marzocchi, W., Spassiani, I, Stallone, A., & Taroni, M. (2020). How to be fooled searching for significant variations of the *b*-value. *Geophys. J. Int., 220* (3), 1845–1856, doi: 10.1093/gji/ggz541.

Monterroso, D. A., & Kulhánek, O. (2003). Spatial variations of b-values in the subduction zone of Central America. *Geofísica Internacional 42* (4), 575-587.

Nanjo, K., Hirata, N., Obara, K., & Kasahara, K. (2012). Decade-scale decrease in *b* value prior to the M9-class 2011 Tohoku and 2004 Sumatra quakes. *Geophys. Res. Lett. 39*, L20304. https://doi.org/10.1029/2012GL052997

Nava, F., Márquez-Ramírez, V., Zúñiga, F., Ávila-Barrientos, L., & Quinteros-Cartaya, C. (2017). Gutenberg-Richter b-value maximum likelihood estimation and sample size. *J. Seismol., 21*, 127-135. DOI: 10.1007/s10950-016-9589-1.

Nava, F., Despaigne, G., & Glowacka, E. (2021). Poisson renormalized entropy as a possible precursor to large earthquakes. *Journal of Seismology, 25*, 1407-1425. https://doi.org/10.1007/s10950-021-10041-0.

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

Ohsawa, Y. (2018). Regional seismic information entropy to detect earthquake activation precursors. *Entropy, 20*(11), 861.

Ogata, Y., & Yamashina, K. (1986). Unbiased estimate for *b*-value of magnitude frequency. *Journal of Physics of the Earth*, *34*(2), 187-194.

Olsson, R. (1999). An estimation of the maximum *b*-value in the Gutenberg-Richter relation. *Journal of Geodynamics*, *27*(4-5), 547-552.

Posadas, T. Hitara, T., & Vidal, F. (2002). Information theory to characterize spatiotemporal patterns of seismicity in the Kanto Region. *Bull. Seismol. Soc. Am., 92*, 600–610.

Posadas, A., Morales, J., & Posadas-Garzon, A. (2021). Earthquakes and entropy: Characterization of occurrence of earthquakes in southern Spain and Alboran Sea. Chaos, 31(4), 043124. doi: 10.1063/5.0031844

Ramírez-Rojas, A., Flores-Márquez, E., Sarlis, N., & Varotsos, P. (2018). The complexity measures associated with the fluctuations of the entropy in natural time before the deadly México M8. 2 earthquake on 7 September 2017. *Entropy*, *20*(6), 477. doi:10.3390/e20060477

Richter, C. (1958). *Elementary seismology*. W.H.Freeman and Co., USA, 768pp.

Rundle, J., Giguere, A., Turcotte, D., Crutchfield, J., & Donnellan, A. (2019). Global seismic nowcasting with Shannon information entropy. *Earth and Space Science*, *6* (1), 191-197.

Sarlis, N., Skordas, E., & Varotsos, P. (2018). A remarkable change of the entropy of seismicity in natural time under time reversal before the super-giant M9 Tohoku earthquake on 11 March 2011. *Europhysics Letters, 124*(2), 29001. doi: 10.1209/0295-5075/124/29001

Scholtz, C. (2015). On the stress dependence of the earthquake *b* value. *Geophys. Res. Lett., 42*:1399–1402. https://doi. org/10.1002/2014GL062863

Shannon, C. (1948). A mathematical theory of communication. *Bell Syst. Tech. J., 27*, 379-423, 623-656.

Schorlemmer, D., Wiemer, S., & Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes, *Nature, 437*, 539–542.

Shen, P. Y., & Mansinha, L. (1983). On the principle of maximum entropy and the earthquake frequency-magnitude relation, *Geophys. J. International, 74*, 777–785.

Shi, Y., & Bolt, B. (1982). The standard error of the magnitude-frequency b value. *Bull. Seismol. Soc. Am., 72*, 1677–1687.

Smith, S. (1976). Determination of maximum earthquake magnitude. *Geophysical Research Letters*, *3*(6), 351-354.

Sornette, D. (2009). Dragon-kings, black swans and the prediction of crises. *Swiss Finance Inst. Res. Paper 09-36*, 1-18.

Telesca, L., Lapenna, V., & Lovallo, M. (2004). Information entropy analysis of Umbria-Marche region (central Italy), *Nat. Hazards Earth Syst. Sci., 4*, 691–695.

Tinti, S., & Mulargia, F. (1987). Confidence intervals of b-values for grouped magnitudes, *Bull. Seismol. Soc. Am., 77*, 2125–2134.

Utsu, T. (1965). A method for determining the value of b in a formula log n= a - bM showing the magnitude frequency relation for earthquakes. *Geophys. Bull. Hokkaido Univ.*, *13*, 99-103.

Utsu, T. (2002). Statistical features of seismicity. In *International Handbook of Earthquake & Engineering Seismology*, Lee, W., Kanamori, H., Jennings, P., Kisslinger, C. (Eds.), 43, 719-732.

Varotsos, P., Sarlis, N., Skordas, E., & Lazaridou, M. (2004). Entropy in natural time domain. *Phys. Rev. E, 70*, 011106.

Varotsos, P., Sarlis, N., Skordas, H., Tanaka, H., & Lazaridou, M. (2006). Entropy of seismic electric signals: Analysis in natural time under time reversal. *Physical Review E, 73*, 031114. DOI: 10.1103/PhysRevE.73.031114

Varotsos, P., Sarlis, N., Skordas, E., Nagao, T., Kamogawa, M., Flores-Márquez, E., et al. (2023). Improving the Estimation of the Occurrence Time of an Impending Major Earthquake Using the Entropy Change of Seismicity in Natural Time Analysis. *Geosciences, 13* (8), 222.

Vogel, E., Brevis, F., Pastén, D., Muñoz, V., Miranda, R., & Chian, A. (2020). Measuring the seismic risk along the Nazca–South American subduction front: Shannon entropy and mutability. *Natural Hazards and Earth System Sciences, 20*(11), 2943-2960.

Wang, J. (2016). A mechanism causing *b*-value anomalies prior to a mainshock. *Bull. Seismol. Soc. Am., 106*, 1663-1671, DOI:10.1785/0120150335

Wiener, N. (1948). *Cybernetics*. MIT Press, Cambridge, Massachussets.

Wyss, M. (1973). Towards a Physical Understanding of the Earthquake Frequency Distribution. *Geophysical Journal of the Royal Astronomical Society, 31*(4):341–359. DOI: 10.1111/j.1365- 246X.1973.tb06506.x.

Wyss, M., Sammis, C., Nadeau, R., & Wiemer, S. (2004). Fractal dimension and b-value on creeping and locked patches of the San Andreas fault near Parkfield, California. *Bull. Seismol. Soc. Am., 94* (2), 410–421.