J Seismol
https://doi.org/10.1007/s10950-021-10041-0

ORIGINAL ARTICLE

Poisson renormalized entropy as a possible precursor to large earthquakes

F. Nava · G. Despaigne · E. Glowacka

Received: 15 October 2020 / Accepted: 11 August 2021

Abstract The concept of Poisson renormalized entropy is presented as an observable of background seismicity with precursory possibilities. The usual way of estimating entropies evaluating probabilities directly from the normalized number of incidences is shown to be lacking in sensitivity and seriously subject to saturation. As an example, the concept of Poisson renormalized entropy and the associated method are applied to the seismicity previous to the 2011 Tohoku M_W = 9.1 megaearthquake, and shown to be able to identify precursory changes in the seismicity. Results are compared with those from fractal dimension and Gutenberg-Richter *b*-value studies.

Keywords Seismic precursors · Entropy · Hazard · Poisson

1 Introduction

In the endeavor to attain useful and reliable seismic forecasts, essential for the diminution of seismic risk, it is necessary to use seismic precursors, anomalies in observables usually indicative of high-stress levels (Evison 1999). To date, there is no instance of a *perfect precursor*, i.e., one that appears if and only if a large earthquake is to occurr in the near future. Furthermore, precursors may be difficult to identify (e.g., Adamaki and Roberts 2016, 2017), and the significance, reliability, and other characteristics of proposed precursors are poorly known; hence, it makes sense to try to use the presence or absence of as many precursors as possible to correctly estimate seismic hazard.

Among the commonly used precursors based on background seismicity caracteristics are changes in fractal dimension (e.g., Main 1996; Bressan et al. 2017); it has been observed that fractal dimensions diminish; i.e., hypocenter clustering increases, before large earthquakes (e.g., Márquez et al. 2012); hence, we can expect the spatial entropy of the system to be less than for the case when events are distributed uniformly (Ling-ren 1988; Goltz,C., 1997; Turcotte 1997; Berrill and Davis 1980; Al-Kindy and Main 2003; Rundle et al. 2003, Bressan et al. 2017).

Billio et al. (2015) use entropy as an indicator for systemic risk. Main and Al-Kindy (2002) examined the question of proximity of the global earthquake population to the critical point, basing their analyses on the entropy and energy of annual seismic frequency from catalog data (but see also the comments by Chen and Chang (2004), with reply by Main and Al-Kindy (2004)). Goltz and Böse (2002) concluded that the Earth's crust is in a state of intermittent criticality, supported by significant fluctuations of ±20% in configurational entropy.

F. Nava (✉) · G. Despaigne · E. Glowacka
Seismology Department, Centro de Investigación Científica y de Educación Superior de Ensenada, B.C., Baja California
22860 Ensenada, Mexico
e-mail: fnava@cicese.mx

G. Despaigne
Universidad Tecnológica de la Mixteca, Instituto de Minería, Huahuapan de León, Oaxaca 69000, México

Published online: 17 September 2021

We will explore the possibility of using a new measure of entropy, the *Poissonian renormalized entropy*, as a precursory observable, because it has a greater sensitivity than the entropy measure commonly used.

1.1 Information and entropy

Let there be a system comprised of a set of K discrete states, let p_j be the probability of the j th state, such that

$$\sum_{j=1}^{K} p_j = 1 \tag{1}$$

The Shannon (1948) entropy of the system is defined as

$$S = -h \sum_{j=1}^{K} p_j \log_2 p_j, \tag{2}$$

where h is some conventional positive constant that relates entropies calculated using logarithms with different bases; we will assume $h = 1$. Since the Shannon (or *surprise*) information of the jth state is

$$I_j = -\log_2 p_j,$$

where use of the base 2 logarithm results in the information having units of bits, it is clear from (2) that the entropy is the expected value of the information over the system.

$$S = E(I) \tag{3}$$

and will thus be expressed in bits.

Figure 1 illustrates the contribution of each p_j to the sum in (2), and shows that very high and very low probabilities are the ones that contribute less to the sum. In particular, the information from $p_j = 1$ is zero and, in spite of having infinite information, a state with $p_j = 0$ contributes nothing to the entropy.

1.2 Entropy and seismicity

We will now apply the concept of entropy to the characterization of seismicity. Let the system be a seismic region, and let the states be the K cells, 2D or 3D, defined by a grid covering the observed region. The probability of each state will be related to the number of *events*, seismic epicenters or hypocenters, that locate within the corresponding cell. The seismic *events* are a sample from a seismic catalog for the region, which usually lists the time, hypocentral location, and

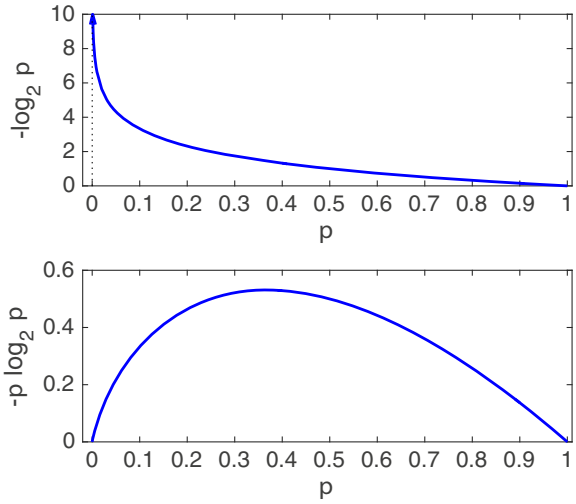

Fig. 1. Top: information, $I = -\log_2 p$ as a function of p; the arrowhead indicates that information is infinite for $p = 0$. Bottom: contribution to entropy of each term of (2) as a function of p; the function has a maximum for $p = 0.3679$, and is zero for both $p = 0$ and $p = 1$

magnitude of events. A sample consists of events occurring within given space, time, and magnitude windows. Here, we are concerned with the spatial distribution of events in a given time window; these events constitute a point process in space, and the problem is to relate this spatial distribution to a probability one, so that the system entropy can be evaluated.

As will be shown below, the measured entropy for a given sample will depend on the number of cells (or, conversely, on their size), so that there will be different values of entropy for any given sample, and the problem arises of which entropy value, or values, to use in order to characterize the system. We will come back to this problem later on.

1.3 The incidence entropy

It is common practice (e.g., Nicholson et al. 2000; Goltz and Böse 2002) to consider that the probabilities associated with the cells are

$$p_j^I = \frac{n_j}{N} \tag{4}$$

where n_j is the number of events in the jth cell and N is the total number of earthquakes in the sample. Borel's Law of Large Numbers tells us that (4) will yield the true probability for $N \to \infty$, but for real seismic catalogs and for time windows short enough to be useful, the typical N values are not large enough to result in reliable

probability estimates. We will call these probabilities *incidence probabilities*, and the corresponding entropy will be the *incidence entropy* S^I:

$$S^I = -\sum_{j=1}^{K} p_j^I \log_2 p_j^I \tag{5}$$

According to (4), an empty cell has probability $p_j^I=0$, and contributes nothing to the entropy. Thus, the summation in (5) can be restricted to the occuppied cells.

As K grows, the cells get smaller, so that, on the average, the contents of each cell diminish and the information of non-empty cells grows so that the incidence entropy increases.

However, for a finite number of events, the incidence entropy cannot increase indefinitely, because when the cells are small enough there will be only one event per cell (assuming that no locations are exactly alike), and the incidence entropy will have a maximum value of

$$S^I_{\max} \equiv S^I_X = -\sum_{j=1}^{K} \frac{1}{N}\log_2\frac{1}{N} = \log_2 N, \tag{6}$$

where m ranges over the indices of the N occupied cells ($m = j \quad \forall \quad n_j > 0$). If events are spaced more or less regularly, S^I_x will be attained soon after $K \geq N$ (we will show some examples of this), but if some events are extremely close together then some cells may contain more than one event and S^I will tend asymptotically to S^I_x.

1.4 The uniform entropy

If we suppose that the entropy of an observed epicentral (or hypocentral) distribution gives information about the system, then it should be compared with the entropy corresponding to the null hypothesis, which is the hypothesis that events are distributed randomly with uniform probability in space.

For K cells with uniform probability

$$p_j^U = \frac{1}{K}, \quad \forall j, \tag{7}$$

and the uniform entropy will be

$$S^U = -\sum_{j=1}^{K} \frac{1}{K}\log_2\frac{1}{K} = \log_2 K, \tag{8}$$

so that as K grows the uniform entropy increases indefinitely. It is well-known that for any given number of cells and events, the uniform distribution is the distribution having the maximum entropy (Shannon 1948).

Note that, while S^U depends only on K (there is an implicit assumption of an infinite number of events so that there is always an equal number of events in each and every cell), S^I depends on both N and K but its limiting value depends solely on N.

1.5 The Poisson entropy

Here we propose a measure of entropy based on how the number and distribution of events differ from the null hypothesis of a uniform distribution: a measure that can complement the information from S^I and S^U.

For a total N events, for K cells the average number of events per cell is

$$\lambda = \frac{N}{K}; \tag{9}$$

if the events were distributed uniformly throughout the region, the probability of finding n_j events in the jth cell would be given by the well-known Poisson distribution (Figure 2)

$$\Pr(n_j) = \frac{\lambda^{n_j} e^{-\lambda}}{n_j!} = P_j^P; \tag{10}$$

Thus, for a given K, after determining the n_j, the probability for each cell will be computed according to (10) instead of to (4) or (7). In order to consider relative sizes among the probabilities and to satisfy (1), the probability in each cell is renormalized as

$$p_j^P = \frac{P_j^P}{\sum_{j=1}^{K} P_j^P} \tag{11}$$

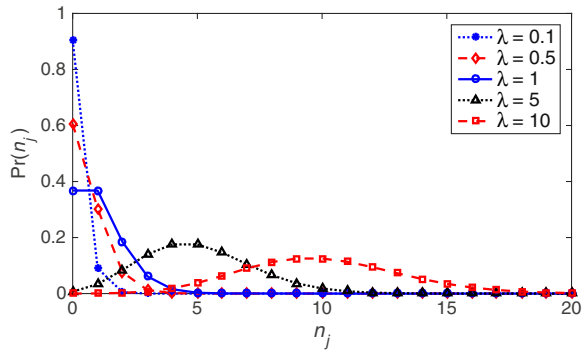

Fig. 2. Some examples of Poisson probabilities for various values of λ

we will refer to these probabilities as *Poisson renormalized probabilities*.

The *Poisson entropy* is calculated from (11) as

$$S^P = -\sum_{j=1}^{K} p_j^P \log_2 p_j^P \tag{12}$$

For $K < N$, empty cells have low probability and the highest probabilities are for cells with $n_j \approx \lambda$, while for $K > N$, empty cells have the highest probabilities and large n_j values have low probabilities (Figure 2), which is a behavior opposite to that of the incidence probabilities. Note that Poisson probabilities are always non-zero and less than one, so that both empty and occupied cells contribute to the Poissonian entropy.

For $K > > N$ $K \gg N$ i.e., extremely small λ, unoccupied cells will have probabilities $p_j^P \cong 1$, while occupied cells will have probabilities $p_j^P \cong 0$; and if K is large enough so that there is no more than one event in any occupied cell, then $p_j^P \cong 1/(K-N)$ for the $K - N$ unoccupied cells and close to zero for the N occupied cells. Hence, for very large K,

$$S^P = -\sum_m \frac{1}{K-N} \log_2 \frac{1}{K-N} = \log_2(K-N), \tag{13}$$

where m ranges over the $K - N$ indices of unoccupied cells $m = j \quad \forall \; n_j \quad > 0)$), so that when $\mathrm{K} \rightarrow \infty$, $\mathrm{S^p} \rightarrow \mathrm{S^u}$.

After we began working with the Poisson entropy, we found out that this name had already been used (in a different context) in signal processing (Adell et al. 2010), but, to our knowledge, the concept has not been used in seismology, nor renormalized, nor in conjuction with the uniform entropy as will be described below.

2 Entropy estimation and conventions

Let the study region have lengths X and Y if working in 2D, or X, Y, and Z if in 3D, and let each length be divided into k segments so that there will be $K = k^3$ cells, each cell with area $a = XY/k^2$ or volume$v = XYZ/k^3$, respectively.

We compute the abovementioned entropies for k ranging from 2 (the smallest meaningful value) to $k = \lceil \sqrt{3N} \rceil$ or$k = \lceil \sqrt[3]{3N} \rceil$, so that the largest value of K is $\approx 3N$ and the effects of saturation can be clearly shown.

We will now illustrate the calculation of the entropies described above, using a sample of 1000 events occuring just before the M_W = 9.1 great Tohoku earthquake occurred in 2011 in the Honshu area of Japan. The particulars of the space-time window containing the sample will be discussed below in the section about application. We will illustrate the process for 2D, because it is easy to visualize only in two dimensions, and only a few steps in the process will be shown for reasons of space.

Figure 3 shows the epicenters constituting the data, the cell grids, and the p_j^P and p_j^I probabilities for K = 25 cells and λ = 40.0 events/cell; note the completely different distribution of values for the different probabilities.

Figures 4, 5, and 6 show the grids and the cell probability matrices for$K$ = 100, 225, and 1024 cells, respectively; the last one is the closest approximation to λ = 1.0, and illustrates the result that for K of the order or greater than N, the Poisson and incidence probabilities appear as a negative photograph of each other, i.e. values for one are high where the values of the other are low, and viceversa.

The entropy values obtained from the calculations for all K values are shown in Figure 7, where, for large K, it is clear that $S^P \rightarrow S^U$ while S^I tends to the abovementioned limit. This difference between S^I and S^U begins quite early, for $K \sim N/4$, and could be interpreted as giving information about differences between the uniform and the observed distributions but, as will be seen below, it is largely due to saturation effects.

Since direct plotting of S vs K does not allow to see the details of what happens at low K values (high λ), it is better to use$\log_{10}K$; and, in order to be able to compare entropies from samples having different numbers of events, we will normalize K by the number of events, N, and S by the S_x^I limit, and plot S/S_x^I vs. $\log_{10}(K/N)$as shown in Figure 8, where the differences between $S^U S^I$, and S^P can be clearly appreciated. It is evident that for small K/N, S^I does not differ very much from S^U; i.e., S^I is not very sensitive for small K/N and departs from S^U quite early and tends to its saturation value as K increases. It remains to be seen whether this departure from S^Uconveys useful information.

3 Null hypothesis and other distributions

It is essential to be able to determine whether the measured entropies are not due to chance variations of an

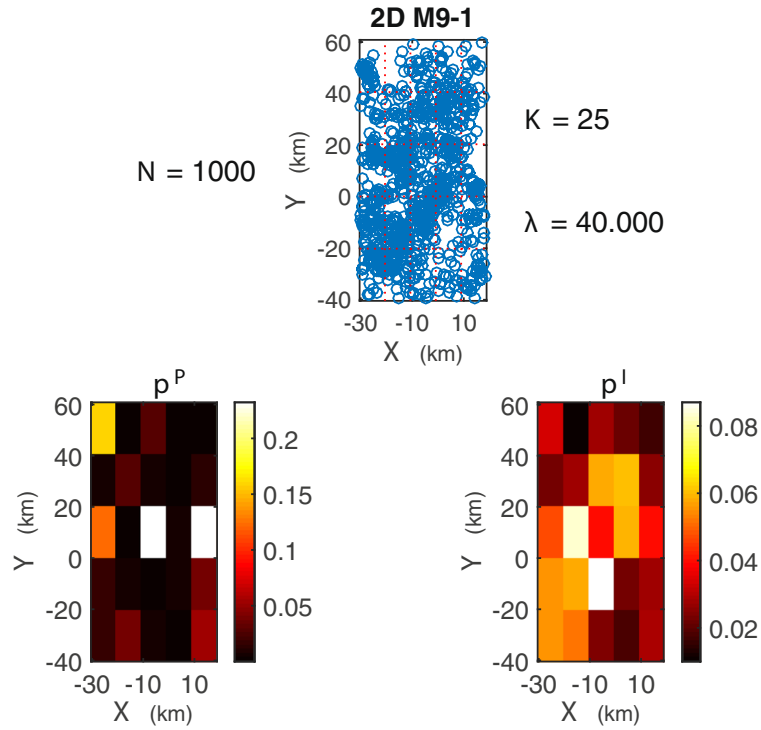

Fig. 3. Epicenters of 1000 events occuring just before the M_W = 9.1 great Tohoku earthquake occurred in 2011 in the Honshu area of Japan and grid for $K = 25$ cells and $\lambda = 40.0$ events/cell (top), and color-coded p_j^P and p_j^I probabilities (bottom)

otherwise informationless uniform spatial distribution. In order to do this, we made tests on what we call null hypothesis H0 distributions and on regularly spaced distributions. H0 distributions are synthetic distributions

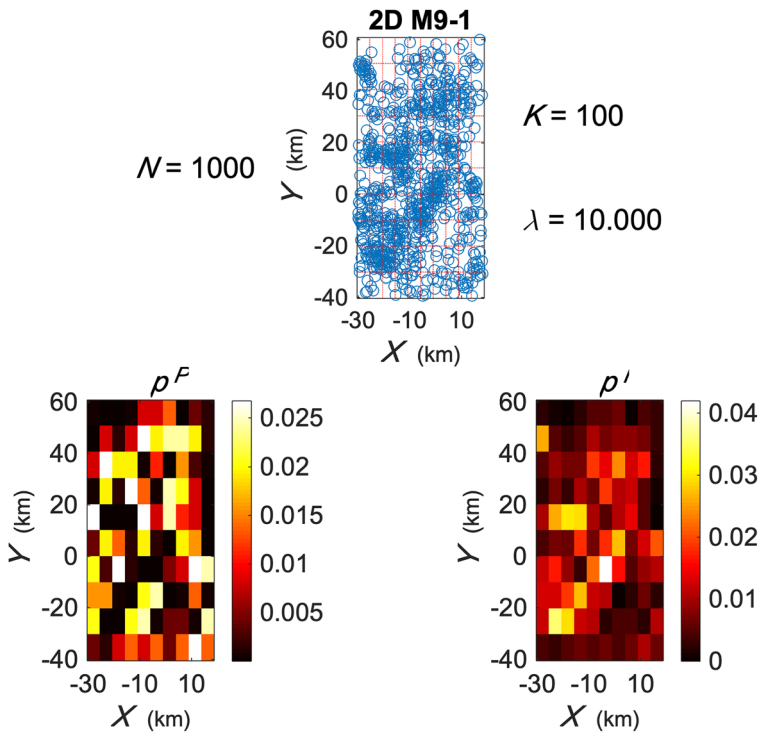

Fig. 4. Epicenters of 1000 events occuring just before the M_W = 9.1 great Tohoku earthquake occurred in 2011 in the Honshu area of Japan and grid for $K = 100$ cells and $\lambda = 10$ events/cell (top), and color-coded p_j^P and p_j^I probabilities (bottom)

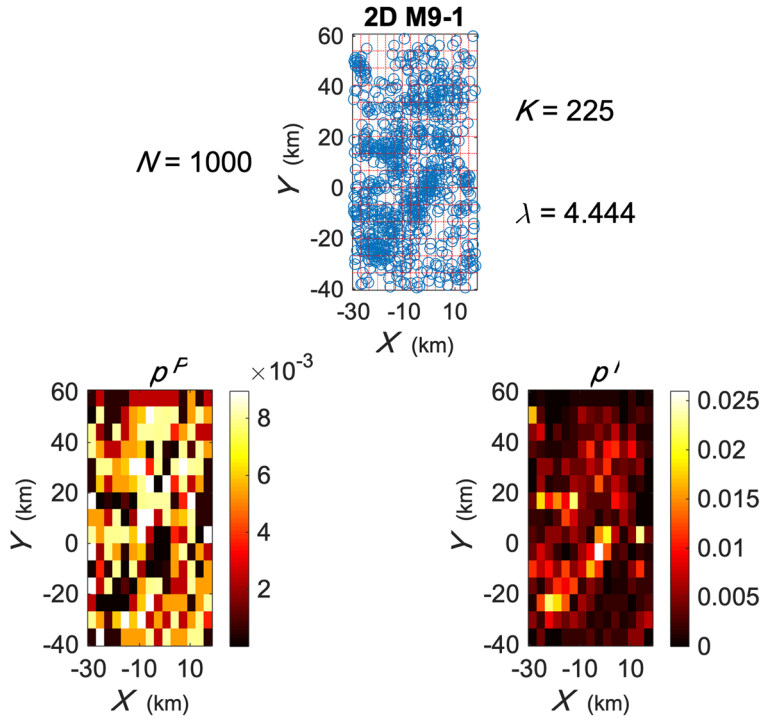

Fig. 5. Epicenters of 1000 events occuring just before the M_W = 9.1 great Tohoku earthquake occurred in 2011 in the Honshu area of Japan. Epicenters and grid (above) and color-coded p_j^P and p_j^I probabilities (below), for K = 225 cells and λ = 4.44 events/cell

having the same spatial dimensions and the same number of events as those of the observed region, but events are randomly distributed with uniform probability over the whole space.

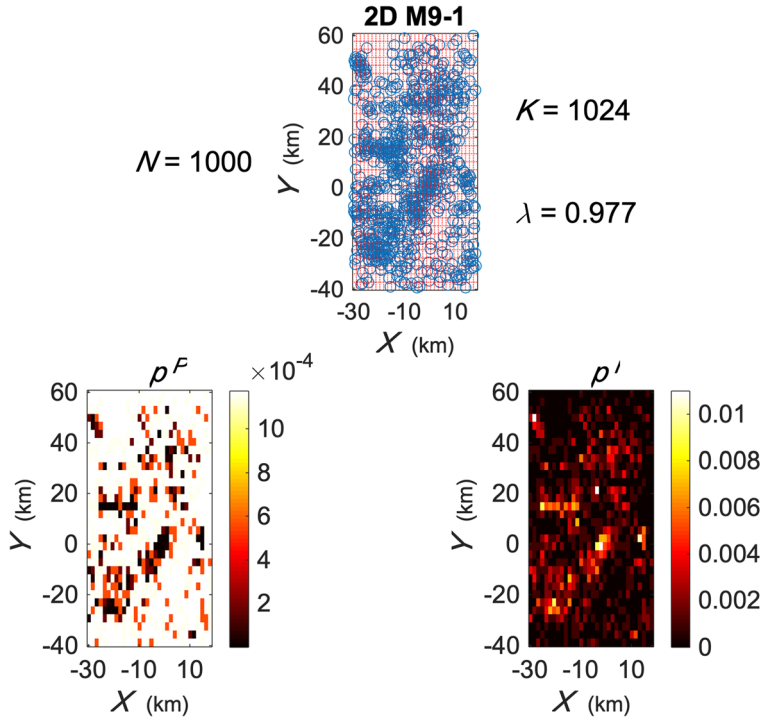

Fig. 6. Epicenters of 1000 events occuring just before the M_W = 9.1 great Tohoku earthquake occurred in 2011 in the Honshu area of Japan. Epicenters and grid (top) and color-coded p_j^P and p_j^I probabilities (bottom), for K = 1024 cells and λ = 0.977 events/cell. This is the closest instance to λ = 1 in the process

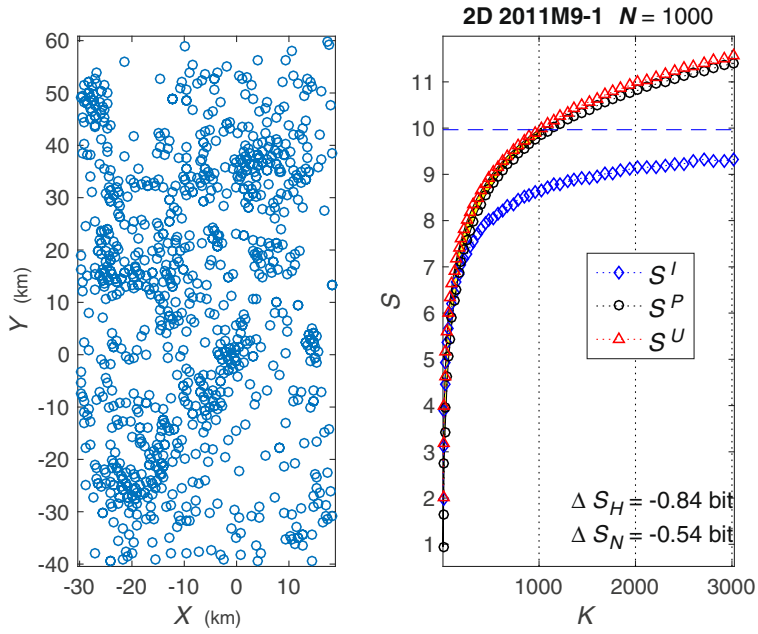

Fig. 7. Epicentral distribution of 1000 earthquakes occurred just before the Tohoku earthquake (left), and the corresponding 2D entropies vs K (right); details in the right panel will be discussed in Figure 8

Figure 9 illustrates an H0 distribution corresponding to the observed one in Figure 8, and the results for our various entropy measures, results that are typical for all H0 distributions. These results are very important, because they show that if the studied spatial distribution is indeed uniform, S^P recognizes it as such and coincides (within numerical error bounds) all the way with S^U. Thus, differences between S^P and S^U do measure departures from the null hypothesis of the uniform distribution.

On the other hand, Figure 9 shows again that S^I deviates from S^U quite soon, and it should not, because the distribution is indeed uniform; for $\log_{10}(K/N)$ larger than approximately −0.6 ($K \sim 0.25N$), S^I is measuring

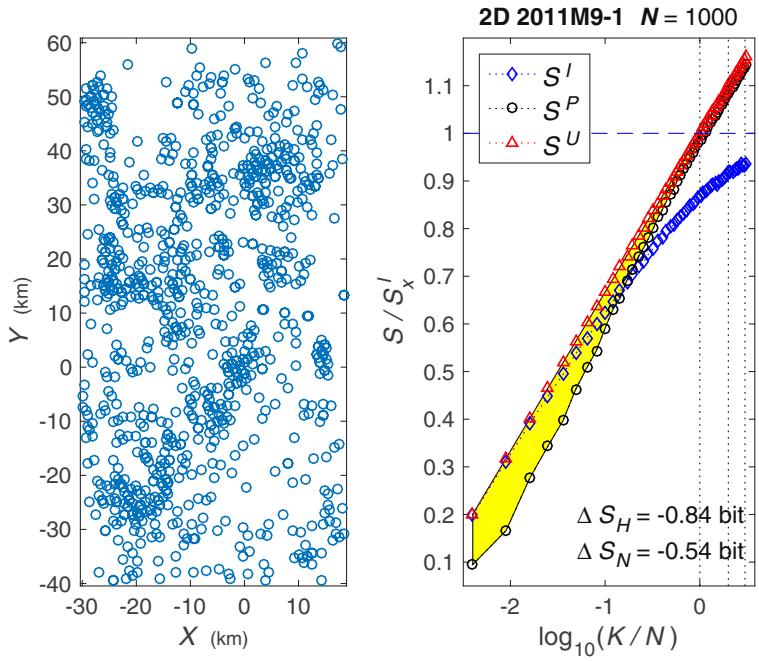

Fig. 8. Epicentral distribution of 1000 earthquakes occurred just before the Tohoku earthquake (left) and the corresponding 2D entropies normalized as S/S_x^I vs $\log_{10}(K/N)$; the vertical lines are for reference and correspond to $K = N$ and $K = 2N$; the shaded (yellow) area indicates the difference between S^U and S^P; ΔS_H and ΔS_N are the areas defined in (15) and (16)

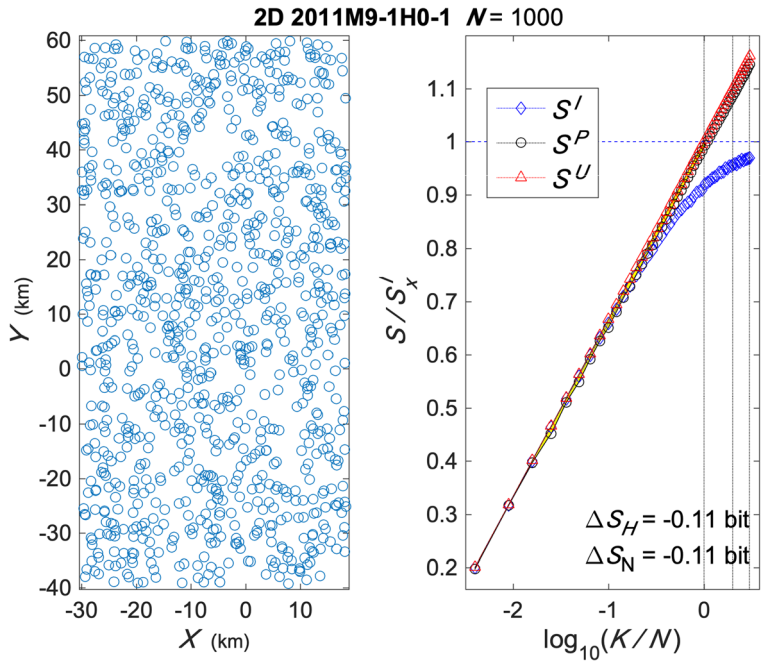

Fig. 9. Example of a null hypothesis uniform distribution corresponding to the actual distribution in Figures 3 to 9 (left), and the corresponding 2D entropies; conventions as in Figure 8

low entropy, not because it is low, but because the measure saturates. Let this be a caveat for people using the incidence entropy for characterizing seismicity distributions.

Figure 10 shows an example of the measured entropies for a regularly spaced rectangular 2D distribution. It is clear that all measured entropies agree with the theoretical uniform one for small K/N, while for slightly larger K/N values S^I starts showing saturation and then reaches S_x^I and stays there, while S^P correctly identifies the distribution as uniform and agrees with S^U for all K/N.

For square grids (not shown here), S^Icoincides with S^U all the way to $K/N = 1$ where $S^I = S_x^I$ and then stays at this maximum value. This indicates that the

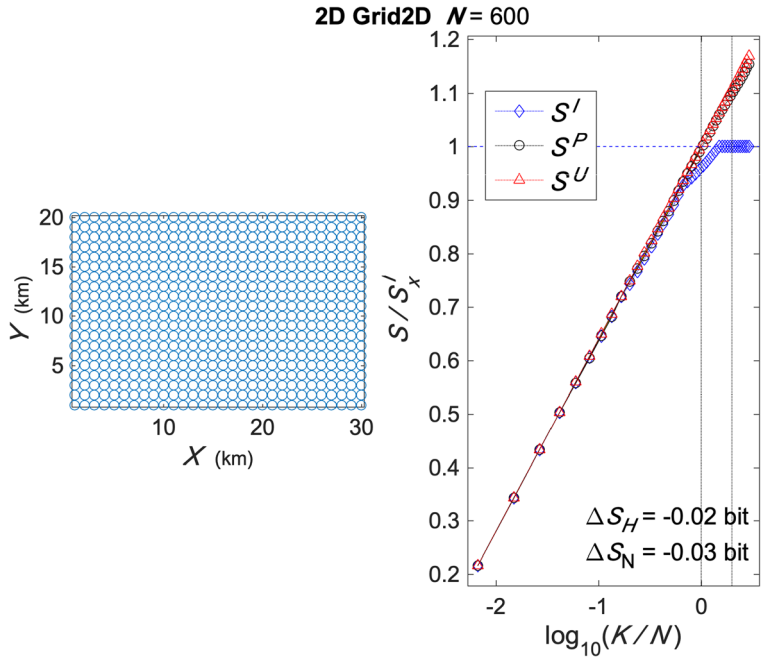

Fig. 10. Example of entropy measures for a rectangular evenly spaced grid distribution; conventions as in Figure 8

departures of S^I from S^U shown in Figures 7 to 10 is largely due to the differences in length of the sides of the studied region and do not reflect actual departures from the entropy of a uniform distribution.

In all cases, S^I is less sensitive than S^P and saturates quite early, so that we will use only the Poisson entropy, S^P, to characterize the observed epicentral or hypocentral distributions, and use for reference purposes only.

4 Cells and measure

Working with cells sets limitations on the number of possible entropy observations, because the number of partitions, k, on each side of the study region has to be integer. $k = 1$ gives no information at all, so that $k = 2$ is the smallest meaningful number of partitions. On the other hand, as illustrated above and below, for k such that K is larger than N, S^I $S^I$is saturated and $S^P$is closely approaching S^U, so that for $K \geq N$ the difference between these two last measures does not carry significant information.

Hence, in what follows, we will plot entropies from $k = 2$ to $k = \lceil \sqrt{3N} \rceil$ or $k = \lceil \sqrt[3]{3N} \rceil$ in order to clearly show the abovementioned saturations.

Now we address the problem of quantifying the information given by S^P. The entropy values change with the number of cells, K, and there is no a priori number to be used for reference; however, for each number of cells, there is a natural reference entropy measure: the uniform one, S^U, that depends only on the number of cells. Hence, we will measure the information given by S^P by comparing it withS^U; the measure will be related to the difference between these entropies, shown in yellow in the preceding figures, for all the useful K range.

The difference between $S^U$and S^Pshown in yellow in the figures, corresponds to entropy values for K ranging from

$$K_1 = 2^2 \text{or} K_1 = 2^3$$

corresponding to index $j = 1$, to

$$K_{\max} = \left\lceil \sqrt{N} \right\rceil^2 \text{ or } K_{\max} = \left\lceil \sqrt[3]{N} \right\rceil^3 \tag{14}$$

for which $K_{\max} \approx N$, corresponding to index$j = j_N$.

One measure is the average difference over the whole range

$$\Delta S_N = \frac{1}{j_N} \sum_{j=1}^{j_N} \left(S_j^P - S_j^U \right) \tag{15}$$

but, due to the fact that differences tend to get smaller as j increases, we will also use the more sensitive measure of the average over the first half of the j range:

$$\Delta S_H = \frac{1}{j_H} \sum_{j=1}^{j_H} \left(S_j^P - S_j^U \right) \tag{16}$$

where j_H is the index for which K is closest to $N/2$.

Let us remember that each entropy, besides being a measure of the orderliness of the system, is the expected surprise information at the corresponding K; thus, a Poisson renormalized entropy smaller than the corresponding uniform entropy, i.e., $\Delta S < 0$, besides identifying a more orderly distribution, is a measure in bits of the decrease in the expected surprise information of the system.

5 Application example

As an example of application, we will apply the measure to the seismicity associated with the $M_W = 9.1$ great Tohoku (or Tohoku-Oki) earthquake occurred in 2011 in the Honshu-Hokkaido region in Japan. We chose that region because of its high seismicity, excellent seismographic coverage, and reliable hypocentral locations. For reasons of space, we will not delve in this work into the complicated tectonic setting that causes the high seismicity of Japan; a succint coverage of the theme can be found in the NUMO-TR-04-04 report (NUMO 2004).

A working catalog, containing 632,777 $M \geq 1.5$ earthquakes, ocurred between 1960 and 2016 was made from data reported in the catalogs published by the Japan Metereological Agency (JMA) and the International Seismological Centre (ISC), taking care to include all events while eliminating repeated ones. When magnitudes did not agree, those reported by the JMA were used; actually, magnitudes were used only for determining catalog completeness, but do not play an explicit role in entropy determinations.

The method was illustrated above with a 2D epicentral analysis of one time window (we will return later to 2D analyses); now the application examples will be 3D hypocentral analyses, as appropriate for seismicity distributed within a volume. The relationship between the 3D and 2D analyses will be discussed later.

The idea behind this analysis is to see whether the entropy in a volume located around the hypocenter of a large earthquake changes as the time to the earthquake occurrence approaches, so that the entropy measures could be used as a forecasting factor.

The data used for the analysis are the hypocenters within a 3D spatial window around the hypocenter of a main event (for retrospective analysis) or a spatial window that represents a target area; the window is chosen small enough so that the processes of organization may be considered to affect all events within it, but large enough to have enough data to work with. For the seismicity within the spatial window, several time windows are chosen so that they represent different stages in the earthquake preparation process while containing enough events for entropy estimation. Different sizes and locations of the space window can be tried and the one resulting in the most sensitive and stable results chosen for the final analysis.

5.1 The March 11, 2011, $M_W = 9.1$ earthquake

This megaearthquake, known as the Great Tohoku (or Tohoku-Oki) earthquake, which we will refer to as M9, was a relatively shallow, 29.0-km-deep (USGS 2011), thrust earthquake, which caused a large number of fatalities, extensive damage, and a large tsunami that also caused further damage that included a meltdown at the Fukushima nuclear power plant. This earthquake has been amply studied (e.g., Iinuma et al. 2011; Ito et al. 2011; Lee et al. 2011; Meng et al. 2011; Suzuki et al. 2011; Hayes et al. 2017), so that we will not go into the process or consequences of the earthquake, but will only mention references that are relevant to the application of our proposed method.

This earthquake was chosen because it seemed an ideal example of stress accumulation, since no earthquakes larger than $M \geq 7.4$ had ocurred in the region since the September 25, 2003, $M_W = 8.0$ Hokkaido earthquake that occurred some 401km NNE of M9; none of the other $M \geq 7.5$ earthquakes that occurred before in the region (Figure 11) were close enough to the hypocenter of M9 to substantially decrease the accumulated stress in the hypocentral region.

Figure 11 shows the seismicity for $M \geq 2.5$ in the Honshu-Hokkaido region of Japan from 1960 to 2016; (yellow) stars indicate earthquakes with $M \geq 7.5$, with symbol sizes proportional to the magnitude.

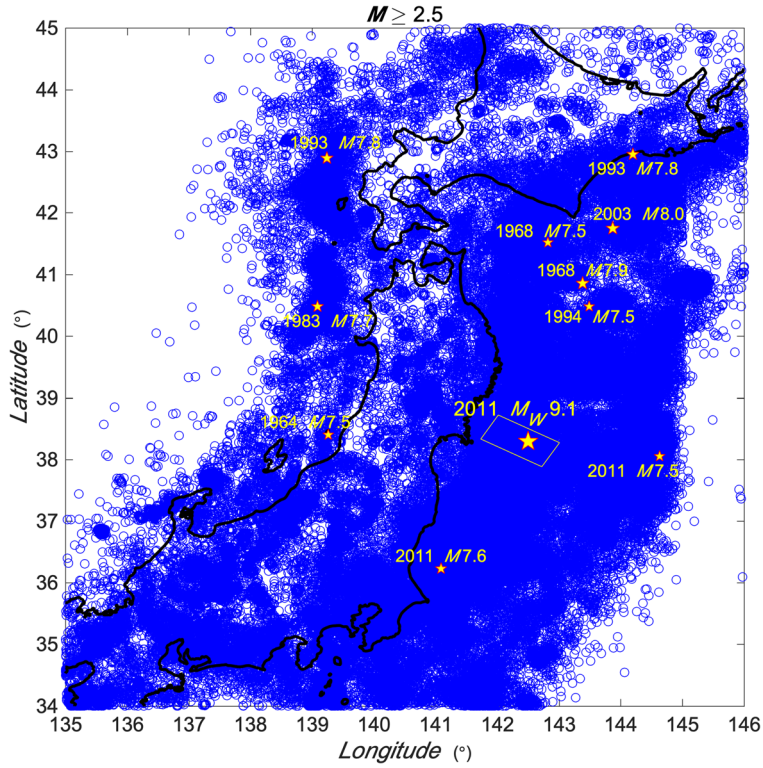

Fig. 11. Circles and stars represent the seismicity for $M \geq 2.5$ in the Honshu-Hokkaido region of Japan from 1960 to 2016; stars indicate earthquakes with $M \geq 7.5$, with sizes proportional to the magnitude. The study area for the earthquake used as example, which is a rectangle when plotted in kilometers, is shown as a polygon

5.2 Space and time windows

For all events in the hypocentral region of the main event to be studied, M9 in this case, longitudes and latitudes are transformed to distances (in km) referred to the hypocenter of the main event, and then the X and Y coordinate axes are rotated so that the chosen volumes are rectangular boxes having their vertical sides parallel or perpendicular to the local direction of the trench. Figure 12 shows how the study rectangular area was chosen so as to try to avoid large lacunae in the epicentral distribution. See Figure 11 for the horizontal projection of the study volume in geographic coordinates.

Next, the cumulative number of earthquakes in the volume, N_c, is plotted vs. time, and the minimum magnitude threshold is gradually increased until the long-term slopes in the cumulative have more or less the same slope, which indicates sampling homogeneity over time (for low thresholds, slopes increase with time, reflecting monitoring improvements). We achieved reasonable homogeneity for threshold magnitude M = 2.3 (Figure 13).

In order to see how the entropy changes with time, time windows are chosen, each window extends backwards in time until it contains a given number of earthquakes, N, so that results from each window can be compaired easily; we used N = 1,000. The first window begins just before the occurrence of the main event, and each following window begins a given time, Δt, before the beginning of the previous one; in this example, we used time shifts Δt = − 2yr. The time windows we chose for this example are shown in Figure 13, where vertical continuous lines indicate the beginning of each window and the number right to the left of it identifies the window; succesive vertical dashed lines indicate where windows end. Note that the two last windows (7 and 8) have much larger durations than the other windows, which means that small magnitude coverage is not as good before about year 2000.

5.3 Poisson renormalized entropy 3D analysis

The events in each window file were analyzed as described above; for reasons of space, only the analyses for windows 1, 2, 4, and 6 are printed in Figures 14, Figures 15 16, 17 and 18; the figures showing the analyses for the other windows are presented as Online Resources 1 to 4 (Files ESM_1.pdf, ESM-2.pdf,

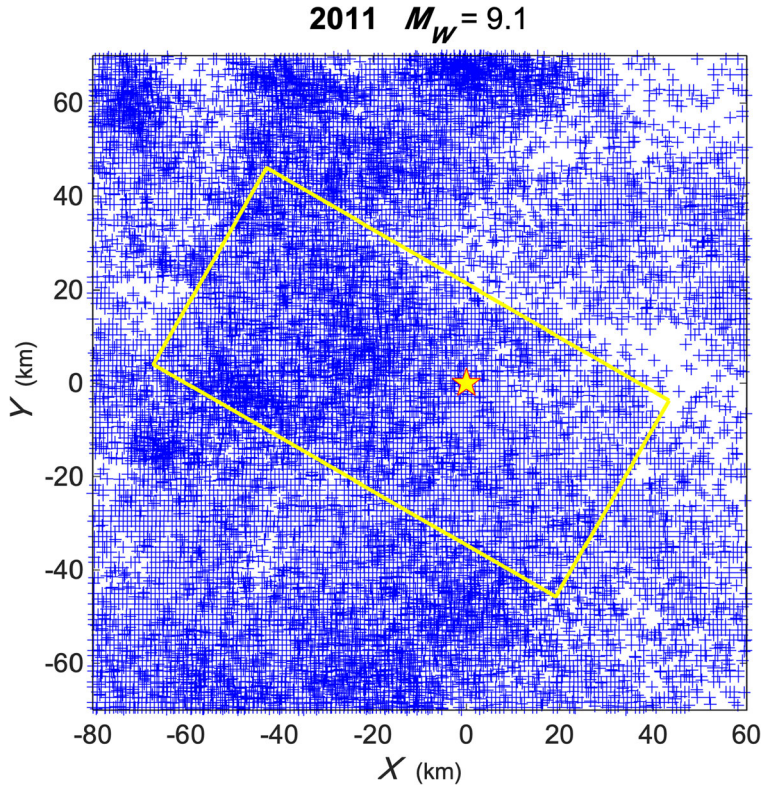

Fig. 12. Study region (rectangle) and M9 epicenter (star); crosses represent epicenters

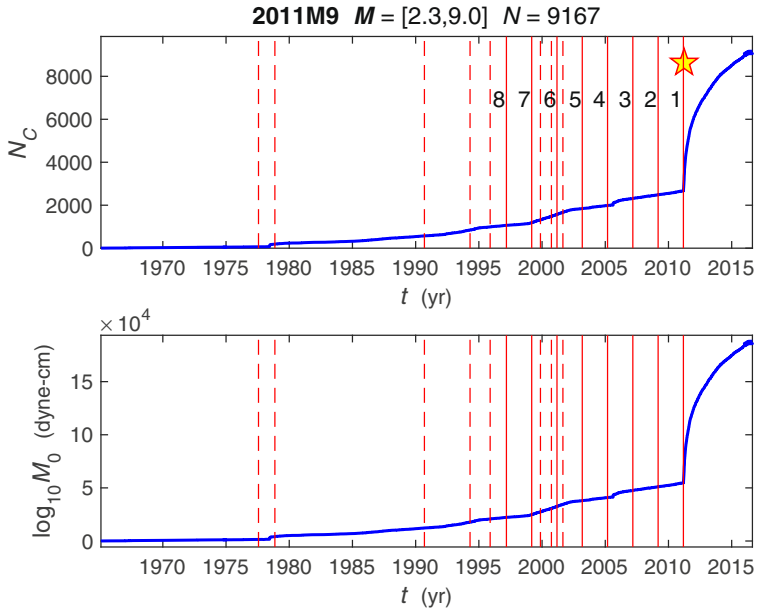

Fig. 13. Cumulative number of earthquakes in the study region vs. time (top) and the corresponding cumulative seismic moment vs. time (bottom). Both graphs show the time windows before the 2011M9 earthquake (star); each window starts at the time of the vertical continuous line immediately to the right of the window number, and extends backwards in time until it contains 1,000 events (dashed vertical lines)

ESM_3.pdf, and ESM_4.pdf) and folllow the same conventions as in Figure 8.

Figure 15 shows the analysis of the H0 distribution corresponding to the observed one in the first window, and it is clear when compared to the analyses of the observed distribution that S^P can correctly identify a uniform distribution, while S^Isaturates.

The results of the analysis are summarized in Table 1 and Figure 19. In this figure, the extension of each horizontal thick line indicates the duration of the corresponding time window and its vertical position indicates the value of ΔS_H; the vertical dashed lines indicate the middle time of each window, and the numbers above them identify each window; the inverted triangles joined

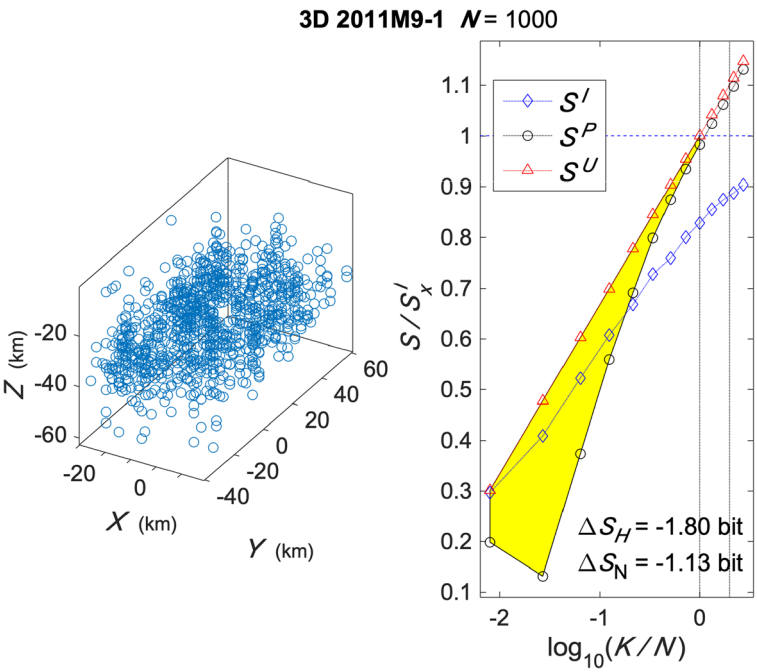

Fig. 14. 2011M9 Window no. 1. On the left is shown a view of the hypocentral spatial distribution as seen from 80° azimuth and 35° elevation; on the right, the corresponding entropy analysis described above is shown for the three entropies. The (yellow) shaded area indicates the difference between S^U andS^P; conventions as in Figure 8

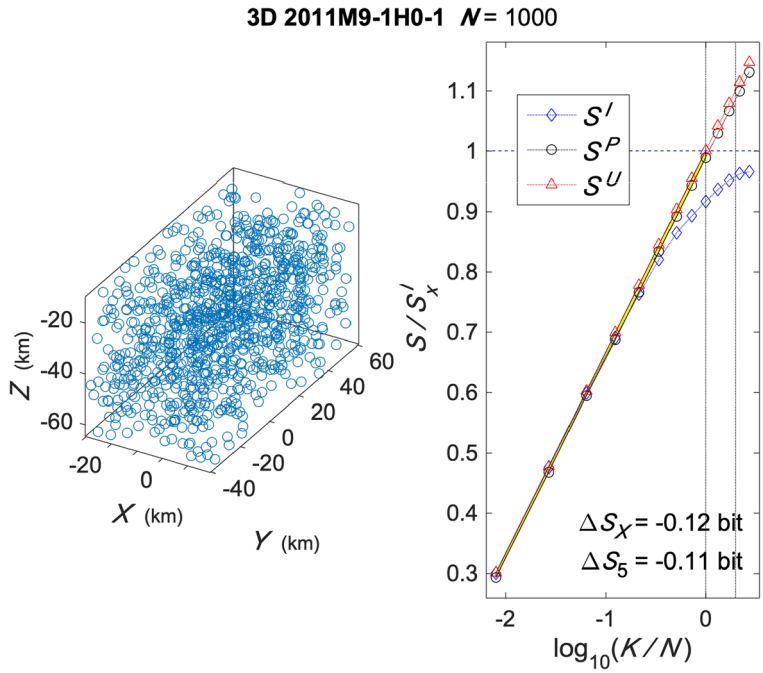

Fig. 15 2011M9 window no. 1 H0 distribution. On the left is shown a view of the hypocentral spatial distribution as seen from 80° azimuth and 35° elevation; on the right, the corresponding entropy analysis described above is shown for the three entropies. The (yellow) shaded area indicates the difference between S^U andS^P; conventions as in Figure 8

to the horizontal lines by a thin vertical one indicates the corresponding ΔS_N entropy change. The occurrence time of the M9 main event is indicated by a thich dashed vertical line and a star.

Our results clearly show entropy decrease before the 2011M9 event; the Poisson renormalized entropy measure ΔS_H in window no. 7 was about 0.927 bit lower than that of the previous no. 6 window, and −1.029 bit lower than the previous, but less reliable, no. 8 window. with respect to the level it had from about twenty and a half to ten years before the M9 earthquake. The ΔS_H and ΔS_N

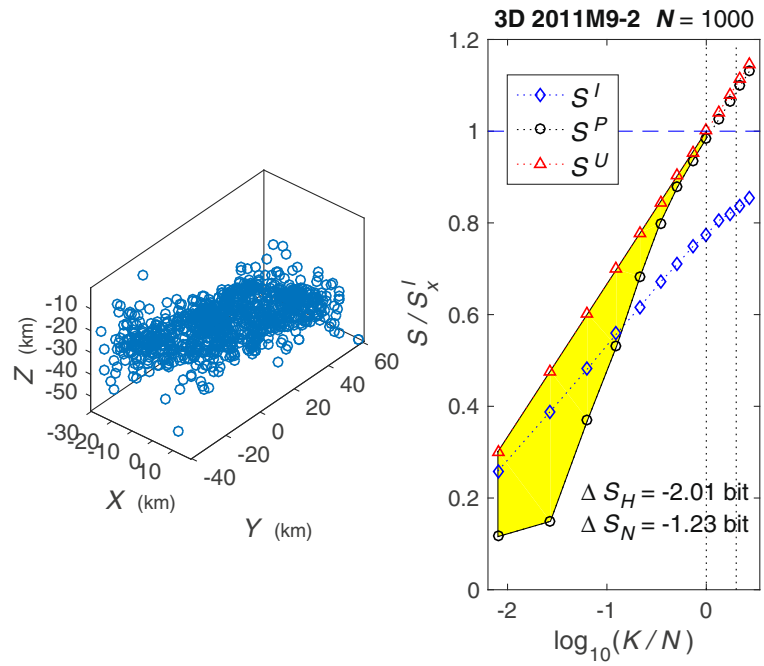

Fig. 16. 2011M9 window no. 2. On the left is shown a view of the hypocentral spatial distribution as seen from 80° azimuth and 35° elevation; on the right, the corresponding entropy analysis described above is shown for the three entropies. The (yellow) shaded area indicates the difference between S^U andS^P; conventions as in Figure 8

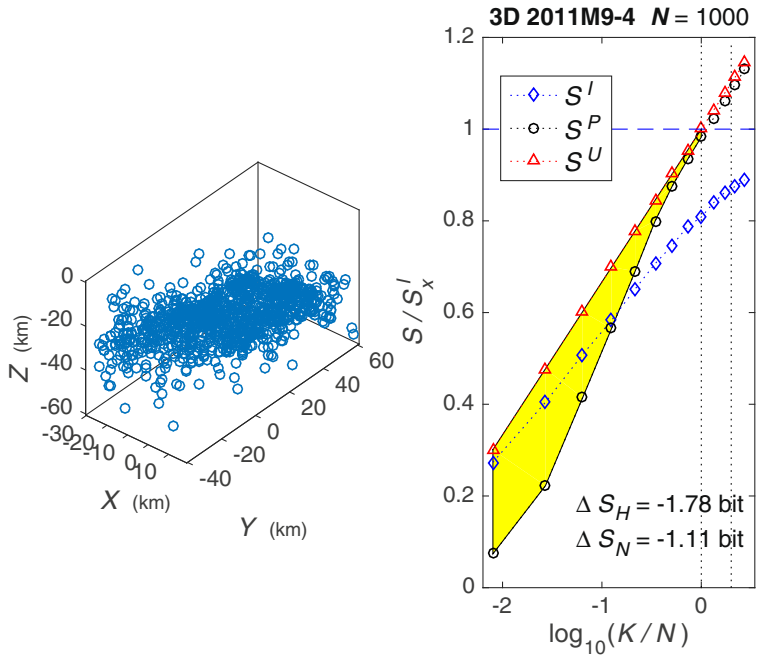

Fig. 17. 2011 M9 window no. 4. On the left is shown a view of the hypocentral spatial distribution as seen from 80° azimuth and 35° elevation; on the right, the corresponding entropy analysis described above is shown for the three entropies. The (yellow) shaded area indicates the difference between S^U andS^P; conventions as in Figure 8

values are just slightly higher for window no. 2 than for window no. 3, and are somewhat a higher value for window no. 7 than for window no. 8 (both of these not quite reliable), but the entropy decreases before the 2011M9 event is quite clear.

5.4 Poisson renormalized entropy 2D analysis

The 2D analysis for the first time window is shown in Figure 9 illustrating the method; for reasons of space, we will not show here the 2D analyses for the other

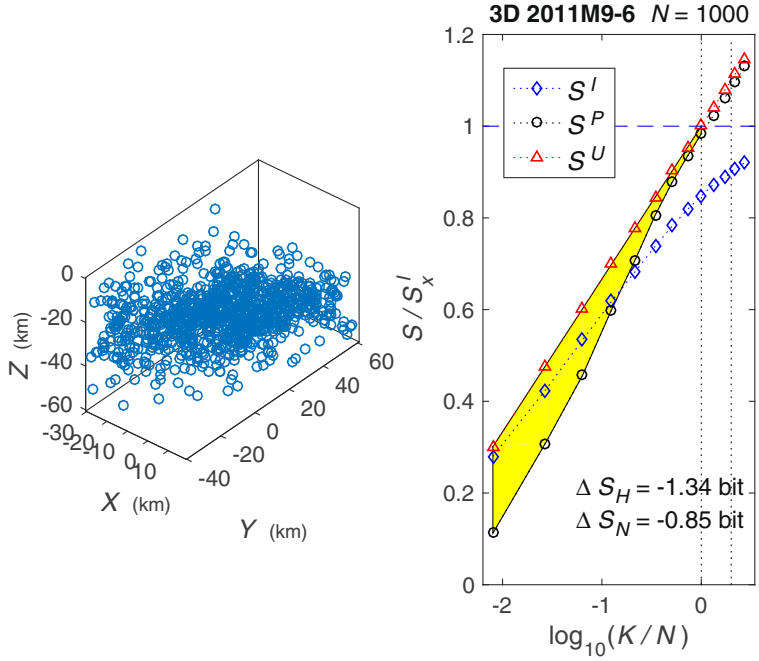

Fig. 18. 2011M9 window no. 6. On the left is shown a view of the hypocentral spatial distribution as seen from 80° azimuth and 35° elevation; on the right, the corresponding entropy analysis described above is shown for the three entropies. The (yellow) shaded area indicates the difference between S^U andS^P; conventions as in Figure 8

Table 1 Starting and ending times, and measured entropy differences for each time window

Window no.	t_{start},t_{end}	ΔS_H	ΔS_N
1	2001.65, 2011.18	−2.265	−1.378
2	2000.74, 2009.18	−2.009	−1.233
3	1999.87, 2007.18	−2.029	−1.245
4	1995.89, 2005.19	−1.777	−1.108
5	1994.30, 2003.19	−1.510	−0.955
6	1990.69, 2001.18	−1.338	−0.854
7	1978.86, 1999.19	−1.147	−0.738
8	1977.57, 1997.17	−1.236	−0.786

windows; they are available as electronic supplement Figures ESM_5 to ESM_11; here, we show only the results of the analysis in Figure 20.

This figure is very important, because it clearly shows that, although less sensitive than the 3D analysis, the 2D analysis also measures the entropy decrease in time before the occurrence of the 2011M9 earthquake; and its importance lies in that a serious limitation for the 3D analysis is the correct determination of hypocentral depths. Of course, a single example is not enough to guarantee that that 2D analysis will be effective in all cases, but it suggests that, if confirmed, Poissonian entropy analysis could be applicable to many catalogs with unreliable depth determinations.

6 Comparison with results from fractal dimension and G-R *b*-value

Since the entropy, the fractal dimension, and the *b*-value are supposed to be related to the stress level in a region (e.g., Scholtz 2015), we estimated these last two parameters for the data in our time windows. For reasons of space, the figures from the analyses of fractal dimensions and *b*-value will be relegated to the electronic supplement as Figures ESM_12.pdf and ESM_13.pdf, respectively.

6.1 Fractal dimensions

Fractal dimensions and entropy are related (Rényi 1959); fractal dimensions indicate organization and clustering, and have been observed to decrease before some large earthquakes (e.g., Goltz 1996, 1997; Rundle et al. 2003).

Fractal dimensions D_0, D_1, and D_2 were estimated for each of the space-time windows used for entropy estimations, using the correlation method of Grassberger and Procaccia (Grassberger and Procaccia 1983; Grassberger 2007) as described in Márquez et al. (2012), from

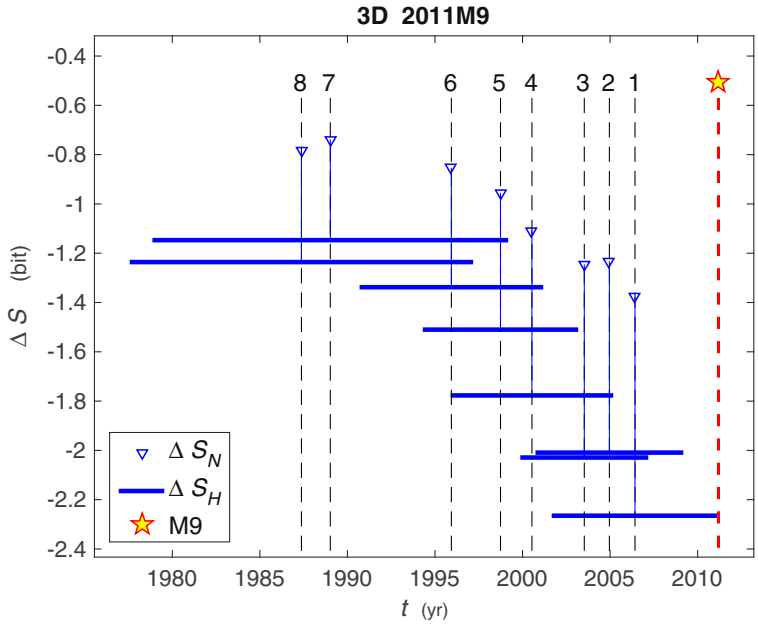

Fig. 19. 3D Poisson renormalized entropy measures vs. time before the 2011M9 earthquake. Thick (blue) horizontal lines indicate the duration of each time window identified by a dotted vertical line passing though its center and a number; the vertical position of each horizontal line indicates the ΔS_H value, while the corresponding ΔS_N value is shown as an inverted triangle. The occurrence time of the 2011M9 earthquake is shown by a thick, dashed vertical line and a star on top of it

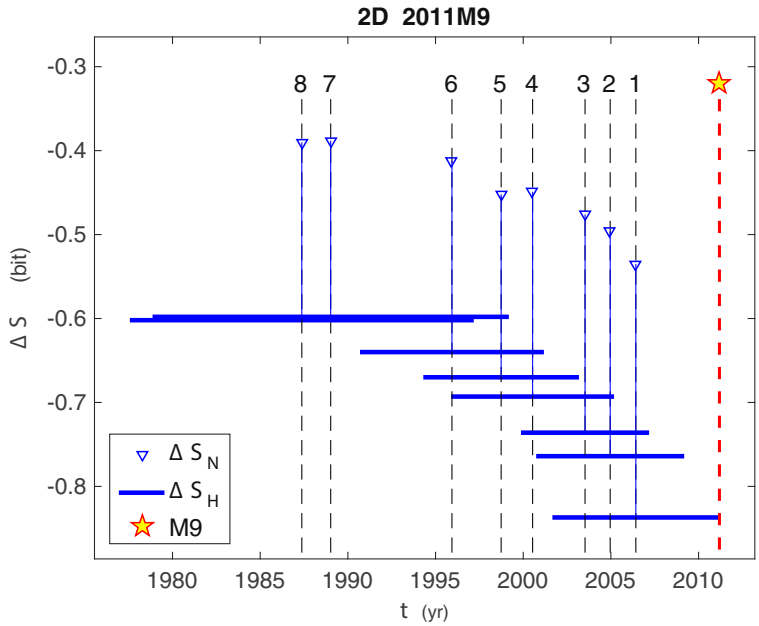

Fig. 20. 2D Poisson renormalized entropy measures vs. time before the 2011M9 earthquake. Thick (blue) horizontal lines indicate the duration of each time window identified by a dotted vertical line and a number; the vertical position of each horizontal line indicates the ΔS_H value, while the corresponding ΔS_N value is shown as an inverted triangle. The occurrence time of the 2011M9 earthquake is shown by a thick, dashed vertical line and a star on top of it

$$C_q(r) = \lim_{N\to\infty} \sum_{i=1}^{N} \left[\frac{1}{N-1} \sum_{j\neq i} H\left(r - |x_j - x_i|\right) \right]^{q-1} \propto r^{D_q}, D_q$$

$$= \frac{1}{q-1} \lim_{r\to 0} \frac{\log C_q(r)}{\log r},$$

and

$$D_1 = \lim_{r\to 0} \frac{\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \log \left[\frac{1}{N-1} \sum_{j\neq i} H\left(r - |x_j - x_i|\right) \right]}{\log r}.$$

Electronic supplement Figures ESM_12 and ESM_13 show the estimation of the fractal dimensions for the several time windows.

Figure 21 shows the values for the time windows for D_0, D_1, and D_2; D_2 behaves somewhat erratically, but the other two more reliable dimension estimators are seen to decrease in time and reach a minimum for the next-to-last time window before the main event. It may be speculated whether, if this behavior, different from those of the entropy and the *b*-value, is observed before other large earthquakes, could have precursory value.

6.2 G-R *b*-value

Several studies have reported *b*-value decreases before large earthquakes (e.g., Nanjo et al. 2012), and, as mentioned above, changes in the Gutemberg-Richter *b*-value are commonly interpreted as related to changes in the stress of state in the region (Aki 1981). De Santis et al. (2011) showed that the entropy of the magnitude distribution decreases when *b* increases. We calculated the most likely source *b*-value, b_x, for the time windows from the Aki-Utsu estimate

$$b_m = \frac{\log_{10} e}{\overline{M} - (M_c - 0.05)},$$

where $\overline{M}$ is the mean observed magnitude and is the completeness magnitude (rounded to 0.1), together with the number of observations and the range of linearity, using the technique described in Nava et al. (2018).

We present here the results for the first six time windows before M9 (Figure 22), because small magnitude coverage was not adequate for the earliest two windows. A figure showing the G-R distributions and the source *b* histograms can be found as electronic supplement Figure ESM_14.

Figure 22 shows the behavior of b_x before M9, the thick horizontal lines show the time window and the values of b_x, and the minimum and máximum limits for

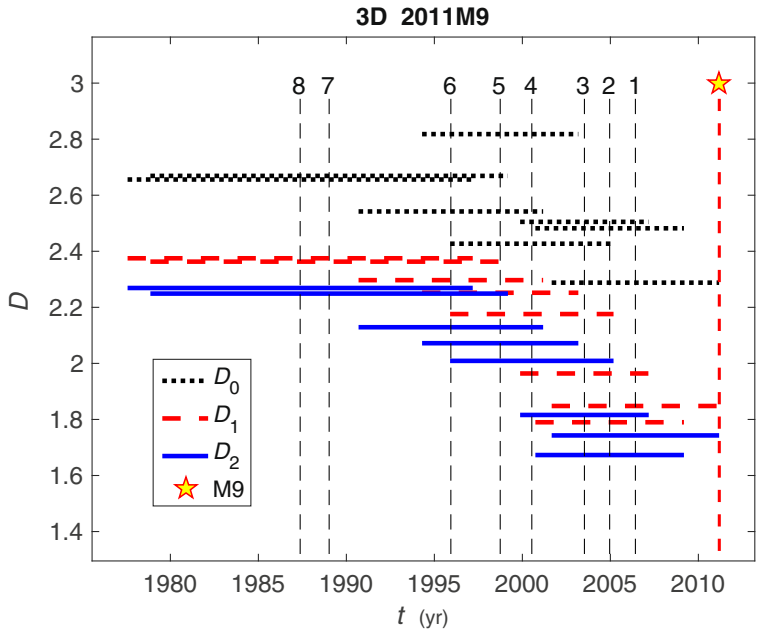

Fig. 21. Fractal dimensions vs. time. Thick, horizontal lines indicate the time span and dimension value for each time window; windows are identified by a vertical dotted line going through their middle and by the window number on top of the line. The occurrence time of the 2011M9 earthquake is shown by a thick, dashed vertical line and a star on top of it

90% confidence are shown as triangles and inverted triangles, respectively.

For our example, the b_x values do decrease as the time windows approach the time of the main event, indicated by a dashed thick line and a star; a behavior that completely agrees with that observed by Nanjo et al. (2012) and with that of our entropy estimates.

7 Discussion and conclusions

The proposed renormalized Poisson entropy is more sensitive and stable than the usual incidence entropy. It clearly identifies uniform distributions and does not saturate as incidence entropy does.

As shown by our example, the difference between uniform and renormalized Poisson entropy appears to be

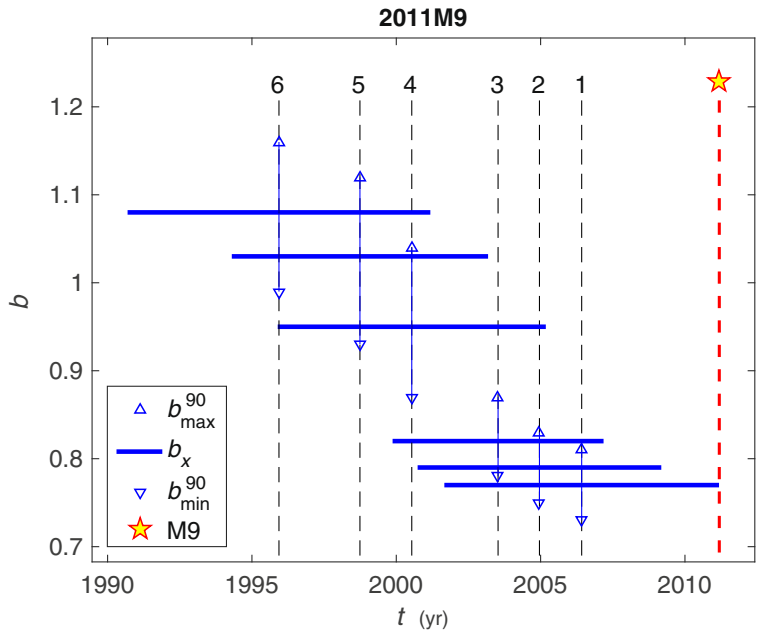

Fig. 22. Temporal variation of the G-R *b*-value for the time windows before M9. Each horizontal line represents the time window, and its vertical location indicates de corresponding most likely source *b* value, while the triangles joined by thin vertical lines represent the corresponding uncertainty. The occurrence time of the 2011M9 earthquake is shown by a thick, dashed vertical line and a star on top of it

sensitive to differences in the stress level in a region where a large earthquake is close to occurr. Hence, renormalized Poisson entropy appears to have precusory value and, together with other indicators of stress levels, such as fractal dimension or G-R *b*-value, could be used for earthquake forecasting. The renormalized Poisson entropy analysis requires less data than either fractal or *b*-value studies, so that its role in seismic hazard studies may be more than complementary.

The main limitations in using renormalized Poisson entropy are length and homogeneity of catalogs, which are also limitations for all other statistical methods. Another limitation is hypocentral location accuracy, but in regions where depth determinations are uncertain but epicentral accuracy is acceptable, 2D analysis can be used; 2D analysis is not as sensitive as 3D analysis but gives acceptable results.

We tried the idea of using the renormalized Poisson entropy to detect entropy changes before a large earthquake in the retrospective analysis presented here, and in some other retrospective analyses, not shown here, where it also gives good results; however, it needs to be tested through prospective analyses (e.g., Rhoades et al. 2011; Schorlemmer et al. 2018; Taroni et al. 2013) to assess its reliablity and ultimate usefulness.

Acknowledgements Our sincere thanks to the late Vicente Ferreira and to José Frez and Beatriz Cordero for useful comments, and to Matteo Taroni for constructive suggestions; we gratefully acknowledge technical support from José Mojarro and Humberto Benítez.

Funding This research was partially funded by Conacyt grant 222795 and by Conacyt scholarship 587073 (G.D).

References

Adamaki K, Roberts R (2016) Evidence of precursory patterns in aggregated time series. Bull Geol Soc Greece 50(3):1283–1292. https://doi.org/10.12681/bgsg.11834

Adamaki A, Roberts R (2017) Precursory activity before larger events in Greece revealed by aggregated seismicity data. Pure Appl Geophys 174:1331–1343. https://doi.org/10.1007/s00024-017-1465-6

Adell A, Lekuona A, Yu Y (2010) Sharp bounds for the entropy of the Poisson Law and related quantities. IEEE Trans Inf 56(5):2299–2306. https://doi.org/10.1109/TIT.2010.2044057

Aki K (1981) A probabilistic synthesis of precursory phenomena. In: Simpson DW, Richards PG (eds) M. Ewing seriesEarthquake Prediction, an International Review, vol 4. American Geophysical Union, Washington, D. C., pp 566–574

Al-Kindy F, Main I (2003) Testing self-organized criticality in the crust using entropy: a regionalized study of the CMT global earthquake catalogue. J Geophys Res 108(B11):2521. https://doi.org/10.1029/2002JB002230

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

Billio M, Casarin R, Costola M, Pasqualini A (2015) An entropy-based early warning indicator for systemic risk. SYRTO Working Paper Series, 12

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

Chen C, Chang C (2004) Comment on "Entropy, energy, and proximity to criticality in global earthquake populations" by Ian G. Main and Fahad H Al-Kindy. Geophys Res Lett 31. https://doi.org/10.1029/2003GL019008

De Santis A, Cianchini G, Favali P, Beranzoli P, Boschi E (2011) The Gutenberg-Richter law and entropy of earthquakes: two case studies in central Italy. Bull Seismol Soc Am 101:1386–1395. https://doi.org/10.1785/0120090390

Evison F (1999) On the existence of earthquake precursors. Ann Geofis 42:763–770

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

Goltz C (1997) Fractal and chaotic properties of earthquakes. Springer-Verlag, Lecture notes in Earth sciences 77. https://doi.org/10.1007/BFb0028315

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

Grassberger P, Procaccia I (1983) Measuring the strangeness of strange attractors. Physica D 9(1-2):189–208

Grassberger P (2007) Grassberger- Procaccia algorithm. Scholarpedia 2(5):3043, revision #91330. https://doi.org/10.4249/scholarpedia.3043

Hayes G, Myers E, Dewey J, Briggs R, Earle P, Benz H, Smoczyk G, Flamme H, Barnhart W, Gold R, Furlong K (2017) Tectonic summaries of magnitude 7 and greater earthquakes from 2000 to 2015. U.S. Geological Survey Open-File Report 2016–1192, 148 p. https://doi.org/10.3133/ofr20161192

Iinuma T, Ohzono M, Ohta Y, Miura S (2011) Coseismic slip distribution of the 2011 off the Pacific coast of Tohoku earthquake (M 9.0) estimated based on GPS data – was the asperity in Miyagi-oki ruptured? Earth Planets Space 63(24):643–648. https://doi.org/10.5047/eps.2011.06.013

Ito Y, Tsuji T, Osada Y, Kido M, Inazu D, Hayashi Y, Tsushima H, Hino R, Fujimoto H (2011) Frontal wedge deformation near the source region of the 2011 Tohoku-Oki earthquake. Geophys Res Lett 38:L00G05. https://doi.org/10.1029/2011GL048355

Lee S, Huang B, Ando M, Chiu H, Wang J (2011) Evidence of large scale repeating slip during the 2011 Tohoku-Oki earthquake. Geophys Res Lett 38:L19306. https://doi.org/10.1029/2011GL049580

Ling-ren Z (1988) Anomaly of information entropy before great earthquake the case study of Wuqia earthquakes. Seismol Geol 10:29–38

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

Main I, Al-Kindy F (2002) Entropy, energy, and proximity to criticality in global earthquake populations. Geophys Res Lett 29(7):1121. https://doi.org/10.1029/2001GL014078

Main I, Al-Kindy F (2004) Reply to "Comment on Entropy, energy, and proximity to criticality in global earthquake populations" by Chien-Chih Chen and Chun-Ling Chang. Geophys Res Lett 31(6) L06609:1–3. https://doi.org/10.1029/2004GL019497

Márquez V, Nava F, Reyes G (2012) Multifractality in seismicity spatial distributions: Significance and possible precursory applications as found for two cases in different tectonic environments. Pageoph 169(12):2091–2105. https://doi.org/10.1007/s00024-012-0473-9

Meng L, Inbal A, Ampuero J (2011) A window into the complexity of the dynamic rupture of the 2011 Mw 9 Tohoku-Oki earthquake. Geophys Res Lett 38:L00G0. https://doi.org/10.1029/2001GL048118

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, Ávila-Barrientos L, Márquez-Ramírez V, Torres I, Zúñiga F (2018) Sampling uncertainties and source b likelihood for the Gutenberg-Richter b-value from the Aki-Utsu method. J Seismol 22:315–324. https://doi.org/10.1007/s10950-017-9707-8

Nicholson T, Sambridge M, Gudmundsson Ó (2000) On entropy and clustering in earthquake hypocentre distributions. Geophys J Int 142:37–51. https://doi.org/10.1046/j.1365-246x.2000.00113.x

NUMO (Nuclear Waste Management Organization of Japan) (2004) Development of repository concepts for volunteer siting environments. Report NUMO-TR-04-03. https://www.numo.or.jp/en/reports/pdf/RC_040901_FNL.pdf

Rényi A (1959) On the dimension and entropy of probability distributions. Acta Math 10:193–215

Rhoades D, Schorlemmer D, Gerstenberger M, Christophersen A, Zechar J, Imoto M (2011) Efficient testing of earthquake forecasting models. Acta Geophys 59:728–747

Rundle J, Turcotte D, Shcherbakov R, Klein W, Sammis C (2003) Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. Rev Geophys 41(4). https://doi.org/10.1029/2003RG000135

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

Schorlemmer D, Werner M, Marzocchi W, Jordan T, Ogata Y, Jackson D, Mak S, Rhoades D, Gerstenberger M, Hirata N, Liukis M, Maechling P, Strader A, Taroni M, Wiemer S, Zechar J, Zhuang J (2018) The collaboratory study of earthquake predictability: achievements and priorities. Seismol Res Lett 89(4):1305–1313. https://doi.org/10.1785/0220180053

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

Suzuki W, Aoi S, Sekiguchi H, Kunugi T (2011) Rupture process of the 2011 Tohoku-Oki mega-thrust earthquake (M9.0) inverted from strong-motion data. Geophys Res Lett 38:L00G16. https://doi.org/10.1029/2012GL049136

Taroni M, Zechar J, Marzocchi W (2013) Assessing annual global M 6+ seismicity forecasts. Geophys J Int 196(1):422–431. https://doi.org/10.1093/gji/ggt369

Turcotte D (1997) Fractals and chaos in geology and geophysics. Cambridge Univ. Press, Cambridge

USGS (2011) "M 9.1 – near the east coast of Honshu, Japan". Significant Earthquakes Archive. https://earthquake.usgs.gov/earthquakes/browse/significant.php. Retrieved May 18, 2020