The authors present an innovative relationship between the earthquake hazard and the Second Law of Thermodynamics by using the Shannon entropy, H, as a indicator of the changes in the seismic activity of a region.

The introduction is well structured, and it shows the evolution of the entropy concept from the very beginning to the current use in seismology as an indicator of the evolution of a system. The theoretical framework allows to understand the relationship between the entropy and the Gutenberg-Richter law although some steps should be explained better (see comments below). The methodology clearly explains how the variables needed to compute the entropy are obtained. The chosen datasets (northern Chilean seismicity) look adequate to probe the hypothesis although it should be explained better (see comments). The results are well presented and explained although some figures can be improved (see comments). The discussions are clear and meaningful.

The overall quality of the paper is very good, and the results can be interesting to the scientific community.

First, we would like to thank referee number 1 for his/her valuable and constructive comments.

Specific questions/issues

1. The authors compute in the first part of the results the entropy for the whole catalogue (all depths). However, the number of earthquakes are not equally distributed for all the different depths and the seismic activity distribution is not the same for cortical and subduction or deep earthquakes so why is the reason of computing entropy using all of them if the physics is going to be different for the three depth regions?. The authors should explain why they do this better and if the idea is to demonstrate that the catalogue has to be spitted by depth region maybe even to carry out the fast fourier transform in Figure 6 may arrive to three main frequencies associated to the different entropy behaviour in each one of the three depth regions.

Referee number 1 is right in identifying different physical behaviours depending on the depth of the selected events. The behaviour of the shallow part of the crust (the relationship between stress, strain and fracture) is different from those occurring in the upper mantle. Our idea is to show how the methodology developed in the paper can identify large earthquakes even using the entire catalogue. This can be seen in figure 6 and it is necessary to emphasize that it is a first approximation to the study of the seismicity and the earthquake hazard characterization of the area. As can be seen in figure 6, other minor variations in entropy (concerning the Tocopilla and Iquique earthquakes) appear and this motivates us to analyse seismicity by depths as a second approximation.

To explain our approximation, we have added a new paragraph at the beginning of the Results section as referee # 1 has suggested.

2. Additionally, in Figure 11 you compute the fast fourier transform only to the intermediate depth region. Why? You argue about the apparent periodicity of the entropic signal here but why it has to be periodic? And why only in the intermediate region? I would also see an apparent periodicity in the deep region with a period of about 1500 days. Why there is no periodicity for shallow earthquakes?. The stress loading rate is usually not uniform in time and a large earthquake may change the stress on the adjacent segment changing the seismic activity behaviour. Also the stress drop may change from event to event and the strength of the crust is not usually constant in time either. Therefore, the author should also try to address this.

Indeed, the periodic behaviour could be in the three regions A, B and C. In fact, as referee #1 correctly points out, in region C a period of 1500 days can be appreciated, although, as the authors indicated, maybe it would be necessary a longer time series be able to conclude the existence of this periodicity (the "sample" only has 3000 days). In short, in region A there are very small and irregular periods, possibly associated with a stress-loading rate usually not uniform in time because, as is well known, the strength of the crust is not constant. What happens in the deeper regions, B and C, should be associated with the extremely complex mechanisms of subduction and slag detachment. However, the periodicity that we had better observe is that of region B and, for this reason, we have carried out the fast Fourier transform only to the intermediate depth region.

To explain better why only use the fast Fourier transform in region B, we have added a new paragraph at line 256 as referee # 1 has suggested.

3. How the uncertainty in magnitude and epicentral and depth location is taken into consideration during the analysis? The authors should also try to explain in the detail about this uncertainty and treatment in the catalogue and in the analysis.

IPOC catalogue used in our work is available from the important and recent paper by Sippl et al. (2018). In that work, cited in our manuscript, authors presented a regional earthquake catalogue containing 101,601 double-difference relocated earthquakes to show high-resolution seismicity images of the northern Chile subduction zone forearc. The dataset is extended to 8 years of continuous seismic waveform data using automatic event detection and phase-picking routines. The uncertainty in hypocentral location was computed by using the probabilistic routine NonLinLoc (Lomax et al., 2000); authors proved that location errors inside the network are typically small (< 5 km) and hypocentral depth errors appear to be systematically larger than horizontal location errors, although this effect is small inside the network. Following Sippl et al. (2018), local magnitudes were determined from maximum amplitudes on the horizontal components after Hutton and Boore (1987) and they retrieved a total of 1,200,404 P and 688,904 S phase picks with average uncertainties of 0.11 and 0.37 s, respectively. Definitively, IPOC catalogue is a high-quality database. All these details are included in the Sippl et al. (2018) paper and we refer to readers to that work (line 192 in our paper). In our opinion, the IPOC catalogue represented an effort to minimize sources of error. However, these data are not absolutely free of

error, which is not uniform in time or place. Our approach has been to begin with a general, kind of average approach and later illustrate differences with respect to depth, which is probably the less precise parameter since it is strongly model dependent especially for deeper seisms. The latitude and longitude coordinates are more precise for shallow earthquakes. Errors in magnitude can be of limited character only, since they handle 2 significant figures at the most. Time determination is almost error free, since it is determined from the arrival of the P waves to several stations. The kind of errors present in these problems are known and they should be bore in mind when analysing the present and other results.

Hutton, L. K., & Boore, D. M. (1987). The Ml Scale in Southern California. Bulletin Of The Seismological Society Of America, 77(6), 2074–2094.

Lomax, A., Virieux, J., Volant, P., & Berge-thierry, C. (2000). Chapter 5 Probabilistic earthquake location in 3D and layered models. In A. Lomax, et al. (Eds.), Advances in Seismic Event Location (pp. 101–134). Amsterdam: Kluwer Academic Publishers.

Sippl, C., Schurr, B., Asch, G., Kummerow, J., Seismicity structure of the northern Chile forearc from >100,000 double-difference relocated hypocenters, J. Geophys. Res. Solid Earth 123, 4063–4087; doi: 10.1002/2017JB015384 (2018).

Additionally, I would suggest you take into consideration the following corrections:

Paragraph 40. Sentence no.4. change entropy is by entropy (S) is

Done

Equation (1) define in paragraph 45 what l and n means (you only have done it for k and W)

We are sorry. "l" and "n" must be joined as "ln" , the natural logarithm; we have corrected it.

Equation (4). As in this section you are still in speaking about entropy, I would suggest saying that assuming k=1 then it is possible to define the Shannon information entropy (H) combining equation (2) and (3). That will allow the reader to know exactly why H(p)=-I(p). Can you explain why k=1 is assumed? In this equation also you should explain why you have changed W to W (also in equation 5) and define W if it is what you want to write.

We believe that your suggestion is right and we have rewritten the paragraph following your recommendation. Moreover, a new reference was added (e.g., Truffet, 2018) to better explain why k = 1. Finally, we have unified the symbols W and Ω.

L. Truffet, Shannon Entropy Reinterpreted, Reports on Mathematical Physics, 81, 3, 2018, 303-319, https://doi.org/10.1016/S0034-4877(18)30050-8.

Equation (5) in the sentence where t is a real number called the entropic index I suggest you improve the sentence saying where t is named the entropic index and can, in principle, be any real number.

This suggestion have been incorporated in the manuscript

Equation (6) and (8). Sometimes you use x to represent the multiplication but others you do not use it, so I recommend removing the symbol x in all the equations.

Of course, you are right and the "x" symbol has been removed.

Paragraph 95. In the sentence: On the other hand, if we have N earthquakes and n denotes the number of earthquakes with magnitude M you have to say On the other hand, if we have N earthquakes and n denotes the number of earthquakes with magnitude equal to or larger than M because it is needed to match with the Gutenberg-Richter relationship where n has that meaning.

Done.

Page 6. First sentence. After the words the calculation of entropy I would add between brackets a reference to the equation you are going to calculate)

Done.

Figure 1. It would be nice if you can add to the right of this figures two histograms (one with the magnitude frequency distribution and other with the depth frequency distribution)

Done. Figure 4 and figure 8 were added to figure 1.

Figure 8. Add dashed lines to the figure to separate the three mentioned regions.

Done

Figure 10. I would remove this figure because it it the same information as in Figure 11 and Figure 11 is more illustrative than the previous one

Done