

## **Response to comments by the reviewers and editor**

We appreciate the comments from two reviewers and a community member, which have helped us improve the manuscript. Overall, we have followed all the suggestions. In the revised version of the manuscript, we highlighted changes in yellow. We now give a response to the individual points raised:

**Reviewer #1: The papers analyzes the statistical characteristics of Non-Volcanic Tremor sequences that took place in the Mexican Subduction Zone. Its originality lies in the subject selection for analysis, the inclusion of all possible cases and the relevant results presented for the first time for the study area. It is very contributable work and worth publishing.**

**The text is perfectly written and clear in all aspects. It is difficult to find any point that needs further clarification.**

**I could comment the linear fit of larger magnitude range when the  $b$  – value is estimated. By fitting 3 points for example in Sequence 3, it rather affects the estimations. In this case I would prefer a larger magnitude cutoff for fitting more points and in particular the larger magnitude earthquakes that play more significant role in the local seismicity. This would of course lead to large  $b$  – value (steeper slope), which in this case already large. If the authors tried to avoid that latter, it would be better to look again at their data sample.**

We would like to clarify that the  $b$ -values in this study were not estimated using a simple linear fit to a limited number of points, but rather using a maximum likelihood method, which is less sensitive to binning and to the number of data points in the frequency–magnitude distribution. In addition, the estimation was performed using a magnitude of completeness ( $M_c$ ) threshold for each sequence, ensuring that only statistically complete portions of the catalog were considered. This approach minimizes biases associated with small sample sizes or incomplete detection of low-magnitude events,

while retaining all available events above  $M_c$ , including the larger magnitudes. Furthermore, the bootstrap resampling used in this study provides uncertainty bounds, demonstrating the robustness of the estimated parameters. We have clarified this point in the manuscript to avoid any confusion regarding the fitting procedure.

**Better representation quality for Figure 4 – is there any reason to start the magnitude axis from number where there is not any observation?**

The figure was drawn according to the data in Table 1, Figure 3 was corrected.

**Before final acceptance, however, the paper could be improved regarding the presentation of the work and the outcomes, for the reader's ease. What I mean is that the calculations of the different data samples (i.e., for each sequence) along with for each statistical parameter and tool are continuously and consequently given, and the reader needs to go back and forth to see the similarities, the differences or the possible peculiarities in each case (and subcase). I encourage the authors to examine any possibility to make their manuscript more easily following, perhaps by putting all number given, compared and discussed inside the text, in corresponding tables (more than the tables already shown).**

We agree that the presentation of results across multiple sequences and statistical analyses can make comparisons less straightforward. Following the reviewer's advice, to improve readability, we have reorganized parts of the manuscript to reduce repetition and enhance the flow of the discussion. We believe these changes significantly improve the clarity and accessibility of the manuscript.

**Reviewer #2: In this paper, the authors apply different statistical methodologies (the Gutenberg–Richter relation, non-extensive statistics, and multifractal detrended moving average analysis) to**

**characterize the behavior of the magnitude and interevent-time distributions of non-volcanic tremor along the Mexican subduction zone. I found this to be an interesting work in which the authors successfully apply these techniques to different sequences, providing a robust analysis of the nonlinear properties of the analyzed data.**

**I recommend acceptance of this work after minor revisions.**

**1) In the abstract, the authors mention that multifractality may be associated with three different properties: long-term correlations, the probability distribution of the data, and the presence of nonlinearities. However, in the conclusions it is not clear which of these properties are responsible for the multifractal behavior of the sequences. I suggest clearly stating in the conclusions the possible cause(s) of the multifractality found in the analyzed sequences.**

We agree with the reviewer that the conclusions should more explicitly state the identified sources of multifractality. Following this suggestion, we have rephrased the conclusions section to clarify that the multifractal behavior in NVT sequences is primarily driven by the broad probability distributions (apparent multifractality) and, to a lesser extent, by long-term correlations (intrinsic multifractality), as evidenced by our surrogate data analysis. We have also addressed the role of nonlinearities as a contributing factor in specific cases. In the revised manuscript, we have clarified this point by explicitly linking the multifractal behavior to the underlying factors also mentioned in the abstract.

**2) Define the acronym NVT in the abstract before using it again.**

We defined the acronym of NVT in the abstract.

**3) At the beginning of the Data subsection, I suggest adding a clear explanation of how the NVT sequences were extracted from the database.**

We clarified this point in the manuscript. Catalogs for sequences 1-3 can be downloaded directly from the World Tremor Database via its website (see the data availability section). On this website, different catalogs are displayed.

**4) Page 5, line 7: Why are the magnitudes not reported for sequence 6?**

We clarified this point in the manuscript. For sequence 6, Chen et al. (2025) used an algorithm based on envelope correlation and matched filter, obtaining location, duration, and average seismic energy rate (instead of magnitude). In many cases, NVTs are a superposition of many low-frequency earthquakes (LFEs) of low amplitude; for this reason, some authors prefer to use the average energy or the root-mean-square (RMS) of the amplitude.

**5) Page 5, line 10: Please add a reference or provide more details about the temporal seismic networks.**

We added more information regarding the temporary seismic networks.

**6) From Figure 3, how did the authors compute  $M_c$ ? Please add the methodology used to estimate it.**

$M_c$  was calculated using the maximum curvature method which is now explicitly stated.

**7) Section 3.2.3: I suggest adding references to other papers in which MDFA has been applied, such as:**

-Telesca, L., & Lapenna, V. (2006). Measuring multifractality in seismic sequences. *Tectonophysics*, 423(1–4), 115–123.

-Monterrubio-Velasco, M., Lana, X., Martínez, M. D., Zúñiga, F. R., & Puente, J. D. L. (2020). Evolution of the multifractal parameters along different steps of seismic activity: The example of Canterbury 2000–2018 (New Zealand). *AIP Advances*, 10(11).

-Alam, A., Nikolopoulos, D., Cantzos, D., Tahir, M., Iqbal, T., Petraki, E., & Rafique, M. (2023). Regional multifractal variability of overall seismic activity in Pakistan from 1820 to 2020 via the application of MDFA on earthquake catalogs. *Fractal and Fractional*, 7(12), 857.

We incorporated an introductory paragraph regarding the use of the MF DFA in seismology.

**8) Page 7, line 13: Please modify the parameter  $M(t)$  to avoid confusion with the magnitude parameter.**

We agree with this observation. We changed  $M(t)$  for  $D(t)$ .

**9) Page 8, lines 1–5: Please rewrite this paragraph to improve clarity. It seems that a symbol in Equations (6) and (7), as well as in the corresponding paragraphs, is missing in the printed version.**

We rewrote these lines. The symbols in equations 6 and 7 are correct. Here,  $\lfloor x \rfloor$  denotes the floor function (the largest integer less than or equal to  $x$ ), while  $\lceil x \rceil$  represents the ceiling function (the

smallest integer greater than or equal to  $x$ ). We kindly ask the reviewer to check the reference to Gu and Zhou (2010) for an extended explanation.

**10) Page 8, line 18: The term “q-th order parameter” may be confused with the non-extensive parameter q. Please clarify this distinction.**

This is a crucial distinction to make, as using the same letter  $q$  for two different concepts in statistical physics and complexity science is a frequent source of "notational whiplash." The confusion usually arises because both parameters highlight different aspects of a system, but they belong to entirely different mathematical frameworks, fractal geometry and statistical thermodynamics, respectively. In the context of MFDFA,  $q$  is the order of the fluctuation and is used to scan different fluctuation scales. The "non-extensive parameter" ( $q$ -value) comes from Tsallis entropy and is used to quantify non-additivity and entropy. We mentioned this in the text.

**11) Page 9, line 17: Why did the authors use the input parameters  $N = 30$ ,  $\theta = 0$ , and  $q \in [-5, 5]$ ? Were other values tested? How might the selection of these parameters affect the results? From a statistical point of view, is  $N = 30$  sufficient?**

To ensure methodological consistency across the different datasets, a unified parameter set of  $N = 30$ ,  $\theta = 0$ , and  $q \in [-5, 5]$  was adopted. The choice of  $N = 30$  provides an adequate scale to capture local trends while maintaining sufficient statistical reliability, particularly for the shortest series, while still allowing detailed local detrending in the longer series. Although this choice represents different relative scales for different lengths, it ensures a uniform detrending procedure and avoids introducing additional variability associated with scale-dependent parameter tuning. The parameter  $\theta = 0$  was adopted, corresponding to the backward moving average, which preserves causality and avoids forward-looking

bias. The range  $q \in [-5, 5]$  was chosen to probe multifractal properties across different fluctuation magnitudes while maintaining numerical stability, particularly in shorter datasets. This parameter configuration ensures consistent, robust estimation of scaling behavior across time series with heterogeneous lengths.

**12) Page 10, line 22: Which singularities are being referred to?**

In the context of multifractal time series analysis, the term "singularities" refers to the local complexity, or "roughness," of the data. The presence of a long left tail indicates a system dominated by strong singularities, which manifest as rapid, intense structural changes. These high-magnitude fluctuations characterize the 'rougher' or more volatile components of the multifractal signal. The presence of a long right tail signifies the dominance of weak singularities. These represent the smaller, smoother variations in the data. According to Ihlen (2012), this indicates that the signal's complexity is concentrated within its low-amplitude, subtle fluctuations.

**13) Page 11, line 15: Please add references for the shuffling and surrogate data procedures.**

We added the references in the new version of the manuscript.

Shuffle:

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., and Stanley, H. E.: Multifractal detrended fluctuation analysis of nonstationary time series, *Physica A*, 316, 87-114, [https://doi.org/10.1016/S0378-4371\(02\)01383-3](https://doi.org/10.1016/S0378-4371(02)01383-3), 2002.

Surrogate:

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. D.: Testing for nonlinearity in time series: the method of surrogate data, *Physica D*, 58, 77-94, [https://doi.org/10.1016/0167-2789\(92\)90102-S](https://doi.org/10.1016/0167-2789(92)90102-S), 1992.

**14) In the Conclusions section (page 18), is there any relationship between the  $b$ -values and the  $q$ -values obtained in the results?**

We found that the relationship between the  $q$ - and  $b$ -values is as follows: down-dip NVT sequences have the highest  $b$ -values. In contrast,  $q$ -values in down-dip regions are lower than those in coastal areas. These results are consistent with the relation between the  $q$ - and  $b$ -values proposed by Silva et al. (2006) and Telesca (2012):  $b = 2(2-q)/(1-q)$ . Although our results and  $b$ -value estimates from this theoretical relationship differ in numerical values, both agree in identifying that high  $b$ -values correspond to low  $q$ -values.

### **Community member observations**

**1. Sample Sizes and Statistical Validity.** The manuscript analyzes sequences with dramatically different sample sizes ( $N = 101$  to 23,408 events), yet applies identical statistical methods without addressing the implications. Sequence 5 ( $N = 101$ ) shows notably different behavior in multiple analyses, which may reflect statistical artifacts rather than genuine tectonic differences. The authors acknowledge this briefly in the discussion of interevent-time distributions but do not systematically evaluate how sample size affects their multifractal parameters,  $b$ -values, or  $q$ -values.

**Recommendation:** Include bootstrap resampling analyses to assess the stability of estimates for small-sample sequences, or clearly delineate which results may be sample-size dependent.

To assess the stability of our estimates, we applied bootstrap resampling with 1000 iterations for all sequences. The updated results (Tables 2 and 3) report b-values,  $M_c$ , q-values, and related parameters with bootstrap-derived uncertainties, allowing a quantitative evaluation of the effect of sample size. Even for the small-sequence 5 ( $N = 101$ ), the parameters remain statistically robust, and observed differences reflect genuine tectonic variations rather than artifacts. We have added a note in the manuscript highlighting that smaller sequences exhibit larger relative uncertainties, now explicitly quantified by the bootstrap.

**2. Magnitude Completeness ( $M_c$ ) Determination.** Table 2 reports  $M_c$  values ranging from 1.10 to 1.80, yet the manuscript provides no information about the method used to determine these thresholds or their uncertainties. Given that sequences include magnitudes as low as -0.8, and that b-value estimates are susceptible to  $M_c$  selection (Woessner & Wiemer, 2005), this omission is problematic. The significant variation in  $M_c$  across spatially proximate sequences (e.g., sequences 1 and 3, both in the Guerrero down-dip region, with  $M_c = 1.80$  and 1.10, respectively) suggests potential inconsistencies in catalog completeness or detection capabilities.

**Recommendation:** Report the method for  $M_c$  estimation, include uncertainty bounds, and discuss how temporal variations in network geometry may affect completeness thresholds.

$M_c$  values were estimated using the maximum curvature method with bootstrap-derived uncertainties (1000 iterations), suitable for low-magnitude NVT catalogs. Temporal variations in network geometry affected detection: sequences 1–3 (Ide, 2012; Idehara et al., 2014) used temporary networks, sequence 4 (Husker et al., 2019) had limited coverage, and sequences 5–6 (Plata-Martínez et al., 2021; Chen et al., 2025) used submarine seismometers, though Chen et al. (2025) did not report magnitudes. These

updates are now discussed in the manuscript, and Table 2 includes  $M_c$  with uncertainties, providing a robust basis for interpreting  $b$ -values and Gutenberg-Richter relationships for NVT.

**3. Multifractal Analysis: Statistical Significance of Differences**The multifractal analysis reveals that shuffled and IAAFT surrogate procedures "cannot destroy the multifractality" (lines 281-283), yet Table 4 shows that many p-values are inconclusive or suggest apparent rather than intrinsic multifractality. The interpretation that "long-term correlations, the probability distribution of the data, and the presence of nonlinearities are present" appears overstated. For example:

**1) Magnitude sequence 5: all p-values (0.39-0.83) suggest apparent multifractality.**

**2) Interevent sequences 1, 4, 5: high p-values indicate PDF characteristics may dominate multifractality.**

The authors correctly identify these cases individually but then make generalized statements about all sequences exhibiting these features.

**Recommendation: Distinguish clearly between sequences showing intrinsic versus apparent multifractality in the conclusions, and avoid generalizations that apply only to subsets of the data.**

We agree that some of our original wording may have overgeneralized the presence of intrinsic multifractality. As noted, several sequences (e.g., magnitude sequence 5 and interevent time sequences 1, 4, and 5) show high p-values, indicating apparent multifractality, while others (magnitude sequence 1; interevent sequences 2, 3, 6) exhibit statistically significant deviations from surrogate data, reflecting contributions from long-range correlations and possible nonlinear dynamics. Accordingly, we have revised the manuscript to clearly distinguish intrinsic, apparent, and inconclusive multifractality and to avoid generalizations. The Conclusions now explicitly reflect the sequence-dependent nature of

multifractality, consistent with Table 4 and the results section. We appreciate the reviewer's suggestion, which improved the clarity and consistency of our interpretation.

**4. Hurst Exponent Interpretations.** The reported Hurst exponents (0.65 to 1.06, with most  $H > 0.95$ ) are interpreted as indicating "exceptionally high persistent memory" and attributed to "relatively limited volume of perturbed regions where fluids are present" (lines 306-308).

**However:**

- 1.  $H > 1.0$  theoretically indicates non-stationary, trending behavior rather than persistent memory**
- 2. The comparison with aftershock sequences ( $H = 0.95$  for Izmit) does not account for fundamental differences between aftershock relaxation and NVT generation mechanisms**
- 3. The physical interpretation linking high  $H$  values to fluid-filled volume is speculative without supporting evidence.**

**Recommendation:** Acknowledge that  $H > 1.0$  may indicate non-stationarity, discuss potential contributions from catalog heterogeneity or temporal clustering, and either provide references supporting the fluid-volume interpretation or present it explicitly as a hypothesis.

We thank the reviewer for these insightful observations, which have helped us refine the statistical and physical interpretation of our results.

1)  $H > 1.0$  and non-stationarity: We acknowledge that  $H > 1.0$  technically represents a transition from stationary persistent memory to non-stationary, "trending" behavior (e.g., fractional Brownian motion). In the revised manuscript, we clarify that values exceeding unity likely reflect long-term non-stationarities or temporal clustering within the NVT catalogs, rather than simple persistence. This interpretation is supported by our statistical tests using surrogate data, which highlight sequence-dependent deviations indicative of complex temporal organization.

2) Comparison with aftershocks: We agree that the physical mechanisms driving aftershock relaxation (stress transfer) and NVT generation (often fluid-driven) are fundamentally distinct. We have revised the text to clarify that the comparison with the Izmit sequence ( $H = 0.95$ ) is intended strictly as a statistical benchmark for scaling behavior, rather than implying shared mechanical origins. The purpose of this comparison is to provide context for the magnitude of  $H$  values observed in our sequences relative to well-characterized seismic catalogs.

3) Fluid-volume interpretation: We concede that the proposed link between high  $H$  values and fluid-filled volumes is a hypothesis. We have rephrased this discussion to present it explicitly as speculative and have added alternative contributors, including catalog heterogeneity and structural complexity. We also emphasize that this hypothesis is consistent with the sequence-dependent patterns observed in our surrogate-based statistical analyses, rather than being a direct physical conclusion.

**5. Duration-Magnitude Scaling Relationships. Figure 11 and Table 6 present duration scaling with remarkably low  $R^2$  values (0.03-0.34), indicating that magnitude explains only 3-34% of duration variance. The authors acknowledge this scatter but attribute it ambiguously to either "genesis of NVT or detection process" (lines 294-295). Given that:**

- 1) All sequences show similar scatter regardless of location or time period.**
- 2) The relationship is fundamental to understanding NVT source physics.**
- 3) Poor scaling contrasts with better-established scaling for regular earthquakes.**

**The dismissive treatment of this negative result undervalues its potential significance. Poor scaling may indicate that NVT magnitude and duration are governed by different physical processes, challenging the assumption that NVT follows earthquake-like scaling.**

**Recommendation: Expand the discussion of duration-magnitude decoupling, potentially including analysis of whether specific subsets (e.g., depth ranges, different tectonic settings) show improved scaling.**

We agree that the low coefficients of determination ( $R^2 = 0.03\text{--}0.34$ ) are not merely a limitation, but may reflect an important physical feature of non-volcanic tremor (NVT). In the revised manuscript, we have expanded the discussion to emphasize that the weak scaling between duration and magnitude suggests a possible decoupling between these variables. This behavior contrasts with the more robust scaling relationships observed in regular earthquakes and may indicate that different physical processes control NVT duration and magnitude. Furthermore, the observation that all sequences exhibit similarly low  $R^2$  values, regardless of location or time period, supports the idea that this decoupling is more likely an intrinsic property of NVT than an artifact of data quality or detection methods. We have clarified this point in the discussion and conclusions.

**Minor Suggestions. Data and Methodology.**

**1. Line 92: The study period extends to 2019, yet the manuscript was submitted in December 2025. Were more recent data unavailable, or were there specific reasons to limit the analysis period?**

We do not know exactly, but we used all the publicly available data.

**2. Lines 124-125: The statement "NEVER use local Storage or session Storage" appears to be artifact-creation instructions that should not be in the manuscript text.**

We did not find this statement in the manuscript.

**3. Table 1: Sequences 4 and 6 lack magnitude data, limiting comparative analyses. The rationale for including these incomplete catalogs should be stated explicitly.**

We understand the reviewer's point, but these sequences provide information about interevent behavior.

**4. Figure 2: The spatial distribution of sequences varies dramatically (concentrated vs. dispersed). Has spatial clustering been considered as a factor affecting statistical properties?**

We agree with the reviewer. The spatial analysis is outside the scope of this study, but could be a subject of study in the near future.

#### **Minor Suggestions. Results Presentation.**

**1. Tables 2-3: The b-value and q-value show inverse correlation, described as "apparent." Statistical testing of this correlation would strengthen the observation.**

Using a Monte Carlo simulation based on the reported uncertainties in b-values and q-values (1000 iterations), we evaluated the robustness of the apparent inverse correlation. The results confirm a strong negative correlation, with mean Pearson  $r = -0.94$  and a 95% confidence interval of  $-0.999$  to  $-0.743$ . This supports the inverse relationship observed in Tables 2–3, and we have added a statement in the Results section reflecting this finding.

**2. Figure 10: The probability density functions would benefit from logarithmic y-axes to better visualize tail behavior, which is critical for distinguishing between distributions.**

We plot the results as loglog.

**3. Lines 313-316: The interpretation that Lognormal distributions suggest "mixed behavior with characteristics compatible with tectonic earthquakes and volcanic seismicity" requires justification, as Lognormal distributions are also observed in purely tectonic settings (Mesimeri et al., 2019, cited by the authors).**

We acknowledge that lognormal interevent-time distributions can also occur in purely tectonic settings (Mesimeri et al., 2019), and therefore the presence of a lognormal distribution alone does not uniquely imply mixed tectonic–volcanic behavior. In our study, the interpretation of “mixed behavior” is supported by additional indicators, including magnitude ranges,  $q$ -values, and variability in short and long interevent times. As shown in Table 5, sequences 1–4 are best described by lognormal distributions, while sequences 5–6 are better fitted by Gamma distributions, with smaller sample sizes contributing to higher uncertainty in the fits. This approach is consistent with global and regional studies (Bantidi, 2022; Mesimeri et al., 2019) and highlights that the lognormal fit is one part of a broader set of observations supporting complex NVT behavior, rather than being interpreted in isolation.

#### Data Availability

The original NVT catalogs cannot be redistributed due to copyright restrictions. The codes used for analysis are publicly available as cited in the Code Availability section.