Earthquake parameters comparabilityPremonitory earthquakes clustering process in an equivalent dimensions space before the Mw8.2 Tehuantepec, Mexico, 2017 mainshock

Stanisław Lasocki[1], Vasileios G. Karakostas[2], F. Ramón Zúñiga[3], and Eleftheria E. Papadimitriou[2]

[1]Institute of Geophysics, Polish Academy of Sciences

[2]Geophysics Department, Aristotle University of Thessaloniki

[3]Instituto de Geofísica, Universidad Nacional Autónoma de México

Corresponding authors: Stanisław Lasocki (lasocki@igf.edu.pl)

Key Points:

- The average distance between earthquakes in the interevent time, interevent distance, and magnitude space systematically changed in time.
- Statistically significant up-down trends of that distance occurred in the twelve years preceding the mainshock.
- These premonitory changes resulted from the earthquake clustering process evolution.

Abstract

A new approach to monitoring stochastic features of earthquake series was applied to track the preparatory process of the 2017 Tehuantepec, Mexico, Mw8.4 earthquake. The seismicity was parameterized by elapsed times and epicentral distances from the directly preceding earthquakes and by earthquake magnitudes. The transformation to equivalent dimensions ensured the comparability of these parameters. We were calculating the average distance between earthquakes in the transformed parameter space in moving in time data windows, each consisting of 100 events. The average distance exhibited a consistent upward trend from ten to two years before the mainshock. Then, it declined until the mainshock. This precursory up-down signal was highly significant statistically. We showed that the detected time changes of the average distance resulted from the evolution of the earthquake clustering in the considered parameters' equivalent dimensions space.

Plain Language Summary

Despite intense research, so far, there are no methods, which decipher with usable confidence that a seismic process leads to large and great earthquakes, and the problem awaits novel approaches. We parameterized the seismicity preceding the 2017 Tehuantepec, Mexico, Mw8.4 earthquake by elapsed times and epicentral distances from the directly preceding events and earthquake magnitudes. The transformation to equivalent dimensions, an innovative methodology to investigate earthquake clustering, ensured the comparability of these parameters. The average distance between earthquakes in the space of the transformed parameters exhibited a consistent upward trend from ten to two years before the mainshock. Then it declined until the mainshock. This highly significant statistically precursory up-down signal resulted from the evolution of earthquake clustering in the used parameter space.

1 Introduction

Non-random, stochastic features of seismic series express the dynamics of seismic processes. Identifying and deciphering these features significantly improve the capabilities for earthquake forecasting. Commonly used earthquake parameters are the occurrence time, epicenter location, hypocentral depth, and magnitude. Due to that, many studies of the earthquake series non-randomness investigate stochastic properties – short- and long–range clustering, memory, etc. of the series of occurrence times, locations, and magnitude, or derived parameters – interoccurrence / interevent / recurrence time and interevent spatial distance.

In an early study of earthquake clustering in time, Gardner and Knopoff (1974), based on the chi-square test results, concluded that earthquake series without aftershocks was Poissonian. Kagan and Jackson (1991) drew the opposite conclusion based on investigations of interevent time distribution in earthquake series deprived of short-term temporally clustered fragments. However, the authors admitted that mainshocks' long–term clustering features were not readily identifiable. Since then, much evidence of the persistence and long–term memory in the time series of seismicity parameterizations has accumulated.

E.g., Corral (2004) compared the shapes of the earthquake recurrence time probability densities for different spatial areas and magnitude ranges and indicated that the clustering became self-similar in the localization–time–magnitude space after rescaling the time with the event occurrence rate. Livina et al. (2005) showed that the interevent time was positively correlated with the previous interevent time. Batac and Kantz (2014) found a certain threshold of

spatial separation of consecutive events. Events separated less than that threshold value were correlated, while the more distant events were not. Moreover, the correlated events were clustered in both time and space. Zhang et al. (2020) affirmed long–term memory for interevent times and interevent distances from the estimates of lagged conditional density functions of these parameters.

The Hurst exponent, *H*, parameterizes the long–term memory amount in a time series (Hurst, 1951 and, e.g., Mandelbrot & Wallis, 1969). *H* estimated for the series of times between consecutive earthquakes was often trending downward before moderate–strong earthquakes (Liu et al., 1995). Studying *H* for the same parameterization Alvarez–Ramirez et al. (2012) indicated that seismicity was clustered in time and the occurrences of more significant earthquakes were positively correlated with *H*. The *H* estimates for the interevent time series, magnitude series, and the series of inter-event epicentral distances of successive earthquakes in two different seismogenic zones in Greece evidenced clustering in time and space in both areas, but the magnitude series was internally correlated only in one (Gkarlaouni et al., 2017). *H* estimation documented also a long–range memory in cumulative seismic moment series (Barani et al., 2018).

Whereas space-time clustering of earthquakes is generally accepted, the views on the stochasticity of earthquake size (magnitude, seismic moment, energy) vary. In many studies, the magnitude was considered memoryless and modeled by an exponential distribution resulting from the Gutenberg-Richer relation (e.g., Kagan & Knopoff, 1981; Ogata, 1988; Sornette & Helmstetter, 2002; Baiesi & Paczuski, 2004; Zaliapin et al., 2008; Zaliapin & Ben-Zion, 2013; Batac et al., 2017; Michas et al., 2021), even though rigorous statistical tests undermined the universality of the Gutenberg-Richter relation (Lasocki & Papadimitriou, 2006; Lasocki, 2007). In recent years, however, there appeared studies evidencing the existence of memory in magnitude series (e.g., Lennartz et al., 2008; Lippiello et al., 2008; Węglarczyk & Lasocki, 2009; Gkarlaouni et al., 2017; Barani et al., 2018; 2021; Chen et al., 2022).

We studied the clustering of seismicity before the M8.2 Tehuantepec, Mexico, earthquake. It occurred on September 8, 2017, at 04:49:18 (UTC) and is the largest intraslab earthquake documented along the Mexican subduction zone.

The above-mentioned earthquake clustering studies indicate that interdependencies among earthquakes in time and space are more easily discernible when examining interevent times and interevent distances. For this reason, we parameterized the earthquakes from the 15 years before the M8.2 Tehuantepec earthquake by these two parameters and magnitudes. We transformed these parameters to equivalent dimensions (Lasocki, 2014) to get the Euclidean distance function in their 3D space. Then, in time-moving data windows, we calculated the average distance between the earthquakes in this space, and we found statistically significant time trends of this distance, which preceded the mainshock.

2 Data and Methods

The M8.2 Tehuantepec earthquake from September 8, 2017 exhibited normal faulting (Melgar et al., 2018; Suárez et al., 2019) within the subducting Cocos Plate ~70 km landward (Ye, et al., 2017) from the Middle American Trench beneath the Tehuantepec gap (Kelleher, et al., 1973; Kelleher & McCann, 1976) (Figure 1).

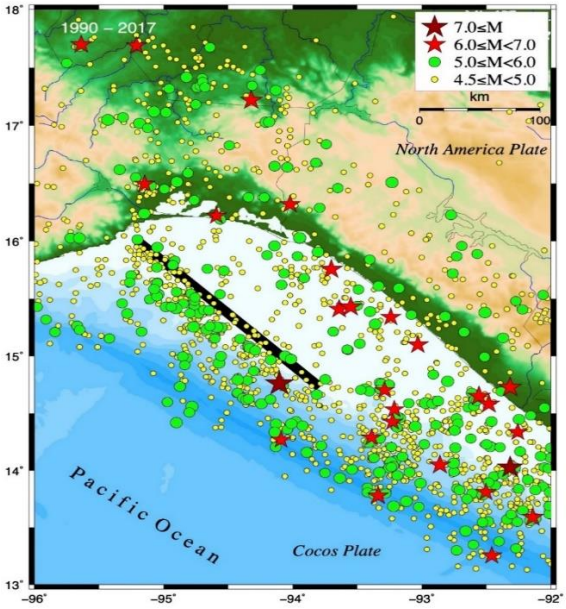

Figure 1. The study area where the epicentral distribution of earthquakes that occurred from 1990 – 2017 is shown. Different symbols show different magnitude ranges, as indicated in the legend. The thick black line shows the inferred surface extension of the main rupture.

The study area is bounded between 13°-18° N and 92°-96°W, with its length along the subduction front taken about three times the fault length of the 2017 rupture. We used the data since 1999 from the earthquake catalog compiled by the Universidad Nacional Autónoma de México, whose magnitudes were homogeneously calculated (Zúñiga et al., 2000). Because the selected data exhibited a sharp, physically unjustified increase of the event occurrence rate that began in 2015 but concerned only smaller events of magnitudes up to 4.5, we decided to use earthquakes of M>4.5. The studied dataset comprised 1048 events and the mainshock within this earthquake series was event #1049. The earthquakes were parameterized by the occurrence time, focal latitude, longitude and depth, and magnitude.

We re-parameterized these earthquakes by the occurrence time, *t*, magnitude, *M*, and the interevent time, *dt*, and orthodromic epicentral distance (interevent distance), *dr*, between this event and the preceding event. The re-parameterized series comprised 1047 eartquakes. Subsequently, we transformed the *dt*, *dr*, and *M* into their equivalent dimensions (EDs), *DT*, *DR*, and *MC*, respectively.

The transformation to EDs (Lasocki, 2014; 2021), based on a probabilistic equivalence of parameters of objects under study, replaces parameter values with their cumulative distribution function (CDF) values. Unknown CDF-s are estimated from data using the non–parametric kernel estimators. We used the data of all 1047 earthquakes under study to estimate the CDF-s of

124 *dt*, *dr*, and *M* (See: Supporting Information Text S1 on the transformation applied in the present
125 work).

126 All parameters transformed to EDs are uniformly distributed in [0, 1], hence comparable,
127 and the space of ED–transformed parameters has the Euclidean metric. After the transformation,
128 the earthquakes became points [*DT*, *DR*, *MC*] in the 3D cube [0, 1].

129 From the Euclidean metric of ED-s spaces, it comes that the average distance between *n*
130 earthquakes in the {*DT*, *DR*, *MC*} space reads

131 $$d_c = \frac{2}{(n-1)n}\sum_{k=1}^{n-1}\sum_{j=k+1}^{n} D(k,j) =$$

132 $$\frac{2}{(n-1)n}\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}\sqrt{\left(DT_k - DT_j\right)^2 + \left(DR_k - DR_j\right)^2 + \left(MC_k - MC_j\right)^2}. \quad (1)$$

133 Starting with event #8, we overlaid the event series 10–1047 with 48 windows of 100
134 consecutive events each, being shifted by 20 events, and calculated d_c in windows. Unless the
135 event generation process is stationary and ergodic, taking parameters of *n* consecutive events
136 from a time series of *N* events, $n<N$, is a non-random draw. Therefore, while in the population of
137 all data, the transformed parameters, *DT*, *DR*, and *MC*, were uniformly distributed in [0, 1], their
138 distributions in the sliding window were not and were changeable, likewise d_c, reflecting the
139 earthquake process's properties in the window's period. We linked d_c-s to the occurrence times of
140 the last events in the windows obtaining the 48-element time series $d_c(t)$.

141 The statistical significance of non–random features of $d_c(t)$ was studied in two ways.
142 First, we calculated d_c-values in 100,000 samples of size 100, each randomly drawn from the
143 initial series of [*DT*, *DR*, *MC*] triples from #8 to #1047 obtaining the empirical distribution of
144 $\{d_c|$*random sampling of [DT, DR, MC] triples*$\}$. From this distribution, we estimated the p_α and
145 $p_{1-\alpha}$ percentiles of d_c.

146 If a series of d_c comprises *n* values that were calculated from randomly drawn triples
147 [*DT*, *DR*, *MC*], the probability that *k* values > p_α equals the probability that *k* values < $p_{1-\alpha}$ and is:

148 $$\Pr(n,k,p_\alpha) = 1 - \sum_{m=0}^{k-1}\binom{n}{m}(1-\alpha)^m\alpha^{n-m}. \quad (2)$$

149 Using the significance α=0.995 we counted the number of times, *k*, the $d_c(t)$-values were
150 greater than p_α / less than $p_{1-\alpha}$ and applied Eq. 2 to calculate $\Pr(48,k,p_{0.995})$ / $\Pr(48,k,p_{0.005})$.

151 In the second test, we shuffled a hundred times the original series of triples [DT_k, DR_k,
152 MC_k], k=1, …, 1047. Then, using windows of 100 triples consecutive in shuffled series and
153 sliding them by 20 triples, we calculated one hundred $d_c(t|$*shuffled series of [DT, DR, MC]*
154 *triples*) series and compared qualitatively their shapes with the shape of actual $d_c(t)$ series.

155

3 Results and Discussion

The obtained time series of the average distances between earthquakes in the consecutive data windows, $d_c(t)$, is shown in Figure 2. From d_c=0.64 in window #20, attributed to 20 March 2007, $d_c(t)$ increases systematically until 28 June 2015 (d_c=0.78). Afterward, the trend of $d_c(t)$ exhibits a rapid decline, which continues from 6 May 2017 (d_c=0.68) to 7 September 2017 (d_c=0.68). The mainshock occurred on 8 September 2017.

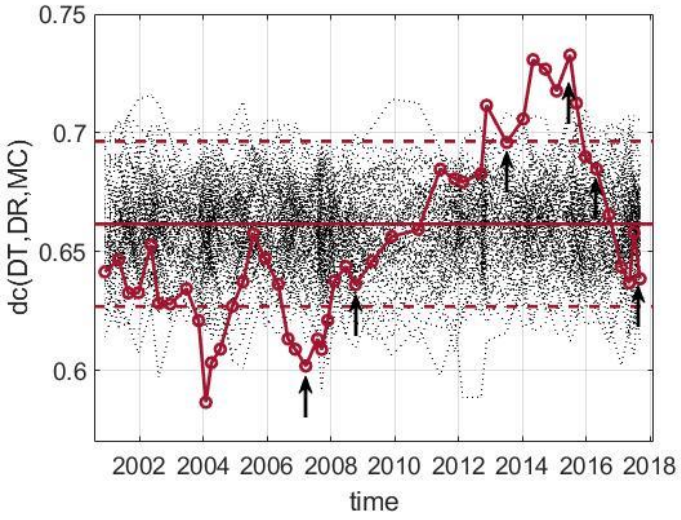

Figure 2. Time series of the average distance in the {DT, DR, MC} space between a hundred consecutive earthquakes. Horizontal solid - the mean dc for one hundred perfectly unclustered [DT, DR, MC] points. Dashed lines - the 90–per cent confidence interval of dc for a single draw of unclustered points. Dotted lines - one hundred d_c-series obtained from the shuffled [DT, DR, MC] triples. Arrows - the points on dc(t) used for an insight into the evolution of earthquake clustering.

From the empirical distribution of {d_c|*random sampling of [DT, DR, MC] triples*} obtained from 100,000 draws, we established the 0.005 and 0.995 percentiles to be $p_{0.005}$=0.6141 and $p_{0.995}$=0.7041. In the $d_c(t)$ series, seven values were greater than $p_{0.995}$ and eight values were less than $p_{0.005}$. These results yielded highly low probabilities, $\Pr(48,7,p_{0.995}) = 4.81 \times 10^{-9}$ and $\Pr(48,8,p_{0.005}) = 1.23 \times 10^{-10}$, for the thesis that the peaks of $d_c(t)$ were due to the statistical scatter of d_c obtained from random triples [*DT, DR, MC*].

However, due to the overlaps of windows in which we calculated $d_c(t)$, these d_c-s were correlated. Consequently, the probabilities $\Pr(48,7,p_{0.995})$ and $\Pr(48,8,p_{0.005})$ were not as from Eq. (2). To study the consequences of the window overlaps, we used the one hundred $d_c(t|$*shuffled series of [DT, DR, MC] triples*) series (see: Data and Methods), counting in each series the number of times its values were greater than $p_{0.995}$ / less than $p_{0.005}$. Proportions of the $d_c(t|$*shuffled series of [DT, DR, MC] triples*) series in which d_c–values were greater/less than $p_{0.995}/p_{0.005}$ k=1, 2, 3, 4, 5 times are compared with the respective probabilities from Eq. (2) in Table 1.

For k>1, the non-zero proportions are greater than the respective probabilities for truly randomly drawn triples [*DT, DR, MC*]. A very rough estimate of a multiplication factor of this probability increase is 10^{k-1}. If so, the probability of a random origin of the number of times that $d_c(t)$ exhibits the values above/below $p_{0.995}/p_{0.005}$ would still be on the order of 10^{-3}. Furthermore, the maximum number of instances where the values of $d_c(t|$*shuffled series of [DT, DR, MC]*

186 *triples*) series deviated from the interval [$p_{0.005}$, $p_{0.995}$] was six. In contrast, the original $d_c(t)$
187 series had fifteen instances where its values fell outside this interval.
188

Table 1. Comparison of the shuffled series proportions, in which d_c-values k times were outside $[p_{0.005}, p_{0.995}]$, with the probability of such cases according to Eq. (2).

k	Proportion of *dc(t|shuffled series of [DT, DR, MC] triples)* in which at least *k*-times:		$\Pr(48, k, p_{0.995}) = \Pr(48, k, p_{0.005})$ (Eq. (2))
	d_c – values ≥ $p_{0.995}$	**d_c – values ≤ $p_{0.005}$**	
1	0.17	0.22	0.214
2	0.08	0.06	0.024
3	0.02	0.02	0.0018
4	0.02	0	0.0001
5	0	0	0.000004

We also took every fifth window. Such windows do not overlap and the probability that random scatter caused d_c-values to be outside the range is as in Eq. (2). The respective $d_c(t)$ series consisted of ten elements. In this series, there were two d_c-s greater than $p_{0.995}$ and two d_c-s less than $p_{0.05}$. According to Eq. (2), the probability to get such results at random was $\Pr(10,2,p_{0.995}) = \Pr(10,2,p_{0.005}) = 0.0011$.

Not only the extreme amplitudes, non-random as shown above, are characteristic of $d_c(t)$, but also the shape of $d_c(t)$ from March 2007 to the mainshock. This shape is compared in Figure 2 with the shapes of the one hundred $d_c(t|$*shuffled series of [DT, DR, MC] triples*$)$ series (dotted lines). The $d_c(t)$ series clearly differs from all $d_c(t|$*shuffled series of [DT, DR, MC] triples*$)$ series in the trends' lengths and widths of peaks.

The minimum achievable value of d_c - the average distance between earthquakes in {*DT, DR, MC*} space, is zero. In such a case, the earthquake process is perfectly regular: interevent time and distance are constant, and magnitudes are the same. Earthquakes are clustered in one point in the {*DT, DR, MC*} space. In this context, a d_c increase reflects the diminishing regularity of the earthquake process.

Having 8x*N* points in the [0, 1] {*DT, DR, MC*} cube, d_c becomes maximum if there are *N* points at each cube vertex. For 96 points (8x12), this maximum of d_c=1.13364. For 100 points, the d_c takes the maximum when 96 points are distributed among eight vertices of the cube in equal parts, and the separation of the remaining four points is maximum, e.g., they are located at (0,0,0), (1,1,1), (0,1,0), (1,0,1) or the like. For that distribution, d_c=1.13303. Hence in case of one hundred earthquakes windows, an increase of d_c from zero to 1.13303 means moving from a one-point cluster to eight best-balanced and maximally distant one-point clusters. On the way, there is a perfectly unclustered state.

Points-earthquakes are unclustered in [0, 1] cube when their components, [*DT, DR, MC*], are drawn from the three–dimensional standard uniform distribution. To get properties of the unclustered state, we drew one million samples, one hundred elements each, from this distribution. Based on this data, we have established that:

- The mean mass center of the unclustered points is at (0.5, 0.5, 0.5), and the standard deviation of this location is (0.0289, 0.0289, 0.0289) with a precision of 10^{-5};

- The mean sum of distances from such one hundred points to their mass center is 48.03, with a standard deviation of 1.39;
- The mean value of one million d_c values is $<d_c>=0.661714$, and the standard deviation is $\sigma=0.017731$. Hence, the 90–percent confidence interval of a single draw is (0.626960; 0.696467). The mean value and the 90–percent confidence interval limits for a single draw are marked in Figure 2 by the horizontal solid and two dashed lines, respectively.

The part of the $d_c(t)$ curve from March 2007 to the mainshock (8/09/2017) begins significantly below $<d_c>$, then crosses $<d_c>$ and increases to values significantly above $<d_c>$ (Figure 2). Next, it goes down again below $<d_c>$. This behavior suggests that the earthquakes in windows linked to smaller $d_c(t)$ values may have tended to form one cluster, and those in windows linked to greater $d_c(t)$ values may have built more than one cluster.

To look closer at the evolution of earthquake clustering, we took the data from six data windows linked to the points marked by arrows in Figure 2. Out of these six, only two windows, No.s 4 and 5, partially overlapped. We estimated the optimal number of clusters in every window. Because the problem of dividing points into clusters is equivocal, different optimality criteria provide different clustering results. Out of the five tested criteria (See: Supporting Information Text S2 for the choice of the optimal number of clusters criterion), we used the gap criterion with the PCA reference data generation method (Tibshirani et al., 2001). The accepted "optimal" numbers of data clusters in the studied windows, the division of earthquakes in the windows into the clusters, and the coordinates of the resultant clusters' centroids are shown in Table 2.

The window #1 mass center was at (0.51, 0.54, 0.45). Hence, the probability that the deviations of window #1 mass center components from 0.5 were due only to statistical scatter are 0.365, 0.083, 0.042, respectively, and the probability that the mass center deviated from the point (0.5, 0.5, 0.5) due to statistical scatter is 1.3×10^{-3}. The total distance of the points in windows #1 to their mass centers was 43.18. The probability that a single draw of one hundred unclustered points would provide this distance value is 2.4×10^{-4}.

The window #6 mass center was at (0.45, 0.46, 0.53). The probabilities that only statistical scatter caused the deviations of mass center components from 0.5 and the mass center deviation from (0.5, 0.5, 0.5) are 0.042, 0.083, 0.150, and 5.2×10^{-4}, respectively. The total distance of window #6 points to the mass center was 45.85. The probability that such a result would be obtained for an unclustered points draw is 0.058.

These results indicate that the points in windows #1 and #6 were not unclustered but built single clusters. This fact and the results from Table 2 suggest that the up-down trends of $d_c(t)$ preceding the Tehuantepec mainshock resulted from the following earthquake clustering process in the {*DT*, *DR*, *MC*} [0, 1] cube. At the beginning (18/08/2005 – 20/03/2007), the earthquakes formed one cluster. The decay of this cluster in separate groups – four in windows #2 and #3 increased $d_c(t)$. The maximum $d_c(t)$ was linked to separating of one hundred earthquakes into as many as six clusters. Next, the number of clusters was gradually reduced – four clusters in window #5 and only one in the last window before the mainshock (window #6). This decrease in clusters caused the systematic decrease of $d_c(t)$. The centroids' coordinates from Table 2 suggest that this clustering process could schematically look like in Figure 3.

***Table 2**. Results of the division of earthquakes in windows into the optimal number of clusters. Centroid coordinates for 'all data' cases refer to the mass centers of all points in windows.*

No.	Window time point, t	$d_c(t)$	Optimal number of clusters	Cluster No	Centroid coordinates			Number of cluster members
					DT	*DR*	*MC*	
1	20/03/2007	0.601	1	I	0.51	0.54	0.45	100
2	9/10/2008	0.636	4	I	0.15	0.19	0.37	25
				II	0.75	0.34	0.34	21
				III	0.41	0.77	0.17	25
				IV	0.38	0.59	0.73	29
3	4/07/2013	0.696	4	I	0.15	0.17	0.72	29
				II	0.63	0.59	0.22	22
				III	0.83	0.30	0.79	24
				IV	0.55	0.80	0.77	25
4	28/06/2015	0.733	6	I	0.40	0.16	0.25	13
				II	0.34	0.86	0.24	14
				III	0.16	0.18	0.73	21
				IV	0.85	0.71	0.27	16
				V	0.76	0.36	0.82	22
				VI	0.52	0.85	0.85	14
5	27/04/2016	0.685	4	I	0.53	0.21	0.22	20
				II	0.16	0.20	0.74	19
				III	0.52	0.77	0.22	29
				IV	0.67	0.55	0.82	32
6	7/09/2017	0.639	1	I	0.45	0.46	0.53	100

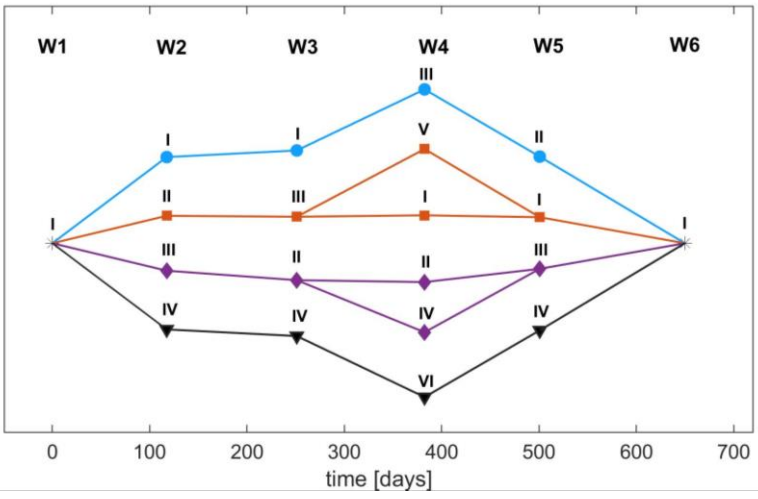

Figure 3. The proposed scheme of the earthquake clustering process preceding the Tehuantepec M8.2 mainshock. W1-W6 mark windows, and roman numbers denote the windows' clusters as in Table 2. The windows locations on time axis agree with periods between times of the windows' middles. The vertical separations of the nodes/clusters in the same window are proportional to the distance between centroids of neighboring nodes.

4 Conclusions

The re-parameterization of seismicity into interevent times and distances and the use of the transformation to equivalent dimensions to ensure comparability of parameters shed new light on the preparatory process to the Tehuantepec earthquake. The average distance between earthquakes in the space of equivalent: interevent time, interevent distance, and magnitude time series exhibited distinct and systematic time changes. Its part from March 2007 until the mainshock (9/09/2017) was a statistically significant precursory signal. This signal included a well-developed rise from March 2007 to June 2015, followed by a decrease until the mainshock. These premonitory changes were linked to the earthquake clustering process evolution. Further research across other great earthquake cases will determine the usefulness of the approach presented here in the earthquake forecasting problem.

Acknowledgments

The work was supported by a subsidy from the Polish Ministry of Education and Science for the Institute of Geophysics, Polish Academy of Sciences, and Project no. 2021/WK/09 subsidized by the Polish Ministry of Education and Science.

Open Research

Data used in this work is the catalogue from the Mexican Seismological Service of Mexico, SSN (2023): Universidad Nacional Autónoma de México, Instituto de Geofísica, Servicio Sismológico Nacional, México: http://www2.ssn.unam.mx:8080/catalogo/. The transformation to equivalent dimensions was done using the Transformation to Equivalent Dimensions application (Leptokaropoulos et al., 2020), available as an application implemented on the EPISODES platform - https://EpisodesPlatform.eu (Orlecka et al., 2020).

References

Baiesi, M., & Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. Physics Review E, 69, 066106. doi: 10.1103/PhysRevE.69.066106

Barani, S., Mascandola, C., Riccomagno, E., Spallarossa, D., Albarello, D., Ferretti, G., Scafidi, D., Augliera, P., & Massa, M. (2018). Long-range dependence in earthquake-moment release and implications for earthquake occurrence probability. Scientific Reports, 8, 5326. doi: 10.1038/s41598-018-23709-4.

Barani, S., Cristofaro, L., Taroni, M., Gil-Alaña, L.A., & Ferretti, G. (2021). Long memory in earthquake time series: the case study of the Geysers Geothermal Field. Frontiers of Earth Science, 9, 563649. doi: 10.3389/feart.2021.563649

Batac, R.C., & Kantz, H. (2014). Observing spatio-temporal clustering and separation using interevent distributions of regional earthquakes. Nonlinear Process in Geophysics, 21, 735–744. doi: 10.5194/npg-21-735-2014

Batac, R.C., Paguirigan Jr., A.A., Tarun, A.B., & Longjas, A.G. (2017). Sandpile-based model for capturing magnitude distributions and spatiotemporal clustering and separation in regional earthquakes. Nonlinear Process in Geophysics, 24, 179–187. doi: 10.5194/npg-24-179-2017

Calinski, T., & Harabasz, J. (1974). A Dendrite Method for Cluster Analysis. Communications in Statistics. Theory and Methods, 3, 1-27. doi: 10.1080/03610927408827101

Chen, Ku-Cheng., Kim, K-H., Wang, J-H., & Chen, Kuo-Chang (2022). Dominant periods and memory effect of the 2021 earthquake swarm in Hualien, Taiwan. Terrestrial Atmospheric Oceanic Sciences, 33, 24, doi: 10.1007/s44195-022-00022-2

Corral, Á. (2004). Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. Physics Review Letters, 92 (10), 108501. doi:10.1103/PhysRevLett.92.108501

Davies, D.L., & Bouldin, D.W. (1979). A Cluster Separation Measure. in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-1, no. 2, 224-227, doi: 10.1109/TPAMI.1979.4766909.

Gardner, J.K., & Knopoff, L. (1974). Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? Bulletin Seismological Society America, 64, 1363-1367. doi: 10.1785/BSSA0640051363

Gkarlaouni, C., Lasocki, S., Papadimitriou, E., & Tsaklidis, G. (2017). Hurst analysis of seismicity in Corinth rift and Mygdonia graben (Greece). Chaos, Solitons and Fractals, 96, 30–42. doi: 10.1016/j.chaos.2017.01.001

Hurst, H. E. (1951). Long-Term Storage Capacity of Reservoirs. Transactions American Society Civil Engineers, 116, 770-799. doi: 10.1061/TACEAT.0006518

Kagan, Y.Y., & Jackson, D.D. (1991). Long-term earthquake clustering. Geophysical Journal International, 104, 117-133. doi: 10.1111/j.1365-246X.1991.tb02498.x

Kagan, Y.Y., & Knopoff, L. (1981). Stochastic synthesis of earthquake catalogs. Journal Geophysical Research, 86 (B4), 2853-2862. doi: 10.1029/JB086iB04p02853

Kelleher, J. A., & McCann, W. R. (1976). Buoyant zones, great earthquakes, and unstable boundaries of subduction. Journal Geophysical Research, 81, 4885–4896. doi: 10.1029/JB081i026p04885

Kelleher, J. A., Sykes, L. R., & Oliver, J. (1973). Possible criteria for predicting earthquake locations and their applications to major plate boundaries of the Pacific and Caribbean region. Journal Geophysical Research, 78, 2547–2585. doi: 10.1029/JB078i014p02547

Lasocki, S. (2007). Evidences of complexity of magnitude distribution obtained from a non-parametric testing procedure, in Proceedings of the 5th International Workshop of Statistical Seismology “Physical and Stochastic Modelling of Earthquake Occurrence and Forecasting” (StatSei 5), Erice, Italy.

Lasocki, S. (2014). Transformation to equivalent dimensions – a new methodology to study earthquake clustering. Geophysical Journal International, 197 (2), 1224-1235. doi: 10.1093/gji/ggu062

Lasocki, S. (2021). Kernel Density Estimation in Seismology. Chapter 1 in Statistical Methods and Modelling of Seismogenesis, Limnios, N., Papadimitriou, E., Tsaklidis, G., eds., ISTE, Wiley, pp. 1-26.

Lasocki, S., & Papadimitriou, E. E. (2006). Magnitude distribution complexity revealed in seismicity from Greece. Journal Geophysical Research, 111, B11309, doi: 10.1029/2005JB003794

Lennartz, S., Livina, V.N., Bunde, A., & Havlin, S. (2008). Long-term memory in earthquakes and the distribution of interoccurrence times. Earth Planetary Letters, 81, 69001. doi: 10.1209/0295-5075/81/69001

Leptokaropoulos, K., Lasocki, S., & Urban, P. (2020) Transformation to Equivalent Dimensions (version September 2020) [Software]. https://episodesplatform.eu/?lang=en#app:T2ED

Lippiello, E., de Arcangelis, L., & Godano, C. (2008). Influence of time and space correlations on earthquake magnitude. Physics Review Letters, 100, 038501, doi: 10.1103/PhysRevLett.100.038501

Liu, CH., Liu, YG., & Zhang, J. (1995). R/S analysis of earthquake time interval. Acta Seismologica Sinica, 8, 481–485. doi: 10.1007/BF02650577

Livina, V. N., Havlin, S., Bunde, A. (2005). Memory in the occurrence of earthquakes. Physics Review Letters, 95, 208501. doi: 10.1103/PhysRevLett.95.208501

Mandelbrot, B., & Wallis, J. (1969). Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. Water Resources Research, 5, 967-988, doi: 10.1029/WR005i005p00967.

Melgar, D., Ruiz-Angulo, A., Garcia, E. S., Manea, M., Manea, V. C., Xu, X., et al. (2018). Deep embrittlement and complete rupture of the lithosphere during the M8.2 Tehuantepec earthquake. Nature Geoscience. doi: 10.1038/s41561-018-0229-y

Michas, G., Kapetanidis, V., Kaviris, G., & Vallianatos, F. (2021). Earthquake diffusion variations in the Western Gulf of Corinth (Greece). Pure Applied Geophysics, 178, 2855–2870. doi: 10.1007/s00024-021-02769-0

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. Journal of the American Statistical Association, 83 (401), 9-27. doi: 10.1080/01621459.1988.10478560

Orlecka-Sikora B., Lasocki S., Kocot J., Szepieniec T., Grasso J-R., Garcia-Aristizabal A., et al., (2020). An open data infrastructure for the study of anthropogenic hazards linked to georesource exploitation. Scientific Data 7, 89. doi:10.1038/s41597-020-0429-3

Rousseeuw, P.J. (1987) Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53-65. doi: 10.1016/0377-0427(87)90125-7

Silverman, B. W. (1986). Density Estimation for Statistics and Data Analysis. Chapman and Hall, London, pp. 176.

Sornette, D., & Helmstetter, A. (2002). Occurrence of finite-time singularities in epidemic models of rupture, earthquakes, and starquakes. Physics Review Letters, 89 (15), 158501. doi: 10.1103/PhysRevLett.89.158501

Suárez, G., Santoyo, M. A., Hjorleifsdottir, V., Iglesias, A., Villafuerte, C., & Cruz–Atienza, V. M. (2019). Large scale lithospheric detachment of the downgoing Cocos plate: The 8 September 2017 earthquake (Mw8.2). Earth Planetary Science Letters, 509, 9–14. doi: 10.1016/j.epsl.2018.12.018

Tibshirani, R., Walther, G. & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. Journal Royal Statistical Society: Series B (Statistical Methodology), 63, 411-423. doi: 10.1111/1467-9868.00293

Węglarczyk, S., & Lasocki, S. (2009) Studies of short and long memory in mining-induced seismic processes. Acta Geophysica, 57, 696-715. doi: 10.2478/s11600-009-0021-x

Ye, L., Lay, T., Bai, Y., Cheung, K. F., & Kanamori, H. (2017). The 2017 Mw 8.2 Chiapas, Mexico, earthquake: Energetic slab detachment. Geophysical Research Letters, 44, 11,824–11,832. doi: 10.1002/2017GL076085

Zaliapin, I., & Ben-Zion, Y. (2013), Earthquake clusters in southern California I: Identification and stability. Journal Geophysical Research, 118, 2847–2864. doi: 10.1002/jgrb.50179

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., & Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. Physics Review Letters 101, 018501. doi: 10.1103/PhysRevLett.101.018501

Zhang, Y., Fan, J., Marzocchi, W., Shapira, A., Hofstetter, R., Havlin, S., & Ashkenazy, Y. (2020). Scaling laws in earthquake memory for interevent times and distances. Physis Review Research, 2, 013264. doi: 10.1103/PhysRevResearch.2.013264

Zúñiga, F. R., Reyes, M. A., & Valdés, C. (2000). A general overview of the catalog of recent seismicity compiled by the Mexican Seismological Survey. Geofísica Internacional, 39, 161–170. doi: 10.22201/IGEOF.00167169P.2000.39.2.273