The magnitude threshold and missing and pseudo links in Markov chains

Fidencio Alejandro Nava (fnava@cicese.mx)
Center for Scientific Research and Higher Education at Ensenada
Quiriat Jearim Gutiérrez
Center for Scientific Research and Higher Education at Ensenada

Research Article

Keywords: Markov process, magnitude threshold, fuzzy thresholds, seismic hazard

Posted Date: November 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2282011/v1

The magnitude threshold and missing and pseudo links in Markov chains.

F. A. Nava, Q. J. Gutiérrez

Fidencio Alejandro Nava

Quiriat Jearim Gutiérrez

Centro de Investigación Científica y de Educación Superior de Ensenada, B. C.

Seismology Department

Ensenada, Baja California, Mexico.

Corresponding author: F. A. Nava fnava@cicese.mx

F. A. Nava ORCID: 0000-0001-6778-2017

Q. J. Gutiérrez ORCID: 0000-0002-1432-9507

Keywords: Markov process, magnitude threshold, fuzzy thresholds, seismic hazard

Acknowledgements

We are grateful to the International Seismological Centre for the use of their catalog.

Abstract

A crisp step function is not an adequate threshold for studies of Markovian occurrence of large earthquakes, because it can lead to missing or pseudo links in an observed sequence that should be a Markov chain. A more realistic threshold is a fuzzy one where there is a magnitude band, located between those magnitudes that are too small for the earthquakes to be part of a Markovian process and those who are certainly large enough for the earthquakes to be part of it, where earthquakes may or may not be part of the process. This fuzzy threshold is described by a membership function that gives the probability of an earthquake with a given magnitude belonging to the process. We propose a counting strategy to evaluate the empirical transition probabilities when considering a fuzzy magnitude threshold, and justify it through Monte Carlo simulations.

We also propose a membership function with probabilities in the transition band proportional to the seismic moment, and apply it to data from a seismic study of the Japan area. By comparing the results of three Markovianity measures for the observed data with those of Monte Carlo simulations, we find that a fuzzy threshold going from zero probability for magnitudes $M \leq 6.9$ to probability one for $M \geq 7.2$ is the best model for the study area.

1. Introduction

The Heaviside step function is a mathematical concept and is not observed in natural phenomena, yet stepwise magnitude thresholds are a common feature of many seismological studies and in most cases they play a quasi-magical role: earthquakes with magnitudes above the threshold are the subject of the study and they are the source of information, while those with magnitudes below the threshold are discarded, which implies that they contain no valuable information. It would seem that earthquakes with magnitudes above the threshold are a different physical phenomenon from those with magnitudes below it, and yet in many cases the magnitude threshold is set because of data limitations (data may be incomplete for magnitudes below a given threshold, or not enough if only magnitudes above some other given threshold are considered), and in some cases may be set completely arbitrarily.

Another complication associated with step-wise thresholds is that seismic magnitude determination is subject to a large uncertainty due to factors such as radiation pattern, directivity, paths through different media, and many others, including measuring and even numerical errors; this uncertainty is the reason behind the customary rounding of magnitudes to $\Delta M = 0.1$. And rounding complicates the issue, because an earthquake with unrounded magnitude 6.94999... (rounded to 6.9) would be discarded by a step-wise threshold $M \geq 7.0$, yet it cannot be physically very different from one with unrounded magnitude 6.9500... (rounded to 7.0). Considering the uncertainties, an earthquake with rounded magnitude $\widetilde{M}$ (the tilde will denote rounded magnitudes henceforward) cannot, generally, be certainly very different from earthquakes with magnitudes $\widetilde{M} \pm 0.1$.

Also, when considering seismic magnitudes it is necessary to take into account that sometimes an earthquake with magnitude below the threshold may be accompanied, close in time and space, by large fore- and/or aftershocks having seismic moment releases that summed to that of the main event result in a moment release episode (Quinteros et al., 2014) with equivalent magnitude above the threshold, that should be considered as an event to be included among those earthquakes having magnitudes above it.

Hence, the commonly used Heaviside function threshold is clearly seen to be inappropriate when applied to magnitudes in seismic studies.

We will here consider the problem of magnitude thresholds as applied to Markovian studies of seismic hazard, because many seismic hazard studies consider Markovian systems (e.g. Nishioka and Shah, 1980; Patwardhan et al., 1980; Anagnos and Kiremidjian, 1988; Fujinawa, 1991; Alvarez, 2005; Nava et al., 2005; Herrera et al., 2006; Votsi et al., 2010, 2013; Ünal and Celebioglu, 2011; Cavers and Vasudevan, 2014, 2015).

The motivation for our study was an article by Gutiérrez et al. (2021) who made a Markovian study of the seismic hazard in an area around Japan, and we will use their results and data to illustrate our arguments.

Their system consists of four seismogenic regions (Figure 1), and is considered to change to a new state whenever a "large" earthquake, i.e. an earthquake with magnitude above a given threshold magnitude M^T, occurs within one of the regions, and the state corresponds to the region it occurred in, so that the system has $N_s = 4$ states. The data set comprises $N = 450$ earthquakes with magnitudes in the $6.5 \leq M \geq 9.2$ range, occurred from June 2, 1905 to November 13, 2015, reported in the International Seismological Centre (ISC) catalog. The seismotectonics of the area and the reasons for choosing the four regions, are discussed in Gutiérrez et al. (2021), so that they will not be discussed here.

The system was expected to be Markovian based on the premise that large earthquakes liberate enough stress and strain to significantly modify the stress field in the area and locally influence the tectonic plates motion and, hence, the occurrence of future earthquakes in neighboring areas (e.g. Lehner et al., 1981; Tsapanos and Papadopoulou, 1999; Márquez et al., 2002; Melbourne et al., 2002; Venkatamaran and Kanamori, 2004; Riga and Balocchi, 2016; Spagnotto et al., 2018).

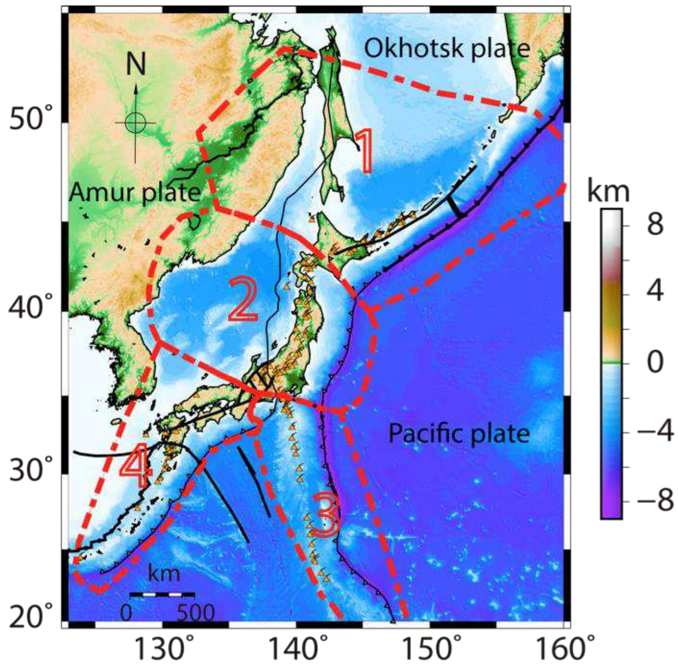

Fig. 1 Map of the study area. The dashed lines enclose the four regions corresponding to the four states of the system (modified from Gutiérrez et al. (2021))

Gutiérrez et al. (2021) obtained transition probability matrices for several threshold magnitudes using stepwise trial thresholds and, in order to evaluate the results for the different thresholds, they proposed measures of reliability, robustness and *Markovianity* (how much the hazard estimates differ from what would be expected from the stationary probabilities discussed below). They found that earthquakes with large magnitudes do occur in a Markovian way, and that Markovianity is larger for threshold magnitudes between 7.0 and 7.4, with a maximum around 7.1; they speculated that

Markovianity was low for magnitudes below 7.0 because data sets included too many events that were not large enough to influence the process and were not part of a Markovian chain, and above 7.4 because too many events that did influence the process were excluded.

Here, we propose that, even assuming magnitudes to be correctly estimated, earthquakes within a given magnitude range may or may not be part of a Markovian process, depending on the local distribution of stress and strength and, possibly, on characteristics of each particular rupture and other imponderables. Hence the set of Markovian earthquakes will be fuzzy (Zadeh, 1965; Aziz and Parthiban, 1996), and the threshold should be a gradual membership function (Zadeh, 1988) instead of a crisp, stepwise one. We propose a way to estimate empirical transition probabilities for fuzzy Markovian thresholds, and use Monte Carlo methods to test the estimation method.

2. Markovian systems and chains

We will now briefly review a few concepts of Markovian systems that will be useful for our study (extensive treatments are found in many texts, e.g. Parzen, 1960; Barucha-Reid, 1960; Gnedenko, 1962; Feller, 1968; Ching and Ng, 2006; Battaglia, 2007, and many more).

A *finite Markov process* is a stochastic process, with a finite number N_s of states, $\{s_k;\ k = 1, \cdots, N_s\},\ N_s > 1$, for which the probability of transition from the current state to a given state in the next trial depends only on the current state, i.e., the system has no memory about the states that occurred before the present one. Let the state at step *m* be $s_{[m]} = s_i$ and that at the next step be $s_{[m+1]} = s_j$, then if the transition probabilities between these states are always the same, independently of the trial or time interval, the Markov process is *homogeneous*, and we can write

$$\Pr[s_{[m+1]} = s_j \mid s_{[m]} = s_i] = \hat{p}_{ij} , \tag{1}$$

where each $\hat{p}_{ij}$ is an element of a square $N_s \times N_s$ transition probability matrix (TPM) $\hat{P}$ that contains the true transition probabilities of the Markovian system. A sequence of state occurrences distributed according to $\hat{P}$ is commonly referred to as a Markov chain, because each occurring state is linked to the previous one according to (1).

If the state at step *m* is $s_{[m]} = s_i$, then the probability that *n* steps later the system will be in state $s_{[m+n]} = s_j$, the *n*-step transition probability, is given by

$$\Pr[s_{[m+n]} = s_j \mid s_{[m]} = s_i] = \hat{p}_{ij}^{(n)}, \tag{2}$$

where $\hat{p}_{ij}^{(n)}$ is an element of matrix $\hat{P}^n$ (Chapmann-Kolmogorov equation).

If $\hat{P}$ is ergodic, then

$$\lim_{n\to\infty} \hat{P}^n = \hat{\Pi} = \begin{bmatrix} \hat{\pi} \\ \vdots \\ \hat{\pi} \end{bmatrix} \tag{3}$$

where $\hat{\pi} = [\hat{\pi}_1, \hat{\pi}_2, \cdots, \hat{\pi}_{N_s}]$, $\sum_{j=1}^{N_s} \hat{\pi}_j = 1$; the elements of $\hat{\pi}$, called *limiting* or *stationary probabilities*, are the same for any initial state and depend only on the total relative number of times each state occurs. Hence, the stationary probabilities are not Markovian, yet are representative of what would be expected if the system were non-Markovian, and will be useful for comparison with the *Markovianity* measures presented below.

The empirical Markovian transition probabilities from state *i* to state *j* are usually estimated as

$$p_{ij} = \frac{\theta_{ij}}{\xi_i}, \tag{4}$$

where θ_{ij} is the observed number of transitions from state *i* to state *j*, and

$$\xi_i = \sum_{j=1}^{N_s} \theta_{ij} \tag{5}$$

is the

total number of transitions that originated from state i. The empirical Markovian transition probabilities are expressed as a square $N_s \times N_s$ TPM P

3. Missing and pseudo links

Picture a geographic area under study, where earthquakes are being generated according to a Markovian process and the sequence of these earthquakes constitutes a Markovian chain; however, other earthquakes not forming directly part of the Markovian process are concurrently occurring in the area. Hence, the sequence of all earthquakes from the area will be a combination of events constituting the links of the chain, minus *missing links*, events missing from the chain possibly because they were discarded for being below some threshold or because of the uncertainties in magnitude determination, or for other reasons, plus *pseudo links* that are events extraneous to the chain.

Missing or pseudo links will influence the estimation of the Markovian transition probabilities as follows:

A missing link, has two effects: if a segment of the true chain is $\cdots, s_i, s_j, s_k, \cdots$, then when counting transitions both θ_{ij} and θ_{jk} will be correctly increased by 1 each, but if state s_j is missing, the resulting sequence is $\cdots, s_i, s_k, \cdots$, which will result in θ_{ik} incorrectly increased by 1, and in θ_{ij} and θ_{jk} incorrectly not increased.

On the other hand, if the true chain is $\cdots, s_i, s_k, \cdots$, then when counting transitions θ_{ik} will be correctly increased by 1, but if a pseudo state s_j occurs in between, the observed sequence will be $\cdots, s_i, s_j, s_k, \cdots$ and both θ_{ij} and θ_{jk} will be each incorrectly increased by 1, and θ_{ik} will be incorrectly not increased.

The error caused by missing (A) and pseudo (B) links for a typical $N_s = 4$ TPM and a $N = 450$ long series is illustrated in Figure 2; it is clear that the mean error per transition probability is significant even for small ratios of missing or pseudo links to the number of data, and that errors due to pseudo links are larger than those due to missing links.

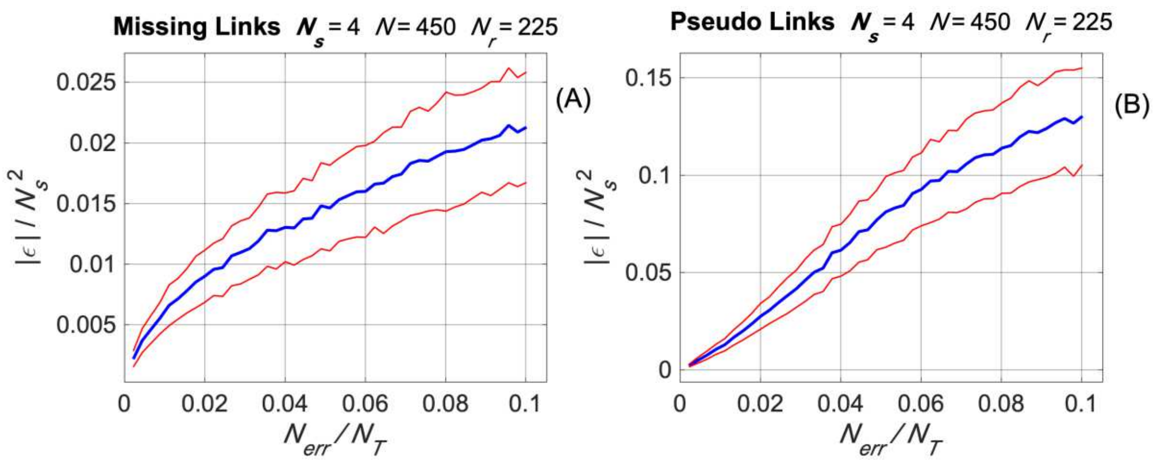

Fig. 2 Average error per transition probability vs. proportion of missing (A) or pseudo (B) links; the blue thick line is the mean error for 225 Monte Carlo realizations and the thin lines are the mean plus/minus one standard deviation

4. Fuzzy thresholds and the membership function

We will use fuzzy magnitude thresholds in order to model the mixture of Markovian and non-Markovian events that can be expected in real seismicity, and will propose a counting strategy, based on the fuzzy threshold, to diminish the effects of missing or pseudo links in the estimation of transition probabilities.

The magnitude fuzzy thresholds will be defined by a membership function $V(M)$ (Zadeh, 1988), which is the probability that an event with magnitude M belongs to the Markovian set.

Because magnitude data in seismic catalogs are usually rounded to $\Delta M = 0.1$, we will characterize all thresholds by two rounded magnitudes: $\widetilde{M}_{K0}$ and $\widetilde{M}_{K1}$, such that events with rounded magnitudes $\widetilde{M} \leq \widetilde{M}_{K0}$ are definitely too small to be part of a Markovian chain and events with rounded magnitudes $\widetilde{M} \geq \widetilde{M}_{K1}$ are definitely large enough to be links in a Markovian chain. Since from a physical point of view it is reasonable to define a membership function in terms of unrounded magnitudes, and since we will need such a function for the Monte Carlo simulation of catalogs with fuzzy thresholds, we note that the above limits correspond to unrounded limits $M_{K0}^{U} \equiv \widetilde{M}_{K0} +$

$\Delta M/2$ and $M_{K1}^{U} \equiv \widetilde{M}_{K1} - \Delta M/2$. These limits define a transition range where events with unrounded magnitudes $M_{K0}^{U} < M < M_{K1}^{U}$ may stochastically belong to the Markovian set, or not, with probability given by the membership function in that range.

The threshold bandwidth

$$\omega = (\widetilde{M}_{K1} - \widetilde{M}_{K0})/\Delta M - 1 = (M_{K1}^{U} - M_{K0}^{U})/\Delta M \tag{6}$$

is the number of M intervals spanned by the threshold. For $\widetilde{M}_{K1} = \widetilde{M}_{K0} + 0.1$, $\omega = 0$, the threshold becomes a "crisp" step function.

It is not known what $V(M)$ is like in the transition band, but to have a transition with a physical meaning, we heuristically propose that it is not unreasonable to suppose the membership probability to be proportional to the released seismic moment M_0, related to the moment magnitude M_W (Hanks and Kanamori, 1979) as

$$\log_{10} M_0 = 16.5 + 1.5 M_W \;\; ; \tag{7}$$

in what follows we will suppose that all magnitudes are moment magnitudes and represent them by simply M. Thus, the moment membership probability is given by

$$V(M) = \begin{cases} 0, M \leq M_{K0}^{U} \\ \dfrac{10^{1.5M} - 10^{1.5 M_{K0}^{U}}}{10^{1.5 M_{K1}^{U}} - 10^{1.5 M_{K0}^{U}}}; \;\; M_{K0}^{U} < M < M_{K1}^{U} \\ 1, M \geq M_{K1}^{U} \end{cases} \tag{8}$$

illustrated in Figure 3.

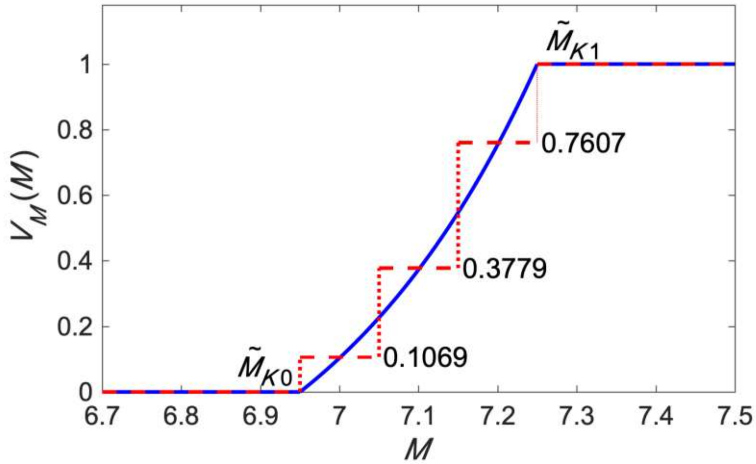

Fig. 3 The moment fuzzy threshold membership probability function with an $\quad = 3$ transition band. The dashed horizontal lines represent the average probabilities for each magnitude step in the transition range

Other possible membership functions may have physical significance, like a transition band proportional to magnitude, or could be just handy commonly used threshold functions, such as the widely-used S-shaped cosine-based function which gives a reasonable and smooth gradual threshold without discontinuities or sharp changes in slope. Different membership functions can be tried to find which one gives best results for a particular data set and thus represents best what is actually happening in a given region. Here we will use only the moment threshold membership function.

5. A counting strategy for fuzzy thresholds

Since when analyzing an observed sequence it is not known a priori which events with magnitudes in the transition range belong to the Markovian set, we will propose a way of reducing the effects of missing or pseudo links for a given trial threshold. Next, we will try different trial thresholds to find which gives the best results and thus represents best what is the actually threshold. We will now present our counting strategy and later discuss how to measure and compare results.

We assume that the magnitude of each event in a catalog has been rounded to ΔM, and let the rounded magnitude be denoted by $\widetilde{M}$; since the original unrounded magnitude can have any value in the $[\widetilde{M}-\Delta M/2, \widetilde{M}+\Delta M/2)$ range we will consider the probability of and event with magnitude $\widetilde{M}$ belonging to the Markovian set to be the average of the probabilities within the range

$$v(\widetilde{M}) = \frac{1}{\Delta M} \int_{\widetilde{M}-\Delta M/2}^{\widetilde{M}+\Delta M/2} V(M)\, \mathrm{d}M \,, \tag{9}$$

shown as dashed horizontal lines in Figure 3. Thus, the *m*'th event of the sequence will be characterized by its observed state $s_{[m]}$, its rounded magnitude $\widetilde{M}_{[m]}$, and its Markovian membership probability $v_{[m]} = v(\widetilde{M}_{[m]})$.

The counting strategy for events with rounded magnitudes consists of four points:

I. Each transition between consecutive events will be counted according to the joint probability of both starting and ending states belonging to the Markovian set

$$w_{[m,m+1]} = v_{[m]} v_{[m+1]}. \tag{10}$$

Obviously only events with non-zero probabilities need be considered.

II. When $v_{[m+1]} = 1$ the transition is achieved, whatever the value of $w_{[m,m+1]}$, but if $v_{[m+1]} < 1$ then there is the possibility that state $s_{[m+1]}$ is not a true link, and that the Markovian transition is really from state $s_{[m]}$ to state $s_{[m+2]}$ with probability

$$w_{[m,m+2]} = v_{[m]}(1-v_{[m+1]})v_{[m+2]} \,,$$

because for this transition to be a link it is necessary to include the probability that event $m+1$ is not Markovian and, hence, not a link.

In general, transitions between non-consecutive states will have a probability given by the product of probabilities of the first and last events times the non-occurrence probabilities of all intermediate events

$$w_{[m,m+n]} = v_{[m]}(1-v_{[m+1]}) \cdots (1-v_{[m+n-1]})v_{[m+n]}. \tag{11}$$

It is clear from (11) that there cannot be a path that has a true link as an intermediate event.

III. The transition probabilities of all possible paths leading to an event with unitary probability should be considered and counted, because it is not known which was the one that actually was followed; there are 2^{n-1} different possible paths from an initial state s_m to a final state s_{m+n}.

IV. Each transition from a state $s_{[m]}$ to a state $s_{[m+n]}$ will be counted as

$$\theta_{s_{[m]}s_{[m+n]}} = \theta_{s_{[m]}s_{[m+n]}} + w_{[m,m+n]}. \tag{12}$$

Now θ_{ij} will not be the number of observed transitions between states i and j, it will be the sum of the probabilities corresponding to each observed possible transition from state i to state j, calculated according to (11). After all transitions have been taken into account, the transition probabilities will be estimated according to (4).

6. Monte Carlo validation of the counting strategy for fuzzy thresholds and exploration of three Markovianity measures

We will use Monte Carlo methods in order to explore the effects of a fuzzy threshold on Markovian studies, by generating a large number, N_r, of realizations of synthetic catalogs consisting of sequences of "events" each one characterized by a magnitude and a state.

As mentioned above, we will illustrate the application of fuzzy thresholds and the counting scheme using the same Markovian system used by Gutiérrez et al. (2021), which consists of $N_s = 4$ states, a sequence of $N = 450$ earthquakes with $6.5 \leq M \geq 9.2$. From all earthquakes reported by the same ISC catalog for the same period mentioned before, we determined a G-R b-value $b = 0.928$ for use in (14).

For each realization, we first generate a sequence of N Gutenberg-Richter distributed unrounded magnitudes,

$$\log_{10} N^C(M) = a_0 - b(M - M_{K0}^U), \tag{13}$$

where $N^C(M)$ is the number of earthquakes with magnitudes $\leq M$, so that

$$\Pr(M) = \beta \mathrm{e}^{-\beta(M-M_{K0}^U)};\ \ M \geq M_{K0}^U\,, \tag{14}$$

where $\beta = b \ln 10$.

Events with unrounded magnitudes below the transition range are considered to be non-Markovian, and events with unrounded magnitudes above the transition range are automatically included in the Markovian set.

For each unrounded magnitude in the transition range a uniformly distributed pseudo-random number in the (0,1) range is generated using the Matlab rand.m function; if according to the membership function (8) the magnitude's probability of belonging to the Markovian set is greater than the random number, then the corresponding event is accepted as belonging to the Markovian set; otherwise it is deemed non-Markovian.

Non-Markovian events are randomly assigned any state, with uniform probability.

Each event in the Markovian set is assigned a state according to the state of the previous Markovian event and a postulated "true" TPM $\hat{P}$, so that the Markovian events in the set constitute a Markov chain (the state before the first Markovian event is chosen randomly). We present results using as $\hat{P}$ for each postulated ω and M_{K0} the corresponding observed P for the Gutiérrez et al. (2021) data set, because we felt they would be the more appropriate to interpret the real data analysis and were adequate to test the counting strategy. The counting strategy was also tested using other $\hat{P}$ TPMs with uniformly satisfactory results.

When constructing a synthetic catalog it is known which events are Markovian, so for each realization a reference "empirical" TPM P_R is built according to (4) using only and all events in the Markovian chain. According to Borel's law of large numbers, P_R should tend to $\hat{P}$ when the chain length tends to infinity but, due to the relatively short span covered by seismic catalogues, observed sequences of large earthquakes cannot be very long, so that P_R will not, in general, equal $\hat{P}$, yet it is the best possible estimation obtainable from a given sequence, since it includes all events in the Markovian set and no events outside it. Hence, P_R will be used as the reference to evaluate the performance of the counting strategy that does not "know" which events in the transition band are links and which are not.

Finally, magnitudes are rounded to one decimal place, as usual in seismological catalogs, and one realization of a synthetic catalog is ready to be analyzed in the same way as a real catalog.

For each model, consisting of given bandwidth ω, $\widetilde{M}_{K0}$, and "true" TPM $\hat{P}$, $N_r = 1{,}000$ realizations of a synthetic catalog were generated, each realization consisting of $N = 450$ events with magnitudes that after rounding were in the [6.5, 9.2] range; the upper limit is the magnitude of the great 11 March 2011 Tohoku earthquake, the largest earthquake recorded in the area. Each realization was analyzed, using the counting strategy described above, for each of the trial bandwidths $\omega^T = 0,1,2,3$, and for trial lower bounds $\widetilde{M}_{K0}^T = 6.4, 6.5, \cdots, 7.0, 7.1$.

For each realization, the "observed" TPM P was estimated according to the counting strategy, and the root-mean-square difference between P and P_R, $\Delta_{rms}P$, was estimated; this difference is a measure of how well the counting strategy is performing.

Three of the Markovianity measures used by Gutiérrez et al. (2021) were also evaluated for each realization, to test their usefulness in identifying characteristics of the model. The considered measures are: M_6, which is the power to which a TPM has to be elevated to converge to stationary state probabilities, ie. to have all rows equal, to 6 decimal places; the difference between the Shannon (1948) entropies of the system and of the reference probability distribution $\Pi = [\pi_1, \pi_2, \cdots, \pi_{N_s}]^T$ (3), defined as

$$S = S^P - S^\Pi, \tag{15}$$

where

$$S^P = -\sum_{i=1}^{N_s} \sum_{j=1}^{N_s} p_{ij} \log_2 p_{ij} \tag{16}$$

and

$$S^P = -\sum_{i=1}^{N_s}\sum_{j=1}^{N_s} p_{ij}\log_2 p_{ij} \;\; ; \tag{17}$$

and the mean Relative Entropy or Kullback-Leibler Distance (Kullback and Leibler, 1951), between the probability distribution and the reference probability distribution Π:

$$\kappa = \frac{1}{N_s}\sum_{i=1}^{N_s}\sum_{j=1}^{N_s} p_{ij}\log_2\left(\frac{p_{ij}}{\pi_j}\right) \tag{18}$$

which is always well-defined because Π has no null elements; $\kappa = 0$ when P and Π are identical. The other two measures in Gutiérrez et al. (2021) were not evaluated here because they are closely related to S and κ.

Mean values of the above mentioned measures, obtained from the analysis of the N_r realizations, are used to see how well the Markovianity measures identify the characteristics of the different thresholds.

Here, for reasons of space, we will show in Figures 3 to 6 only the results for the postulated thresholds that best agree with the results for the real data. In each figure the postulated threshold is shown at top left in (a) and other graphs show mean values of the results of the N_r realizations for each trial M_{K0}^T; graph (b) shows $\Delta_{rms}P$, and graphs (A) to (D) show measures corresponding to those resulting from the analyses of real data: N_{K0}, M_6, S, and κ, respectively.

Graph (b) is very important, because it shows in Figures 4 to 6 that for all actual fuzzy thresholds, the counting strategy, applied to the rounded magnitudes series, for which it is not known which events are links in the chain and which are not, results, for the right combination of ω and M_{K0}^T, in the best mean estimate, i.e. the one closest to the optimal, which indicates that this strategy is indeed working as it should.

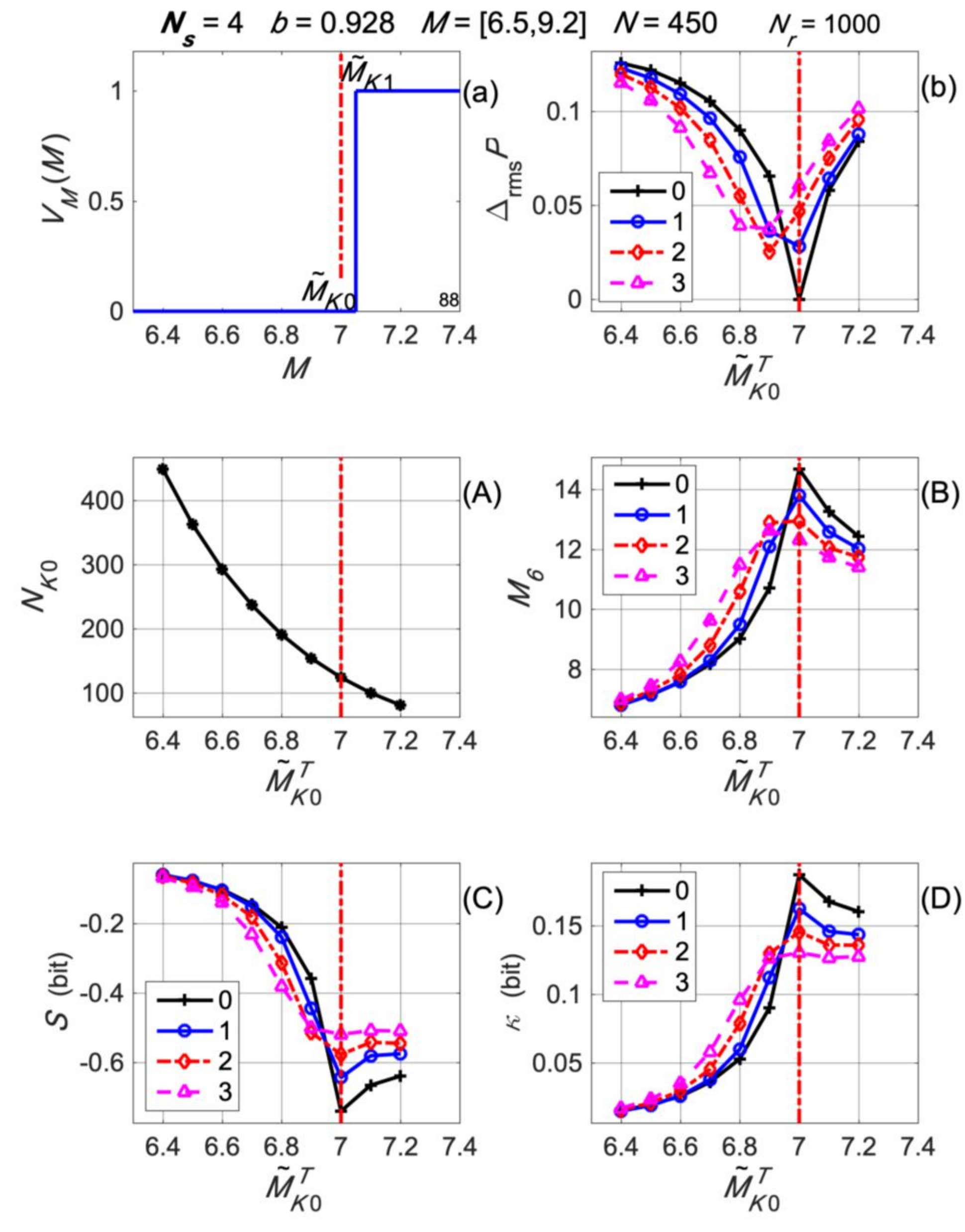

Fig. 4 Results of Monte Carlo simulation of a crisp, $\omega = 0$ threshold at $\widetilde{M}_{K0} = 7.0$ (a) for $N_r = 500$ realizations of $N = 450$ events each; mean values for N_r realizations are shown for $\Delta_{rms}P$ (b), N_{K0} (A), M_6 (B), S (C), and κ (D), plotted versus the trial $\widetilde{M}_{K0}^T$. The different lines correspond to different trial threshold bandwidths ω as indicated in the legends

Results for an actual crisp $\omega = 0$ threshold at $M_{K0} = 7.0$ are shown in Figure 4, where trial $\omega^T = 0$ has a clear peak for $M_{K0}^T = 7.0$ for all measures, which is to be expected because both threshold and measure are crisp. The response for wider bandwidths is not so clear because the wider the bandwidth, the more links are being discarded. The general behavior for all measures confirms the speculation of Gutiérrez et al. (2021) that trial thresholds with too low M_{K0}^T include too many pseudo links, while too large M_{K0}^T loses too many real links. The asymmetric behavior on the sides of the peak shows that the effect of pseudo links is larger than that of missing links.

Figure 5 shows results for an actual threshold bandwidth $\omega = 1$ and $M_{K0} = 7.0$; all measures show the largest extreme values for $M_{K0}^T = M_{K0} + 0.1$ for all trial bandwidths, except for M_6, where $\omega^T = 2$ and $\omega^T = 3$ peak at $M_{K0}^T = M_{K0}$. The peaks in Markovianity measures are larger for $\omega^T = 0$ than for larger bandwidths, including the correct $\omega^T = 1$.

For $\omega = 2$ and $\omega = 3$, shown in Figures 6 and 7, respectively, both for $M_{K0} = 6.9$, all measures peak at $M_{K0}^T > M_{K0}$, indicating that their peak values are not useful for identifying M_{K0} and suggesting that the measure feature that identifies the correct M_{K0} might be an inflection point rather than a (positive or negative) peak.

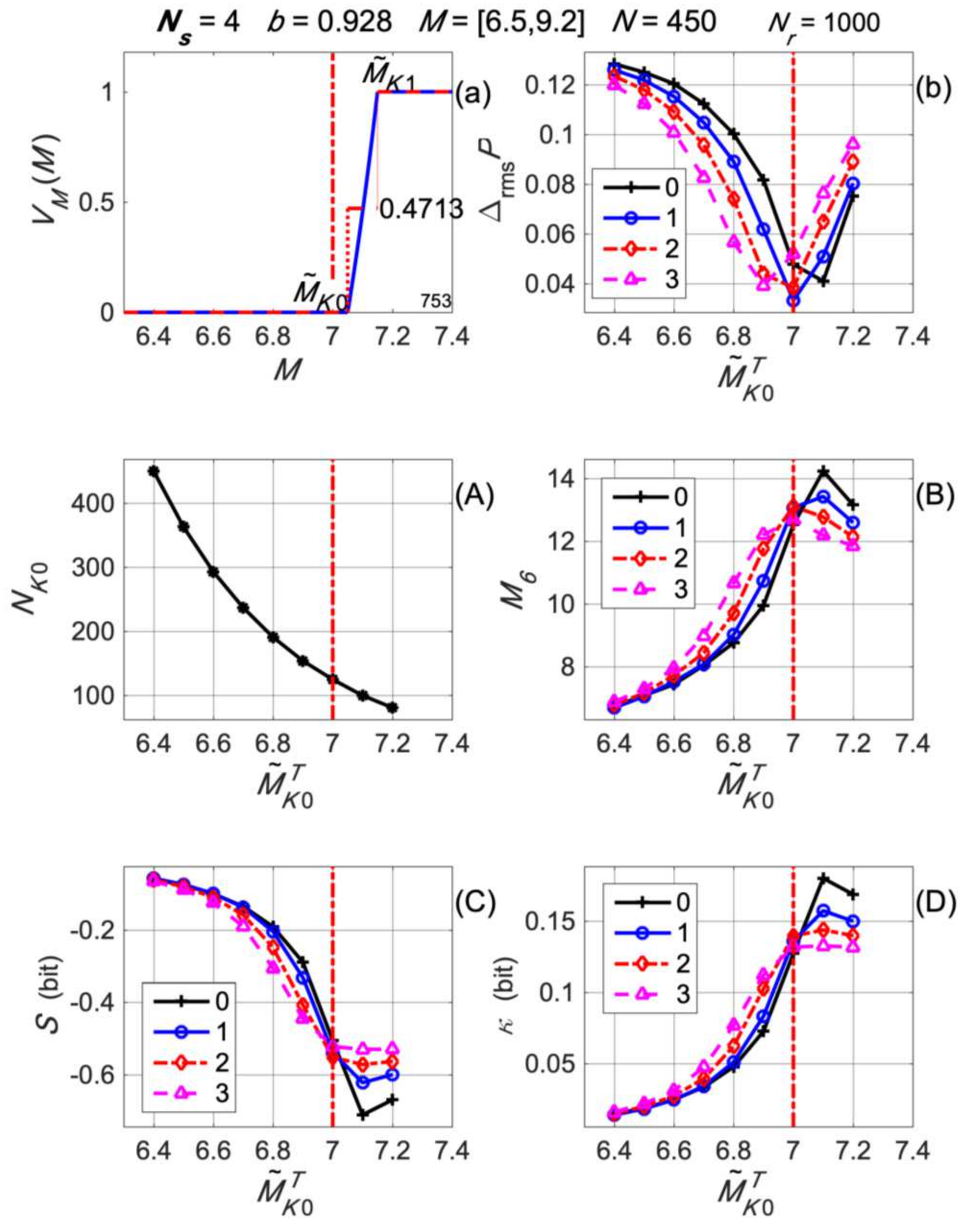

Fig. 5 Results of Monte Carlo simulation of a $\omega = 1$ threshold at $\widetilde{M}_{K0} = 7.0$ (a) for $N_r = 500$ realizations of $N = 450$ events each; mean values for N_r realizations are shown for $\Delta_{rms}P$ (b), N_{K0} (A), M_6 (B), S (C), and κ (D), plotted versus the trial $\widetilde{M}_{K0}^T$. The different lines correspond to different trial threshold bandwidths ω as indicated in the legends

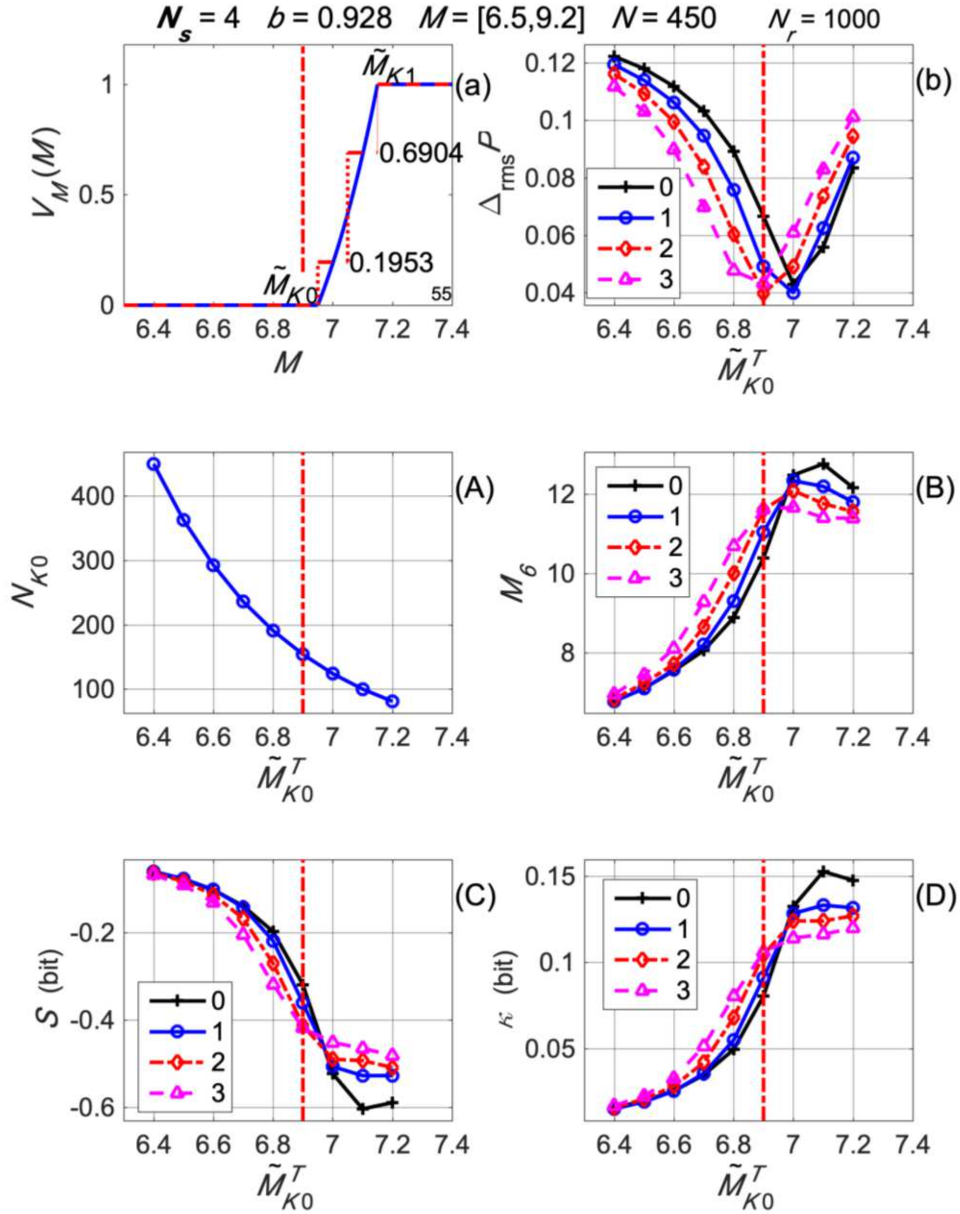

Fig. 6 Results of Monte Carlo simulation of a $\omega = 2$ threshold at $\widetilde{M}_{K0} = 6.9$ (a) for $N_r = 500$ realizations of $N = 450$ events each; mean values for N_r realizations are shown for $\Delta_{rms}P$ (b), N_{K0} (A), M_6 (B), S (C), and κ (D), plotted versus the trial $\widetilde{M}^T_{K0}$. The different lines correspond to different trial threshold bandwidths ω as indicated in the legends

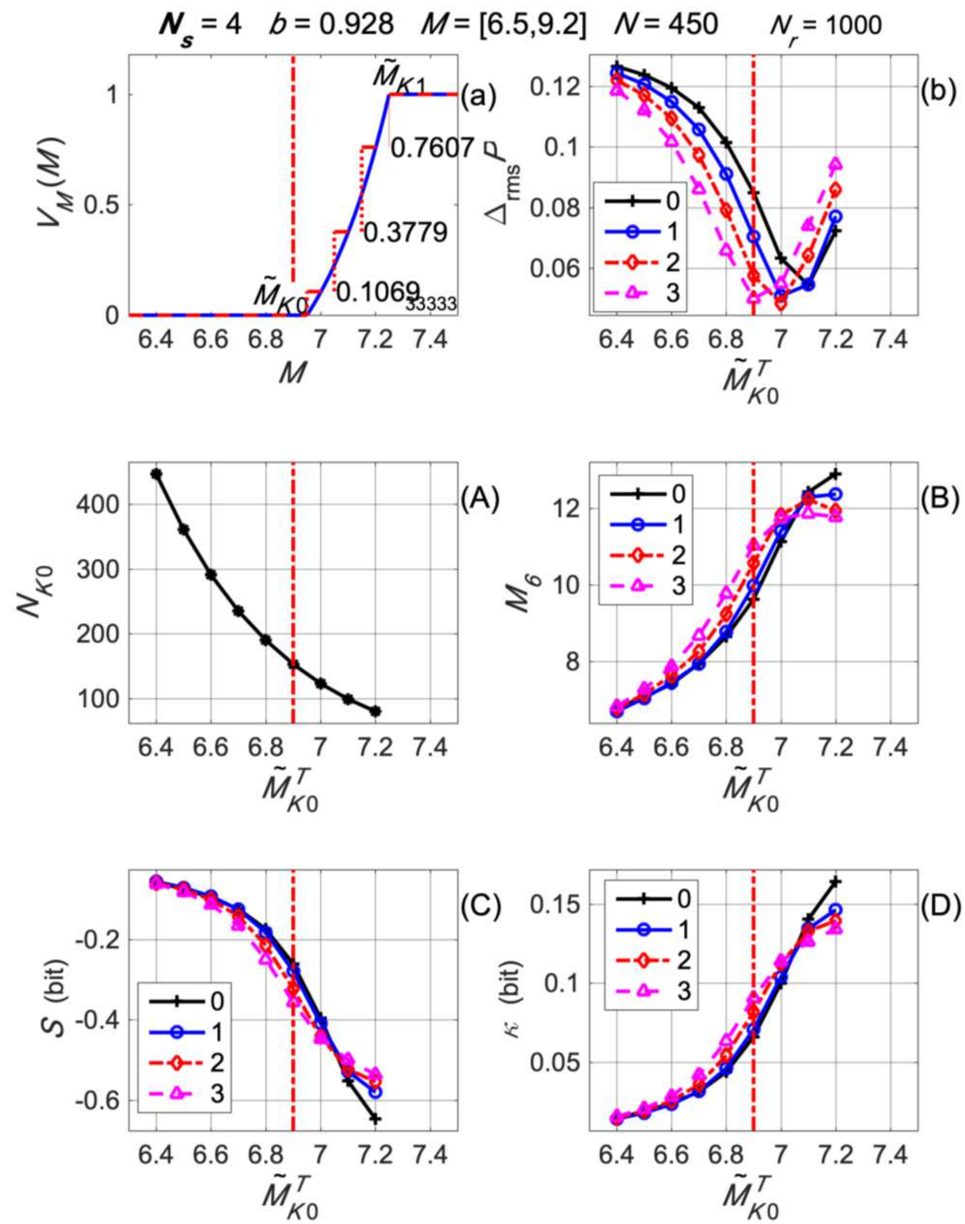

Fig. 7 Results of Monte Carlo simulation of a $\omega = 3$ threshold at $\widetilde{M}_{K0} = 6.9$ (a) for $N_r = 1{,}000$ realizations of $N = 450$ events each; mean values for N_r realizations are shown for $\Delta_{rms}P$ (b), N_{K0} (A), M_6 (B), S (C), and κ (D), plotted versus the trial $\widetilde{M}_{K0}^T$. The

different lines correspond to different trial threshold bandwidths ω as indicated in the legends

7. Application of the counting strategy to a real catalog

Figure 8 shows the results of applying our counting strategy to the data set used in Gutiérrez et al. (2021) for the same trial bandwidths and thresholds employed above; as mentioned above, the set comprises $N = 450$ earthquakes with magnitudes in the $6.5 \leq M \geq 9.2$ range. The lines for $\omega = 0$ correspond to those of Figure 4 of Gutiérrez et al. (2021), who used only a crisp threshold; trial threshold magnitudes, M^T, are equivalent to our $\widetilde{M}_{K1}^T = \widetilde{M}_{K0}^T + 1$.

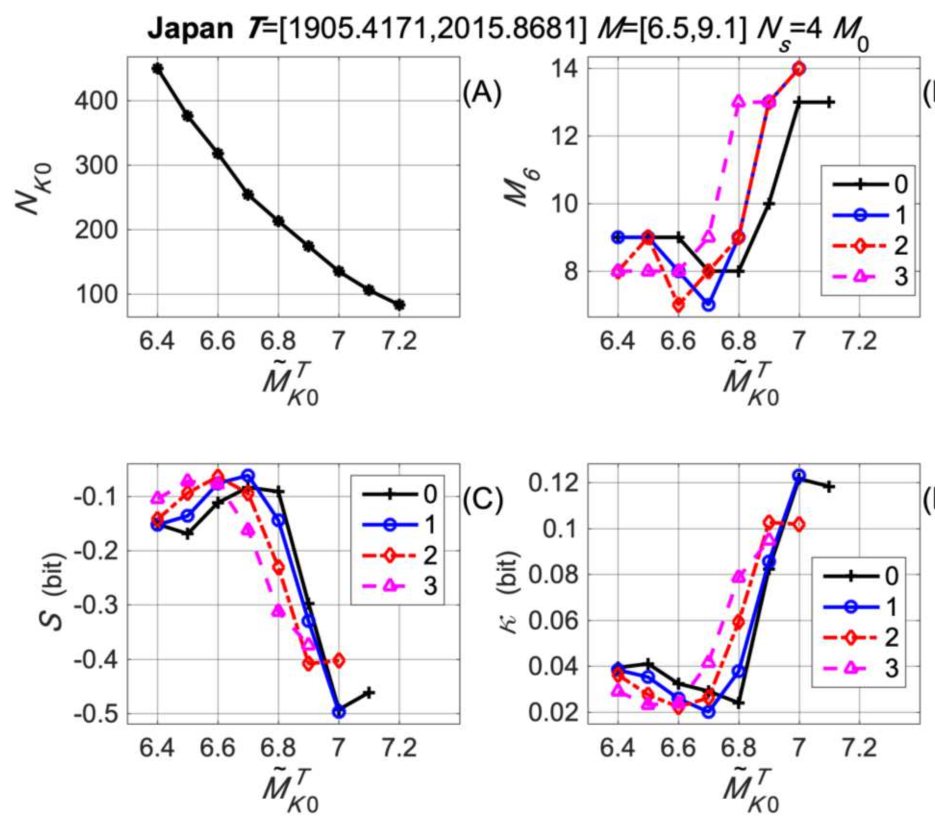

Fig. 8 Application of the counting strategy to real data from Japan. N_{K0}, is shown in (A); M_6, S, and κ are shown in (B), (C), and (D), respectively, plotted versus the trial $\widetilde{M}_{K0}^T$.

The different lines correspond to different trial threshold bandwidths ω^T indicated in the legends

The problem now is how to interpret the results shown in Figure 8; comparison with Figures 4 to 7 show similar overall behaviors, except for $\widetilde{M}_{K0}^T < 6.7$ where values are small but trends are different, but there is not a clear resemblance for higher $\widetilde{M}_{K0}^T$ that could directly and conclusively identify the actual threshold.

It should be emphasized at this point that the data represent only one (very short) realization of a stochastic process, so that the results will not necessarily conform to those for means shown above for the synthetics. Hence, we will use the means as guides and take into account the deviations from them found during the Monte Carlo simulations.

Figures 9 to 12 show, each for a given bandwidth ω and for the corresponding best fitting $\widetilde{M}_{K0}$, a comparison between the observed measures and the synthetic means together with their standard deviations.

Comparison for the crisp threshold for $\widetilde{M}_{K0} = 7.0$ shown in Figure 9, shows that for S and κ the observed measures peak for the same $\widetilde{M}_{K0}^T$ as the synthetics, but peak values are at the end of the standard deviation bands, and for $\widetilde{M}_{K0}^T$ below 6.9 observed values range outside the bands.

For $\omega = 1$ and $\widetilde{M}_{K0} = 7.0$ (Fig. 10) there were not enough observed data to go beyond $\widetilde{M}_{K0}^T = 7.0$; S and κ fit reasonably well for $\widetilde{M}_{K0}^T$ between 6.7 and 7.0, but M_6 does not fit as well.

Figure 11 shows $\omega = 2$ and $\widetilde{M}_{K0} = 6.9$ and it seems that measures for $\widetilde{M}_{K0}^T = 7.0$ were determined from too few data, but fit is acceptable for most of the $\widetilde{M}_{K0}^T$ range.

Our widest bandwidth $\omega = 3$ is shown in Figure 12 for $\widetilde{M}_{K0} = 6.9$; it was only possible to measure up to $\widetilde{M}_{K0}^T = 6.9$ and S and κ fit fairly well, but M_6 does not.

Considering the above, we find that $\omega = 2$ results in the best fit, followed closely by $\omega = 1$; more data would be needed to make a reliable choice.

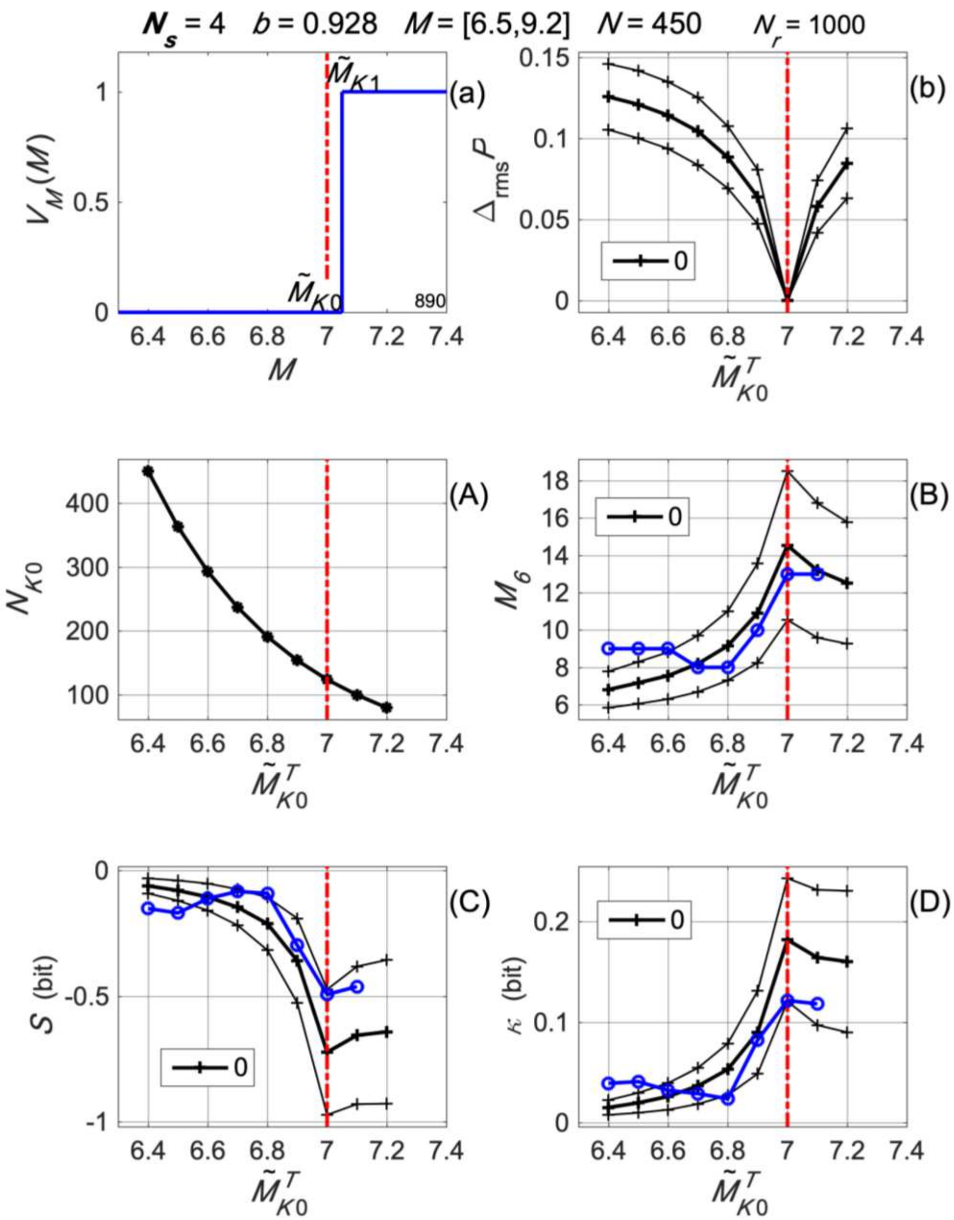

Fig. 9 Comparison of observed measures, blue lines with circles in (B) to (D), with synthetic means, thick black lines, and means plus/minus one standard deviation, thin black lines, for the synthetic threshold (a) with $\omega = 0$ and $\widetilde{M}_{K0} = 7.0$

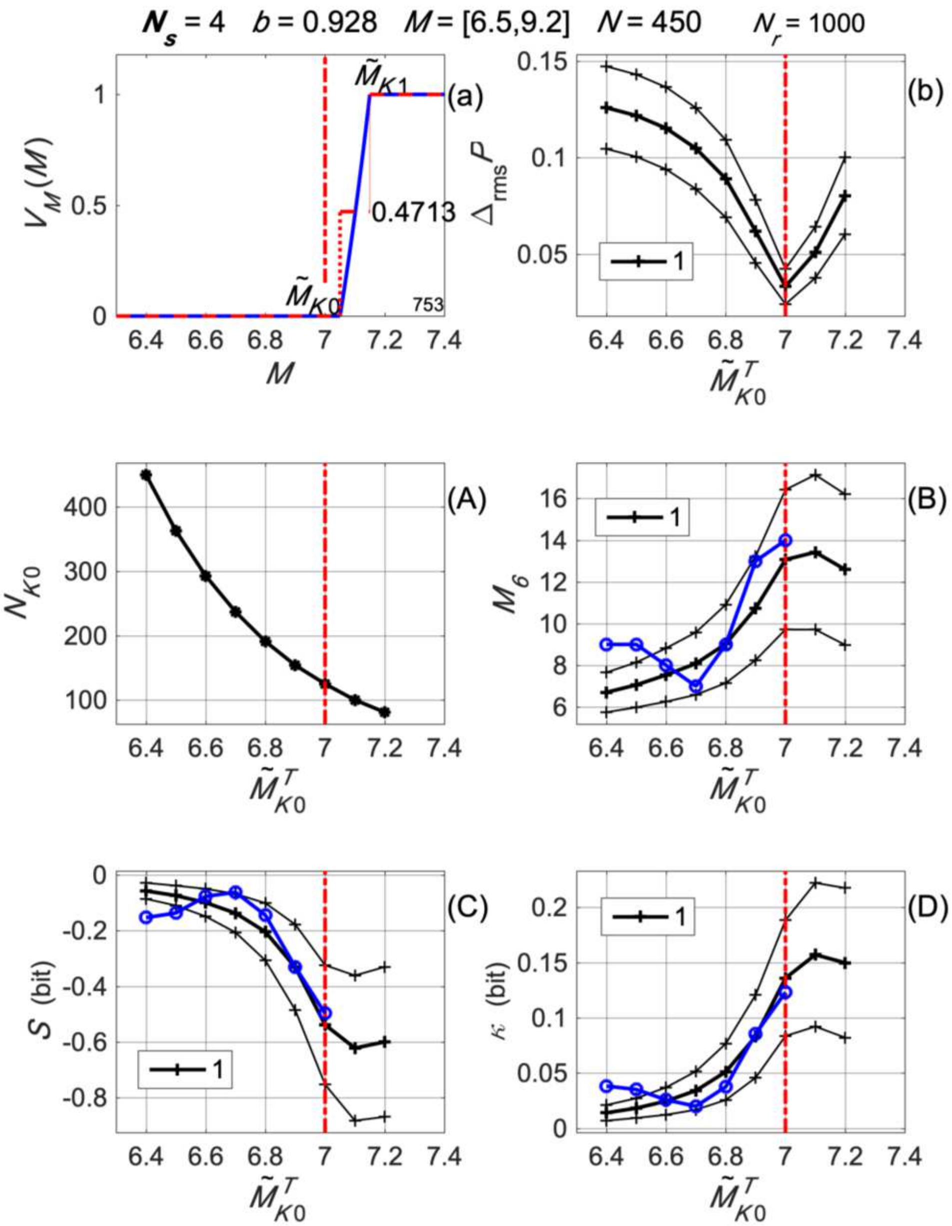

Fig. 10 Comparison of observed measures, blue lines with circles in (B) to (D), with synthetic means, thick black lines, and means plus/minus one standard deviation, thin black lines, for the synthetic threshold (a) with $\omega = 1$ and $\widetilde{M}_{K0} = 7.0$

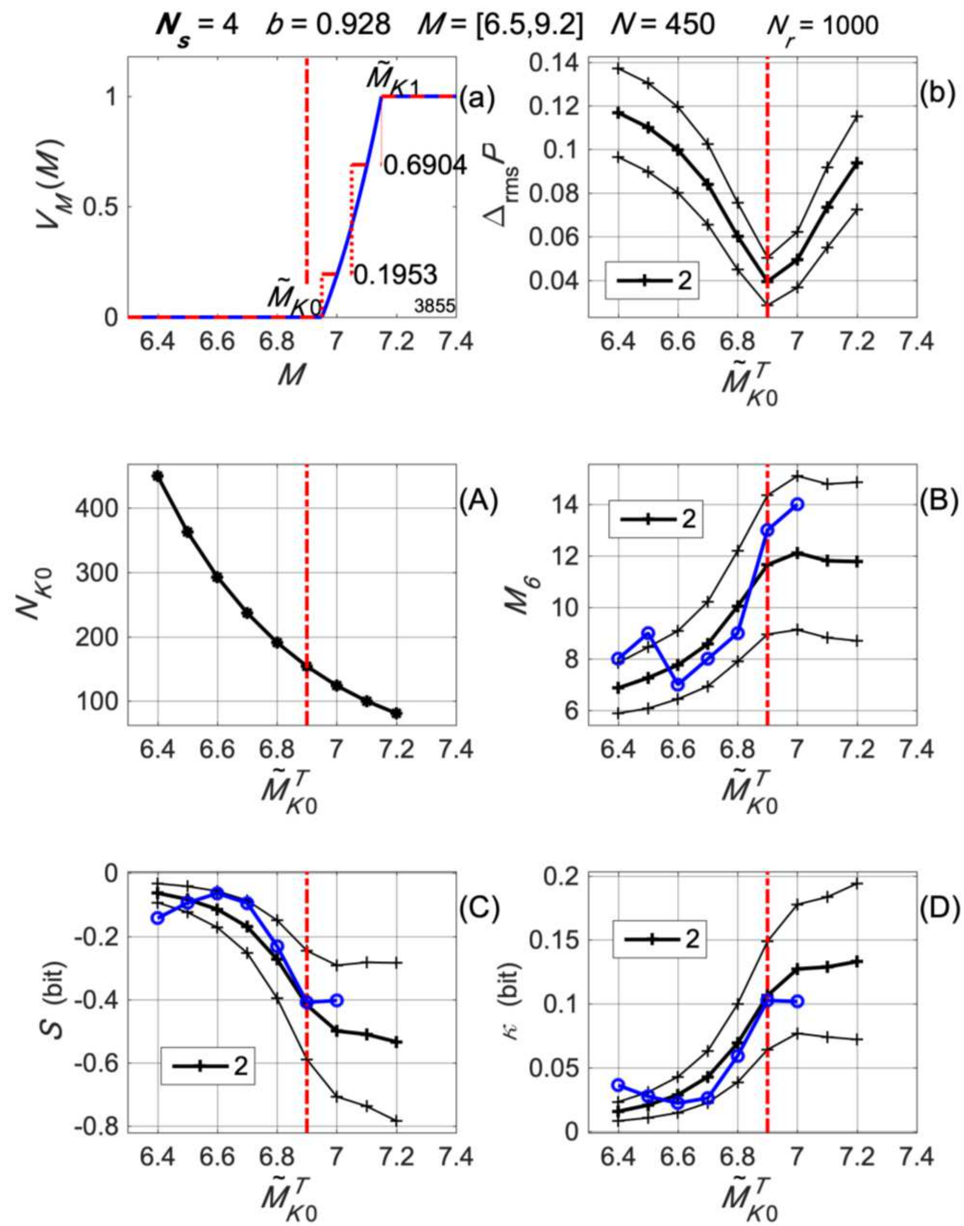

Fig. 11 Comparison of observed measures, blue lines with circles in (B) to (D), with synthetic means, thick black lines, and means plus/minus one standard deviation, thin black lines, for the synthetic threshold (a) with $\omega = 2$ and $\widetilde{M}_{K0} = 6.9$

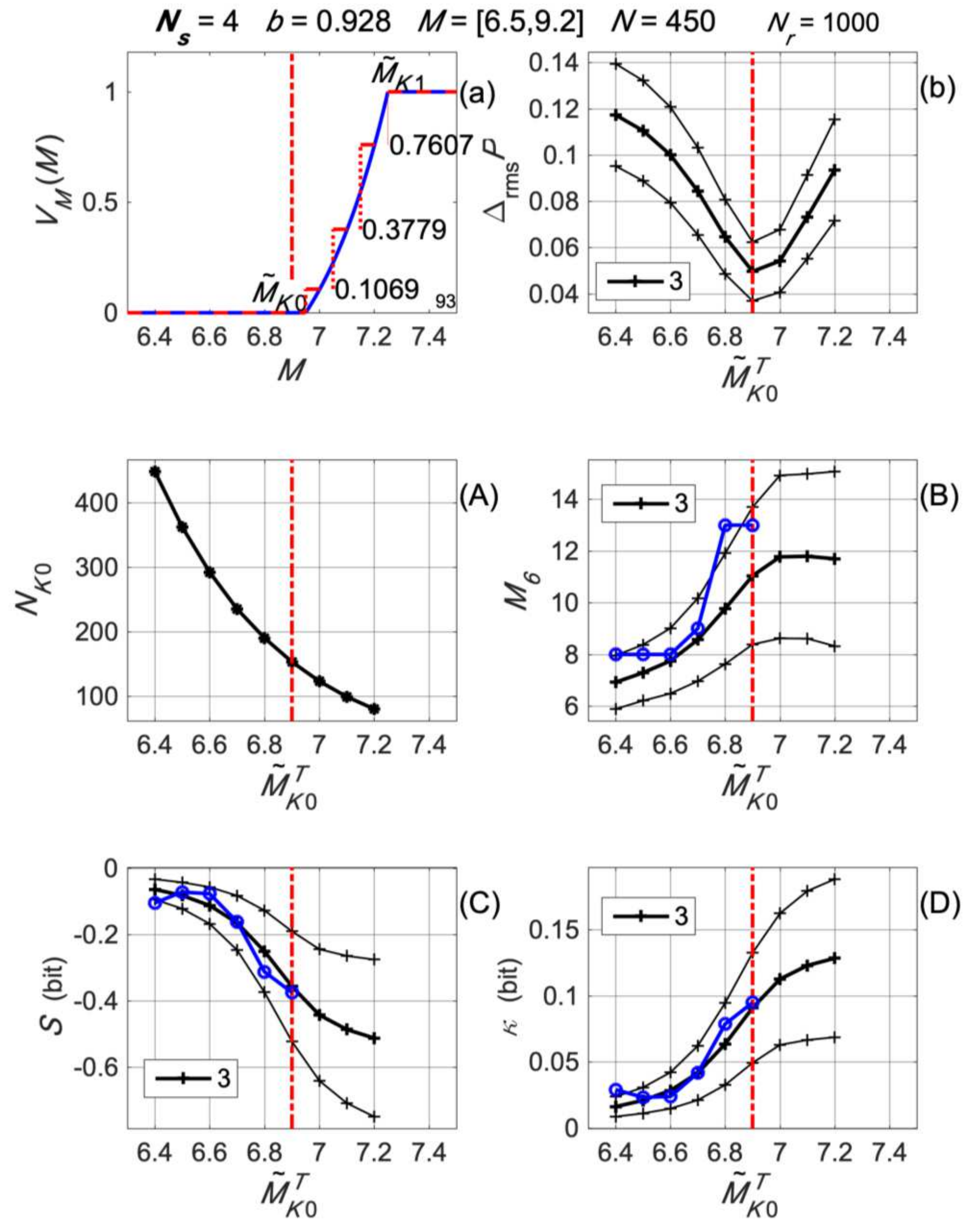

Fig. 12 Comparison of observed measures, blue lines with circles in (B) to (D), with synthetic means, thick black lines, and means plus/minus one standard deviation, thin black lines, for the synthetic threshold (a) with $\omega = 3$ and $\widetilde{M}_{K0} = 6.9$

8. Discussion and conclusions

It is unreasonable to suppose that the threshold between seismic magnitudes that constitute a Markovian process and those that do not can be adequately modeled by a crisp step-function, particularly when considering the uncertainties in magnitude determination and the effects of rounding. Hence it is necessary to use fuzzy thresholds to model the process. We heuristically propose a membership function with probability proportional to seismic moment in the transition band.

The usual method for evaluating empirical transition probabilities assumes all events in the observed sequence are links in a Markov chain; we propose a counting strategy for fuzzy thresholds based on the probabilities of being Markovian of the events in the sequence, and justify the strategy through Monte Carlo simulations. The counting strategy can be applied to any membership function.

The results from the real data show features resembling some of the simulations, but do not correspond closely and unequivocally to any one of them; which is not surprising since the observed data constitute only one very short realization of a random process, so that large differences can be expected from realization to realization, and any one realization can be expected to differ from the mean of many realizations. Taking the above into account, we find that the empirical results resemble most the simulations for $M_{K0} = 6.9$ and $\omega = 2$; so we conclude that, within the limitations of the data, the Markovian behavior of the seismicity in the Japanese area is best modeled by a fuzzy threshold 0.2 magnitude units wide above $M_{K0} = 6.9$.

Use of this model to estimate the Markovian seismic hazard in the Japan area instead of the Gutiérrez et al. (2021) crisp threshold model, changes the Markovian transition probabilities from their

$$P_G = \begin{bmatrix} 0.20833 & 0.47917 & 0.14583 & 0.16667 \\ 0.36667 & 0.26667 & 0.18333 & 0.18333 \\ 0.30000 & 0.20000 & 0.33333 & 0.16666 \\ 0.20000 & 0.40000 & 0.08571 & 0.31429 \end{bmatrix}$$

to

$$P = \begin{bmatrix} 0.24568 & 0.43339 & 0.15040 & 0.17052 \\ 0.32478 & 0.27180 & 0.18293 & 0.22048 \\ 0.27301 & 0.22185 & 0.41513 & 0.09001 \\ 0.26966 & 0.42882 & 0.04776 & 0.25376 \end{bmatrix}.$$

The largest changes are +0.08180 for $p_{3\,3}$ and -0.07665 for $p_{3\,4}$, which make region 3 the most likely to have repetitions of large earthquakes, and the mean absolute change over all transition probabilities is 0.03627. The significance of these changes is apparent when considering that they are 0.3272, 0.3066, and 0.1451, respectively, of the uniform probability 0.25.

References

Alvarez, E. (2005). Estimation in stationary Markov renewal processes, with application to earthquake forecasting in Turkey. *Methodology and Computing in Applied Probability, 7*, 119-130.

Anagnos, T., & Kiremidjian, A. (1988). A review of earthquake occurrence models for seismic hazard analysis. *Probabilistic Engineering Mechanics*, *3*, 3-11.

Aziz, S., & Parthiban, J. (1996). *Fuzzy logic*. http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/sbaa/report.fuzzysets.html [25 October 2007]

Barucha-Reid, A. (1960). *Elements of the theory of Markov processes and their applications.* Dover 0-486-69539-5.

Battaglia, F. (2007). *Metodi di previsione statistica.* Springer-Verlag Italia, Milano, 323pp.

Cavers, M., & Vasudevan, K. (2014). Spatio-temporal complex Markov chain (SCMC) model using directed graphs: earthquake sequencing. *PAGEOPH* DOI 10.1007/s00024-014-0850-7

Cavers, M., & Vasudevan, K. (2015). Brief communication: Earthquake sequencing: analysis of time series constructed from the Markov chain model. *Nonlin. Processes Geophys.*, *22*, 589-599.

Ching, W., & Ng, M. (2006). *Markov chains: Models, algorithms and applications.* Springer Science+Business Media, Inc. , 205pp.

Feller, W. (1968). *An introduction to probability theory and applications.* (3dEd) John Wiley & Sons, USA, 2 Vols. 509 & 669pp.

Fujinawa, Y. (1991). A method for estimating earthquake occurrence probability using first- and multiple-order Markov chain models, *Natural Hazard*, *4*, 7–22.

Gnedenko, B. (1962). *The theory of probability and the elements of statistics*. (Translated by B. Seckler) AMS Chelsea Publishing, 5th Ed..

Gutiérrez, Q., Nava, F., Glowacka, E., Castro, R., & Márquez, V. (2021). Assessing Markovian models for seismic hazard and forecasting. *PAGEOPH*, *178*, 847-863. https://doi.org/10.1007/s00024-021-02686-2

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

Herrera, C., Nava, F., & Lomnitz, C. (2006). Time-dependent earthquake hazard evaluation in seismogenic systems using mixed Markov chains: An application to the Japan area, *EarthPlanetsSpace*, *58*, 973-979.

Kullback, S., & Leibler, R. (1951). On information and sufficiency. *Ann. Math. Statist.*, *22*(1), 79-86. https://projecteuclid.org/euclid.aoms/1177729694

Lehner, F., Li, V., & Rice, J. (1981). Stress diffusion along rupturing plate boundaries. *J. Geophys. Res.*, *86* (B7), 6155-6169.

Márquez Azúa, B., DeMets, C., & Masterlark, T. (2002). Strong interseismic coupling, fault afterslip, and viscoelastic flow before and after the Oct. 9, 1995 Colima-Jalisco earthquake: Continuous GPS measurements from Colima, Mexico. *Geophys. Res. Lett.*, *29* (8) 10.11029/2002GL014702.

Melbourne, T., Webb, F., Stock, J., & Reigber, C. (2002). Rapid postseismic transients in subduction zones from continuous GPS. *J. Geophys. Res.*, *107* (B10), 2241.

Nava, F., Herrera, C., Frez, J., & Glowacka, E. (2005). Seismic hazard evaluation using Markov chains; Application to the Japan area. *PAGEOPH*, *162*, 1347-1366.

Nishioka, T., & Shah, H. (1980). Application of Markov chain on probability of earthquake occurrence. *Proceeds of JSCE*, *298*, 137-145.

Parzen, E. (1960). *Modern Probability Theory and its Applications.* J. Wiley and Sons, Japan.

Patwardhan, A., Kulkarni, R., & Tocher, D. (1980). A semi-Markov model for characterizing recurrence of great earthquakes. *Bull. Seism. Soc. Am.*, *70*, 323–347.

Quinteros, C., Nava, F., Glowacka, E., & Frez, J. (2014). Semi-periodic sequences and extraneous events in earthquake forecasting: II Application, forecasts for Japan and Venezuela. *PAGEOPH*, *171*(7), 1367-1383, DOI 10.1007/s00024-013-0678-6.

Riga, G., & Balocchi, P. (2016). Seismic sequences' branching structures: long-range interactions and hazard levels. *Open Journal of Earthquake Research*, *5*, 189-205.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*: 379-423, 623-656.

Spagnotto, S., Alvarez, O., & Folguera, A. (2018). Static stress increase in the outer forearc produced by MW 8.2 September 8, 2017 Mexico earthquake and its relation to the gravity signal. *PAGEOPH*, *175*, 2575–2593, https://doi.org/10.1007/s00024-018-1962-2

Tsapanos T., & Papadopoulou, A. (1999). A discrete Markov Model for earthquake occurrences in Southern Alaska and Aleutian Islands. *J. Balkan Geophys. Soc.,* *2*(3), 75-83.

Ünal, S., & Celebioglu, S. (2011). A Markov chain modeling of the earthquakes occurring in Turkey, *Gazi University Journal of Science*, *24*(2), 263–274.

Votsi, I., Limnios, N., Tsaklidis, G., & Papadimitriou, E. (2010). Semi-markov models for seismic hazard assessment in certain areas of Greece. *Bulletin of the Geological Society of Greece*, *43*, 2200–2209.

Votsi, I., Tsaklidis, G., Limnios, N., Papadimitriou, E., & Vallianatos, F. (2013). A Markov model for seismic hazard analysis along the Hellenic subduction Zone (Greece). *Bulletin of the Geological Society of Greece*, *47*(3), 1376–1385. https://doi.org/ 10.12681/bgsg.10934

Zadeh, L. (1965). Fuzzy sets. *Information and Control*, *8*, 338–353.

Zadeh, L. (1988). Fuzzy logic. *IEEE Computer*, *21*, 83–93. 3

Statements and Declarations

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

Author Contributions: All authors contributed to the study conception and design. The first draft of the manuscript was written by Fidencio Alejandro Nava. All authors read and approved the final manuscript.”