Pure Appl. Geophys.

https://doi.org/10.1007/s00024-024-03534-9

Pure and Applied Geophysics

The Magnitude Threshold and Missing and Pseudo Links in Markov Chains

F. A. Nava[1] and Q. J. Gutiérrez[1]

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 transition 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 membership function with probabilities in the transition band proportional to the seismic moment. To estimate empirical transition probabilities when considering a fuzzy magnitude threshold, we propose a counting strategy for the observed transitions and justify it through Monte Carlo simulations. The counting strategy is illustrated by application to the model from a previous seismic study of the Japan area by testing, through Monte Carlo simulations, how well the counting strategy results resemble optimum estimations of the transition probabilities. The simulations are also used to study the behavior of three Markovianity measures, and it is found that the peak values of these measures are not useful in identifying the true transition band, but that this band may be better identified by using the whole set of values taken by each measure for different transition band models. As an illustration, the measures were applied to real data from the previous study, a short set corresponding to a single realization, and found that the behavior of the measures does not agree with those expected from a crisp threshold, but agree, within the limitations of the data, with either a fuzzy threshold going from zero probability for magnitudes $M \leq 6.9$ to probability one for $M \geq 7.2$ or from zero probability for magnitudes $M \leq 7.0$ to probability one for $M \geq 7.2$.

Keywords: Markov process, Magnitude threshold, Fuzzy thresholds, Seismic hazard.

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 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 (e.g., Leptokaropoulos et al., 2018; Ringdal, 1976; Werner & Sornette, 2008); 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 $6.94\overline{999}$ (rounded to 6.9) would be discarded by a stepwise threshold $M \geq 7.0$, yet it cannot be physically very different from one with unrounded magnitude $6.95\overline{000}$ (rounded to 7.0). Considering the uncertainties, an earthquake with rounded magnitude M cannot, generally, be certainly very different from earthquakes with magnitudes $M \pm 0.1$.

[1] Seismology Department, Centro de Investigación Científica y de Educación Superior de Ensenada, B. C., Ensenada, Baja California, Mexico. E-mail: fnava@cicese.mx

Published online: 13 July 2024

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 may result in a moment release episode (Quinteros et al., 2014) with equivalent magnitude above the threshold, and the episode 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 Markovian systems are considered in many seismic hazard studies (e.g., Alvarez, 2005; Anagnos & Kiremidjian, 1988; Cavers & Vasudevan, 2014, 2015; Fujinawa, 1991; Herrera et al., 2006; Nava et al., 2005; Nishioka & Shah, 1980; Patwardhan et al., 1980; Ünal & Celebioglu, 2011; Votsi et al., 2010, 2013).

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 (Fig. 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 $6.5 \leq M \leq 9.2$ 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 in depth here; to give a rough picture, regions one and two are defined by the seismicity due to the Japan and Kuriles trenches, and the limit between these was determined from seismicity sections that show different dips for the Pacific plate subduction. Region three is characterized by deep significant earthquakes and a different trench azimuth from region two. Region four contains the seismicity from the Ryuku and Nankai trenches.

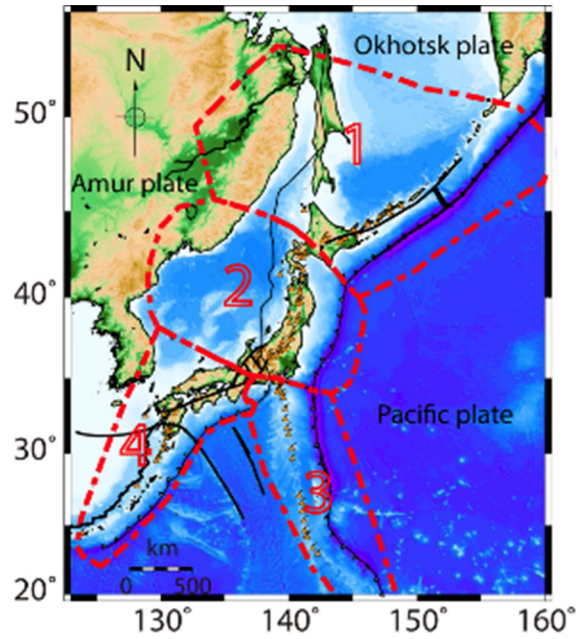

Figure. 1
Map of the study area. The dashed lines enclose the four regions corresponding to the four states of the system (modified from Gutiérrez et al. (2021))

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 & Papadopoulou, 1999; Márquez et al., 2002; Melbourne et al., 2002; Venkatamaran and Kanamori, 2004; Riga & Balocchi, 2016; Spagnotto et al., 2018).

Gutiérrez et al. (2021) obtained transition probability matrices for several threshold magnitudes using stepwise trial thresholds and, 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 when magnitudes are 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 (Aziz & Parthiban, 1996; Zadeh, 1965), 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., Barucha-Reid, 1960; Battaglia, 2007; Ching & Ng, 2006; Feller, 1968; Gnedenko, 1962; Parzen, 1960, 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

$$\Pr[s_{[m+1]} = s_j | s_{[m]} = s_i] = \widehat{p}_{ij} \tag{1}$$

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

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

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

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

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

where $\widehat{\pi} = [\widehat{\pi}_1, \widehat{\pi}_2, \cdots, \widehat{\pi}_{N_s}], \sum_{j=1}^{N_s} \widehat{\pi}_j = 1$; the elements of $\widehat{\pi}$, 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 directly forming 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:

If a segment of the true chain is $\cdots, s_i, s_j, s_k, \cdots$, when counting transitions both θ_{ij} and θ_{jk} will be correctly increased by 1 each, but if the link consisting of 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 intrudes 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 and pseudo links for a typical $N_s = 4$ TPM and a $N = 450$ long series is illustrated in Fig. 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 considerably larger than those due to missing links.

4. Fuzzy Thresholds and the Membership Function

We will use fuzzy magnitude thresholds 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.

Magnitude data in seismic catalogs are usually rounded to $\Delta M = 0.1$, in what follows we shall denote unrounded magnitudes by m and rounded magnitudes by M. 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.

We will characterize all thresholds by two rounded magnitudes: M_{K0} and M_{K1}, such that events with rounded magnitudes $M \leq M_{K0}$ are definitely too small to be part of a Markovian chain and events with rounded magnitudes $M \geq M_{K1}$ 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 $m^U_{K0} \equiv M_{K0} + \Delta M/2$ and $m^U_{K1} \equiv M_{K1} - \Delta M/2$. These limits define a transition range where events with unrounded magnitudes $m^U_{K0} < m < m^U_{K1}$ may stochastically belong to the Markovian set, or not, with probability given by the membership function in that range.

The threshold bandwidth

$$\omega = (M_{K1} - M_{K0})/\Delta M - 1 = (m^U_{K1} - m^U_{K0})/\Delta M \tag{6}$$

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

It is not known what $V(m)$ is like in the transition band, but given that we associate the size of the earthquakes with their probability of belonging to a Markovian chain, 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 & 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

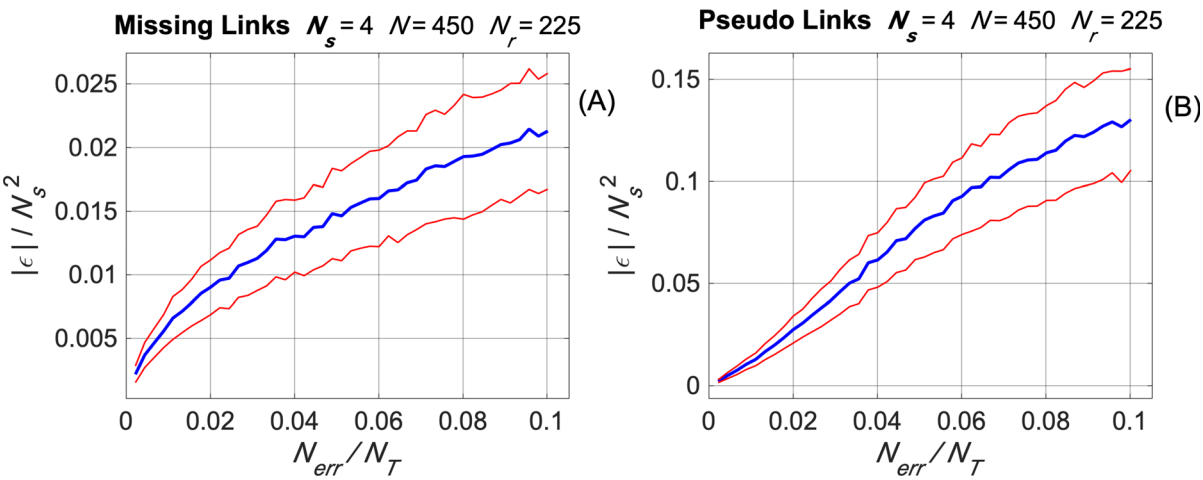

Figure. 2
Average error per transition probability vs. proportion of missing (**A**) or pseudo (**B**) links; the thick blue line is the mean error for 225 Monte Carlo realizations and the thin lines are the mean plus/minus one standard deviation

simply m or M. Thus, the moment membership probability is given by

$$V(m) = \begin{cases} 0; & m \le m_{K0}^U \\ \frac{10^{1.5m} - 10^{1.5M_{K0}^U}}{10^{1.5M_{K1}^U} - 10^{1.5M_{K0}^U}}; & m_{K0}^U < m < m_{K1}^U \\ 1,; & m \ge m_{K1}^U \end{cases} \quad (8)$$

illustrated in Fig. 3.

Other possible membership functions may not have physical significance, like a transition band proportional to magnitude (because a magnitude is not proportional to the earthquake size but to its logarithm), or could be just handy commonly used

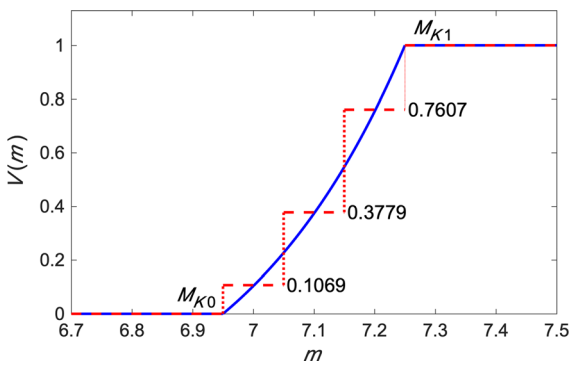

Figure. 3
The moment fuzzy threshold membership probability function for ω=3 transition band. The dashed horizontal lines represent the average probabilities for each rounded magnitude step in the transition range

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 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 one gives the best results and thus represents best the actual 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 denote the rounded magnitude by M; since the original unrounded magnitude can have any value in the $[M - \Delta M/2, M + \Delta M/2)$ range we will consider the probability of an event with magnitude M belonging

to the Markovian set to be the average of the probabilities within the range

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

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

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

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

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

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

$$w_{[n,n+2]} = v_{[n]}(1 - v_{[n+1]})v_{[n+2]} \tag{11}$$

because for this transition to be a link it is necessary to include the probability that event $n+1$ 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

$$w_{[n,n+k]} = v_{[n]}(1 - v_{[n+1]}) \cdots (1 - v_{[n+k-1]})v_{[n+k]} \tag{12}$$

It is clear from eqn. (12) 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^{k-1} different possible paths from an initial state s_n to a final state s_{n+k}.

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

$$\theta_{s_{[n]}s_{[n+k]}} = \theta_{s_{[n]}s_{[n+k]}} + w_{[n,n+k]} \tag{13}$$

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 eqn. (12). After all transitions have been counted, the transition probabilities will be estimated according to eqn. (4).

6. *Monte Carlo Validation of the Counting Strategy for Fuzzy Thresholds and Exploration of Three Markovianity Measures*

We will use Monte Carlo methods 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 eqn. (14).

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

$$\log_{10} N^C(m) = a_0 - b(m - m_{K0}^U) \tag{14}$$

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

$$\Pr(m) = \beta \mathrm{e}^{-\beta(m - m_{K0}^U)}; m \geq m_{K0}^U \tag{15}$$

where $\beta = b \ln 10$.

Events with unrounded magnitudes below the transition range are considered 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 $\widehat{P}$, 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 $\widehat{P}$ 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 $\widehat{P}$ 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 eqn. (4) using only and all events in the Markovian chain. According to Borel's law of large numbers, P_R should tend to $\widehat{P}$ 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 $\widehat{P}$, 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 an optimal estimate 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 threshold (called henceforward simply model), consisting of given bandwidth ω, M_{K0}, and "true" TPM $\widehat{P}$, $N_r = 1000$ 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 $M_{K0}^T = 6.4{,}6.5, \cdots, 7.0{,}7.1$.

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, i.e., 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 $\Pi = [\pi_1, \pi_2, \cdots, \pi_{N_s}]^T$ eqn. (3), defined as

$$S = S^P - S^\Pi \tag{16}$$

where

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

and

$$S^\Pi = -4\sum_{j=1}^{N_s} \pi_j \log_2 \pi_j \tag{18}$$

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

$$\kappa = \frac{1}{N_s}\sum_{i=1}^{N_s}\sum_{j=1}^{N_s} p_{ij}\log_2\left(\frac{p_{ij}}{\pi_j}\right) \tag{19}$$

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 Figs. 4, 5, 6, 7 how the probabilities are adjusted and how the measures behave for $M_{K0} = 6.9$ only, to illustrate the effects of different bandwidths. In each figure the postulated "real" threshold M_{K0} and M_{K1} are 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 N_{K0}, M_6, S, and κ, respectively.

Graph (b) is very important, because it shows in Figs. 4, 5, 6, 7 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 estimate, i.e. the one closest to the optimal, which indicates that this strategy is indeed working as it should. These results will be summarized in Fig. 8.

Results for an actual crisp $\omega = 0$ threshold at $M_{K0} = 6.9$ and $N = 450$ are shown in Fig. 4; we chose this combination of parameters because they correspond to the preferred result in Gutiérrez et al. (2021), for a crisp threshold at their $M^T = 7.0 = M_{K1}$. Trial $\omega^T = 0$ has a sharp peak at $M_{K0}^T = 6.9$ for all measures, which is to be expected because both real and trial thresholds are crisp. These peaks tell us that the measures suggested heuristically by Gutiérrez et al. (2021) do work for a crisp threshold and supports their speculation that measurements for too low M_{K0}^T include too many pseudo links, while those for too large M_{K0}^T lose too many real links. The asymmetric behavior on the sides of the peak suggests that the effect of pseudo links is larger than that of missing links.

Since the actual threshold width is not known a priori, different trial bandwidths should be used, and the measurements for wider trial bandwidths are also shown in Fig. 4. Although these measurements are not appropriate for the actual $\omega = 0$, they also show smaller peaks for the right M_{K0}^T, except for M_6 with $\omega = 3$ that peaks at $M_{K0}^T = 6.8$.

Figure 5 illustrates the effects of a larger threshold bandwidth, $\omega = 1$, for the same $M_{K0} = 6.9$. The measures for $\omega^T = 1$ peak for $M_{K0}^T = 7.0$, halfway between M_{K0} and M_{K1}. $\omega^T = 0$ also peaks for $M_{K0}^T = 7.0$ with larger extreme values than $\omega^T = 1$, and could be incorrectly interpreted as corresponding to a crisp threshold with $M_{K0} = 7.0$.

The $\omega = 2$ threshold for $M_{K0} = 6.9$ is shown in Fig. 6. For M_6 the correct $\omega^T = 2$ peaks at $M_{K0}^T = 7.0$, but S and κ feature no peaks. This is an effect of the small number of events used; the same calculations using $N = 1000$ events (not shown here) feature clear peaks for $M_{K0}^T = 7.0$, except for $\omega^T = 0$ that peaks for $M_{K0}^T = 7.1$. Also notable is the fact that the peak values do not identify the correct M_{K0}.

Figure 7 shows the $\omega = 3$ threshold for $M_{K0} = 6.9$. Measures show no peaks for $N = 450$, but this is again an artifact of the small number of events; using $N = 1000$ events (not shown here) results in peaks for $M_{K0}^T = 7.0$ for the correct $\omega^T = 3$, but also in larger peaks for $M_{K0}^T = 7.1$ for the smaller trial bandwiths.

Repetition of the Monte Carlo simulations introducing small variations in the various P's, does not substantially change the results.

Figure 8 summarizes the result, mentioned above, that in all cases, the correct trial bandwidth results in the smallest differences between measured P and optimal P_r for $M_{K0}^T = M_{K0}$, which supports the appropriateness of the proposed counting strategy for probability estimation. Note that the second-best value corresponds in each case to the model that is closest to the correct one.

7. *Application of the Counting Strategy to a Real Catalog*

Having validated our counting strategy through Monte Carlo simulation, we will now illustrate its application to a real set of data. In what follows, it should be kept in mind that a serious limitation in

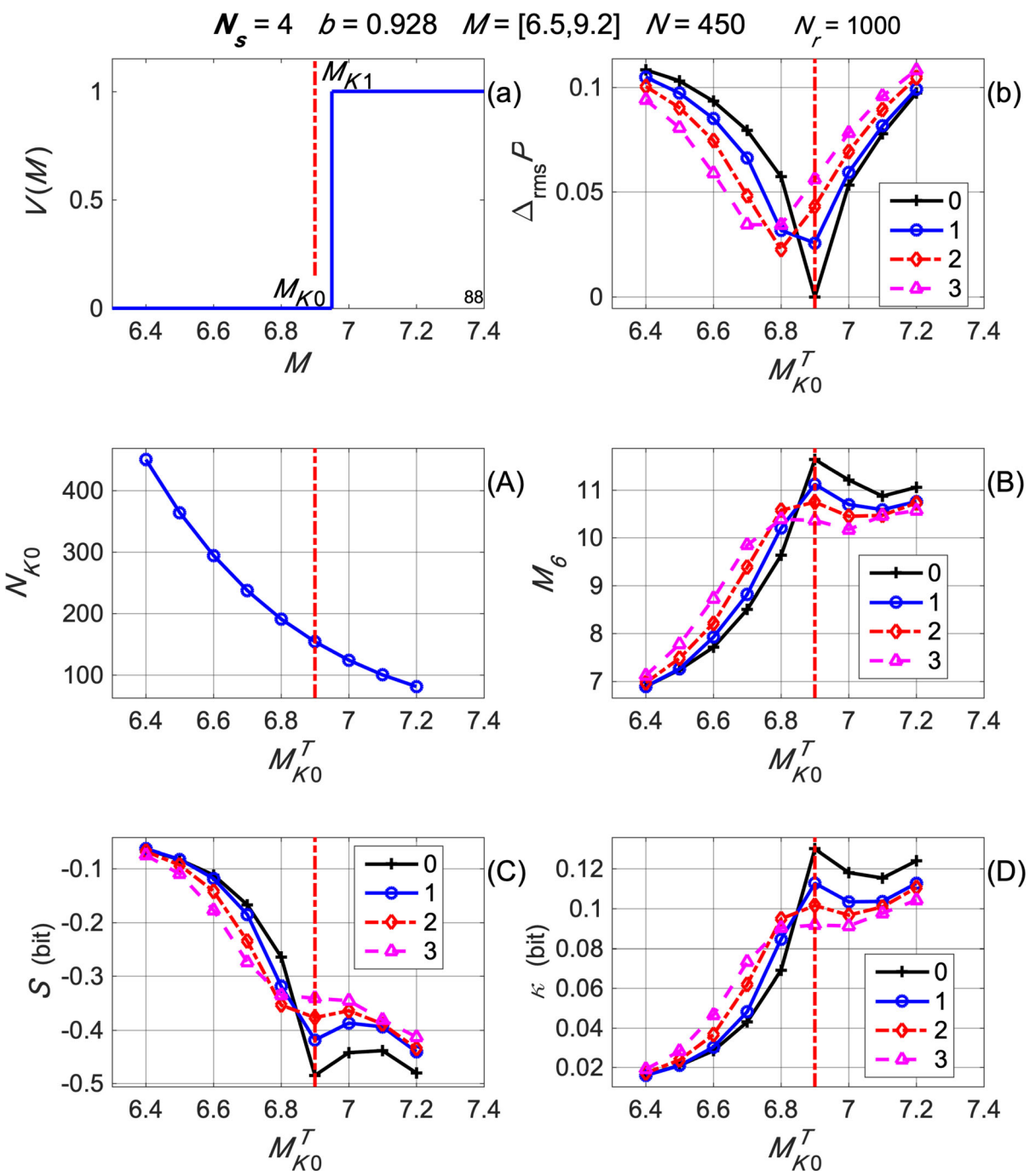

Figure. 4
Results of Monte Carlo simulation of a crisp, $\omega = 0$ threshold at $M_{K0} = 6.9$ (**a**) for $N_r = 1{,}000$ realizations of $N = 450$ events each; mean values for N_r realizations are shown for $\Delta_{rms}P$ (**b**), N_{K0} (**A**), M_6 (**B**), S (**C**), and κ (**D**), plotted versus the trial M^T_{K0}. The different lines correspond to different trial threshold bandwidths ω^T as indicated in the legends

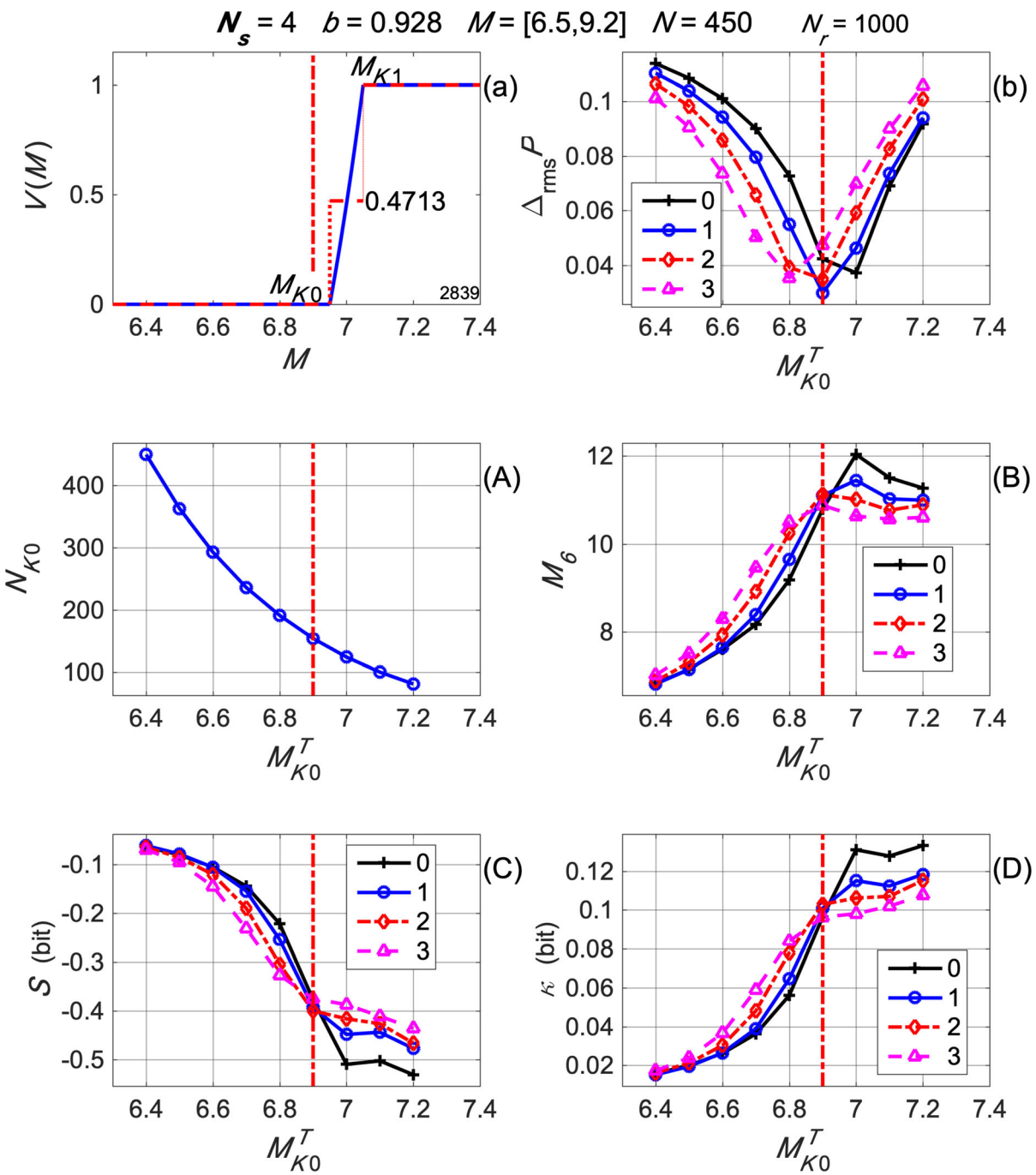

Figure. 5
Results of Monte Carlo simulation of a $\omega = 1$ threshold at $M_{K0} = 6.9$ (**a**) for $N_r = 1000$ realizations of $N = 450$ events each; mean values for N_r realizations are shown for $\Delta_{rms}P$ (**b**), N_{K0} (**A**), M_6 (**B**), S (**C**), and κ (**D**), plotted versus the trial M^T_{K0}. The different lines correspond to different trial threshold bandwidths ω^T as indicated in the legends

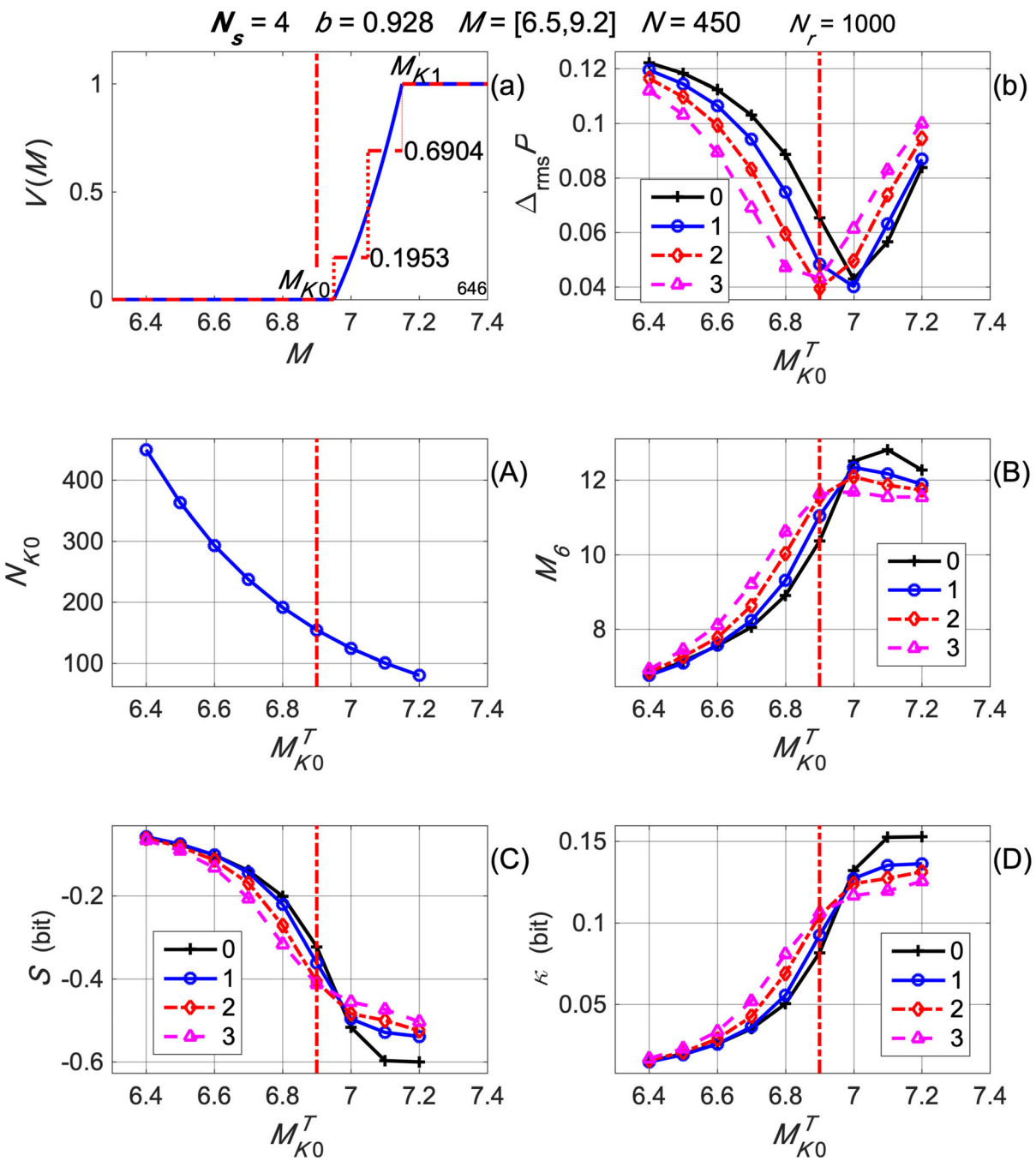

Figure. 6
Results of Monte Carlo simulation of a $\omega = 2$ threshold at $M_{K0} = 6.9$ (**a**) for $N_r = 1000$ realizations of $N = 450$ events each; mean values for N_r realizations are shown for $\Delta_{rms}P$ (**b**), N_{K0} (**A**), M_6 (**B**), S (**C**), and κ (**D**), plotted versus the trial M_{K0}^T. The different lines correspond to different trial threshold bandwidths ω^T as indicated in the legends

statistical studies of large earthquakes is the size of the sample; as larger M_{k0} values are considered the number of pertinent events, i.e. the effective sample length, N_{K0}, shown in plot (A) of the figures,

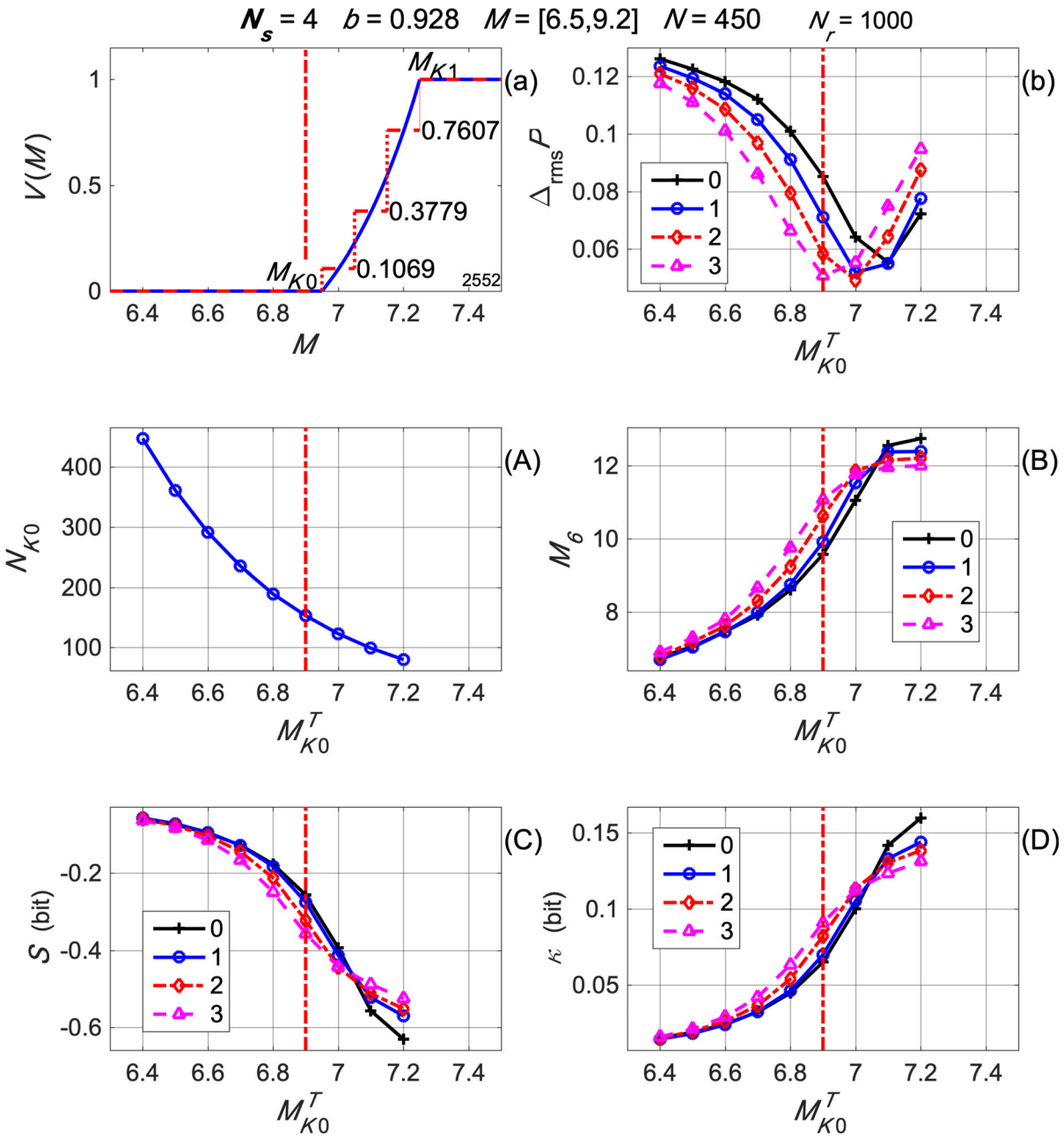

Figure. 7
Results of Monte Carlo simulation of a $\omega = 3$ threshold at $M_{K0} = 6.9$ (**a**) for $N_r = 1{,}000$ realizations of $N = 450$ events each; mean values for N_r realizations are shown for $\Delta_{rms}P$ (**b**), N_{K0} (**A**), M_6 (**B**), S (**C**), and κ (**D**), plotted versus the trial M^T_{K0}. The different lines correspond to different trial threshold bandwidths ω^T as indicated in the legends

decreases rapidly, so there are not enough data to get reliable results for large M_{k0} and/or ω. This is the case for the data set we are considering, hence, since no method can extract information from unavailable data, a definitive threshold determination can hardly be expected, but application of the method to the available data gives suggestive results, to be ratified or rebutted when more data are available and is an

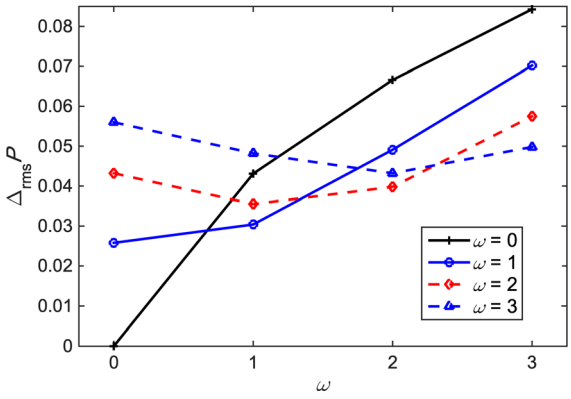

Figure. 8
Comparison of $\Delta_{rms}P$ for $M_{k0} = 6.9$ and different bandwidths ω

illustration of the capabilities of the proposed method.

Figure 9 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 $\omega^T = 0$ correspond to those of Fig. 4 of Gutiérrez et al. (2021), who used only a crisp threshold; their trial threshold magnitudes, M^T, are equivalent to our $M^T_{K1} = M^T_{K0} + 0.1$ for this bandwidth.

We are now faced with the problem of choosing the best model, i.e., the values M_{k0} and ω, that would best fit the observed data. As mentioned above, except for the case of the unrealistic crisp threshold, the peak values of the measured statistics do not occur for M_{k0}, and the correct ω is unknown. However, the synthetics show that the way in which the transition from non-Markovian to Markovian occurs for each trial bandwidth as M^T_{K0} increases is different for different true bandwidths. Hence, we will try to identify the correct parameters for the observed real system, by comparing the whole curves in Fig. 9 with those of synthetic models. In what follows we will only show the comparison with two crisp thresholds: the preferred one of Gutiérrez et al. (2021) and one with the next higher M_{K0}, and the comparisons with the three fuzzy models that best fit the data.

It should be emphasized at this point that the data correspond to only one (very short) realization of a stochastic process, so that the results will not necessarily conform to those shown above for means from the synthetics. Hence, we will use the means from Monte Carlo simulations that used the same amount of data as guides and consider their standard deviations. A further problem is that for the largest ω^T values the results for large M^T_{K0} are not reliable due to the small amount of data.

Figures 11, 12, 13, 14, 15 (see appendix) show, each for a given bandwidth ω and for the trial M^T_{K0} values used above, a comparison between the observed measures and the synthetic means together with their standard deviations.

Figure 11 shows, for the crisp threshold $\omega = 0$ for $M_{K0} = 6.9$, the comparison between the synthetics (Fig. 4) and the observed data (Fig. 9). All observed measures peak for $M^T_{K0} = 7.0$ and the measures for the smaller M^T_{K0} values behave in completely different manner from the synthetics. Clearly, this model is not appropriate to the data.

Since the observed measures for $\omega = 0$ peak at $M^T_{K0} = 7.0$, it is necessary to try the crisp model for $M_{K0} = 7.0$. The comparison is shown in Fig. 12 where peaks now coincide (although the peak for M_6 is not well defined) but the value for S is too high while that for κ is too low, and the behaviors at lower M^T_{K0} values do not agree. This model is not appropriate to the data either.

Other crisp threshold models do not agree with the data either, so larger bandwidths will be tried now.

The observed measures for $\omega = 1$ (Fig. 9) do not fit at all the model for $M_{K0} = 6.9$ (not shown here), but the measures for $M_{K0} = 7.0$ s with $\omega = 1$ (Fig. 13) fit the data reasonably in the $M^T_{K0} = 6.6$–7.0 range, after which there are not enough data. There are not enough data for comparison with larger M_{K0} values for this bandwidth.

Figure 14 shows the comparison for the $\omega = 2$ and $M_{K0} = 6.9$ threshold. Agreement is reasonable in the $M^T_{K0} = 6.6$–7.0 range, but measured values are not reliable for $M^T_{K0} = 7.0$.

Our widest bandwidth $\omega = 3$ is shown in Fig. 15 for $M_{K0} = 6.9$; although fit is much better for the lower M^T_{K0} range, M_6 is a very bad fit and there were not enough data for observed estimations above $M^T_{K0} = 6.9$.

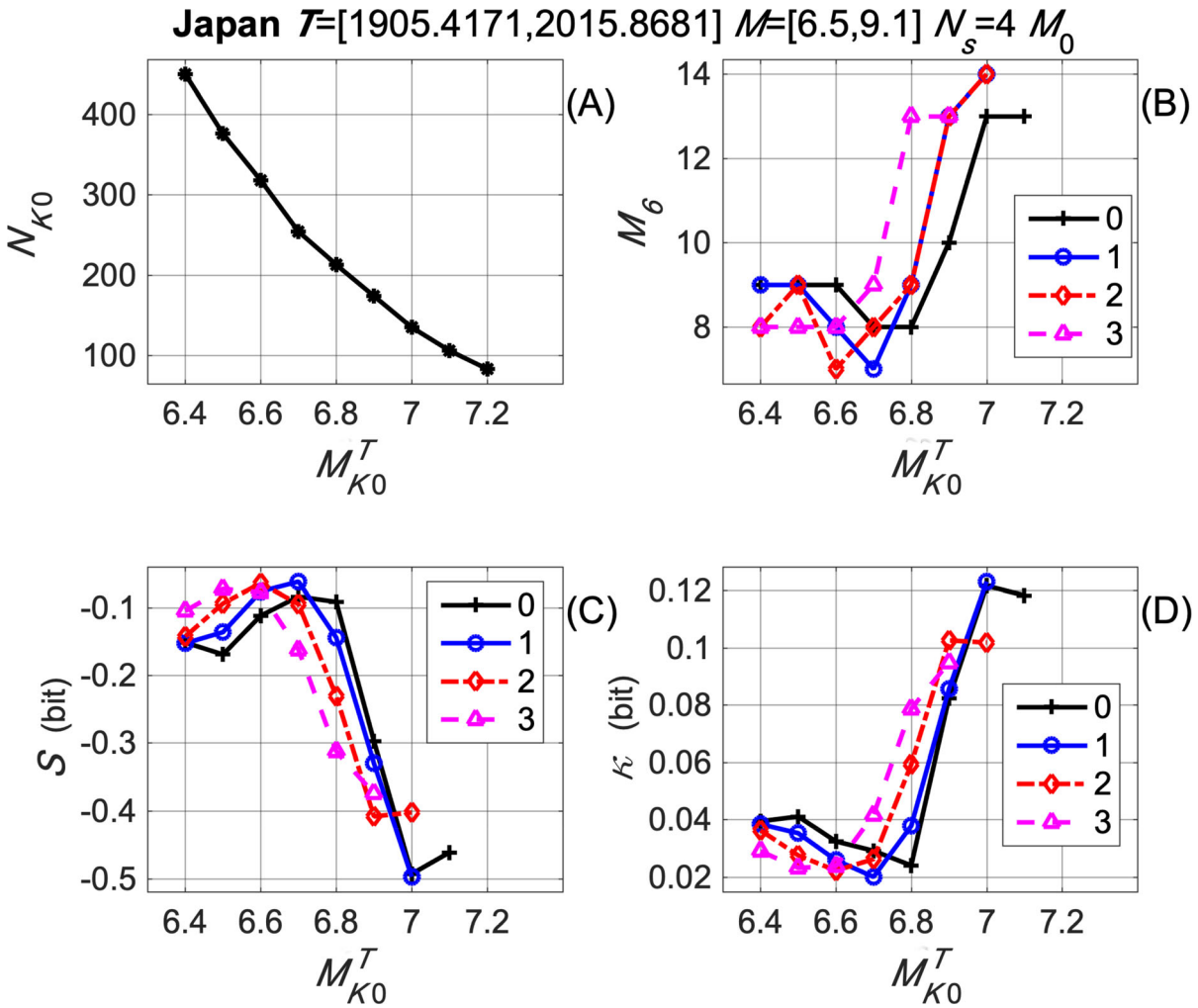

Figure. 9
Application of the counting strategy to real data from Japan. N_{K0}, is shown in (**A**); M_6, S, and κ are shown in (**B**), (**C**), and (**D**), respectively, plotted versus the trial M^T_{K0}. The different lines correspond to different trial threshold bandwidths ω^T indicated in the legends

Thus, from the comparisons we can conclude that a crisp threshold is not compatible with the observed measures, which agrees with our contention that sharp thresholds are not reasonable in nature.

Within the limitations imposed by the small number of data, we tentatively identify models $M_{K0} = 6.9$ with $\omega = 2$ and $M_{K0} = 7.0$ with $\omega = 1$, as those that best adjust the observed measurements. More data would be needed to make a reliable choice between these two models, but it is clear that a fuzzy threshold is preferable to a crisp one.

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.

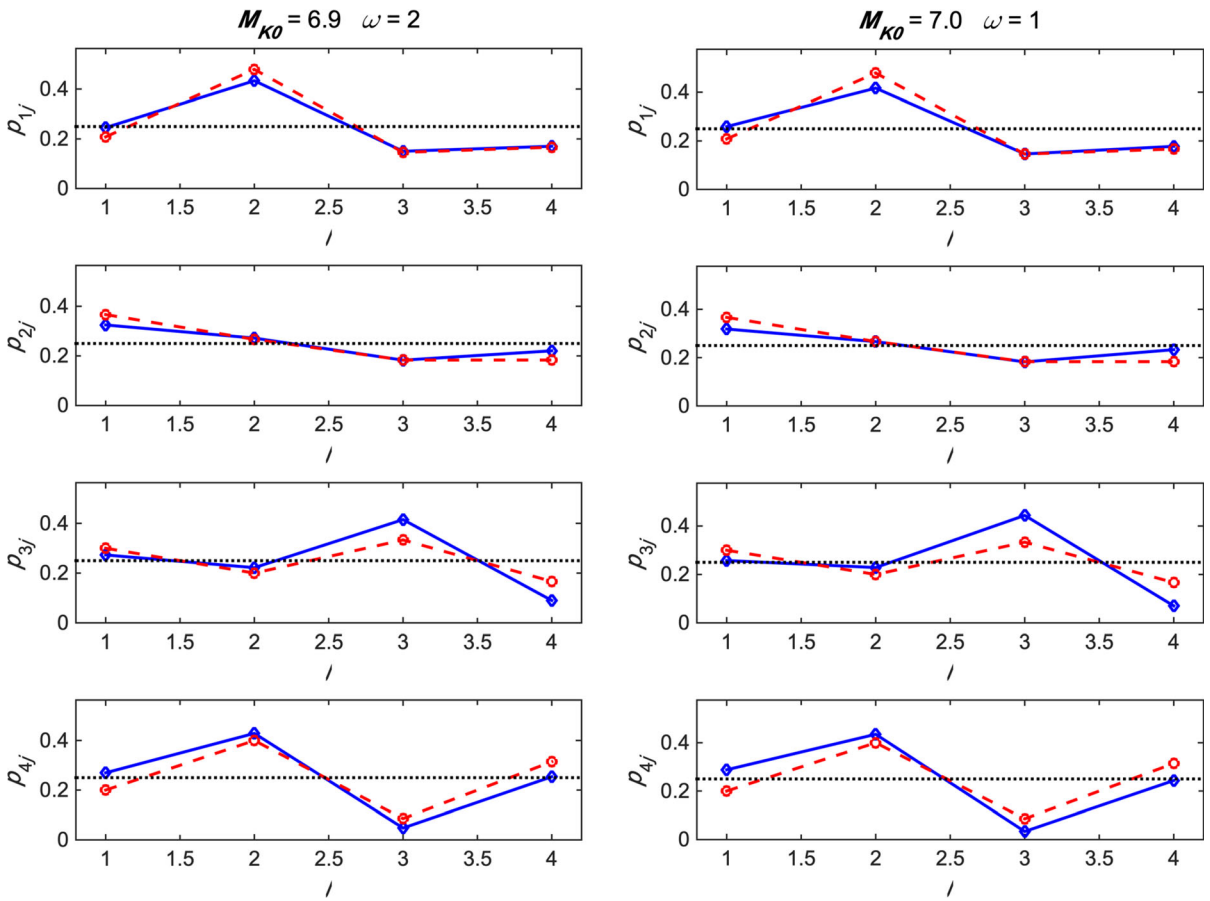

Figure. 10
Transition probability matrices shown row by row; P_G is shown as the dashed red line with circles, P is the blue line with diamonds for $M_{K0} = 6.9$ and $\omega = 2$ (left) and for $M_{K0} = 7.0$ and $\omega = 1$ (right), and the reference uniform probability $p_U = 0.25$ is the dotted black line

The usual method for evaluating empirical transition probabilities assumes all events in the observed sequence above a given threshold 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 appropriateness of the counting strategy was validated by applying it to synthetic state chains created according to different threshold models and, in all cases, the difference between estimated and optimal TPMs was smallest for the combination of trial parameters that matched those of the model. The Monte Carlo simulations show that peaks in the Markovianity measures are not good indicators of threshold magnitudes, so the whole shape of the measure vs. M_{K0}^T should be used to identify the threshold.

Next, we applied the counting strategy to a set of real data, and found that the measures 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, and any one realization can be expected to differ from the mean of many realizations. The very short length of the data chain precludes exploring models with large threshold magnitudes and/or wide acceptance bandwidths.

Taking the above into account, we find that the empirical results resemble most the simulations for

$M_{K0} = 6.9$ with $\omega = 2$ and $M_{K0} = 7.0$ with $\omega = 1$ 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 0.1 to 0.2 magnitude units wide.

Use of these models 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

$$P_G = \begin{bmatrix} 0.2083 & 0.4792 & 0.1458 & 0.1667 \\ 0.3667 & 0.2667 & 0.1833 & 0.1833 \\ 0.3000 & 0.2000 & 0.3333 & 0.1667 \\ 0.2000 & 0.4000 & 0.0857 & 0.3143 \end{bmatrix}$$

to, for $M_{K0} = 6.9$ and $\omega = 2$,

$$P = \begin{bmatrix} 0.2457 & 0.4334 & 0.1504 & 0.1705 \\ 0.3248 & 0.2718 & 0.1829 & 0.2205 \\ 0.2730 & 0.2219 & 0.4151 & 0.0900 \\ 0.2697 & 0.4288 & 0.0478 & 0.2538 \end{bmatrix},$$

or, for $M_{K0} = 7.0$ and $\omega = 1$,

$$P = \begin{bmatrix} 0.3000 & 0.3667 & 0.1333 & 0.2000 \\ 0.3429 & 0.2286 & 0.1714 & 0.2571 \\ 0.1667 & 0.2778 & 0.4444 & 0.1111 \\ 0.2727 & 0.4545 & 0.0454 & 0.2273 \end{bmatrix}.$$

Figure 10 illustrates the probabilities for our preferred models as compared with the probabilities for the crisp threshold PG, and with the reference uniform probabilities. Note that, although the fuzzy thresholds are somewhat similar, but not equal, their TPMs are very much alike and differ in the same ways from the PG probabilities. Hence, in describing the probability changes from crisp to fuzzy we will use the plural to refer to both fuzzy threshold probabilities.

After an event in Region 1, the probabilities of it repeating in the same region are increased, while those of being followed by an event in Region 2 (the largest in the TPM) are decreased; those for transitions to the other farther regions remain essentially the same.

After an event in Region 2, the probabilities of a repetition decrease slightly, while those of a transition to Region 4 increase slightly.

After an event in Region 3, the probabilities of a repetition, the largest, increase considerably, while those of a transition to Region 4 decrease significantly.

After an event in Region 4, the probabilities of a repetition or to a transition to Region 3 decrease significantly, while, surprisingly, those for a transition to Region 1 increase.

We submit that, despite the limitations of working with a single short observed chain, the estimates of transition probabilities from fuzzy thresholds are preferable to those from a crisp threshold, because their models are more realistic and because they fit the synthetic results better.

The counting strategy might be useful in some other applications where crisp thresholds are not acceptable.

Acknowledgements

We are grateful to the International Seismological Centre for the use of their catalog. Our thanks to an anonymous reviewer and to the Editor Costas Papazachos.

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."

Funding

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

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Appendix

See Figs. 11, 12, 13, 14, 15.

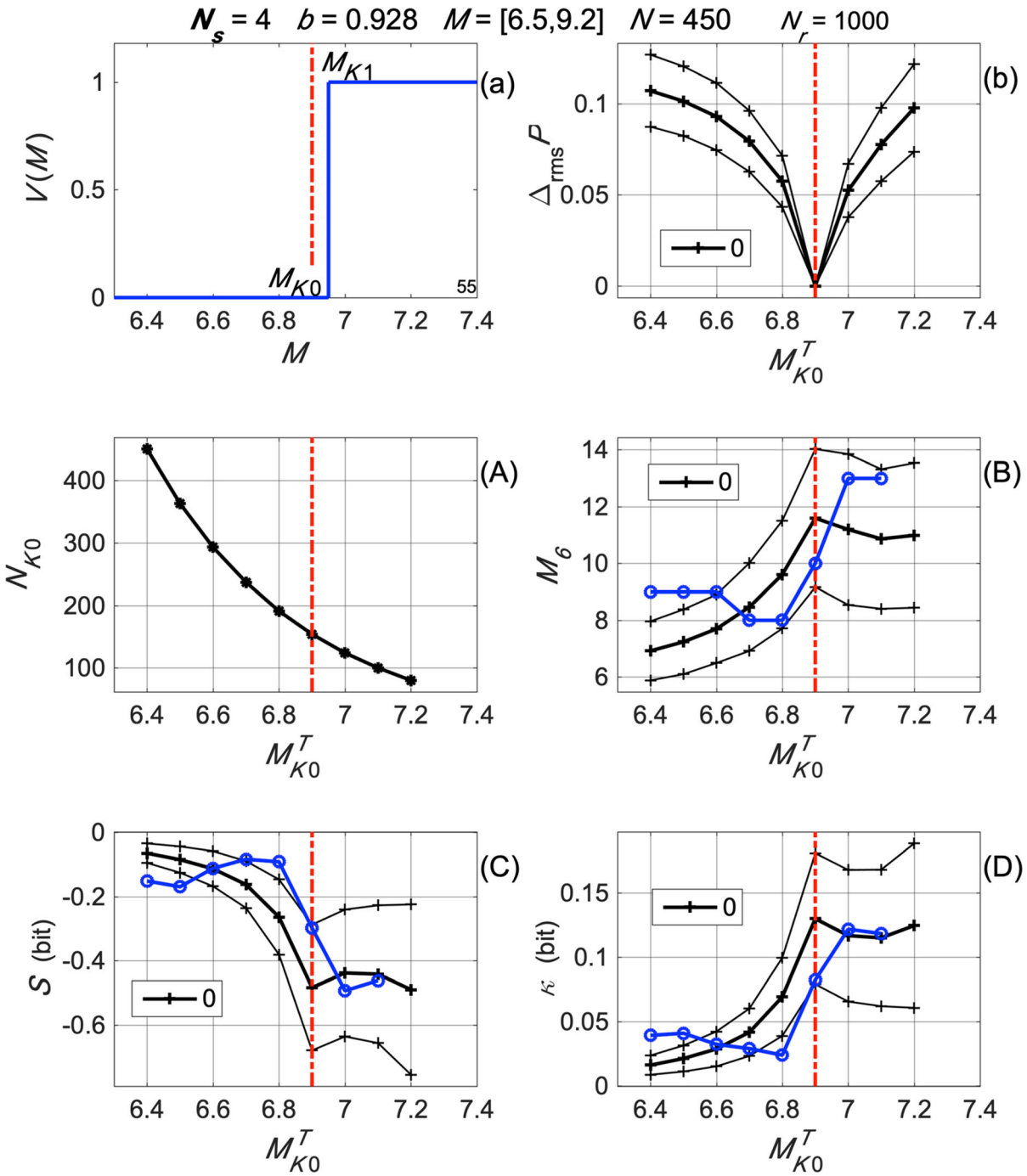

Figure. 11

Synthetic threshold with $\omega = 0$ and $M_{K0} = 6.9$ (**a**) and differences between estimated TPMs and optimum TPMs (**b**). Number of magnitudes used in the analysis as a function of trial initial magnitude M^T_{K0} (**A**). Comparison of observed measures, blue lines with circles in (**B**) to (**D**), with synthetic Monte Carlo means (thick black lines) and means plus/minus one standard deviation (thin black lines) for the threshold

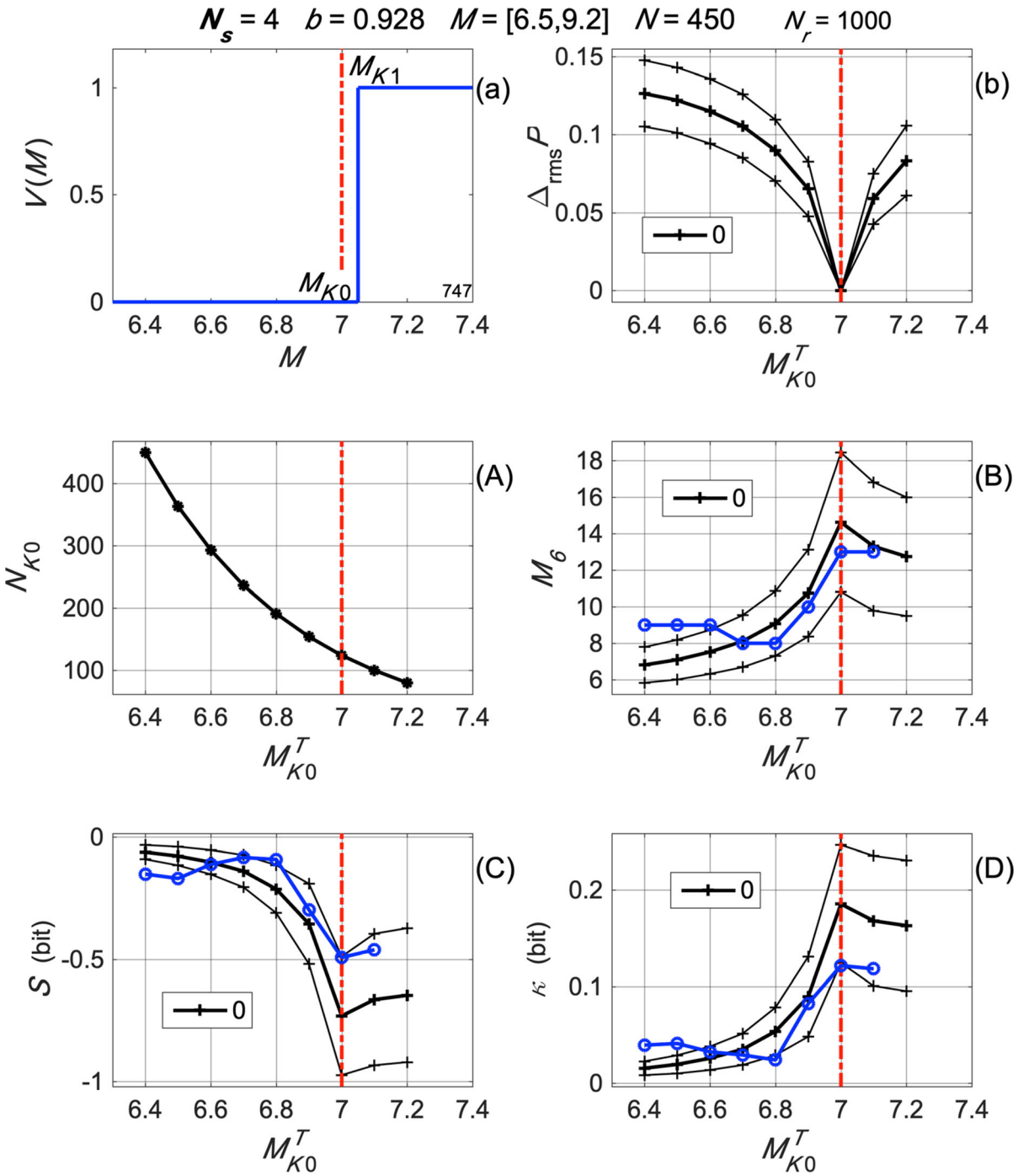

Figure. 12

Synthetic threshold with $\omega = 0$ and $M_{K0} = 7.0$ (**a**) and differences between estimated TPMs and optimum TPMs (**b**). Number of magnitudes used in the analysis as a function of trial initial magnitude M_{K0}^{T} (**A**). Comparison of observed measures, blue lines with circles in (**B**) to (**D**), with synthetic Monte Carlo means (thick black lines) and means plus/minus one standard deviation (thin black lines) for the threshold

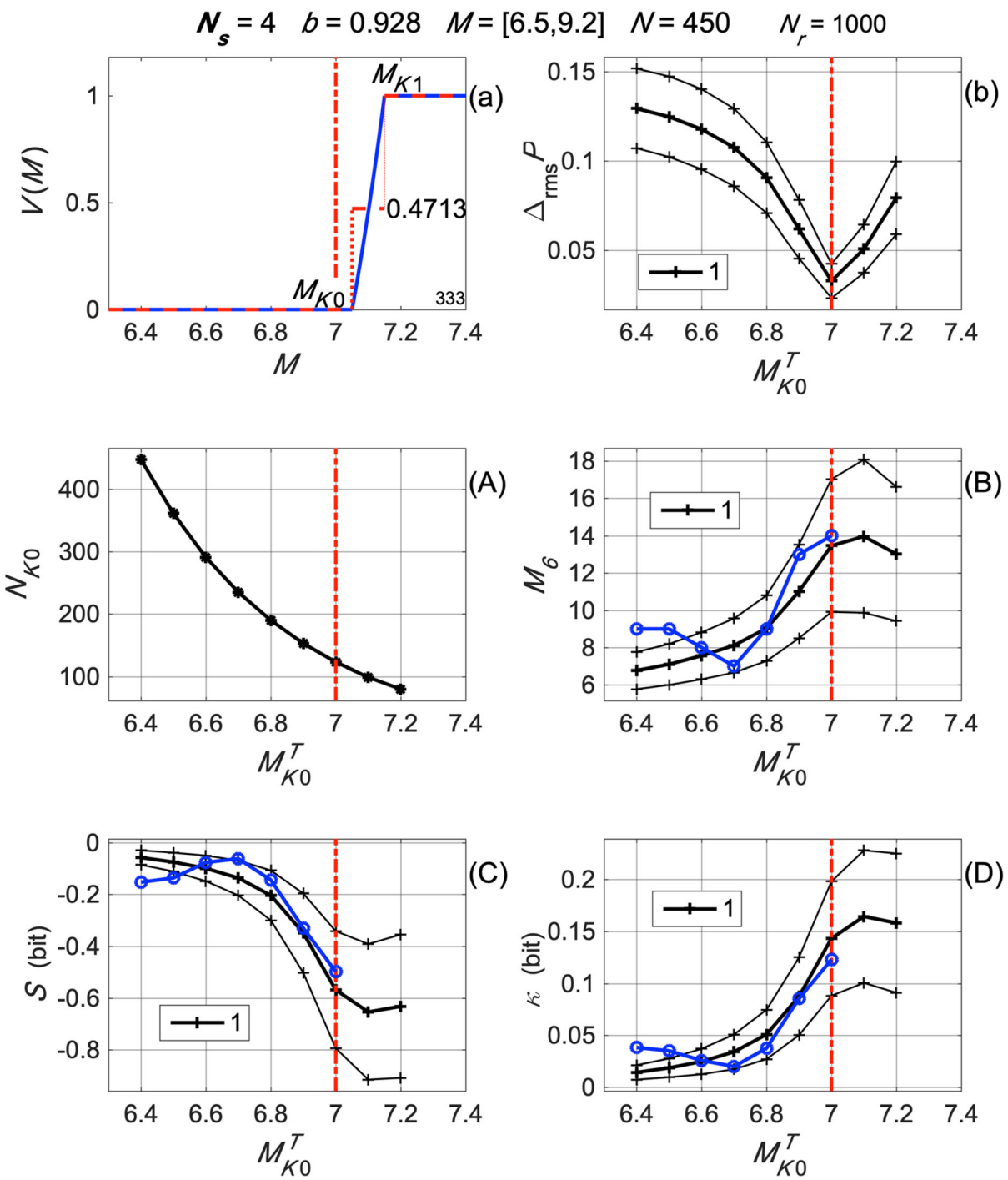

Figure. 13

Synthetic threshold with $\omega = 1$ and $M_{K0} = 7.0$ (**a**) and differences between estimated TPMs and optimum TPMs (**b**). Number of magnitudes used in the analysis as a function of trial initial magnitude M_{K0}^{T} (**A**). Comparison of observed measures, blue lines with circles in (**B**) to (**D**), with synthetic Monte Carlo means (thick black lines) and means plus/minus one standard deviation (thin black lines) for the threshold

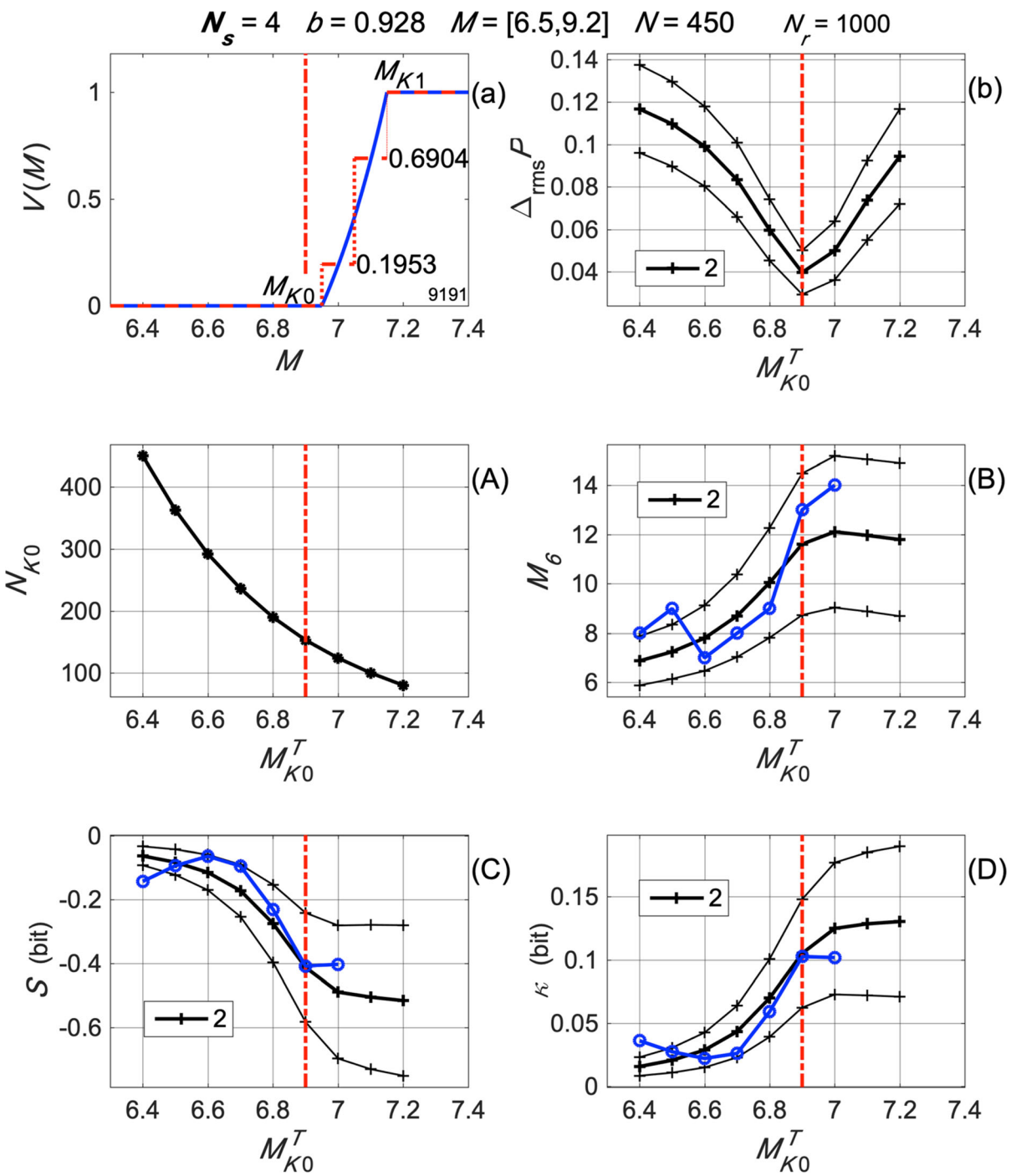

Figure. 14

Synthetic threshold with $\omega = 2$ and $M_{K0} = 6.9$ (**a**) and differences between estimated TPMs and optimum TPMs (**b**). Number of magnitudes used in the analysis as a function of trial initial magnitude M_{K0}^T (A). Comparison of observed measures, blue lines with circles in (**B**) to (**D**), with synthetic Monte Carlo means (thick black lines) and means plus/minus one standard deviation (thin black lines) for the threshold

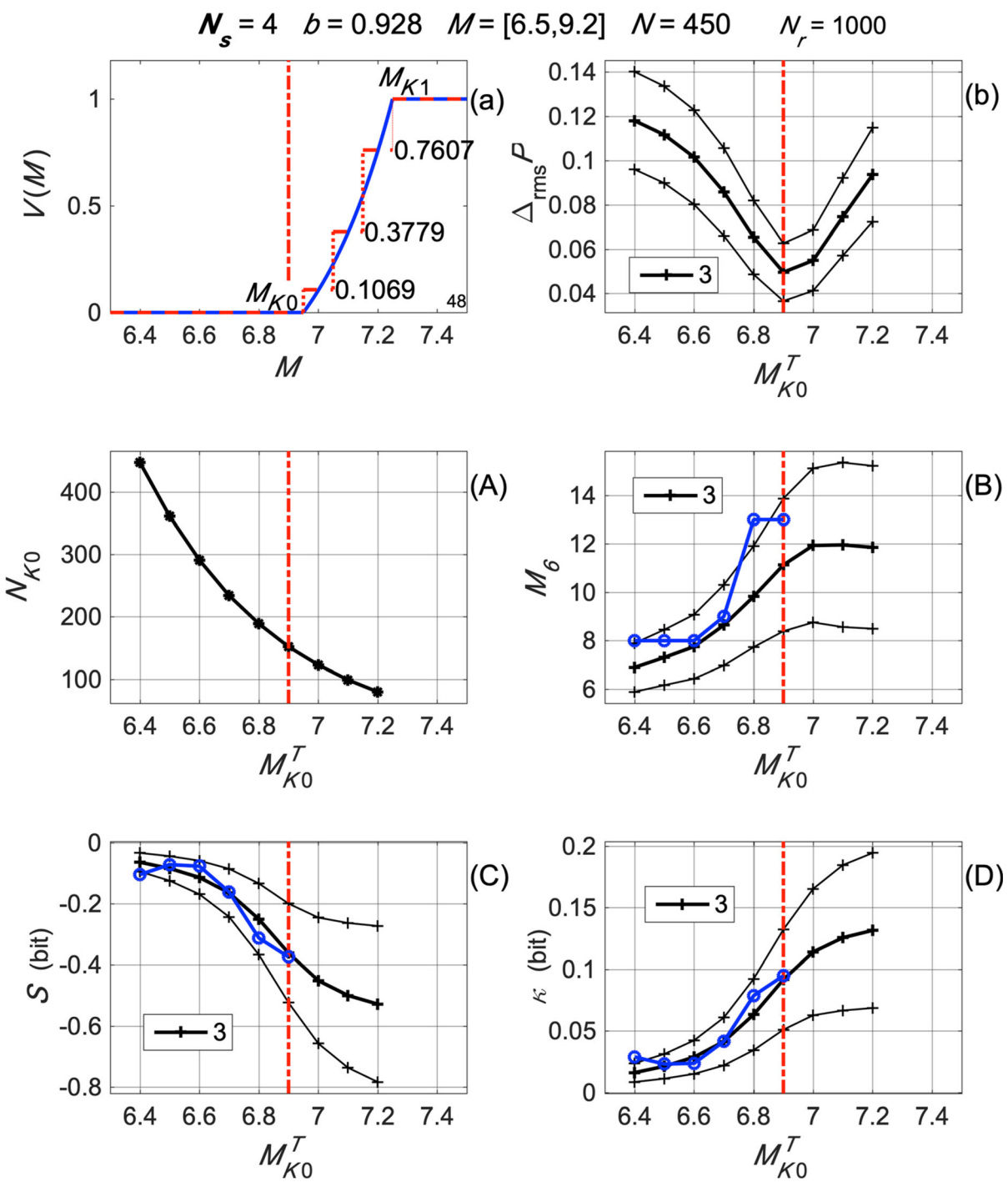

Figure. 15

Synthetic threshold with $\omega = 3$ and $M_{K0} = 6.9$ (**a**) and differences between estimated TPMs and optimum TPMs (**b**). Number of magnitudes used in the analysis as a function of trial initial magnitude M_{K0}^T (**A**). Comparison of observed measures, blue lines with circles in (**B**) to (**D**), with synthetic Monte Carlo means (thick black lines) and means plus/minus one standard deviation (thin black lines) for the threshold

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 Oct 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* (p. 323). Italia, Milano: Springer-Verlag.

Cavers, M., & Vasudevan, K. (2014). Spatio-temporal complex Markov chain (SCMC) model using directed graphs: Earthquake sequencing. *PAGEOPH*. https://doi.org/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* (p. 205). Springer Science+Business Media Inc.

Feller, W. (1968). *An introduction to probability theory and applications* (Vol. 2, pp. 509–669). Wiley.

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)* (5th ed.). AMS Chelsea Publishing.

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. *Journal of Geophysical Research, 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. *Earth, Planets and Space, 58*, 973–979.

Kullback, S., & Leibler, R. (1951). On information and sufficiency. *Annals of Mathematical Statistics, 22*(1), 79–86.

Lehner, F., Li, V., & Rice, J. (1981). Stress diffusion along rupturing plate boundaries. *Journal of Geophysical Research, 86*(B7), 6155–6169.

Leptokaropoulos, K., Adamaki, A., Roberts, R., & Gkarlaouni, C. (2018). Impact of magnitude uncertainties on seismic catalogue properties. *Geophysical Journal International, 213*, 940–951.

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. *Geophysical Research Letters*. https://doi.org/10.11029/2002GL014702

Melbourne, T., Webb, F., Stock, J., & Reigber, C. (2002). Rapid postseismic transients in subduction zones from continuous GPS. *Journal of Geophysical Research, 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.* Wiley.

Patwardhan, A., Kulkarni, R., & Tocher, D. (1980). A semi-Markov model for characterizing recurrence of great earthquakes. *Bulletin of the Seismological Society of America, 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. https://doi.org/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.

Ringdal, F. (1976). Maximum-likelihood estimation of seismic magnitude. *Bulletin of the Seismological Society of America, 66*(3), 789–802.

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

Werner, M., & Sornette, D. (2008). Magnitude uncertainties impact seismic rate estimates, forecasts, and predictability experiments. *Journal of Geophysical Research: Solid Earth*, *113*(B8), B08302. https://doi.org/10.1029/2007JB005427

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

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

(Received November 16, 2022, revised May 9, 2024, accepted July 1, 2024)