Geophysical Journal International

Geophys. J. Int. (2019) **216,** 820–839
Advance Access publication 2018 October 31
GJI Seismology

doi: 10.1093/gji/ggy459

Empirical evaluation of the magnitude-independence assumption

A. Stallone [1,2] and W. Marzocchi [1]

[1] *Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy. E-mail: angela.stallone@ingv.it*
[2] *University of Roma TRE, Department of Earth Sciences, Rome, Italy*

Accepted 2018 October 30. Received 2018 October 29; in original form 2018 April 09

SUMMARY

The most common earthquake forecasting models assume that the magnitude of the next earthquake is independent from the past. This feature severely limits the capability to forecast large earthquakes with high probabilities. Here we investigate empirically on the magnitude-independence assumption, exploring if: (i) background and triggered earthquakes have the same frequency–magnitude distribution, (ii) variations of seismicity in the space–time–magnitude domain encode some information on the future earthquakes size. For this purpose, and to verify the stability of the findings, we consider seismic catalogues covering different space–time–magnitude windows, such as the Alto Tiberina Near Fault Observatory (TABOO), the California and Japanese seismic catalogues. Our approach is inspired by the nearest-neighbour method proposed by Baiesi & Paczuski and elaborated by Zaliapin *et al.* to distinguish between triggered and background earthquakes. Here we implement the same metric-based correlation to identify the precursory seismicity of any triggered earthquake; this allows us to analyse, for each triggered earthquake, the space–time–magnitude distribution of the seismicity that likely contributed to its occurrence. Our results show that the magnitude-independence assumption holds reasonably well in all catalogues, with a remarkable exception that is consistent with a previous independent study; this departure from the magnitude-independence assumption shows that larger events tend to nucleate at a higher distance from the ongoing sequence. We also notice that the reliability of this assumption may depend on the spatial scale considered; it holds for seismic catalogues of large areas, but we identify possible departures in small areas, reflecting different ways to release locally seismic energy. Finally, we come across an important issue that may lead to misleading results in similar studies, that is, if a seismic catalogue appears overall complete above a fixed magnitude threshold, it may still yield spurious signals into the analysis. Specifically, we show that some significant departures from the magnitude-independence assumption do not survive when considering spatiotemporal variations of the magnitude of completeness.

Key words: Persistence, memory, correlations, clustering; Probability distributions; Spatial analysis; Statistical methods; Earthquake interaction, forecasting, and prediction; Statistical seismology.

1 INTRODUCTION

The most reliable operational earthquake forecasting models are rooted in the idea that earthquakes (any of them or those above a certain magnitude) can generate other earthquakes, and that the triggering probability depends on time, space and magnitude (Ogata 1998; Gerstenberger *et al.* 2005; Marzocchi *et al.* 2017). The mathematical description of such a clustering process consists of a few well-known empirical scaling laws such as the Gutenberg & Richter (1944), the Omori (1894) and the Utsu (1957) relationships (Ogata 1988; Ogata 1998; Reasenberg & Jones 1989; Gerstenberger *et al.* 2005). These clustering models assume that the magnitude of any earthquake is independent from the past (the magnitude-independence assumption hereafter) (Ogata 1988; Kagan & Knopoff 1987; Reasenberg & Jones 1989; Zhuang *et al.* 2002). This implies that, in such models, the intensity function is magnitude separable, that is the magnitude distribution (usually, but not necessarily described by the Gutenberg–Richter law) is not dependent on the past (Schoenberg 2003) . As an example, in the ETAS model (Ogata 1988; Ogata 1998), the conditional seismic intensity λ can be written as:

$$\lambda(t, \overline{x}, m|H_t) = f(m)\left[\mu(\overline{x}) + \Sigma g(t - t_i, \|\overline{x} - \overline{x_i}\|, m_i)\right], \quad (1)$$

where $H_t = (t_i, \overline{x_i}, m_i) : t_i < t$ represents the history of the process up to time t, $f(m)$ is the function describing the magnitude-frequency distribution, $\mu(\overline{x})$ represents the background seismicity rate, and g is the triggering function describing the productivity of an event of magnitude m_i and the spatiotemporal distribution of its offspring. Eq. (1) clearly shows that the separability of the magnitude distribution [$f(m)$ is not influenced by past events] is implied by the magnitude-independence assumption, stating that all the magnitudes are randomly drawn from the same magnitude-frequency distribution, no matter what the triggering event magnitude (m_i) is.

As a matter of fact, this assumption represents the strongest limitation to the capability to forecast large earthquakes in the short term with clustering models (Michael 2012), because the probability of a large earthquake is constrained by the probability of sampling the right tail of an exponential distribution. More in general, if the magnitude of an earthquake is purely random, that is, independent from the past and from the boundary conditions, this would hamper any attempt to forecast large earthquakes with high probability.

The reliability of this assumption has been investigated in previous studies, returning somehow diverging results. For instance, Felzer *et al.* (2004) and Helmstetter *et al.* (2003) show no difference in the frequency–magnitude distributions of triggered and background events, as is implied by the magnitude-independence assumption. Conversely, more recent papers highlight some possible discrepancies with respect to the magnitude-independence assumption. Zhuang *et al.* (2004) find different frequency–magnitude distribution for triggered and background events, whereas Lippiello *et al.* (2008), Lippiello *et al.* (2012a) and Spassiani & Sebastiani (2016) find a pairwise correlation between consecutive magnitudes. The physical motivation of such discrepancies with the magnitude-independence assumption is still unclear (e.g. Shearer 2012; de Arcangelis *et al.* 2016). We note, however, that these results, though being statistically significant, do not seem to carry out large probability gains with respect to the magnitude-independence assumption (Ogata *et al.* 2018). Moreover, the pairwise magnitude correlation is around 0.1 or less and observed only over short time intervals (of the order of tens of minutes), after which it almost vanishes (Corral 2006; Lippiello *et al.* 2012a). Also for the spatial distribution of the seismicity that precedes large shocks, the results seem to be controversial. On one hand, Helmstetter & Sornette (2003) and Helmstetter *et al.* (2003) do not find any clear correlation between foreshock features and the Felzer *et al.* (2004) find that the foreshocks area does not depend on the mainshock magnitude. On the other hand, Lippiello *et al.* (2012b) and Lippiello *et al.* (2017) find possible evidence that the spatial distribution of the foreshock activity encodes information on the magnitude of the following main shocks. Likewise, Shearer (2012) shows some differences in the spatial distribution of earthquakes following larger shocks, with respect to what suggested by classical clustering models. The debate is still ongoing, and it is not yet clear whether these discrepancies may eventually lead to significant improvement of the earthquake forecasting skill with respect to the use of the magnitude-independence assumption.

In this work, we investigate on the reliability of the magnitude-independence assumption by analysing seismicity features in time-space windows that are relevant for operational earthquake forecasting. Specifically, instead of investigating possible pairwise correlations between consecutive earthquakes that may be very close in time, we analyse the characteristics of the whole seismic sequence that precedes earthquakes with different magnitudes; the goal is to check if these sequences have similar features independent from the magnitude of the triggered earthquakes. We use an innovative empirical approach, which has some features worth being remarked. First, it does not consider the arbitrary distinction among aftershocks, mainshocks and foreshocks, but only between independent and triggered earthquakes; this is done through a technique that minimizes the choice of arbitrary thresholds and parameters (these choices will be described in Section 3), avoiding the subjective definition of space and time binnings. Second, it explores several dimensions of the triggering seismic activity, namely the magnitude and the space–time distribution, through formal statistical testing methods. Third, it allows a check of the robustness and stability of the results considering three independent catalogues that cover a different space–time–magnitude range, that is, the inland Japanese catalogue, the South California catalogue, and the catalogue of a small and well monitored part of Central Italy (TABOO).

Specifically, we aim to address the following scientific questions: (i) Are the magnitude distributions of triggered and background earthquakes equivalent, as predicted by the magnitude-independence assumption? (ii) Is the preceding phase of larger triggered earthquakes characterized by different time–space–magnitude features with respect to the preceding phase of smaller triggered earthquakes? The first step of the analysis consists of separating the triggered and background earthquakes of a seismic catalogue. For each triggered event, we then identify the set of one or more earthquakes which are thought to directly contribute to its occurrence. We address the first question by statistically comparing the magnitude distribution of triggered and background earthquakes, whereas we address the second question by analysing the space–time–magnitude seismicity patterns that precede small and large triggered earthquakes. In both cases, we apply our analysis to a set of synthetic ETAS catalogues as well. Since the ETAS model is rooted in the magnitude-independence assumption (eq. 1), the analysis of these catalogues is a sort of sanity check: the null hypothesis of each test, which is related to the magnitude-independence assumption, should never be rejected for the ETAS catalogues.

2 SEISMIC CATALOGUES

The simultaneous use of different seismic catalogues allows us to explore the stability of the results on a wide space–time–magnitude range. Specifically, we consider three seismic catalogues covering different space–time–magnitude windows: (1) the Alto Tiberina Near Fault Observatory earthquake catalogue (TABOO, Chiaraluce *et al.* 2014), (2) the southern California catalogue (SOCAL, Hauksson *et al.* 2012), (3) the Japanese catalogue (JMA catalogue). The TABOO earthquake catalogue is a high-quality and high-resolution data set recorded by a dense seismic network in the area of the Alto Tiberina fault (northern Apennines, Italy), covering the period 2010–2015 and including 43 067 earthquakes. We have estimated a magnitude of completeness M_c of 0.7 (which leaves 4661 recorded earthquakes) and a b-value

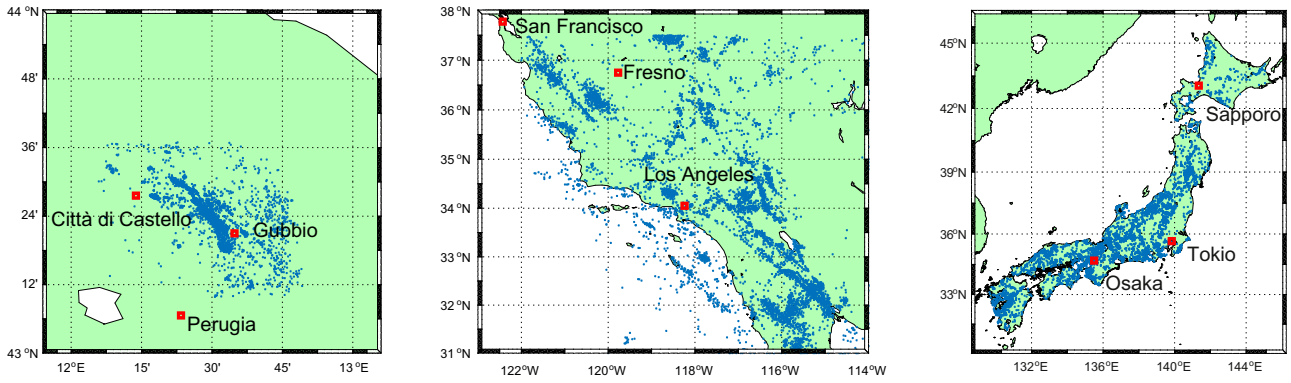

Figure 1. Seismicity maps of the three catalogues, from left to right: (1) TABOO catalogue, (2) SOCAL catalogue and (3) JMA catalogue.

equals to 0.96. The maximum magnitude is 3.8. The SOCAL data set is a relocated catalogue provided by SCSN for the period 1981–2014, including 567 258 earthquakes. The estimated value for M_c is 3 (which leaves 14 472 earthquakes), whereas the estimated b-value is 1.02. The maximum magnitude is 7.3. The JMA catalogue is provided by the Japan Meteorological Agency and includes 1 831 412 earthquakes; we have extracted all the shallower earthquakes (less than 30 km; mostly inland) occurred after 1990. We estimate $M_c = 2.5$ (which leaves 18 004 earthquakes) and a b-value equals to 0.84. The maximum magnitude is 7.3.

For all the three catalogues, M_c is estimated as the first magnitude value such that the hypothesis of exponentially distributed data cannot be rejected at a significance level of 0.05 (Lilliefors test, Lilliefors 1969; Clauset *et al.* 2009). The b-value has been estimated for the exponentially distributed magnitudes by implementing the maximum-likelihood formula provided by Marzocchi & Sandri (2003). As a matter of fact, the M_c estimated for each catalogue is an average value, which could not be truly representative for each space and time subset of the whole seismic catalogue. This variability may introduce spurious signals in the data (Kagan 2004). Several factors can contribute to increase M_c for specific portions of the catalogue:

(1) underreporting of small events after large earthquakes (short term aftershock incompleteness, STAI; Kagan 2004);
(2) lower local density of seismic stations (spatial variation of M_c, Wiemer & Wyss 2000);
(3) lower-quality and lower-density seismic stations available at older times (time-varying quality of the catalogue).

In this work, we address all these points in order to prevent the possibility that our findings are biased by the possible space and time variability of the real magnitude of completeness. As regards the STAI issue, we remove all the events occurring within a short space–time window from $M6+$ events. To limit any subjectiveness, we use different criteria for the definition of this window (see Appendix B for more details). Depending on the catalogue under consideration, we handle the issues 2 and 3 by applying our analysis to events occurred within a region characterized by a high seismic station density and/or to an excerpt of the catalogue which does not include old records that have poorer quality. As an additional test, we verify the stability of the results imposing a more conservative overall minimum threshold magnitude ($M_c^\star$), that is, $M_c^\star > M_c$ (TABOO: $M_c^\star = 1.1$ in place of $M_c = 0.7$; SOCAL: $M_c^\star = 3.5$ in place of $M_c = 3$; JMA: $M_c^\star = 3$ in place of $M_c = 2.5$). Using $M_c^\star$ should minimize all the problems described above; at the same time, if the discrepancy between the magnitude distribution of triggered and background earthquakes stems from a different physical process, then it should be preserved after this modification. If it is not, it is likely a consequence of the underreporting of the smaller events caused by any of the three issues described above. Note that increasing M_c of about 0.5 units lowers the number of data of about of factor of 3; this means that the number of data are still enough to carry out meaningful statistical tests.

It is worth of noting that the three catalogues are optimal for the purpose of our work, for they cover different space, time and magnitude windows. Fig. 1 shows the seismicity maps of the three recorded catalogues. For the sake of comparison with the expectation of typical clustering models and to check if the adopted declustering technique introduces some spurious signals in our analysis, we also carry out our analysis on a synthetic ETAS catalogue (see Appendix A for details about how we simulate it). Since we use a stochastic simulator program, we generate multiple datasets. In Section 5, we show the findings relative to only one ETAS catalogue (7955 events with magnitudes in the range 3–7.3), while we summarize the outputs from the other synthetic catalogues in the Supporting Information.

3 IDENTIFYING TRIGGERED EARTHQUAKES AND THEIR PRECEDING SEISMICITY

3.1 Declustering method

Background earthquakes represent the independent component of seismicity, which is generally related to the regional tectonic strain rate. On the contrary, triggered earthquakes occur in space–time clusters and are mostly caused by previous earthquakes. It is unlikely that any earthquake can be either entirely due to tectonics or entirely triggered; therefore any method used to separate triggered and background

earthquakes unavoidably contains some degree of subjectivity (Zhuang *et al.* 2002). Although we argue that any reasonable choice to make this distinction can hardly introduce some meaningful signals into the data set, we check this possibility by analysing the performance of the declustering procedure applied to synthetic seismic catalogues where we know which earthquake belongs to the background and which one is triggered. Here we implement the original statistical method proposed by Baiesi & Paczuski (2004) and elaborated by Zaliapin *et al.* (2008) afterward, namely the Nearest-Neighbour (NN) Clustering Analysis. This choice is mostly motivated by the elegance of the procedure, its simplicity, and its capability to identify the link between triggered and triggering earthquakes in a more objective way with respect to the commonly used methods that involve subjective choices on the dimension of the space–time–magnitude windows for the identification of foreshocks, main shocks and aftershocks (Van Stiphout *et al.* 2012). The approach invokes the use of a distance defined in the space–time–magnitude domain to quantify the correlation between any couple of events. For an earthquake j, the distance with respect to any of its ancestors i can be written as:

$$n_{ij} = t_{ij} r_{ij}^{d_f} 10^{-bm_i}, \tag{2}$$

where t_{ij} is the difference between i and j occurrence times, r_{ij} is the distance between the two epicentres, d_f is the epicentral fractal dimension, b is the Gutenberg-Richter b-value and m_i is the magnitude of the ancestor. The Nearest-Neighbour distance (NND) is then defined as:

$$\mathrm{NND}_{i^* j} = \min(n_{ij}), \tag{3}$$

where $i*$ is the one, among the ancestors, which minimizes n_{ij} and is defined as the nearest-neighbour of earthquake j. Triggered events are then identified as those being unexpectedly close, that is strongly correlated, to their nearest neighbour. A threshold value τ for the distance must be defined, such that all the NND values lower than τ pinpoint the strongly correlated earthquake pairs. As shown by Zaliapin *et al.* (2008), the NND distribution is clearly bimodal, with one mode representing the triggered earthquakes and the other one the background seismicity. The fit of a two-component Gaussian mixture model to the distribution returns a threshold value which separates the two subpopulations. The clustered portion of seismicity is therefore extracted by simply keeping all those earthquakes whose distance with respect to their nearest neighbour falls below τ. Hereafter we use the term distance referring to the distance among earthquakes in the space–time–magnitude domain. Though the quality and stability of the NN method has already been deeply investigated (see the Auxiliary material for Zaliapin & Ben-Zion 2013), we verify the stability of the method by running the same analysis with different values of the parameters b and d_f.

3.2 The *K*-NN method

Although the NN procedure is sufficient to identify clustered and background seismicity, in this work we boost this method by introducing the K-Nearest Neighbours (K-NN hereafter) variation that allows us to identify any possible distinctive patterns of the seismicity which precedes triggered earthquakes of different magnitude. The rationale of K-NN technique is rooted in the fact that, from a physical point of view, it is hardly conceivable that one triggered earthquake can be linked only to one single ancestor and the contribution of all other ancestors is negligible; the idea that more ancestors contribute, at a different degree, to the triggering of one earthquake stands also behind the class of ETAS models, which consider the cumulative triggering effect of the past earthquakes, not only of the most influential. In the NN approach, the prerequisite for an earthquake to be triggered is to have a distance from its NN which is less than the identified threshold τ: when this condition is met, we say that the two events are correlated. We use the same threshold-based approach to identify all the prior events that are correlated to a given triggered event j: if K ($K \geq 1$) ancestors are unexpectedly close to j ($d_{ij} \leq \tau$, where i is the ith of the K-NN), we can assume that they all have contributed to the occurrence of that given event. This enables us to use the identified K-NN as good candidates for the definition of the preceding seismicity of a given triggered earthquake. Throughout this paper we will refer to these ancestors as triggering or K-NN earthquakes, and to a triggered earthquake with more than one triggering earthquake as a highly connected earthquake. We intentionally avoid the common nomenclature which refers to foreshocks–main shock–aftershocks, for we do not impose any limitations on the relative size, time, or location of earthquakes. Two conditions only must be met: (1) all the K-NN of a given earthquake must have occurred before it; (2) their distance (in space–time–magnitude domain) from the triggered event must be less than τ. This approach should minimize possible misinterpretations linked to arbitrary definitions of foreshocks–main shock–aftershocks. As regards the distinction between triggered and background earthquakes, our method is equivalent to the NN method. The K-NN modification comes into play when identifying the precursory seismicity of a given triggered earthquake: while the NN method only focuses on the correlation between an event and its nearest neighbour (which is sufficient for declustering purposes), its K-NN variation focuses on all the strongly correlated K neighbours of a given triggered event (as an alternative to foreshocks identification methods). In other words, by implementing the same metric-based approach, we separate background and triggered events (NN method), and we identify the precursory seismicity of the latter (our add-on).

4 TESTING THE MAGNITUDE-INDEPENDENCE ASSUMPTION

In this section we describe the strategy to test the magnitude-independence assumption. Specifically, we aim to give answer to the following two questions:

(1) Q1: are the magnitude distributions of triggered and background earthquakes equivalent, as predicted by the magnitude-independence assumption, or are they statistically different?

(2) Q2: Is the preceding phase of triggered larger earthquakes characterized by different time–space–magnitude features with respect to the preceding phase of triggered smaller earthquakes?

4.1 Q1 – Looking for possible differences on the magnitude distribution of background and triggered earthquakes

The magnitude-independence assumption implies that the magnitude distribution of triggered and background earthquakes must be the same. To verify this hypothesis, we implement multiple statistical tests, each one characterized by a different power (i.e. they are sensitive to different kinds of variability between the two distributions). First, we apply the non-parametric two-sample Kolmogorov–Smirnov test (KS2 test, Smirnov 1939), which is aimed to detect any kind of difference between the two distributions; specifically, we test the null hypothesis that the two distributions are samples from the same population by comparing the p-value of the test with the significance level $\alpha = 0.01$ (notice that we use this low significance level across the paper to minimize the finding of spurious significant results that are more likely when multiple tests are carried out). Then, we inspect whether the discrepancy, if any, is caused by the two distributions being different (one of them is not exponential), by implementing the Lilliefors test ($\alpha = 0.01$). If the null hypothesis of exponential distribution is not rejected, we implement the t-test to verify whether the two distributions have different parameters (i.e. different b-values). More specifically, we use the bootstrap t-test (Amorese *et al.* 2010), which is based on the t-test for samples with unequal variances (Welch 1947), the only difference being that the standard errors are estimated by bootstrap. Following Amorese *et al.* (2010), the statistic is:

$$t = \frac{b_1 - b_2}{\sqrt{SE_1^2 + SE_2^2}}, \tag{4}$$

where b_1 and b_2 are the b-values being compared, SE_1 and SE_2 are the b-values standard errors estimated by bootstrap. If $t \geq 2.58$, the null hypothesis of equal b-values is rejected at a significance level of 0.01.

4.2 Q2 – Exploring the space–time–magnitude patterns: the *K*-NN analysis

We then proceed by analysing some statistical characteristics of the identified K-NN to inspect whether their space–time–magnitude distribution depends on the magnitude of the triggered earthquake they precede. We implement a statistical approach based on a stacking method, which allows us to combine information from many different sequences. For each K-NN property, we group its values by the triggered earthquake magnitude (sorted in an ascending order); this results in one distribution of the property values of K-NN earthquakes for each magnitude of the triggered events. Then, we summarize each distribution with the following statistics:

(1) the median of the K-NN magnitudes (median of the maximum K-NN magnitudes as stability test);
(2) the median of the K-NN elapsed times from the triggered earthquake;
(3) the median of the K-NN unnormalized and normalized spatial distances from the triggered earthquake;
(4) the interquartile (75–25 percentile) range of the K-NN unnormalized and normalized spatial distances from the triggered earthquake.

For cases (1)–(3), we consider the median of such distributions because it is the most robust estimator of the central value for unknown parent distributions markedly asymmetrical, and with a small dataset.

For case (4), we consider the interquartile range as a robust estimator of the K-NN spatial dispersion: in fact, a lower interquartile range implies a lower variability in the distances (the K-NN illuminate a narrower area) and vice versa. In this sense, it could be thought of as a proxy for the K-NN area. Note that this is different from case (3), where we look at the median distance between the triggered earthquake and its K-NN, no matter how the latter are dispersed. As regards the spatial distances, we account for them as they are (unnormalized), and we also normalize them to the dimension of the fault of the largest K-NN for each triggered earthquake. With this normalization we aim to remove the spatial signal due to the occurrence of triggering earthquakes of different magnitude, that is, we avoid to have a large spatial distribution of the K-NN simply because one of them has a large magnitude. In this way, we focus only on the possibility that the spatial distribution of the K-NN, independently from their magnitude, encodes some information about the magnitude of the triggered earthquake. Let us consider a specific triggered earthquake j, for which we have identified K nearest-neighbours. The normalized distance of the triggered earthquake with respect to its ith nearest-neighbour ($i \in [1, K]$) is calculated as:

$$\mathrm{ND}(i, j) = \frac{\mathrm{dist}(i, j)}{10^{a + b * log M_o^{i*}}}, \tag{5}$$

where $\mathrm{dist}(i, j)$ is the distance, in km, between the triggered earthquake and the i-th nearest-neighbour, and the index $i*$ indicates the largest nearest-neighbour with seismic moment M_o^{i*} . The latter has been derived by rearranging eq. (7) in Hanks & Kanamori (1979), which relates moment magnitude to seismic moment [N m]:

$$M_o = 10^{\frac{3}{2}(M_w + 10.7)} \times 10^{-7} \tag{6}$$

The expression at the denominator of eq. (5) is the relation proposed by Mai & Beroza (2000) which provides an estimation of the

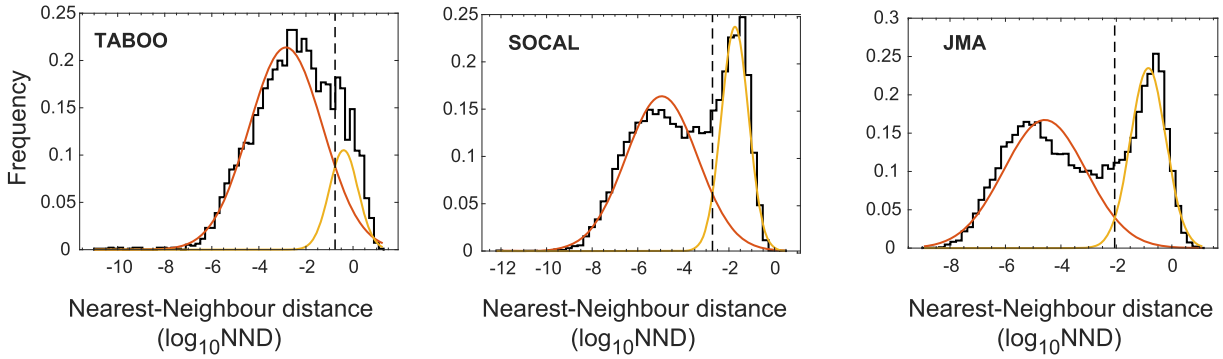

Figure 2. Histogram of the nearest-neighbour distance (NND) values with superimposed the two-component Gaussian mixture model and the estimated threshold value (dashed line). The latter separates the triggered earthquakes (NND values lower than the threshold) from the background earthquakes (NND values higher than the threshold). As regards the TABOO catalogue, the bimodality is less marked compared to the other two data sets, for the background events are poorly represented. This is expected, due to the short time window covered by this catalogue. Nevertheless, a two-component Gaussian mixture model returns a better fit than a single-component Gaussian mixture model: in fact, the Akaike information criterion (AIC) is 1.8044×10^4 and 1.8236×10^4, respectively.

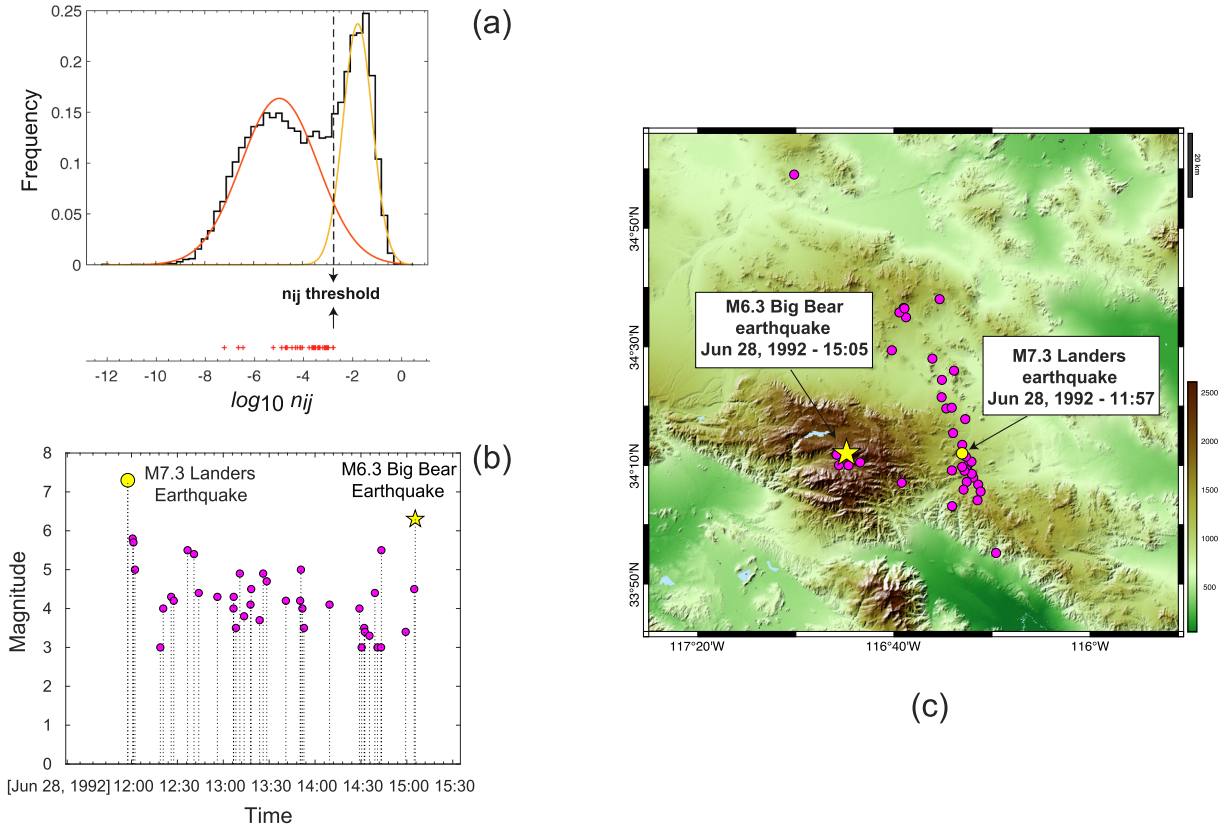

Figure 3. An example of the application of the *K*-NN method. The figure shows the space–time domain of the *K*-NN identified for the *M* 6.3 Big Bear earthquake. (a) The detection of the *K*-NN relies on the same threshold-based technique rooted in the NN method: they are all strongly correlated—their distance fall below the threshold in the NND distribution—to the triggered earthquake they precede (each red cross represents the distance between Big Bear earthquake and one of its *K*-NN). (b) Magnitude and temporal domain of the *K*-NN: the first, in order of time, is the *M*7.3 Landers earthquake. (c) *K*-NN spatial domain. .

Table 1. Proportion of highly connected events out of triggered events.

Catalogue	Highly connected events (per cent)
TABOO	89
SOCAL	80
JMA	82
ETAS	89

Percentage of triggered earthquakes strongly correlated to more than one ancestor.

Table 2. Statistical comparison of the distributions of background and triggered earthquake magnitudes.

			Triggered			Background		
Catalogue	*p*-value KS2 Test	Bootstrap *t*	*N*	*p*-value Lilliefors Test	*b*-value	*N*	*p*-value Lilliefors Test	*b*-value
TABOO	**9.1561e-06**	**5.7104**	3772	0.1120	0.9213	889	0.1100	1.1496
SOCAL	**1.4144e-10**	**6.8615**	8917	0.2312	0.9752	5555	0.3942	1.1043
JMA	**5.3010e-09**	**6.4581**	10984	0.5271	0.8091	7020	0.0360	0.8929
ETAS	0.3310	0.5719	5291	0.3152	1.0034	2664	0.1113	1.0171

For each catalogue, the table displays: the *p*-value of the two-sample Kolmogorov–Smirnov test, the bootstrap *t* statistics, the number of data in the two distributions (*N*), the *p*-value of the Lilliefors test, and the estimated *b*-value. Bold values for the *p*-value of the two-sample Kolmogorov–Smirnov test indicate a significant difference between the magnitude distributions of triggered and background events. Bold values for the bootstrap *t* statistics are indicative of a statistically significant difference between the two *b*-values.

Table 3. Statistical comparison of the distributions of background and triggered earthquake magnitudes after STAI removal.

			Triggered			Background		
			Helmstetter method					
Catalogue	*p*-value KS2 Test	Bootstrap *t*	*N*	*p*-value Lilliefors Test	*b*-value	*N*	*p*-value Lilliefors Test	*b*-value
SOCAL	**3.1974e-06**	**4.8087**	8218	0.1416	1.0128	5555	0.3399	1.1043
JMA	**6.6144e-04**	**3.8094**	8883	0.4633	0.8393	7020	0.0270	0.8929
			t > 1 day					
Catalogue	*p*-value KS2 Test	Bootstrap *t*	*N*	*p*-value Lilliefors Test	*b*-value	*N*	*p*-value Lilliefors Test	*b*-value
SOCAL	**6.2993e-05**	**4.4870**	7360	0.2822	1.0188	5555	0.3697	1.1043
JMA	**0.0016**	**3.4410**	8766	0.8218	0.8444	7019	0.0410	0.8930
			r > Fault Length & t > 1 week					
Catalogue	*p*-value KS2 Test	Bootstrap *t*	*N*	*p*-value Lilliefors Test	*b*-value	*N*	*p*-value Lilliefors Test	*b*-value
SOCAL	**1.1162e-05**	**4.7494**	6182	0.2117	1.0108	5555	0.2081	1.1043
JMA	**0.0017**	**3.3836**	7727	0.3796	0.8441	7020	0.0250	0.8929
			r > Fault Length & t > 3 months					
Catalogue	*p*-value KS2 Test	Bootstrap *t*	*N*	*p*-value Lilliefors Test	*b*-value	*N*	*p*-value Lilliefors Test	*b*-value
SOCAL	**4.6191e-06**	**4.8139**	4790	0.3041	1.0018	5554	0.2495	1.1043
JMA	**0.0018**	**3.6459**	6463	0.7545	0.8375	7016	0.0490	0.8928

For each catalogue, the table displays: the *p*-value of the two-sample Kolmogorov–Smirnov test, the bootstrap *t* statistics, the number of data in the two distributions (*N*), the *p*-value of the Lilliefors test, and the estimated *b*-value. The methods used for removing the STAI issue are described in Appendix B. Bold values have the same meaning as in Table 2.

Table 4. Statistical comparison of the distributions of background and triggered earthquake magnitudes after having removed multiple sources of magnitude incompleteness.

			Triggered			Background		
Catalogue	*p*-value KS2 Test	Bootstrap *t*	*N*	*p*-value Lilliefors Test	*b*-value	*N*	*p*-value Lilliefors Test	*b*-value
SOCAL (Region 1)	0.5417	2.4294	2301	0.0432	0.9741	1014	0.4512	1.0704
JMA (Year ≥2000)	**8.9018e-07**	**5.3895**	10984	0.4068	0.8163	7020	0.2519	0.9159
			Triggered			Background		
			Helmstetter method					
Catalogue	*p*-value KS2 Test	Bootstrap *t*	*N*	*p*-value Lilliefors Test	*b*-value	*N*	*p*-value Lilliefors Test	*b*-value
SOCAL (Region 1)	0.5768	2.4518	2295	0.0480	0.9729	991	0.6366	1.0678
JMA (Year ≥2000)	0.0143	**2.8748**	5067	0.8723	0.8596	3716	0.1090	0.9159
			t > 1 day					
Catalogue	*p*-value KS2 Test	Bootstrap *t*	*N*	*p*-value Lilliefors Test	*b*-value	*N*	*p*-value Lilliefors Test	*b*-value
SOCAL (Region 1)	0.5386	2.4031	2292	0.0281	0.9722	966	0.5792	1.0667
JMA (Year ≥2000)	0.0286	2.4995	5023	0.5524	0.8668	3716	0.2276	0.9159
			r > Fault Length & t > 1 week					
Catalogue	*p*-value KS2 Test	Bootstrap *t*	*N*	*p*-value Lilliefors Test	*b*-value	*N*	*p*-value Lilliefors Test	*b*-value
SOCAL (Region 1)	0.9996	1.1002	1344	0.4615	1.0229	1014	0.6947	1.0704
JMA (Year ≥2000)	0.0477	2.1630	4265	0.7991	0.8710	3716	0.2548	0.9159
			r > Fault Length & t > 3 months					
Catalogue	*p*-value KS2 Test	Bootstrap *t*	*N*	*p*-value Lilliefors Test	*b*-value	*N*	*p*-value Lilliefors Test	*b*-value
SOCAL (Region 1)	0.9989	1.1284	907	0.9430	1.0167	1014	0.7505	1.0704
JMA (Year ≥2000)	0.1185	2.1627	3345	0.2831	0.8698	3715	0.1909	0.9158

For each catalogue, the table displays: the *P*-value of the two-sample Kolmogorov–Smirnov test, the bootstrap *t* statistics, the number of data in the two distributions (*N*), the *p*-value of the Lilliefors test, and the estimated *b*-value. The methods used for removing the STAI issue are described in Appendix B. The spatial variability of magnitude incompleteness in SOCAL is handled by selecting events within Region 1, while in the case of the JMA catalogue the temporal variation of the record quality is handled by considering only the events occurred after 2000 (see Section 5.2). Bold values have the same meaning as in Table 2.

Table 5. Statistical comparison of the distributions of background and triggered earthquake magnitudes with higher M_c.

		Triggered			Background		
Catalogue	*p*-value KS2	*N*	*p*-value Lilliefors	*b*-value	*N*	*p*-value Lilliefors	*b*-value
TABOO	0.0333	1608	0.2824	0.9183	287	0.4401	1.1622
SOCAL	0.1390	2752	0.6735	0.9843	1692	0.5532	1.0600
JMA	0.3625	4279	0.3414	0.8135	2521	0.3297	0.8501

For each catalogue, the table displays: the *p*-value of the two-sample Kolmogorov–Smirnov test, the number of data in the two distributions (N), the *p*-value of the Lilliefors test, and the estimated *b*-value. Here we have imposed a minimum threshold magnitude which is higher than the magnitude of completeness estimated for the whole catalogue (see Section 2).

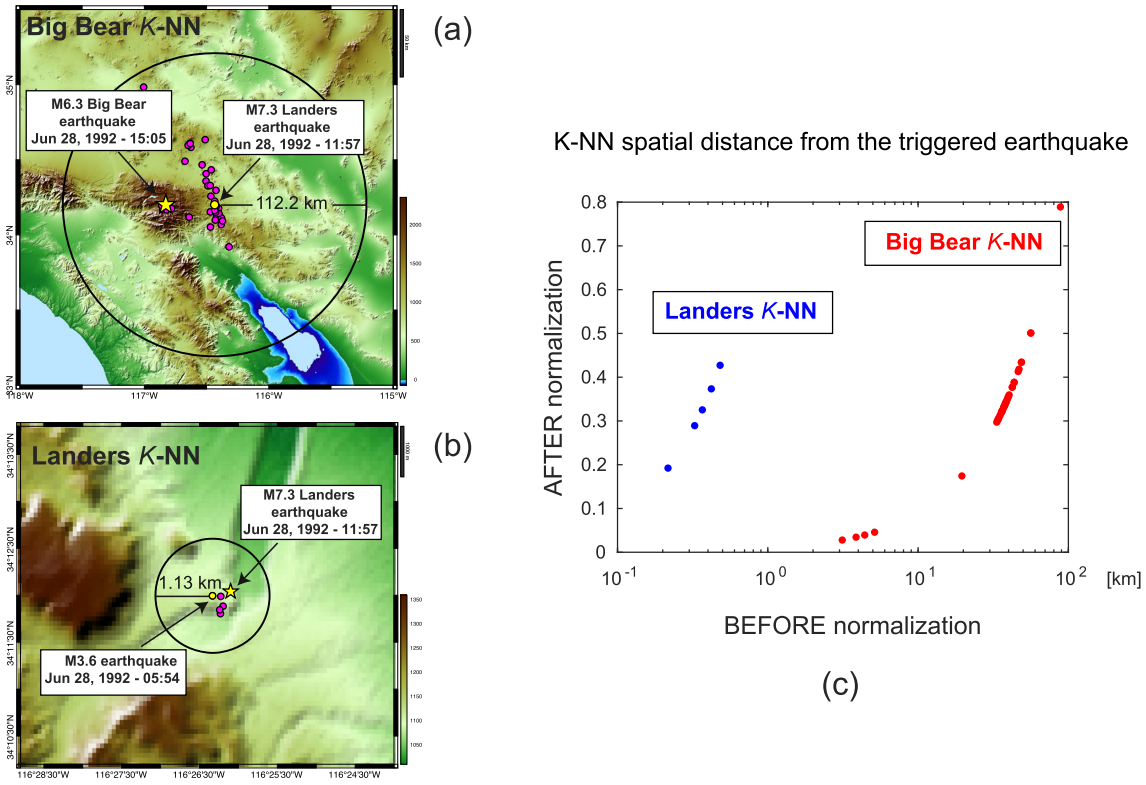

Figure 4. Effect of the normalization applied to the distances between the *K*-NN and their triggered event. (a) Same as Fig. 3. Now, the area of influence of the largest among the *K*-NN (M7.3 Landers earthquake) is displayed: it corresponds to a circle of radius equal to the Landers fault rupture length [estimated with eqs (5) and (6)]. All the distances between Big Bear earthquake and its *K*-NN are normalized by this value. (b) *K*-NN of the *M*7.3 Landers earthquake. Now, the largest precursor has a magnitude equals to 3.6, hence its fault rupture length is much lower. (c) Effect of the normalization for cases depicted in (a) and (b). The occurrence of a large precursor (Landers event) strongly affects the distance values after the normalization: most of the normalized distances for case (a) are now mapped on the same range of case (b). By removing the footprint of large *K*-NN, we make the distances comparable. .

subsurface rupture length through source-scaling laws based on finite-fault source inversion. The coefficients a and b, the least-square fits intercept and slope respectively, change depending on the fault type (see table 2 in Mai & Beroza 2000). In this work, we use $a = -5.5$ and $b = 0.36$ for the SOCAL catalogue (dominant fault type: strike-slip), $a = -5.71$ and $b = 0.38$ for TABOO and JMA catalogues (dominant fault type: dip-slip) and $a = -5.20$ and $b = 0.35$ for ETAS catalogue (all fault types). In order to test the stability of our findings, we also consider a different scaling law linking fault rupture length to earthquake size (Wells & Coppersmith 1994). According to the magnitude-independence assumption, the medians and the interquartile range considered in the four cases should not depend on the magnitude of the triggered earthquakes. For each case, we verify this assumption using the Kendall-tau test (Kendall 1948), a non-parametric test that verifies if the quantity of interest is somehow correlated with the magnitude of the triggered earthquake (Appendix C). In order to prevent any bias caused by the lack of data, we calculate the correlation coefficient and the relative *p*-value only for those distributions with at least 10 data (the minimum sample size required for a correct estimation of the median of a distribution). We reject the null hypothesis of no-correlation if the calculated *p*-value is lower than $\alpha = 0.01$, which represents the significance level adopted. When dealing with *K*-NN spatial distances, we opt for the one-tail test, as we are only interested into a positive correlation with respect to the triggered event magnitude (a negative one would be not supported by physical evidence). We also remark again that the stability of the results across the seismic catalogues is also important here, because it minimizes the problem of finding 'significant' results simply because we are carrying out many tests.

Table 6. *K*-NN analysis.

	Kendall's tau	Catalogues			
		TABOO	SOCAL	JMA	ETAS
K-NN magnitude	τ	0.2523	0.3577	0.4662	–0.1859
	p-value	0.0624	0.0112	**8.8895e-05**	0.1904
K-NN elapsed times from the triggered earthquake	τ	0.1011	–0.2903	–0.2065	0.0689
	p-value	0.4375	0.0218	0.0657	0.5727
K-NN unnormalized spatial distances from the triggered earthquake	τ	–0.1828	0.3978	0.3954	–0.2336
	p-value	0.9276	**6.9777e-04**	**1.4867e-04**	0.9769
K-NN normalized spatial distances from the triggered earthquake	τ	–0.3161	0.4710	0.3954	–0.1294
	p-value	0.9945	**6.1966e-05**	**1.4867e-04**	0.8649
K-NN spatial dispersion (unnormalized distances)	τ	–0.2817	0.0323	–0.0202	–0.2538
	p-value	0.9880	0.4067	0.5762	0.9849
K-NN spatial dispersion (normalized distances)	τ	–0.4065	0.1398	–0.4359	–0.1496
	P-value	0.9995	0.1396	1	0.8986

For each catalogue and for each *K*-NN property, the table displays the correlation coefficient and the *p*-value returned by the Kendall-tau test.

5 RESULTS AND DISCUSSION

5.1 Separating background and triggered seismicity

In the first part of our analysis, we implement the NN Clustering Analysis within each catalogue. As input parameters for eq. (2), we use an average value for the epicentral fractal dimension ($d_f = 1.2$, as found in Kagan 1991) and the maximum-likelihood estimate of the *b*-value for each catalogue. We verify the stability of the method by running our analysis with different values for the input parameters: *b*-value = 1 and $d_f = 1.6$, as estimated for Southern California (Corral 2003). The stability test findings are available in the Supporting Information. Fig. 2 shows the results of the NN procedure applied to the different seismic catalogues. For all the three cases, the distribution of the nearest-neighbour distance values is apparently bimodal, with two local maxima. By fitting a two-component Gaussian mixture model, we have estimated the threshold value which separates the two modes (intersection point of the two Gaussian components). The mode corresponding to values of the nearest-neighbour distance lower than the identified threshold is identified as the clustered portion of seismicity; the other one corresponds to the background earthquakes (events being unusually far from their nearest-neighbour).

As a next step, we implement our *K*-NN modification: for each triggered earthquake, we identify all its K ($K \geq 1$) strongly correlated ancestors, whose distance is lower than the identified threshold. As an example, Fig. 3 depicts the space–time domain of *K*-NN identified for the *M*6.3 Big Bear earthquake. In Table 1 we show the percentage of highly connected events out of triggered events for all the three catalogues. It is clear that a large part of triggered events is strongly correlated to more than one ancestor, which motivates us to implement the *K*-NN analysis in Section 5.3. As regards the skill of the declustering method, we apply the NN method to the synthetic ETAS catalogue for which we know whether an earthquake belongs to the background or it is triggered. For the synthetic ETAS catalogue used in this paper, the NN method identifies correctly the 97.7 per cent of the earthquakes. As regards the other synthetic catalogues that are used to check the stability of the results, this proportion ranges from 96.8 to 98.6 per cent. This result further proves that the NN procedure is a valuable method for separating triggered and background earthquakes in a catalogue.

5.2 Magnitude distribution of triggered and background earthquakes

Table 2 shows the *p*-values derived from the KS2 test, which verifies if the magnitude distributions of triggered and background earthquakes are the same. Additionally, it displays the *p*-values of the Lilliefors test for testing the hypothesis of exponentially distributed data and the *t*-statistics for the bootstrap *t*-test. We find a statistically significant difference between the distributions of triggered and background magnitudes (with a significance level $\alpha = 0.01$) for all the three seismic catalogues, but not for the ETAS catalogue. All background and triggered catalogues are overall exponentially distributed, and the difference is only due to different *b*-values, as suggested by the bootstrap *t*-test (*t* is always $\geq$2.58). Specifically, the background events are characterized by higher *b*-values, in agreement with Zhuang *et al.* (2004); if we interpret the *b*-value as a stress indicator (Scholz 1968), this would imply that the stress is higher during seismic sequences than for background earthquakes.

To verify whether the observed discrepancy is attributable to physical processes or to some kind of magnitude incompleteness, we handle the latter through a comprehensive analysis. When analysing SOCAL and JMA catalogues, STAI should properly taken into account, as they both are characterized by a maximum magnitude of 7.3. To do that, we implement all the methods described in Appendix B. Results are shown in Table 3: though the *p*-value of KS2 test has strongly increased, the difference between the magnitude distributions of triggered and background earthquakes remains significant. This implies that the STAI after a large earthquake may affect the results but it cannot justify entirely the difference between the two distributions. We now investigate on other types of magnitude incompleteness, analysing the impact of the spatial variability of the network density in SOCAL, and the time variability of the quality of JMA. In this respect, we apply our analysis to the following catalogue subsets:

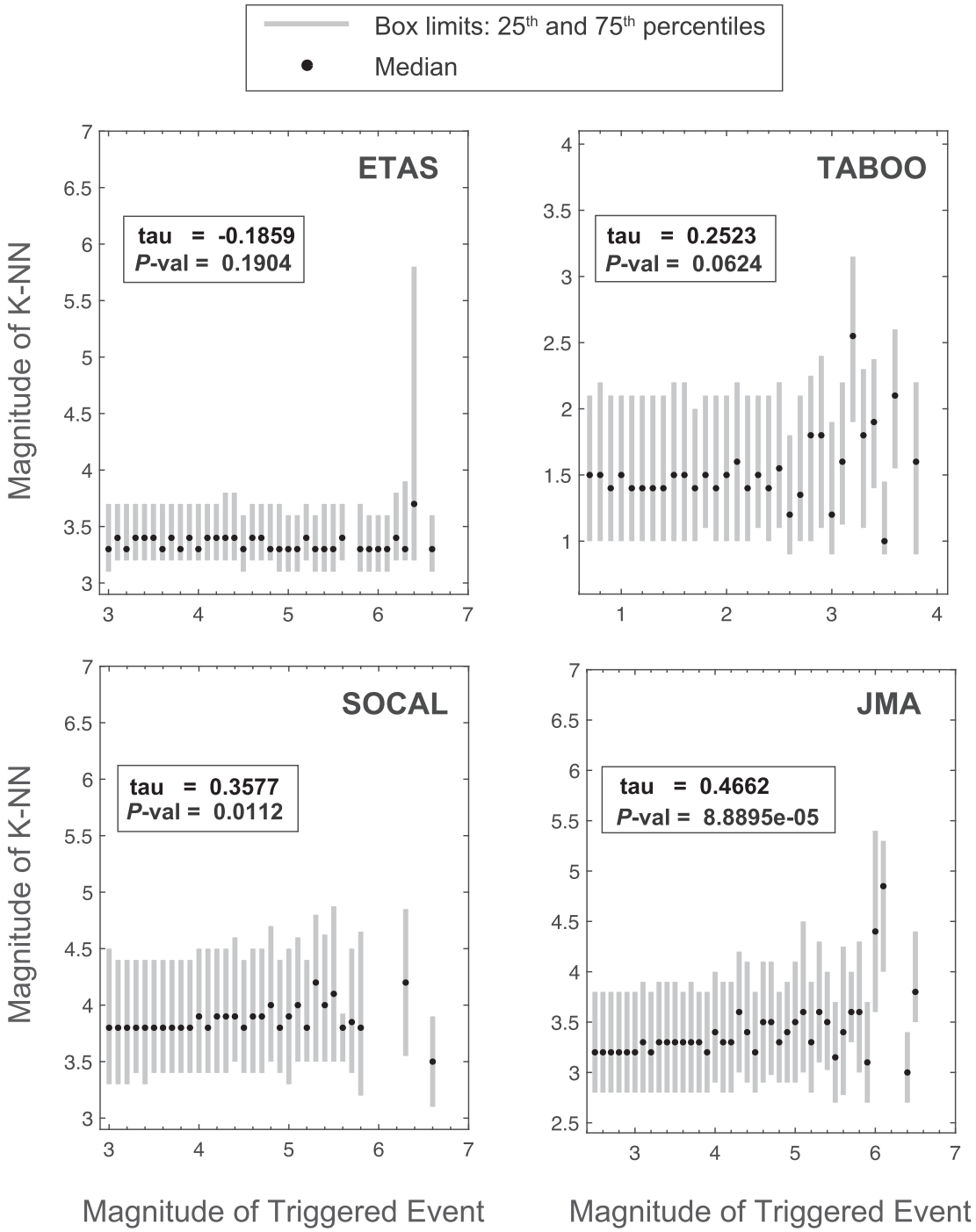

Figure 5. *K*-NN magnitudes analysis for the ETAS and the instrumental catalogues. For each unique value of the triggered event magnitude, the corresponding distribution of *K*-NN magnitude values is summarized by its 25th and 75th percentiles (lower and upper edge of the grey box, respectively). The black dots represent the median of the distributions. The correlation coefficient and the *p*-values returned by the Kendall-tau test are also shown. A *p*-value $\leq$0.01 indicates a significant correlation between the median of the distributions and the triggered event magnitude. .

Table 7. *K*-NN magnitude analysis after STAI removal.

Catalogues	Kendall's tau	Methods of STAI removal			
		Helmstetter method	$t > 1$ day	$r >$ Fault length and $t > 1$ week	$r >$ Fault length and $t > 3$ months
SOCAL	τ	0.1698	0.0591	–0.1000	0.1118
	p-value	0.2532	0.7084	0.5061	0.4747
JMA	τ	0.3211	0.3034	0.3362	0.1182
	p-value	**0.0098**	0.0151	**0.0094**	0.3837

For SOCAL and JMA catalogues, the table displays the correlation coefficient and the *P*-value returned by the Kendall-tau test for the analysis of the *K*-NN magnitudes.

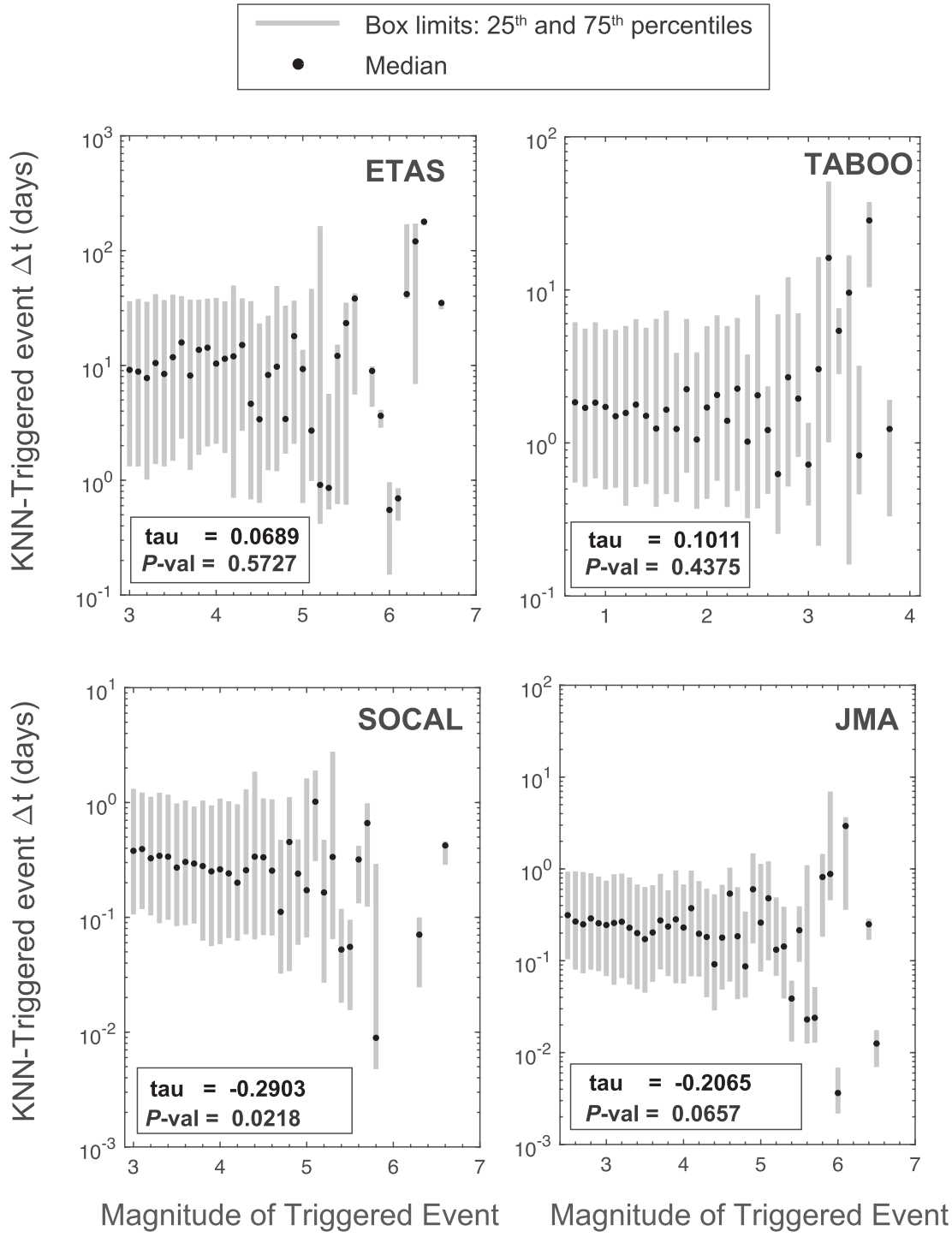

Figure 6. Analysis of the *K*-NN elapsed times from the triggered earthquake for the ETAS and the instrumental catalogues. For each unique value of the triggered event magnitude, the corresponding distribution of *K*-NN Δt is summarized by its 25th and 75th percentiles (lower and upper edge of the grey box, respectively). The black dots represent the median of the distributions. The correlation coefficient and the *p*-values returned by the Kendall-tau test are also shown. A *p*-value $\leq$0.01 indicates a significant correlation between the median of the distributions and the triggered event magnitude. .

(1) [SOCAL] region with a higher density of seismic stations (we select Region 1 defined in Lippiello *et al.* 2012a);

(2) [JMA] subcatalog starting from 2000 January 1 (starting year of the installation of modern seismic networks in Japan (Nanjo *et al.* 2010).

Results are summarized in Table 4. As it can be noticed, by properly taking into account all the possible sources of magnitude incompleteness that apply to a specific catalogue, the difference between the distributions of triggered and background magnitudes is not significant anymore. As suggested by the bootstrap *t*-test, the two *b*-values are now comparable. We reach the same conclusion by using $M_c^\star$

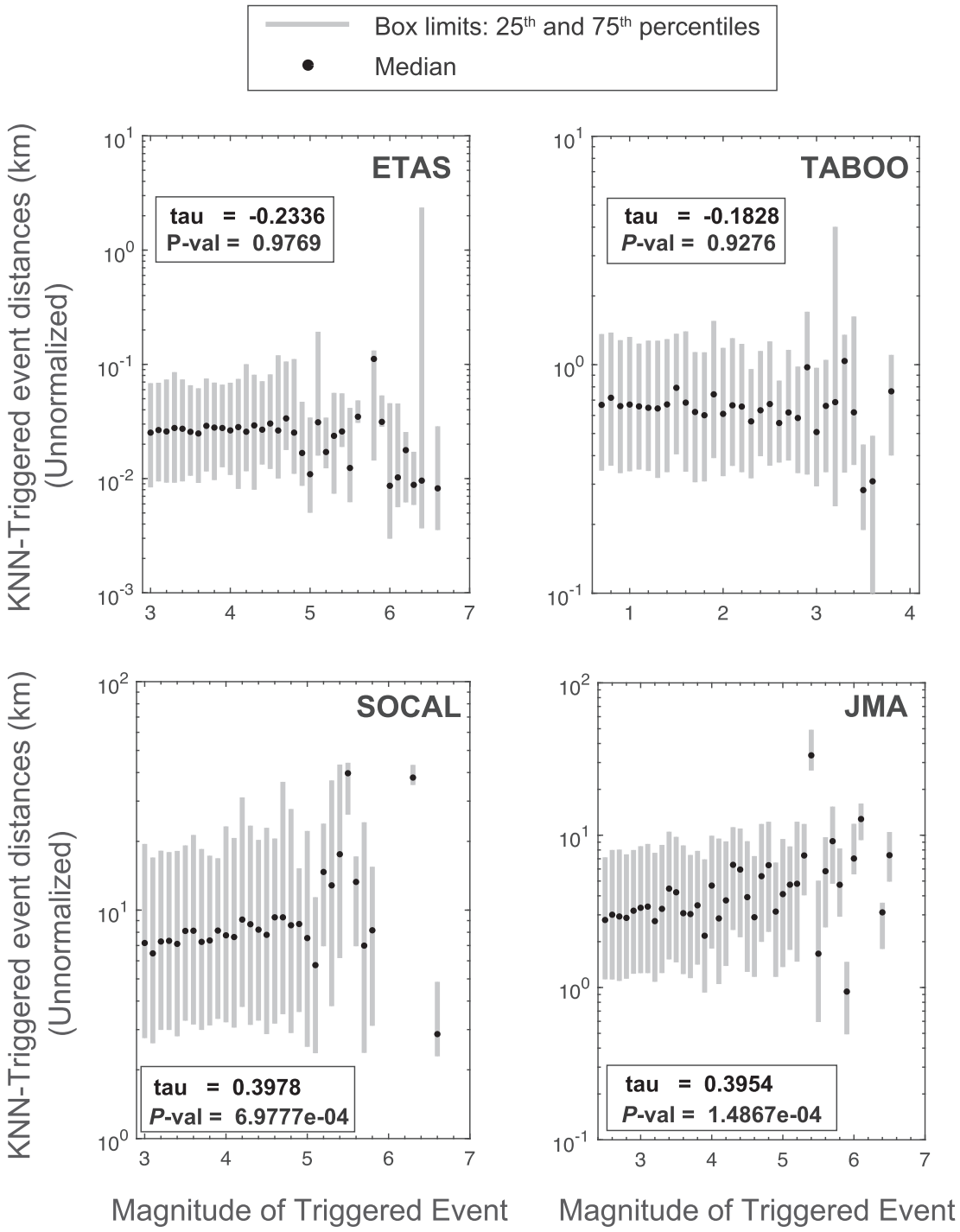

Figure 7. Analysis of the *K*-NN unnormalized spatial distances from the triggered earthquake for the ETAS and the instrumental catalogues. For each single value of the triggered event magnitude, the corresponding distribution of distances is summarized by its 25th and 75th percentiles (lower and upper edge of the grey box, respectively). The black dots represent the median of the distributions. The correlation coefficient and the *p*-values returned by the Kendall-tau test are also shown. A *p*-value $\leq$0.01 indicates a significant correlation between the median of the distributions and the triggered event magnitude. .

instead of M_c (see Section 2). Results are shown in Table 5. Although some authors advocate different physical motivations to explain this difference (Zhuang *et al.* 2004; Shearer 2012; Lippiello *et al.* 2017), the results of this paper show that a discrepancy from the magnitude-independence assumption may be due to different sources of magnitude incompleteness, which cannot be detected by looking at the catalogue as a whole. A significant difference between magnitude distribution of triggered and background earthquakes persists only for the TABOO catalogue, even when addressing the STAI issue (we have removed all the earthquakes occurred 1 and 7 d after $M \geq 3$ events). This result may indicate that the spatial scale is important to evaluate the reliability of the magnitude-independence assumption. Specifically, in the area

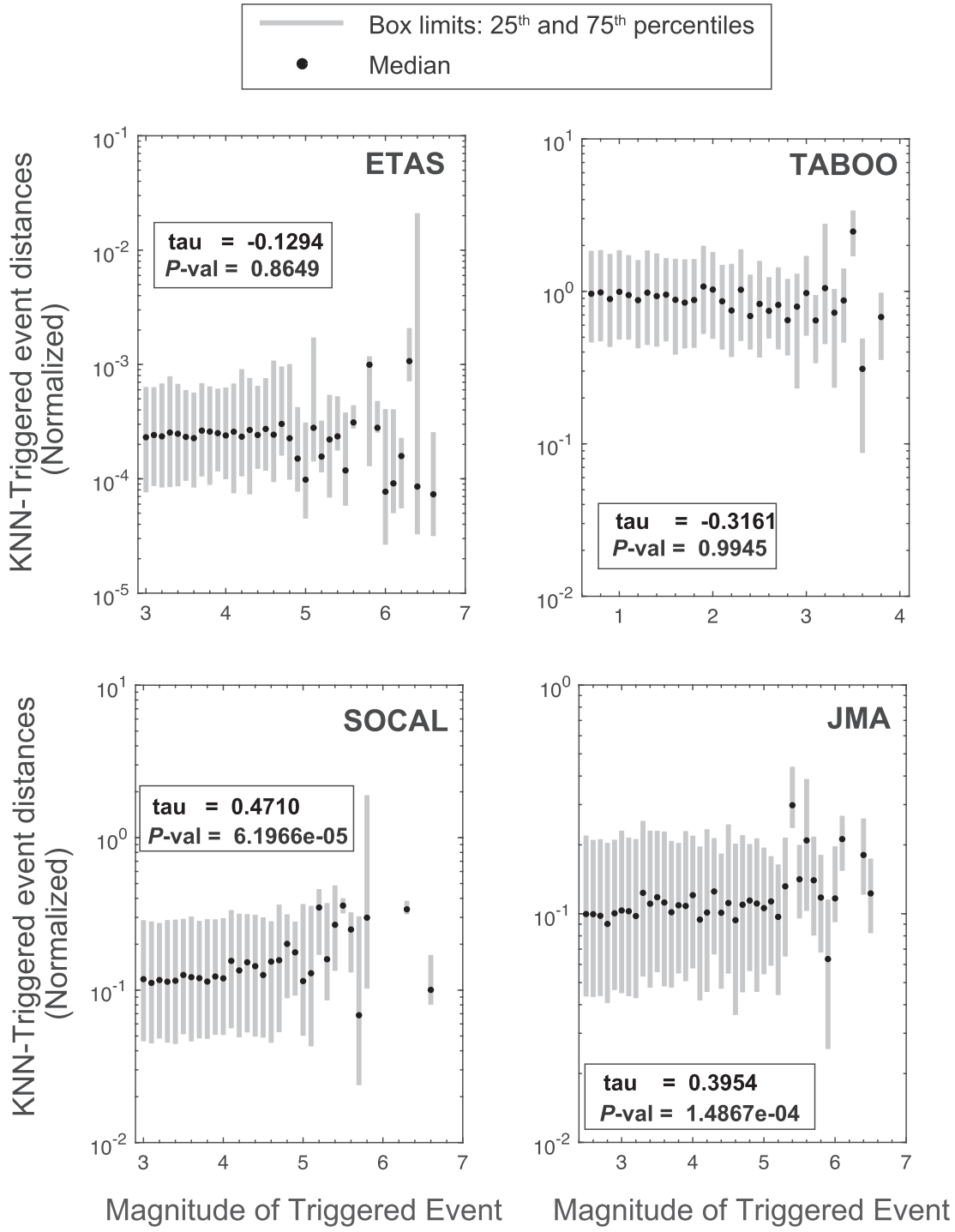

Figure 8. Analysis of the *K*-NN normalized spatial distances from the triggered earthquake for the ETAS and the instrumental catalogues. For each unique value of the triggered event magnitude, the corresponding distribution of distances is summarized by its 25th and 75th percentiles (lower and upper edge of the grey box, respectively). The black dots represent the median of the distributions. The correlation coefficient and the *p*-values returned by the Kendall-tau test are also shown. A *p*-value $\leq$0.01 indicates a significant correlation between the median of the distributions and the triggered event magnitude. .

covered by the TABOO catalogue, the difference between the magnitude distributions of triggered and background earthquakes may stem from the existence of two different modes to release seismic energy, as proposed by Chiaraluce *et al.* (2007).

5.3 Analysis of the seismicity that precedes triggered earthquakes

For each *K*-NN feature, that is, each feature of the seismicity that precedes triggered earthquakes, we group its values by the triggered event size, a stacking process which results in as many distributions as the single values of triggered events magnitudes. Then, we summarize each

distribution by its median (or by its interquartile range in the case of the spatial dispersion) and we inspect whether it is significantly correlated to the triggered event magnitude. This is done by implementing the Kendall-tau test, whose results (the correlation coefficient *tau* and the estimated *p*-value) are reported in Table 6. As regards the normalization of the spatial distances, Fig. 4 shows an example related to Landers and Big Bear earthquakes.

5.3.1 K-NN magnitudes

For the JMA catalogue, the Kendall-tau test returns a significant correlation ($\alpha = 0.01$) between the median of the *K*-NN magnitudes and the magnitude of the triggered earthquake, while for the SOCAL catalogue such a positive correlation has a *p*-value of 0.012; conversely, TABOO does not show any significant correlation. Fig. 5 shows the plot of the *K*-NN magnitude as a function of the magnitude of the triggered earthquake for the ETAS catalogue and the three seismic catalogues. To investigate whether such a correlation is a real physical signal or it is due to short-term incompleteness after a large earthquake, we apply the STAI procedures described previously and we rerun the Kendall-tau test. After removing the STAI periods, we notice that the correlation between the median of the *K*-NN magnitudes and the magnitude of the triggered earthquake mostly vanishes (see Table 7). As an additional test, we also consider the median of the maximum *K*-NN magnitudes, which may be a better indicator of the energy released by the seismic sequence and which is barely affected by the short-term incompleteness after large earthquakes; in this case, we do not find any significant correlation (see Table S2 in the Supporting Information). This further corroborates the hypothesis that the magnitude distribution of the seismicity that precedes a triggered earthquake does not have a prognostic value for the magnitude of the latter. Although this result appears in contradiction with findings by Lippiello *et al.* (2012a), Spassiani & Sebastiani (2016) and Corral (2006), which found a pairwise correlation between the magnitude of consecutive earthquakes, we stress again that in this work we are investigating a different time interval and this may justify the differences. Specifically, the pairwise correlation found by Lippiello *et al.* (2012a) is more evident for time intervals of few tens of minutes (see fig. 3 in Lippiello *et al.* 2012a), whereas, considering the whole seismic sequence which precedes an earthquake, we are investigating on a time interval of the order of hours or more (Fig. 6).

5.3.2 K-NN elapsed times from the triggered earthquake

We do not find any significant correlation in this case, as expected by the magnitude-independence assumption. Fig. 6 shows the plot of the *K*-NN time intervals from the triggered earthquake as a function of the magnitude of the latter, for the ETAS catalogue and the three real seismic catalogues.

5.3.3 K-NN spatial distances from the triggered earthquake

Figs 7 and 8 show the plot of the median of the *K*-NN unnormalized and normalized spatial distances from the triggered earthquake as a function of the magnitude of the latter. The results highlight a positive significant correlation for the *K*-NN unnormalized distances for SOCAL and JMA catalogues. The analysis of the *K*-NN normalized distances confirm this trend for both the catalogues. This correlation, not explained by the magnitude-independence assumption (as showed by the plot of the ETAS catalogue), implies that the largest triggered earthquakes preferentially nucleate farther away from the ongoing seismic sequence. This is in agreement with the findings of van der Elst & Shaw (2015), who show that the largest aftershocks tend to nucleate in the outer regions of the parent aftershock zone.

This empirical result may also provide a possible answer to the scientific question posed by Field *et al.* (2017) in their Fig. 7: what is the likelihood that a large earthquake nucleates from within the fault area that has just ruptured? According to our results, the larger triggered earthquakes tend preferentially to nucleate in the outer region of the seismic sequences, that is, outside the blue area in the fig. 7 of Field *et al.* (2017). Finally, the lack of such a correlation for the TABOO catalogue could be due to the much smaller events (and so smaller dimension of the ruptures that becomes comparable to the location uncertainty) with respect to the other two catalogues.

5.3.4 K-NN spatial dispersion

The results of the test are showed in Figs 9 and 10. For all the catalogues, the Kendall-tau test has returned no significant correlations between the spatial dispersion (interquartile range) of the *K*-NN and the magnitude of the triggered earthquake, using either unnormalized or normalized distances. This is in agreement with the expected behaviour under the magnitude-independence assumption (ETAS). It is worth noting that these results are not in contradiction with those reported in 5.3.3: in fact, in that case we assess how close the triggered earthquake nucleates with respect to the region occupied by the *K*-NN, no matter how this region is extended. Now, we focus our attention on the spatial dispersion of the *K*-NN. Differently from Lippiello *et al.* (2012a) and Lippiello *et al.* (2017), we do not find any evidence that the *K*-NN spatial distribution encodes information on the magnitude of the triggered earthquake. However, we highlight that any comparison

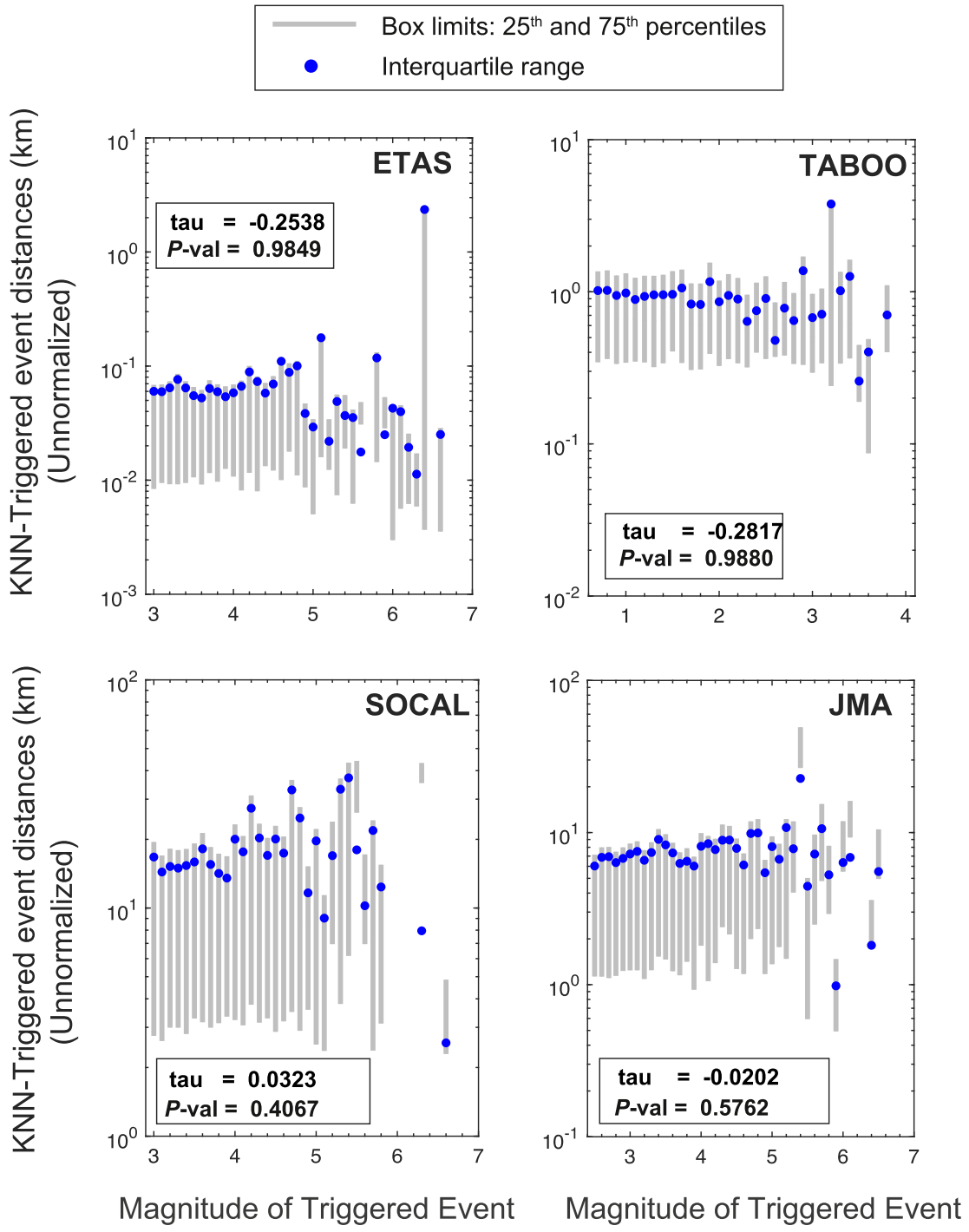

Figure 9. Analysis of the *K*-NN unnormalized spatial distances from the triggered earthquake for the ETAS and the instrumental catalogues. For each unique value of the triggered event magnitude, the corresponding distribution of distances is summarized by its 25th and 75th percentiles (lower and upper edge of the grey box, respectively). The blue dots represent the interquartile range of the distributions. The correlation coefficient and the *p*-values returned by the Kendall-tau test are also shown. A *p*-value ≤ 0.01 indicates a significant correlation between the interquartile range of the distributions and the triggered event magnitude. .

of these results should take into account the different spatial windows considered: in Lippiello *et al.* (2012a) and Lippiello *et al.* (2017), the differences with respect to the magnitude-independence assumption are mostly noticeable for areas with less than 1 km of radius (see fig. 5 in Lippiello *et al.* 2017), whereas in this work we are exploring much wider areas with interquartile differences of about 10 km (see Fig. 9).

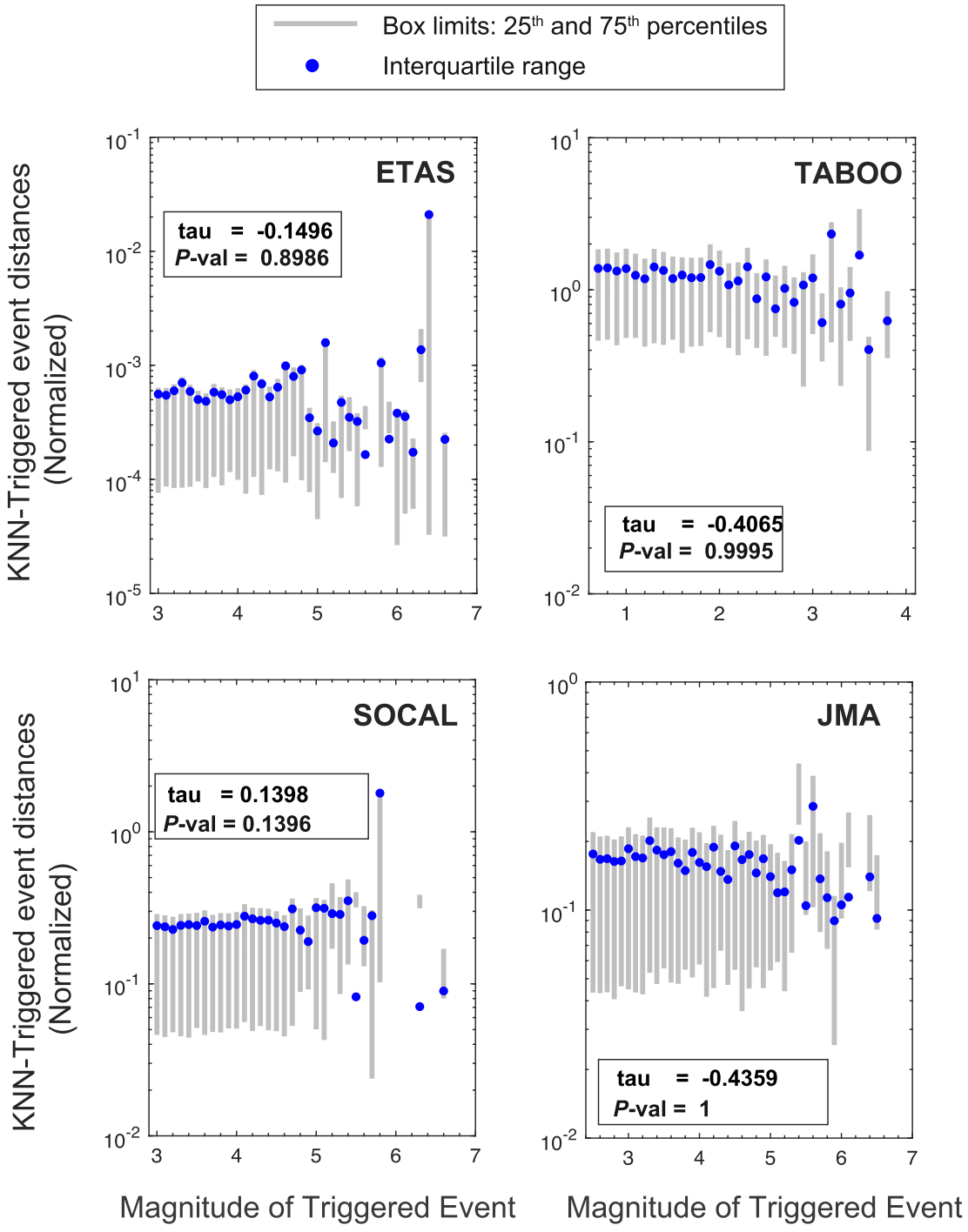

Figure 10. Analysis of the *K*-NN normalized spatial distances from the triggered earthquake for the ETAS and the instrumental catalogues. For each unique value of the triggered event magnitude, the corresponding distribution of distances is summarized by its 25th and 75th percentiles (lower and upper edge of the grey box, respectively). The blue dots represent the interquartile range of the distributions. The correlation coefficient and the *p*-values returned by the Kendall-tau test are also shown. A *p*-value ≤ 0.01 indicates a significant correlation between the interquartile range of the distributions and the triggered event magnitude. .

6 CONCLUSIONS

We have investigated on the magnitude-independence assumption that stands behind the most common short-term earthquake forecasting models. This assumption states that the preceding phase of large and small earthquakes is not distinguishable. We have used a well-established procedure to separate background and triggered earthquakes. However, as any method relying on a declustering technique, the latter may influence the statistical analysis of instrumental catalogues. Though we have done multiple tests to assure an unbiased description of the observed seismicity, the method we have used brings unavoidably a certain degree of uncertainty. We have minimized this potential source

of bias considering only results that were found in independent seismic catalogues and survived to a careful analysis of the possible technical problems that may have induced them. Overall, the magnitude-independence assumption holds reasonable well across different seismic catalogues and using different observations, with one important departure that may be used to improve earthquake forecasting. In particular, after separating background and triggered earthquakes using the *K*-NN method, we may give an answer to the two questions posed in the Section 4.

(1) Q1: The magnitude distributions of the triggered and background earthquakes are not statistically different, if all possible sources of magnitude incompleteness are properly taken into account. Those consist of short-term incompleteness and spatiotemporal variations in the magnitude of completeness, the latter being attributable to an heterogeneous seismic stations density and to a time-dependent catalogue quality. A significant difference is found only for the TABOO catalogue, indicating that the spatial scale may play a role; specifically, in the TABOO area this difference may be due to two different modes to release seismic energy.

(2) Q2: The statistical analysis of the *K*-NN earthquake properties in the space–time–magnitude domain has shown an overall lack of evidence against the magnitude-independence assumption, with one remarkable exception. In particular, we have found a statistically significant correlation of both the unnormalized and normalized distances between the triggered earthquake and the preceding *K*-NN earthquakes, which is not implied in the magnitude-independence assumption. Such a correlation may be explained by a preference of larger events to nucleate at a higher distance from the ongoing sequence. Noteworthy, this feature may be implemented in operational earthquake forecasting models, leading to a better spatial definition of the probability of occurrence of the largest earthquakes during a seismic sequence.

As a final remark, we have found that the statistical analysis of a seismic catalogue with an overall M_c may be still plagued by some undesired fictitious signals. In fact, the analysis of the SOCAL and JMA seismic catalogues has returned some spurious departures from the magnitude-independence assumption that have vanished after a careful consideration of the space–time variation of magnitude of completeness. This issue may be of some guidance for future similar analyses.

ACKNOWLEDGEMENTS

We are grateful to the Editor and to the two anonymous reviewers, who contributed to improve significantly the quality and readability of the manuscript. We also wish to thank Lauro Chiaraluce and Eugenio Lippiello for their feedbacks. Angela Stallone was partly supported by the EU project SERA (Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe), and Warner Marzocchi by the Seismic Hazard Center of the Istituto Nazionale di Geofisica e Vulcanologia. The Southern California catalogue is available here. The Japan Meteorological Agency (JMA) unified catalogue is available at the JMA website. Figure 3 and Figure 4 include NASA SRTM data files (available here), which have been downloaded and plotted in MATLAB thanks to the code provided by Francois Beauducel, 2016.

REFERENCES

Amorese, D., Grasso, J.R. & Rydelek, P.A., 2010. On varying *b*-values with depth: results from computer-intensive tests for Southern California, *Geophys. J. Int.,* **180**(1), 347–360.

Baiesi, M. & Paczuski, M., 2004. Scale-free networks of earthquakes and aftershocks, *Phys. Rev. E,* **69**(6), 066106.

Chiaraluce, L., Chiarabba, C., Collettini, C., Piccinini, D. & Cocco, M., 2007. Architecture and mechanics of an active low-angle normal fault: Alto Tiberina fault, northern Apennines, Italy, *J. Geophys. Res.: Solid Earth,* **112**(B10).

Chiaraluce, L., Amato, A., Carannante, S., Castelli, V., Cattaneo, M., Cocco, M. & Marzorati, S., 2014. The Alto Tiberina Near Fault Observatory (northern Apennines, Italy), *Ann. Geophys.,* **57**(3).

Clauset, A., Shalizi, C.R. & Newman, M.E., 2009. Power-law distributions in empirical data, *SIAM Rev.,* **51**(4), 661–703.

Corral, A., 2003. Local distributions and rate fluctuations in a unified scaling law for earthquakes, *Phys. Rev. E,* **68**(3), 035102.

Corral, A., 2006. Dependence of earthquake recurrence times and independence of magnitudes on seismicity history, *Tectonophysics,* **424**(3-4), 177–193.

de Arcangelis, L., Godano, C., Grasso, J.R. & Lippiello, E., 2016. Statistical physics approach to earthquake occurrence and forecasting, *Phys. Rep.,* **628,** 1–91.

Felzer, K.R. & Brodsky, E.E., 2006. Decay of aftershock density with distance indicates triggering by dynamic stress, *Nature,* **441**(7094), 735.

Felzer, K.R., Becker, T.W., Abercrombie, R.E., Ekström, G. & Rice, J.R., 2002. Triggering of the 1999 Mw 7.1 Hector Mine earthquake by aftershocks of the 1992 Mw 7.3 Landers earthquake, *J. Geophys. Res.: Solid Earth,* **107**(B9).

Felzer, K.R., Abercrombie, R.E. & Ekström, G., 2004. A common origin for aftershocks, foreshocks, and multiplets, *Bull. Seism. Soc. Am.,* **94**(1), 88–98.

Field, E.H., Milner, K.R., Hardebeck, J.L., Page, M.T., van der Elst, N., Jordan, T.H. & Werner, M.J., 2017. A spatiotemporal clustering model for the third uniform California Earthquake Rupture Forecast (UCERF3-ETAS): toward an operational earthquake forecast, *Bull. Seism. Soc. Am.,* **107**(3), 1049–1081.

Gerstenberger, M.C., Wiemer, S., Jones, L.M. & Reasenberg, P.A., 2005. Real-time forecasts of tomorrow's earthquakes in California, *Nature,* **435**(7040), 328.

Gutenberg, B. & Richter, C.F., 1944. Frequency of earthquakes in California, *Bull. Seism. Soc. Am.,* **34**(4), 185–188.

Hanks, T.C. & Kanamori, H., 1979. A moment magnitude scale, *J. Geophys. Res.: Solid Earth,* **84**(B5), 2348–2350.

Hauksson, E., Yang, W. & Shearer, P.M., 2012. Waveform relocated earthquake catalogue for southern California (1981 to June 2011), *Bull. Seism. Soc. Am.,* **102**(5), 2239–2244.

Helmstetter, A. & Sornette, D., 2003. Foreshocks explained by cascades of triggered seismicity, *J. Geophys. Res.: Solid Earth,* **108**(B10).

Helmstetter, A., Sornette, D. & Grasso, J.R., 2003. Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws, *J. Geophys. Res.: Solid Earth,* **108**(B1).

Helmstetter, A., Kagan, Y.Y. & Jackson, D.D., 2006. Comparison of short-term and time-independent earthquake forecast models for southern California, *Bull. Seism. Soc. Am.,* **96**(1), 90–106.

Kagan, Y.Y., 1991. Fractal dimension of brittle fracture, *J. Nonlin. Sci.,* **1**(1), 1–16.

Kagan, Y.Y., 2004. Short-term properties of earthquake catalogues and models of earthquake source, *Bull. Seism. Soc. Am.,* **94**(4), 1207–1228.

Kagan, Y.Y. & Knopoff, L., 1987. Statistical short-term earthquake prediction, *Science,* **236**(4808), 1563–1567.
Kalbfleisch, J.G., 2012. *Probability and Statistical Inference,* **Vol. 1: Probability**. Springer Science & Business Media.
Kendall, M.G., 1948. *Rank Correlation Methods,* C. Griffin.
Lilliefors, H.W., 1969. On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown, *J. Am. Stat. Assoc.,* **64**(325), 387–389.
Lippiello, E., de Arcangelis, L. & Godano, C., 2008. Influence of time and space correlations on earthquake magnitude, *Phys. Rev. Lett.,* **100**(3), 038501.
Lippiello, E., Godano, C. & Arcangelis, L., 2012. The earthquake magnitude is influenced by previous seismicity, *Geophys. Res. Lett.,* **39**(5).
Lippiello, E., Marzocchi, W., De Arcangelis, L. & Godano, C., 2012. Spatial organization of foreshocks as a tool to forecast large earthquakes, *Scient. Rep.,* **2,** 846.
Lippiello, E., Giacco, F., Marzocchi, W., Godano, G. & de Arcangelis, L., 2017. Statistical features of foreshocks in instrumental and ETAS catalogues, *Pure Appl.Geophys.,* **174**(4), 1679–1697.
Mai, P.M. & Beroza, G.C., 2000. Source scaling properties from finite-fault-rupture models, *Bull. seism. Soc. Am.,* **90**(3), 604–615.
Marzocchi, W. & Sandri, L., 2003. A review and new insights on the estimation of the b-value and its uncertainty, *Ann. Geophys,* **46**(6), doi:10.4401/ag-3472.
Marzocchi, W., Taroni, M. & Falcone, G., 2017. Earthquake forecasting during the complex Amatrice-Norcia seismic sequence, *Sci. Adv.,* **3**(9), e1701239.
Michael, A.J., 2012. Fundamental questions of earthquake statistics, source behavior, and the estimation of earthquake probabilities from possible foreshocks, *Bull. seism. Soc. Am.,* **102**(6), 2547–2562.
Nanjo, K.Z., Ishibe, T., Tsuruoka, H., Schorlemmer, D., Ishigaki, Y. & Hirata, N., 2010. Analysis of the completeness magnitude and seismic network coverage of Japan, *Bull. seism. Soc. Am.,* **100**(6), 3261–3268.
Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.,* **83**(401), 9–27.
Ogata, Y., 1998. Space-time point-process models for earthquake occurrences, *Ann. Inst. Stat. Math.,* **50**(2), 379–402.
Ogata, Y., Katsura, K., Tsuruoka, H. & Hirata, N., 2018. Exploring magnitude forecasting of the next earthquake, *Seismol. Res. Lett.,* **89,** (4), 1298–1304.
Omori, F., 1894. *On the After-Shocks of Earthquakes,* **Vol. 7**. The University.
Reasenberg, P.A. & Jones, L.M., 1989. Earthquake hazard after a mainshock in California, *Science,* **243**(4895), 1173–1176.
Schoenberg, F.P., 2003. Multidimensional residual analysis of point process models for earthquake occurrences, *J. Am. Stat. Assoc.,* **98**(464), 789–795.
Scholz, C.H., 1968. The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes, *Bull. Seism. Soc. Am.,* **58**(1), 399–415.
Shearer, P.M., 2012. Space-time clustering of seismicity in California and the distance dependence of earthquake triggering, *J. Geophys. Res.: Solid Earth,* **117**(B10).
Smirnov, N.V., 1939. Estimate of deviation between empirical distribution functions in two independent samples, *Bull.Moscow Univ.,* **46**(253) 2(2), 3–16.
Spassiani, I. & Sebastiani, G., 2016. Exploring the relationship between the magnitudes of seismic events, *J. Geophys. Res.: Solid Earth,* **121**(2), 903–916.
Utsu, T., 1957. Magnitude of earthquakes and occurrence of their aftershocks, *Zisin, Ser.,* **2**(10), 35–45.
van der Elst, N.J. & Shaw, B.E., 2015. Larger aftershocks happen farther away: Nonseparability of magnitude and spatial distributions of aftershocks, *Geophys. Res. Lett.,* **42**(14), 5771–5778.
Van Stiphout, T., Zhuang, J. & Marsan, D., 2012. Seismicity declustering, *Commun. Online Resour. Stat. Seism. Anal*., doi:10.5078/corssa-52382934.
Welch, B.L., 1947. The generalization of student's' problem when several different population variances are involved, *Biometrika,* **34**(1/2), 28–35.
Wells, D.L. & Coppersmith, K.J., 1994. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. seism. Soc. Am.,* **84**(4), 974–1002.
Wiemer, S. & Wyss, M., 2000. Minimum magnitude of completeness in earthquake catalogues: examples from Alaska, the western United States, and Japan, *Bull. Seism. Soc. Am.,* **90**(4), 859–869.
Zaliapin, I. & Ben-Zion, Y., 2013. Earthquake clusters in southern California I: identification and stability, *J. Geophys. Res.: Solid Earth,* **118**(6), 2847–2864.
Zaliapin, I., Gabrielov, A., Keilis-Borok, V. & Wong, H., 2008. Clustering analysis of seismicity and aftershock identification, *Phys. Rev. Lett.,* **101**(1), 018501.
Zhuang, J., Ogata, Y. & Vere-Jones, D., 2002. Stochastic declustering of space-time earthquake occurrences, *J. Am. Stat. Assoc.,* **97**(458), 369–380.
Zhuang, J., Ogata, Y. & Vere-Jones, D., 2004. Analyzing earthquake clustering features by using stochastic reconstruction, *J. Geophys. Res.: Solid Earth,* **109**(B5).

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Table S1. *K*-NN analysis ($b - value = 1$).
Table S2. *K*-NN analysis ($d_f = 1.6$).
Table S3. *K*-NN analysis (higher magnitude of completeness).
Table S4. Kendall-tau test results - Maximum magnitude of *K*-NN.
Table S5. Kendall-tau test results - Spatial distance from the target event and *K*-NN spatial dispersion.
Table S6. Statistical comparison of the distributions of background and triggered earthquake magnitudes (ETAS simulations).
Table S7. *K*-NN analysis (ETAS simulations).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: ETAS CATALOGUE SIMULATION

We generate the synthetic catalogue by implementing the stochastic program described in Felzer *et al.* (2002), which simulates the ETAS model (Ogata 1988) as a branching process. The code implements the following empirical relationships to model the observed seismicity: the Gutenberg–Richter law for earthquake magnitudes, the Omori's law for earthquake times of occurrence and the law describing the decay of aftershock density with distance (Felzer & Brodsky 2006). Monte Carlo techniques are used to draw samples from the empirical distributions

for simulation purpose. The spatial distribution of background earthquakes is a probability density map based on a density grid derived from a real data set. Here we use the Southern California background earthquake rate grid provided with the code. The times of occurrence of background events are randomly generated and their magnitudes are randomly drawn from the Gutenberg–Richter law. In the original code, earthquakes with a magnitude larger than 6.5 are modelled as planar sources. We change that by modelling all the events as point sources (the space–time–magnitude metric that we use assumes that earthquakes are point sources). We use the program to simulate a 20-yr-long catalogue in Southern California, with magnitudes ranging from 3 to 7.9 (but any earthquake with magnitude $\geq$2 can trigger its own offspring). We impose [0.001 100] as the range of possible distances, in km, between triggered and triggering earthquakes, based on the evidence from the SOCAL catalogue. We leave unchanged the remaining parameters needed for the simulation as indicated in the code. These are summarized in Table A1.

Table A1. Input parameters for the stochastic simulator program.

Gutenberg–Richter Law	Omori's Law			Aftershocks decay with distance
b-value	*c*	*p*	*a*	*n*
1	0.095	1.34	0.008	1.37

APPENDIX B: METHODS FOR STAI REMOVAL

In this work, we implement several methods to remove the STAI issue. The first two only define the temporal window of events to be removed after a large shock ($M6+$ earthquake), while the third specifies the critical window in both the temporal and spatial domain:

(1) Rearranging eq. (15) in Helmstetter *et al.* (2006), we find the time (in days) after a large shock such that the catalogue can be considered complete:

$$t = 10^{\frac{M^* - 4.5 - M_c}{0.75}}, \tag{B1}$$

where $M*$ is the main shock magnitude and M_c is the magnitude of completeness estimated for the whole catalogue;

(2) We remove all the events occurred 1 d after a large shock;

(3) We remove all the events occurred 1 week/3 months and within one fault length from a large shock. The fault length is estimated with the Mai & Beroza (2000) empirical scaling relationships.

APPENDIX C: KENDALL-TAU TEST

The Kendall-tau test (Kendall 1948) is a non-parametric test which returns a measure of rank correlation between the variables x and y. Unlike the Pearson correlation test, it is not based on any assumptions about the distributions of x and y and about their relationship. The value τ is found as the difference between the number of concordant pairs and the number of discordant pairs. An observation pair is concordant when both the variables go in the same direction (namely they both increase or decrease). If this is not the case, the pair is discordant. A high value of τ implies that x and y tend to be correlated, as they both go in the same direction. The Kendall-tau-b variation includes an adjustment for the ties (identical values in the vector) issue.

Let us consider two variables x and y. First, we rank the values of the two variables; then, we find the concordant and discordant pairs among all the possible 2-tuples. For $i < j$, a pair is concordant if:

$$\begin{cases} x_i > x_j \ y_i > y_j \\ x_i < x_j \ y_i < y_j \end{cases}. \tag{C1}$$

On the contrary, it is discordant if:

$$\begin{cases} x_i > x_j \ y_i < y_j \\ x_i < x_j \ y_i > y_j \end{cases}. \tag{C2}$$

We define:

$$V = (x_i - x_j)(y_i - y_j). \tag{C3}$$

The Kendall score S is calculated by adding 1 when V is positive, –1 when V is negative, and 0 (ties) if $V = 0$. We have that:

$$S = \sum_{i<j} sgn(x_i - x_j) * sgn(y_i - y_j). \tag{C4}$$

The Kendall's rank correlation coefficient τ is then simply derived by normalizing the score S by the number of possible pairs combinations:

$$\tau = \frac{S}{\frac{N(N-1)}{2}}, \tag{C5}$$

where N is the number of observations.

When there are ties in the data, the Kendall-tau-b variation is considered::

$$\tau_b = \frac{S}{\left[(D - T_x)(D - T_y)\right]^{\frac{1}{2}}}, \tag{C6}$$

where $D = \frac{N(N-1)}{2}$, $T_x = \sum_i \frac{t_i(t_i-1)}{2}$ and t_i is the number of ties in the x variable. T_y is calculated in a similar manner.

The value of τ spans from -1 to 1 and gives a measure of the correlation between the x and y variables. For $N < 50$, the true distribution of the Kendall score can be derived and the relative p-value (termed exact p-value) can be calculated. For $N > 50$, the asymptotic normal probabilities are determined knowing that

$$Z = \frac{S}{\left[\mathrm{var}(S)\right]^{\frac{1}{2}}} \tag{C7}$$

follows approximately a normal standard distribution for large N (e.g. Kalbfleisch 2012), where here:

$$\begin{aligned}\mathrm{var}(S) = {} & \frac{N(N-1)(2N+5) - \sum_x t_i(t_i-1)(2t_i+5) - \sum_y t_i(t_i-1)(2t_i+5)}{18} \\ & + \frac{\left[\sum_x t_i(t_i-1)(t_i-2)\right]\left[\sum_y t_i(t_i-1)(t_i-2)\right]}{9N(N-1)} \\ & + \frac{\left[\sum_x t_i(t_i-1)\right]\left[\sum_y t_i(t_i-1)\right]}{2N(N-1)}.\end{aligned} \tag{C8}$$

The derived p-value can be therefore used to assess if the two variables are statistically independent or not. In this work, we have implemented the Kendall-tau-b method (to handle ties) in order to look for a possible correlation between the medians of the distributions of the K-NN properties values and the target event magnitudes.