The Effect of Declustering on the Size Distribution of Mainshocks

Leila Mizrahi*[1], Shyam Nandan[1], and Stefan Wiemer[1]

Abstract

Declustering aims to divide earthquake catalogs into independent events (mainshocks), and dependent (clustered) events, and is an integral component of many seismicity studies, including seismic hazard assessment. We assess the effect of declustering on the frequency–magnitude distribution of mainshocks. In particular, we examine the dependence of the *b*-value of declustered catalogs on the choice of declustering approach and algorithm-specific parameters. Using the catalog of earthquakes in California since 1980, we show that the *b*-value decreases by up to 30% due to declustering with respect to the undeclustered catalog. The extent of the reduction is highly dependent on the declustering method and parameters applied. We then reproduce a similar effect by declustering synthetic earthquake catalogs with known *b*-value, which have been generated using an epidemic-type aftershock sequence model. Our analysis suggests that the observed decrease in *b*-value must, at least partially, arise from the application of the declustering algorithm on the catalog, rather than from differences in the nature of mainshocks versus fore- or aftershocks. We conclude that declustering should be considered as a potential source of bias in seismicity and hazard studies.

Cite this article as Mizrahi, L., S. Nandan, and S. Wiemer (2021). The Effect of Declustering on the Size Distribution of Mainshocks, *Seismol. Res. Lett.* **92**, 2333–2342, doi: 10.1785/0220200231.

Supplemental Material

Introduction

Models for probabilistic seismic hazard analysis (PSHA; see e.g., Pace *et al.*, 2006; Wiemer, Giardini, *et al.*, 2009; Petersen *et al.*, 2018; Gerstenberger *et al.*, 2020), are commonly based on the approach described by Cornell (1968), which assumes earthquake occurrence times to be representable by a stationary Poisson process. The long-term seismicity rate in a region is considered to be constant in time, reflecting a constant deformation rate and hence constant energy input at any given location, driven by plate tectonics. In reality, earthquakes trigger aftershocks, which in turn trigger their aftershocks, and so on, leading to intense clustering of earthquakes in space and time (Ogata, 1998; Jackson and Kagan, 1999; Helmstetter and Sornette, 2003). Earthquakes can also occur in swarms (Hainzl and Fischer, 2002; Hainzl, 2004), lasting days to months, sometimes comprising thousands of earthquakes in one location, which are followed by long periods of quiescence. Consequently, the recorded earthquake catalogs, especially modern instrumental ones that are complete down to small magnitudes, always show conspicuous deviations from Poissonianity. Average seismicity rates in regions with recent large sequences are therefore not representative of the long-term seismic hazard, indicating a potentially substantial location-dependent bias of seismicity rates.

Aims and challenges of declustering

So-called declustering algorithms aim to divide earthquake catalogs into clusters of dependent events and retain only the independent event of each such cluster. Although Luen and Stark (2012) find that Poissonianity depends on "the declustering method, the catalog, the magnitude range, and the statistical test," it is generally assumed that a properly declustered earthquake catalog satisfies the condition of being Poissonian (Gardner and Knopoff, 1974; van Stiphout *et al.*, 2012). Because of the requirement of Poissonianity for the current approach to PSHA, rate estimation for hazard assessment is often done on the basis of declustered catalogs (Pace *et al.*, 2006; Wiemer, Giardini, *et al.*, 2009; Beauval *et al.*, 2013; Field *et al.*, 2014; Woessner *et al.*, 2015; Meletti *et al.*, 2017; Akinci *et al.*, 2018; Petersen *et al.*, 2018; Sesetyan *et al.*, 2018; Waseem *et al.*, 2019; Drouet *et al.*, 2020). In this sense, PSHA approaches estimate mainshock rates rather than total seismicity rates.

Although Poissonianity of the declustered catalog is necessary for a declustering method to serve its purpose, this condition does not ensure a unique solution to the declustering problem. To avoid inadvertently rewarding the excessive removal of events from the catalog, an additional criterion is required. However, because the actual triggering processes are not currently known and nature does not provide us with labels such as "mainshock," "aftershock," "foreshock," or "swarm member," we lack an objective criterion for the performance

1. Swiss Seismological Service, ETH Zürich, Zürich, Switzerland

*Corresponding author: leila.mizrahi@sed.ethz.ch

evaluation of declustering methods. Several algorithms have been proposed and used in the past (Gardner and Knopoff, 1974; Gruenthal, 1985; Reasenberg, 1985; Uhrhammer, 1986; Zhuang *et al.*, 2002; Marsan and Lengline, 2008; Zaliapin *et al.*, 2008); see van Stiphout *et al.* (2012) for an overview.

Effects of declustering on PSHA

In a study on the effect of declustering on hazard results for the city of Istanbul, Azak *et al.* (2018) found that peak ground acceleration values vary by up to 20% depending on the declustering method. Marzocchi and Taroni (2014) discuss the need for declustering for PSHA, concluding that it is only necessary to avoid a bias in the spatial distribution of earthquake occurrences. Furthermore, considering that aftershocks can also cause considerable damage, they find that the neglecting of aftershock effects due to declustering may lead to significant underestimation of seismicity rates and hence of seismic hazard. In this regard, Iervolino *et al.* (2018) and Iervolino (2019) have proposed a generalization of the hazard integral to reintroduce aftershock hazard in PSHA. Moreover, van Stiphout *et al.* (2011) found that the choice of declustering method has a major effect on seismicity rate-change estimations. On the other hand, sensitivity studies to different declustering approaches in Switzerland have shown that the impact of declustering on the hazard is often negligible (Wiemer, García-Fernández, and Burg, 2009). The need for, potential biases introduced by, and alternatives to declustering have also been discussed in the context of seismicity forecasting (see Schorlemmer and Gerstenberger, 2007; Nandan, Ouillon, Sornette, and Wiemer, 2019). In particular, the issue is raised that a mainshock forecast can only be tested against a mainshock "truth" which is inherently dependent on the somewhat arbitrary choice of declustering method, yielding full seismicity forecasts the only objectively testable type of forecast.

Effects of declustering on the *b*-value

A major role in the calculation of seismicity rates is played by the *b*-value of the empirical Gutenberg–Richter (GR) law (Gutenberg and Richter, 1944), which describes the frequency distribution of earthquake magnitudes. Typically, *b*-values of earthquake catalogs lie close to 1 (Kagan, 1999; Kamer and Hiemer, 2015) but have been found to vary with time, region, depth, and stress regime. Several studies have also reported higher *b*-values during swarms or in volcanic areas (Henderson *et al.*, 1992; Main *et al.*, 1992; Frohlich and Davis, 1993; Wiemer and Wyss, 1997; Wyss *et al.*, 1997; Schorlemmer *et al.*, 2005; Petruccelli *et al.*, 2019). Kagan (1999), Kamer and Hiemer (2015), and Marzocchi *et al.* (2020) discussed a variety of potential technical causes of *b*-value variations, such as magnitude binning, network coverage, catalog incompleteness, or the finiteness of data. Moreover, imposing a GR law on declustered catalogs, as is commonly done in seismic hazard analysis, often results in a significantly lower *b*-value compared to full catalogs (Kagan, 2010; Christophersen *et al.*, 2011; Field *et al.*, 2014; Petersen *et al.*, 2018). Some argue that this behavior is a property naturally inherent to mainshocks (Knopoff, 2000). On a similar note, Gulia *et al.* (2018) suggested that the *b*-value of typical aftershock sequences is on average 20% higher than the mainshock *b*-value, and that this increase in *b*-value is a long-lasting effect for several years.

However, it is debatable whether the *b*-value of a declustered catalog is at all meaningful. Most declustering methods define a mainshock as the largest event of an independent cluster. If one assumes a GR law-type Pareto distribution of magnitudes on the full catalog, one should not at the same time assume a GR law-type Pareto distribution of mainshock magnitudes. The distribution of the maximum of a set of independent and identically distributed random variables, that is, the distribution of mainshock magnitudes, can be derived from fundamental principles of probability theory (Kolmogoroff, 1934). Lombardi (2003) gave a mathematical description of how mainshock magnitude distribution follows from the full-catalog GR law. She showed that the difference in *b*-value between mainshocks and all events becomes minimal when a corrected log-likelihood function is used in the maximum-likelihood estimation of the mainshock *b*-value. The mainshock magnitude distribution she proposed depends on the empirical distribution of cluster sizes emerging from the declustering process. Given this result and assuming that different declustering algorithms will lead to different cluster size distributions, it is expected that different declustering methods will also lead to different mainshock magnitude distributions. Hence, *b*-values of mainshocks, when estimated in the usual way, are expected to be biased as an artifact of declustering.

Similarly, Zhuang and Ogata (2006) found that the magnitude distribution of mainshocks defined via the epidemic-type aftershock sequence (ETAS) model (see their article for the definition or Text S4 in the supplemental material to this article for details on the ETAS model) departs from the GR law and that the full catalog *b*-value is valid for mainshocks in the asymptotic case in which $m \to \infty$. For lower magnitude mainshocks, one could argue that a GR law with lower *a*- and *b*-values than those of the full catalog presents an acceptable approximation of the true, non-Pareto distribution of mainshock magnitudes. However, when the logarithms of the numbers $N(m)$ and $N_{\mathrm{main}}(m)$ of earthquakes and mainshocks of magnitude $M > m$ are both described by linear terms of the form:

$$\log_{10} N(m) = a - b \times m, \tag{1}$$

$$\log_{10} N_{\mathrm{main}}(m) = a_{\mathrm{main}} - b_{\mathrm{main}} \times m, \tag{2}$$

in which $b_{\mathrm{main}} \neq b$, the two lines intersect at a point

$$m_x = \frac{a - a_{\mathrm{main}}}{b - b_{\mathrm{main}}}. \tag{3}$$

If $b_{main} < b$, this means that the expected number of mainshocks of magnitude $M > m_x$ is larger than the expected number of total earthquakes of magnitude $M > m_x$, even though the observed number of mainshocks can never be larger than the observed number of earthquakes.

Paper outline

Considering the importance of the b-value for seismicity studies and seismic hazard estimates, we here systematically assess the influence of declustering on mainshock size distribution. To do so, we first verify that imposing a GR law on mainshocks yields a b-value that does indeed depend on the choice of the declustering method applied. Then, we show that a similar effect is observed for synthetic catalogs with known b-value, whose magnitude distribution by design does not distinguish mainshocks and other events. Furthermore, we illustrate the consequences of approximating mainshock magnitude distribution with a Pareto distribution and calculate the tipping point magnitude m_x, above which the bias introduced by declustering cannot be interpreted as mainshock-specific behavior.

The rest of the article is organized as follows. In the Data section, we describe the earthquake catalog used for this study and discuss the coupled estimation of completeness magnitude and b-value. In the Method section, we describe the declustering methods and corresponding parameter choices. There, we also describe the ETAS model, which is used for the simulation of synthetic catalogs and furthermore serves as the basis for two of the declustering methods. We then present and discuss our main results in the Results and Discussion section and state our Conclusion. The supplemental material to this article contains a more detailed description of all methods and algorithms used. Moreover, it contains analyses of the sensitivity of full catalog b-value and mainshock b-value on the completeness magnitude M_c.

Data

In this study, we use the Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat) provided by the U.S. Geological Survey (see Data and Resources) with "preferred" magnitudes as defined in ComCat, in the collection area around the state of California as in the regional earthquake likelihood models (RELM) testing center (Schorlemmer and Gerstenberger, 2007). The choice of the study region is motivated mainly by the high seismicity in the area and by completeness at low magnitudes of the catalog for several decades (see Hutton *et al.*, 2006), both ensuring that a large and representative amount of data can be used in our study. We consider events of magnitude $M \geq 0.0$, with magnitudes rounded into bins of size $\Delta M = 0.2$. Figure S1 shows that the b-value is insensitive to bin size for reasonable choices of M_c. It also shows that $b(M_c)$ is more stable for $\Delta M = 0.2$ compared to $\Delta M = 0.1$. The time frame used is 1 January 1970 until 30 September 2019, of which only the events on or after 1 January 1980 are used for the estimation of b-values. We subsequently call this set of events the incomplete primary catalog. The earlier events make up the incomplete auxiliary catalog. As earthquake clusters may occur close to the start of the primary time period, ignoring auxiliary events could lead to unwanted deficiencies in cluster detection and mainshock identification (Schoenberg *et al.*, 2010; Wang *et al.*, 2010; Nandan, Ouillon, and Sornette, 2019). Our choice of time periods aims to achieve balance between long enough primary and auxiliary periods, and low completeness magnitude in the primary catalog thanks to seismic network configuration (see e.g., Hutton *et al.*, 2006).

b-value estimation and completeness magnitude

Estimating the b-value of a catalog requires knowledge of its completeness magnitude M_c, the magnitude threshold above which all events are assumed to be detected. Assuming too low values for M_c can cause severe underestimation of the b-value (see Fig. S1). On the other hand, assuming overly conservative values for M_c leads one to discard a large portion of the data, making b-value estimates imprecise. In reality, M_c is not known and has to be estimated itself. Several methods to do so have been proposed; see Mignan and Woessner (2012) for an overview. Commonly, M_c is estimated by defining it as the magnitude threshold above which earthquakes follow the GR law. In this sense, the estimation of b-value and M_c becomes a coupled problem; one cannot be estimated without knowledge of the other. In Texts S1 and S2 and Figure S2, we adapt the method proposed by Clauset *et al.* (2009) to jointly estimate M_c and b-value, ultimately arriving at a value of 3.6 for M_c. A sensitivity analysis (see Fig. S3) shows that the results presented in the following sections are insensitive to reasonable choices of M_c.

Setting the value of M_c to 3.6 implies that we subsequently use the subset of events with magnitude $M \geq 3.6$ of the previously described (binned) catalog. This filter is applied to both the incomplete primary and the incomplete auxiliary catalog, yielding the (complete) primary and (complete) auxiliary catalog. Figure 1a,c shows the spatial and temporal distribution of events in the catalog with magnitude $M \geq 3.6$, for which auxiliary events are highlighted in yellow. For our primary catalog we obtain a b-value of 1.01, as illustrated in Figure 1b.

Method

To better understand the influence of declustering on the b-value, we first apply five often used declustering techniques with different parameter and window choices to the same real catalog. We then apply the same declustering methods, with standard parameters, to a set of 2000 synthetic catalogs. The synthetic catalogs are generated using a basic ETAS model (see Ogata, 1998; Veen and Schoenberg, 2008; Nandan *et al.*, 2017; Nandan, Ouillon, Sornette, and Wiemer, 2019), which is described in Text S4. Table 1 shows the parameters used in the simulation of synthetic catalogs. They were obtained by

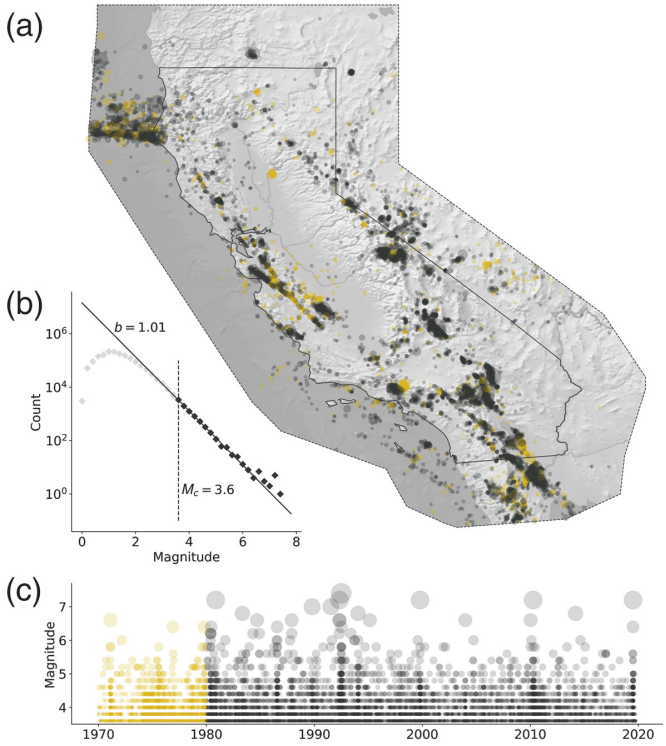

Figure 1. Earthquake catalog used in this analysis. (a) Seismicity map. Dots represent earthquakes in the catalog with $M \geq 3.6$, in which dot size indicates magnitude. Events between 1970 and 1980, which serve as auxiliary data, are marked in yellow. Solid black line marks the California state boundary, dotted line marks the boundary of the considered region. (b) Absolute frequency distribution of magnitudes above and below M_c (black versus gray diamonds). Solid black line shows the Gutenberg–Richter (GR) law fitted to the catalog of events with $M \geq 3.6$. (c) Temporal distribution of the events shown in (a), with identical size and color coding.

applying expectation maximization (Veen and Schoenberg, 2008; Nandan *et al.*, 2017) to the primary and auxiliary California catalog, to support the comparability of real and synthetic catalogs. For a detailed description of the ETAS model, as well as the algorithms used for inversion and simulation, see Text S4. We use the synthetic catalogs to test whether declustering introduces any systematic bias to the mainshock size distribution. As in the case of synthetic catalogs, the distribution from which magnitudes are drawn is known and is assumed to be the same for mainshocks and aftershocks, any changes in b-value observed after declustering must have their origin in the application of declustering algorithms.

To further understand the consequences of approximating mainshock magnitude distribution with a lower-b-value GR law, we compare the ratio

$$r(m) = \frac{N_{\text{main}}(m)}{N(m)} \tag{4}$$

of mainshocks among earthquakes of magnitude $M > m$ between observation and approximation, for the different declustering methods with standard parameter settings applied to the Californian catalog. We calculate m_x (see equation 3), above which $r(m) > 1$ implies that the introduced bias can impossibly be supported by observations.

We examine the declustering methods proposed by Reasenberg (1985), Zaliapin *et al.* (2008), and window methods as proposed by Gardner and Knopoff (1974), Gruenthal (1985), and Uhrhammer (1986). We also consider two versions of declustering based on the ETAS model (Zhuang *et al.*, 2002). For the detailed description of all declustering algorithms and parameter ranges applied, see Texts S3 and S4 and Tables S1–S3; we give a short description of each method subsequently. The nonparametric stochastic declustering algorithm proposed by Marsan and Lengline (2008) is not used here. This is because of its similarity to the already considered parametric stochastic declustering alternative provided by the ETAS model. The main difference to ETAS declustering is that the triggering rate, described as $g(t, x, y, m)$ in Text S4, is there obtained empirically, without presuming the laws described by equations (S12–S14). In their analysis of southern California seismicity, they observe that their empirically derived triggering rate follows laws similar to equations (S12–S14), which, similarly, were originally discovered empirically.

Short descriptions of the declustering methods applied in this article

1. Reasenberg (1985) introduced an algorithm that has been used in numerous studies and recent PSHA, for example, in Ecuador (Beauval *et al.*, 2013) or Afghanistan (Waseem *et al.*, 2019). It defines earthquake interaction zones in space and time. Here, we apply the spatial interaction relationships proposed by Reasenberg (1985), and Wells and Coppersmith (1994), and the parameter ranges for temporal interaction zones recommended by Schorlemmer and Gerstenberger (2007).
2. Window methods, as first described by Gardner and Knopoff (1974), use space–time windows around large events to identify their fore- and aftershocks. Different formulations of such window boundaries have been suggested and are applied in this study (Gardner and Knopoff, 1974; Gruenthal, 1985; Uhrhammer, 1986; see van Stiphout *et al.*, 2012). We use the original formulation by Gardner and Knopoff (1974) as the standard window, which is used in the Uniform California Earthquake Rupture Forecast, Version 3 (Field *et al.*, 2014). Generally, window methods are widely used in modern regional and national seismic hazard models, see Drouet *et al.* (2020) for France, Meletti *et al.* (2017) for Italy, Sesetyan *et al.* (2018) for Turkey, and Woessner *et al.* (2015) for Europe (ESHM13).

3. The Zaliapin *et al.* (2008) alternative approach applies a Gaussian mixture model on space–time nearest-neighbor distances between events to distinguish independent from dependent events.
4. The ETAS model is used here in two ways. First, it is used to simulate synthetic earthquake catalogs upon which declustering methods are applied to study their effects. Second, the ETAS model induces an alternative, parametric approach to declustering, which was introduced by Zhuang *et al.* (2002). We consider two versions of declustering based on the ETAS model, which differ in their definition of mainshocks and are described in detail in Text S4. "ETAS-main" defines the largest event of a cluster to be the mainshock, whereas "ETAS-background" defines events to be mainshocks if they are not triggered. The definition used in ETAS-background is in the spirit of the ETAS model, for which background earthquakes of any size can trigger cascades of aftershocks. ETAS-main, on the other hand, imposes the mainshock definition used in the other methods, in the interest of comparability. We subsequently call those methods that define mainshock as the largest events "mainshock methods." Because of its different definition of mainshocks, ETAS-background is unsuited to be applied in the standard PSHA approach, which is designed to work with mainshock methods.

TABLE 1
ETAS Parameters Used for Catalog Simulation, Obtained by Expectation Maximization

Parameter	Value
$\log_{10}(k_0)$	−2.49
a	1.69
$\log_{10}(c)$	−2.95
ω	−0.03
$\log_{10}(\tau)$	3.99
$\log_{10}(d)$	−0.35
γ	1.22
ρ	0.51
$\log_{10}(\mu)$	−7.17

ETAS, epidemic-type aftershock sequence.

Results and Discussion

The disparity between declustering methods

The cumulative number of mainshocks for different declustering methods with standard parameter and window choices, compared to the full California catalog, is shown in Figure 2a. The diversity among the resulting declustered catalogs is remarkable. Mainshock rates vary by a factor of 6.1 between the most and least "aggressive" algorithm. Moreover, although the removal of temporal clusters is the primary goal of the declustering process, some are still clearly visible after declustering with Reasenberg's method, and still recognizable, though less pronounced, after applying Zaliapin's method. Gardner-Knopoff and ETAS (main and background) appear to be more successful at achieving temporal Poissonianity.

The observed and fitted complementary cumulative frequency functions (CCFFs) of mainshock magnitudes are shown in Figure 2b,c. Observed absolute frequencies of large events ($M \geq 6.4$) are somewhat similar for all declustering methods, with the exception of ETAS-background. Relative frequencies of large events versus small events vary strongly between methods, which manifests itself in slope differences between the CCFFs. For the mainshock methods, the aggressiveness of the method coincides with the extent of slope decrease. This effect can be explained as a consequence of the methods' mainshock definition. Because small events are less likely to be identified as mainshocks, they are more likely to be removed from the catalog, increasing relative frequencies of large events. ETAS-background, in contrast, does not seem to preferentially remove events from specific magnitude ranges.

Figure 2d illustrates the consequences of estimating seismic hazard based on a mainshock GR law with lower b-value. For the California catalog, observed and approximated evolutions of $r(m) = \frac{N_{\text{main}}(m)}{N(m)}$ are shown for mainshocks obtained by Gardner-Knopoff declustering. The magnitude m_x (see equation 3), above which rate overestimation cannot be denied, is given also for Reasenberg, Zaliapin, and ETAS-main declustered catalogs (see Fig. S4 for the corresponding plots). m_x varies between 6.9 and 8.8, in which higher values of m_x are predominantly observed for declustering methods that do not succeed at achieving Poissonianity in time. Furthermore, most declustering methods show considerable deviations of observed $r(m)$ from its approximation already at lower magnitudes. For instance, the approximation of $r_{6.6} = r(6.6)$ lies between 0.72 and 0.86, depending on mainshock definition, even though all definitions except ETAS-main classified all $M > 6.6$ events to be mainshocks. ETAS-background is excluded from this part of the analysis due to its inapplicability in the standard PSHA approach.

b-value of declustered catalogs

Observations on real data. The relative frequency increase of large events translates into a lower b-value when a GR law is imposed on the frequency–magnitude distribution of mainshocks. In Figure 3a, b-value is plotted against a-value of the declustered California catalog, comparing the effects of varying declustering methods and parameters. We find that b-values of declustered catalogs vary strongly with declustering algorithms. Values between 0.73 and 1.00 are attained without any significant gap. A general trend is recognizable among the mainshock methods: removal of more events correlates with

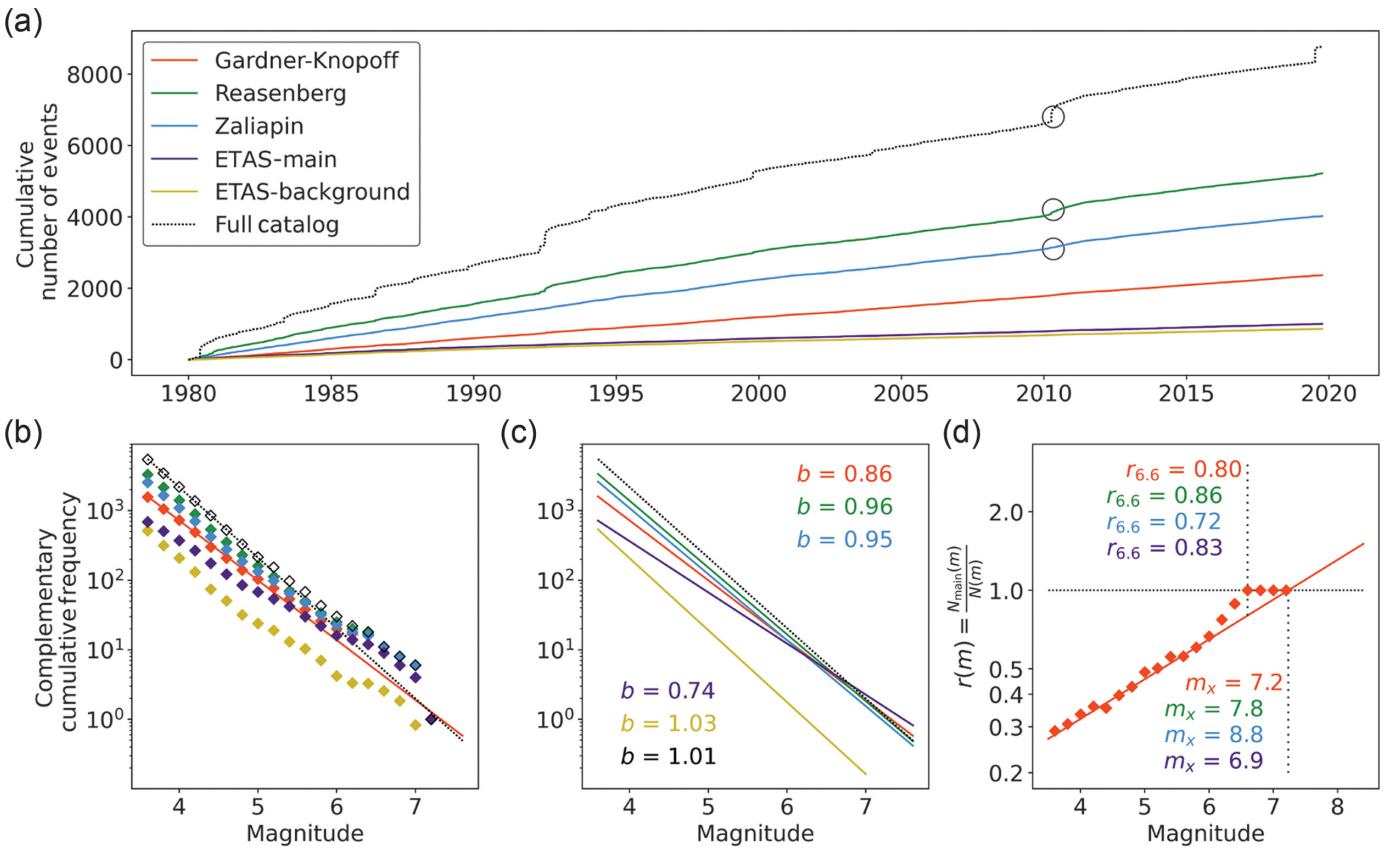

Figure 2. Properties of the California catalog of mainshocks larger than or equal to M 3.6, depending on declustering method. Standard parameter settings (and standard window) of each method are used for declustering. (a) Cumulative number of mainshocks. Dotted black line represents the full (nondeclustered) catalog. The rapid increase in seismicity highlighted in circles corresponds to the 2010 M_w 7.2 El Mayor–Cucapah earthquake in Baja California, Mexico. (b) Empirical complementary cumulative frequency function (CCFF, diamonds) of mainshock magnitudes. Empty black diamonds represent the full California catalog. The two lines are the fitted CCFF for Gardner-Knopoff declustered and full catalog. (c) Fitted CCFF for declustered catalogs compared to full catalog GR law fit. Fitted b-values are given. (d) Observed (diamonds) and approximated (line) evolution of $r(m) = \frac{N_{main}(m)}{N(m)}$ for the Gardner-Knopoff declustered catalog. Black dotted line marks $r(m) \equiv 1$. m_x and $r_{6.6}$ are given, also for Reasenberg, Zaliapin, and epidemic-type aftershock sequence (ETAS)-main declustered catalogs. Note the different x axis for (d) compared to (b) and (c).

lower b-values, indicating a penchant of these methods to relatively remove more smaller events than larger ones. The b-value obtained with ETAS-background does not significantly differ from the full catalog b-value. These observations are in line with the explanation given earlier, which describes the b-value decrease as a consequence of the mainshock definition, and are expected knowing the results by Lombardi (2003), Zhuang and Ogata (2006), Kagan (2010), and van Stiphout *et al.* (2011). A sensitivity analysis of the b-value to the completeness magnitude M_c (see Fig. S3) shows that the b-value decrease after declustering is an effect that is observed regardless of the reasonable choice of M_c, with the extent of the decrease being characteristic of each method.

Observations on simulated data. Synthetic catalogs, for which all magnitudes are drawn from one single distribution, show lower b-values after declustering. In Figure 3b, the distribution of mainshock b-values of 2000 ETAS-simulated catalogs is shown for different declustering methods with standard parameter settings, aligned according to the median observed a-value of the respective method. The mainshock b-values of the same methods applied to the California catalog are indicated with stars; error bars mark the estimated standard error. If no declustering, or ETAS-background declustering, is applied, the estimated b-value of synthetic catalogs is consistent with the b-value used in their simulation. At the same time, b-values of synthetic catalogs declustered with mainshock methods are always lower than the b-value used in their simulation. Comparing the extent of the effect across different declustering methods, synthetic and real data have the same qualitative behavior. Similarly, the a-value decrease is observed to be method-characteristic.

The effect of declustering on the b-value is more pronounced in synthetic data, for all methods. A possible explanation for this is that all declustering methods assume

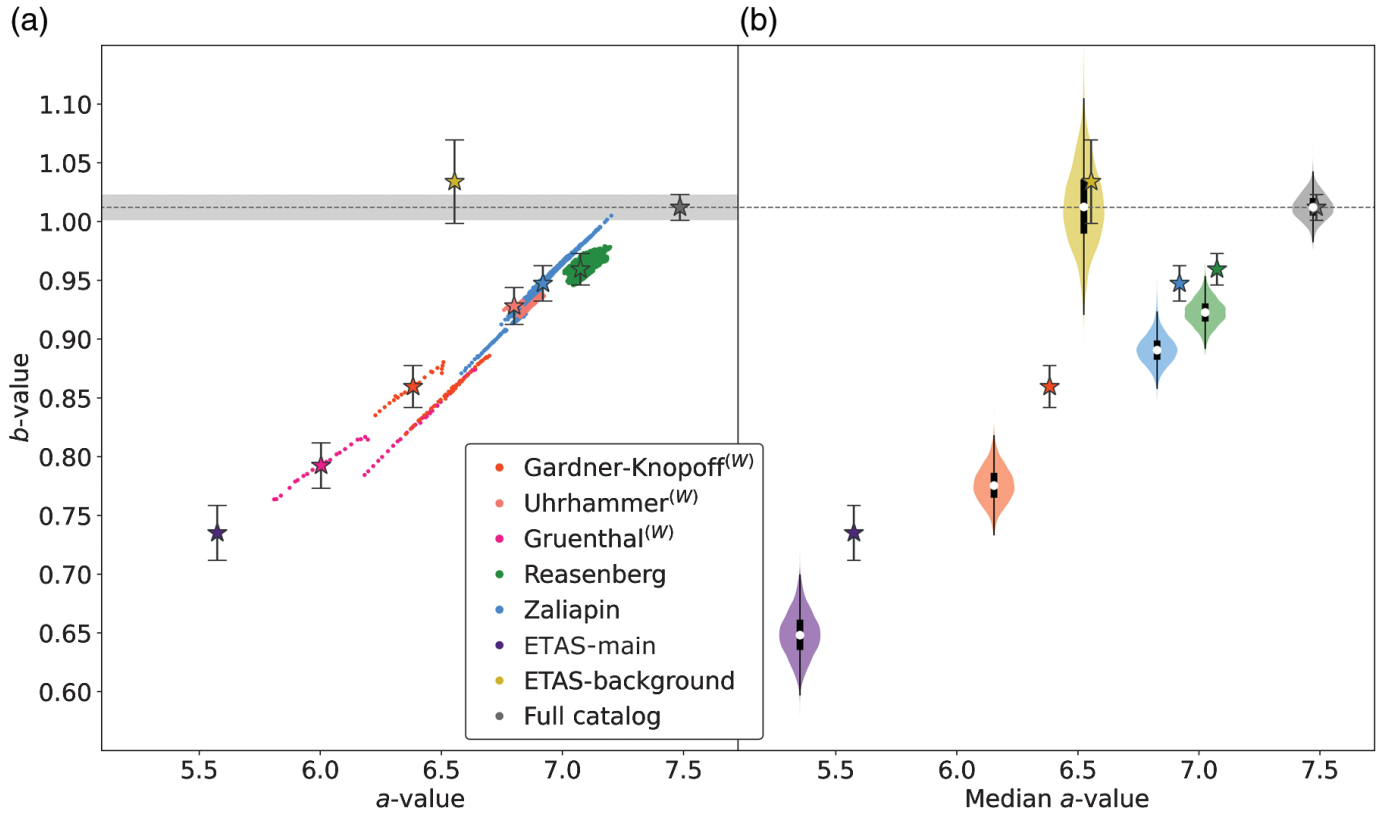

Figure 3. (a) *a*-value versus *b*-value of the declustered California catalog depending on declustering method and parameters. Each dot represents one variation of parameter settings, stars with error bars represent standard parameter settings. Marked with (W) are window methods. The dotted gray line and gray area indicate the *b*-value of the nondeclustered catalog and its uncertainty. (b) Distribution of mainshock *b*-values of 2000 simulated catalogs, depending on declustering method (with standard parameter settings and standard (Gardner-Knopoff) window), plotted against median *a*-value per method. Stars with error bars represent the *a*- and *b*-value of the regional earthquake catalog from (a) for the respective methods. White dot, black box, and black line represent median, interquartile range, and adjacent values, respectively. The dotted line displays the *b*-value used for catalog generation, which corresponds to the full-catalog *b*-value observed in the Californian primary catalog.

isotropic spatial distribution of aftershocks, which is known to be wrong in reality, but valid for synthetic catalogs. Hence, cluster detection is facilitated in synthetic catalogs, resulting in more effective removal of small events compared to the real catalog.

Remarks regarding ETAS declustering

1. Despite ETAS being the generative process of synthetic catalogs, a large difference in *b*-value is observed after ETAS-main declustering. This is not a flaw in the generative process or the parameter inversion. On the contrary, this behavior is expected. The low *b*-value of ETAS-main-declustered catalogs is due to the imposed definition of mainshock as the largest event of a cluster, not to be confused with ETAS' notion of background events. This concept of mainshocks is not relevant in the generation of catalogs. Imposing such a definition leads to selective removal of small events rather than to the removal of aftershocks in the true ETAS sense. With ETAS, aftershocks are temporally restricted to occur after their triggering events but may have larger magnitudes.
2. Declustering with ETAS-background allows a comparison between the *b*-value of background events according to the ETAS definition and the full catalog. No significant difference is observed, both in the case of synthetic and real catalogs. The difference only arises when the rule of maximum magnitude is applied.
3. For synthetic data, the underlying branching structure is known by design of the experiment. In contrast, cluster detection for real data requires the lengthy process of inversion and probabilistic cluster assignment. Thus, compared to synthetic data, cluster detection is intrinsically less correct for real data, and declustering is inclined to be less effective. It is reasonable to assume that this circumstance explains the particularly pronounced difference in *b*-value in the case of ETAS-main declustering.

ETAS simulations do not distinguish the magnitude distribution of mainshocks versus aftershocks. A difference in pre- and

post-declustering b-value of a declustered catalog that was generated using ETAS can, therefore, only have its origin in the systematic selection of large events as mainshocks. The purely declustering-induced and strongly method-dependent decrease in b-value suggests that other potential causes, such as a different nature of mainshocks compared to fore- or aftershocks, have negligible effects on the mainshock b-value. Although the possibility cannot be precluded that a part of the effect is due to the change in stress state before and after major events (e.g., Gulia *et al.*, 2018), the notably arbitrary effect of declustering should not be ignored. The mere observation of a lower b-value of mainshocks is no proof of its meaningfulness; the observation of artifactual effects of declustering on the mainshock b-value, however, is a reason to doubt its meaningfulness.

Conclusion

We demonstrate that a decrease in overall b-value of the California catalog after declustering is observed for a variety of declustering methods and parameter settings. Furthermore, the extent of the decrease is highly dependent on the algorithm applied. A general trend is observed, suggesting that more "aggressive" algorithms tend to be accompanied by a more pronounced b-value decrease, ETAS-background being the only exception to this rule. With a medial resulting a-value among the methods considered, it leaves the b-value unchanged. Finally, we find that all the previously described effects can be reproduced in synthetic data, which was generated using a constant b-value for all events.

Our results indicate that declustering substantially affects the earthquake size distribution. Imposing a GR law on declustered catalogs, therefore, leads to b-values that are biased to a somewhat arbitrary and not immediately apparent extent. This bias leads to an overestimation of seismic hazard above a certain magnitude m_x. Thus, we can conclude that the current state of practice of equating seismic hazard with mainshock rates that follow a GR law can be accused of three deficiencies. One is the nonverifiability of any mainshock definition. Second, fore- or aftershocks can be large and devastating. Neglecting aftershock effects may give rise to a substantial underestimation of seismic hazard (Marzocchi and Taroni, 2014). And finally, the earthquake size distribution resulting from the procedure causes hazard overestimation for events above a certain size.

One may argue that increasing the relative frequency of large events and decreasing the absolute frequency of all events have antagonistic effects on the absolute frequency of large events, justifying any choice of declustering method. Indeed, most hazard studies ignore previous findings and continue to calculate hazard in the usual way. However, we believe that two wrongs do not make a right. To be precise, two wrongs make a right only for one particular magnitude m_x. Our analysis suggests that above m_x, classical hazard studies overestimate the seismic hazard, whereas below m_x, they underestimate it. Although the accusation of underestimation could partially be rejected by insisting that declustering reveals the true mainshocks and that aftershock effects are deliberately excluded from the scope of hazard assessment, we have shown that the overestimation cannot be similarly attributed to a true b-value that is revealed by declustering, but that this resulting b-value is biased to a nonnegligible extent.

It is crucial to be aware of this issue when estimating seismic hazard. Although analysis of earthquake dependency is inevitable to eliminate spatial bias for the calculation of seismicity rates (Marzocchi and Taroni, 2014), basing calculations solely on declustered catalogs is not an appropriate approach. One alternative possibility is to use ETAS models to assess seismicity rates (see e.g., Field *et al.*, 2015). In a pseudoprospective forecasting experiment on the Californian catalog conducted by Nandan, Ouillon, Sornette, and Wiemer (2019), ETAS models generally outperform all competing smoothed seismicity models and models based on strain rates. Using hundreds of thousands of simulations of possible scenarios as the basis for a forecast, they intrinsically account for the spatiotemporal clustering of earthquakes. This approach incorporates the non-Poissonian nature of reality while reducing the spatial bias encountered in undeclustered catalogs. At the same time, ETAS relies only on the GR law of the full catalog and therefore avoids making assumptions on the frequency–magnitude distribution of somewhat arbitrarily selected large events.

Other ways to address this matter may exist. What is essential is to recognize the problematic aspects of doing hazard assessment based on declustered catalogs and to find a way to address the issues presented here.

Data and Resources

The Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat) provided by the U.S. Geological Survey (USGS) was searched using https://earthquake.usgs.gov/data/comcat/ (last accessed November 2019). The supplemental material to this article contains a more detailed description of all methods and algorithms used. Moreover, it contains analyses of the sensitivity of full catalog b-value and mainshock b-value on the completeness magnitude M_c. Finally, observed and approximated ratio of mainshocks among earthquakes of magnitude $M > m$ is shown using different declustering methods for mainshock definition.

Acknowledgments

The authors wish to thank Celso Reyes for providing Python implementations of the declustering algorithms for Reasenberg, Zaliapin, and window method declustering, as well as Laurentiu Danciu, Arnaud Mignan, as well as two anonymous reviewers for their helpful feedback on an earlier version of this article. This work has received funding from the Eidgenössische Technische Hochschule (ETH) research grant for project number 2018-FE-213, "Enabling dynamic earthquake risk assessment (DynaRisk)" and from the European Union's Horizon 2020 research and innovation program under

Grant Agreement Number 821115, real-time earthquake risk reduction for a resilient Europe (RISE).

References

Akinci, A., M. P. Moschetti, and M. Taroni (2018). Ensemble smoothed seismicity models for the new Italian probabilistic seismic hazard map, *Seismol. Res. Lett.* **89,** no. 4, 1277–1287.

Azak, T. E., D. Kalafat, K. Şeşetyan, and M. B. Demircioğlu (2018). Effects of seismic declustering on seismic hazard assessment: A sensitivity study using the Turkish earthquake catalogue, *Bull. Earthq. Eng.* **16,** no. 8, 3339–3366.

Beauval, C., H. Yepes, P. Palacios, M. Segovia, A. Alvarado, Y. Font, J. Aguilar, L. Troncoso, and S. Vaca (2013). An earthquake catalog for seismic hazard assessment in Ecuador,*Bull. Seismol. Soc. Am.* **103,** no. 2A, 773–786.

Christophersen, A., M. C. Gerstenberger, D. A. Rhoades, and M. W. Stirling (2011). Quantifying the effect of declustering on probabilistic seismic hazard, *Proc. of the Ninth Pacific Conf. on Earthquake Engineering: Building an Earthquake-Resilient Society*, Auckland, New Zealand, 14–16 April 2011, Paper Number 206.

Clauset, A., C. R. Shalizi, and M. E. Newman (2009). Power-law distributions in empirical data, *SIAM Rev.* **51,** no. 4, 661–703.

Cornell, C. A. (1968). Engineering seismic risk analysis, *Bull. Seismol. Soc. Am.* **58,** no. 5, 1583–1606.

Drouet, S., G. Ameri, K. Le Dortz, R. Secanell, and G. Senfaute (2020). A probabilistic seismic hazard map for the metropolitan France, *Bull. Earthq. Eng.* **18,** 1865–1898.

Field, E. H., R. J. Arrowsmith, G. P. Biasi, P. Bird, T. E. Dawson, K. R. Felzer, D. D. Jackson, K. M. Johnson, T. H. Jordan, C. Madden, *et al.* (2014). Uniform California earthquake rupture forecast, version 3 (UCERF3)—The time-independent model, *Bull. Seismol. Soc. Am.* **104,** no. 3, 1122–1180.

Field, E. H., G. P. Biasi, P. Bird, T. E. Dawson, K. R. Felzer, D. D. Jackson, K. M. Johnson, T. H. Jordan, C. Madden, A. J. Michael, *et al.* (2015). Long-term time-dependent probabilities for the third Uniform California Earthquake Rupture Forecast (UCERF3), *Bull. Seismol. Soc. Am.* **105,** no. 2A, 511–543.

Frohlich, C., and S. D. Davis (1993). Teleseismic b values; or, much ado about 1.0, *J. Geophys. Res.* **98,** no. B1, 631–644.

Gardner, J. K., and L. Knopoff (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bull. Seismol. Soc. Am.* **64,** no. 5, 1363–1367.

Gerstenberger, M. C., W. Marzocchi, T. Allen, M. Pagani, J. Adams, L. Danciu, E. H. Field, H. Fujiwara, N. Luco, K. F. Ma, *et al.* (2020). Probabilistic seismic hazard analysis at regional and national scales: State of the art and future challenges, *Rev. Geophys.* **58,** no. 2, e2019RG000653, doi: 10.1029/2019RG000653.

Gruenthal, G. (1985). The up-dated earthquake catalogue for the German Democratic Republic and adjacent areas - statistical data characteristics and conclusions for hazard assessment, *3rd International Symposium on the Analysis of Seismicity and Seismic Risk*, Liblice/Czechoslovakia, 17–22 June 1985, Vol. I, 19–25.

Gulia, L., A. P. Rinaldi, T. Tormann, G. Vannucci, B. Enescu, and S. Wiemer (2018). The effect of a mainshock on the size distribution of the aftershocks, *Geophys. Res. Lett.* **45,** no. 24, 13,277–13,287.

Gutenberg, B., and C. F. Richter (1944). Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.* **34,** no. 4, 185–188.

Hainzl, S. (2004). Seismicity patterns of earthquake swarms due to fluid intrusion and stress triggering, *Geophys. J. Int.* **159,** no. 3, 1090–1096.

Hainzl, S., and T. Fischer (2002). Indications for a successively triggered rupture growth underlying the 2000 earthquake swarm in Vogtland/NW Bohemia, *J. Geophys. Res.* **107,** no. B12, ESE-5.

Helmstetter, A., and D. Sornette (2003). Importance of direct and indirect triggered seismicity in the ETAS model of seismicity, *Geophys. Res. Lett.* **30,** no. 11, doi: 10.1029/2003GL017670.

Henderson, J., I. Main, P. Meredith, and P. Sammonds (1992). The evolution of seismicity at Parkfield: Observation, experiment and a fracture-mechanical interpretation, *J. Struct. Geol.* **14,** nos. 8/9, 905–913.

Hutton, K., E. Hauksson, J. Clinton, J. Franck, A. Guarino, N. Scheckel, D. Given, and A. Yong (2006). Southern California seismic network update, *Seismol. Res. Lett.* **77,** no. 3, 389–395.

Iervolino, I. (2019). Generalized earthquake counting processes for sequence-based hazard, *Bull. Seismol. Soc. Am.* **109,** no. 4, 1435–1450.

Iervolino, I., E. Chioccarelli, and M. Giorgio (2018). Aftershocks' effect on structural design actions in Italy, *Bull. Seismol. Soc. Am.* **108,** no. 4, 2209–2220.

Jackson, D. D., and Y. Y. Kagan (1999). Testable earthquake forecasts for 1999, *Seismol. Res. Lett.* **70,** no. 4, 393–403.

Kagan, Y. Y. (1999). Universality of the seismic moment-frequency relation, in *Seismicity Patterns, Their Statistical Significance and Physical Meaning*, M. Wyss, K. Shimazaki, and A. Ito (Editors), Pageoph Topical Volumes, Birkhäuser, Basel, Switzerland, doi: 10.1007/978-3-0348-8677-2_16.

Kagan, Y. Y. (2010). Earthquake size distribution: Power-Law with exponent $\beta \equiv 1/2$? *Tectonophysics* **490,** nos. 1/2, 103–114.

Kamer, Y., and S. Hiemer (2015). Data-driven spatial *b* value estimation with applications to California seismicity: To *b* or not to *b*, *J. Geophys. Res.* **120,** no. 7, 5191–5214.

Knopoff, L. (2000). The magnitude distribution of declustered earthquakes in Southern California, *Proc. Natl. Acad. Sci.* **97,** no. 22, 11,880–11,884.

Kolmogoroff, A. (1934). Grundbegriffe der Wahrscheinlichkeitsrechnung, *Bull. Am. Math. Soc.* **40,** 522–523.

Lombardi, A. M. (2003). The maximum likelihood estimator of b-value for mainshocks, *Bull. Seismol. Soc. Am.* **93,** no. 5, 2082–2088.

Luen, B., and P. B. Stark (2012). Poisson tests of declustered catalogues, *Geophys. J. Int.* **189,** no. 1, 691–700.

Main, I. G., P. G. Meredith, and P. R. Sammonds (1992). Temporal variations in seismic event rate and b-values from stress corrosion constitutive laws, *Tectonophysics* **211,** nos. 1/4, 233–246.

Marsan, D., and O. Lengline (2008). Extending earthquakes' reach through cascading, *Science* **319,** no. 5866, 1076–1079.

Marzocchi, W., and M. Taroni (2014). Some thoughts on declustering in probabilistic seismic-hazard analysis, *Bull. Seismol. Soc. Am.* **104,** no. 4, 1838–1845.

Marzocchi, W., I. Spassiani, A. Stallone, and M. Taroni (2020). How to be fooled searching for significant variations of the b-value, *Geophys. J. Int.* **220,** no. 3, 1845–1856.

Meletti, C., W. Marzocchi, D. Albarello, V. D'Amico, L. Luzi, F. Martinelli, B. Pace, M. Pignone, A. N. Rovida, F. Visini, *et al.*

(2017). The 2016 Italian seismic hazard model, *16th World Conf. on Earthquake Engineering*, Santiago, Chile, 9–13 January 2017.

Mignan, A., and J. Woessner (2012). Estimating the magnitude of completeness for earthquake catalogs, *Community Online Resource for Statistical Seismicity Analysis*, 1–45.

Nandan, S., G. Ouillon, and D. Sornette (2019). Magnitude of earthquakes controls the size distribution of their triggered events, *J. Geophys. Res.* **124,** no. 3, 2762–2780.

Nandan, S., G. Ouillon, D. Sornette, and S. Wiemer (2019). Forecasting the rates of future aftershocks of all generations is essential to develop better earthquake forecast models, *J. Geophys. Res.* **124,** no. 8, 8404–8425.

Nandan, S., G. Ouillon, S. Wiemer, and D. Sornette (2017). Objective estimation of spatially variable parameters of epidemic type aftershock sequence model: Application to California, *J. Geophys. Res.* **122,** no. 7, 5118–5143.

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences, *Ann. Inst. Stat. Math.* **50,** no. 2, 379–402.

Pace, B., L. Peruzza, G. Lavecchia, and P. Boncio (2006). Layered seismogenic source model and probabilistic seismic-hazard analyses in central Italy, *Bull. Seismol. Soc. Am.* **96,** no. 1, 107–132.

Petersen, M. D., C. S. Mueller, M. P. Moschetti, S. M. Hoover, K. S. Rukstales, D. E. McNamara, R. A. Williams, A. M. Shumway, P. M. Powers, P. S. Earle, *et al.* (2018). 2018 one-year seismic hazard forecast for the central and eastern United States from induced and natural earthquakes, *Seismol. Res. Lett.* **89,** no. 3, 1049–1061.

Petruccelli, A., P. Gasperini, T. Tormann, D. Schorlemmer, A. P. Rinaldi, G. Vannucci, and S. Wiemer (2019). Simultaneous dependence of the earthquake-size distribution on faulting style and depth, *Geophys. Res. Lett.* **46,** no. 20, 11,044–11,053.

Reasenberg, P. (1985). Second-order moment of central California seismicity, 1969–1982, *J. Geophys. Res.* **90,** no. B7, 5479–5495.

Schoenberg, F. P., A. Chu, and A. Veen (2010). On the relationship between lower magnitude thresholds and bias in epidemic-type aftershock sequence parameter estimates, *J. Geophys. Res.* **115,** no. B4, doi: 10.1029/2009JB006387.

Schorlemmer, D., and M. C. Gerstenberger (2007). RELM testing center, *Seismol. Res. Lett.* **78,** no. 1, 30–36.

Schorlemmer, D., S. Wiemer, and M. Wyss (2005). Variations in earthquake-size distribution across different stress regimes, *Nature* **437,** no. 7058, 539–542.

Sesetyan, K., M. B. Demircioglu, T. Y. Duman, T. Can, S. Tekin, T. E. Azak, and Ö. Z. Fercan (2018). A probabilistic seismic hazard assessment for the Turkish territory—part I: The area source model, *Bull. Earthq. Eng.* **16,** no. 8, 3367–3397.

Uhrhammer, R. A. (1986). Characteristics of northern and central California seismicity, *Earthq. Notes* **57,** no. 1, 21.

van Stiphout, T., D. Schorlemmer, and S. Wiemer (2011). The effect of uncertainties on estimates of background seismicity rate, *Bull. Seismol. Soc. Am.* **101,** no. 2, 482–494.

van Stiphout, T., J. Zhuang, and D. Marsan (2012). Seismicity declustering, *Community Online Resource for Statistical Seismicity Analysis*, 10, no. 1, doi: 10.5078/corssa-52382934.

Veen, A., and F. P. Schoenberg (2008). Estimation of space–time branching process models in seismology using an em–type algorithm, *J. Am. Stat. Assoc.* **103,** no. 482, 614–624.

Wang, Q., D. D. Jackson, and J. Zhuang (2010). Missing links in earthquake clustering models, *Geophys. Res. Lett.* **37,** no. 21, doi: 10.1029/2010GL044858.

Waseem, M., A. Lateef, I. Ahmad, S. Khan, and W. Ahmed (2019). Seismic hazard assessment of Afghanistan, *J. Seismol.* **23,** no. 2, 217–242.

Wells, D. L., and K. J. Coppersmith (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Am.* **84,** no. 4, 974–1002.

Wiemer, S., and M. Wyss (1997). Mapping the frequency-magnitude distribution in asperities: An improved technique to calculate recurrence times? *J. Geophys. Res.* **102,** no. B7, 15,115–15,128.

Wiemer, S., M. García-Fernández, and J. P. Burg (2009). Development of a seismic source model for probabilistic seismic hazard assessment of nuclear power plant sites in Switzerland: The view from PEGASOS expert group 4 (EG1d), *Swiss J. Geosci.* **102,** no. 1, 189–209.

Wiemer, S., D. Giardini, D. Fäh, N. Deichmann, and S. Sellami (2009). Probabilistic seismic hazard assessment of Switzerland: Best estimates and uncertainties, *J. Seismol.* **13,** no. 4, 449.

Woessner, J., D. Laurentiu, D. Giardini, H. Crowley, F. Cotton, G. Grünthal, G. Valensise, R. Arvidsson, R. Basili, M. B. Demircioglu, *et al.* (2015). The 2013 European seismic hazard model: Key components and results, *Bull. Earthq. Eng.* **13,** no. 12, 3553–3596.

Wyss, M., K. Shimazaki, and S. Wiemer (1997). Mapping active magma chambers by b values beneath the off-Ito volcano, Japan, *J. Geophys. Res.* **102,** no. B9, 20,413–20,422.

Zaliapin, I., A. Gabrielov, V. Keilis-Borok, and H. Wong (2008). Clustering analysis of seismicity and aftershock identification, *Phys. Rev. Lett.* **101,** no. 1, 018501.

Zhuang, J., and Y. Ogata (2006). Properties of the probability distribution associated with the largest event in an earthquake cluster and their implications to foreshocks, *Phys. Rev. E* **73,** no. 4, 046134.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2002). Stochastic declustering of space-time earthquake occurrences, *J. Am. Stat. Assoc.* **97,** no. 458, 369–380.

Manuscript received 2 July 2020
Published online 17 February 2021