

Comparison of earthquake catalogs for the Korean Peninsula declustered using three different methods

Sung Kyun Kim¹

⁽¹⁾ Chonnam National University, 77 Yongbong-ro, Buk-gu, Gwangju 61186 Korea

Article history: received March 6, 2023; accepted September 6, 2023

Abstract

Earthquake catalogs include dependent earthquakes, which are spatiotemporally related, and independent or background earthquakes. In order to predict long-term seismicity or conduct a seismic hazard assessment, the dependent earthquakes must be removed to generate a declustered earthquake catalog. Several declustering methods have been proposed to date; however, the result of seismic hazard assessment may vary depending on which declustering methods are selected. In the present study, the catalog of earthquakes that were observed between 2016 and 2021 in and around the Korean Peninsula is declustered using the methods proposed by Gardner and Knopoff [1974], Reasenberg [1985], and Zhuang et al. [2002], and the resultant catalogs are compared. The seismicity parameters (a - and b -values) in the Gutenberg–Richter relationship are found to vary among the three declustered catalogs, thus affecting long-term earthquake predictions and seismic hazard assessment. The raw (original) and three declustered catalogs are also tested to see whether they follow the Poisson process. The minimum magnitude (M_p) above which the null hypothesis of the Poisson process cannot be rejected in the earthquake catalogs ranges from 1.6 to 2.2, depending on the declustered catalog. Further, the M_p obtained herein shows a large value compared to the completeness magnitude estimated in the present study. Comparing the curves representing the cumulative number of background earthquakes against the elapsed time for the declustered catalogs shows that the method by Zhuang et al. [2002] produces the result in the closest agreement with the real background seismicity curve.

Keywords: Earthquake catalog; Seismicity declustering method; Poisson process test; Korean Peninsula

1. Introduction

An earthquake catalog is required when studying a particular area's seismicity or hazards. The catalog includes dependent and background (or independent) earthquakes [van Stiphout et al., 2012]. Background earthquakes, referred to as mainshocks or parent earthquakes, are generally caused by permanent structural stresses relating to plate tectonic movements. In contrast, dependent earthquakes include foreshocks caused by temporary changes

in stress before a mainshock and aftershocks caused by slippage after the fault motion related to a mainshock. As such, dependent earthquakes are spatiotemporally related to mainshocks; thus, their removal from earthquake catalogs, called seismicity declustering, is a crucial consideration in seismology [van Stiphout et al., 2012]. The ultimate goal of seismicity declustering is to obtain a pure background seismicity rate based only on background earthquakes that are spatiotemporally independent by completely eliminating dependent earthquakes from an observed earthquake catalog. For example, a commonly-used method in the prediction of long-term seismicity and seismic hazard research is that of Cornell [1968], which considers that the long-term seismicity rate within a particular area is constant concerning time, i.e., a stationary Poisson process is assumed [Mulgaria et al., 2017; Taroni and Atkinci, 2021]. Therefore, a pure background seismicity rate that completely removes the dependent earthquake from the observed earthquake catalog is required for seismic hazard assessment. The characteristics of dependent earthquakes vary according to region. For this reason, it can be said that different declustering methods have been proposed in various regions [van Stiphout et al., 2012].

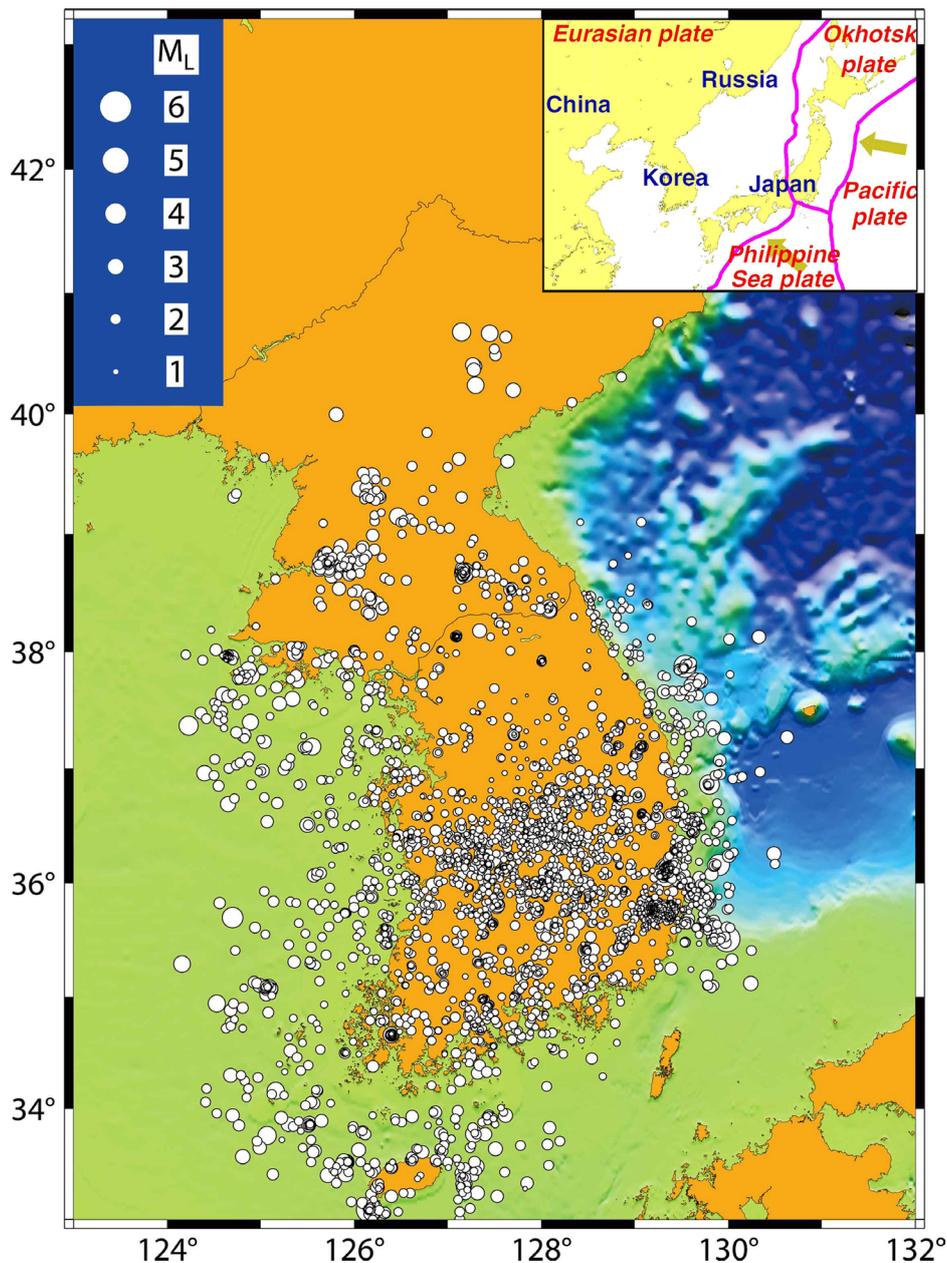


Figure 1. The epicenter distribution observed by the Korea Meteorological Administration between 2016 and 2021. The index map modified from Park et al. [2021] in the upper right corner represents the tectonic environment of the Korean Peninsula.

Comparison of earthquake catalogs for the Korean Peninsula declustered using three different methods

Since the Korean Peninsula lies on the stable intraplate region, the eastern margin of the Eurasian Plate (see the index map in Figure 1), its seismicity is much more inactive than neighboring southwest Japan, located on the plate boundary. Seismological research in Korea was launched in earnest due to the need to evaluate earthquake hazards following the construction of industrial complexes and nuclear power plants. Though instrumental earthquake observations started in 1905, digital observations did not begin until the end of the 2000s, meaning that there has been a general lack of research on earthquake hazards in this area. However, on 12 September 2016, a magnitude (M_L) 5.8 earthquake struck Gyeongju in the southeastern Korean Peninsula [Kim et al., 2016; Son et al., 2018]. About a year later, a magnitude (M_L) 5.4 earthquake occurred in Pohang, close to Gyeongju, which was thought to have been triggered by water injection into a borehole for geothermal power generation [Kim et al., 2018; Lim et al., 2020]. These two earthquakes were accompanied by many foreshocks and aftershocks [Kim and Lee, 2019], causing considerable damage to the surrounding area [Eem et al., 2018; Ghim et al., 2018; Lee et al., 2018; Jin et al., 2020]. As a result, national and social interests in earthquakes have increased, and research on seismic hazards has become more active.

The present study selected the earthquake catalog compiled by the Korea Meteorological Administration [KMA] for the Korean Peninsula between 2016 and 2021 for declustering analysis. Although six years is insufficient for declustering studies, this period is chosen because micro-earthquakes with a magnitude of 2.0 or less have been included in the catalog since 2016, along with dependent earthquakes. The epicenter distribution for the catalog is presented in Figure 1, showing that the offshore epicenters are concentrated along the southeast coast of the Korean Peninsula and from the central to the southern region of the west coast. In contrast, the inland epicenters are primarily distributed from the central western region to the southeastern region of the peninsula. The KMA catalog is declustered using three different methods. First, the resulting declustered catalogs are compared regarding their completeness magnitude (M_c) and seismicity parameters (*Gutenberg-Richter* a - and b -values). As noted above, whether a declustered earthquake catalog follows a temporal Poisson process is a long-standing area in seismology [Wyss and Toya, 2000; Shearer and Stark, 2012]. Thus, this is examined for the three declustered catalogs using several statistical test methods.

2. Seismicity declustering methods

In the present study, the seismicity declustering methods proposed by Gardner and Knopoff [1974], Reasenber [1985], and Zhuang et al. [2002] are compared. The first of these is also known as the window method and is the simplest one for identifying mainshocks and aftershocks. After an earthquake occurs, earthquakes within specific space and time windows relative to that earthquake are recognized as aftershocks. The sizes of the spatial and temporal windows are usually expressed as functions of the magnitude of the mainshock (i.e., the largest event in the cluster). Smaller earthquakes that occurred before the mainshock are regarded as foreshocks. The secondary and higher-order aftershocks (i.e., aftershocks that are triggered by or related to previous aftershocks) are not considered in the Gardner and Knopoff [1974]. Since this method assumes a circular form as a space window, the distribution of aftershocks implying the extension of the fault is not considered. While Gardner and Knopoff [1974] presented the window's spatial length and temporal duration in a table, van Stiphout et al. [2012] approximated the window as Eq. (1):

$$r = 10^{0.1238M+0.983} \text{ [km]}, \quad t = \begin{cases} 10^{0.032M+2.7389}, & \text{if } M \geq 6.5 \\ 10^{0.5409M-0.547}, & \text{else} \end{cases} \text{ (days)}, \quad (1)$$

where r is the spatial radius, t is the time, and M is the earthquake's magnitude. These equations are used for declustering in the present work.

The Reasenber method links earthquakes within the space-time interaction zones and groups them as an aftershock cluster. For example, when earthquake A is the aftershock of B, and B is the aftershock of C, earthquakes A, B, and C are grouped into a common aftershock cluster. It means that secondary and higher-order aftershocks are considered in treating aftershocks. The largest earthquake in the cluster is regarded as the mainshock of that cluster. The spatial interaction zone is defined by the size of the seismic source. The size of the temporal interaction zone is given by Omori's law for aftershock statistics. In the present work, declustering by the Reasenber

method is performed using the CLUSTER2000 computer program [United States Geological Survey, 2021], and the standard input parameters of the original Reasenberg [1985] algorithm are adopted besides the size of the spatial interaction zone.

Tibi et al. [2011] used an alternative input for the size of the spatial interaction zone based on the work by Kanamori and Anderson [1975], who represented radius a of a circular rupture area (in km) as a function of the moment magnitude (M_w) and the static stress drop ($\Delta\sigma$). Taking the stress drop to be 3 MPa, which is appropriate for most shallow earthquakes, produces the simple relationship illustrated in Eq. (2):

$$a = 10^{0.5M_w - 2.25}. \quad (2)$$

Then, the spatial interaction zone R consists of the rupture area radius of the aftershock multiplied by the factor F , plus the rupture radius of the largest earthquake in the cluster, as given by Eq. (3) [Tibi et al., 2011]:

$$R = a \times F + 10^{0.5M_w^0 - 2.25}, \quad (3)$$

where M_w^0 is the moment magnitude of the mainshock. The above alternative input of Tibi et al. [2011] is adopted in the present study. Factor F is assumed to be 10, a commonly used value.

In the methods proposed by Gardner and Knopoff [1974] and Reasenberg [1985], the constants representing the size of the space-time windows or the size of the space-time interaction zones are subjective. Therefore, the distinction between background and dependent earthquakes changes if the constant is adjusted. In response to this, Zhuang et al. [2002] proposed a method for the probabilistic determination of background and dependent earthquakes using an epidemic-type aftershock sequence (ETAS) model. Their ETAS model can be expressed using the conditional intensity function presented as Eq. (4) [Zhuang et al., 2002]:

$$\lambda(t, x, y) = \mu(x, y) + \sum_{k: t_k < t} \chi(m_k) g(t - t_k) f(x - x_k, y - y_k | m_k), \quad (4)$$

where $\mu(x, y)$ is the time-independent intensity function for a background earthquake, $g(t)$ and $f(x, y | m_k)$ are the normalized response functions concerning the epicenter and the time of occurrence, respectively. $\chi(m_k)$ represents the number of dependent earthquakes expected to occur due to earthquakes of magnitude m_k . Eq. (4) defines the probability that the j -th earthquake was triggered by the i -th earthquake as the relative contribution of the i -th earthquake to the occurrence of the j -th earthquake ($\rho_{i,j}$), which can be expressed as Eq. (5):

$$\rho_{i,j} = \frac{\chi(m_i) g(t_j - t_i) [f(x_j - x_i, y_j - y_i | m_i)]}{\lambda(t_j, x_j, y_j)}. \quad (5)$$

In the same way, the probability that an earthquake is a background earthquake (ϕ_j) or a dependent earthquake (ρ_j) can be expressed using Eqs. (6) and (7), respectively.

$$\phi_j = \frac{\mu(x_j, y_j)}{\lambda(t_j, x_j, y_j)} \quad (6)$$

$$\rho_j = 1 - \phi_j = \sum_{i=1}^{j-1} \rho_{i,j}. \quad (7)$$

Comparison of earthquake catalogs for the Korean Peninsula declustered using three different methods

In the present study, the probabilities for equations (5), (6), and (7) are estimated using the computer program of Zhuang [2017] and the following algorithm proposed by Zhuang et al. [2002] to identify background and dependent earthquakes is used:

- 1) For each pair of earthquakes $i, j = 1, 2, \dots, N (i < j)$, calculate ϕ_j in Eq. (6) and ρ_{ij} in Eq. (7).
- 2) Set $j = 1$.
- 3) Generate a uniform random number U_j in $(0, 1)$.
- 4) If $U_j < \phi_j$, then the j -th event is considered a background event.
- 5) Otherwise, select the smallest I such that $U_j < \phi_j + \sum_{i=1}^I \rho_{ij}$. Then the j -th event is considered a dependent of the i -th event.
- 6) If $j = N$, terminate the algorithm; otherwise, set $j = j + 1$ and go to step 3.

3. Results of declustering

The results of declustering according to the three methods are presented in Table 1, where the total number of events in the raw earthquake catalog is 7952. The Zhuang method removed the most significant number of dependent earthquakes and followed by the Gardner-Knopoff method. The Reasenber method, which considers the links between dependent earthquakes, detects relatively few. In the case of the Reasenber method, it is characteristic that the number of clusters of dependent earthquakes is relatively small compared to the number of earthquakes removed. It means that the Reasenber method has a more significant number of earthquakes per cluster on average than other methods. All earthquakes, including the raw earthquake catalog's dependent and background earthquakes, are plotted in Figure 2a (white circles). In contrast, the mainshocks (white circles) and dependent earthquakes (red dots) detected using the Gardner-Knopoff, Reasenber, and Zhuang methods are plotted in Figure 2b, c, and d, respectively. The horizontal and vertical axes represent time (year) and latitude, respectively. The dependent earthquakes following the Gyeongju earthquake in September 2016 and the Pohang earthquake in November 2017 (denoted by the yellow star and triangle, respectively) are well-separated on the time axis, resembling a dotted line. The largest number of dependent earthquakes is detected and removed using the Zhuang method (Figure 2d, Table 1), followed by the Gardner-Knopoff method (Figure 2b, Table 1). The dependent earthquakes in Figure 2b and d have almost the same spatial range. In contrast, the temporal extent of Fig. 2d is broader than that of Figure 2b. The Reasenber method detects the lowest number of dependent earthquakes in both the spatial and temporal range (Figure 2c, Table 1).

Method	No. of events	No. of clusters	Removed events (%)	Remaining events
Gardner and Knopoff [1974]	7952	463	4557 (57.2)	3395
Reasenber [1985]	7952	187	4144 (52.1)	3808
Zhuang et al. [2002]	7952	454	5606 (70.5)	2346

Table 1. The numbers of events, clusters, removed, and remaining events obtained using the declustering methods of Gardner and Knopoff [1974], Reasenber [1985], and Zhuang et al. [2002].

A well-declustered earthquake catalog is required to predict future seismicity or evaluate seismic hazards, from which the seismicity parameters a and b and the occurrence rate above a certain magnitude can be estimated using the Gutenberg-Richter model for the frequency distribution of earthquake magnitudes. The quantity and quality of the earthquake data collected within a certain period and region usually depend on the capability of the seismic observation network. In particular, denser observation networks with high-performance instruments can detect

even small-scale events; thus, the number of detected events increases. In an earthquake catalog, a completeness magnitude (M_c) is usually defined as the minimum magnitude above which all earthquakes within a specific region and period are reliably recorded [Rydelek and Sacks, 1989; Wiemer and Wyss, 2000]. A correct estimate of M_c is

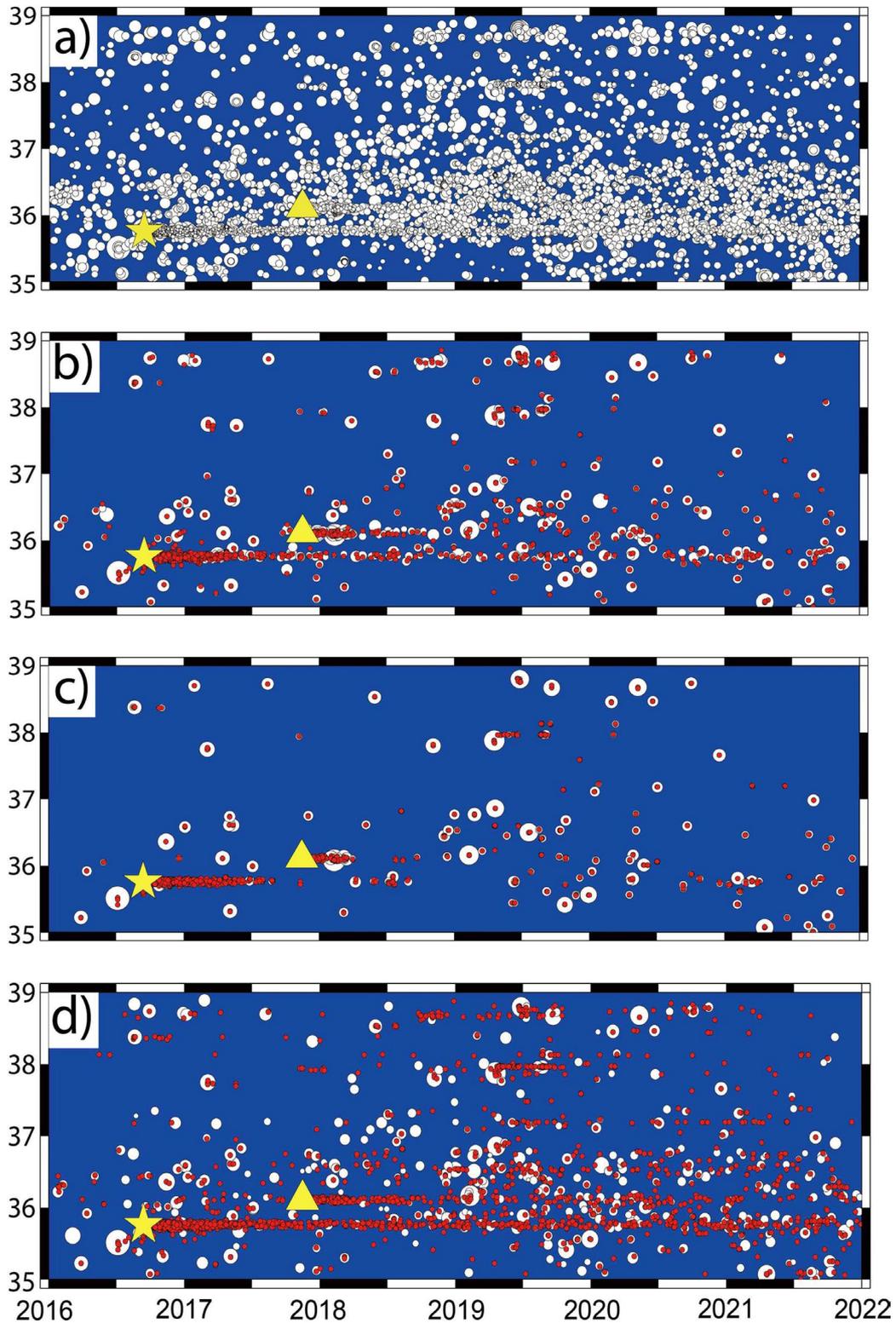


Figure 2. The time (year) – latitude plots for the mainshocks (white circles) and clustered earthquakes (red dots): (a) the raw catalog, and (b-d) the declustered catalogs using the methods of (b) Gardner and Knopoff [1974], (c) Reasenberg [1985], and (d) Zhuang et al. [2002]. In each case, the yellow star and triangle symbols represent the Gyeongju earthquake in September 2016 and the Pohang earthquake in November 2017, respectively.

Comparison of earthquake catalogs for the Korean Peninsula declustered using three different methods

crucial because a value that is too high leads to undersampling, while a value that is too low leads to erroneous values for the seismicity parameters and subsequently biased analyses [Mignan and Woessner, 2012]. Consequently, the determination of M_c is the starting point for estimating the seismicity parameters from an earthquake catalog. While many approaches for estimating M_c have been proposed (e.g., Ogata and Katsura, 1993; Woessner and Wiemer, 2005), the widely-used method introduced by Wiemer and Wyss [2000] is used in the present study.

Once the M_c has been determined, the events with magnitudes equal to or greater than M_c are selected from the earthquake catalog to estimate the seismicity parameters. The seismicity parameter b -value can be estimated from the Gutenberg-Richter curve using either the least-squares or maximum-likelihood methods. However, the maximum-likelihood approach is more reasonable because the cumulative event numbers are not independent, and the number of earthquakes over time is better represented by a Poisson distribution than a Gaussian distribution [Weichert, 1980]. Several methods for estimating the b -value based on the maximum likelihood method have been proposed [Aki, 1965; Page, 1968; Weichert, 1980]. It is well known that the b -value is greatly affected by the method used to derive it [Bender, 1983; Amorese et al., 2009]. Through simulations, Lee [2013] demonstrated that the Aki-Utsu formula [Aki, 1965; Utsu, 1965] generates a more stable b -value than other methods. Thus, the Aki-Utsu formula is used in the present study to estimate the b -value. In addition, instead of the a -value, the annual occurrence rate of earthquakes with magnitudes equal to or greater than 5.0 is used, following most seismic hazard studies.

The estimated b -values, their standard deviation, and the annual occurrence rate for magnitude 5.0 or larger earthquakes for the raw catalog and the three declustered catalogs are presented in Table 2. M_c is estimated to be 1.5 for the raw catalog, decreasing to 1.4 for each of the three declustered catalogs. Similarly, a b -value of 0.97 is obtained from the raw catalog, which decreases after removing the dependent earthquakes. It is thought to be because many small-scale earthquakes are removed from the raw catalog. However, unlike the M_c values, the b -values of the declustered catalogs vary according to the method, being 0.91, 0.93, and 0.91 for the Gardner-Knopoff, Reasenber, and Zhuang methods, respectively. Similarly, the annual occurrence rate of magnitude 5.0 or larger earthquakes also varies between the declustering methods, with estimates of 0.2708 for the raw catalog and 0.2188, 0.2209, and 0.2122 for the catalogs by the Gardner-Knopoff, Reasenber, and Zhuang methods, respectively.

Catalog	M_c	$b \pm SD$	M5.0 / year
Raw	1.5	0.97 ± 0.003	0.2708
Gardner and Knopoff [1974]	1.4	0.91 ± 0.004	0.2188
Reasenber [1985]	1.4	0.93 ± 0.004	0.2209
Zhuang et al. [2002]	1.4	0.91 ± 0.006	0.2122

Table 2. The completeness magnitude (M_c), b -value and its standard deviation (SD), and annual occurrence rate of $M \geq 5.0$ for the raw catalog and declustered catalogs using the methods of Gardner and Knopoff [1974], Reasenber [1985], and Zhuang et al. [2002].

4. Poisson process tests

For long-term predictions of seismicity or seismic hazard research, it is assumed that seismicity follows the Poisson process. Therefore, it is necessary to verify this assumption for declustered earthquake catalogs. Since Gardner and Knopoff [1974] attempted the Poisson test for the declustered catalog in California, many researchers have conducted such studies in various areas [Wyss and Toya, 2000; Luen and Stark, 2012; Shearer and Stark, 2012]. In particular, Noh [2016] performed a Poisson test on earthquakes observed in and around the Korean Peninsula between 1980 and 2014, finding that earthquakes with magnitudes greater than or equal to 2.7 that occurred on land followed the Poisson process.

In the present study, three Poisson tests are applied to the three declustered earthquake catalogs: (i) multinomial chi-square (MC), (ii) conditional chi-square (CC), and (iii) Kolmogorov-Smirnov (KS) tests. The MC and CC tests focus on the distribution of earthquakes within a specific period, while the KS test evaluates the distribution of the time intervals between events. The method of Shearer and Stark [2012] described as follows is adopted for the MC test. First, the observation period is divided into several equal intervals (N_i), and the number of events occurring within each interval is counted. The number of intervals with a certain number of events is then assessed to determine whether this is consistent with the theoretical Poisson distribution. Assuming that seismicity follows the Poisson process, the average event rate (λ) per interval is N/N_i , where the total number of events is N . When the expected number of intervals is at least five, the smallest and largest intervals (K^- and K^+) can be expressed as Eqs. (8) and (9) [Shearer and Stark, 2012]:

$$K^- \equiv \min \left(k: N_i e^{-\lambda} \sum_{j=0}^k \lambda^j / j! \geq 5 \right), \quad (8)$$

$$K^+ \equiv \max \left[k: N_i \left(1 - e^{-\lambda} \sum_{j=0}^{k-1} \lambda^j / j! \geq 5 \right) \right]. \quad (9)$$

The test statistic χ_M^2 is obtained with $(K^+ - K^- + 1) - 2$ degrees of freedom using Eq. (10):

$$\chi_M^2 = \sum_{k=K^-}^{K^+} (X_k - E_k)^2 / E_k, \quad (10)$$

where X_k denotes the number of intervals, including k events, and E_k is defined as Eq. (11):

$$E_k \equiv \begin{cases} N_i e^{-\lambda} \sum_{j=0}^{k^-} \lambda^j / j!, & k = K^-, \\ N_i e^{-\lambda} \lambda^k / k!, & k = K^- + 1 \dots K^+ - 1, \\ N_i \left(1 - e^{-\lambda} \sum_{j=0}^{K^+ - 1} \lambda^j / j! \right), & k = K^+. \end{cases} \quad (11)$$

Because the MC test does not distinguish between the number of intervals with small and large numbers of events, it is not sensitive to variance in the data. For this reason, the MC test is not as sensitive as other tests to fluctuations in the seismicity rate [Luen and Stark, 2012]. This problem can be overcome using the CC test, also known as the Poisson dispersion test. This test divides the entire observation period into several equal intervals and finds the variance between the number of events within each interval along with the average number of events. When the number of events during the k -th observation period is N_k , the test statistic χ_c^2 can be expressed as Eq. (12):

$$\chi_c^2 = \sum_{k=1}^M \frac{(N_k - \lambda)^2}{\lambda}. \quad (12)$$

Comparison of earthquake catalogs for the Korean Peninsula declustered using three different methods

The KS test assesses the distribution of the intervals between earthquake occurrences. For example, for an earthquake catalog with $(n + 1)$ events, the intervals between earthquake occurrences are organized in ascending order as (x_1, x_2, \dots, x_n) , and their average is defined as \bar{x} . The time interval distribution function $F(x_i)$ is then defined using Eq. (13):

$$F(x_i) = z_i = 1 - \exp(x_i/\bar{x}), \quad i = 1, 2, \dots, n \quad (13)$$

The statistic for the KS test is then expressed using Eq. (14) [Stephens, 1974]:

$$D = \max(D^+, D^-);$$

$$D^+ = \max(i/n - z_i), \quad D^- = \max[z_i - (i - 1)/n], \quad 1 \leq i \leq n. \quad (14)$$

To conduct a hypothesis test using the table provided by Stephens [1974], Eq. (14) must be corrected to give Eq. (15):

$$(D - 0.2/n)(\sqrt{n} + 0.28 + 0.5/\sqrt{n}) \rightarrow D. \quad (15)$$

The null hypothesis is then used to test whether the earthquake catalogs that are declustered using the three methods follow the Poisson process. For each catalog, each Poisson test (MC, CC, and KS) is conducted with a significance level of 0.05 for increasing values of earthquake magnitude from 1.0 to 3.0 in steps of 0.1. Generally, the null hypothesis is rejected at a small magnitude but not above a certain limiting magnitude. Noh [2016] defined the minimum magnitude for which the null hypothesis is not rejected as M_p , and this term is also adopted in the present study. The MC and CC tests are conducted for 72 and 216 intervals, obtained by dividing the 2016-2021 observation period into intervals of 30 and 10 days, respectively. The results for the Poisson tests are shown in Table 3. The M_p values range from 1.6 to 2.2 depending on the declustered catalog and the test method. The difference in the M_p according to the number of intervals is also insignificant. The M_p values obtained from the Poisson tests are also larger than the M_c values (see Table 2), with the largest M_p obtained for the catalog declustered using Reasenber method, which removes the smallest number of dependent earthquakes. In addition, the CC test produces a larger M_p for the catalogs obtained using the Gardner-Knopoff and Reasenber methods than the other tests. Thus, careful consideration must be paid to selecting the cutoff magnitude for earthquake catalogs in seismic hazard assessment. In this respect, the M_p value provides a helpful reference with M_c .

Catalog	Test method		
	Multinomial chi-square (MC)	Conditional chi-square (CC)	Kolmogorov-Smirnov (KS)
Gardner and Knopoff [1974]	1.7	1.9	1.6
Reasenber [1985]	1.8	2.2	1.9
Zhuang et al. [2002]	1.6	1.6	1.6

Table 3. The M_p estimated from the three types of Poisson hypothesis test (multinomial chi-square, conditional chi-square, and Kolmogorov-Smirnov) for the declustered catalogs using the methods of Gardner and Knopoff [1974], Reasenber [1985], and Zhuang et al. [2002].

5. Discussion

In the present study, seismicity parameters and Poisson tests are used to compare three methods for seismicity declustering. However, it is difficult to determine which method is the most effective because there is no inherently unique method for removing dependent earthquakes, and the evaluation criteria for the removal results are not absolute. In the ETAS model [Ogata, 1988; Utsu et al., 1995], which generalizes Omori's law for aftershock occurrence, the seismicity rate observed in an area is expressed as the sum of the constant background seismicity rate over time and the aftershock activity rate triggered by other earthquakes [Ogata, 1988]. Therefore, if dependent earthquakes are removed using a seismicity declustering method, the cumulative background seismicity rate to time should be linear.

The cumulative numbers of earthquakes obtained from the raw catalog and the three declustered earthquake catalogs are plotted against time in Figure 3. Here, the cumulative number of earthquakes includes only earthquakes with an M_c greater than or equal to 1.5. Including earthquakes smaller than M_c can distort the seismicity in space and time. Notably, the cumulative number of earthquakes with $M_c \geq 1.5$ from the raw catalog (the solid black line in Figure 3) increased significantly after the Gyeongju earthquake (G.E.; $M_L = 5.8$) on 12 September 2016 and the Pohang earthquake (P.E.; $M_L = 5.4$) on 15 November 2017. However, as noted above, if the dependent earthquakes are completely removed from a catalog, only the background earthquakes with a constant seismicity rate will remain; hence, the cumulative number of earthquakes will appear as a straight line with a constant slope. Therefore, the optimal declustering method can be identified based on the curve closest to being a straight line with a constant slope. By identifying the straight line that best fits each of the curves in Figure 3 and calculating the slope using the least-square method, values of 314.7, 343.4, and 236.5 are obtained for the catalogs declustered using the Gardner-Knopoff, Reasenberg, and Zhuang methods, respectively. These values represent the annual mean occurrence numbers of earthquakes with magnitudes greater than or equal to 1.5. The Reasenberg method produces the highest slope (the dot-dashed line b of Figure 3), possibly due to fewer dependent earthquakes being removed than the other two declustering methods.

As shown in Table 1, the number of removed dependent earthquakes increases in the order of the Reasenberg, Gardner-Knopoff, and Zhuang methods. Curves in Figure 3 for the catalogs declustered using the Reasenberg and Gardner-Knopoff methods (the solid red line) look like a line bent slightly upward rather than a straight line. In contrast, the curve by the Zhuang method (the blue dotted line d of Figure 3) appears to be the closest to a straight line. However, based on this evidence alone, it is difficult to confidently conclude that the Zhuang method is optimal

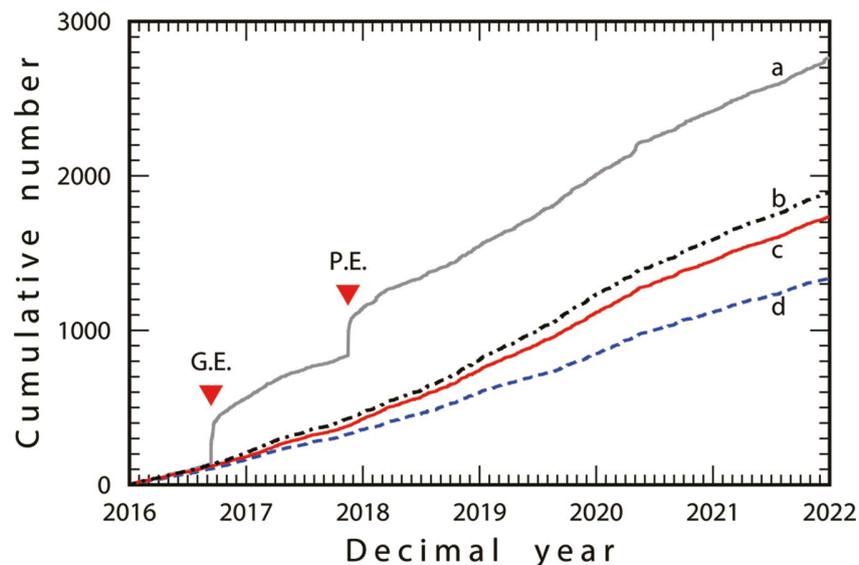


Figure 3. The cumulative numbers of earthquakes between 2016 and 2021 from (a) the raw catalog and (b-d) the declustered catalogs using the methods of (b) Reasenberg [1985], (c) Gardner and Knopoff [1974], and (d) Zhuang et al. [2002]. The symbols G.E. and P.E. denote the occurrence times of the 2016 Gyeongju earthquake ($M_L = 5.8$) and the 2017 Pohang earthquake ($M_L = 5.4$), respectively.

Comparison of earthquake catalogs for the Korean Peninsula declustered using three different methods

because the background seismicity rate can only be regarded as constant over a sufficiently long period. However, the study period for the present research is only six years.

In Figure 4, the Gardner-Knopoff method for removing dependent earthquakes within fixed space-time windows is compared with the Zhuang method. The black dotted line represents the upper limit of the spatial window for the Gardner-Knopoff method. In contrast, the solid red line represents the upper limit of the spatial window for the method proposed by Uhrhammer [1986]. The crosses indicate the dependent earthquakes removed probabilistically using the Zhuang method. The spatial windows for Gardner and Knopoff [1974] and Uhrhammer [1986] are a function of the magnitude, with the spatial window becoming wide as the magnitude increases. Here, the size of the spatial window of Zhuang et al. [2002] appears similar to that of Gardner and Knopoff [1974]. Figure 5 compares the range of time windows between the Gardner and Knopoff [1974] and Uhrhammer [1986] methods for removing dependent

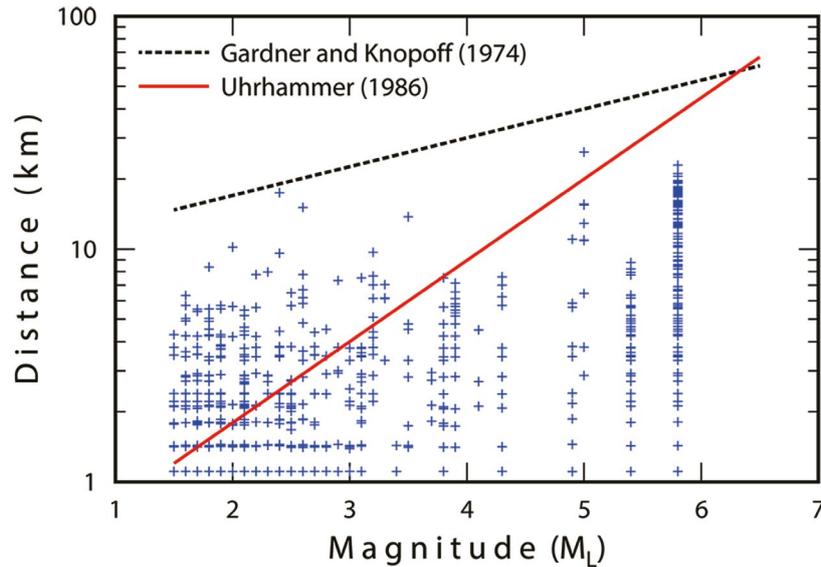


Figure 4. The ranges of the spatial windows for removing dependent earthquakes by the methods of Gardner and Knopoff [1974] and Uhrhammer [1986]. The cross marks represent the dependent earthquakes removed probabilistically using the method of Zhuang et al. [2002].

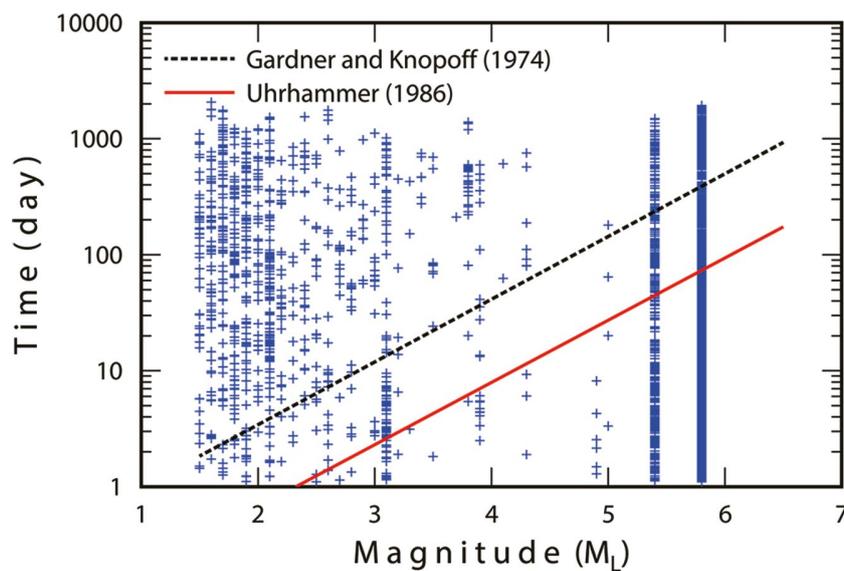


Figure 5. The ranges of the time windows for removing the dependent earthquakes by the methods of Gardner and Knopoff [1974] and Uhrhammer [1986]. The cross marks represent the dependent earthquakes removed probabilistically using the method of Zhuang et al. [2002].

earthquakes with those that are removed probabilistically using the Zhuang method. As with the spatial windows, the time windows for the methods of Gardner and Knopoff [1974] and Uhrhammer [1986] are a function of the magnitude, with the time window becoming wide as the magnitude increases. However, the Zhuang method shows an unlimited time window regardless of the earthquake magnitude, allowing it to remove many more dependent earthquakes.

Nevertheless, no inherently unique method for removing dependent earthquakes exists, as noted above. It is difficult to determine which method is the most effective because of the subjective evaluation criteria for the removal results. It is thought that the range of the space-time window for dependent earthquakes is associated with the extent of the fault motion depending on the size of the seismic source and the physical properties of the subsurface. Therefore, the range of the space-time window may vary from region to region. In addition, the range can vary according to the specific definitions of foreshocks and aftershocks. For example, whether or not earthquakes triggered by the reactivation of an old fault should be regarded as aftershocks remains unresolved.

6. Conclusions

In the present study, the seismicity declustering methods proposed by Gardner and Knopoff [1974], Reasenbergl [1985], and Zhuang et al. [2002] were compared for a catalog of earthquakes for 2016-2021 in and around the Korean Peninsula. It was found that the Zhuang method detected the largest number of dependent earthquakes, accounting for up to 70.5% of all earthquake events. In contrast, the Reasenbergl approach detected the lowest number of dependent earthquakes. Notably, while the Zhuang and Gardner-Knopoff methods exhibited similar spatial windows, the Zhuang method had an unrestricted time window independent of the earthquake magnitude. The M_c of the raw earthquake catalog was 1.5, which fell to 1.4 after removing the dependent earthquakes. The b -value of the raw catalog was estimated to be 0.97, decreasing to 0.91 at the maximum for the declustered catalogs. In addition, the annual occurrence rate of earthquakes with magnitudes greater than or equal to 5.0 differed between the declustering methods. These observations demonstrate that long-term earthquake prediction or seismic hazard assessment results can vary depending on the declustering method.

Null hypothesis tests were performed using MC, CC, and KS tests to determine whether the declustered catalogs followed the Poisson process. The M_p above which the null hypothesis of the Poisson process could not be rejected ranged between 1.6 and 2.2 depending on the declustered catalog and the test method. For the MC test, the M_p ranged from 1.6 to 1.8, and there was no significant difference between the three declustered catalogs. The KS test also produced a similar M_p range to the MC test. However, the CC test showed significant variations in the M_p values from 1.6 to 2.2 depending on the declustered catalog. The M_p values obtained from the Poisson tests were generally larger than M_c ; thus, M_p represents a valuable reference for selecting the cutoff magnitude.

Only the background earthquakes with a constant seismicity rate remain when the dependent earthquakes are entirely removed from a catalog. Consequently, the cumulative number of earthquakes over time will appear as a straight line. Although this was confirmed by the plot of the cumulative number of earthquakes over time for the Zhuang method, it cannot be guaranteed that this is the optimal method based only on this evidence. However, while the Zhuang and Gardner-Knopoff methods employ similar spatial windows, the former utilizes an unlimited time window independent of the earthquake magnitude. It indicates that the Zhuang approach can remove much more dependent earthquakes than the other two methods. It is important to note that there is no consensus method for removing dependent earthquakes, and the criteria used to evaluate the removal results remain subjective.

Acknowledgements. The author thanks editor Dr. Andrea Bizzarri and the anonymous reviewers for their detailed and constructive comments.

References

- Aki, K. (1965). Maximum likelihood estimate of b in the formula $\log_{10}N=a-bm$ and its confidence limits, Bull. Earthquake Res. Inst., Tokyo University, 43, 237-239.
- Amorese, D., J.R. Grasso and P.A. Rydelek (2009). On varying b -values with depth: Result from computer-intensive tests for southern California, Geophys. J. Int., 180, 347-360.

Comparison of earthquake catalogs for the Korean Peninsula declustered using three different methods

- Bender, B. (1983). Maximum likelihood estimation of b values for magnitude grouped data, *Bull. Seism. Soc. Am.*, 73, 831-851.
- Cornell, C.A. (1968). Engineering seismic risk analysis, *Bull. Seism. Soc. Am.*, 58, 1583-1606.
- Eem, S.-H., B. Yang and H. Jeon (2018). Earthquake damage assessment of buildings using open data in the Pohang and the Gyeongju Earthquakes, *Journal of the Earthquake Engineering Society of Korea*, 22, 121-128 (in Korean with English Abstract).
- Gardner, J.K., and L. Knopoff (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?. *Bull. Seism. Soc. Am.*, 64, 1363-1367.
- Gihm, Y.S., S.W. Kim, K. Ko, J.-H. Choi, H. Bae, P.S. Hong, Y. Lee, H. Lee, K. Jin, S.-J. Choi, J.C. Kim, M.S. Choi and S.R. Lee (2018). Paleoseismological implications of liquefaction-induced structures caused by the 2017 Pohang earthquake, *Geosci. J.*, 22, 871-880.
- Kanamori, H. and D.L. Anderson (1975). Theoretical basis of some empirical relations in seismology, *Bull. Seism. Soc. Am.*, 65, 1073-1095.
- Kim, K.H., T.S. Kang, J. Rhie, Y.H. Kim, Y. Park, S.U. Kang, M. Han, J. Kim, J. Park, M. Kim, C.H. Kong, D. Heo, H. Lee, E. Park, H. Park, S.J. Lee, S. Cho, J.U. Woo, S.H. Lee and J. Kim (2016). The 12 September Gyeongju earthquakes: 2. Temporary seismic network for monitoring aftershocks, *Geosci. J.*, 20, 753-757.
- Kim, K.H., J.-H. Rhie, Y.H. Kim, S. Kim, S.U. Kang and W. Seo (2018). Assessing whether the 2017 Mw 5.4 Pohang earthquake in South Korea was an induced event, *Science*, 360, 1007-1009.
- Kim, S.K. and J.M. Lee (2019). Comparison of the aftershock activities of the 2016 M5.8 Gyeongju and 2017 M5.4 Pohang earthquakes. *Korea, Journal of the Geological Society of Korea*, 55, 207-218 (in Korean with English Abstract).
- Korea Meteorological Administration (2022). <https://data.kma.go.kr> (last access: 21 January 2022).
- Jin, K., J. Lee, K.-S. Lee, J.B. Kyung and Y.S. Kim (2020). Earthquake damage and related factors associated with the 2016 ML = 5.8 Gyeongju earthquake, southeast Korea, *Geosci. J.*, 24, 141-157.
- Lee, C.-H., S.-Y. Kim, J.-H. Park, G.-K. Kim and T.-J. Kim (2018). Comparative analysis of structural damage potentials observed in the 9.12 Gyeongju and 11.15 Pohang Earthquakes. *Journal of the Earthquake Engineering Society of Korea*, 22, 175-184 (in Korean with English Abstract).
- Lee, H.-K. (2013). Estimation of Gutenberg-Richter b-value and Mmax using instrumental earthquake catalog from the southern Korean Peninsula. MS thesis, Chonnam National University, 71, (in Korean with English abstract).
- Lim, H., K. Deng, Y.H. Kim, J.-H. Ree, T.-R.A. Song and H. Kim (2020). The 2017 Mw 5.5 Pohang Earthquake, South Korea, and poroelastic stress changes associated with fluid injection, *J. Geophys. Res.: Solid Earth*, 125, <https://doi.org/10.1029/2019JB019134>.
- Luen, B. and P.B. Stark (2012). Poisson tests of declustered catalogues, *Geophys. J. Int.*, 189, 691-700.
- Mignan, A. and J. Woessner (2012). Estimating the magnitude of completeness for earthquake catalogs, *Community Online Resource for Statistical Seismicity Analysis*, doi:10.5078/corssa-00180805.
- Mulargia, F., P.B. Stark and R.J. Geller (2017). Why is Probabilistic Seismic Hazard Analysis (PSHA) still used ?, *Phys. Earth Planet. Int.*, 264, 63-75.
- Noh, M. (2016). On the Poisson process of the Korean earthquakes, *Geosci. J.*, 20, 775-779.
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Ass.*, 83, 9-27.
- Ogata, Y. and K. Katsura (1993). Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues, *Geophys. J. Int.*, 113, 727-738.
- Page, R. (1968). Aftershock and microaftershocks of the Great Alaska earthquake of 1964, *Bull. Seism. Soc. Am.*, 58, 1131-1168.
- Park, S., T.-K. Hong and G. Ra (2021). Seismic hazard analysis for the Korean Peninsula. *Bulletin of the Seism. Soc. Am.*, 111, 2696-2719.
- Reasenberg, P. (1985). Second-order moment of central California seismicity, 1969-82, *J. Geophys. Res.: Solid Earth*, 90, 5479-5495.
- Rydelek, P.A. and I.S. Sacks (1989). Testing the completeness of earthquake catalogues and the hypothesis of self-similarity, *Nature*, 337, 6204, 251-253.
- Shearer, P.M. and P.B. Stark (2012). Global risk of big earthquakes has not recently increased, *Earth, Atmosph., Planet. Sci.*, 109, 717-721
- Son, M., C.S. Cho, J.S. Shin, H.M. Rhee and D.H. Sheen (2018). Spatio-temporal distribution of events during the first three months of the 2016 Gyeongju, Korea, earthquake sequence, *Bull. Seism. Soc. Am.*, 108, 210-217.

Sung Kyun Kim

- Stephens, M.A. (1974). EDF statistics for goodness of fit and some comparisons, *J. Am. Stat. Ass.*, 69, 730-737.
- Taroni, M. and A. Akinci (2021). Good practices in PSHA: declustering, b-value estimation, foreshocks and aftershocks inclusion; a case study in Italy, *Geophys. J. Int.*, 224, 1174-1187.
- Tibi, R., J. Blanco and A. Fatehi (2011). An alternative and efficient cluster-link approach for declustering of earthquake catalogs, *Seism. Res. Lett.*, 82, 509-518.
- United States Geological Survey (2021). <https://www.usgs.gov/software/cluster2000> (last access: 30 January 2022).
- Utsu, T. (1965). A method for determining the value of b in a formula $\log n = a - bM$ showing the magnitude-frequency relation for earthquakes, *Geophys. Bull. Hokkaido Univ.*, 13, 99-103.
- Utsu, T., Y. Ogata and R.S. Matsu'ura (1995). The Centenary of the Omori formula for a decay law of aftershock activity, *J. Phys. Earth*, 43, 1-33.
- Uhrhammer, R. (1986). Characteristics of northern and central California seismicity, *Earthquake Notes*, 57, 21-21.
- van Stiphout, T., J. Zhuang and D. Marsan, (2012). Seismicity declustering, Community Online Resource for Statistical Seismicity Analysis, doi:10.5078/Corssa-52382934.
- Weichert, D.H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes, *Bull. Seism. Soc. Am.*, 70, 1337-1346.
- Wiemer, S. and M. Wyss (2000). Minimum magnitude of complete reporting in earthquake catalogs: Examples from Alaska, the western United States, and Japan, *Bull. Seism. Soc. Am.*, 90, 859-869.
- Woessner, J. and S. Wiemer (2005). Assessing the quality of earthquake catalogs: Estimating the magnitude of completeness and its uncertainty, *Bull. Seism. Soc. Am.*, 95, 684-698.
- Wyss, M. and Y. Toya (2000). Is background seismicity produced at a stationary Poisson rate ?, *Bull. Seism. Soc. Am.*, 90, 1174-1187.
- Zhuang, J., Y. Ogata and D. Vere-Jones (2002). Stochastic declustering of space-time earthquake occurrences, *J. Am. Stat. Ass.*, 97, 369-380.
- Zhuang, J. (2017). <http://bemlar.ism.ac.jp/zhuang/software.html> (last access: 30 January 2022).

CORRESPONDING AUTHOR: Sung KYUN KIM,

Chonnam National University, 77 Yongbong-ro, Buk-gu, Gwangju 61186, Korea
e-mail: kimsk3454@gmail.com