Investigating the Distributions of Differences between Mainshock and Foreshock Magnitudes

by Christine Smyth, Jim Mori, and Masumi Yamada

Abstract   Previous research produces seemingly contradictory statements about the distribution of differences between mainshock and foreshock magnitudes. Specifically, some authors find that the magnitude difference between the mainshock and the foreshock is equally likely to be large as to be small. However, other authors find that the distribution of the magnitude differences between the foreshock–mainshock pairs is not uniform. We consider foreshock–mainshock pairs within the recent Japanese earthquake catalog and worldwide data to explore the discrepancies between these seemingly contradictory studies. The results of the previous studies differ because of the different foreshock–mainshock earthquake pairs that are considered by the two sets of authors. We show that using the definitions employed by either type of study permits the found distributions to be derived analytically and further explains how the results are dependent upon the assumed definition of foreshocks and the data selection.

Introduction

It is currently impossible to determine whether an arbitrary earthquake is a foreshock to a larger event in a real-time setting. However, it is possible to search historical earthquake catalogs of foreshock–mainshock pairs to determine whether particular patterns occur. Previous research produces seemingly contradictory statements about the distribution of differences between mainshock and foreshock magnitudes.

The results of Agnew and Jones (1991), Michael and Jones (1998), and Reasenberg (1999) show that the magnitude difference between the mainshock and the foreshock is equally likely to be large as to be small. Although these authors consider different datasets, they all produce results that suggest a uniform distribution of the difference in magnitude between mainshocks and foreshocks. We represent this distribution schematically as the dotted line in Figure 1a, which shows the relative number of pairs at each magnitude difference. We convert this schematic count to a cumulative proportion, shown as the dotted line in Figure 1b. This distribution is incorporated into a model that calculates the probability of a major earthquake characteristic to a particular fault segment following the occurrence of an arbitrary earthquake near the fault (Agnew and Jones, 1991).

However, other authors using different datasets find that the magnitude difference between a mainshock and a foreshock is not uniformly distributed but instead is more likely to be small than large (Jones, 1985; Savage and DePolo, 1993; Tormann et al., 2008). This distribution is also shown schematically as a relative count by the solid line in Figure 1a and as a cumulative proportion by the solid line in Figure 1b.

This result seemingly contradicts the previous results that the magnitude difference between the foreshock and the mainshock is equally likely to be large as to be small.

The distribution of differences between foreshock and mainshock magnitudes is important for understanding foreshock occurrence and estimating foreshock probabilities. A uniform distribution (dotted lines in Fig. 1) implies that there should be many more small magnitude foreshocks in seismicity catalogs. We hypothesize that the two contradictory lines may simply reflect dissimilarities in study design between the two sets of authors. Agnew and Jones (1991) suggest that the Jones (1985) result is a consequence of using an equal magnitude threshold for foreshocks and mainshocks. In this research we critically evaluate this suggestion and other dissimilarities we find between the two study designs to explain the contradictory results.

Previous and Present Studies

The research by Reasenberg (1999) clearly shows that the distribution of differences between foreshock and mainshock magnitudes is uniform (the equivalent of the dotted lines in Fig. 1). In contrast, the results of Jones (1985) show the same distribution to be the equivalent of the solid lines in Figure 1. We choose to compare these two studies because they both explicitly produce graphs that show the empirical cumulative proportion of differences between mainshock and foreshock magnitudes. We first outline the approaches taken by the two studies. We explain how each study is conducted in the context of finding the distribution of differences
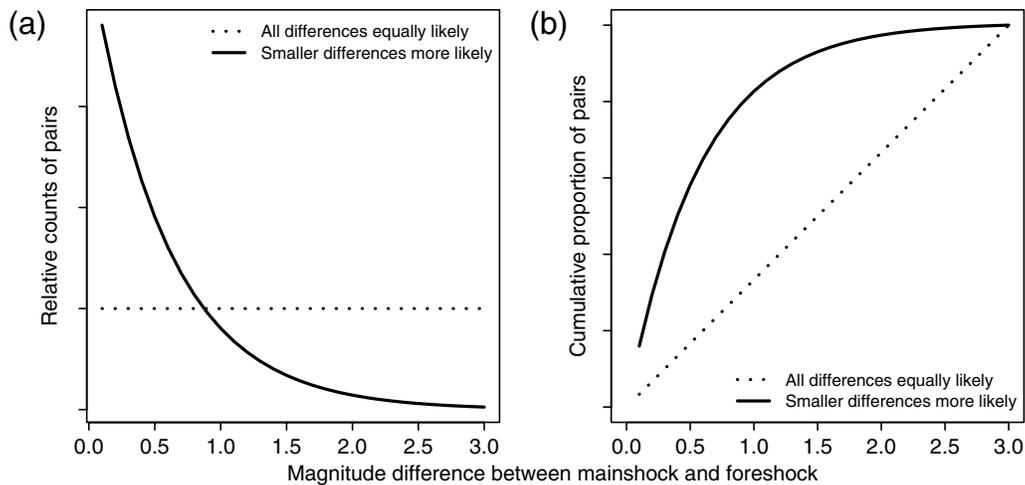
Figure 1. (a) The dotted line shows the relative counts of pairs as a function of magnitude difference between the mainshock and the foreshock, assuming all differences are equally likely. The solid line shows the relative counts of pairs as a function of magnitude difference between the mainshock and the foreshock, assuming smaller differences are more likely. (b) The dotted line shows the cumulative proportion of pairs as a function of magnitude difference between the mainshock and the foreshock, assuming all differences are equally likely. The solid line shows the cumulative proportion of pairs as a function of magnitude difference between the mainshock and the foreshock, assuming smaller differences are more likely.

between the mainshock and the foreshock magnitudes. We try to keep notation as close as possible to the original. Then we explain how we design our study to investigate which difference between the historical approaches is driving their contrasting results.

The Reasenberg (1999) approach uses the Harvard catalog of Centroid Moment Tensor (CMT) solutions (now the Global CMT Project catalog; see Data and Resources) for large earthquakes (Dziewonski *et al.*, 1994), with a cutoff magnitude M_c of five and depth shallower than 50 km. Specific details are given in the original reference. The foreshock and mainshock pairs are then found according to the following five rules:

1. The magnitude of the foreshock M^* and the magnitude of the mainshock M^{main} must be greater than M^*_{min} and $M^{\mathrm{main}}_{\mathrm{min}}$, respectively.
2. The mainshock must be larger than the foreshock.
3. The epicentral distance between the foreshock and the mainshock, dX, is less than 75 km, and the interevent time dT is less than 10 days.
4. If two mainshocks are clustered within these dX and dT limits, only the larger mainshock is considered.
5. The foreshock must be within dI units of the mainshock; that is, $M^* \geq M^{\mathrm{main}} - dI$.

The difference between the magnitudes of each admissible foreshock–mainshock pair is then calculated. Reasenberg (1999) tests various values of $M^{\mathrm{main}}_{\mathrm{min}}$ and dI and finds the results are consistent across these parameters.

The Jones (1985) approach uses the southern California catalog compiled by the California Institute of Technology (Hileman *et al.*, 1973). All events with a magnitude equal to or greater than three are used. The data are declustered with a

windowing algorithm, and the specific parameters are described in the original paper. Then, each earthquake followed by a larger earthquake within $dT = 5$ days and $dX = 10$ kilometers is considered a foreshock. The secondary earthquake is considered a mainshock. No magnitude thresholds are placed upon the foreshocks or the mainshocks. Similarly, there is no requirement that the foreshock has to be within dI units of the mainshock. The procedure allows for mainshocks to be preceded by more than one foreshock. The difference between the magnitudes of each resulting admissible foreshock–mainshock pair is calculated.

Table 1 shows the main differences between the studies. We believe that at least one of these variables is causing the results of the two studies to disagree. We investigate these variables to see which is causing the contradictory results. Although it is highly unlikely that the different data are forcing the contradictory results, we consider the effect of different data by examining two contrasting datasets.

The first dataset is taken from the Japanese Meteorological Agency (JMA) catalog (see Data and Resources). We use all earthquakes within the area of the Japanese islands ($22.5°–47.0°$ N and $122.0°–148.0°$ E). We use recent data from 1 January 2000 to 31 December 2008 (inclusive). To obtain a uniform dataset of shallow earthquakes, we use events with depth less than 50 km and magnitude (M_{JMA}) greater than or equal to 3.0; see the work by Nanjo *et al.* (2010) for more information about magnitude of completeness values for Japan.

The second dataset contains data from the Global CMT Project catalog (Dziewonski *et al.*, 1994; see Data and Resources). We use earthquakes within this catalog that have been compiled by the National Earthquake Information Center (NEIC). We use data from 1 January 1977 to 31 December

Table 1
The Differences between the Two Approaches in Previous Studies

	Reasenberg (1999) Approach	Jones (1985) Approach
Catalog	Harvard	Southern California
Declustering	None	Windowing
Search parameters (dT, dX)*	(10 days, 75 km)	(5 days, 10 km)
Magnitude parameters (M_{min}^{main}, dI)†	(7,1.5); (6.5,1)	(None, None)

*dT is the maximum interevent time between the foreshock and the mainshock. dX is the maximum epicentral distance between the foreshock and the mainshock.

†M_{min}^{main} is the minimum magnitude of the mainshock. dI is the maximum magnitude difference between the mainshock and the foreshock.

1996 (inclusive), the same time period as the original Reasenberg (1999) study. We use events of magnitude 5.0 or greater, using the surface-wave magnitude M_s when it is given and the body wave magnitude m_b when M_s is not available, and with depths less than 50 km. These conditions are identical to the original study. We refer to this dataset from here on as the Global dataset.

For each catalog we find the foreshock–mainshock pairs using all possible permutations of the variable values shown in Table 2. We consider the results when no declustering has been applied to the data and when a windowing declustering algorithm has been applied to the data. To apply the declustering process, we use the windowing procedure included in a statistical seismology package for R (Harte and Brownrigg, 2010) (see Data and Resources). The procedure deletes all smaller magnitude events that occur within a specified time and space window following an earthquake. The time and space parameters are dependent on the magnitude of the earthquake under consideration. We use the default parameters for the space and time windows specified in the package. This declustering procedure implies that any earthquake followed by a larger earthquake within the specified windows is considered a foreshock. Therefore, multiple foreshocks to a single event are permitted to remain in our data, identical to the original Jones (1985) study. When the windowing declustering procedure is not applied to the data, we employ the approach of Reasenberg (1999) so that multiple foreshocks to a single event are not permitted; only the largest foreshock is used. The other parameter choices in Table 2 are self-explanatory. Calculated distances between events are epicentral distances. There are 16 possible permutations of

the variable values for the JMA dataset. Similarly, there are eight possible permutations of the variable values for the Global dataset.

Results

In Figures 2 and 3 we show the cumulative proportions of earthquake pairs as a function of magnitude difference between the mainshock and the foreshock for each set of variable values. We use a cumulative proportion rather than a simple count to enable comparison across the parameter values.

Figure 2 shows the cumulative proportion of earthquake pairs as a function of magnitude difference between the mainshock and the foreshock for the JMA dataset as a thick black line. The thinner gray line shows the expected cumulative proportion of earthquake pairs using the theory we discuss in the following section (Validating the Empirical Results). Figure 3 shows the same graphs for the Global catalog. The parameters used to find the pairs are shown in the lower right corner of each smaller plot. The parameters are (in order) declustering type (N, none; W, windowing), dT, dX, M_{min}^{main} (N, none; M_{min}^{main}), and dI (N, none; dI). For example, the black line in the top left graph of Figure 2 shows the cumulative proportion of foreshock–mainshock pairs as a function of magnitude difference using the JMA dataset with windowing declustering, where the mainshock is within five days and ten kilometers of the foreshock, has magnitude greater than or equal to five, and $M^* \geq M^{main} - 2$. We can see that the results are stable for different values of dX and dT. The application of a declustering algorithm also does

Table 2
Values of the Variables Tested in This Research for the JMA Catalog and the Global Catalog*

	JMA Catalog ($M_c \geq 3$)	Global Catalog ($M_c \geq 5$)
Declustering	None; windowing	None; windowing
Search parameters (dT, dX)†	(5 days, 10 km); (10 days, 75 km)	(5 days, 10 km); (10 days, 75 km)
Magnitude parameters (M_{min}^{main}, dI)‡	(5,2); (4,1); (4, None); (None, None)	(6.5,1.5); (None, None)

*All possible permutations for each catalog are tested.

†dT is the maximum interevent time between the foreshock and the mainshock. dX is the maximum epicentral distance between the foreshock and the mainshock.

‡M_{min}^{main} is the minimum magnitude of the mainshock. dI is the maximum magnitude difference between the mainshock and the foreshock.
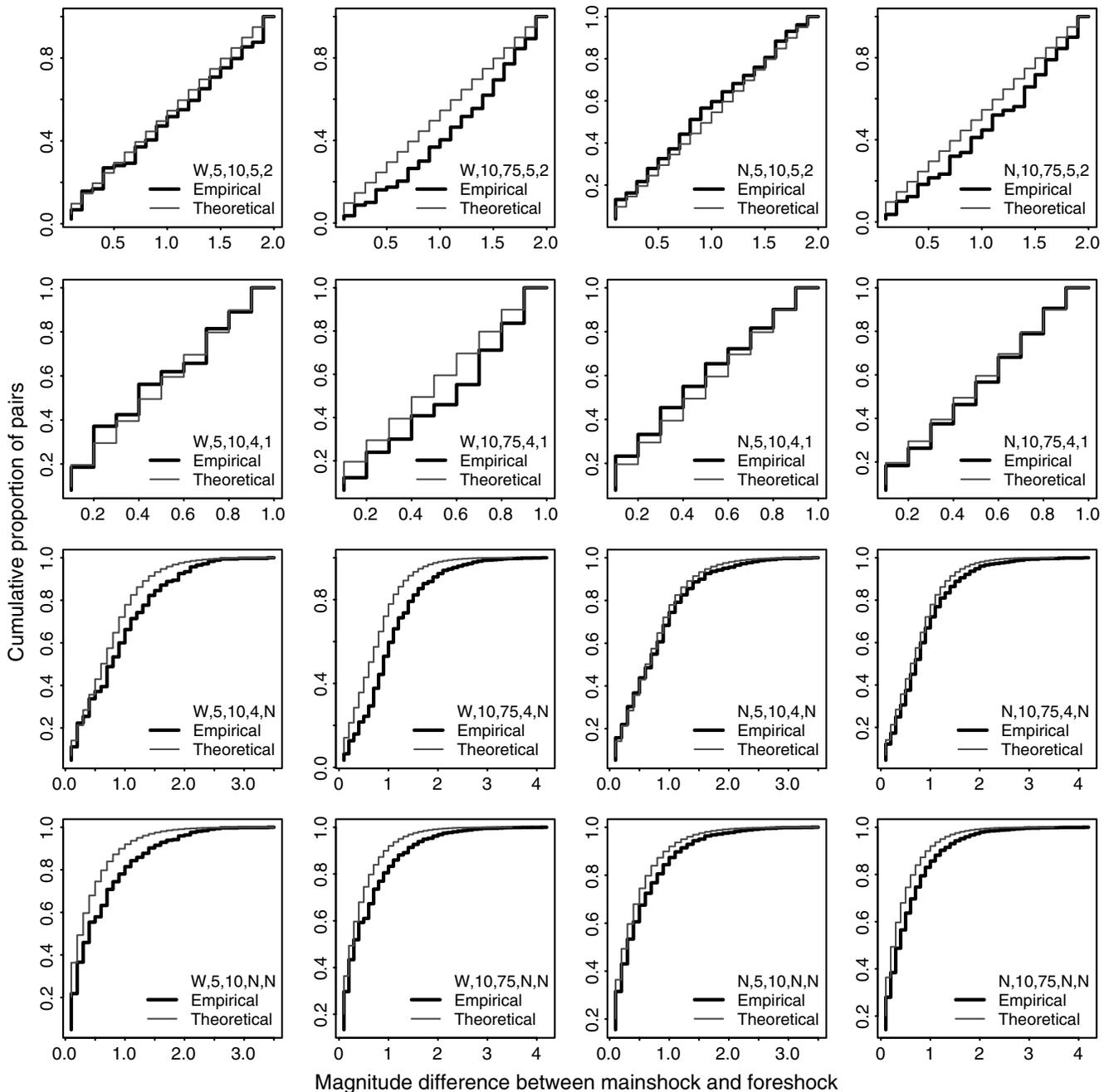
Figure 2. Plots obtained using the JMA dataset. The thick, black line shows the observed cumulative proportion of pairs as a function of magnitude difference between the mainshock and the foreshock for each set of variable values [declustering type (N, none; W, windowing), dT, dX, M_{min}^{main} (N, none; M_{min}^{main}), and dI (N, none; dI)] shown in the lower right corner of each plot. The thin, gray line shows the estimated cumulative proportion of pairs as a function of magnitude difference between the mainshock and the foreshock, as obtained with the theory described in the text.

not affect the shape of the graphs. We start to notice a difference in shape when we do not use identical minimum magnitude thresholds for mainshocks and foreshocks, as is suggested by Agnew and Jones (1991).

The real difference in shape of the distributions is created by the dI variable, which is the maximum difference between the mainshock and foreshock magnitude. If we assume the foreshock to be within dI units of the mainshock

(upper two rows of Fig. 2 and upper row of Fig. 3), we obtain the uniform shape of the distribution found by Reasenberg (1999). However, when dI is unrestricted (lower two rows of Fig. 2 and lower row of Fig. 3), we obtain shapes for the distribution that are curved and similar to the dotted line in Figure 1b.

We therefore believe that the contrasting distributions of the previous studies are a consequence of the use
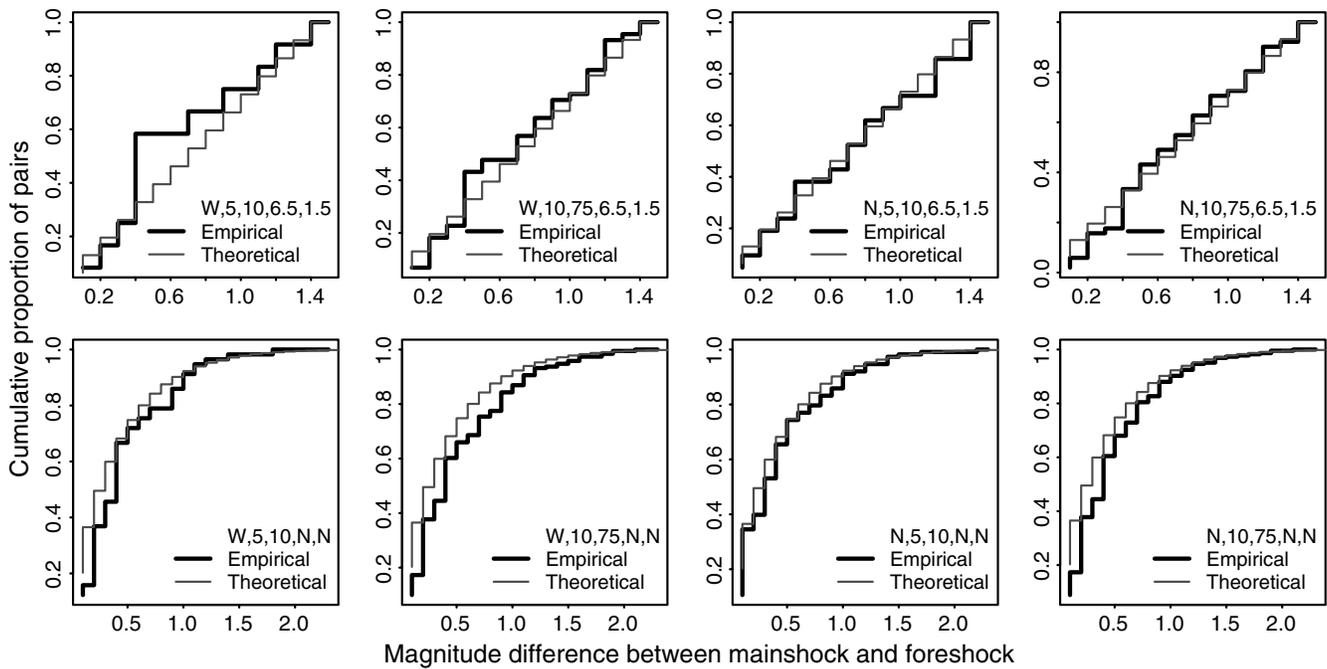
Figure 3. Plots obtained using the Global dataset. The thick, black line shows the observed cumulative proportion of pairs as a function of magnitude difference between the mainshock and the foreshock for each set of variable values [declustering type (N, none; W, windowing), dT, dX, $M_{\min}^{\mathrm{main}}$ (N, none; $M_{\min}^{\mathrm{main}}$), and dI (N, none; dI)] shown in the lower right corner of each plot. The thin, gray line shows the estimated cumulative proportion of pairs as a function of magnitude difference between the mainshock and the foreshock, as obtained with the theory described in the text.

of the mainshock magnitude threshold, in combination with forcing the foreshock magnitude to be within dI units of the mainshock. Simply, the different pairs that are allowed into the respective studies are forcing the different results.

Figure 4a shows the event pairs that were allowed by the Reasenberg (1999) study with $M_{\min}^{\mathrm{main}} = 7$ and $dI = 1.5$. We assume that the maximum magnitude of an earthquake in the dataset is 8, however the shape of the graph will not change if we increase this upper magnitude limit. The y axis shows the possible magnitudes of the foreshock events, the x axis

shows the possible magnitudes of the mainshocks as a magnitude difference from the initiating event. For example, the lowest dot in the left-most column shows that an initiating event of M 6.9 can be followed by an M 7 earthquake (magnitude difference of 0.1). If we follow this row horizontally, we can see that the M 6.9 foreshock can be followed by an M 7 through to an M 8 (magnitude difference of 1.1) earthquake. A dot on the graph indicates that the foreshock–mainshock pair is admissible using the definitions employed by the original study, where a mainshock magnitude has to
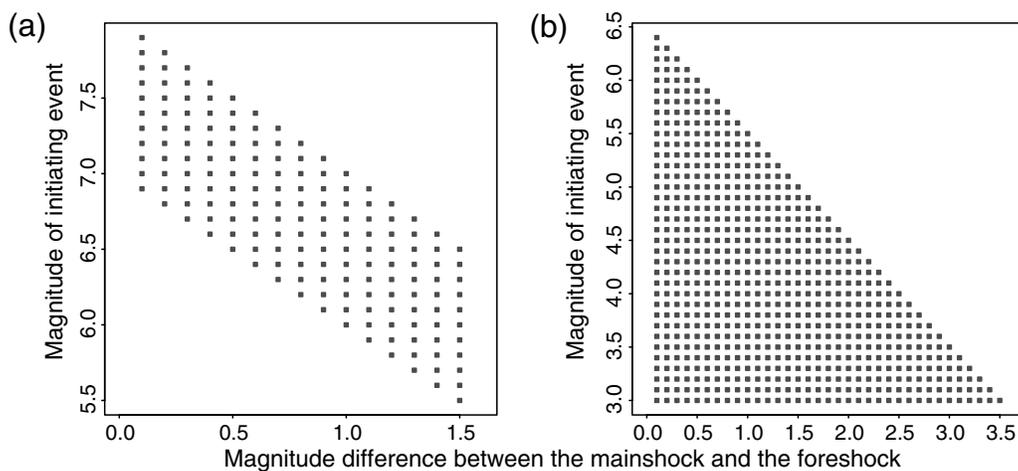

Figure 4. (a) The pairs that are allowed by the Reasenberg (1999) study with $M_{\min}^{\mathrm{main}} = 7$ and $dI = 1.5$. The y axis shows the possible magnitudes of the initiating events, the x axis shows the possible magnitudes of the mainshocks, as a magnitude difference from the initiating event. (b) The same as (a) but showing the pairs allowed by the Jones (1985) study.

be larger than 7 and the foreshock has to be within 1.5 units of the mainshock.

Similarly, Figure 4b shows the pairs that are allowed by the Jones (1985) study with no thresholds for $M_{\min}^{\text{main}}$ and dI. We assume here a maximum magnitude of 6.5, although the shape would not change if we increase this upper magnitude limit. As previously mentioned, the Jones (1985) study does not assign a threshold level for the mainshock magnitude, and any size earthquake can be a foreshock to a larger earthquake.

The main difference between the allowed pairs shown in Figure 4 is the lack of points in the lower left corner of Figure 4a. This means that for the Reasenberg study, smaller magnitude foreshocks and mainshock pairs are not included. This difference in admissible foreshock–mainshock pairs is the driving factor producing the difference in the resultant shapes of the curves shown in Figure 1.

It is somewhat counterintuitive that the more restrictive criteria of Reasenberg (1999) give a more uniform (constant) shape for the distribution of the differences between mainshock and foreshock magnitudes. However, Jones (1985) finds many more foreshocks close in magnitude to the mainshock because smaller-magnitude mainshocks can only have foreshocks slightly lower in magnitude as a result of the magnitude cutoff of the data; and the many smaller mainshocks tend to dominate the results. Finally, we point out that by definition the Jones (1985) pairs include the Reasenberg (1999) pairs. If, as assumed by the Reasenberg (1999) definition, smaller foreshocks do not precede larger earthquakes, we would not observe large-magnitude differences within our catalog data. We do, however, observe small initiating events to larger mainshocks.

Validating the Empirical Results

We can investigate further the effect of the permitted foreshock–mainshock pairs shown in these graphs. We can calculate the probability of each foreshock–mainshock pair, represented as dots within Figure 4, by closely following the work of Vere-Jones *et al.* (2006). These authors show how to calculate the probability of any magnitude difference x between an initiating event with magnitude M^* and the largest event in the sequence. We explain how to incorporate their original derivation here, and we keep our notation as close as possible to the original. First, it is assumed that the magnitudes of events in a sequence are independently and identically distributed as

$$\Pr(M \geq m) = \exp[-\beta(m - M_0)]. \quad (1)$$

Here, $\beta = 2.3b$, where b is the usual b-value of the Gutenberg–Richter distribution (Gutenberg and Richter, 1944) and M_0 is an arbitrary origin. We use $M_0 = M_c$, the cutoff magnitude. Then, if M^{max} is the magnitude of the largest event in the sequence of N events following the initiating event, the probability that M^{max} is less than m is given by

$$\Pr(M^{\text{max}} < m|N) = \{1 - \exp[-\beta(m - M_c)]\}^N. \quad (2)$$

We take the expectation of equation (2) over N, assuming that N has a Poisson distribution with parameter λ, $N \sim \text{Pois}(\lambda)$, so

$$\Pr(M^{\text{max}} < m) = \exp\{-\lambda \exp[-\beta(m - M_c)]\}. \quad (3)$$

Now, let $m = M^* - x$ so that

$$\Pr[M^{\text{max}} < (M^* - x)] = \exp\{-\lambda \exp[-\beta(M^* - x - M_c)]\}. \quad (4)$$

Finally,

$$\begin{aligned} \Pr[(M^* &- M^{\text{max}}) \leq x] \\ &= 1 - \exp\{-\lambda \exp[-\beta(M^* - x - M_c)]\}. \quad (5) \end{aligned}$$

Equation (5) shows the cumulative distribution function $F_\Delta(x)$, where $\Delta = M^* - M^{\text{max}}$. Therefore, we can use

$$\Pr(x_1 \leq \Delta \leq x_2) = \int_{x_1}^{x_2} f_\Delta(x)dx = F_\Delta(x_2) - F_\Delta(x_1) \quad (6)$$

to calculate the probability that the magnitude difference between foreshock and mainshock is between $[x_1, x_2]$. We use $b = 1$ and

$$\lambda = A \exp[\alpha(M^* - M_c)], \quad (7)$$

where $A = 0.05$ and $\alpha = \beta$. We see that the natural logarithm of the expected number of earthquakes triggered by the initiating event is linearly proportional to its size.

To calculate the probability of each admissible pair in Figure 4, we calculate the probability of the appropriate interval for Δ using equation (6) and multiply by the probability of obtaining the corresponding magnitude of the initiating event. Then, we obtain the relative probabilities of each magnitude difference by the summation of the probabilities within the column. The summation is normalized with respect to the other column totals, and this gives us the relative proportions of each admissible magnitude difference.

Figure 5a shows the estimated relative proportions of each magnitude difference using the Reasenberg pairs in Figure 4a, and the estimated relative proportions of each magnitude difference using the Jones pairs in Figure 4b are shown in Figure 5b. We see that the results agree with those found empirically and represented purely schematically in Figure 1. Also, using the theory described previously in this paper, we plot the estimated relative proportions of each magnitude difference for each of the possible sets of variables in Figures 2 and 3 as thin gray lines. We see excellent agreement between those that are found empirically (thick black line) and those that are found by considering the admissible pairs and using the mathematics described previously. We also considered other published distributions
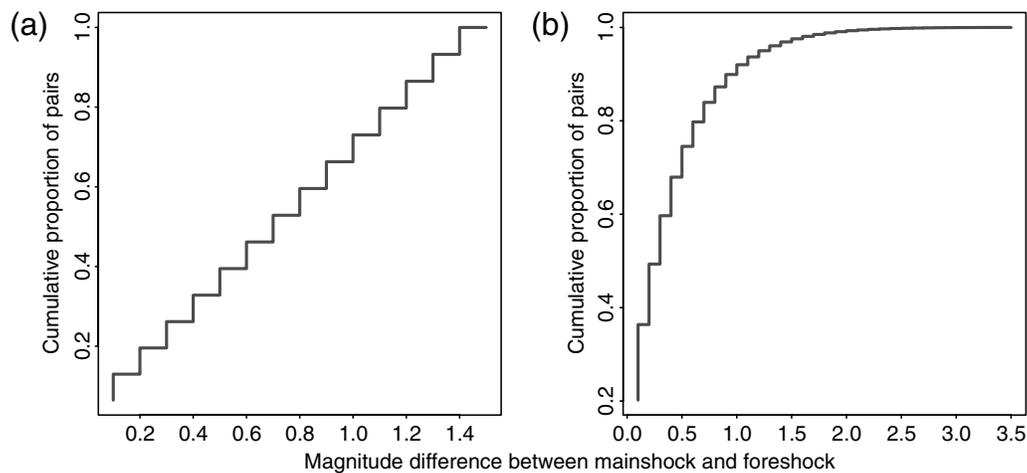
Figure 5. Estimated relative cumulative proportions of each magnitude difference using the theory described in the text for (a) the Reasenberg (1999) pairs and (b) the Jones (1985) pairs.

such as those by Michael and Jones (1998) and found that, by considering their permissible foreshock–mainshock pairs, we could replicate the published empirical results.

Conclusions

We have shown that the results of the previous historical studies differ because of the different criteria for foreshock–mainshock earthquake pairs that are considered by the two sets of authors. That is, the assumption of what constitutes a valid foreshock dictates the structure of the distribution of the differences between mainshock and foreshock magnitudes.

To verify this, we employed the theory, similar to the epidemic-type aftershock sequence (Ogata, 1988) and the Reasenberg–Jones (Reasenberg and Jones, 1989) models, that an earthquake triggers N earthquakes, where N is dependent on the magnitude of the initiating event M^*, and that each triggered earthquake grows to a magnitude chosen randomly from the Gutenberg–Richter distribution (Gutenberg and Richter, 1944). From this model, we can derive theoretical cumulative proportion graphs for any set of defined foreshock–mainshock pairs. Our derived theoretical graphs closely resemble those obtained empirically for both fore-shock–mainshock definitions. Therefore, we believe that the results of the previous studies do not actually contradict each other, as would be assumed from an initial inspection. Rather, these results are complementary in the sense that either can be derived from the same set of assumptions about the underlying process we have described in this research.

Data and Resources

We used hypocentral information from the Japanese Meteorological Agency. We also used hypocentral information from the Global Centroid Moment Tensor Project catalog, available publicly at http://www.globalcmt.org/CMTfiles.html (last accessed July 2011). All functions were coded by us (the authors) in the open-source statistical software package R, available at http://www.r-project.org/ (last accessed July 2011). The ssM8 library for R was also used (Harte and Brownrigg, 2010).

References

Agnew, D. C., and L. M. Jones (1991). Prediction probabilities from foreshocks, *J. Geophys. Res.* **96,** 11,959–11,971.

Dziewonski, A. M., G. Ekström, and M. P. Salganik (1994). Centroid moment tensor solutions for January–March 1994, *Phys. Earth. Planet. In.* **86,** 253–261.

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

Harte, D., and R. Brownrigg (2010). ssM8: M8 earthquake forecasting algorithm, http://www.statsresearch.co.nz/software.html (last accessed July 2011).

Hileman, J., C. Allen, and J. Nordquist (1973). *Seismicity of the Southern California Region, 1 January, 1932–31 December 1972*, Seismological Laboratory, California Institute of Technology, Pasadena, California.

Jones, L. M. (1985). Foreshocks and time-dependent earthquake hazard assessment in southern California, *Bull. Seismol. Soc. Am.* **75,** 1669–1679.

Michael, A. J., and L. M. Jones (1998). Seismicity alert probabilities at Parkfield, California, revisited, *Bull. Seismol. Soc. Am.* **88,** 117–130.

Nanjo, K. Z., T. Ishibe, H. Tsuruoka, D. Schorlemmer, Y. Ishigaki, and N. Hirata (2010). Analysis of the completeness magnitude and seismic network coverage of Japan, *Bull. Seismol. Soc. Am.* **100,** 3261–3268.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.* **83,** 9–27.

Reasenberg, P. A. (1999). Foreshock occurrence before large earthquakes, *J. Geophys. Res.* **104,** 4755–4768.

Reasenberg, P. A., and L. M. Jones (1989). Earthquake hazard after a mainshock in California, *Science* **243,** 1173–1176.

Savage, M. K., and D. M. DePolo (1993). Foreshock probabilities in the western Great Basin, eastern Sierra Nevada, *Bull. Seismol. Soc. Am.* **83,** 1910–1938.

Tormann, T., M. K. Savage, E. G. C. Smith, M. W. Stirling, and S. Wiemer (2008). Time-, distance-, and magnitude-dependent foreshock probability model for New Zealand, *Bull. Seismol. Soc. Am.* **98,** 2149–2160.

Vere-Jones, D., J. Murakami, and A. Christophersen (2006). A further note on Båth's Law, *The 4th International Workshop on Statistical Seismology (Statsei4)*, Kanagawa, Japan, 9–13 January 2006.

Disaster Prevention Research Institute
Kyoto University
Gokasho, Uji
Kyoto 611-0011
Japan
christine@eqh.dpri.kyoto-u.ac.jp