Geophysical Journal International

Geophys. J. Int. (2019) **219,** 1590–1612
Advance Access publication 2019 August 23
GJI Seismology

doi: 10.1093/gji/ggz381

Maximum magnitude of subduction earthquakes along the Japan-Kuril-Kamchatka trench estimated from seismic moment conservation

Fuyuki Hirose,[1] Kenji Maeda[2] and Yasuhiro Yoshida[3]

[1] *Seismology and Tsunami Research Department, Meteorological Research Institute,* 1-1 *Nagamine, Tsukuba, Ibaraki* 305-0052*, Japan. E-mail: fhirose@mri-jma.go.jp*
[2] *Seismology and Volcanology Department, Japan Meteorological Agency,* 1-3-4 *Otemachi, Chiyoda-ku, Tokyo* 100–8122*, Japan*
[3] *Meteorological College, Japan Meteorological Agency,* 7-4-81 *Asahi-cho, Kashiwa, Chiba* 277–0852*, Japan*

Accepted 2019 August 20. Received 2019 July 23; in original form 2019 January 22

SUMMARY

We estimated the maximum magnitude of earthquakes in the Japan-Kuril-Kamchatka trench subduction zone with a method based on the conservation of seismic moment and the record of interplate seismicity from 1977 to 2017. The key point of this method is to base calculations on the tectonic moment rate instead of the total seismic moment rate. We modeled a seismic-moment-frequency distribution for the Japan-Kuril-Kamchatka trench on the basis of the truncated Gutenberg–Richter (G–R) law, the formula published by Utsu in 1974, the gamma distribution, and the tapered G–R law. We estimated the maximum magnitude along the Japan-Kuril-Kamchatka trench as ∼10 under the truncated G–R law and ∼11 under Utsu's formula, although the latter may be an overestimate. Therefore, the 2011 Tohoku earthquake, of moment magnitude 9.2, may not be the largest possible event in this area. The recurrence interval for magnitude 10 events based on the truncated G–R law is 4000 yr. Although these two models perform equally well in terms of Akaike Information Criterion, the range of the 95 per cent confidence level is consistently narrower for the truncated G–R law than for Utsu's formula. The estimated maximum magnitude depends not only on the model used, but also on the parameters that constitute the tectonic moment. It is essential to accumulate more seismic data and achieve more precise estimates of tectonic moment to improve estimates of maximum magnitude.

Key words: Earthquake hazards; Seismicity and tectonics; Subduction zone processes.

1 INTRODUCTION

It is very important in efforts related to earthquake disaster prevention to estimate the maximum magnitude of events that will occur in a region. Precise estimates of this kind generally require data from time periods longer than the average recurrence interval of the largest earthquake. However, the recurrence time of great earthquakes is far longer than human life spans, and available data are limited.

Various attempts have been made to estimate the maximum moment magnitude of earthquakes in the subduction zone along the Japan-Kuril-Kamchatka trench. (To avoid confusion, this paper uses m for moment magnitude and M for seismic moment.) Matsuzawa (2013) used the scaling law between fault area and magnitude to estimate this maximum magnitude as m 10, while admitting it was an extremely rough estimate. Moreover, this estimate did not utilize data of actual seismicity. Kagan & Jackson (2013) estimated maximum magnitudes for all subduction zones on the basis of the principle of conservation of moment, with some assumptions. They estimated the maximum magnitude of earthquakes off the Tohoku district in Japan as m 9.26 ± 0.29 by considering one of the parameters that prescribes the gamma distribution to be the maximum magnitude (a corner magnitude, see section 4.4), and they pointed out that at m 9.2, the 2011 Tohoku earthquake (Hirose *et al.* 2011) was within expectations for this region. However, they also showed that, for the Andaman Island-Sumatra region, the estimated maximum magnitude changed when the data period was extended to include an earthquake with the observed maximum magnitude. Thus, it is worth examining whether the estimated maximum magnitude may change depending on the data with or without of the 2011 Tohoku earthquake and its aftermath. Rong *et al.* (2014) used the tapered Gutenberg–Richter (G–R) law (Kagan 2002a) to estimate maximum magnitudes of m 9.2 for 1000 yr and m 9.3 for 10 000 yr in the wide region extending from the Japan trench to the Kuril-Kamchatka trench.

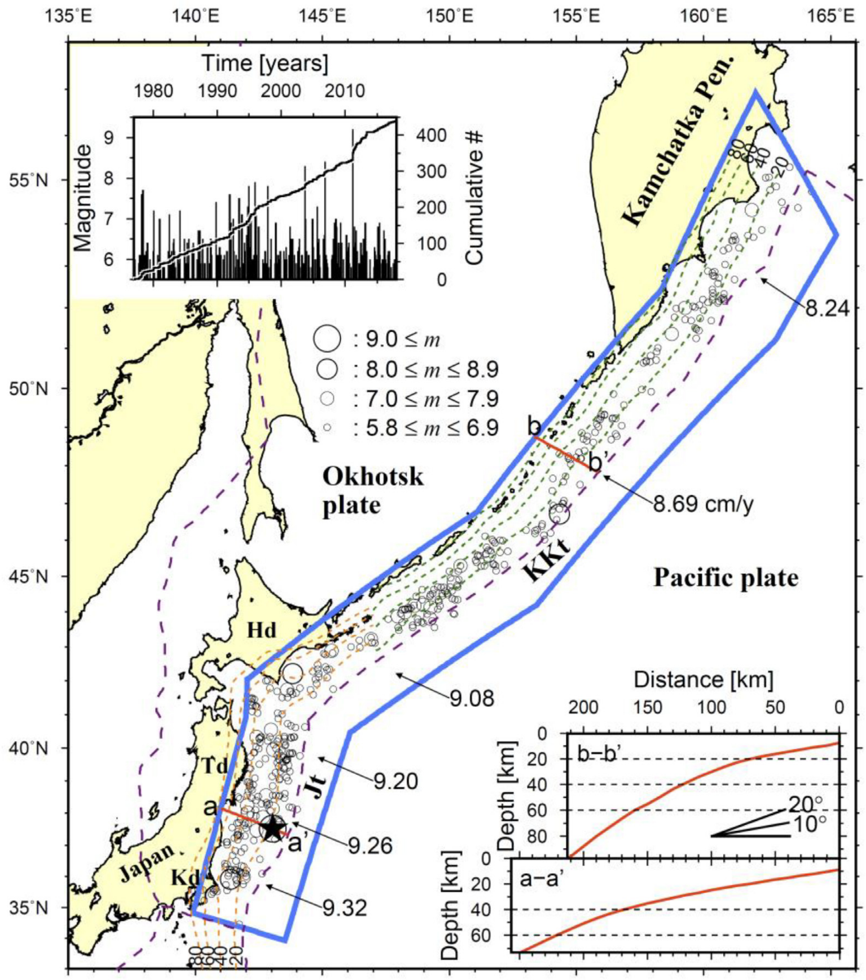

Figure 1. Epicenters of interplate earthquakes of $m \geq 5.8$ at depths of 0–70 km in the Japan-Kuril-Kamchatka trench, 1977–2017. Broken purple lines are plate boundaries from the PB2002 compilation (Bird 2003). Broken orange lines and broken green lines are isodepth contours of 20, 40, 60 and 80 km for the upper boundary of the Pacific slab from Nakajima & Hasegawa (2006) and Kita *et al.* (2010); Hayes *et al.* (2012), respectively. Arrows indicate plate convergence rates between the Pacific and Okhotsk plates (DeMets *et al.* 2010). The inset at upper left is a plot of earthquake magnitudes (vertical bars) and cumulative earthquake number (curve) versus time. The insets at lower right show vertical cross sections of the plate boundary along lines a–a′ and b–b′; no vertical exaggeration. The star denotes the epicenter of the 2011 Tohoku earthquake. Jt, Japan trench; KKt, Kuril-Kamchatka trench; Hd, Hokkaido district; Td, Tohoku district; Kd, Kanto district.

However, neither the gamma distribution nor the tapered G–R law includes a parameter specifying the maximum magnitude, thus these laws allow an event with an infinite magnitude to occur inevitably. On the other hand, two other formulations, the truncated G–R law (Utsu 1978) and Utsu's formula (Utsu 1974), do include such a parameter specifying the maximum magnitude. These may thus be better choices for estimating maximum magnitudes.

In this study, we investigated the maximum magnitude of earthquakes along the Japan-Kuril-Kamchatka trench by applying the truncated G–R law and the Utsu's formula (and both based on the moment conservation principle) to the seismic record of 1977–2017, which includes 7 yr more data than analysed by Kagan & Jackson (2013). We also applied the gamma distribution and the tapered G–R law for reference purposes.

2 DATA

We extracted three data sets of interplate earthquakes with $m \geq 5.8$ from the GCMT catalogue (Dziewonski *et al.* 1981; Ekström *et al.* 2012) from 1977 through 2017 (Fig. 1). We defined interplate events as those within specific ranges of strike angle (150–270°), dip angle (0–45°), rake angle (45–135°) and depth (0–70 km). The range of strike angles allows for the strike of the trench, which ranges from 180° to 240°, plus the error of the GCMT solutions. Our datasets all started on 1 January 1977, but ended on 31 December 2010 (Period 1), 31 December 2013 (Period 2) or 31 December 2017 (Period 3). Period 1 was the same period used by Kagan & Jackson (2013, and Rong *et al.* 2014) and preceded the 2011 Tohoku earthquake. Periods 2 and 3, both of which included the 2011 Tohoku earthquake and aftershocks, were used to check the dependence of estimated results on data periods.

Table 1. Parameters used in this study. The width W of fault was estimated at 173 km along the Kuril-Kamchatka trench and 249 km along the Japan trench in accordance with dip angles (10°, 20°, 30°) of the configuration of the Pacific Plate (inset in Fig. 1). The length L of the fault was totaled 2990 km from 2200 km along the Kuril-Kamchatka trench and 790 km along the Japan trench.

χ (per cent)	μ (GPa)	Dip (°)	W (km)	L (km)	V_{pl} (cm yr^{-1})	$\dot{M}_T$ (10^{20} Nm yr^{-1})
70	49	10/20/30	173/249	2200/790	8.83	17.48

3 METHOD AND PARAMETER SETTING

Like Kagan & Jackson (2013), we used the moment conservation principle to estimate the maximum magnitude. The key point of this method is to restrict the total seismic moment release rate $\dot{M}_s$ to an upper limit set by the tectonic moment rate $\dot{M}_T$, which is the product of the plate convergence rate, the interplate coupling rate, the modulus of rigidity, and the width and length of the fault. We set each of these parameters of $\dot{M}_T$ as follows.

We calculated the plate convergence rate (V_{pl}) from the relative plate motion between the Okhotsk Plate and the subducting Pacific Plate in the MORVEL compilation (DeMets *et al.* 2010) by using the Plate Motion Calculator on the UNAVCO website (https://www.unavco.org/software/geodetic-utilities/geodetic-utilities.html). The resulting rate has a range of 8.24–9.32 cm yr^{-1} along the full set of trenches (arrows in Fig. 1). For simplicity, this rate was set at 8.83 cm yr^{-1}, the average of representative convergence rates near the centre of the Japan trench ($V_{pl} = 9.26$ cm yr^{-1}) and Kuril-Kamchatka trench ($V_{pl} = 8.69$ cm yr^{-1}), weighted in proportion to the length of each trench (i.e. $9.26 \times 1/4 + 8.69 \times 3/4$).

The interplate coupling rate (χ) may be calculated from V_{pl} and the slip deficit rate. We determined the slip deficit rate from Hashimoto *et al.* (2012), who used GNSS data from 1996 to 2000 to estimate slip deficit rates from off Hokkaido through off Kanto before the 2011 Tohoku earthquake. From the spatially heterogeneous results of Hashimoto *et al.* (2012), we adopted a mean slip deficit rate of 6 cm yr^{-1} for the aftershock area of the 2011 Tohoku earthquake. The value of χ, defined by the ratio of slip deficit and V_{pl} (6/9.26 cm yr^{-1}), was thus 65 per cent.

Generally, in estimates of slip deficit from GNSS data, all of the crustal displacement is assumed to occur at the plate boundary (Savage 1983); that is, the stress release caused by small events between and within the plates is ascribed to stable sliding, including slow slip on the plate boundary, and is not considered separately when estimating the tectonic moment. Our purpose, however, is to estimate maximum magnitudes while taking into consideration the seismic moment (energy) released by all interplate earthquakes (see Section 4.6). Accordingly, although it is unknown what proportion of the stable sliding estimated by the GNSS analysis is slip displacement of small interplate earthquakes, a χ value of 65 per cent should be considered to be an underestimate.

Uchida & Matsuzawa (2011) estimated a mean χ of 66 per cent in the aftershock area of the 2011 Tohoku earthquake from their analysis of small repeating earthquakes ($m \sim 3$) from 1993 to March 2007. Note that because they used a V_{pl} value of 7.2 cm yr^{-1} (Shella *et al.* 2002), the stable sliding rate was 2.4 cm yr^{-1}. If instead we assume a V_{pl} of 9.26 cm yr^{-1} based on MORVEL, the slip deficit rate would be 6.86 cm yr^{-1} and the resulting χ would be 74 per cent (=6.86/9.26). Small repeating earthquakes are considered to play a partial role in stable sliding on the plate boundary which, we defined, is consist of seismic slip by small repeating earthquakes and slow slip. The value of χ depends not only on energy released by large earthquakes and stable sliding on the plate boundary, but also on energy released by intraplate earthquakes. However, this analysis of repeating earthquakes does not treat slow slip and earthquakes off the plate boundary, which would result in overestimating χ. Note that Igarashi *et al.* (2003) confirmed that the scaling relationship between seismic moment and seismic slip in California is applicable to interplate earthquakes beneath northeastern Japan. The slip rate of small repeating earthquakes during 1996–2000, the period of the GNSS analysis, is the same as that estimated from small repeating earthquakes from 1993 to March 2007 (Uchida & Matsuzawa 2013).

The difference in χ between the small repeating earthquakes analysis (74 per cent) and the GNSS analysis (65 per cent) is considered to arise from strain release by small interplate and intraplate earthquakes. However, the proportion of the strain release represented by interplate and intraplate earthquakes is unknown. Furthermore, χ is unknown for most of the Kuril-Kamchatka trench. Therefore, we adopted a value of 70 per cent, an intermediate value between 65 and 74 per cent. We also explored this issue for reference purposes by varying χ from 10 to 100 per cent.

We set the modulus of rigidity (μ) at 49 GPa (Bird & Kagan 2004), the same value used by previous studies (Kagan & Jackson 2013; Rong *et al.* 2014).

Because the mean water depth of the Japan-Kuril-Kamchatka trench is about 7 km and the base of the interplate seismogenic zone is 50–60 km deep (e.g. Kita *et al.* 2010), we specified a fault plane with a depth range between 7 and 60 km and a width W of 173 km (=(20–7)/sin 10 + (40–20)/sin 20 + (60–40)/sin 30) along the Kuril-Kamchatka trench and 249 km (=(40–7)/sin 10 + (60–40)/sin 20) along the Japan trench in accordance with the configuration of the Pacific Plate (Nakajima & Hasegawa 2006; Kita *et al.* 2010; Hayes *et al.* 2012) (inset in Fig. 1). The length L of the fault was measured along the strike of the trench and totaled 2990 km from 2200 km along the Kuril-Kamchatka trench and 790 km along the Japan trench. Table 1 shows the list of parameters assigned and the resulting value of $\dot{M}_T$ for $\chi = 70$ per cent.

We applied the truncated G–R law, Utsu's formula, gamma distribution and tapered G–R law to the cumulative frequency-moment distributions of our three seismic datasets. These laws are represented by two parameters, β (the representative parameter of slope of a frequency-moment distribution) and M_c (the characteristic moment corresponding to the maximum magnitude). The M_c terms under each of the four laws are determined uniquely from the corresponding β terms when $\dot{M}_T$ is assigned a priori. The next section presents detailed numerical formulas and parameters for each law.

4 NUMERICAL FORMULAS

Moment (M) is linked to magnitude (m) by the relations

$$\log M = 1.5m + 9.0 \tag{1}$$

$$m = \frac{\log M}{1.5} - 6.0 \tag{2}$$

$$\frac{\partial m}{\partial M} = \frac{1}{1.5M \ln 10}, \tag{3}$$

where log and ln indicate common and natural logarithms, respectively. After Kagan & Jackson (2013) and Rong *et al.* (2014), we adopted 9.0 as the coefficient on the right side in eq. (1), although 9.05 (Hanks & Kanamori 1979) and 9.1 (Kanamori 1977) have also been suggested. This choice of coefficient puts the moment magnitude of the 2011 Tohoku earthquake at 9.2, rather than 9.0 as estimated by the Japan Meteorological Agency (Hirose *et al.* 2011).

The total seismic moment release rate $\dot{M}_s$ as restricted by the tectonic moment rate $\dot{M}_T$ was obtained analytically from the probability density function $\phi(M)$ and the complementary cumulative distribution function $\Phi(M)$ of moment. We show the derivation process for the magnitude and moment representations of each law in Sections 4.1–4.5 and the total seismic moment release rate $\dot{M}_s$ in Section 4.6.

4.1 G–R law

4.1.1 Magnitude representation of G–R law

When the number of earthquakes with magnitudes from m to $m + dm$ in a given region and a given period is defined as $n_m(m)\mathrm{d}m$, their size distribution is approximated by the G–R law (Gutenberg & Richter 1944):

$$\log n_m(m) = a - bm \tag{4}$$

$$n_m(m) = 10^a\, 10^{-bm} = 10^a\, e^{-Bm}, \tag{5}$$

where a and b are constants and $B = b \ln 10$. The total number of earthquakes $N_m(m)$ greater than or equal to m is given by

$$N_m(m) = \int_m^\infty n_m(m')\,\mathrm{d}m' = \frac{10^a e^{-Bm}}{B}. \tag{6}$$

When we set $x = m - m_t$, where m_t is the completeness magnitude, the probability density function $f(x)$ of x is given by

$$f(x) = \frac{n_m(m)}{N_m(m_t)} = \frac{10^a e^{-Bm}}{\frac{10^a e^{-Bm_t}}{B}} = B\, e^{-Bx} = B10^{-bx}. \tag{7}$$

The complementary cumulative distribution function $\mathrm{F}(x)$ of x is given by

$$\mathrm{F}(x) = \int_x^\infty f(x')\,\mathrm{d}x' = e^{-Bx} = 10^{-bx}. \tag{8}$$

$N_m(m)$ is also represented using $\mathrm{F}(x)$ as follows:

$$N_m(m) = N_m(m_t)\,\mathrm{F}(x) = N\mathrm{F}(x), \tag{9}$$

where $N = N_m(m_t)$. Note that the relation between $f(x)$ and $\mathrm{F}(x)$ is

$$f(x) = -\frac{\partial \mathrm{F}(x)}{\partial x}. \tag{10}$$

Because the likelihood L of observed x_i ($i = 1, 2, \ldots, N$) is represented as

$$L = \prod_{i=1}^{N} f(x_i) = f(x_1) \times f(x_2) \times \cdots \times f(x_N), \tag{11}$$

the log-likelihood $\ln L$ becomes

$$\ln\ L = \ln\ f(x_1) + \ln\ f(x_2) + \cdots + \ln\ f(x_N)$$
$$= \sum_{i=1}^{N} \ln f(x_i) = \sum_{i=1}^{N} \ln\left(Be^{-Bx_i}\right) = N\ \ln\ B - B \sum_{i=1}^{N} x_i, \tag{12}$$

and the maximum-likelihood estimate of B becomes

$$\frac{\partial \ln L}{\partial B} = \frac{N}{B} - \sum_{i=1}^{N} x_i = 0 \tag{13}$$

$$\therefore B = \frac{N}{\sum_{i=1}^{N} x_i} = \frac{1}{\mathrm{E}[x]}, \tag{14}$$

where $\mathrm{E}[x]$ is the mean of x_i.

Therefore, the maximum-likelihood estimate of b becomes

$$\therefore b = \frac{1}{\mathrm{E}[x] \ln 10} = \frac{\log e}{\mathrm{E}[x]} = \frac{\log e}{\mathrm{E}[m] - m_t}. \tag{15}$$

4.1.2 *Moment representation of G–R law*

Here we convert the magnitude representation shown in Section 4.1.1 to a moment representation. Again, $n_m(m)\mathrm{d}m$ is the number of earthquakes with magnitudes from m to $m + \mathrm{d}m$. The number of earthquakes with moments from M to $M + \mathrm{d}M$, $n_M(M)\mathrm{d}M$, is derived from eqs (2), (3) and (5) as follows:

$$n_M(M)\,\mathrm{d}M = n_m(m)\,\mathrm{d}m = n_m(m)\frac{\partial m}{\partial M}\mathrm{d}M$$
$$= 10^a\ 10^{-b\left(\frac{\log\ M}{1.5} - 6.0\right)} \frac{1}{1.5\ M\ \ln\ 10}\mathrm{d}M = \frac{10^{a+6.0b}}{1.5\ \ln\ 10} M^{-\beta-1}\mathrm{d}M, \tag{16}$$

where $\beta = b/1.5$. Therefore, the moment representation for $n_m(m)$ in eq. (5), $n_M(M)$, is given by

$$\therefore n_M\ (M) = \frac{10^{a+6.0b}}{1.5 \ln 10} M^{-\beta-1}. \tag{17}$$

From eq. (16), the moment representation for $N_m(m)$ in eq. (6), $N_M(M)$, is given by

$$N_M\ (M) = \int_M^{\infty} n_M\left(M'\right) \mathrm{d}M' = \frac{10^{a+6.0b}}{1.5 \ln 10}\left[\frac{1}{-\beta}\left(M'\right)^{-\beta}\right]_M^{\infty} = \frac{10^{a+6.0b}}{1.5 \ln 10}\frac{1}{\beta} M^{-\beta}. \tag{18}$$

From eqs (17) and (18), the probability density function $\phi(M)$ of M is given by

$$\phi\ (M) = \frac{n_M(M)}{N_M(M_t)} = \frac{\frac{10^{a+6.0b}}{1.5 \ln 10} M^{-\beta-1}}{\frac{10^{a+6.0b}}{1.5 \ln 10}\frac{1}{\beta} {M_t}^{-\beta}} = \beta {M_t}^{\beta} M^{-\beta-1}, \tag{19}$$

and the complementary cumulative distribution function $\Phi(M)$ of M is given by

$$\Phi\ (M) = \int_M^{\infty} \phi\left(M'\right) \mathrm{d}M' = {M_t}^{\beta}\ M^{-\beta} \qquad (\mathrm{M}_t \le M < \infty). \tag{20}$$

$N_M(M)$ is also represented using $\Phi(M)$ as follows:

$$N_M\ (M) = N_M\ (M_t)\,\Phi\ (M) = N\Phi(M). \tag{21}$$

Note that $\phi(M)$ and $\Phi(M)$ are related as

$$\phi\ (M) = -\frac{\partial \Phi(M)}{\partial M}. \tag{22}$$

The log-likelihood $\ln L$ of observed M_i ($i = 1, 2, \ldots, N$) becomes

$$\ln L = \sum_{i=1}^{N} \ln \phi(M_i) = \sum_{i=1}^{N} \ln\left(\beta {M_t}^{\beta} {M_i}^{-\beta-1}\right) = N\left[\ln \beta + \beta \ln M_t\right] - (1+\beta)\sum_{i=1}^{N} \ln M_i, \tag{23}$$

and the maximum-likelihood estimate of B becomes

$$\frac{\partial \ln L}{\partial \beta} = \frac{N}{\beta} + N \ln M_t - \sum_{i=1}^{N} \ln M_i = 0 \tag{24}$$

$$\therefore \beta = \frac{N}{\sum_{i=1}^{N} \ln \frac{M_i}{M_t}}. \tag{25}$$

4.2 Truncated G–R law

4.2.1 Magnitude representation of truncated G–R law

The truncated G–R law with upper magnitude limit c_{tr} (e.g. Utsu 1978) is given by

$$\log n_m(m) = a_{tr} - b_{tr} m \qquad (m \le c_{tr}) \tag{26}$$

$$n_m(m) = 10^{a_{tr}}\ 10^{-b_{tr} m} = 10^{a_{tr}}\ e^{-B_{tr} m} \qquad (m \le c_{tr}) \tag{27}$$

$$n_m(m) = 0 \qquad (m > c_{tr}), \tag{28}$$

where a_{tr}, b_{tr} and c_{tr} are constants and $B_{tr} = b_{tr} \ln 10$. The number of earthquakes with magnitudes from m to c_{tr}, $N_m(m)$, is given by

$$N_m(m) = \int_m^{c_{tr}} n_m(m')\,\mathrm{d}m' = \frac{10^{a_{tr}}}{B_{tr}}\left(e^{-B_{tr} m} - e^{-B_{tr} c_{tr}}\right) \qquad (m \le c_{tr}). \tag{29}$$

The probability density function $f(x)$ of $x(= m - m_t)$ is given by

$$\begin{aligned} f(x) &= \frac{n_m(m)}{N_m(m_t)} = \frac{10^{a_{tr}} e^{-B_{tr} m}}{\frac{10^{a_{tr}}}{B_{tr}}\left(e^{-B_{tr} m_t} - e^{-B_{tr} c_{tr}}\right)} \\ &= \frac{B_{tr}}{e^{B_{tr}(m-m_t)} - e^{B_{tr}(m-c_{tr})}} = \frac{B_{tr}}{e^{B_{tr} x} - e^{B_{tr}(x-C_{tr})}} \\ &= \frac{B_{tr}}{1 - e^{-B_{tr} C_{tr}}}\, e^{-B_{tr} x}, \end{aligned} \tag{30}$$

where $C_{tr} = c_{tr} - m_t$.

The complementary cumulative distribution function $\mathrm{F}(x)$ of x is given by

$$\mathrm{F}(x) = \int_x^{C_{tr}} f(x')\,\mathrm{d}x' = \frac{N_m(m)}{N_m(m_t)} = \frac{\frac{10^{a_{tr}}}{B_{tr}}\left(e^{-B_{tr} m} - e^{-B_{tr} c_{tr}}\right)}{\frac{10^{a_{tr}}}{B_{tr}}\left(e^{-B_{tr} m_t} - e^{-B_{tr} c_{tr}}\right)} = \frac{e^{-B_{tr} x} - e^{-B_{tr} C_{tr}}}{1 - e^{-B_{tr} C_{tr}}}. \tag{31}$$

Here, note that the upper limit of the integration is C_{tr} rather than c_{tr}.

The log-likelihood $\ln L$ becomes

$$\ln L = \sum_{i=1}^{N} \ln f(x_i) = \sum_{i=1}^{N} \ln\left(\frac{B_{tr}}{1 - e^{-B_{tr} C_{tr}}} e^{-B_{tr} x_i}\right) = N \ln \frac{B_{tr}}{1 - e^{-B_{tr} C_{tr}}} - B_{tr} \sum_{i=1}^{N} x_i. \tag{32}$$

See Utsu (1978, 1999b) for details of the estimation procedure for parameters B_{tr} and C_{tr}.

4.2.2 Moment representation of truncated G–R law

Here we convert the magnitude representation shown in Section 4.2.1 to a moment representation. The number of earthquakes with moments from M to $M + \mathrm{d}M$, $n_M(M)\mathrm{d}M$, is derived from eqs (2), (3) and (27) as follows:

$$\begin{aligned} n_M(M)\,\mathrm{d}M &= n_m(m)\,\mathrm{d}m = n_m(m)\frac{\partial m}{\partial M}\mathrm{d}M \\ &= 10^{a_{tr}}\ 10^{-b_{tr}\left(\frac{\log M}{1.5} - 6.0\right)} \frac{1}{1.5 M \ln 10}\mathrm{d}M \\ &= \frac{10^{a_{tr}+6.0 b_{tr}}}{1.5 \ln 10} M^{-\beta_{tr}-1}\mathrm{d}M \qquad (M \le M_{ctr}), \end{aligned} \tag{33}$$

where $\beta_{tr} = b_{tr}/1.5$ and M_{ctr} is the upper limit of moment, corresponding to the upper limit of magnitude c_{tr}. Therefore, the moment representation for $n_m(m)$ in eq. (27), $n_M(M)$, is given by

$$\therefore n_M(M) = \frac{10^{a_{tr}+6.0 b_{tr}}}{1.5 \ln 10} M^{-\beta_{tr}-1} \qquad (M \le M_{ctr}). \tag{34}$$

From eq. (33), the moment representation for $N_m(m)$ in eq. (29), $N_M(M)$, is given by

$$
\begin{aligned}
N_M\ (M) &= \int_M^{M_{ctr}} n_M\,(M')\,\mathrm{d}M' = \frac{10^{a_{tr}+6.0b_{tr}}}{1.5\ \ln\ 10}\left[\frac{1}{-\beta_{tr}}\left(M'\right)^{-\beta_{tr}}\right]_M^{M_{ctr}} \\
&= \frac{10^{a_{tr}+6.0b_{tr}}}{1.5\ \ln\ 10}\frac{1}{\beta_{tr}}\left(M^{-\beta_{tr}} - {M_{ctr}}^{-\beta_{tr}}\right).
\end{aligned}
\tag{35}
$$

From eqs (34) and (35), the probability density function $\phi(M)$ of M is given by

$$
\begin{aligned}
\phi\,(M) &= \frac{n_M\,(M)}{N_M\,(M_t)} = \frac{\dfrac{10^{a_{tr}+6.0b_{tr}}}{1.5\ \ln\ 10}M^{-\beta_{tr}-1}}{\dfrac{10^{a_{tr}+6.0b_{tr}}}{1.5\ \ln\ 10}\dfrac{1}{\beta_{tr}}\left({M_t}^{-\beta_{tr}} - {M_{ctr}}^{-\beta_{tr}}\right)} \\
&= \beta_{tr}\frac{M^{-\beta_{tr}-1}}{{M_t}^{-\beta_{tr}} - {M_{ctr}}^{-\beta_{tr}}} = \frac{{M_{ctr}}^{\beta_{tr}}{M_t}^{\beta_{tr}}}{{M_{ctr}}^{\beta_{tr}} - {M_t}^{\beta_{tr}}}\beta_{tr}M^{-\beta_{tr}-1} \\
&= \frac{\beta_{tr}{M_t}^{\beta_{tr}}M^{-\beta_{tr}-1}}{1-\left(\dfrac{M_{ctr}}{M_t}\right)^{-\beta_{tr}}} \qquad (M_t \le M \le M_{ctr}),
\end{aligned}
\tag{36}
$$

and the complementary cumulative distribution function $\Phi(M)$ of M is given by

$$
\begin{aligned}
\Phi\,(M) &= \int_M^{M_{ctr}} \phi\,(M')\,\mathrm{d}M' = \frac{N_M\,(M)}{N_M\,(M_t)} = \frac{M^{-\beta_{tr}} - {M_{ctr}}^{-\beta_{tr}}}{{M_t}^{-\beta_{tr}} - {M_{ctr}}^{-\beta_{tr}}} \\
&= {M_t}^{\beta_{tr}}\ M^{-\beta_{tr}}\frac{1-\left(\dfrac{M_{ctr}}{M}\right)^{-\beta_{tr}}}{1-\left(\dfrac{M_{ctr}}{M_t}\right)^{-\beta_{tr}}} \qquad (M_t \le M \le M_{ctr}).
\end{aligned}
\tag{37}
$$

Comparing this equation with eq. (20) of the G–R law, we can find a taper with power. Note that this equation accords with eq. (20) when M_{ctr} becomes ∞. The log-likelihood $\ln L$ becomes

$$
\begin{aligned}
\ln\ L &= \sum_{i=1}^{N}\ln\ \phi\,(M_i) = \sum_{i=1}^{N}\ln\left(\frac{{M_{ctr}}^{\beta_{tr}}{M_t}^{\beta_{tr}}}{{M_{ctr}}^{\beta_{tr}} - {M_t}^{\beta_{tr}}}\beta_{tr}{M_i}^{-\beta_{tr}-1}\right) \\
&= N\left[\beta_{tr}\ \ln M_{ctr} + \beta_{tr}\ \ln\ M_t + \ln\ \beta_{tr} - \ln\left({M_{ctr}}^{\beta_{tr}} - {M_t}^{\beta_{tr}}\right)\right] - (1+\beta_{tr})\sum_{i\,=\,1}^{N}\ln\ M_i.
\end{aligned}
\tag{38}
$$

4.3 Utsu's formula

4.3.1 Magnitude representation of Utsu's formula

Utsu (1974) suggested a law with a different upper magnitude limit, c_u, in order to represent a decrease of frequency near the upper limit of magnitude, than that of the truncated G–R law in Section 4.2. Utsu's formula is given by

$$\log n_m\,(m) = a_u\ - b_u m + \log\,(c_u - m) \qquad (m < c_u) \tag{39}$$

$$n_m\ (m) = 10^{a_u}\ 10^{-b_u m}\ 10^{\log(c_u-m)} = 10^{a_u}\ e^{-B_u m}\,(c_u - m) \qquad (m < c_u) \tag{40}$$

$$n_m\ (m) = \ 0 \qquad (m \ge c_u), \tag{41}$$

where a_u, b_u, and c_u are constants and $B_u = b_u\ \ \ln 10$. We can find that eq. (39) was expressed by the addition of eq. (4) of the G–R law and logarithmic taper. The number of earthquakes with magnitudes from m to c_u, $N_m(m)$, is given by

$$
\begin{aligned}
N_m\,(m) &= \int_m^{c_u} n_m\,(m')\,\mathrm{d}m' \ = 10^{a_u}\ \int_m^{c_u} e^{-B_u m'}\,(c_u - m')\,\mathrm{d}m' \\
&= 10^{a_u}\ \left\{\left[(c_u - m')\frac{1}{-B_u}e^{-B_u m'}\right]_m^{c_u} - \int_m^{c_u}(-1)\frac{1}{-B_u}e^{-B_u m'}dm'\right\} \\
&= \frac{10^{a_u}}{B_u}\left\{\left(c_u - m - \frac{1}{B_u}\right)e^{-B_u m} + \frac{1}{B_u}e^{-B_u c_u}\right\} \qquad (m < c_u).
\end{aligned}
\tag{42}
$$

The probability density function $f(x)$ of x is given by

$$
\begin{aligned}
f\ (x) &= \frac{n_m\,(m)}{N_m\,(m_t)} = \frac{10^{a_u}e^{-B_u m}\,(c_u - m)}{\dfrac{10^{a_u}}{B_u}\left\{\left(c_u - m_t - \dfrac{1}{B_u}\right)e^{-B_u m_t} + \dfrac{1}{B_u}e^{-B_u c_u}\right\}} \\
&= \frac{{B_u}^2 e^{-B_u x}\,(C_u - x)}{B_u C_u - 1 + e^{-B_u C_u}} = \frac{{B_u}^2 e^{-B_u x}\,(C_u - x)}{P},
\end{aligned}
\tag{43}
$$

where $C_u = c_u - m_t$ and $P = B_u C_u - 1 + e^{-B_u C_u}$.

The complementary cumulative distribution function $\mathrm{F}(x)$ of x is given by

$$\begin{aligned} \mathrm{F}(x) &= \int_x^{C_u} f(x')\,\mathrm{d}x' = \frac{B_u{}^2}{P}\int_x^{C_u} e^{-B_u x'}(C_u - x')\,\mathrm{d}x' \\ &= \frac{B_u}{P}\left\{\left(C_u - x - \frac{1}{B_u}\right)e^{-B_u x} + \frac{1}{B_u}e^{-B_u C_u}\right\}. \end{aligned} \tag{44}$$

Here, note that the upper limit of the integration is C_u rather than c_u. Furthermore, note that the opening brace '{' is incorrectly placed in eq. (11.105) of Utsu (1999b).

The log-likelihood $\ln L$ becomes

$$\begin{aligned} \ln\ L &= \sum_{i=1}^{N} \ln\ f(x_i) = \sum_{i=1}^{N} \ln\left\{\frac{B_u{}^2 e^{-B_u x_i}(C_u - x_i)}{P}\right\} \\ &= N(2\ \ln\ B_u - \ln\ P - B_u \mathrm{E}[x]) + \sum_{i=1}^{N} \ln(C_u - x_i), \end{aligned} \tag{45}$$

and the maximum-likelihood estimates of B_u and C_u can be obtained from $\partial \ln L/\partial B_u = \partial \ln L/\partial C_u = 0$ (e.g. Utsu 1999b; Mabuchi *et al.* 2002).

4.3.2 *Moment representation of Utsu's formula*

Here we convert the magnitude representation shown in Section 4.3.1 to a moment representation. The number of earthquakes with moments from M to $M + \mathrm{d}M$, $n_M(M)\mathrm{d}M$, is derived from eqs (2), (3) and (40) as follows:

$$\begin{aligned} n_M(M)\,\mathrm{d}M &= n_m(m)\,\mathrm{d}m = n_m(m)\frac{\partial m}{\partial M}\mathrm{d}M \\ &= 10^{a_u}\ 10^{-b_u\left(\frac{\log\ M}{1.5} - 6.0\right)}\ 10^{\log\left\{\left(\frac{\log\ M_{cu}}{1.5} - 6.0\right) - \left(\frac{\log\ M}{1.5} - 6.0\right)\right\}}\ \frac{1}{1.5M\ \ln\ 10}\mathrm{d}M \\ &= \frac{10^{a_u + 6.0b_u}}{1.5\ \ln\ 10}M^{-\beta_u - 1}\frac{1}{1.5}\log\frac{M_{cu}}{M}dM \qquad (M < M_{cu}), \end{aligned} \tag{46}$$

where $\beta_u = b_u/1.5$ and M_{cu} is the upper limit of moment corresponding to the upper limit of magnitude c_u. Therefore, the moment representation for $n_m(m)$ in eq. (40), $n_M(M)$, is given by

$$\therefore n_M\ (M) = \frac{10^{a_u + 6.0b_u}}{1.5 \ln 10}M^{-\beta_u - 1}\frac{1}{1.5}\log\frac{M_{cu}}{M} \qquad (M < M_{cu})\ . \tag{47}$$

From eq. (46), the moment representation for $N_m(m)$ in eq. (42), $N_M(M)$, is given by

$$\begin{aligned} N_M\ (M) &= \int_M^{M_{cu}} n_M(M')\,\mathrm{d}M' = \frac{1}{1.5}\frac{10^{a_u + 6.0b_u}}{1.5\ \ln\ 10}\int_M^{M_{cu}} M'^{-\beta_u - 1}\log\frac{M_{cu}}{M'}\mathrm{d}M' \\ &= \frac{1}{1.5}\frac{10^{a_u + 6.0b_u}}{1.5\ \ln\ 10}\frac{1}{\beta_u}\left(M^{-\beta_u}\log\frac{M_{cu}}{M} - \frac{M^{-\beta_u} - M_{cu}{}^{-\beta_u}}{\beta_u\ \ln\ 10}\right). \end{aligned} \tag{48}$$

From eqs (47) and (48), the probability density function $\phi(M)$ of M is given by

$$\begin{aligned} \phi(M) &= \frac{n_M(M)}{N_M(M_t)} = \frac{\frac{10^{a_u + 6.0b_u}}{1.5\ \ln\ 10}M^{-\beta_u - 1}\frac{1}{1.5}\log\frac{M_{cu}}{M}}{\frac{1}{1.5}\frac{10^{a_u + 6.0b_u}}{1.5\ \ln\ 10}\frac{1}{\beta_u}\left(M^{-\beta_u}\log\frac{M_{cu}}{M_t} - \frac{M_t{}^{-\beta_u} - M_{cu}{}^{-\beta_u}}{\beta_u \ln 10}\right)} \\ &= \beta_u\ M_t{}^{\beta_u}M^{-\beta_u - 1}\frac{\log\frac{M_{cu}}{M}}{\log\frac{M_{cu}}{M_t} - \frac{1 - \left(\frac{M_{cu}}{M_t}\right)^{-\beta_u}}{\beta_u \ln 10}} \qquad (M_t \le M \le M_{cu}), \end{aligned} \tag{49}$$

and the complementary cumulative distribution function $\Phi(M)$ of M is given by

$$\Phi(M) = \int_M^{M_{cu}} \phi(M')\,\mathrm{d}M' = \frac{N_M(M)}{N_M(M_t)} = \frac{M^{-\beta_u}\log\dfrac{M_{cu}}{M} - \dfrac{M^{-\beta_u} - M_{cu}{}^{-\beta_u}}{\beta_u \ln 10}}{M_t{}^{-\beta_u}\log\dfrac{M_{cu}}{M_t} - \dfrac{M_t{}^{-\beta_u} - M_{cu}{}^{-\beta_u}}{\beta_u \ln 10}}$$

$$= M_t{}^{\beta_u}\, M^{-\beta_u} \frac{\log\dfrac{M_{cu}}{M} - \dfrac{1-\left(\dfrac{M_{cu}}{M}\right)^{-\beta_u}}{\beta_u \ln 10}}{\log\dfrac{M_{cu}}{M_t} - \dfrac{1-\left(\dfrac{M_{cu}}{M_t}\right)^{-\beta_u}}{\beta_u \ln 10}} \qquad (M_t \le M \le M_{cu})\ . \tag{50}$$

The log-likelihood $\ln L$ becomes

$$\ln L = \sum_{i=1}^{N} \ln\phi(M_i) = \sum_{i=1}^{N} \ln\left[\beta_u M_t{}^{\beta_u} M_i{}^{-\beta_u - 1} \frac{\log\dfrac{M_{cu}}{M_i}}{\log\dfrac{M_{cu}}{M_t} - \dfrac{1-\left(\dfrac{M_{cu}}{M_t}\right)^{-\beta_u}}{\beta_u \ln 10}}\right]. \tag{51}$$

Although maximum-likelihood estimates of β_u and M_{cu} can be obtained from $\partial\ln L/\partial\beta_u = \partial\ln L/\partial M_{cu} = 0$, the formulas are complicated. Therefore, for practical applications we recommend converting parameters B_u and C_u obtained by eq. (45) to the moment representations β_u and M_{cu} using eq. (1). In this study, β_u was estimated numerically by a grid search for the maximum value of eq. (51) because M_{cu} is represented by β_u by setting $\dot{M}_T$ *a priori* (see Section 4.7).

4.4 Gamma distribution

4.4.1 Magnitude representation of gamma distribution

The gamma distribution is an example of a law stipulating that there is no upper limit of magnitude or moment, although the number of earthquakes decreases sharply when m or M becomes large (Saito *et al.* 1973; Kagan 1991). In this section, we treat expressions that Utsu called the generalized Saito *et al.* equation [eq. (15) in Utsu 1999a]. Kagan (2002a, Section 2.2.1) used a version of the gamma distribution that was based on the generalized Saito *et al.* equation, but did not present a magnitude representation of it. The generalized Saito *et al.* equation is given by

$$\log n_m(m) = a_g - b_g m - k10^{1.5m} \tag{52}$$

$$n_m(m) = 10^{a_g}\, 10^{-b_g m}\, 10^{-k10^{1.5m}} = 10^{a_g}\, e^{-B_g m} e^{-\gamma e^{\frac{B_g}{\beta_g} m}}, \tag{53}$$

where a_g, b_g and k are constants, $B_g = b_g \ln 10$, $\beta_g = b_g/1.5$ and $\gamma = k\ln 10$. The number of earthquakes $N_m(m)$ with magnitudes greater than or equal to m is given by

$$N_m(m) = \int_m^{\infty} n_m(m')\,\mathrm{d}m' = \frac{10^{a_g}\beta_g\gamma^{\beta_g}}{B_g}\Gamma\left(-\beta_g,\ \gamma e^{\frac{B_g}{\beta_g}m}\right), \tag{54}$$

where $\Gamma(-\beta_g,\ \gamma e^{\frac{B_g}{\beta_g}m})$ is an upper incomplete gamma function defined as

$$\Gamma(a,\ x) \equiv \int_x^{\infty} t^{a-1} e^{-t}\mathrm{d}t \qquad (x > 0)\ . \tag{55}$$

The probability density function $f(x)$ of x is given by

$$f(x) = \frac{n_m(m)}{N_m(m_t)} = \frac{10^{a_g} e^{-B_g m} e^{-\gamma e^{\frac{B_g}{\beta_g}m}}}{\dfrac{10^{a_g}\beta_g\gamma^{\beta_g}}{B_g}\Gamma\left(-\beta_g,\ \gamma e^{\frac{B_g}{\beta_g}m_t}\right)} = \frac{B_g e^{-B_g x} e^{-De^{\frac{B_g}{\beta_g}x}}}{\beta_g D^{\beta_g}\Gamma(-\beta_g,\ D)}, \tag{56}$$

where $D = \gamma e^{\frac{B_g}{\beta_g}m_t}$. Note that eq. (56) is consistent with eq. (11.138) of Utsu (1999b) with $\beta_g = 1/2$ (note that c in Utsu's eq. [11.138] is a misprint of C which is the same as D in this paper).

The complementary cumulative distribution function $\mathrm{F}(x)$ of x is given by

$$\mathrm{F}\ (x) = \int_x^{\infty} f\left(x'\right) \mathrm{d}x' \ = \frac{\Gamma\left(-\beta_g,\ De^{\frac{B_g}{\beta_g}x}\right)}{\Gamma\left(-\beta_g,\ D\right)}. \tag{57}$$

4.4.2 Moment representation of gamma distribution

Here we convert the magnitude representation shown in Section 4.4.1 to a moment representation. The number of earthquakes with moments from M to $M + \mathrm{d}M$, $n_M(M)\mathrm{d}M$, is derived from eqs (2), (3) and (53) as follows:

$$\begin{aligned} n_M\,(M)\,\mathrm{d}M &= n_m\,(m)\,\mathrm{d}m \ = n_m\ (m)\,\frac{\partial m}{\partial M}\mathrm{d}M \\ &= 10^{a_g}\ 10^{-b_g\left(\frac{\log\ M}{1.5} - 6.0\right)}\ 10^{-k10^{1.5\left(\frac{\log\ M}{1.5} - 6.0\right)}}\ \frac{1}{1.5M\ \ln\ 10}\mathrm{d}M \\ &= \frac{10^{a_g+6.0b_g}}{1.5\ \ln\ 10} M^{-\beta_g-1} e^{-10^{-9.0}\gamma M}\mathrm{d}M \\ &= \frac{10^{a_g+6.0b_g}}{1.5\ \ln\ 10} M^{-\beta_g-1} e^{-\frac{M}{M_{cg}}}\,\mathrm{d}M, \end{aligned} \tag{58}$$

where M_{cg} is a corner moment parameter that characterizes the frequency-moment distribution, which we set at $M_{cg} = (10^{-9.0}\gamma)^{-1}$. Note that M_{cg} is not an upper limit of moment like M_{ctr} in the truncated G–R law and M_{cu} in Utsu's formula. Therefore, the moment representation for $n_m(m)$ in eq. (53), $n_M(M)$, is given by

$$\therefore n_M\ (M) = \frac{10^{a_g+6.0b_g}}{1.5\,\ln 10} M^{-\beta_g-1} e^{-\frac{M}{M_{cg}}}\ . \tag{59}$$

From eqs (55) and (58), the moment representation for $N_m(m)$ in eq. (54), $N_M(M)$, is given by

$$\begin{aligned} N_M\ (M) &= \int_M^{\infty} n_M\,(M')\,\mathrm{d}M' \ = \frac{10^{a_g+6.0b_g}}{1.5\ \ln\ 10} \int_M^{\infty} M'^{-\beta_g-1} e^{-\frac{M'}{M_{cg}}}\,\mathrm{d}M' \\ &= \frac{10^{a_g+6.0b_g}}{1.5\ \ln\ 10} M_{cg}{}^{-\beta_g} \int_{\frac{M}{M_{cg}}}^{\infty} t^{-\beta_g-1} e^{-t}\mathrm{d}t \\ &= \frac{10^{a_g+6.0b_g}}{1.5\ \ln\ 10} M_{cg}{}^{-\beta_g}\,\Gamma\left(-\beta_g,\ \frac{M}{M_{cg}}\right), \end{aligned} \tag{60}$$

where we substituted t for $\frac{M'}{M_{cg}}$ for simplicity.

From eqs (59) and (60), the probability density function $\phi(M)$ of M is given by

$$\begin{aligned} \phi\,(M) &= \frac{n_M\,(M)}{N_M\,(M_t)} = \frac{\frac{10^{a_g+6.0b_g}}{1.5\ \ln\ 10} M^{-\beta_g-1} e^{-\frac{M}{M_{cg}}}}{\frac{10^{a_g+6.0b_g}}{1.5\ \ln\ 10} M_{cg}{}^{-\beta_g}\,\Gamma\left(-\beta_g,\ \frac{M_t}{M_{cg}}\right)} \\ &= \frac{M^{-\beta_g-1} e^{-\frac{M}{M_{cg}}}}{M_{cg}{}^{-\beta_g}\,\Gamma\left(-\beta_g,\ \frac{M_t}{M_{cg}}\right)} \qquad (M_t \le M \le \infty)\,, \end{aligned} \tag{61}$$

and the complementary cumulative distribution function $\Phi(M)$ of M is given by

$$\Phi\ (M) = \int_M^{\infty} \phi\left(M'\right) \mathrm{d}M' \ = \frac{N_M\,(M)}{N_M\,(M_t)} \ = \frac{\Gamma\left(-\beta_g,\ \frac{M}{M_{cg}}\right)}{\Gamma\left(-\beta_g,\ \frac{M_t}{M_{cg}}\right)} \qquad (M_t \le M \le \infty)\ . \tag{62}$$

The log-likelihood $\ln L$ becomes

$$\ln L = \sum_{i=1}^{N} \ln \phi(M_i) = \sum_{i=1}^{N} \ln \left(\frac{M_i{}^{-\beta_g - 1} e^{-\frac{M_i}{M_{cg}}}}{M_{cg}{}^{-\beta_g} \Gamma\left(-\beta_g, \frac{M_t}{M_{cg}}\right)} \right)$$
$$= N\left[\beta_g \ln M_{cg} - \ln \Gamma\left(-\beta_g, \frac{M_t}{M_{cg}}\right)\right] - (1+\beta_g) \sum_{i=1}^{N} \ln M_i - \frac{1}{M_{cg}} \sum_{i=1}^{N} M_i. \tag{63}$$

Note that

$$\Gamma(a+1, x) = a\Gamma(a, x) + x^a e^{-x}. \tag{64}$$

Through eq. (64), eq. (61) becomes the following complicated form that was presented as eqs (16) and (17) by Kagan (2002a):

$$\phi(M) = \frac{M^{-\beta_g - 1} e^{-\frac{M}{M_{cg}}}}{M_{cg}{}^{-\beta_g} \Gamma\left(-\beta_g, \frac{M_t}{M_{cg}}\right)}$$
$$= \frac{\beta_g M^{-\beta_g - 1} M_t{}^{\beta_g} e^{\frac{M_t - M}{M_{cg}}}}{1 - \left(\frac{M_t}{M_{cg}}\right)^{\beta_g} e^{\frac{M_t}{M_{cg}}} \Gamma\left(1 - \beta_g, \frac{M_t}{M_{cg}}\right)}$$
$$= G^{-1} \frac{\beta_g}{M} \left(\frac{M_t}{M}\right)^{\beta_g} e^{\frac{M_t - M}{M_{cg}}} \qquad (M_t \le M \le \infty) \tag{65}$$

$$G = 1 - \left(\frac{M_t}{M_{cg}}\right)^{\beta_g} e^{\frac{M_t}{M_{cg}}} \Gamma\left(1 - \beta_g, \frac{M_t}{M_{cg}}\right). \tag{66}$$

Similarly, through eq. (64), eq. (62) becomes the complicated form presented as eq. (19) by Kagan (2002a):

$$\Phi(M) = \int_M^{\infty} \phi(M')\, \mathrm{d}M' = \frac{\Gamma\left(-\beta_g, \frac{M}{M_{cg}}\right)}{\Gamma\left(-\beta_g, \frac{M_t}{M_{cg}}\right)}$$
$$= \frac{-\beta_g{}^{-1} \Gamma\left(1 - \beta_g, \frac{M}{M_{cg}}\right) - \left(\frac{M}{M_{cg}}\right)^{-\beta_g} e^{-\frac{M}{M_{cg}}}}{-\beta_g{}^{-1} \Gamma\left(1 - \beta_g, \frac{M_t}{M_{cg}}\right) - \left(\frac{M_t}{M_{cg}}\right)^{-\beta_g} e^{-\frac{M_t}{M_{cg}}}}$$
$$= G^{-1} \left(\frac{M_t}{M}\right)^{\beta_g} e^{\frac{M_t - M}{M_{cg}}} \left\{1 - \left(\frac{M}{M_{cg}}\right)^{\beta_g} e^{\frac{M}{M_{cg}}} \Gamma\left(1 - \beta_g, \frac{M}{M_{cg}}\right)\right\} \tag{67}$$

Similarly, through eqs (64) and (65), eq. (63) becomes the form shown as eq. (38) by Kagan (2002a):

$$\ln L = \sum_{i=1}^{N} \ln \phi(M_i) = \sum_{i=1}^{N} \ln \left(G^{-1} \frac{\beta_g}{M} \left(\frac{M_t}{M}\right)^{\beta_g} e^{\frac{M_t - M}{M_{cg}}} \right)$$
$$= N\left[\beta_g \ln M_t + \ln \beta_g - \ln G\right] + \frac{N M_t - \sum_{i=1}^{N} M_i}{M_{cg}} - (1+\beta_g) \sum_{i=1}^{N} \ln M_i. \tag{68}$$

Note that the plus sign of $\log C$ in eq. (38) of Kagan (2002a) is changed to a minus sign, that is, $-\ln G$, in eq. (68) of this study. It appears that Kagan (2002a) used the complicated forms of eqs (65)–(67) rather than the simpler forms of eqs (61) and (62) for easier comparison with his presentation of the tapered G–R law (see Section 4.5) in the same paper.

4.5 Tapered G–R law

4.5.1 Moment representation of tapered G–R law

The distribution functions shown in Sections 4.1–4.4 were based on magnitude representations. However, no magnitude representation of the tapered G–R law (Vere-Jones *et al.* 2001; Kagan 2002a) has been published. Therefore, in Section 4.5.2 we present a derivation based on the moment representation of Kagan (2002a).

According to eq. (12) of Kagan (2002a), the probability density function $\phi(M)$ of M is given by

$$\phi\ (M) = \left[\frac{\beta_{ta}}{M} + \frac{1}{M_{cta}}\right]\left(\frac{M_t}{M}\right)^{\beta_{ta}} e^{\dfrac{M_t - M}{M_{cta}}} \qquad (M_t \le M \le \infty)\,, \tag{69}$$

where β_{ta} and M_{cta} are constants. M_{cta} is a corner moment parameter that characterizes the frequency-moment distribution, just as M_{cg} defines the gamma distribution and unlike M_{ctr} in the truncated G–R law and M_{cu} in Utsu's formula, M_{cta} is not an upper limit of moment. From eq. (69), the complementary cumulative distribution function $\Phi(M)$ of M is given by

$$\begin{aligned}\Phi\,(M) &= \int_M^\infty \phi\,(M')\,\mathrm{d}M' = {M_t}^{\beta_{ta}}\, e^{\dfrac{M_t}{M_{cta}}} \int_M^\infty \left(\frac{\beta_{ta}}{M'} + \frac{1}{M_{cta}}\right) M'^{-\beta_{ta}} e^{-\dfrac{M'}{M_{cta}}}\,\mathrm{d}M' \\ &= {M_t}^{\beta_{ta}} M^{-\beta_{ta}} e^{\dfrac{M_t - M}{M_{cta}}} \qquad (M_t \le M \le \infty)\,.\end{aligned} \tag{70}$$

We could confirm that eq. (70) was consistent with eq. (11) of Kagan (2002a). We can find eq. (70) was expressed by the product of eq. (20) of the G–R law and exponential taper. Note that this equation accords with eq. (20) when M_{cta} becomes ∞.

The log-likelihood $\ln L$ becomes

$$\begin{aligned}\ln\ L &= \sum_{i=1}^{N} \ln\ \phi\,(M_i) = \sum_{i=1}^{N} \ln\left(\left[\frac{\beta_{ta}}{M_i} + \frac{1}{M_{cta}}\right]\left(\frac{M_t}{M_i}\right)^{\beta_{ta}} e^{\dfrac{M_t - M_i}{M_{cta}}}\right) \\ &= N\beta_{ta} \ln M_t + \frac{1}{M_{cta}}\left(N M_t - \sum_{i\,=\,1}^{N} M_i\right) - \beta_{ta} \sum_{i\,=\,1}^{N} \ln M_i + \sum_{i\,=\,1}^{N} \ln\left(\frac{\beta_{ta}}{M_i} + \frac{1}{M_{cta}}\right).\end{aligned} \tag{71}$$

4.5.2 Magnitude representation of tapered G–R law

Here we convert the moment representation shown in Section 4.5.1 to a magnitude representation. The relation between $x = m - m_t$ and M is derived from eq. (2) as follows:

$$\frac{M}{M_t} = 10^{1.5x}. \tag{72}$$

When we set $x_{cta} = m_{cta} - m_t$, where m_{cta} is the magnitude representation of M_{cta},

$$\frac{M_{cta}}{M_t} = 10^{1.5x_{cta}}. \tag{73}$$

From eqs (70), (72), and (73),

$$\begin{aligned}\Phi\,(M) &= {M_t}^{\beta_{ta}} M^{-\beta_{ta}} e^{\dfrac{M_t - M}{M_{cta}}} = \left(\frac{M}{M_t}\right)^{-\beta_{ta}} e^{\dfrac{1 - \dfrac{M}{M_t}}{\dfrac{M_{cta}}{M_t}}} \\ &= \left(10^{1.5x}\right)^{-\beta_{ta}} e^{\dfrac{1 - 10^{1.5x}}{10^{1.5x_{cta}}}} = 10^{-b_{ta}x} e^{\dfrac{1 - 10^{1.5x}}{10^{1.5x_{cta}}}} = e^{-B_{ta}x} e^{\dfrac{1 - 10^{1.5x}}{10^{1.5x_{cta}}}} = \mathrm{F}\,(x)\,,\end{aligned} \tag{74}$$

where $\beta_{ta} = b_{ta}/1.5$ and $B_{ta} = b_{ta}\ \ln 10$.

The probability density function $f(x)$ of x is given by

$$f\ (x) = -\ \frac{\partial \mathrm{F}\,(x)}{\partial x} = e^{-B_{ta}x}\, e^{\dfrac{1 - 10^{1.5x}}{10^{1.5x_{cta}}}} \left(B_{ta} + \frac{10^{1.5x}}{10^{1.5x_{cta}}} 1.5 \ln 10\right). \tag{75}$$

This is consistent with eq. (7) in the G–R law when x_{cta} is ∞.

$N_m(m)$ was obtained from eqs (9) and (74), $N_M(M)$ was obtained from eqs (21) and (74), $n_m(m)$ was obtained from eqs (7) and (75), and $n_M(M)$ was obtained from eqs (22) and (69).

4.6 Seismic moment release rate

In Sections 4.1–4.5 we explained how the functions $\phi(M)$ and $\Phi(M)$ are derived under each of the five laws used in this study. In this section, we show how to estimate the total seismic moment release rate $\dot{M}_s$ from observed frequency–magnitude distributions using $\phi(M)$ and $\Phi(M)$ under each law. The total seismic moment release $M_s(M_t)$ from events greater or equal to M_t is given by

$$M_s\ (M_t) = \sum_{i=1}^{N} M_i\ = N_M\ (M_t)\,\mathrm{E}\ [M] = N_M\ (M_t)\, I_1\,(M_t)\,, \tag{76}$$

where $\mathrm{E}[M]$ is the mean of M_i and $I_1(M_t)$ is the first-order statistical moment of seismic moment from events greater or equal to M_t:

$$I_1\,(M_t) \approx \int_{M_t}^{M_{\max}} M'\phi\left(M'\right)\mathrm{d}M', \tag{77}$$

where the upper limit of integration $M_{\max}$ is M_{ctr} for the truncated G–R law, M_{cu} for Utsu's formula, and infinite for the G–R law, gamma distribution, and tapered G–R law.

From eq. (76), the total seismic moment release rate $\dot{M}_s(M_t)$ from events greater or equal to M_t is given by

$$\dot{M}_s\ (M_t) = \dot{N}_M\ (M_t)\, I_1\,(M_t)\,, \tag{78}$$

where $\dot{N}_M(M_t)$ is the occurrence rate of earthquakes greater or equal to M_t. Introducing an arbitrary seismic moment M_0 from eq. (21), $\dot{N}_M(M_t)$ is given by

$$\dot{N}_M\ (M_t) = \frac{\dot{N}_M\,(M_0)}{\Phi\,(M_0)}\ = \frac{\alpha_0}{\Phi\,(M_0)}\ , \tag{79}$$

where we set $\dot{N}_M\ (M_0) = \alpha_0$. The reason for incorporating M_0 and α_0 here is to refine the formulation of the total seismic moment release rate $\dot{M}_s$ as explained later.

Eq. (77) is combined with eqs (19), (36), (49), (61), (65) and (69), respectively, to obtain $I_1(M_t)$ under each law as follows.

For the G–R law,

$$\begin{aligned} I_1\,(M_t) &= \int_{M_t}^{\infty} M'\phi\,(M')\,\mathrm{d}M' = \int_{M_t}^{\infty} M'\beta {M_t}^{\beta} M'^{-\beta-1} dM' \\ &= \beta {M_t}^{\beta}\int_{M_t}^{\infty} M'^{-\beta}\mathrm{d}M' = \begin{cases} \beta {M_t}^{\beta}\left[\dfrac{1}{1-\beta}M'^{1-\beta}\right]_{M_t}^{\infty} = \dfrac{\beta}{\beta-1}M_t & (\beta > 1) \\ M_t\,[\ln\,M']_{M_t}^{\infty} = \infty & (\beta = 1) \\ \beta {M_t}^{\beta}\left[\dfrac{1}{1-\beta}M'^{1-\beta}\right]_{M_t}^{\infty} = \infty & (\beta < 1) \end{cases}. \end{aligned} \tag{80}$$

For the truncated G–R law,

$$\begin{aligned} I_1\,(M_t) &= \int_{M_t}^{M_{ctr}} M'\phi\,(M')\,\mathrm{d}M' = \int_{M_t}^{M_{ctr}} M' \frac{\beta_{tr} {M_t}^{\beta_{tr}} M'^{-\beta_{tr}-1}}{1-\left(\dfrac{M_{ctr}}{M_t}\right)^{-\beta_{tr}}}\mathrm{d}M' \\ &= \frac{\beta_{tr} {M_t}^{\beta_{tr}}}{1-\left(\dfrac{M_{ctr}}{M_t}\right)^{-\beta_{tr}}}\int_{M_t}^{M_{ctr}} M'^{-\beta_{tr}}\,\mathrm{d}M' = \begin{cases} \dfrac{\beta_{tr}}{1-\beta_{tr}}\dfrac{{M_t}^{\beta_{tr}} M_{ctr} - M_t {M_{ctr}}^{\beta_{tr}}}{{M_{ctr}}^{\beta_{tr}} - {M_t}^{\beta_{tr}}} & (\beta_{tr} \neq 1) \\ \dfrac{M_t M_{ctr}}{M_{ctr} - M_t}\ln\dfrac{M_{ctr}}{M_t} & (\beta_{tr} = 1) \end{cases}. \end{aligned} \tag{81}$$

For Utsu's formula,

$$
\begin{aligned}
I_1(M_t) &= \int_{M_t}^{M_{cu}} M'\phi(M')\,\mathrm{d}M' = \int_{M_t}^{M_{cu}} M'\beta_u M_t^{\beta_u} M'^{-\beta_u-1} \frac{\log \dfrac{M_{cu}}{M'}}{\log \dfrac{M_{cu}}{M_t} - \dfrac{1-\left(\dfrac{M_{cu}}{M_t}\right)^{-\beta_u}}{\beta_u \ln 10}} \mathrm{d}M' \\
&= \frac{\beta_u M_t^{\beta_u}}{\log \dfrac{M_{cu}}{M_t} - \dfrac{1-\left(\dfrac{M_{cu}}{M_t}\right)^{-\beta_u}}{\beta_u \ln 10}} \int_{M_t}^{M_{cu}} M'^{-\beta_u} \log \frac{M_{cu}}{M'} \mathrm{d}M' \\
&= \begin{cases} \dfrac{\beta_u}{1-\beta_u} \dfrac{\dfrac{M_t^{\beta_u} M_{cu}^{1-\beta_u} - M_t}{(1-\beta_u)\ln 10} - M_t \log \dfrac{M_{cu}}{M_t}}{\log \dfrac{M_{cu}}{M_t} - \dfrac{1-\left(\dfrac{M_{cu}}{M_t}\right)^{-\beta_u}}{\beta_u \ln 10}} & (\beta_u \neq 1) \\[2ex] \dfrac{M_t}{\log \dfrac{M_{cu}}{M_t} - \dfrac{1-\left(\dfrac{M_{cu}}{M_t}\right)^{-1}}{\ln 10}} \left\{ -\ln M_t \log \dfrac{M_{cu}}{M_t} + \dfrac{(\ln M_{cu})^2 - (\ln M_t)^2}{2 \ln 10} \right\} & (\beta_u = 1) \end{cases}
\end{aligned}
\tag{82}
$$

For the gamma distribution,

$$
\begin{aligned}
I_1(M_t) &= \int_{M_t}^{\infty} M'\phi(M')\,\mathrm{d}M' \\
&= \int_{M_t}^{\infty} M' \frac{M'^{-\beta_g-1} e^{-\frac{M'}{M_{cg}}}}{M_{cg}^{-\beta_g}\Gamma\left(-\beta_g, \dfrac{M_t}{M_{cg}}\right)} \mathrm{d}M' = \frac{M_{cg}\Gamma\left(1-\beta_g, \dfrac{M_t}{M_{cg}}\right)}{\Gamma\left(-\beta_g, \dfrac{M_t}{M_{cg}}\right)} \\
&= \frac{\beta_g M_t^{\beta_g}}{1-\beta_g} M_{cg}^{1-\beta_g} \Gamma\left(2-\beta_g, \frac{M_t}{M_{cg}}\right) e^{\frac{M_t}{M_{cg}}} G^{-1} - \frac{\beta_g}{1-\beta_g} M_t G^{-1} \quad (\beta_g \neq 1)
\end{aligned}
\tag{83}
$$

Here, the last complicated formulation was converted by using eq. (64). We omit the case of $\beta_g = 1$.

For the tapered G–R law,

$$
\begin{aligned}
I_1(M_t) &= \int_{M_t}^{\infty} M'\phi(M')\,\mathrm{d}M' \\
&= \int_{M_t}^{\infty} M' \left[\frac{\beta_{ta}}{M'} + \frac{1}{M_{cta}}\right] \left(\frac{M_t}{M'}\right)^{\beta_{ta}} e^{\frac{M_t - M'}{M_{cta}}} \mathrm{d}M' \\
&= \frac{M_t^{\beta_{ta}}}{1-\beta_{ta}} M_{cta}^{1-\beta_{ta}} \Gamma\left(2-\beta_{ta}, \frac{M_t}{M_{cta}}\right) e^{\frac{M_t}{M_{cta}}} - \frac{\beta_{ta}}{1-\beta_{ta}} M_t \quad (\beta_{ta} \neq 1)
\end{aligned}
\tag{84}
$$

We omit the case of $\beta_{ta} = 1$.

The total seismic moment release rate $\dot{M}_s$ (including events smaller than M_t) is obtained from eq. (78) by having the limit of M_t approach zero:

$$
\dot{M}_s = \lim_{M_t \to 0} \left[\dot{N}_M(M_t) I_1(M_t)\right] = \lim_{M_t \to 0} \left[\frac{\alpha_0}{\Phi(M_0)} I_1(M_t)\right]. \tag{85}
$$

Thus, we can derive $\dot{M}_s$ under each law as follows. From eqs (20) and (80), $\dot{M}_s$ for the G–R law is given by

$$\begin{aligned}\dot{M}_s &= \lim_{M_t\to 0}\left[\frac{\alpha_0}{\Phi(M_0)} I_1(M_t)\right]\\ &= \alpha_0 \lim_{M_t\to 0}\left[\frac{\int_{M_t}^{\infty} M'^{-\beta} dM'}{\int_{M_0}^{\infty} M'^{-\beta-1} dM'}\right]\\ &= \begin{cases}\alpha_0 \lim_{M_t\to 0}\left[\dfrac{\left[\dfrac{1}{1-\beta}(M')^{1-\beta}\right]_{M_t}^{\infty}}{\left[-\dfrac{1}{\beta}(M')^{-\beta}\right]_{M_0}^{\infty}}\right] = \infty \quad (\beta \neq 1)\\ \alpha_0 \lim_{M_t\to 0}\left[\dfrac{[\ln M']_{M_t}^{\infty}}{\left[-(M')^{-1}\right]_{M_0}^{\infty}}\right] = \infty \qquad (\beta = 1)\end{cases}\end{aligned} \tag{86}$$

In the following, we show cases in which the series of β terms (β_{tr}, β_u, β_g, β_{ta}) is less than 1 because $\dot{M}_s$ diverges when series of $\beta \geq 1$ for other four laws (but not for the G–R law).

From eqs (37) and (81), $\dot{M}_s$ for the truncated G–R law is given by

$$\begin{aligned}\dot{M}_s &= \lim_{M_t\to 0}\left[\frac{\alpha_0}{\Phi(M_0)} I_1(M_t)\right]\\ &= \lim_{M_t\to 0}\left[\frac{\alpha_0}{\dfrac{M_0{}^{-\beta_{tr}} - M_{ctr}{}^{-\beta_{tr}}}{M_t{}^{-\beta_{tr}} - M_{ctr}{}^{-\beta_{tr}}}} \frac{\beta_{tr}}{1-\beta_{tr}} \frac{M_t{}^{\beta_{tr}} M_{ctr} - M_t M_{ctr}{}^{\beta_{tr}}}{M_{ctr}{}^{\beta_{tr}} - M_t{}^{\beta_{tr}}}\right]\\ &= \frac{\alpha_0\beta_{tr}}{1-\beta_{tr}} \lim_{M_t\to 0}\left[\left(\frac{M_t}{M_0}\right)^{-\beta_{tr}} \frac{1-\left(\dfrac{M_t}{M_{ctr}}\right)^{\beta_{tr}}}{1-\left(\dfrac{M_0}{M_{ctr}}\right)^{\beta_{tr}}} M_t{}^{\beta_{tr}} M_{ctr}{}^{\beta_{tr}} \frac{M_{ctr}{}^{1-\beta_{tr}} - M_t{}^{1-\beta_{tr}}}{M_{ctr}{}^{\beta_{tr}} - M_t{}^{\beta_{tr}}}\right]\\ &= \frac{\alpha_0\beta_{tr}}{1-\beta_{tr}} M_0{}^{\beta_{tr}} M_{ctr}{}^{1-\beta_{tr}} \frac{M_{ctr}{}^{\beta_{tr}}}{M_{ctr}{}^{\beta_{tr}} - M_0{}^{\beta_{tr}}} \qquad (\beta_{tr} < 1)\end{aligned} \tag{87}$$

From eqs (50) and (82), $\dot{M}_s$ for Utsu's formula is given by

$$\begin{aligned}\dot{M}_s &= \lim_{M_t\to 0}\left[\frac{\alpha_0}{\Phi(M_0)} I_1(M_t)\right]\\ &= \lim_{M_t\to 0}\left[\frac{\alpha_0}{M_t{}^{\beta_u} M_0{}^{-\beta_u} \dfrac{\log\dfrac{M_{cu}}{M_0} - \dfrac{1-\left(\dfrac{M_{cu}}{M_0}\right)^{-\beta_u}}{\beta_u \ln 10}}{\log\dfrac{M_{cu}}{M_t} - \dfrac{1-\left(\dfrac{M_{cu}}{M_t}\right)^{-\beta_u}}{\beta_u \ln 10}}} \frac{\beta_u}{1-\beta_u} \frac{\dfrac{M_t{}^{\beta_u} M_{cu}{}^{1-\beta_u} - M_t}{(1-\beta_u)\ln 10} - M_t \log\dfrac{M_{cu}}{M_t}}{\log\dfrac{M_{cu}}{M_t} - \dfrac{1-\left(\dfrac{M_{cu}}{M_t}\right)^{-\beta_u}}{\beta_u \ln 10}}\right]\end{aligned} \tag{88}$$

$$
= \frac{\dfrac{\alpha_0\beta_u}{1-\beta_u}{M_0}^{\beta_u}}{\log\dfrac{M_{cu}}{M_0} - \dfrac{1-\left(\dfrac{M_{cu}}{M_0}\right)^{-\beta_u}}{\beta_u\ \ln 10}} \lim_{M_t\to 0}\left[\frac{{M_{cu}}^{1-\beta_u} - {M_t}^{1-\beta_u}}{(1-\beta_u)\ \ln 10} - {M_t}^{1-\beta_u}\log\frac{M_{cu}}{M_t}\right]
$$

$$
= \frac{\alpha_0\beta_u}{1-\beta_u}{M_0}^{\beta_u}\frac{\dfrac{{M_{cu}}^{1-\beta_u}}{(1-\beta_u)\ \ln 10}}{\log\dfrac{M_{cu}}{M_0} - \dfrac{1-\left(\dfrac{M_{cu}}{M_0}\right)^{-\beta_u}}{\beta_u\ \ln 10}}
$$

$$
= \frac{\alpha_0{\beta_u}^2}{(1-\beta_u)^2}{M_0}^{\beta_u}{M_{cu}}^{1-\beta_u}\frac{1}{\beta_u\ \ln 10\cdot\log\dfrac{M_{cu}}{M_0} - \left\{1-\left(\dfrac{M_{cu}}{M_0}\right)^{-\beta_u}\right\}} \qquad (\beta_u < 1)
$$

From eqs (62) and (83), $\dot{M}_s$ for the gamma distribution is given by

$$
\begin{aligned}
\dot{M}_s &= \lim_{M_t\to 0}\left[\frac{\alpha_0}{\Phi(M_0)}I_1(M_t)\right] = \lim_{M_t\to 0}\left[\frac{\alpha_0}{\dfrac{\Gamma\left(-\beta_g,\ \dfrac{M_0}{M_{cg}}\right)}{\Gamma\left(-\beta_g,\ \dfrac{M_t}{M_{cg}}\right)}}\frac{M_{cg}\Gamma\left(1-\beta_g,\ \dfrac{M_t}{M_{cg}}\right)}{\Gamma\left(-\beta_g,\ \dfrac{M_t}{M_{cg}}\right)}\right] \\
&= \frac{\alpha_0 M_{cg}\Gamma\left(1-\beta_g\right)}{\Gamma\left(-\beta_g,\ \dfrac{M_0}{M_{cg}}\right)} \\
&= \frac{\alpha_0\beta_g}{1-\beta_g}{M_0}^{\beta_g}{M_{cg}}^{1-\beta_g}\Gamma\left(2-\beta_g\right)\frac{e^{\frac{M_0}{M_{cg}}}}{1-\left(\dfrac{M_0}{M_{cg}}\right)^{\beta_g}e^{\frac{M_0}{M_{cg}}}\Gamma\left(1-\beta_g,\ \dfrac{M_0}{M_{cg}}\right)} \qquad (\beta_g < 1)
\end{aligned} \tag{89}
$$

The last complicated formulation in eq. (89) is converted by using eq. (64), corresponding to eqs (7) and (10) of Kagan (2002b). (Note that the appearance of C in eq. (10) of Kagan (2002b) is an error, being unnecessary.) The term $\Gamma(1-\beta_g)$ is the gamma function, defined as

$$
\Gamma(a) \equiv \int_0^\infty t^{a-1}e^{-t}dt \qquad (a > 0)\ . \tag{90}
$$

From eqs (70) and (84), $\dot{M}_s$ for the tapered G–R law is given by

$$
\begin{aligned}
\dot{M}_s &= \lim_{M_t\to 0}\left[\frac{\alpha_0}{\Phi(M_0)}I_1(M_t)\right] \\
&= \lim_{M_t\to 0}\left[\frac{\alpha_0}{{M_t}^{\beta_{ta}}{M_0}^{-\beta_{ta}}e^{\frac{M_t-M_0}{M_{cta}}}}\frac{{M_t}^{\beta_{ta}}}{1-\beta_{ta}}{M_{cta}}^{1-\beta_{ta}}\Gamma\left(2-\beta_{ta},\ \frac{M_t}{M_{cta}}\right)e^{\frac{M_t}{M_{cta}}} - \frac{\beta_{ta}}{1-\beta_{ta}}M_t\right] \\
&= \frac{\alpha_0}{1-\beta_{ta}}{M_0}^{\beta_{ta}}{M_{cta}}^{1-\beta_{ta}}e^{\frac{M_0}{M_{cta}}}\Gamma\left(2-\beta_{ta}\right) \qquad (\beta_{ta} < 1)
\end{aligned} \tag{91}
$$

As mentioned earlier, M_0 is an arbitrary seismic moment and $\alpha_0 = \dot{N}_M(M_0)$. In this study, we adopted a completeness magnitude of m 5.8 such that $M_0 = M_{\mathrm{m}5.8}$ and $\dot{N}_M(M_0) = \dot{N}_M(M_{\mathrm{m}5.8})$.

4.7 Estimation of parameters

Section 4.6 introduced the analytical solutions of total seismic moment release rate $\dot{M}_s$ under each law in eqs (86)–(89) and (91), respectively. By assuming that $\dot{M}_s$ and $\dot{M}_T$ are equal, the set of M_c terms (M_{ctr}, M_{cu}, M_{cg}, M_{cta}) can be represented by a corresponding set of β terms

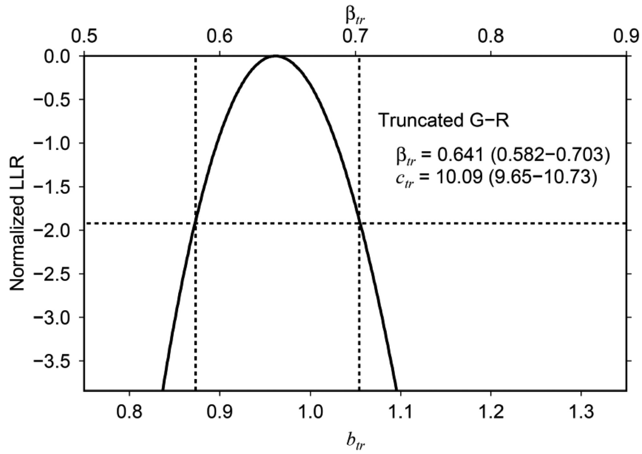

Figure 2. Logarithm of the likelihood ratio (LLR) function, normalized to a maximum of 0.0, for the truncated G–R law. The values in parentheses indicate both ends of the 95 per cent confidence level, also marked by dashed vertical lines for β_{tr}.

$(\beta_{tr},\ \beta_u,\ \beta_g,\ \beta_{ta})$. For example, the truncated G–R law is written as

$$\dot{M}_s = \frac{\alpha_0 \beta_{tr}}{1-\beta_{tr}} M_0{}^{\beta_{tr}} M_{ctr}{}^{1-\beta_{tr}} \frac{M_{ctr}{}^{\beta_{tr}}}{M_{ctr}{}^{\beta_{tr}} - M_0{}^{\beta_{tr}}} = \dot{M}_T \ = \ \chi \mu W L V_{pl}, \tag{92}$$

where all variables except for M_{ctr} and β_{tr} are known. We used a grid search to estimate values of the β term that maximized the log-likelihood $\ln L$ under each law, and then obtained the corresponding set of M_c terms simultaneously. The magnitude representations corresponding to the M_c terms are denoted by $c_{tr},\ c_u,\ c_g,$ and c_{ta}.

Finally, we address estimates of error. Wilks (1962) showed that the logarithm of the likelihood ratio (LLR) was distributed as half of a chi-squared distribution with k degrees of freedom $(\frac{1}{2}\chi_k^2)$. LLR in our case follows $\frac{1}{2}\chi_1^2$, because all of the laws we used, other than the G–R law, have one degree of freedom through $\dot{M}_s = \dot{M}_T$. The value of χ_1^2 is 3.84 when the upper probability becomes 5 per cent; that is, χ_1^2 values no greater than 3.84 correspond to the 95 per cent confidence level. When the LLR $(\frac{1}{2}\chi_1^2)$ is normalized to a maximum of 0.0, the 95 per cent confidence level nearly corresponds to the contour value –1.92 (half of –3.84) (*cf.* Kagan 1997 for two degrees of freedom). Fig. 2 shows the LLR function normalized for applying the truncated G–R law to the 1977–2017 seismicity record.

We judged the relative quality of the resulting models by the Akaike Information Criterion or AIC (Akaike 1974):

$$\mathrm{AIC} \ = \ -2\ln L_{\max} + 2k, \tag{93}$$

where $\ln L_{\max}$ is the maximum log-likelihood and k is the degree of freedom. The AIC grows small as the quality of the model increases.

5 RESULTS

Fig. 3 shows the frequency–magnitude distribution and the approximate cumulative distribution under each law for the periods 1977–2010 (Period 1) and 1977–2017 (Period 3). Note that the approximate cumulative distribution under the G–R law was obtained by using the maximum-likelihood method of eq. (15) because $\dot{M}_s$ diverged in eq. (86). All five laws fit the existing observations well; however, because the Earth is not infinite, the frequency–magnitude distribution departs from the straight line of the G–R law with the accumulation of seismic data. Although it is hard to choose from among these four curves, we can estimate the parameters of each law on the basis of seismic moment conservation even from limited seismic data. Fig. 4 shows the moment-magnitude distribution, in which the vertical axis in Fig. 3 is replaced with the product of the number of earthquakes in each magnitude bin and moment release per year. The area enclosed by each curve corresponds to $\dot{M}_s$ and is equal to $\dot{M}_T$. One can see that although small events greatly outnumber large events in Fig. 3, the large events greatly outweigh the small events in Fig. 4. The difference between these theoretical curves and the observation data above m 7 should become smaller with the accumulation of seismic data. The estimated maximum magnitude based on the 1977–2010 record (Period 1) is m 9.92 under the truncated G–R law and m 10.65 under Utsu's formula, and the corner parameter is 10.00 under the gamma distribution and 9.65 under the tapered G–R law (Fig. 3a). The 2011 Tohoku earthquake occurred just after this period. Because its observed magnitude (m 9.2) was less than the maximum magnitudes estimated by the truncated G–R law (m 9.92) and Utsu's formula (m 10.65), it fits within the expected range. Table 2 lists the parameters of each law for each of the three periods we analysed.

Fig. 5 shows the maximum magnitudes $(c_{tr},\ c_u)$ estimated by the truncated G–R law and Utsu's formula for the three time periods. As shown in Fig. 3, Utsu's formula deviates from the G–R law at smaller magnitude than others. Accordingly, the maximum magnitude (c_u) estimated by Utsu's formula tends to become large in order to compensate the total seismic moment release (Fig. 4). On the other hand, for the truncated G–R law, the seismic moment release rate suddenly falls from certain positive value into exact zero at the maximum magnitude.

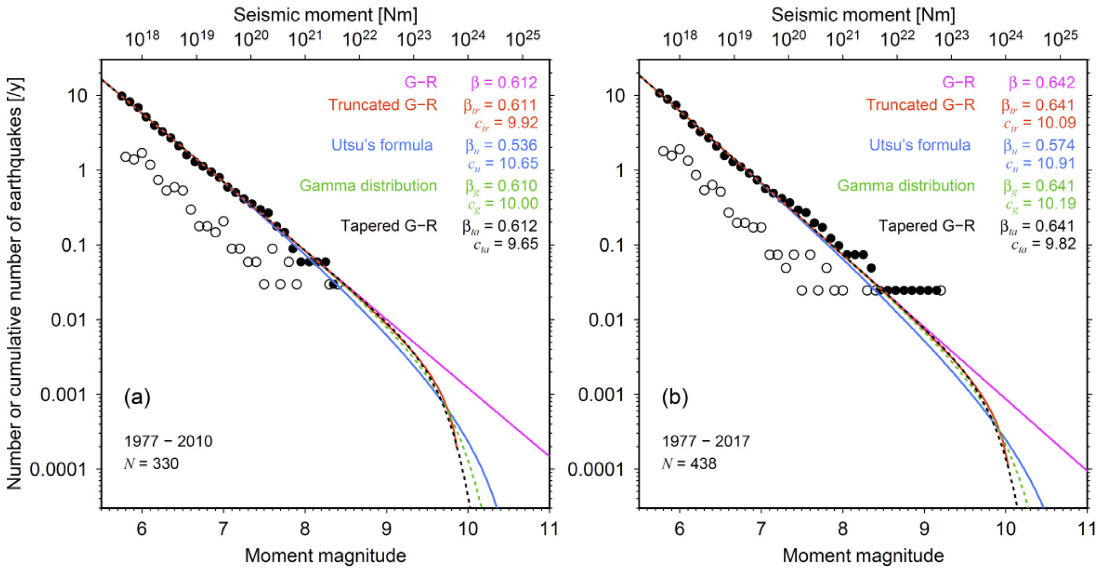

Figure 3. Frequency–magnitude distribution for the periods (a) 1977–2010 and (b) 1977–2017. Open circles denote the number of events per magnitude bin of 0.1 width, and solid circles are the cumulative number of events. Magenta, red, blue, broken green, and broken black lines represent the approximate cumulative event distributions under the G–R law, truncated G–R law, Utsu's formula, gamma distribution and tapered G–R law, respectively. The five models are pretty much indistinguishable from the actual earthquake data. Estimated parameters for these laws are listed at upper right.

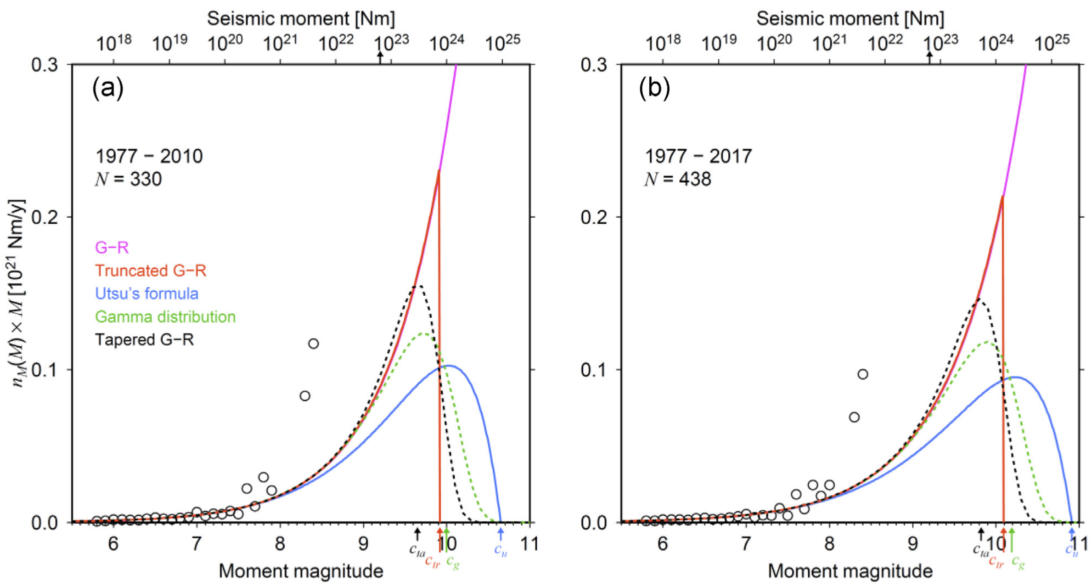

Figure 4. Moment–magnitude distribution for the periods (a) 1977–2010 and (b) 1977–2017. Open circles denote the annual total moment of events per magnitude bin of 0.1 width. Note that the 2011 Tohoku earthquake (*m* 9.2) plots outside this figure (an arrow at the top of the figure). Magenta, red, blue, broken green, and broken black lines represent the approximate cumulative distribution by the G–R law, truncated G–R law, Utsu's formula, gamma distribution and tapered G–R law, respectively.

Therefore, the maximum magnitude estimated by the truncated G–R law may show nearly the infimum of the possible maximum magnitude. The consistent difference of 0.7–0.8 between the results of the two laws (Fig. 5) arises from their difference in seismic moment release curves. The estimated maximum magnitudes changed little after the 2011 Tohoku earthquake, and we cannot confirm the dependence of the maximum magnitude on the data period, as Kagan & Jackson (2013) proposed.

Eq. (92) can be rewritten simply as follows: $\dot{M}_s \sim \alpha_0\, M_{ctr}{}^{1-\beta_{tr}} =$ const when $M_{ctr} \gg M_0$ and $\beta_{tr} \sim 0.6$. Thus, generally, the maximum magnitude ($c_{tr},\ c_u$) becomes smaller when seismicity rate α_0 increases, and the β terms ($\beta_{tr},\ \beta_u$) decrease as a consequence of seismic moment conservation. Because the seismicity rate with $m \geq 5.8$ is 10.7 events yr^{-1} (=438 events/41 yr) in Period 3 and 9.7 events yr^{-1} (=330 events/34 yr) in Period 1, the maximum magnitude tends to decrease, yet because the β terms in Period 3 are larger than those in Period 1, the maximum magnitude tends to increase. The result of these counteracting effects is that the maximum magnitude is only slightly larger in Period 3 than in Period 1.

Table 2. Parameters estimated in this study. Values in parentheses are upper and lower limits of the 95 per cent confidence interval.

Period	1977–2010		1977–2013		1977–2017	
Events	330		408		438	
Seismicity rate (events/y)	9.7		11.0		10.7	
$m_{\max\_obs}$[a]	8.4		9.2		9.2	
	β[b]	c[c]	β[b]	c[c]	β[b]	c[c]
Truncated G–R	0.611 (0.545–0.680)	9.92 (9.49–10.55)	0.630 (0.569–0.693)	9.97 (9.55–10.58)	0.641 (0.582–0.703)	10.09 (9.65–10.73)
Utsu's equation	0.536 (0.457–0.618)	10.65 (10.11–11.44)	0.560 (0.488–0.635)	10.76 (10.22–11.53)	0.574 (0.503–0.647)	10.91 (10.34–11.71)
Gamma distribution	0.610 (0.543–0.679)	10.00 (9.56–10.64)	0.630 (0.571–0.693)	10.07 (9.64–10.68)	0.641 (0.583–0.703)	10.19 (9.74–10.83)
Tapered G–R	0.612 (0.547–0.680)	9.65 (9.20–10.30)	0.629 (0.571–0.693)	9.69 (9.26–10.33)	0.641 (0.582–0.703)	9.82 (9.36–10.48)

[a]Maximum observed magnitude during this time period.
[b]The β terms in the truncated G–R law, Utsu's equation, gamma distribution, and tapered G–R law are β_{tr}, β_u, β_g, and β_{ta}, respectively.
[c]The c terms in the truncated G–R law, Utsu's equation, gamma distribution, and tapered G–R law are c_{tr}, c_u, c_g, and c_{ta}, respectively.

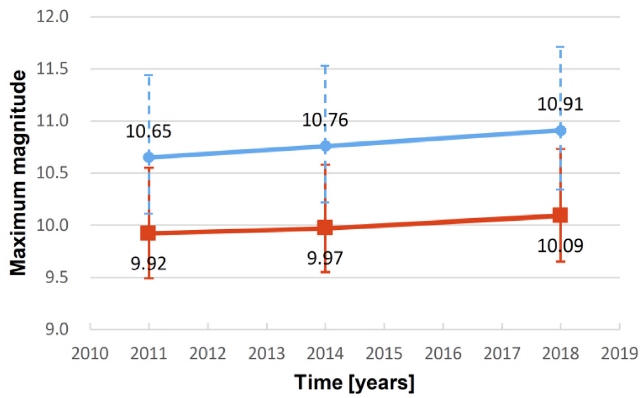

Figure 5. Maximum magnitudes estimated by the truncated G–R law (red) and Utsu's equation (blue) for Periods 1–3. Vertical bars indicate both ends of the 95 per cent confidence level.

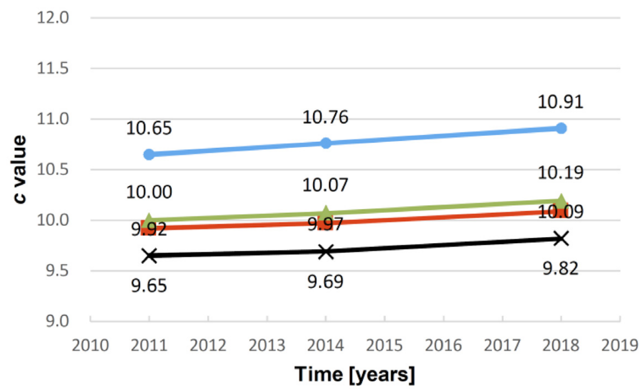

Figure 6. Temporal variation of c values. Red, blue, green and black symbols represent c values (c_{tr}, c_u, c_g, and c_{ta}) estimated by the truncated G–R law, Utsu's formula, gamma distribution and tapered G–R law, respectively. Note that c_g and c_{ta} are mere parameters defining a corner magnitude not the maximum magnitude.

The estimation error of the parameters is little changed after the 2011 Tohoku earthquake (Table 2). The AIC value for Utsu's formula is 0.3 smaller than that of the truncated G–R law before the 2011 Tohoku earthquake, and 0.4 larger than that after the event. Although there is little difference in the AICs, the range of the 95 per cent confidence level of the maximum magnitude for the result of the truncated G–R law is consistently about 0.3 narrower than that for Utsu's formula (Table 2).

Fig. 6 shows the maximum magnitudes estimated by all four laws for the three time periods. The results from the gamma distribution (c_g) are very close to those from the truncated G–R law, and those from the tapered G–R law (c_{ta}) are about 0.3 smaller.

6 DISCUSSION

As mentioned in Section 1, the parameters of the gamma distribution and tapered G–R law do not exactly prescribe the maximum magnitude, whereas the truncated G–R law and Utsu's formula have parameters clearly corresponding to the maximum magnitude. Accordingly, we confine our discussion to the latter two laws, paying particular attention to parameters estimated by the truncated G–R law due to their smaller error.

6.1 Recurrence interval

The maximum magnitude estimated by the truncated G–R law became 10.09 after the 2011 Tohoku earthquake (Fig. 5). The recurrence interval of earthquakes with $m \geq 9.95$ is about 4000 yr along the Japan-Kuril-Kamchatka trench because the expected occurrence rate is 0.01 event in the 41-yr period 1977–2017 (Fig. 3b). Although the tectonic moment accumulated over 4000 yr corresponds to an earthquake of m 10.6, we can avoid overestimating the maximum magnitude by taking into consideration the seismic moment release of smaller earthquakes, as noted in this study.

The recurrence interval of giant earthquakes of $m \geq 9.15$, similar to the 2011 Tohoku earthquake, is about 200 yr along the Japan-Kuril-Kamchatka trench because the occurrence rate is 0.2 events in 41 yr (Fig. 3b). The Earthquake Research Committee (2011) estimated a mean recurrence interval of 600 yr for such earthquakes because studies of tsunami deposits indicate that five great earthquakes comparable to the 2011 Tohoku earthquake, such as the 869 Jogan earthquake (Namegaya & Satake 2014), occurred in the past 2500 yr. On the other hand, because the area off the Tohoku district (500 km × 200 km) corresponds to about one-sixth of the length of the Japan-Kuril-Kamchatka trench, the recurrence interval off the Tohoku district would be 1200 yr if the spatial occurrence rate is constant, or twice that estimated by the Earthquake Research Committee (2011). However, there is scant information for a 15th-century earthquake between the 869 and 2011 events. If the magnitude of the 15th-century earthquake was a little less than the Jogan and Tohoku earthquakes, for example m 8.85, the recurrence interval estimated by seismic moment conservation would be about 600 yr, consistent with the estimate of the Earthquake Research Committee (2011).

A great earthquake of m 8.8 occurred off Hokkaido along the Kuril trench in the 17th century (Ioki & Tanioka 2016), and tsunami deposit studies suggest that the mean recurrence interval of events this size is about 400 yr (Earthquake Research Committee 2017). The recurrence interval of earthquakes with $m \geq 8.75$ is about 80 yr for the whole study region because the expected occurrence rate is 0.54 event in 41 yr under the truncated G–R law. Because the area off Hokkaido (300 km × 130 km) is about 1/15 of the whole area, the recurrence interval there would be 1200 yr if the spatial occurrence rate is constant, or three times the estimate by the Earthquake Research Committee (2017). Reversely, the magnitude corresponding to earthquakes whose mean recurrence interval is about 400 yr is estimated 8.30 or larger, which is a little smaller than 8.75, under the truncated G–R law. Thus, there is such a difference between our mode and the investigation of Earthquake Research Committee (2011, 2017), which may indicate the estimation error range of our method.

Murotani *et al.* (2013) introduced various scaling relations from m 7–8 interplate earthquakes around Japan and other global great earthquakes. Their scaling relationship between rupture area S (km^2) and seismic moment, $S = 1.34 \times 10^{-10} M^{2/3}$, indicates that the magnitude of an earthquake that ruptures the entire study region would be m 9.8, which is slightly smaller than our estimate. It is uncertain whether this scaling relationship can be extrapolated to events of m 10 because Murotani *et al.* (2013) used only four $m \geq 9$ events, the largest of which was m 9.2. Furthermore, the scaling relationship increases maximum magnitudes without limit as the rupture area increases, whereas our method does not always result in large magnitudes. The maximum magnitude from Utsu's formula may be an overestimate because it departs steeply from the scaling relationship of Murotani *et al.* (2013).

6.2 Error estimates

In this section, we discuss error estimates. The long-term seismic moment rate $\dot{M}_s$ depends on the seismicity rate α_0, the β value, the shape and temporal stability of the frequency–magnitude distribution and the magnitude uncertainty of the earthquakes in the catalogue, etc. On the other hand, the tectonic moment rate $\dot{M}_T$ depends on downdip width, the coupling coefficient, the 3-D geometry of the subduction zone, the rigidity, the temporal stability of plate convergence rate, etc. All of these are subject to estimation errors and contribute the uncertainty of the maximum magnitude. Here, we discuss about the major error factors among those such as: the interplate coupling rate χ, the plate convergence rate V_{pl}, the seismicity rate α_0, the β value and the width of the seismogenic zone W.

Our results discussed so far have been based on a single interplate coupling rate ($\chi = 70$ per cent). However, χ is highly arbitrary because the observation periods available to estimate it are insufficient. Therefore, we investigated this issue by making calculations with χ values ranging from 10 to 100 per cent at intervals of 10 per cent. Fig. 7 shows that the maximum magnitude increases monotonically with χ, and the maximum magnitude estimated by the truncated G–R law is consistently 0.7–0.9 smaller than that estimated by Utsu's formula. Because the truncated G–R law estimates a maximum magnitude smaller than the largest observed event (m 9.2, the 2011 Tohoku earthquake) for χ values of 10 and 20 per cent, the actual value of χ is probably >30 per cent under the truncated G–R law. This inference may help to constrain χ. In addition, the maximum magnitude under the truncated G–R law is 10.38, even at $\chi = 100$ per cent.

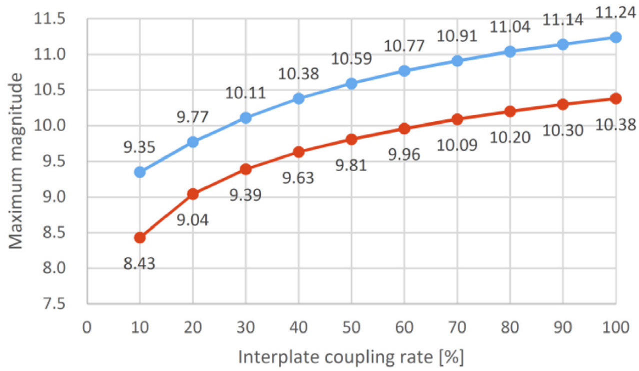

Figure 7. Maximum magnitude for different interplate coupling rates under the truncated G–R law (red) and Utsu's equation (blue) for the period 1977–2017.

We adopted the MORVEL model (DeMets *et al.* 2010) as a source of plate convergence rate V_{pl}, but other models could be proposed. For example, in the REVEL model (Shella *et al.* 2002), V_{pl} off the Tohoku district is 7.22–7.42 cm yr^{-1}, or 80 per cent of its value in MORVEL (9.20–9.32 cm yr^{-1}). Accordingly, $\dot{M}_T$ and the estimated maximum magnitude would also decrease.

As mentioned in Section 5, the maximum magnitude is inversely correlated to seismicity rate. Although the seismicity rate of $m \geq 5.8$ earthquakes has a range of 9.7–11.0 events yr^{-1} during Periods 1–3, it may depart from that range over longer periods. For example, if β_{tr} is set at its value in Period 3 and α_0 ranges from 5 to 15, the maximum magnitude has a range of 10.70–9.82 from eq. (92) corresponding to α_0. Similarly, the maximum magnitude is correlated to β_{tr}. Although β_{tr} has a range of 0.611–0.641 in Periods 1–3, it may depart from that range over longer periods. For example, if α_0 is set at its value in Period 3 and β_{tr} ranges from 0.5 to 0.7, the maximum magnitude has a range of 9.20–10.69.

Kagan & Jackson (2013) applied the gamma distribution to the seismic record before the 2011 Tohoku earthquake and estimated the maximum magnitude (strictly speaking c_g) off the Tohoku district to be 9.26 ± 0.29. They assumed a seismogenic zone with a width W of 104 km and a dip angle of 14° between 14 and 40 km depth. However, many studies have reported that the rupture extended as far as the trench axis (Sakaguchi *et al.* 2011; Sun *et al.* 2017). In addition, as mentioned in Section 3, the lower limit of the interplate seismogenic zone was at 50–60 km depth (e.g. Kita *et al.* 2010). From these considerations, $W = 104$ km must be an underestimate because the width of the seismogenic zone off the Tohoku district is instead 249 km. We can recalculate c_g by adopting the parameters of Kagan & Jackson (2013) ($\chi = 50$ per cent, $= 49$ GPa, $L = 620$ km), adding $V_{pl} = 11.15$ cm yr^{-1} (derived by inversion from $\dot{M}_T = 1.76 \times 10^{20}$ Nm yr^{-1} in Zone number 12a in Kagan and Jackson's Table 1, although this value is overestimated), and adopting a value of 249 km for W. For Period 1 off the Tohoku district, c_g then becomes 9.92 and the range of 95 per cent confidence becomes 9.19–11.43.

As this exercise shows, the maximum magnitude depends not only on the model selected, but also on the parameters (χ, μ, W, L and V_{pl}) that constitute the tectonic moment. It is essential to accumulate more seismic data and also make more precise estimates of tectonic moment to improve our estimates of maximum magnitude.

7 SUMMARY

We estimated the maximum magnitude of events that can be expected to occur in the subduction zone along the Japan-Kuril-Kamchatka trench on the basis of the seismic moment conservation principle. We applied the truncated G–R law and Utsu's formula, which contain a parameter clearly corresponding to the maximum magnitude, to an earthquake dataset that covers a longer period than previous studies and includes the 2011 Tohoku earthquake.

The estimated maximum magnitude was ∼10 under the truncated G–R law and ∼11 under Utsu's formula. The estimated maximum magnitudes increased only slightly when based on the longer data period. The maximum magnitude was about one unit greater under Utsu's formula than under the truncated G–R law because the seismic moment release in Utsu's formula is smaller than that of the truncated G–R law around m 9.

Although the performances of these two models were little different in terms of AIC, the size of the 95 per cent confidence level was consistently smaller for the truncated G–R law than for Utsu's formula. The maximum magnitude under Utsu's formula may be overestimated because it departs from the scaling relationship of Murotani *et al.* (2013).

The estimated recurrence interval for events greater than m 9.95 is about 4000 yr under the truncated G–R law. The recurrence interval of $m \geq 9.15$ earthquakes, the size of the 2011 Tohoku earthquake, off the Tohoku district is estimated to be 1200 yr, or twice the length estimated from tsunami deposits. However, if the magnitude of the 15th-century earthquake is no more than one unit smaller than the 869 Jogan and 2011 Tohoku earthquakes (for example m 8.85), the recurrence interval from the truncated G–R law becomes consistent with the recurrence interval estimated from tsunami deposits at about 600 yr. The recurrence interval of $m \geq 8.75$ earthquakes off Hokkaido along the Kuril trench is estimated to be 1200 yr, or three times the recurrence interval estimated from tsunami deposits. The recurrence interval of $m \geq 8.30$ events becomes 400 yr under the truncated G–R law.

Under the scaling relationship, the maximum magnitude increases without limit as the size of the convergent margin generating interplate earthquakes grows. But our method does not always result in large magnitudes considering the moment release of smaller earthquakes. The maximum magnitude for the Japan-Kuril-Kamchatka trench is no greater than 10.38, even for the unrealistic case of a 100 per cent interplate coupling rate.

The maximum magnitude for a given seismogenic region depends on the parameters that constitute tectonic moment as well as the model used to estimate it. To improve estimates of the maximum magnitude, it is essential to both accumulate seismic data and achieve more precise estimates of tectonic moment.

ACKNOWLEDGEMENTS

This manuscript was greatly improved by careful reviews of anonymous reviewers. We thank the Global CMT Project for providing the CMT catalogue. Isodepth contour data were taken from the U.S. Geological Survey at http://earthquake.usgs.gov/data/slab/and from the Japan Meteorological Agency at http://www.mri-jma.go.jp/Dep/st/member/fhirose/en/en.PlateData.html. Plate convergence rate data were calculated by the Plate Motion Calculator on the UNAVCO website (https://www.unavco.org/). Figures were prepared using Generic Mapping Tools (Wessel *et al.* 2013). The data for this paper are available by contacting the corresponding author at fhirose@mri-jma.go.jp.

REFERENCES

Akaike, H., 1974. A new look at the statistical model identification, *IEEE Trans. Autom. Contr.,* **AC-19,** 716–723.

Bird, P., 2003. An updated digital model of plate boundaries, *Geochem. Geophys. Geosyst.,* **4**.

Bird, P. & Kagan, Y.Y., 2004. Plate-tectonic analysis of shallow seismicity: apparent boundary width, beta, corner magnitude, coupled lithosphere thickness, and coupling in seven tectonic settings, *Bull. seism. Soc. Am.,* **94,** 2380–2399.

DeMets, C., Gordon, R.G. & Argus, D.F., 2010. Geologically current plate motions, *Geophys. J. Int.,* **181,** 1–80.

Dziewonski, A.M., Chou, T.-A. & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.,* **86,** 2825–2852.

Earthquake Research Committee, 2011. Long-term evaluation of earthquakes from off the Boso to off the Sanriku, the second edition (in Japanese), < https://www.jishin.go.jp/main/chousa/kaikou_pdf/sanriku_boso_4.pdf>, (2018-04-17).

Earthquake Research Committee, 2017. Long-term evaluation of earthquakes along the Kuril trench, the third edition (in Japanese), < https://www.jishin.go.jp/main/chousa/kaikou_pdf/chishima3.pdf> (2018-06-19).

Ekström, G., Nettles, M. & Dziewonski, A.M., 2012. The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth planet. Inter.,* **200–201,** 1–9.

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

Hanks, T.C. & Kanamori, H., 1979. A moment magnitude scale, *J. geophys. Res.,* **84,** 2348–2350.

Hashimoto, C., Noda, A. & Matsu'ura, M., 2012. The Mw9.0 northeast Japan earthquake: total rupture of a basement asperity, *Geophys. J. Int.,* **189,** 1–5.

Hayes, G.P., Wald, D.J. & Johnson, R.L., 2012. Slab1.0: a three-dimensional model of global subduction zone geometries, *J. geophys. Res.,* **117,** B01302.

Hirose, F., Miyaoka, K., Hayashimoto, N., Yamazaki, T. & Nakamura, M., 2011. Outline of the 2011 off the Pacific Coast of Tohoku Earthquake (Mw9.0)- Seismicity: foreshocks, mainshock, aftershocks, and induced activity, *Earth Planets Space,* **63,** 513–518.

Igarashi, T., Matsuzawa, T. & Hasegawa, A., 2003. Repeating earthquakes and interplate aseismic slip in the northeastern Japan subduction zone, *J. geophys. Res.,* **108,** doi:10.1029/2002JB001920.

Ioki, K. & Tanioka, Y., 2016. Re-estimated fault model of the 17th century great earthquake off Hokkaido using tsunami deposit data, *Earth planet. Sci. Lett.,* **433,** 133–138.

Kagan, Y.Y., 1991. Seismic moment distribution, *Geophys. J. Int.,* **106,** 123–134.

Kagan, Y.Y., 1997. Seismic moment-frequency relation for shallow earthquakes: regional comparison, *J. geophys. Res.,* **102,** 2835–2852.

Kagan, Y.Y., 2002a. Seismic moment distribution revisited: I. Statistical results, *Geophys. J. Int.,* **148,** 520–541.

Kagan, Y.Y., 2002b. Seismic moment distribution revisited: II. Moment conservation principle, *Geophys. J. Int.,* **149,** 731–754.

Kagan, Y.Y. & Jackson, D.D., 2013. Tohoku Earthquake: A Surprise?, *Bull. seism. Soc. Am.,* **103,** 1181–1194.

Kanamori, H., 1977. The energy release in great earthquakes, *J. geophys. Res.,* **82,** 2981–2987.

Kita, S., Okada, T., Hasegawa, A., Nakajima, J. & Matsuzawa, T., 2010. Anomalous deepening of a seismic belt in the upper-plane of the double seismic zone in the Pacific slab beneath the Hokkaido corner: possible evidence for thermal shielding caused by subducted forearc crust materials, *Earth planet. Sci. Lett.,* **290,** 415–426.

Mabuchi, H., Ohtake, M. & Sato, H., 2002. Global distribution of maximum earthquake magnitude Mc based on a modified G-R model of magnitude-frequency distribution (in Japanese with English abstract), *J. Seismol. Soc. Jpn.,* **55,** 261–273.

Matsuzawa, T., 2013. What is the largest earthquake we should prepare for? (in Japanese), *Report of the Coordinating Committee for Earthquake Prediction,* **89,** 446–449.

Murotani, S., Satake, K. & Fujii, Y., 2013. Scaling relations of seismic moment, rupture area, average slip, and asperity size for M∼9 subduction-zone earthquakes, *Geophys. Res. Lett.,* **40,** 5057–5074.

Nakajima, J. & Hasegawa, A., 2006. Anomalous low-velocity zone and linear alignment of seismicity along it in the subducted Pacific slab beneath Kanto, Japan: reactivation of subducted fracture zone?, *Geophys. Res. Lett.,* **33,** L16309.

Namegaya, Y. & Satake, K., 2014. Reexamination of the A.D. 869 Jogan earthquake size from tsunami deposit distribution, simulated flow depth, and velocity, *Geophys. Res. Lett.,* **41,** 2297–2303.

Rong, Y., Jackson, D.D., Magistrale, H. & Goldfinger, C., 2014. Magnitude limits of subduction zone earthquakes, *Bull. seism. Soc. Am.,* **104,** 2359–2377.

Saito, M., Kikuchi, M. & Kudo, K., 1973. Analytical solution of "Go-game model" of earthquake (in Japanese with English abstract), *J. Seismol. Soc. Jpn.,* **26,** 19–25.

Sakaguchi, A. *et al.*, 2011. Seismic slip propagation to the updip end of plate boundary subduction interface faults: vitrinite reflectance geothermometry on Integrated Ocean Drilling Program NanTro SEIZE cores, *Geology,* **39,** 395–398.

Savage, J.C., 1983. A dislocation model of strain accumulation and release at a subduction zone, *J. geophys. Res.,* **88,** 4984–4996.

Shella, G.F., Dixon, T.H. & Mao, A., 2002. REVEL: a model for recent plate velocities from space geodesy, *J. geophys. Res.,* **107**(B4), 2081.

Sun, T., Wang, K., Fujiwara, T., Kodaira, S. & He, J., 2017. Large fault slip peaking at trench in the 2011 Tohoku-oki earthquake, *Nature Comms.,* **8,** doi:10.1038/ncomms14044.

Uchida, N. & Matsuzawa, T., 2011. Coupling coefficient, hierarchical structure, and earthquake cycle for the source area of the 2011 off the Pacific coast of Tohoku earthquake inferred from small repeating earthquake data, *Earth Planets Space,* **63,** 675–679.

Uchida, N. & Matsuzawa, T., 2013. Pre- and postseismic slow slip surrounding the 2011 Tohoku-oki earthquake rupture, *Earth planet. Sci. Lett.,* **374,** 81–91.

Utsu, T., 1974. A three-parameter formula for magnitude distribution of earthquakes, *J. Phys. Earth,* **22,** 71–85.

Utsu, T., 1978. Estimation of parameters in formulas for frequency-magnitude relation of earthquake occurrence: in cases involving a parameter c for the maximum magnitude (in Japanese with English abstract), *J. Seismol. Soc. Jpn.,* **31,** 367–382.

Utsu, T., 1999a. Representation and analysis of the earthquake size distribution: a historical review and some new approaches, *Pure appl. Geophys.,* **155,** 509–535.

Utsu, T., 1999b. *Seismicity Studies: A Comprehensive Review (in Japanese),* pp. 876, University of Tokyo Press.

Vere-Jones, D., Robinson, R. & Yang, W., 2001. Remarks on the accelerated moment release model: problems of model formulation, simulation and estimation, *Geophys. J. Int.,* **144,** 517–531.

Wessel, P., Smith, W.H.F., Scharroo, R., Luis, J. & Wobbe, F., 2013. Generic mapping tools: improved version released, *EOS, Trans. Am. Geophys. Un.,* **94,** 409–410.

Wilks, S.S., 1962. *Mathematical Statistics,* pp. 644, John Wiley.