



CENTRE DE RECERCA MATEMÀTICA

Title:	<i>Time window to constrain the corner value of the global seismic-moment distribution</i>
Journal Information:	<i>PLoS ONE,</i>
Author(s):	Corral Á., Serra I..
Volume, pages:	14 1, DOI:[10.1371/journal.pone.0220237]

Time window to constrain the corner value of the global seismic-moment distribution

Álvaro Corral^{1,2,3,4*}, Isabel Serra¹

1 Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, E-08193 Barcelona, Spain

2 Departament de Matemàtiques, Facultat de Ciències, Universitat Autònoma de Barcelona, E-08193 Barcelona, Spain

3 Barcelona Graduate School of Mathematics, Edifici C, Campus Bellaterra, E-08193 Barcelona, Spain

4 Complexity Science Hub Vienna, Josefstädter Straße 39, 1080 Vienna, Austria

* alvaro.corral@uab.es

Abstract

It is well accepted that, at the global scale, the Gutenberg-Richter (GR) law describing the distribution of earthquake magnitude or seismic moment has to be modified at the tail to properly account for the most extreme events. It is debated, though, how much additional time of earthquake recording will be necessary to properly constrain this tail. Using the global CMT catalog, we study how three modifications of the GR law that incorporate a corner-value parameter are compatible with the size of the largest observed earthquake in a given time window. Current data lead to a rather large range of parameter values (e.g., corner magnitude from 8.6 to 10.2 for the so-called tapered GR distribution). Updating this estimation in the future will strongly depend on the maximum magnitude observed, but, under reasonable assumptions, the range will be substantially reduced by the end of this century, contrary to claims in previous literature.

Introduction

Statistics of earthquake occurrence, in particular of the most extreme events, must be a fundamental source to assess seismic hazard [1]. The cornerstone model for describing the earthquake-size distribution is the Gutenberg-Richter (GR) law [2, 3]. The original version of the GR law states that earthquake magnitudes follow an exponential distribution, and since this is a perfectly “well-behaved” distribution, with all statistical moments (such as the mean and the standard deviation) being finite, the problem of earthquake sizes would seem a rather trivial one.

However, a physical interpretation of the meaning of the GR law needs a proper understanding of magnitude. In fact, magnitude presents several difficulties as a measure of earthquake size [4], and a true physical quantity is given instead by seismic moment [5, 6]. Due to the logarithmic dependence of magnitude on seismic moment, the GR law for the latter transforms into a power-law distribution, i.e.,

$$f(x) \propto \frac{1}{x^{1+\beta}}, \text{ for } a \leq x < \infty, \quad (1)$$

where x is seismic moment, $f(x)$ the seismic-moment probability density, a a lower cut-off below which the power law does not hold (presumably because of the incompleteness of the considered catalog for small earthquakes), and $1 + \beta$ the power-law exponent, which takes values close to 1.6 or 1.7 (and with the symbol “ \propto ” representing proportionality). It turns out that the solution to the physical interpretation of the GR law has a price to be paid: the power-law distribution, when $1 + \beta$ is smaller than 2 (which is indeed the case), is not “well behaved”, in the sense that the mean value of the seismic moment becomes infinite.

The reason is that, for power-law distributed seismic moments, events in the tail of the distribution, despite having very small probability, bring an enormous contribution to seismic-moment release [7], and the seismic-moment sample mean does not converge, no matter how large the number of data is, due to the inapplicability of the law of large numbers to power-law distributions [8] such as that in Eq. (1). In consequence, as when extended to the whole range of earthquake sizes the GR law is unphysical, the tail of the distribution of seismic moment must deviate from the GR power-law shape [9].

Due to scarcity of data, the problem has to be approached at a global scale, or at least for a large subset of the global data (for instance, for subduction zones as a whole [9]). This approach has been followed by a number of authors [3, 9–17]. Essentially, a new parameter $M_c > 0$ is introduced, providing a scale for the seismic moment of the largest (“non-GR”) earthquakes, in such a way that for $x \ll M_c$ the GR law can be considered to hold but for $x \gg M_c$ the distribution clearly departs from this law, decaying faster than the GR power law. The values of M_c are more easily read in terms of the corresponding (moment) magnitude m_c [5, 6], through the formula $m_c = 2(\log_{10} M_c - 9.1)/3$, where the seismic moment is measured in N·m. As m_c is sometimes referred to as “corner magnitude”, so M_c would be the “corner seismic moment” [9], independently of the specific probabilistic model (in practice, we will use M_c for formulas and m_c for reporting numeric values, and both will be referred to as “corner parameters” or “corner values”).

In this article we aim to further clarify to what extent the available observations can constrain M_c or m_c , and how many more earthquakes (and then, how many more years of recording) would be likely necessary to yield reasonably precise values of such estimates. Before proposing a rigorous statistical way to tackle these issues, we will need first to assess a previously proposed approach [16].

Probabilistic models

We define the probabilistic models in terms of the cumulative distribution function, $F(x)$, which gives the probability that the random variable (seismic moment) is equal or smaller than a value x . This description is totally equivalent to the one in terms of the probability density, as both functions are related as $f(x) = dF(x)/dx$ and $F(x) = \int_a^x f(x')dx'$ (at some point we will use also the complementary cumulative distribution function, $S(x) = 1 - F(x)$).

The distributions of our interests are:

- (i) the truncated power-law (TPL) distribution [16],

$$F_{tpl}(x) = \left[1 - \left(\frac{a}{M_c} \right)^\beta \right]^{-1} \left[1 - \left(\frac{a}{x} \right)^\beta \right], \text{ for } a \leq x \leq M_c; \quad (2)$$

- (ii) the tapered (Tap) GR law [14, 16, 18], also called Kagan distribution [19],

$$F_{tap}(x) = 1 - \left(\frac{a}{x} \right)^\beta e^{-(x-a)/M_c}, \text{ for } a \leq x < \infty; \quad (3)$$

(iii) the truncated gamma (TrG) distribution [12, 20],

$$F_{trg}(x) = 1 - \frac{\Gamma(-\beta, x/M_c)}{\Gamma(-\beta, a/M_c)}, \text{ for } a \leq x < \infty; \quad (4)$$

with $\Gamma(\gamma, z) = \int_z^\infty x^{\gamma-1} e^{-x} dx$ the upper incomplete gamma function, defined when $\gamma < 0$ only for $z > 0$. All three $F(x)$ are zero for $x < a$ and $F_{tpl}(x) = 1$ for $x \geq M_c$. The parameter β has to be greater than zero, except in the TrG model, where it has no restriction. Of course, $M_c > 0$ and $a > 0$.

The three distributions are graphically depicted in Figs. S1-S3 of the supporting information. Note that for the TPL distribution M_c is a truncation parameter, whereas for the Tap and TrG it is a scale parameter (it sets the scale of $F(x)$ in the x -axis) [12, 20]. Namely, $f_{tpl}(x)$ goes abruptly (discontinuously) to zero at $x = M_c$, whereas for the other two distributions this point sets the scale at which the power law transforms smoothly into an exponential decay. So, the physical meaning of M_c in the TPL is quite different than in the other two models. Note also that the TPL is truncated both from below and from above (but the adjective refers to the truncation from above, $x \leq M_c$), whereas the TrG and Tap are truncated only from below ($x \geq a$). Summarizing, all the considered distributions have two free parameters, β and M_c (or β and m_c), with the value of a fixed by the completeness of the earthquake catalog. In all cases, the limit $M_c \rightarrow \infty$ yields the usual power-law (PL) distribution [20], $F_{pl}(x) = 1 - (a/x)^\beta$ for $x \geq a$, which is equivalent to Eq. (1). Other works have considered different distributions, such as the Gumbel in Ref. [21], for which the power-law limit is not so clear.

State of the art

Several authors have addressed the constraining of the value of M_c and related issues. In particular, Ref. [16] studied the TPL and the Tap distributions (called there GR and MGR, respectively). It was claimed that, for global seismicity with magnitude above 5.75 (i.e., seismic-moment lower cut-off $a = 5.31 \times 10^{17}$ N·m), an enormous amount of data would be necessary in order to obtain reliable estimates of M_c or m_c (200,000 years are mentioned for the Tap distribution with $m_c \simeq 8.5$). Reasonable values proposed previously by other authors (for instance $m_c \simeq 9$ in Ref. [14] for the Tap distribution) were discarded.

The analysis was based on a single statistic: the maximum seismic moment Y of the N earthquakes with magnitude above 5.75 contained in the catalog; that is,

$$Y = \max\{X_1, X_2, \dots, X_N\}.$$

Elementary probability theory allows one obtaining the probability distribution of the maximum Y when the N observations are independent [16, 22] (independence is the maximum-entropy outcome when there is no constrain for the dependence between the observations [23]). Namely, the cumulative distribution function of this maximum is given by

$$F_{max}(y) = \text{Prob}[Y \leq y] = [F(y)]^N, \quad (5)$$

where $F(y)$ can be given by any of the distributions in Eqs. (2)-(4), depending on the underlying statistical model. This approach constitutes an “extreme” limit of the classical block-maxima procedure used in extreme-value theory, considering just one single block [24]. Figure 1 provides an illustration for $F_{max}(y)$; Figs. S4-S6 in the supporting information provide a full picture.

Given a set of N observations with empirical maximum $y_{emp} = \max\{x_1, x_2, \dots, x_N\}$ and a modeling probability distribution $F(x)$, Zöller [16] correctly argued that, if the

data come indeed from $F(x)$, then, $F_{max}(y_{emp}) = \text{Prob}[Y \leq y_{emp}]$ should not be too close to 1. The reason is that proximity to 1 would mean that the empirical value y_{emp} is too large in relation to the values of Y that one can expect from the model distribution $F(x)$ and the number of earthquakes observed. Subsequently, this author introduced an *ad-hoc* distinction between what he called “not well-sampled” distributions, characterized by $F_{max}(y_{emp}) = \text{Prob}[Y \leq y_{emp}]$ large (close to 1) and “well-sampled” distributions, characterized by $F_{max}(y_{emp})$ small. The latter can be equivalently characterized by a large value of the complementary cumulative distribution at y_{emp} , that is, $S_{max}(y_{emp}) = 1 - F_{max}(y_{emp}) = \text{Prob}[Y > y_{emp}]$ large (close to one). In practice [16],

$$S_{max}(y_{emp}) = \text{Prob}[Y > y_{emp}] > 0.99 \quad (6)$$

for “well-sampled” distributions [16]. We will explain below that this criterion cannot be sustained from a statistical point of view, and will introduce instead a robust criterion.

Analyzing global data from the centroid moment tensor (CMT) catalog [25,26], from January 1, 1977 to June 30, 2012 (including shallow, intermediate and deep events, $N = 7,585$ for $x \geq a$), Zöller [16] found that the value of the maximum magnitude corresponds to the 2011 Tohoku earthquake, with magnitude 9.1 (note that the 2004 Sumatra earthquake had a combined multiple-source moment magnitude of 9.3, but only 9.0 with the standard CMT determination [27]). In our work, we will analyze the same dataset, for the sake of comparison. Then, this author [16] evaluated the performance of the TPL and the Tap distributions for different fixed values of the parameter M_c . The considered values correspond to $m_c = 8.5, 9, 9.5, \dots, 12$, in addition to $m_c = 9.2$. In contrast, it was stated that β was estimated by maximum likelihood for fixed M_c .

For the TPL model, a value of $m_c = 9.2$ resulted in $\text{Prob}[Y > y_{emp}] = 0.55$ [16], whereas $m_c = 9.5$ and $m_c = 10$ led to $\text{Prob}[Y > y_{emp}]$ very close to one, and even closer-to-one values were obtained for $m_c \geq 10.5$. Following the “well-sampledness” criterion, the value $m_c = 9.2$ was discarded for the TPL model, despite of having the maximum likelihood among all the values of the parameters considered, and values with $m_c \geq 10.5$, with much smaller likelihood, were preferred. However, no preference was shown between $m_c = 10.5$ and any other higher value (for instance $m_c = 12$) and all the models were considered equally likely. For the Tap model, the previous results and the conclusions [16] were similar to those for the TPL model, and in this way the value $m_c = 9$ (proposed in Ref. [14]) was rejected despite of yielding maximum likelihood.

The calculation of the required number of data to perform a reliable estimation of parameter M_c (or m_c) was obtained by imposing a minimum number of events N_m such that the distribution becomes “well-sampled” [16], in the sense of Eq. (6). So, introducing Eq. (5) into Eq. (6),

$$S_{max}(y_{emp}) = \text{Prob}[Y > y_{emp}] = 1 - [F(y_{emp})]^{N_m} > 0.99. \quad (7)$$

Note that, no matter the value of $F(y_{emp})$, if this is strictly smaller than 1, for sufficiently large N_m we will have $[F(y_{emp})]^{N_m} < 0.01$ and the condition will be fulfilled by any model, with any parameter value, if enough data are gathered (except truncated models with $F(y_{emp}) = 1$). Imposing that the previous condition becomes an equality one gets

$$N_m = \frac{|\ln 0.01|}{|\ln F(y_{emp})|} = \frac{7,585 |\ln 0.01|}{|\ln (1 - \text{Prob}[Y > y_{emp}])|}. \quad (8)$$

We will argue below that this equation (8), used (but not made explicit) in previous research [16], does not hold for the problem under consideration.

In this way, for the TPL model with $m_c = 9.2$, accepting the value $\text{Prob}[Y > y_{emp}] = 0.55$, the approach just outlined, Ref. [16] [Eq. (8) here], yields that

N_m has to be higher than 45,000 (corresponding to 212 years of earthquake recording, with about 214 earthquakes with $x \geq a$ per year). For the Tap model with $m_c = 8.5$, for which $\text{Prob}[Y > y_{emp}] = 0.0007$, one obtains that more than 200,000 years would be needed (from $N_m = 50 \times 10^6$, roughly). Note the counterintuitive results that this approach leads to: the larger the corner seismic moment M_c , the less data are required for its estimation, as contained in Eq. (8) (due to the decrease of $F(y_{emp})$ with m_c) and illustrated for the TPL model in Fig. 2.

Proper testing using the maximum seismic moment

First, let us show that the previously used “well-sampledness” criterion [16], reproduced here in Eq. (6), is not appropriate. If the distribution $F(x)$ is a good model for the empirical data, what one expects is that both $\text{Prob}[Y \leq y_{emp}]$ and $\text{Prob}[Y > y_{emp}]$ are not too close to 1, let us say, below $1 - (1 - r)\alpha$ and $1 - r\alpha$, respectively, at significance level α (with $r = 1/2$ in the usual symmetric case and $\alpha = 0.05$ or 0.01). As both probabilities add to one, the conditions can be written as

$$r\alpha < \text{Prob}[Y \leq y_{emp}] < 1 - (1 - r)\alpha. \quad (9)$$

or, equivalently, as

$$(1 - r)\alpha < \text{Prob}[Y > y_{emp}] < 1 - r\alpha,$$

i.e., the random variable Y takes not too extreme values with probability $1 - \alpha$ (e.g. 0.95 or 0.99). Note the profound difference between these conditions and the “well-sampledness” criterion [16], Eq. (6) here.

Note that, following this “new” criterion, previous numerical results for the truncated power-law distribution [16] seem to indicate (in contrast to the conclusions there) that all tested values of m_c should be rejected at the 0.05 significance level (as Ref. [16] reports $\text{Prob}[Y > y_{emp}] > 0.975$), except $m_c = 9.2$ (the value of $\text{Prob}[Y > y_{emp}]$ for $m_c = 9.5$ displayed in Fig. 3 of Ref. [16] seems to be slightly above 0.975 and should be rejected as well, at least in the symmetric case $r = 1/2$). For the Tap distribution, the only values of m_c that should not be clearly rejected from the numerical results of Ref. [16] (again in contrast with the conclusions of that reference) are $m_c = 9$ and $m_c = 9.2$ (for the rest of m_c values Ref. [16] reports $\text{Prob}[Y > y_{emp}]$ above 0.975 or below 0.025). But the numerical results of Ref. [16] are not in correspondence with ours; our maximum-likelihood estimations for β do not lead to $\text{Prob}[Y > y_{emp}] \simeq 1$ when m_c is large ($m_c \geq 10$). What we find for those values is $\text{Prob}[Y > y_{emp}] < 0.975$, see Figs. 1 and S5 (and Fig. S6 for the TrG), so all large values of m_c are allowed, in principle.

Regarding the number of earthquakes required to constrain the corner parameters (M_c or m_c), what is implicit behind Eq. (8) is that a “not well-sampled” distribution (with $\text{Prob}[Y > y_{emp}]$ close to zero) is “not well-sampled” just because of “bad luck”, that is, the largest earthquake had y_{emp} much larger than expected from both the model $F(x)$ and the actual value of N . This bad luck is what leads to the rejection of the null hypothesis in usual statistical testing (and corresponds to the significance level, see Eq. (9)). But, in Ref. [16]’s argument, gathering more data would eventually lead to the accommodation of the theoretical distribution of the maximum to the empirical value y_{emp} , regardless of the model. Thus, in that assumption y_{emp} is considered quenched, i.e., it does not grow despite the fact that the number of data increases. This is hard to justify.

Proper constraining of the corner seismic-moment: TPL case

In this section we derive a proper statistical way to evaluate the number N of earthquakes necessary to constrain the estimated value of M_c or m_c for the TPL distribution. In this case, our approach uses the distribution of the estimator of these quantities (M_c and m_c) to calculate their statistical uncertainty as a function of N , and looks for the value of N that reduces the uncertainty down to a desired range. This will necessarily depend on the true values of the parameters, which are unknown, and is also based on the assumption that the sample is representative of the whole population (otherwise, no inference is possible).

For this purpose, let us focus in the truncated-power-law model, which has the peculiar property that the random variable Y (the maximum seismic moment of the N earthquakes) constitutes the maximum-likelihood estimator, \hat{M}_c , of the truncation parameter M_c , that is $Y = \hat{M}_c$ for the TPL (or, equivalently, \hat{m}_c for the magnitude). Then, inverting $F_{max}(y_p) = p$, with y_p defining the $100p$ -th percentile of the distribution of the maximum seismic moment (i.e., the distribution of \hat{M}_c), one can get the probability of any interval for \hat{M}_c . The limiting points for these intervals are, from Eqs. (5) and (2),

$$y_{p,tpl} = \frac{a}{\sqrt[\beta]{1 - p^{1/N}[1 - (a/M_c)^\beta]}},$$

and in terms of the magnitude,

$$m_{p,tpl} = \frac{2}{3} \left[\log_{10} \left(\frac{a}{\sqrt[\beta]{1 - p^{1/N}[1 - (a/M_c)^\beta]}} \right) - 9.1 \right], \quad (10)$$

using the relation between magnitude and seismic moment, with m_p the $100p$ -th percentile of the distribution of the maximum magnitude. For the true distribution, the resulting 95%-probability intervals, $(m_{p,tpl}, m_{p+0.95,tpl})$, should contain the empirical value of the maximum with a 0.95 probability. These intervals are shown in Fig. 3, using the empirical value of N in the global CMT catalog and different values of M_c , with β fixed to 0.67, and $p = 0.025$ for symmetric intervals (we have checked that the final results do not depend too much on this choice).

Figure 3 shows that the ideal situation happens when the distribution of the maximum-likelihood estimator is very narrow, and then $\hat{m}_c \simeq m_c$, leading to the automatic recovering of the true value (a value very close to it, but below, in fact). When N is equal to the empirical value (considering the case previously studied in the literature [16], up to mid 2012) this happens for $m_c < 8.5$. One could refer to this case as “sampled enough” (in sharp contrast with previous terminology [16]). On the contrary, when the upper limit of the interval, $m_{p+0.95}$, departs clearly from the true value of m_c , we may talk of undersampling (there is no hint of the real maximum m_c after the N observations, again in contrast with previous research [16]). This is the case for $m_c > 10.5$ (for $N = 7,585$), for which the intervals do not include the true value of m_c (for instance, for $m_c = 12$ the interval of the maximum goes from 9 to 11, roughly, see Fig. 3). But note this kind of undersampling still would allow ruling out the values of the parameters of the undersampled distributions, if the empirical value of the maximum were outside the resulting interval (nevertheless, this is not the case for the actual value, see below). In the intermediate case ($8.5 < m_c < 10.5$ for the period under consideration), the intervals are wide but they reach the true value.

We can use the previous argument to find the value of N that leads to narrow 95%-probability intervals for the estimation of M_c or m_c in the TPL model. Using Eq.

(10), the width of the magnitude intervals, $\Delta = m_{p+0.95} - m_p$, is obtained as

$$\Delta_{tpl} = \frac{2}{3\beta} \log_{10} \frac{1 - [1 - (a/M_c)^\beta] p^{1/N}}{1 - [1 - (a/M_c)^\beta] (p + 0.95)^{1/N}}. \quad (11)$$

Isolating N as a function of Δ_{tpl} for given values of M_c and β yields the desired result. Notice that, in contrast to Ref. [16] [Eq. (6)], our approach does not need any empirical information (except the value of β). Going back to Fig. 2, this includes the number of events necessary to obtain intervals of a fixed width after numerical inversion of Eq. (11), as a function of M_c . The results are clearly different to the previous ones [16], as shown in the figure.

Figure 2 is particularly useful for testing a specific value of m_c . If the real value of m_c were 9.5 (the largest earthquake in the historical record [28], but not contained in the CMT catalog) a 95%-probability interval with width $\Delta = 0.4$ (from 9.1 to 9.5, roughly) would be obtained after about $N = 14,000$ events (corresponding to 65 years, reached in 2042). If one wants instead a width of $\Delta = 0.2$ (yielding an interval from 9.3 to 9.5) the necessary N is 36,400, to be reached around the year 2147 (assuming that the TPL were the right model, that there is no dependence between the magnitudes, and that the long-term global earthquake rate and β were constant).

It is important to realize that, in all the cases shown in the figure, the top value of the interval coincides with the real value. Although the probability that the estimated value is between $m_{p+0.95}$ and m_c is $0.05 - p$, the two values are very close, i.e., $m_{p+0.95} \simeq m_c$; this is due to the extreme sharpness of the density of the observed maximum close to m_c (for instance, as in Fig. 1, where the vertical axis is logarithmic). So, the value of N provided in the figure guarantees no undersampling. Note also that a 95%-probability interval is a much more strict requirement than an interval corresponding to one standard deviation.

We have just calculated the number of earthquakes required to estimate M_c with a given uncertainty, for different hypothetical values of the true M_c . This does not make use the empirical value y_{emp} obtained in 35.5 years. A different issue then is how y_{emp} discards or not the possible values of M_c . Figure 3 shows (in addition to the intervals of the maximum magnitude obtained from Eq. (10)) the empirical value obtained for the period 1977-2012.5. If the observed maximum magnitude (9.1 in the global CMT catalog) is inside the interval, there is no reason to reject the parameters of the model (with a 95% confidence); on the contrary, if the empirical value is outside, we should reject the parameters.

The figure shows how, for the TPL model, no value of $m_c \geq 9.1$ can be rejected, i.e., any value of m_c between 9.1 and ∞ is compatible with the empirical result, and therefore the data do not allow to determine an upper bound for m_c , although values of m_c above 10 are close to rejection (with a 95% confidence; if we decreased the confidence or increased the number of data an upper bound would appear). Indeed, considering the most recent data at the time of writing, up to the end of 2017 (where no other earthquake of magnitude larger than 9.1 has taken place) the range of compatible values of m_c turns out to be 9.1–10.8, as reported in Table 1.

As an illustration, we also analyze what an hypothetical y_{emp} corresponding to a 9.1 magnitude in a 71-year period (from 1977 to 2047, let us say) would imply. Table 1 shows that that would constrain m_c to be between 9.1 and 9.5, for 95%-probability intervals, but if the maximum in the same period were 9.3, the allowed range would be between 9.3 and 10.3. In contrast, a maximum empirical value of 9.5 (or higher) in that period would yield m_c unbounded from above again. Needless to say, we need to wait about 30 years to chose between these three answers.

Proper constraining of the corner seismic-moment: Tap and TrG cases

Note that, although the maximum empirical value of the seismic moment is the maximum-likelihood estimator of M_c only for the TPL distribution (out of the three considered models), we can still use the previous procedure to constrain the value of M_c for any distribution, but with the resulting values of M_c not related to maximum likelihood estimation, in general. Thus, for the Tap distribution, the percentiles of the maximum seismic moment turn out to be, using Eq. (3) and (5),

$$y_{p,tap} = \beta M_c W \left(\frac{a e^{a/(\beta M_c)}}{\beta M_c (1 - p^{1/N})^{1/\beta}} \right),$$

with W the Lambert W function [29], fulfilling $z = W(ze^z)$. And for the truncated gamma we get, using Eq. (4),

$$y_{p,trg} = M_c \Gamma_2^{-1}(-\beta, (1 - p^{1/N})\Gamma(-\beta, a/M_c)),$$

with Γ_2^{-1} the inverse, respect to its second argument, of the incomplete gamma function. In the same way as for the TPL, the empirical value y_{emp} leads to an unbounded range of the values of m_c compatible with y_{emp} for the original value of N (7,585). These ranges go from 8.65 to ∞ for the Tap distribution and from 8.8 to ∞ for the TrG, with $\beta = 0.67$. However, when one extends the analysis up to 2017 the ranges become bounded, although large, see Table 1.

This table also explores the values of these ranges in the future, depending on the hypothetical value of the maximum magnitude observed. We see that, in general, the ranges provided by the Tap distribution are somewhat wider than those provided by the TPL, whereas the TrG yields rather larger ranges. This means that the number of data necessary to constrain the value of m_c is larger in the TrG than in the other two distributions. The table also allows us to rule out the scenario that there will be no earthquakes larger than magnitude 9.1 before 2097 for a TPL distribution, as this scenario leads to the implausibility of having events larger than 9.3, contrary to what was observed in the 9.5 1960 event in Chile (although the CMT catalog would probably underestimate the seismic moment of such an event [27]).

Discussion

Before concluding, we briefly explore the implications of our results for the assesment of seismic hazard. Considering as an illustration the case of the tapered model, we have seen (Table 1) how the CMT data, up to 2017, is compatible with a range of values of the corner magnitude, from $m_{cmin} = 8.6$ to $m_{cmax} = 10.2$ (with a 95 % confidence). Therefore, the resulting seismic-moment distribution (or, in the same way, the resulting magnitude distribution) will be a mixture (or combination) of the different $S_{tap}(x|M_c)$ (now we use the complementary cumulative distribution function and make explicit in the notation the dependence on the corner seismic moment M_c), with M_c ranging from M_{cmin} to M_{cmax} . Thus,

$$S_{mix}(x) = \int_{M_{cmin}}^{M_{cmax}} S_{tpl}(x|M_c) \rho(M_c) dM_c, \quad (12)$$

where the resulting distribution $S_{mix}(x)$ is no longer a Tap distribution but a mixture of Tap's with different M_c . The term $\rho(M_c)$ gives weight to the different values of M_c . The same equation holds for any other probabilistic model (such as TPL and TrG).

One could assume a uniform distribution of corner magnitudes (all its values would be equally likely from $m_{c\min}$ to $m_{c\max}$). Interestingly, for the corner seismic-moment distribution, this leads to the Jeffreys prior of a scale parameter, $\rho(M_c) \propto 1/M_c$. Under this choice, the integral in Eq. (12) can be easily evaluated by the Monte-Carlo method. For the Tap model, the probability of an earthquake of magnitude 9.1 or larger (among all earthquakes with magnitude larger than 5.75) turns out to be $S_{mix}(x) = 2.6 \times 10^{-4}$, corresponding to about 1 in 20 years. In comparison with the CMT catalog itself (1 of such events in 35.5 years) this probability seems somewhat large. Even higher values of the magnitude or other models (TPL or TrG) also seem to lead to an overestimation of these probabilities. Naturally, this is the core problem in the statistics of extreme events, one has very few extreme events to contrast estimations. As the result is highly sensitive to the choice of the distribution $\rho(M_c)$, this is a topic that deserves further study.

Our results can also have applications for time-dependent hazard [18]. If we know when the last earthquake of a given seismic moment x or higher happened (a time t ago), we can obtain the probability of recurrence in a given time period Δ from the present as

$$\mathcal{P}_{x,t,\Delta} = \text{Prob}[t < \text{waiting time} \leq t + \Delta \mid \text{waiting time} > t] = 1 - \frac{S_w(t + \Delta \mid x)}{S_w(t \mid x)},$$

where the subindex w denotes that the distribution refers to the waiting time (not to the seismic moment). For a Poisson process S_w is exponential with rate λ_x and then we recover

$$\mathcal{P}_{x,t,\Delta} = 1 - e^{-\lambda_x \Delta} \simeq \lambda_x \Delta = R_a S(x) \Delta,$$

which turns out to be independent on t and becomes essentially the same formula used above for time independent hazard, with $R_a = 213.7 \text{ year}^{-1}$ (we have assumed $\Delta \ll \lambda_x^{-1}$).

In order to obtain time-dependent hazard one needs to go beyond Poisson occurrence. At a global scale it has been pointed out that the gamma distribution can describe well earthquake waiting times [30, 31]; nevertheless, for the sake of simplicity, we are going to illustrate the calculation with the Weibull distribution, which can give similar fits [32]. In this way, from the equation ago we can write

$$\mathcal{P}_{x,t,\Delta} = 1 - \exp \left[\left(\frac{t}{c_x} \right)^\gamma - \left(\frac{t + \Delta}{c_x} \right)^\gamma \right], \quad (13)$$

with γ and c_x the shape and scale parameters of the Weibull distribution, respectively (the latter depending on x). The Poisson case is included in the particular limit $\gamma = 1$.

The scale parameter of the waiting-time distribution can be directly related to the seismic-moment distribution: On the one hand, the number of events per unit time (with seismic moment above x) is $R_a S(x)$. On the other hand, this number is also given by $1/\langle t(x) \rangle$, where $\langle t(x) \rangle$ is the mean waiting time for events above x . In the particular case of the Weibull distribution, this is given by $\langle t(x) \rangle = c_x g(\gamma)$ with $g(\gamma) = \Gamma(1 + \gamma^{-1})$. Thus,

$$c_x = \frac{1}{g(\gamma) R_a S(x)},$$

which substituting into Eq. (13) allows the calculation of the probability $\mathcal{P}_{x,t,\Delta}$. In the case $\Delta \ll t$ this can be simplified to

$$\mathcal{P}_{x,t,\Delta} \simeq 1 - \exp \left[-\gamma \Delta (g(\gamma) R_a S(x))^\gamma t^{\gamma-1} \right].$$

In the context of this article, the seismic-moment distribution $S(x)$ could be substituted by the mixture for different values of M_c given by Eq. (12). Nevertheless, the calculation of these probabilities needs the accurate fitting of the waiting time distributions $S_w(t \mid x)$ (i.e., the fitting of γ and c_x in the case of the Weibull distribution). This is left to future works.

Conclusions

Summarizing the main results of the article, we have reconsidered to what extent the available earthquake record can constrain the parameter that characterizes the tail of the global seismic-moment distribution: a corner seismic moment (M_c , or its corresponding moment magnitude m_c), for three different distributions (truncated power law, tapered GR, and truncated gamma). We have corrected some of the drawbacks of previous literature, regarding the number of events necessary for such a purpose.

The key point in our approach is to obtain the percentiles of the distribution of the maximum seismic moment of N earthquakes, and to derive from there probability intervals that can be compared with the maximum seismic moment observed, y_{emp} . If y_{emp} is inside the interval there is no reason to reject the considered value of the corner parameter. Although currently (up to the end of 2017), the range of values of m_c is rather wide, in 80 years from now these ranges are expected to decrease substantially, but depending crucially on the maximum value to be observed. For instance, if this were 9.3, the tapered model would lead to $m_c \simeq 9.1 \pm 0.3$ (95 % confidence), and the truncated gamma model to 9.35 ± 0.45 (see Table 1 for more hypothetical examples). From here we conclude that the much larger periods of time estimated earlier are not justified. In addition, for the same reasons elaborated in this article, the standard errors of corner parameters that we [20] calculated previously for almost 37 years of shallow global seismicity using asymptotic likelihood theory do not provide a convenient description of the range of uncertainty in those parameters.

Acknowledgements

We appreciate the critical reading of Álvaro González, as well as suggestions from the reviewers. The data used in this research can be downloaded from <http://www.globalcmt.org> [25].

References

1. Mulargia F, Stark PB, Geller RJ. Why is Probabilistic Seismic Hazard Analysis (PSHA) still used? *Phys Earth Planet Int.* 2017;264:63–75.
2. Utsu T. Representation and analysis of earthquake size distribution: a historical review and some new approaches. *Pure Appl Geophys.* 1999;155:509–535.
3. Kagan YY. *Earthquakes: Models, Statistics, Testable Forecasts.* Wiley; 2014.
4. Ben-Zion Y. Collective behavior of earthquakes and faults: continuum-discrete transitions, progressive evolutionary changes, and different dynamic regimes. *Rev Geophys.* 2008;46:RG4006.
5. Kanamori H. The energy release in great earthquakes. *J Geophys Res.* 1977;82(20):2981–2987.
6. Kanamori H, Brodsky EE. The physics of earthquakes. *Rep Prog Phys.* 2004;67:1429–1496.
7. Corral A, Font-Clos F. Criticality and self-organization in branching processes: application to natural hazards. In: Aschwanden M, editor. *Self-Organized Criticality Systems.* Open Academic Press, Berlin; 2013. p. 183–228.

8. Corral A. Scaling in the Timing of Extreme Events. *Chaos Solit Fract.* 2015;74:99–112.
9. Kagan YY. Seismic moment distribution revisited: I. Statistical results. *Geophys J Int.* 2002;148:520–541.
10. Kagan YY. Universality of the Seismic Moment-frequency Relation. *Pure Appl Geophys.* 1999;155:537–573.
11. Godano C, Pingue F. Is the seismic moment-frequency relation universal? *Geophys J Int.* 2000;142:193–198.
12. Main IG, Li L, McCloskey J, Naylor M. Effect of the Sumatran mega-earthquake on the global magnitude cut-off and event rate. *Nature Geosci.* 2008;1:142.
13. Kagan YY. Earthquake size distribution: Power-law with exponent $\beta \equiv 1/2$? *Tectonophys.* 2010;490:103–114.
14. Bell AF, Naylor M, Main IG. Convergence of the frequency-size distribution of global earthquakes. *Geophys Res Lett.* 2013;40:2585–2589.
15. Deluca A, Corral A. Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions. *Acta Geophys.* 2013;61:1351–1394.
16. Zöller G. Convergence of the frequency-magnitude distribution of global earthquakes: Maybe in 200 years. *Geophys Res Lett.* 2013;40:3873–3877.
17. Geist EL, Parsons T. Undersampling power-law size distributions: effect on the assessment of extreme natural hazards. *Nat Hazards.* 2014;72:565–595.
18. Mulargia F, Geller RJ, Earthquake Science and Seismic Risk Reduction. Kluwer, Dordrecht; 2003.
19. Vere-Jones D, Robinson R, Yang W. Remarks on the accelerated moment release model: problems of model formulation, simulation and estimation. *Geophys J Int.* 2001;144(3):517–531.
20. Serra I, Corral A. Deviation from power law of the global seismic moment distribution. *Sci Rep.* 2017;7:40045.
21. Lomnitz-Adler J, Lomnitz C. A modified form of the Gutenberg-Richter magnitude-frequency relation. *Bull Seismol Soc Am.* 1979; 69(4) 1209–1214.
22. Ross SM. *A First Course in Probability.* 8th ed. Prentice Hall, Englewood Cliffs; 2010.
23. Broderick T, Dudík M, Tkacik G, Schapireb RE, Bialek W. Faster solutions of the inverse pairwise Ising problem. *arXiv.* 2007;0712.2437.
24. Coles S. *An Introduction to Statistical Modeling of Extreme Values.* Springer, London; 2001.
25. Ekström G, Nettles M, Dziewoński AM. The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes. *Phys Earth Planet Int.* 2012;200-201:1–9.
26. Dziewonski AM, Chou TA, Woodhouse JH. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *J Geophys Res.* 1981;86:2825–2852.

27. Tsai VC, Nettles M, Ekström G, Dziewonski AM. Multiple CMT source analysis of the 2004 Sumatra earthquake. *Geophys Res Lett.* 2005;32(17):L17304.
28. Satake K, Atwater BF. Long-Term Perspectives on Giant Earthquakes and Tsunamis at Subduction Zones. *Ann Rev Earth Planet Sci.* 2007;35(1):349–374.
29. Corless RM, Gonnet GH, Hare DEG, Jeffrey DJ, Knuth DE. On the Lambert W function. *Adv Comp Math.* 1996;5(1):329–359.
30. Corral A, Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Phys Rev Lett.* 2004; 92:108501.
31. Corral A. Statistical Features of Earthquake Temporal Occurrence. In: Bhattacharyya P, Chakrabarti BK, editors. *Modelling Critical and Catastrophic Phenomena in Geoscience. Lecture Notes in Physics, 705.* Springer, Berlin; 2007. p 191–221.
32. Moriña D, Serra I, Puig P, Corral A. Probability estimation of a Carrington-like geomagnetic storm. *Sci Rep.* 2019; 9:2393.

1 Supporting information

S1 Fig. ccdf $S(x)$ and pdf $f(x)$ of TPL distribution with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend.

S2 Fig. ccdf $S(x)$ and pdf $f(x)$ of Tap distribution with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend.

S3 Fig. ccdf $S(x)$ and pdf $f(x)$ of TrG distribution with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend.

S4 Fig. ccdf $S_{max}(y)$ and pdf $f_{max}(y)$ of the maximum of 7,585 TPL observations with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend. Critical values at the 95% confidence level are shown as horizontal lines. Empirical value of maximum seismic moment observed is shown as a vertical line. Note that this is exactly Fig. 1 of the main text, repeated here for completeness.

S5 Fig. ccdf $S_{max}(y)$ and pdf $f_{max}(y)$ of the maximum of 7,585 Tap observations with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend. Critical values at the 95% confidence level are shown as horizontal lines. Empirical value of maximum seismic moment observed is shown as a vertical line.

S6 Fig. ccdf $S_{max}(y)$ and pdf $f_{max}(y)$ of the maximum of 7,585 TrG observations with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend. Critical values at the 95% confidence level are shown as horizontal lines. Empirical value of maximum seismic moment observed is shown as a vertical line.

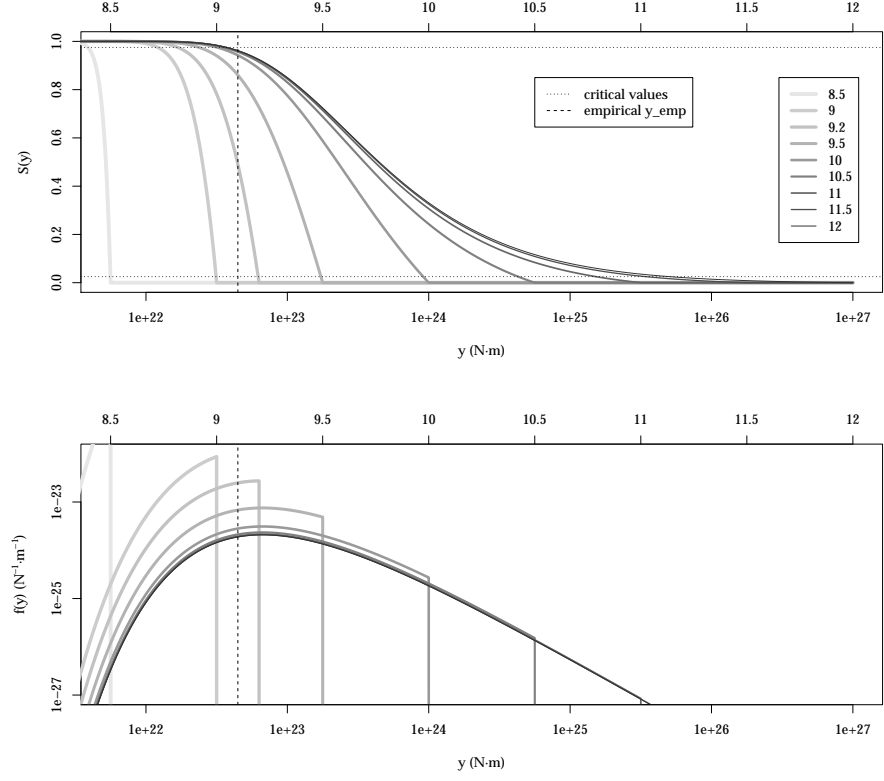


Fig 1. Probability distributions for the maximum of $N = 7,585$ values of seismic moment (as in the global CMT catalog considered), assuming that these are independent and distributed according to truncated power laws with lower cut-off $a = 5.31 \times 10^{17} \text{ N}\cdot\text{m}$ and diverse values of m_c ranging from 8.5 to 12. The value of the exponent is fixed to $1 + \beta = 1.67$, very close to the maximum-likelihood solution. The largest empirical value in the catalog, y_{emp} , is shown as a vertical line. (a) Complementary cumulative distributions $S_{max}(y)$ and critical values 0.025 and 0.975 (horizontal lines). Note that $0.025 < S_{max}(y_{emp}) < 0.975$ at least for $m_c \geq 9.2$, in contrast with the results of Ref. [16], so these values of m_c cannot be ruled out. (b) The corresponding probability densities $f_{max}(y)$.

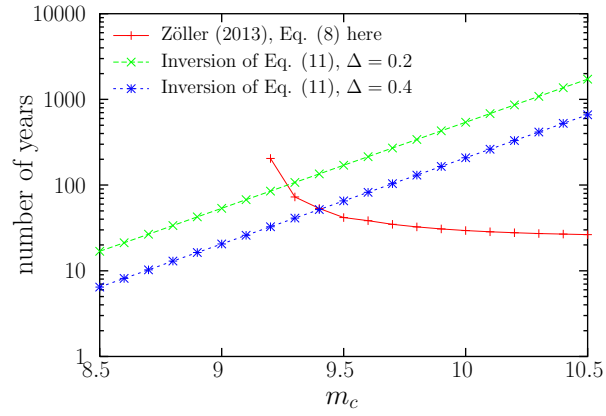


Fig 2. Number of years necessary to obtain a reliable estimation of the truncation parameter M_c for the TPL model with $\beta = 0.67$ as a function of the hypothetical true value of M_c (represented by m_c), according to Ref. [16] (decreasing curve) and according to our results [inverting Eq. (11), increasing curves], assuming an average rate of 213.7 events per year. In the latter case we impose that 95%-probability intervals have magnitude width $\Delta = 0.2$ and 0.4 . The resulting values of N guarantee no undersampling (i.e., $m_{p+0.95} \simeq m_c$, not shown). Note the totally different outcomes of the two approaches.

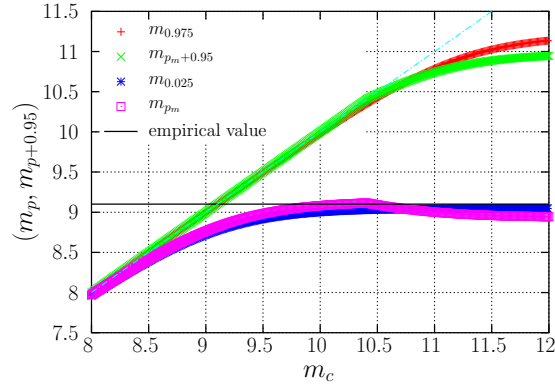


Fig 3. 95%-probability intervals, represented by the starting and ending points $(m_p, m_{p+0.95})$ for the truncation parameter M_c of a TPL distribution with $N = 7,585$ earthquakes (in terms of the corresponding truncation magnitude m_c), as a function of the hypothetical true values of m_c . The value of the exponent is $1 + \beta = 1.67$. Two kinds of intervals are shown: symmetric ($r = 1/2$ in Eq. (9)) and of minimum width (the r that gives minimum width is selected), labeled with m_{p_m} . The empirical value of the maximum observed magnitude in the global CMT catalog for the 7,585 considered earthquakes is shown as a horizontal line. When the line is outside the interval, the parameter value m_c should be rejected.

Table 1. Values of the corner parameter m_c compatible (for 95%-probability intervals) with a maximum observed magnitude $m(y_{emp})$ in a time period starting in 1977 and ending in the indicated final year, for the truncated power law (TPL), tapered (Tap) and truncated gamma (TrG) distributions. The values of $m(y_{emp})$ marked with an asterisk (*) indicate hypothetical values (the rest corresponds to the real observed value, 9.1). The value of β is 0.67. The final year is estimated assuming a global rate of 213.7 earthquakes with moment magnitude ≥ 5.75 per year.

final year	$m(y_{emp})$	m_c TPL	m_c Tap	m_c TrG
2012.5	9.1	9.1– ∞	8.6– ∞	8.8– ∞
2017	9.1	9.1–10.8	8.6–10.2	8.8–11.2
2047	9.1*	9.1– 9.5	8.6– 9.3	8.7– 9.7
2047	9.3*	9.3–10.3	8.8– 9.95	9.0–10.6
2047	9.5*	9.5– ∞	9.1– ∞	9.2– ∞
2097	9.1*	9.1– 9.3	8.6– 9.1	8.7– 9.4
2097	9.3*	9.3– 9.6	8.8– 9.4	8.9– 9.8
2097	9.5*	9.5–10.3	9.0–10.0	9.2–10.6

**Supporting Information for “Time window to
constrain the corner value of the global
seismic-moment distribution”**
Contents of this file

1. Text S1
2. Figures S1 to S6

Introduction. We provide graphical description of the three models for seismic-moment distributions in the paper as well as for the maximum of 7,585 of such events. We provide also the mathematical expressions for the probability density functions (pdf) of the three considered models for seismic moment, x . Figures *S1*, *S2*, and *S3* display the probability density functions and complementary cumulative distribution functions (ccdf) for each model when lower cut-off and shape parameter are fixed, and parameter M_c is varying (as indicated in the legend). Figures *S4*, *S5*, and *S6* correspond to the pdf and ccdf of the maximum of 7,585 observations for each of the previous models.

Text S1.

The mathematical expressions of the probability density functions of the three models follow:

Truncated power-law (TPL) distribution,

$$f_{tpl}(x) = \left[\frac{1}{1 - (a/M_c)^\beta} \right] \frac{\beta}{a} \left(\frac{a}{x} \right)^{1+\beta}, \quad (14)$$

for $a \leq x \leq M_c$, and zero otherwise, with $\beta > 0$.

Tapered (Tap) GR distribution, also called Kagan distribution,

$$f_{tap}(x) = \left[\frac{\beta}{a} \left(\frac{a}{x} \right)^{1+\beta} + \frac{1}{M_c} \left(\frac{a}{x} \right)^\beta \right] e^{-(x-a)/M_c}, \quad (15)$$

for $a \leq x < \infty$, and zero otherwise, with $\beta > 0$.

Truncated gamma (TrG) distribution,

$$f_{trg}(x) = \frac{1}{M_c \Gamma(-\beta, a/M_c)} \left(\frac{M_c}{x} \right)^{1+\beta} e^{-x/M_c}, \quad (16)$$

for $a \leq x < \infty$, and zero otherwise, with $-\infty < \beta < \infty$ and $\Gamma(\gamma, z) = \int_z^\infty x^{\gamma-1} e^{-x} dx$ being the upper incomplete gamma function, defined when $\gamma < 0$ only for $z > 0$.

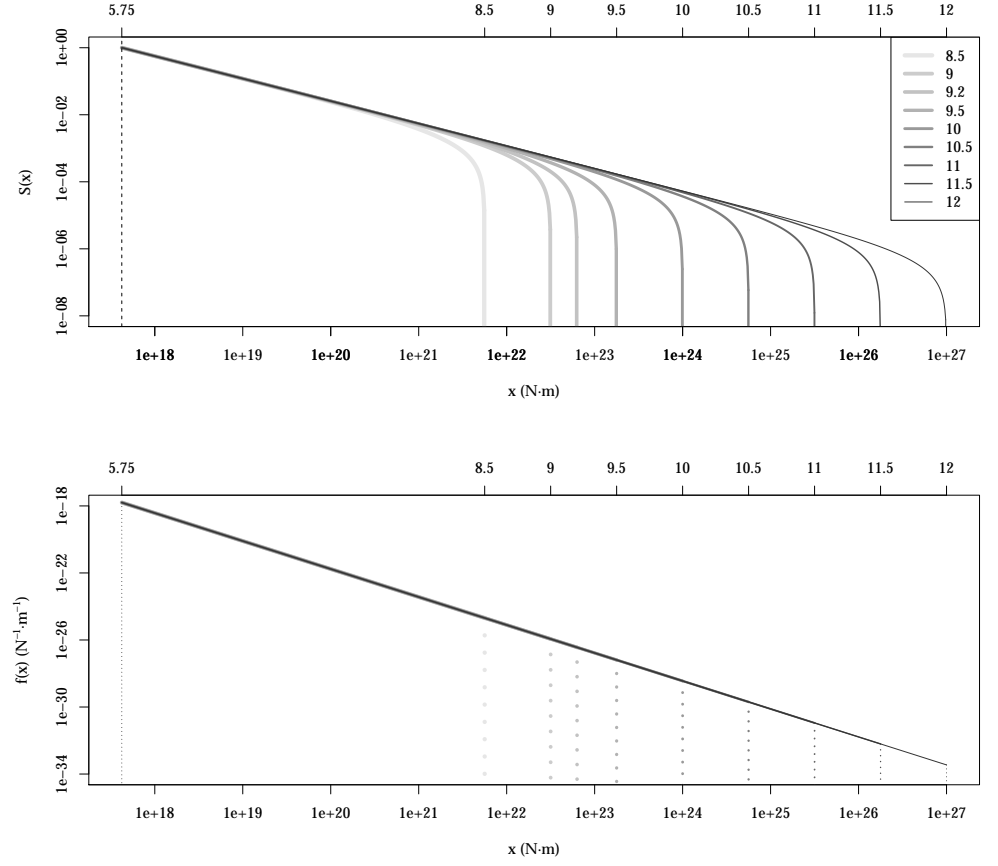


Fig 4. (Fig. S1) cdf $S(x)$ and pdf $f(x)$ of TPL distribution with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend.

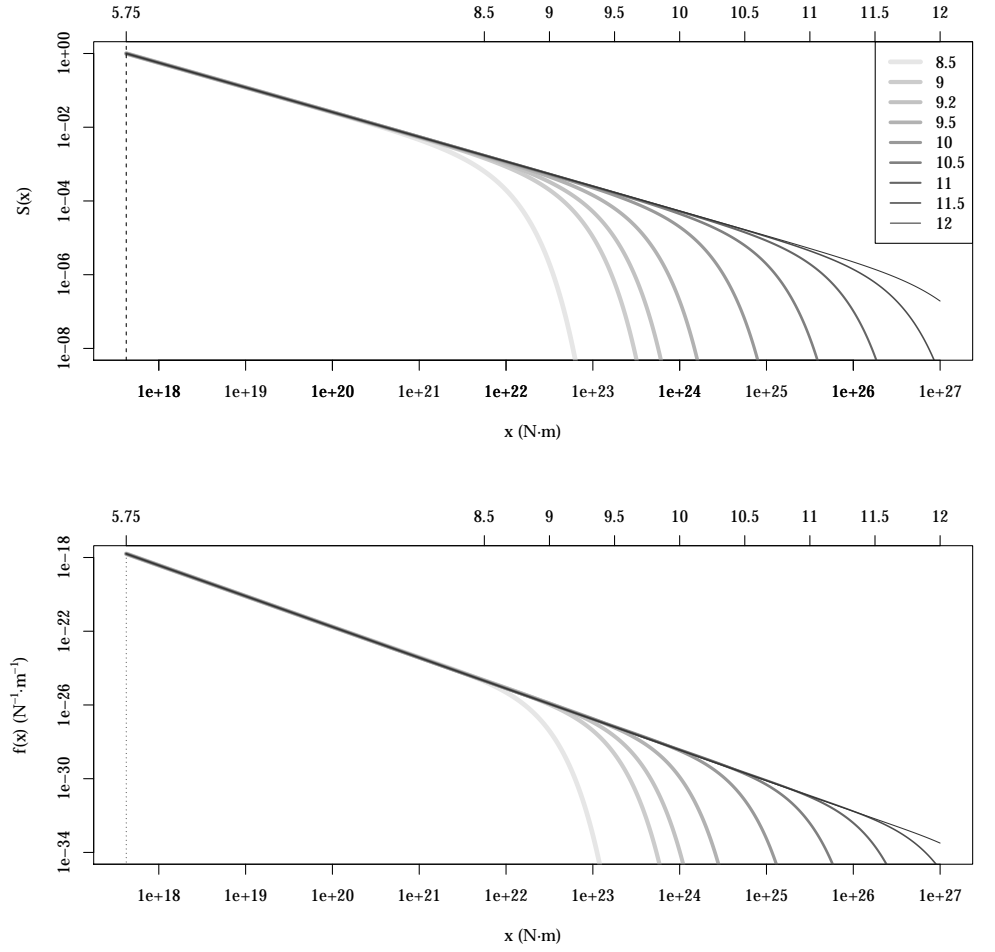


Fig 5. (Fig. S2) ccdf $S(x)$ and pdf $f(x)$ of Tap distribution with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend.

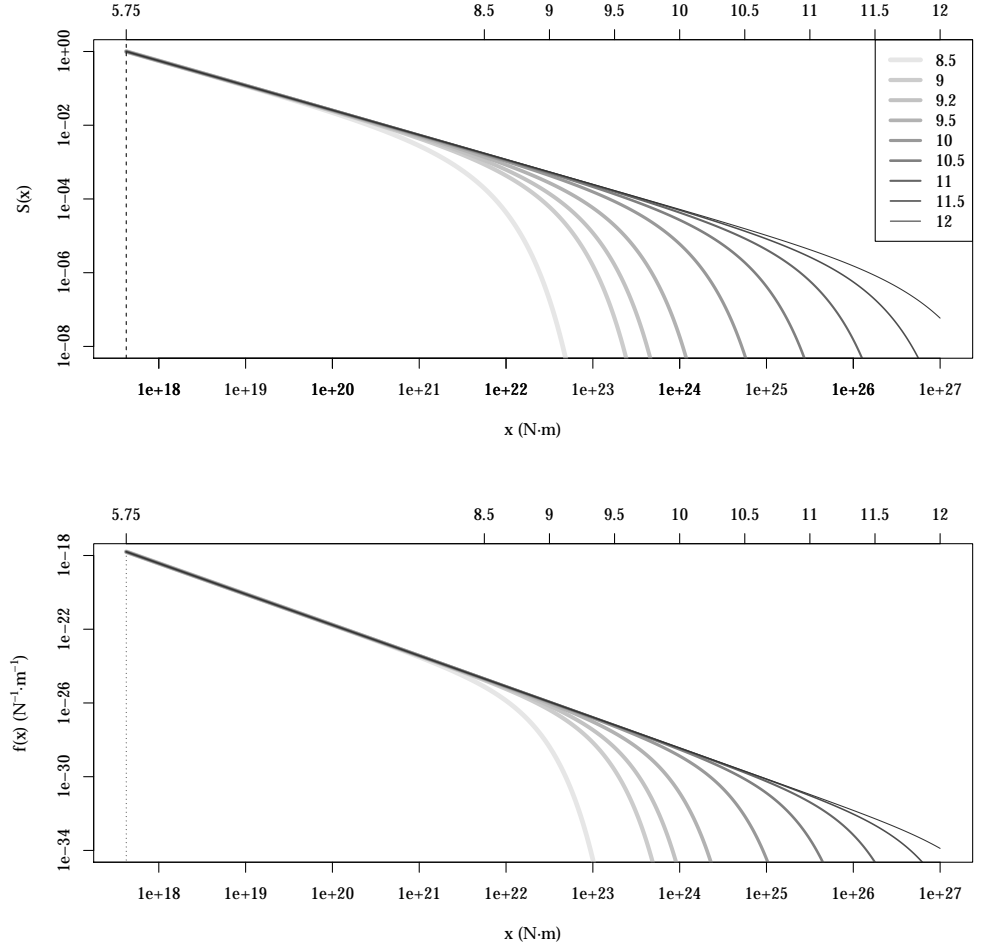


Fig 6. (Fig. S3) cdf $S(x)$ and pdf $f(x)$ of TrG distribution with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend.

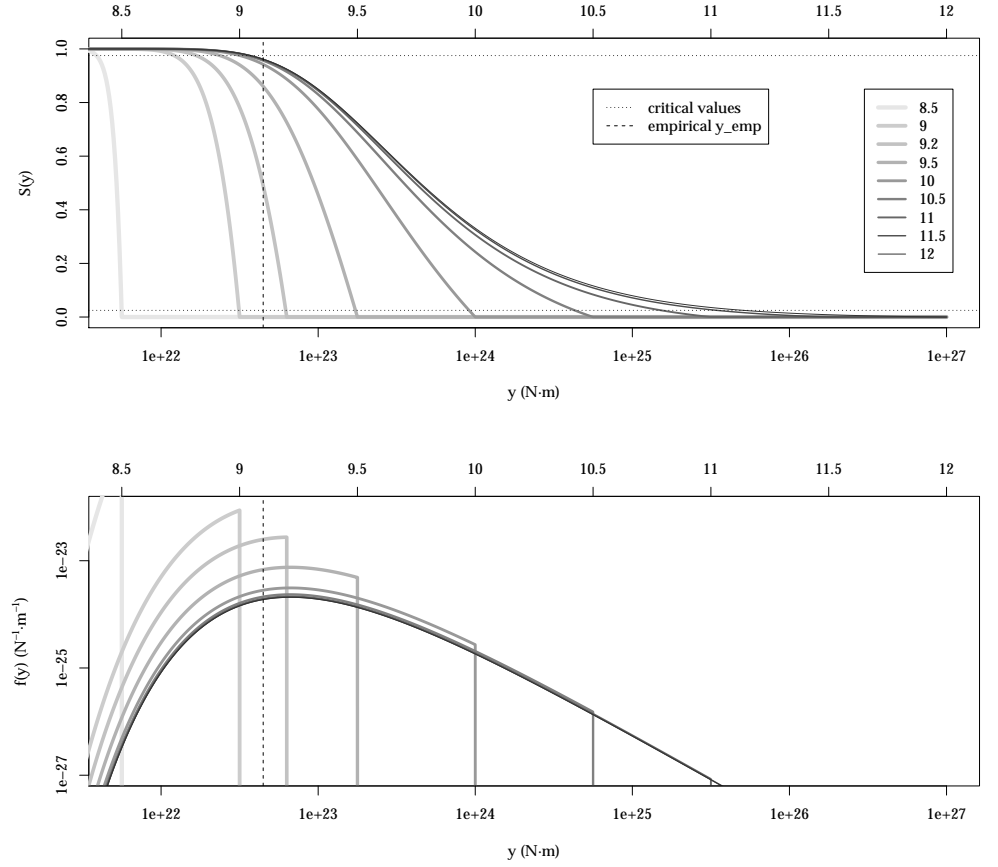


Fig 7. (Fig. S4) cdf $S_{max}(y)$ and pdf $f_{max}(y)$ of the maximum of 7,585 TPL observations with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend. Critical values at the 95% confidence level are shown as horizontal lines. Empirical value of maximum seismic moment observed is shown as a vertical line. Note that this is exactly Fig. 1 of the main text, repeated here for completeness.

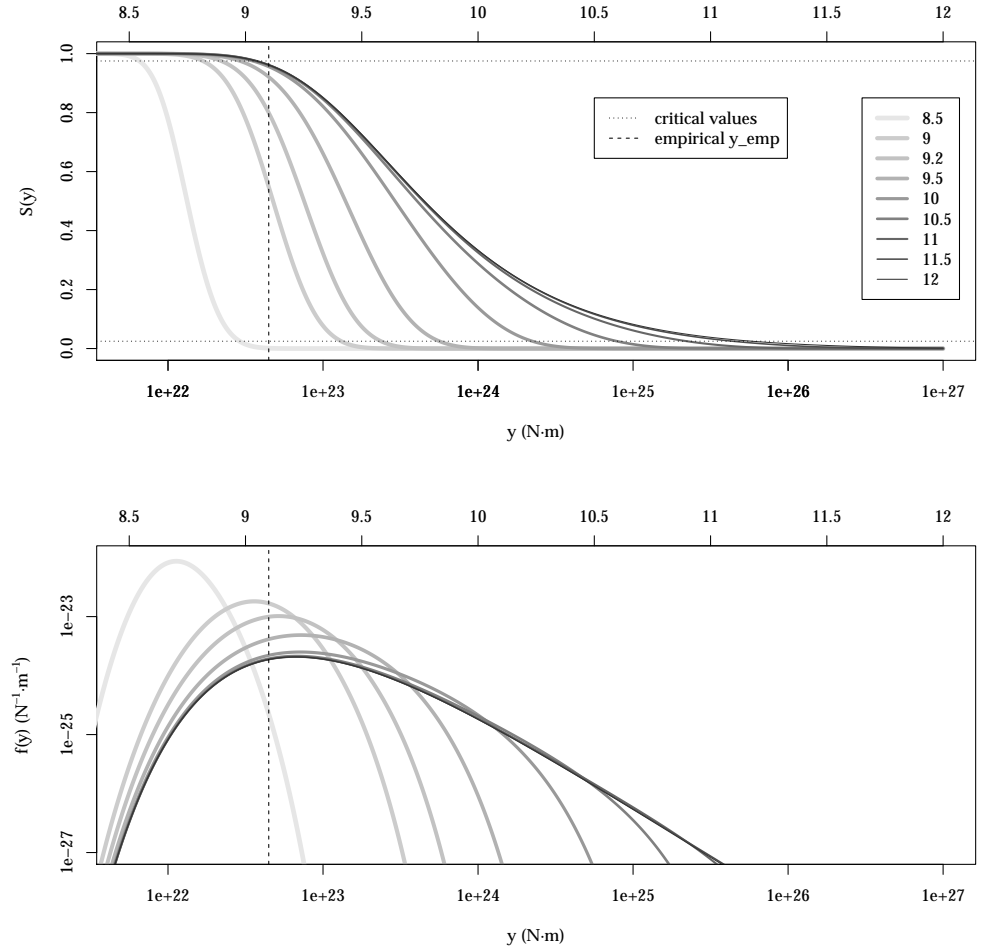


Fig 8. (Fig. S5) cdf $S_{max}(y)$ and pdf $f_{max}(y)$ of the maximum of 7,585 Tap observations with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend. Critical values at the 95% confidence level are shown as horizontal lines. Empirical value of maximum seismic moment observed is shown as a vertical line.

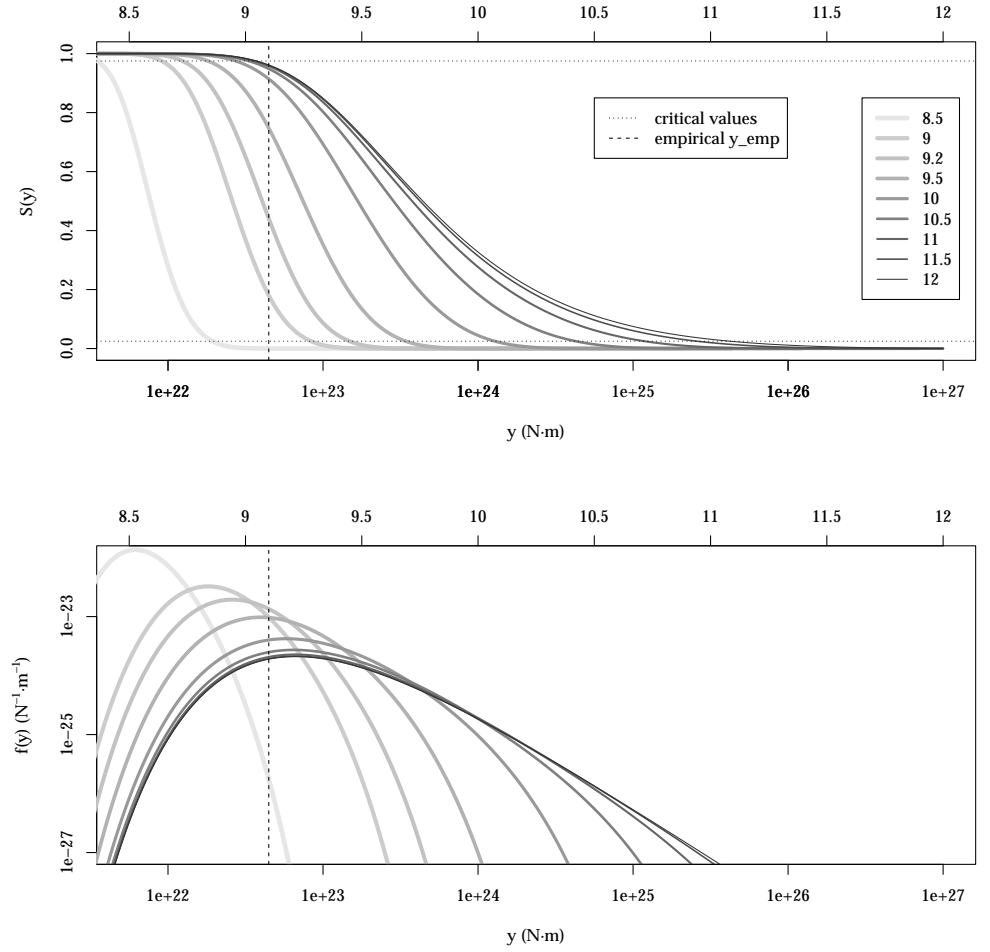


Fig 9. (Fig. S6) cdf $S_{max}(y)$ and pdf $f_{max}(y)$ of the maximum of 7,585 TrG observations with $\beta = 0.67$, a corresponding to moment magnitude 5.75, and M_c corresponding to the values of m_c shown in the legend. Critical values at the 95% confidence level are shown as horizontal lines. Empirical value of maximum seismic moment observed is shown as a vertical line.