УДК 550.34.042.4 DOI: https://doi.org/10.24028/gj.v46i6.306222

Magnitude probabilities for extreme earthquakes around the globe with Rank-Ordering

G. Domej[1,2], 2024

[1]Adam Mickiewicz University, Faculty of Geographical and Geological Sciences, Institute of Geoecology and Geoinformation, Department of Geomorphology, Poznan, Poland
[2]Geological Survey of Slovenia, Department of Regional Geology, Ljubljana, Slovenia
Received 27 June 2024

Earthquake likelihoods have occupied humankind ever since, and the estimation of potential magnitudes is crucial for a multitude of aspects of safety. In this work, we present a probabilistic analysis of extreme magnitudes in 16 regions across the globe characterized by different seismicity to invert the traditional question of «what probability is associated with certain magnitudes». We combine the Gutenberg-Richter Law and Rank-Ordering-Statistics in a methodological approach to estimate what magnitude ranges can be almost certainly (i.e., with 95%) expected, and what magnitudes become reasonably unlikely beyond those ranges.

This approach allows for estimating probabilities for maximal magnitudes per region and comparing thereto the maximal magnitudes (m_r) that appeared in reality. The method explores the maximal magnitudes that could occur or be exceeded with a probability of 95%, if the respective m_r are equal to or greater than these 95%-predictions, and how probable it is, that also these m_r could be reproduced or exceeded. We suspect a lack of great magnitudes in the Alps and a surplus across the Atlantic Ocean from these statistical considerations.

Key words: Gutenberg-Richter Law, Rank-Ordering-Statistics, magnitude-frequency distribution, magnitude probability, b-value, geostatistics.

1. Introduction. From the multitude of geohazards, earthquakes belong to those few phenomena that remain to this day unforeseeable, which, ever since, pose a considerable threat to people and infrastructure. Over centuries, efforts were undertaken to associate earthquakes with other phenomena such as, e.g., astronomical phases and the Earth's tides, animal behavior, and geophysical measurable processes, with the result that — in contrast to other geohazards — earthquakes do not announce themselves. Even if tremors can let expect an event, precursor times are usually too short to take measures to prevent disasters, and magnitudes are incalculable in advance. The literature discusses only a very small number of earthquakes being — possibly — predicted. One of them is the Haicheng Earthquake that struck North-Eastern China on 4th of February 1975 with a moment magnitude (m_w[1]) of 7.0 [USGS, 2023], but, similar to the other cases, vivid discussion is to be found scrutinizing the true success (e.g., [Wang et al., 2006]).

To master the problem of unpredictability, scientists have, so far, resorted to time-series analysis of event records and pro-

[1] For all variables cf. Section S1 in the supplement.

Citation: Domej, G. (2024). Magnitude probabilities for extreme earthquakes around the globe with Rank-Ordering. *Geofizychnyi Zhurnal*, *46*(6), 41—70. https://doi.org/10.24028/gj.v46i6.306222.

bability theory, which nowadays is an entire subfield of seismology with an abundance of literature whose overview would go beyond the scope of this publication.

Typical probabilistic approaches start with magnitude-frequency distributions and derive therefrom likelihoods of occurrence of certain magnitudes. The most popular, and thoroughly legitimate, question of analysis is: «How likely is it for an event of a certain magnitude to take place?» Figuratively the question is, thus, directed from the abscissa (usually representing the magnitudes) to the ordinate (usually representing the respective probabilities per magnitude) in a classic magnitude-frequency graph.

This work presents a statistical approach to earthquake probabilities in the opposite direction; instead of deducing event probabilities to certain magnitudes, another question is expressed from the ordinate towards the abscissa: «Assuming an almost certain event probability (i.e., 95%), what magnitudes can be reached, and which magnitudes beyond that range become, then, reasonably unlikely?» This question simultaneously sheds light on the assumption that the Gutenberg-Richter Law must have limits at certain magnitude ranges, even though formulated as an exponential decrease striving towards but never reaching zero.

The present study consists of a comparative and probabilistic analysis of the extreme magnitudes in 16 regions across the globe characterized by specific tectonic settings and, hence, seismicity, and by different sizes and numbers of earthquakes. We combine the Gutenberg-Richter Law [Gutenberg & Richter, 1944, 1956] and Rank-Ordering-Statistics — a statistical method initially proposed by Zipf [1949] in social science and later adapted for other fields such as geosciences by Sornette et al. [1996]. This methodological approach allows for the estimation of probabilities for maximal magnitudes per region and thereto the comparison of the maximal magnitudes (m) that appeared in reality. The method explores the maximal magnitudes that could occur or be exceeded with a probability of 95%, if the respective m_r are equal to or greater than these 95%-predictions, and how probable it is, that also these m could be reproduced or exceeded.

2. Data. For this study, we retrieved more than one million earthquakes in 16 regions across the globe (Fig. 1, Table 1) from the Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat) of the USGS [2009]; hereafter, regions are abbreviated with the letter «R» and their number. As the Gutenberg-Richter Law should be valid in any geographic and tectonic environment [Gutenberg & Richter, 1944, 1956; Stein & Wysession, 2003], different spatial extents and locations were chosen deliberately for these partly overlapping regions. In principle, all regions appear rectangular according to latitude and longitude — except Antarctica (R16), whose rectangle is reduced to a circle as it covers the South Pole. Different tectonic settings are accounted for: strong seismic activity and volcanism along the «Ring of Fire» around the Pacific Ocean (Japan, R12; Kamchatka, R13; South Asia, R14; New Zealand, R15); convergent tectonics of different intensities and subductions zones (South America, R4; the Alps, R7; Iran, R9; Central Asia, R11); divergent tectonics of different intensities (Atlantic Ocean, R5; Africa, R10); regions dominated by one or several major fault systems (the Western USA, R1; Central America, R3; Algeria, R6; Turkey, R8); and regions that are rarely affected by strong earthquakes (the Eastern USA, R2; Antarctica, R16). Also, the time frame covered by the individual datasets is slightly variable. Initially, a period of 20 years (1990—2009) should have been taken into account in each region. Due to sparse data, e.g., in geographically small regions or regions where the seismic network is not dense, five datasets were enlarged to 49, 50, and 56 years (Table 1) to ensure at least 1000 earthquakes per region.

3. Methodology. The Gutenberg-Richter Law describes the exponential relationship between a particular magnitude m and the number of earthquakes $N(m)$ per year with at least this magnitude [Gutenberg & Richter, 1944, 1956] in any considered region:

Table 1. Regions of different seismicity with their numbers of considered earthquakes, covered areas, and timespans of the datasets

Region		Earth-quakes	Latitude, [°]	Longitude, [°]	Time span
1	Western USA	842,600	50 N—30 N	135 W—100 W	1990—2009
2	Eastern USA	4,745	45 N—25 N	95 W—70 W	1990—2009
3	Central America	38,647	30 N—5 N	120 W—50 W	1990—2009
4	South America	38,029	5 N—55 S	85 W—60 W	1990—2009
5	Atlantic Ocean	4,809	70 N—25 N	55 W—10 W	1960—2009
6	Algeria	1,948	38 N—33 N	2 W—9 E	1954—2009
7	Alps	9,641	49 N—45 N	4 E—19 E	1990—2009
8	Turkey	51,336	45 N—35 N	15 E—50 E	1990—2009
9	Iran	3,298	40 N—25 N	45 E—65 E	1990—2009
10	Africa	2,913	15 N—35 S	20 E—50 E	1961—2009
11	Central Asia	12,874	50 N—20 N	70 E—105 E	1990—2009
12	Japan	28,658	45 N —5 N	125 E—155 E	1990—2009
13	Kamchatka	12,586	73 N—45 N	150 E—180 E	1990—2009
14	South Asia	51,657	20 N—15 S	100 E—160 E	1990—2009
15	New Zealand	6,503	30 S—50 S	165 E—180 E	1954—2009
16	Antarctica	1,272	60 S—90 S	180 W—180 E	1960—2009

Note: Data were retrieved from the Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat [USGS, 2009]).

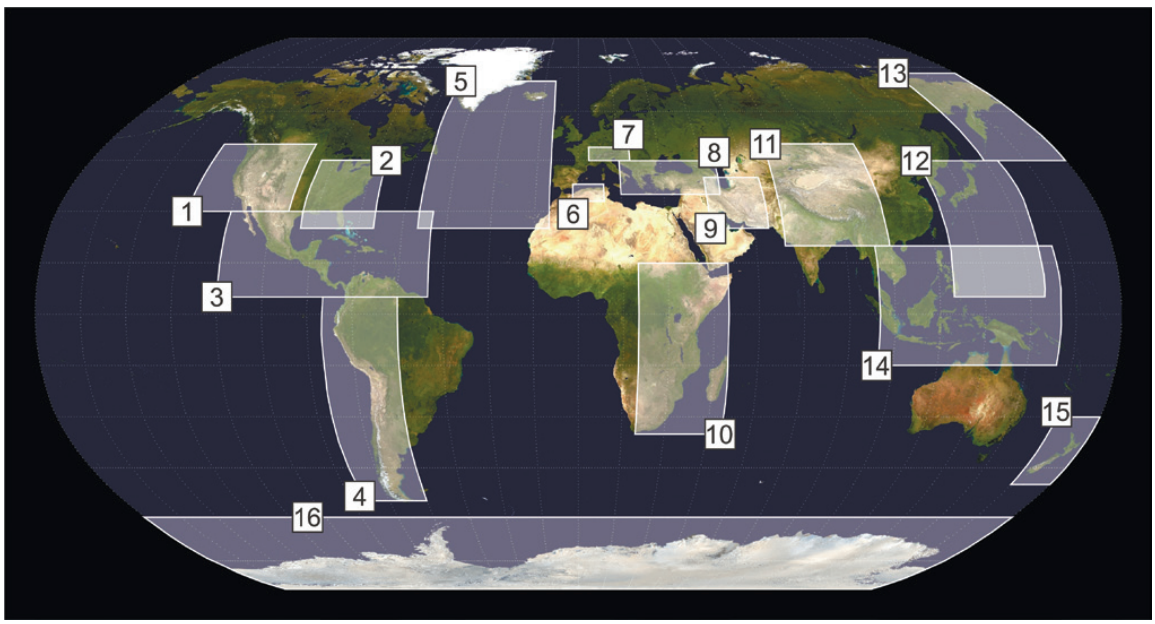

Fig. 1. Regions of different seismicity on a Robinson Projection [NASA, 2020]. Covered areas are given in Table 1.

$$\log_{10} N(m) = a - bm\,. \qquad (1)$$

In this equation, the constant a is sometimes referred to as the seismic activity in a region and time frame, as it directly links to the number of considered earthquakes.

Likewise, the gradient — termed «b-value» — is characteristic for different regions and geotectonic settings (e.g., [Turcotte, 1997; Utsu, 2002a; Spada et al., 2013; Godano et al., 2014]). It usually varies around 1.0 with a range of ±0.2 in seismically active regions [Clauser, 2014]. Substantial deviations from the usual b-value range are reported to be caused by earthquake swarms (e.g., b-values up to 2.5; e.g., [Stein & Wysession, 2003]), related to specific earthquake physics and source environments or geographic areas (e.g., b-values from 0.5 to 2.0; e.g., [Mogi, 1962, 1967; Mori & Abercrombie, 1997; Schorlemmer et al., 2005; Amaro-Mellado & Tien Bui, 2020]), or linked to induced seismicity (e.g., b-values up to 1.9; e.g., [Urban et al., 2016]). Another reason for b-value deviations are incomplete datasets. In this context, it is essential for various types of analysis to determine the magnitude above which a particular earthquake record is complete; hereafter, this magnitude will be called «m_0» and is not to be confounded with the seismic moment (M_0). Although termed «time-frame-independent», the b-value can become altered by earthquake records with a deficit of great magnitudes, e.g., if by chance a short time frame is chosen, during which no great magnitude occurred. Hence, both the b-value as well as m_0 are essential for a correct approximation of the Gutenberg-Richter Law in a particular region; the two parameters will play a significant role in the probability estimation for the maximal magnitudes presented in this work.

Moreover, the b-value of the Gutenberg-Richter Law entails a 10^b-fold decrease of earthquakes per increase of entire magnitudes. With a b-value equal to 1.0 and an exemplary constant a of 6.0, we expect 100 earthquakes with a magnitude ≥4.0, ten with a magnitude ≥5.0, and only one with a magnitude ≥6.0. However, as statistical fluctuations are calculated via the square root of the number of random and independent events (i.e., $\sqrt{n} = n^{0.5}$; [Hergarten, 2002]) — such as, e.g., earthquakes — it becomes apparent that earthquake distributions deliver uncertain statistic statements at great magnitudes. In the example mentioned before (i.e., with b=1 and a=6), the fluctuation for ten earthquakes amounts to three, and for one earthquake, the fluctuation is one just as well. Unfortunately, though, earthquake distributions at great magnitudes are the most significant when aiming to study probabilities of rare extreme events, which usually have amore severe impact in terms of deformation and damage than the more frequent smaller events.

Here, the Rank-Ordering-Statistics [Sornette, 2006] remedies this discrepancy by a conceptual axis-interchange in order to estimate the magnitude fluctuation at a specific number of earthquakes, rather than the fluctuation of the number of earthquakes per magnitude.

Applied for the first time in the context of earthquake distributions, probabilities, and exceedance predictions, Sornette et al. [1996] described Rank-Ordering-Statistics as useful to detect power-law distributions of under sampled data and to characterize sparsely sampled tails of earthquake distributions. Unlike cumulative earthquake statistics, which are usually governed by numerous small events, Rank-Ordering-Statistics consists of placing values x_i (e.g., magnitudes) of a dataset with I entries (e.g., an earthquake dataset) in descending order and attributing to each value a rank i — with rank 1 hosting the highest of all values x_i, where i is in the range of 1 to I. As Sornette [2006] generalizes, the method has a close link to a cumulative distribution $P(x)$, giving the probability of obtaining values equal to or greater than x:

$$P(x) = \int_x^\infty p(x)dx\,, \qquad (2)$$

where $p(x)$ is the probability density function.

Multiplying the probability $P(x_i)$ for the i^{th} value by I, and taking the integer part, provides the number of values equal to or greater than x_i, which corresponds to the rank of x_i as well:

$$IP(x_i) = i\,. \quad (3)$$

If the probability function is known, it is then possible to appoint ranks i and the respective values x_i.

3.1. Implementation of the probability function. Generally expressed, probabilities rely on the «ratio of what is favorable compared to what is possible». Considering that $N(m)$ in the Gutenberg-Richter Law gives the number of earthquakes per year with at least this magnitude, and the fact that earthquake records are complete only after m_0, the probability of randomly picking an earthquake with at least one specific magnitude m out of the entire dataset is:

$$P(m) = N(m)/N(m_0)\,. \quad (4)$$

By taking the logarithm and substituting with the Gutenberg-Richter Law (Eq. 1), we obtain:

$$\log_{10} P(m) = a - bm - \log_{10} N(m_0)\,. \quad (5)$$

Using this general principle, but considering only earthquakes with magnitudes greater than m_0, the probability of randomly picking an earthquake with at least one specific magnitude (but greater than m_0) is:

$$P(m_0) = N(m_0)/N(m_0) = 1\,, \quad (6)$$

$$\log_{10} P(m_0) = 0\,. \quad (7)$$

Remaining uniquely in the range where the earthquake records are complete, we again use the general principle (Eq. 5) and substitute with Eq. 7 before rearranging for the constant a:

$$0 = a - bm_0 - \log_{10} N(m_0)\,, \quad (8)$$

$$a = bm_0 + \log_{10} N(m_0)\,. \quad (9)$$

Stepping back to the general principle, we substitute the constant a in Eq. 5 with Eq. 9:

$$\log_{10} P(m) = bm_0 + \log_{10} N(m_0) - \\ -bm - \log_{10} N(m_0)\,, \quad (10)$$

$$\log_{10} P(m) = b(m_0 - m)\,. \quad (11)$$

This probability function to randomly pick an earthquake with at least one specific magnitude m out of the entire datasethas two essential advantages. First, m_0 can be defined individually for each region, and second, the constant a does not appear any longer in the equation rendering it number- (and hence) time-independent — i.e., scale-invariant.

Moreover, it ensures that the gradient — i.e., the b-value — refers only to the range of magnitudes greater than m_0. It is suggested to use a negative gradient since the equation describes an exponential decay:

$$P(m) = 10^{-b(m-m_0)}\,. \quad (12)$$

Having defined a probability function for picking an earthquake with at least one specific magnitude m out of the entire dataset, we develop thereupon another function that gives the probabilities to pick exactly those earthquakes that are ranked highest according to their magnitude. Therefore, we consider the following four probabilistic expressions for picking:

- one earthquake with a magnitude $\geq m$: $P(m)$;
- one earthquake with a magnitude $< m$: $1–P(m)$;
- all earthquakes with magnitudes $< m$: $(1–P(m))^n$;
- the earthquake with the greatest magnitude $\geq m$: $1–(1–P(m))^n$.

From the last expression, we can directly deduce a probability function to pick out of the entire dataset the earthquake with the greatest magnitude, as there is exactly one of this kind per dataset:

$$P_{1st} = 1 - \left(1 - P(m)\right)^n\,. \quad (13)$$

In a general form [Sornette, 2006], this equation appears as follows, allowing likewise for the computation of the probability functions for the earthquakes with magnitudes of the subsequent ranks r (i.e., 2nd, 3rd, 4th, and 5th in this study):

$$P_r(m) = 1 - \sum_{k=0}^{r-1} \binom{n}{k} P(m)^k \left(1 - P(m)\right)^{n-k}\,, \quad (14)$$

where n is the number of earthquakes starting from m_0.

Finally, we substitute $P(m)$ in Eq. 14 with Eq. 12:

$$P_r(m) = 1 - \sum_{k=0}^{r-1} \binom{n}{k} \left(10^{-b(m-m_0)}\right)^k \times \times \left(1 - 10^{-b(m-m_0)}\right)^{n-k}. \quad (15)$$

Theoretically, any ranking range can be evaluated using Eq. 15. For simplicity, on the one hand, we evaluated only ranks 1 to 5; on the other hand, this choice seemed to reflect well the number of extreme magnitudes per region not following the Gutenberg-Richter Law (Section S6 in the supplement).

One crucial detail is that in this study, the total number of earthquakes throughout the time spans per region (Table 1) was considered, although the Gutenberg-Richter Law suggests a division of the number of earthquakes by the number of years. The motivation behind this decision is, on the one hand, a better representation of the relationship between magnitudes and their frequencies. On the other hand, a division by the number of years should — in theory — imply only a reduction of the constant a, which is not present anymore in the probability functions since Eq. 11 and, therefore, not relevant in this study. Strictly speaking, thus, the Gutenberg-Richter Law is not represented in its original form, and we call its «non-divided» from a magnitude-frequency distribution (MFD) hereafter and where applicable.

3.2. Estimation of m_0 and the b-value. The two unknown terms in the Rank-Ordering probability function (Eq. 15) are m_0 and the b-value. Both variables are linked to each other; if m_0 is chosen incorrectly, it is likely to alter the b-value and, hence, falsify the seismic character of the considered region.

In the first step, earthquake data per region were grouped into magnitude intervals of 0.1 by rounding down all hundredths to their lower tenths — e.g., magnitudes from 2.10 to 2.19 belong to the interval of 2.1. Negative as well as zero magnitudes (i.e., those of the interval of 0.0) were deleted. The total numbers of earthquakes used in the analyses of the individual 16 regions are given in Table 1.

Then, magnitude distributions were represented in a y-semi-logarithmic plot as incremental and cumulative values to approximate a polynomial of the first degree with MATLAB's Curve Fitting Toolbox [MATLAB, 2019] representing the MFD (Fig. 2, *a*). This polynomial approximation in a y-semi-logarithmic plot equals the process of fitting a linear regression in non-logarithmic plots. It should be noted that the MFD can be approximated either to the incremental or the cumulative point sets, as it can be mathematically shown that the resulting curves of both point sets differ only by a constant — provided that only magnitudes above m_0 are considered:

$$\log_{10} N(m) = a - bm + \left(\frac{1}{\ln(10)b}\right). \quad (16)$$

In this study, MFD were approximated for each of the 16 regions between a visually identified m_0 and the greatest magnitude (m_{10}) at which at least ten earthquakes were recorded to prevent b-value alterations caused by the occurrence or lack of particularly great magnitudes. Allowing for better regression fitting due to a higher sampling number, the cumulative curve assisted as a gradient indicator and simultaneously at the correct identification of m_0. Only the b-value was retained from each approximation, as the constant a is not considered in the following probability estimation (Eq. 15).

3.3. Estimation of the probabilities for maximal magnitudes. Having defined the two unknown terms in the Rank-Ordering probability function (Eq. 15) — i.e., m_0 and the b-value — individually for each of the 16 regions, the function is used to compute the probability curves for the maximal magnitudes (Fig. 2, b) based on the complete cumulative point sets, as the Gutenberg-Richter Law and, hence, the MFD originally counts earthquakes with at least a particular magnitude. Plots (Fig. 2, b) explore the following three questions with respect to the maximal magnitudes, on which the interpretation of results is based:

(i) Which are the maximal magnitudes (m) that could occur or be exceeded with a probability of 95%?

(ii) Which are the maximal magnitudes (m_r) that appeared in reality, and are they equal to or greater than these 95%-predictions?

(iii) How probable is it, that also these m_r could be reproduced or exceeded?

3.4. Magnitude types used in this study. One crucial aspect during data preparation was the magnitude type to consider. Usually, the Gutenberg-Richter Law is established for a so-called «unified magnitude» adjusted from body-wave, local, and surface-wave magnitudes [Gutenberg & Richter, 1956]. The ANSS ComCat [USGS, 2009] data retrieved for the 16 regions include, though, a variety of magnitude types (Section S2 in the supplement) such as, amongst others, body-wave magnitudes (m_b), duration magnitudes (m_d), local magnitudes (m_l), surface-wave magnitudes (m_s), and moment magnitudes (m_w).

To answer the question, of whether all magnitudes should be used independently of their type, or if preference should be given for one particular magnitude type, a separate test with different earthquake data and a different geographic framework was conducted. Instrumental earthquake data from the International Seismological Centre (ISC) Bulletin covering a time span from 1900 to 2019 [Storchak et al., 2017, 2020; ISC, 2020] were first tailored to cover the Alps with a 500 km wide buffer zone, resulting in a set of 52,511

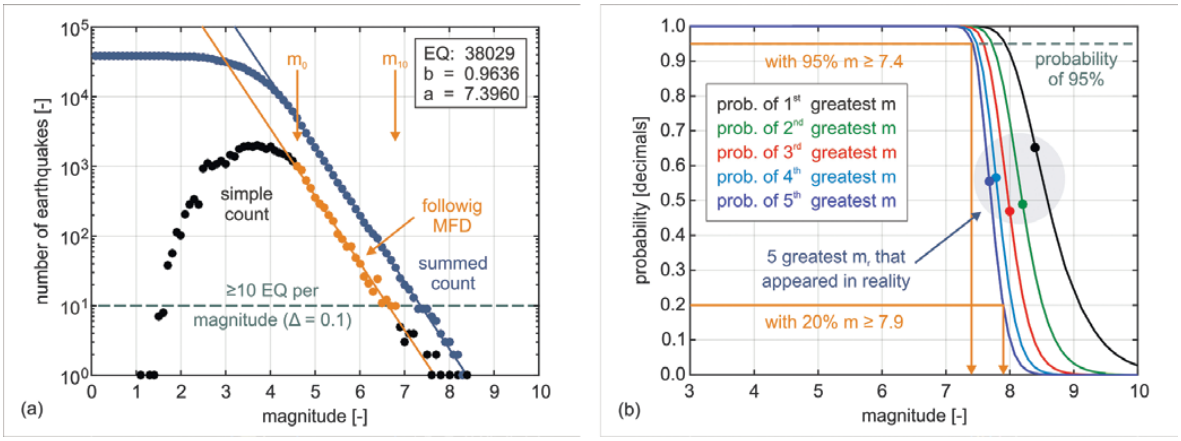

Fig. 2. Example of result curves: approximation of a magnitude-frequency distribution (MFD) via linear regression fitting (a); probability estimation for the maximal magnitudes via Rank-Ordering-Statistics (b). The curves represent the results of South America (R4).

Table 2. Numbers of earthquakes per magnitude type and the respective b-values for the «Alps Test»

Magnitude types for the «Alps Test» (ISC Bulletin)	Number	Set fraction	b-value
All magnitude types (M)	*48,099	*91.60%	0.8930
Undefined magnitude (m)	697	1.33%	1.6470
Body-wave magnitude (m_b)	1,403	2.67%	1.5820
Duration magnitude (m_d)	5,761	10.97%	1.0340
Local magnitude (m_l)	39,727	75.65%	1.9540
Surface-wave magnitude (m_s)	15	0.03%	too few data
Moment magnitude (m_w)	228	0.43%	too few data
Unknown magnitude (m_{ukn})	268	0.51%	too few data
No magnitude given for earthquakes (none)	*4,412	*8.40%	lack of data

Note: The set fractions refer to the entire dataset of 52,511 earthquakes (*). Data were retrieved from the ISC Bulletin [Storchak et al., 2017, 2020; ISC, 2020].

earthquakes, and then grouped according to the prevalent magnitude types. Table 2 shows set fractions in numbers and percent relative to the entire dataset. A small fraction of 8.4% does not report magnitudes at all and is not included in the analyses; the combination of all «usable» magnitudes regardless of their type amounts to the complementary 91.6% of the dataset.

Via linear regression fitting, the MFD was approximated to all subsets filtered by magnitude type — where sufficient data was available — as well as to the set of magnitude type M consisting of a combination of all subsets. The resulting b-values are shown in the last column of Table 2.

Because b-values generally vary closely around 1.0, it is apparent that the subset of local magnitudes (m_l) results in a value that is by far too high (i.e., 1.9540). According to Schorlemmer et al. [2005], extraordinarily high b-values would, moreover, be characteristic of divergent tectonic settings, which is not the case in the Alps. Only the b-value of the subset of duration magnitudes (m_d) and the one of the combined set (i.e., the one of magnitude type M) seem to have a trustworthy b-value (1.0340 and 0.8930, respectively; Fig. 3, a, Table 2). Duration magnitudes (m_d), though, amount to only 10.97% of the entire dataset and are generally rather uncommon to use in statistical studies since they relate to the recorded time span and not to the amplitude of earthquakes.

As the b-value is characteristic of a particular region (e.g., [Godano et al., 2014]), and since filtering a dataset should — in the theory of the Gutenberg-Richter Law — not alter the seismic character of the respective region, it seems more appropriate to use the most complete dataset even though magnitudes are not reported in the same type. Another argument against filtering according to magnitude types is that — being by nature seldom phenomena — earthquakes with great magnitudes are very likely to be discarded (Fig. 3), which is particularly disadvantageous if precisely those should be the target of the probabilistic study; also, their representation in MFD is essential in proper linear regression fitting between m_0 and m_{10}, especially when the event numbers are small (Section S3 in the supplement). Moreover, the logic subsequent step after filtering datasets by magnitude types would be subset conversions into one type of magnitude. However, such conversions are usually empirical and reflect strong regional and/or local characteristics of respective areas of study. For areas as large as the 16 regions of this study, such magnitude conversions would make little sense and even falsify data to a great extent.

Therefore, we expect the most accurate results for approximations of the MFD to earthquake data by comprising various magnitude types by considering the entire set instead of

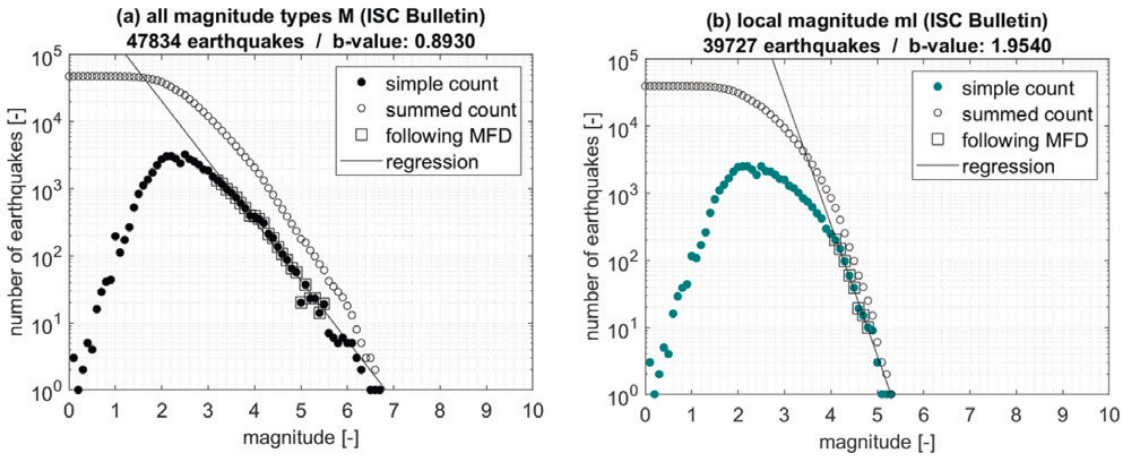

Fig. 3. Approximation of a magnitude-frequency distribution (MFD) via linear regression fitting to all magnitude types (M (a)) and to local magnitudes (m_l (b)). Filtering according to magnitude types is likely to alter the b-value and discard great magnitudes.

a filtered and/or truncated version, even if this solution is not ideal.

We assume that magnitudes of extreme events, which might be reported in different magnitude types and — thus — appear smaller or greater than what their respective moment magnitude (m_w) could indicate, would not bother this approach, as regressions are fitted only up to m_{10}.

Anticipating the results of the b-value estimation using the ANSS ComCat [USGS, 2009] (presented in the next section), the approach of using the set of magnitude type *M* is emphasized by the fact that the estimated b-value for the Alps (R7; Table 3) is 0.9147. In comparison, the b-value estimated using the ISC Bulletin [Storchak et al., 2017, 2020; ISC, 2020] resulted in 0.8930. Both values are very close to each other and, hence, support not only the use of the entire dataset but also the actual b-value computation itself. Gulia & Wiemer [2010] report b-values ranging from 0.75±0.02 to 0.92±0.04 for a large part of the Alps.

4. Results and discussion. Results of the approximation of the MFD via linear regression fitting between m_0 and m_{10} are shown in Table 3. One example is illustrated in Fig. 2, *a* for South America (R4); the other regions are shown in Fig. 7, a and 8, a, and Section S6 (in the supplement). The root-mean-square-errors (RMSE) obtained for the 16 regions indicate particularly satisfying regressions with the greatest deviation from a perfect fit (i.e., RMSE=0) amounting to 0.16 in New Zealand (R15) and to 0.11 and 0.10 in the Atlantic Ocean (R5) and Central Asia (R11), respectively. Here, however, it should be considered, that polynomial curve fitting in MATLAB is based on the Least-Square Method reported to possibly produce bias [Sandri & Marzocchi, 2007].

Reasons for these slightly higher RMSE might be the scatter within greater magnitudes, which — apparently — could not be obliterated by truncation of the datasets to m_{10}. Depending on the considered region, m_0 varies between 2.0 and 5.4, with higher m_0 reflecting either high general seismicity or the need for more complete earthquake records at lower magnitudes — e.g., due to high population densities and the aspect of hazard prevention. The Western USA (R1), the Eastern USA (R2), and the Alps (R7) show m_0 of 2.9, 2.4, and 2.8, respectively, whereas for Antarctica (R16), m_0 fits with 5.4. New Zealand (R15) — with a m_0 of 5.2 — seems to step out of the line, being a seismically very active, well-surveyed, and locally densely populated region.

As apparent from the methodology section, the b-value is the most critical for the subsequent probability estimations for the maximal magnitudes. It is, therefore, of particular interest how well the b-values presented in this study correspond to those of other authors. Here, Cheng & Sun [2018] provide a global reference for comparison, as they computed b-values for almost all 50 Flinn-Engdahl Regions representing the ma-

Table 3. Linear regression parameters for the 16 regions of this study

Region		Linear regression		
		m_0	RMSE	$b\pm\Delta b$
1	Western USA	2.0	0.0546	0.8892±0.02
2	Eastern USA	2.4	0.0876	0.7926±0.10
3	Central America	4.4	0.0799	0.9652±0.06
4	South America	4.6	0.0754	0.9636±0.05
5	Atlantic Ocean	4.5	0.1108	1.3040±0.13
6	Algeria	4.4	0.0666	0.9691±0.32
7	Alps	2.8	0.0428	0.9147±0.04
8	Turkey	3.8	0.0809	1.0570±0.06
9	Iran	4.4	0.0849	1.2610±0.16
10	Africa	4.8	0.0693	1.3250±0.18
11	Central Asia	4.6	0.1039	1.1300±0.10
12	Japan	4.8	0.0890	1.0870±0.07
13	Kamchatka	4.8	0.0865	1.0890±0.08
14	South Asia	4.8	0.0802	1.0750±0.05
15	New Zealand	5.2	0.1611	1.1630±0.35
16	Antarctica	5.4	0.0801	1.0160±0.30

Note: m_0 — magnitudes above which earthquake records are complete; RMSE — root-mean-square-errors of the regressions; b-values with their errors.

jor seismic zones of the Earth [Flinn et al., 1974]. Via an overlay of the 16 regions of this study onto the map established by Cheng & Sun [2018], we found in most of the cases a very satisfying fit of b-values within the proposed ranges taking into account that the 16 regions usually cover parts of one or several Flinn-Engdahl Regions (Fig. 4, Section S4 in the supplement). Only in Iran (R9), our b-value (i.e., 1.2610) is greater than the proposed range of 1.02 to 1.07 by Cheng & Sun [2018], which could be due to the fact, that the overlay covered predominantly the Flinn-Engdahl Region 29, which is much larger than our R9. One conjecture could be, therefore, that the authors had several more great magnitudes in their set, entailing a lower b-value range. In contrast, our b-values for Algeria (R6) and the Alps (R7) — i.e., 0.9691 and 0.9147 — are smaller than the proposed ranges of 1.17—1.21 and 1.17—1.32, respectively. Again, the covered Flinn-Engdahl Regions 31 and 36 are larger than our R6 and R7, which might have brought more low-magnitude earthquakes to the sets of Cheng & Sun [2018]. Another factor contributing to potential imprecisions of b-values might be small sample sizes; according to Kamer [2014], there is a minimum sample size of more than 2000 above m_0 for the correct estimation of the b-value with a resolution of 0.1, which is — at least in the case of Algeria (R6) — not met. For Antarctica (R16), the corresponding main region is the Flinn-Engdahl Region 50, for which there is no data provided by Cheng & Sun [2018].

How sensitive b-values are for particular regions, is exemplified by Fig. 4. Given authors have assessed MFD in a variety of tectonic settings for subregions within the 16 regions of this study (Fig. 1). Some of their b-values correspond very much to those obtained by Cheng & Sun [2018] for the Flinn-Engdahl Regions and those presented in this study, whereas some others divert signifi-

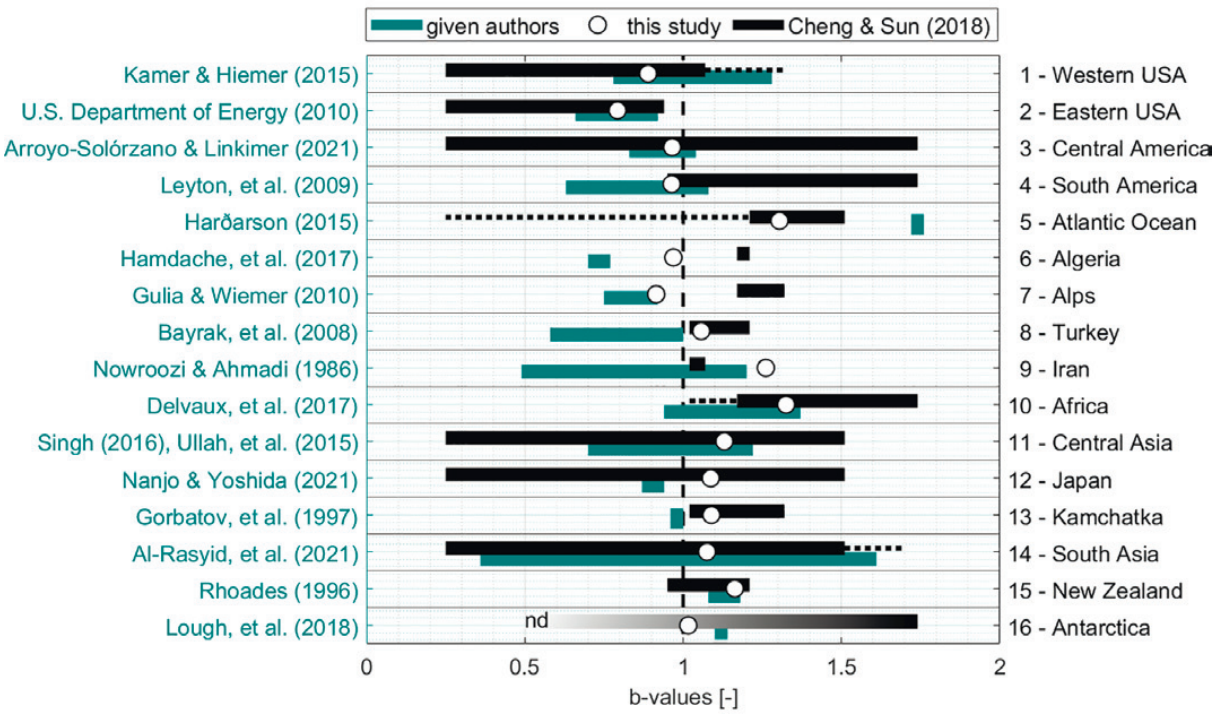

Fig. 4. Comparison of b-values of the 16 regions of this study with ranges for b-values given by various authors for different subregions within those 16 regions and with ranges resulting from overlapping those 16 regions with the 50 Flinn-Engdahl Regions [Flinn et al.,1974], for which Cheng & Sun [2018] computed b-values individually. Range decrease or increase caused by minor overlaps (i.e., when Flinn-Engdahl Regions cover only a very small part of the 16 regions of this study) are marked with dotted lines; «nd» stands for «no data».

cantly. As the target regions of the given authors are generally much smaller than the 50 Flinn-Engdahl Regions and the 16 regions of this study — i.e., possibly creating a bias in b-values —, we considered the better overlap of b-values calculated from the latter two as an indicator for a reasonably correct estimation of MFD and b-values in this study.

Also, the phenomena of b-value increase or decrease according to the tectonic setting can be traced to some extent in this study. Throughout literature — and on regional as well as on global scale — divergent tectonic settings are described to have the highest b-values, convergent settings the lowest, and transform settings some intermediate ones (e.g., [Frohlich & Davis, 1993; Schorlemmer et al., 2005; Gulia & Wiemer, 2010]). In this study, the 16 regions are mainly rectangular and enclose different tectonic settings, which hinders a clear subdivision into divergent, convergent, or transform regimes only. However, it is apparent that the two regions dominated by prominent divergence have by far the highest b-values — i.e., above 1.2: the Atlantic Ocean (R5) with its Mid-Atlantic Ridge, and Africa (R10) with the East African Rift System. The associated errors (Δb; Table 3) could indeed decrease the two b-values but do not change the general tendency compared to other regions. The third highest b-value in Iran (R9) lies still above 1.2; considering the relatively high associated error (Δb) and the fact that the prevalent tectonic setting is rather convergent than divergent, it might be an outlier. All other regions show b-values below 1.2, which might result from the dominance of divergent or transform tectonic settings or a combination of them. The Eastern USA (R2) reveal the lowest of all b-values, where one could have assumed the opposite due to an expected lack of great magnitudes within the continental interior. In this case, though, the rectangle in question also covers the New Madrid Seismic Zone — an intraplate zone of strong seismicity famous for the New Madrid Earthquake Series between 1811 and 1812 with (undefined) magnitudes up to about 7.5 [USGS, 2021].

Tendencies of reflecting the tectonic setting via b-values correspond, thus, to those previously published in the literature. However, it should be noted that b-values can vary considerably depending on the method of determination, the used earthquake data, and the considered magnitude ranges. Citing these three aspects, Frohlich & Davis [1993] report errors (Δb) of up to 30%, entailing a b-value range from 0.72 to 1.34, which comes close to the b-values obtained in this study.

Based on m_0 and the b-values for each of the 16 regions, we estimated the probabilities for maximal magnitudes via Rank-Ordering-Statistics following the methodology described above. One example is illustrated in Fig. 2, b for South America (R4); the other regions are shown in Fig. 7, b and 8, b, and Section S7 (in the supplement). Even though the plots can be used to deduce any magnitude-probability-relation, we focus on the interpretation of the three critical questions mentioned in Section 3.3., as we judge them to be particularly demonstrative and meaningful to interpret.

First, the question of which magnitudes could occur or be exceeded with a probability of 95% is answered by Fig. 5, a (black stars). The logically following question of which magnitudes (m_r) appeared in reality is given in the same figure (white triangles). By simple subtraction of the individual magnitudes, it becomes easy to see if these m_r are indeed equal to or greater than what is probable with 95%. The third question of how probable it is, that also these m_r could be reproduced or exceeded, is answered by Fig. 5, b: the colored points give the probabilities for magnitudes equal to or greater than m_r; the uncolored points include a magnitude step of +0.1 to account for the magnitude binning in steps of 0.1 during data processing.

Following this series of questions, the method predicts in most regions that, with 95% of probability, certain magnitudes are expected to occur or be exceeded; these estimates give a rough idea of the likely range of the maximal magnitudes to be expected in each region. In most regions, the five m_r respect this prediction and are either equal in magnitude or greater by up to one magnitude.

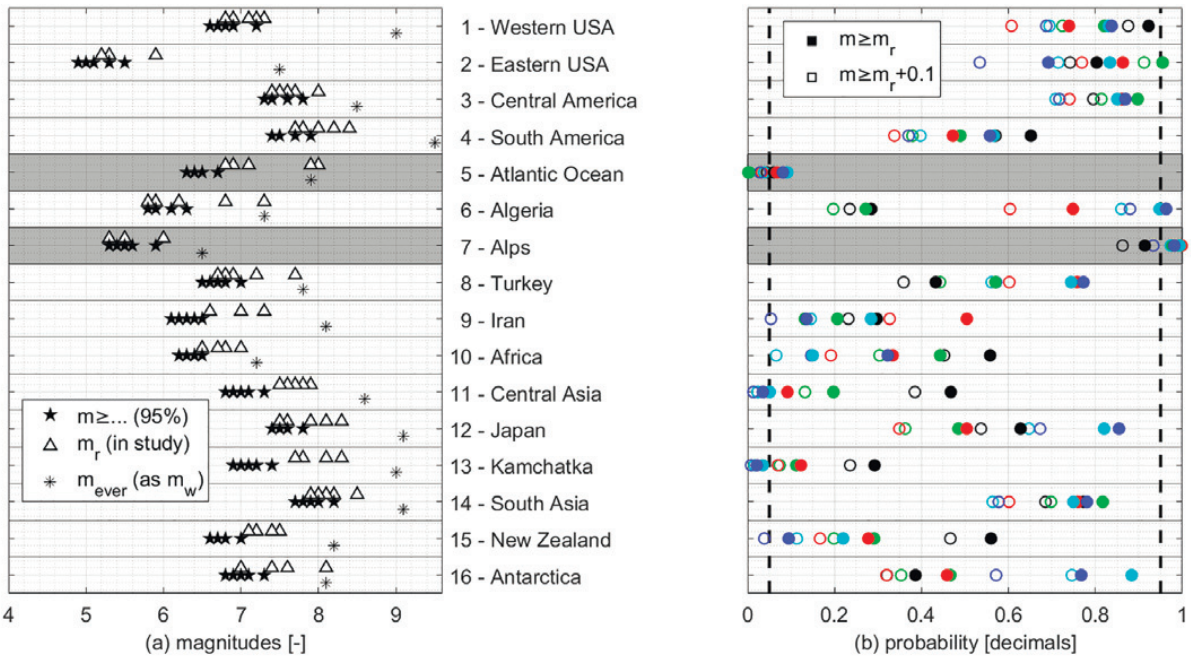

Fig. 5. Maximal magnitudes (m) that could occur or be exceeded with a probability of 95%; maximal magnitudes (m_r) that appeared in reality, and maximal magnitudes (m_{ever}) that occurred since registration up to February 2023 [USGS, 2023] (a). Probabilities for maximal magnitudes (m) to be equal to or greater than those magnitudes (m_r) that appeared in reality (colored points); probabilities for maximal magnitudes (m) to be equal to or greater than those magnitudes (m_r) that appeared in reality including a magnitude step of +0.1 (uncolored points; b), identifying the Atlantic Ocean (R5) and the Alps (R7) as «extreme» cases in terms of probabilities.

A representative example is South America (R4); from Fig. 2, b and Fig. 5, a (black stars), it is apparent that, with a probability of 95%, the maximal magnitudes could be equal to or greater than 7.9, 7.7, 7.5, 7.5, and 7.4, which is indeed the case as the respective m_r are 8.4, 8.2, 8.0, 7.8, and 7.7 (white triangles). According to the magnitude difference between the 95%-predictions and the respective m_r, the five points indicating m_r in the plots (Fig. 2, b) slide up and down the probability functions but stay in most regions somewhere within the center span arbitrarily set between 5% and 95%. The smaller the magnitude difference, the more upwards slide the five points indicating m_r, and the higher is the probability that also the five m_r could be reproduced or exceeded. For the example of South America (R4), we deduce from Fig. 2, b and Fig. 5 that it becomes moderately unlikely for the maximal magnitudes to be equal to or greater than the respective m_r; probabilities decrease for the maximal magnitudes in a series of 65%, 49%, 47%, 57%, and 56%. Including the magnitude step of 0.1, the «percentage-series» is even lower. Although most regions show a similar tendency — with probabilities not necessarily decreasing within one «percentage-series», as the example shows —, some particularities should be mentioned separately.

Assessing the variation in the individual «percentage-series», it emerges that not all of them show a significant range of variation (Fig. 5, b). Some of the probabilities to reproduce or exceed m_r lie in, above or below the center span from 5% and 95%, which leads to the assumption that those m_r should become extraordinarily likely or unlikely to be reproduced or exceeded. Here we should distinguish three cases:

(i) outliers which appear and disappear by adding a magnitude step of 0.1 to m_r (e.g., the Eastern USA (R2), Algeria (R6), and New Zealand (R15)), and, hence, seem to be negligible;

(ii) outliers that appear regardless of the

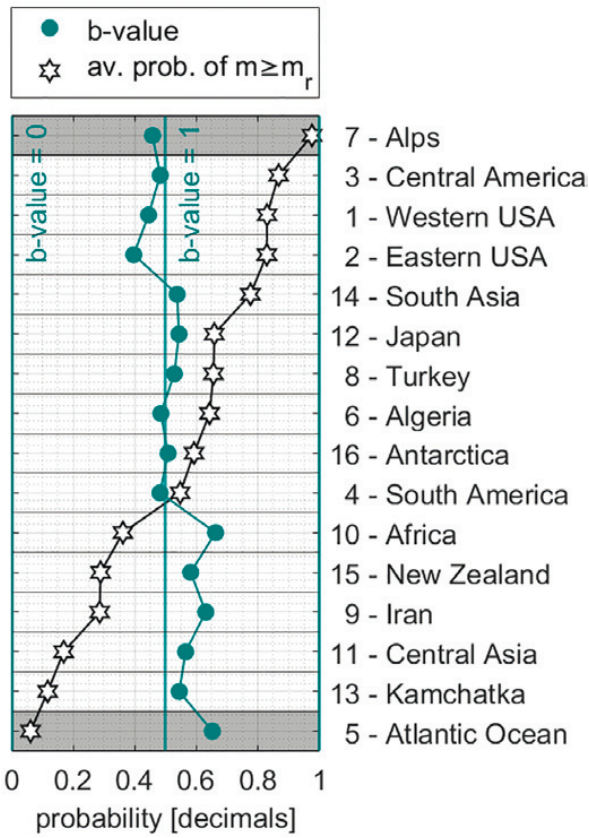

Fig. 6. Regions ordered according to the averages of the probabilities for maximal magnitudes (m) to be equal to or greater than those magnitudes (m_r) that appeared in reality. The horizontal axis is scaled differently for the b-values and the probabilities, respectively. The connecting lines serve only for readability; they do not indicate a relation between regions.

magnitude step of 0.1, but which concern only the lower ranks of the maximal magnitudes (e.g., Central Asia (R11) and Kamchatka (R13)), and, therefore, become gradually less meaningful for Rank-Ordering-Statistics;

(iii) outliers, which persist over almost an entire «percentage-series» regardless of the magnitude step of 0.1 (e.g., the Atlantic Ocean (R5) and the Alps (R7)), and, thus, will be discussed in more detail hereafter.

Surprisingly, probabilities for reproducing or exceeding the five m_r seem to have an opposite trend to the respective b-values per region. Fig. 6 illustrates this phenomenon by ordering regions according to their decrea-

Table 4. Maximal magnitudes (m_{ever}) that occurred since registration up to February 2023 [USGS, 2023]

Region		m_{ever}	Date	Earthquake name	Note
1	Western USA	9.0	27-01-1700	Cascadia	—
2	Eastern USA	7.5	07-02-1812	New Madrid	—
3	Central America	8.5	08-02-1843	Guadeloupe	—
4	South America	9.5	22-05-1960	Valdivia	—
5	Atlantic Ocean	7.9	26-05-1975	North Atlantic	m_{ever} is not m_r (Fig. 5, a, S5)
6	Algeria	7.3	10-10-1980	El Asnam	m_{ever} is m_r (Fig. 5, a)
7	Alps	6.5	06-05-1976	Friuli	—
8	Turkey	7.8	06-02-2023	Turkey-Syria	—
9	Iran	8.1	27-11-1945	Balochistan	>1° S outside of R9 (Fig. 1)
10	Africa	7.2	20-05-1990	South Sudan	m_{ever} is not m_r (Fig. 5, a, S5)
11	Central Asia	8.6	15-08-1950	Assam-Tibet	—
12	Japan	9.1	11-03-2011	Tohoku	—
13	Kamchatka	9.0	04-11-1952	Severo-Kurilsk	—
14	South Asia	9.1	26-12-2004	Sumatra	4° W outside of R14 (Fig. 1)
15	New Zealand	8.2	01-05-1917	Kermadec Islands	1° N outside of R15 (Fig. 1)
16	Antarctica	8.1	27-06-1929	South Sandwich Islands	—

Note: Values are moment magnitudes (m_w). The El Asnam Earthquake in Algeria (R6) is the only one included in the respective dataset with the maximal magnitude (m_r) that appeared in reality being the same as m_{ever}. The data discrepancy between m_{ever} and m_r in the Atlantic Ocean (R5) and Africa (R10) is explained in Section S5 (in the supplement).

sing average percentages (Fig. 5, b) and their increasing b-values.

With regard to the third question presented in Section 3.3. (i.e., how probable it is, that also these m_r could be reproduced or exceeded), we can also expand this question towards a comparison with maximal magnitudes (m_{ever}; Table 4) that occurred since registration up to February 2023 [USGS, 2023] and, hence, to earthquakes that are not included in any of the 16 datasets. The comparison might seem crucial, as time spans ranging from 20 to 56 years (Table 1) might be too short to cover rare extreme events. Fig. 5 shows that in the majority of all 16 regions, the greatest m_r (white triangles in (a)) should be reproduced or exceeded according to high probabilities (colored black points in (b)), what indeed turns out to be true by comparing the respective m_{ever} (asterisks in (a)). For Algeria (R6) and Antarctica (16), the greatest m_r should not be reproduced or exceeded according to rather low probabilities, and the respective m_{ever} and m_r correspond to each other (however not necessarily resulting from the same event). We assume, from these tendencies, that the proposed method is, thus, applicable to the datasets used in this study and — at the same time — not contradictory to MFD covering longer time spans including other extreme events.

4.1. Lack of great magnitudes? The first region, where probabilities for maximal magnitudes differ significantly from the before-described tendencies, is the one of the Alps (R7). From Fig. 5, a, it is evident that, here, solely the first ranked m_r is greater than the magnitude that could be reproduced or exceeded with a probability of 95%, whereas all other m_r ranked 2nd to 5th are even smaller than the respective 95%-predictions. Hence, the five points indicating m_r slide upwards along the probability functions closely towards and above the center span from 5% to 95% (Fig. 7, b). This leads to the conclusion that these m_r should become extremely likely to be reproduced or exceeded. Indeed, Fig. 5, b shows probabilities of up to 100%, and — compared to all other regions — the Alps (R7) are the only region where these strikingly high probabilities for great magnitudes also persist when including a magnitude step of +0.1. Therefore, it seems that — statistically — there is a lack of great magnitudes since the method confidently predicts them, but they do not appear as numerously as in other regions.

On the one hand, one could argue that a period of 20 years is relatively short and that great magnitudes might not have occurred within this period. However, when comparing different time spans, it is evident that earthquakes with great magnitudes in the Alps and their surroundings are indeed sporadic events. The SHARE European Earthquake Catalog [Grünthal & Wahlström, 2012; Grünthal et al., 2013; Stucchi et al., 2013; SHARE,

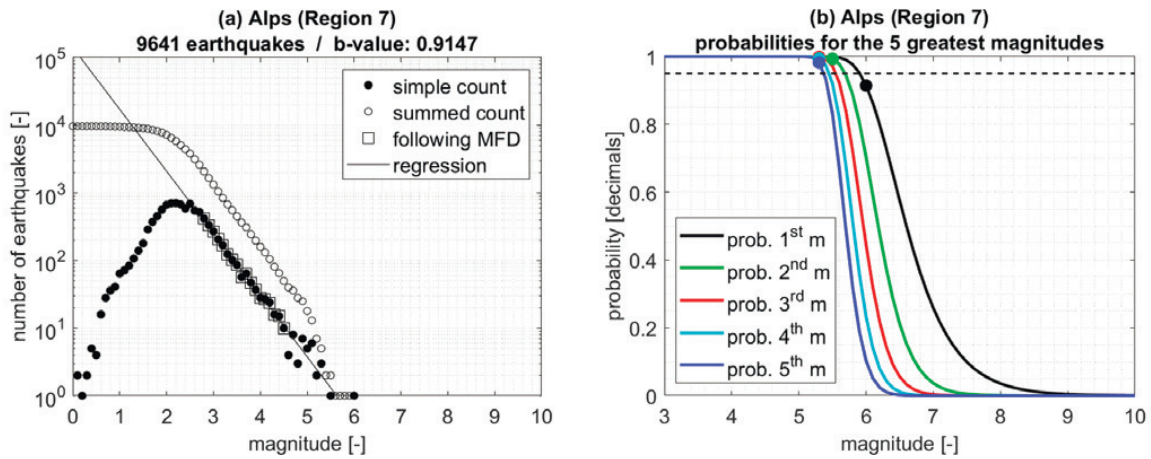

Fig. 7. Result curves of the Alps (R7): approximation of a magnitude-frequency distribution (MFD) via linear regression fitting (a); probability estimation for the maximal magnitudes via Rank-Ordering-Statistics (b) letting assume a lack of great magnitudes. The five colored points mark the maximal magnitudes (m_r) that appeared in reality.

2020] covers the time span from 1000 to 2006, and — after having applied the same 500 km wide buffer zone as for the earthquake data from the ISC Bulletin (Table 2) — the remaining 6,127 moment magnitudes (m_w; 98.55% of the considered set) list 60 moment magnitudes (m_w) ≥6.0 partly in the Alps (R7), and only two ≥7.0 outside: the Avezzano Earthquake (m_w=7.0) on 13th of January 1915 and the Central Italy Earthquake (m_w=7.2) on 5th of December 1456 northeast of Naples. In the here presented earthquake data of the ANSS ComCat [USGS, 2009] for the rectangular area of the Alps (R7) and the timespan from 1990 to 2009, we find only one single earthquake with a moment magnitude (m_w) ≥6.0 (Fig. 7, *a*). The timespan from 2010 to 2020 [EMSC, 2020] would have almost missed out on magnitude records ≥6.0, if not the Petrinja Earthquake (m_w=6.4) had struck Northern Croatia (still within the rectangular area of the Alps; R7) on 29th of December 2020. Extrapolating these datasets to common grounds, all three of them roughly even out to one to two magnitude(s) ≥6.0 and no magnitude ≥7.0 within a period of 20 years.

4.2. Surplus of great magnitudes? The second region, where probabilities for maximal magnitudes differ significantly from the before-described tendencies, is the Atlantic Ocean (R5). From Fig. 5, a, we see that all m_r are greater by up to one magnitude than the magnitude that could be reproduced or exceeded with a probability of 95%. Therefore, the five points indicating m_r slide along the probability functions closely towards and below the center span from 5% to 95% (Fig. 8, b), implying that these m_r should become extremely unlikely to be reproduced or exceeded. Probabilities reach down to 0%, and — including a magnitude step of +0.1 — these distinct low probabilities persist for all maximal magnitudes (Fig. 5, b). From a statistical point of view, there seems to be a surplus of great magnitudes since the method almost certainly rules out the possibility of any further great magnitudes based on their already existing high number.

Likewise, comparisons between different catalogs draw a picture of many great magnitudes in the rectangular area of the Atlantic Ocean (R5). In the here presented earthquake data of the ANSS ComCat [USGS, 2009] for the time span from 1960 to 2009, we find two earthquakes with magnitudes ≥7.0 (i.e., one m_s, one m_{ukn}; Fig. 8, a) and 24 ≥6.0 (i.e., one m_e, four m_b, seven m_s, 12 m_w). The time span from 2010 to 2020 [EMSC, 2020] lists one earthquake with moment magnitudes (m_w) ≥7.0 and two ≥6.0. Extrapolating these datasets again to common grounds, both of them roughly even out to four to ten magnitudes ≥6.0 and one to two magnitude(s) ≥7.0 within a period of 20 years.

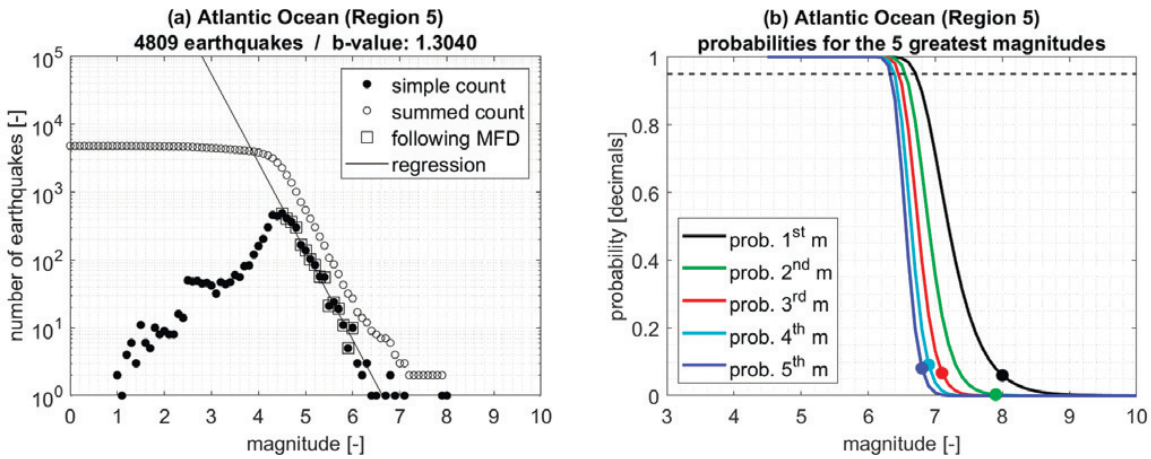

Fig. 8. Result curves of the Atlantic Ocean (R5): approximation of a magnitude-frequency distribution (MFD) via linear regression fitting (a); probability estimation for the maximal magnitudes via Rank-Ordering-Statistics (b) letting assume a surplus of great magnitudes. The five colored points mark the maximal magnitudes (m_r) that appeared in reality.

4.3. Validity of the magnitude-frequency distribution. One question that arises from the «percentage-series» for the maximal magnitudes is whether they bring us closer to the answer of whether the MFD has limited validity in different regions around the globe. Before attempting to answer this question, it is necessary to recall several widely discussed aspects that have partly proven true already.

Generally, the MFD must have a limit of validation for the simple reason that faults — and, hence, the released energy during earthquakes as well as the resulting magnitudes — cannot become unlimitedly great (e.g., [Kagan & Knopoff, 1984]). Although describing a power-law decay never reaching zero, the MFD is supposed to kink at some great magnitude. Various authors commonly assume region-specific kinks, and some also propose other distribution types that might suit earthquake data more appropriately. Sornette et al. [1996], e.g., fitted distribution types to the Harvard Catalog [Dziewonski et al., 1996] and the Southern California Earthquake Catalog [Hutton & Jones, 1993] to find a two-branched power-law distribution for the first but not the second catalog, and that also gamma-distributions would fit well [Kagan, 1994]. Schwartz & Coppersmith [1984], Wesnousky [1994], Wesnousky et al. [1984], and Wu et al. [1995] favor the «characteristic earthquake model», in which some — i.e., the «characteristic» — earthquakes occur more frequently than expected from linear distributions. Also, other bilinear (e.g., [Shah et al., 1975]) and quadratic (e.g., [Merz & Cornell, 1973]) expressions were proposed.

Moreover, kinks in magnitude distributions can be related to the phenomenon of magnitude saturation (e.g., [Shearer, 2009]). Magnitudes derived from amplitudes in seismograms (e.g., m_b, m_c, m_d, m_e, m_l, and m_s) do not relate directly to the physics of earthquakes and have the uncomfortable disadvantage of saturating at greater magnitudes, meaning that their magnitude does not increase any more although the seismic moment (M_0) still does. Drawing, hence, distributions for frequently reported body-wave magnitudes (m_b) [Gutenberg, 1945; Nuttli, 1973], local magnitudes (m_l) [Richter, 1935; Nuttli, 1983], or an assembly of these, might entail a kink with m_b and m_l starting to saturate first at magnitudes about 6.0 followed by m_s at greater magnitudes around 7.5—8.0 [Stein & Wysession, 2003; Shearer, 2009]. In contrast, moment magnitudes (m_w) [Kanamori, 1977; Hanks & Kanamori, 1979] do not saturate as they are directly computed from the seismic moment (M_0) and, thus, do not cause kinks. In the earthquake datasets used in this work, we separately analyzed the magnitude types of those earthquakes having magnitudes greater than m_{10} (Fig. 2, a) in each of the 16 regions. It turned out that for the remaining 20 to 50 earthquakes reported in 14 regions, m_w was the most or the second most frequent magnitude type (Section S8 in the supplement), but not covering exclusively the greatest recorded magnitudes (Section S9 in the supplement). Also, m_b, m_s, and m_e appeared to be the most frequent magnitude types in some regions. Therefore, it is not possible to limit the validity of the MFD in the 16 regions based on argumentations regarding (non-)saturation with these datasets. Besides, magnitudes greater than m_{10} were not considered in any of the 16 regions during the linear regression fitting to the respective datasets.

Nonetheless, one could attempt to — at least — partially answer the question of the validity of MFD, drawing on the results of the probability estimations for the maximal magnitudes in each region. Indirectly, one of the two «extreme» cases — i.e., the Alps (R7) — could let assume a limitation of validity. Here, the probability estimations show that it should become extraordinarily likely for the five m_r (i.e., 6.0, 5.5, 5.3, 5.3, and 5.3; Fig. 5) to be reproduced or exceeded, but records do not show such great magnitudes. Magnitudes seem not becoming greater than about 6.0—6.5; so one could argue that the MFD might be limited in the Alps to roughly that range of maximal magnitudes. In regions where probability estimations draw either moderately likely or extremely unlikely scenarios of reproducing or exceeding the five m_r — such as in the «opposite extreme» case of the Atlantic Ocean (R5) — we cannot presume a

limited validity of the MFD. Here, the five m_r are equal to or greater than the magnitude that could be reproduced or exceeded with a probability of 95%; thus, in those regions, great magnitudes exist — even though with different likelihoods of occurrence — and there are no grounds of assumption that the MFD could be invalid.

5. Conclusion and perspectives. In this work, we present a probabilistic analysis of extreme magnitudes in 16 regions across the globe (Fig. 1) characterized by specific tectonic settings and — hence — seismicity, and by different sizes and numbers of earthquakes (Table 1). Earthquake data was retrieved from the ANSS ComCat [USGS, 2009] for periods of 20 to 56 years; numbers of earthquakes per region range from around 1,000 to 800,000.

The methodological approach combines the Gutenberg-Richter Law [Gutenberg & Richter, 1944, 1956] and Rank-Ordering-Statistics [Zipf, 1949; Sornette et al., 1996; Sornette, 2006]. First, an MFD based on the Gutenberg-Richter Law is approximated via linear regression to each of the 16 datasets between m_0 and m_{10} (Fig. 2, a, 7, a, 8, a, Section S6 in the supplement). For each region, we retrieved the gradient of the regression — i.e., the b-value of the MFD — as well as m_0, as both variables are used in an adapted probability expression (Eq. 15) based on the cumulative earthquake distribution ($P(x)$ in Eq. 2). With this adapted probability expression, it is possible to estimate probabilities for maximal magnitudes and compare — thereto — the maximal magnitudes (m_r) that appeared in reality.

Besides the general information content, from which probabilities ranging from 0% to 100% can be deduced, we approached the series of probability curves (see Fig. 2, b, 7, b, 8, b, Section S7 in the supplement) questioning: (i) which are the maximal magnitudes that could occur or be exceeded with a probability of 95%, (ii) if the respective five m_r are equal to or greater than these 95%-predictions, and (iii) how probable it is, that also these five m_r could be reproduced or exceeded.

From these considerations, we find that in most regions, the five m_r indeed correspond or exceed those magnitudes that are expected from the 95%-predictions (Fig. 5, a). Moreover, we conclude from the respective «percentage-series» that it becomes moderately unlikely that also the five m_r could be reproduced or exceeded (Fig. 5, b). This fits the assumption that with increasing magnitudes, the probabilities of occurrence decrease.

Interestingly, two «extreme» cases break the ranks of this overall behavior.

On the one hand, the «percentage-series» of the Alps (R7; Fig. 7, b) reveal five m_r that lie close or even above the center span from 5% to 95%. Therefore, these m_r should become extremely likely to be reproduced or exceeded, but — although the method confidentially predicts greater magnitudes — they do not appear as numerously as in other regions. Thus, we statistically assume a lack of great magnitudes in the Alps (R7).

On the other hand, the «percentage-series» of the Atlantic Ocean (R5; Fig. 8, b) reveal five m_r that lie close or even below the center span from 5% to 95%. It seems, therefore, extremely unlikely for these five m_r to be reproduced or exceeded. Statistically, we interpret this as a surplus of great magnitudes, as the method almost certainly rules out any greater magnitudes in the Atlantic Ocean (R5).

Here, a perspective for subsequent work might be a detailed geodynamic study of the individual regions with particular focus on their associated tectonic settings in relation to the obtained tendencies of lack and surplus of great magnitudes; from Fig. 6, it appears that the Alps (R7) and the Atlantic Ocean (R5) are the «extreme» cases in terms of probabilities, but other regions might have comparable tendencies — e.g., Central Asia (R11) and Kamchatka (R13), both of which are likewise characterized by convergent tectonics.

Another aspect for future work could be the harmonization of magnitudes per earthquake dataset in each of the regions, which might modify the magnitude distributions and, thus, slightly alter the b-value — the most sensitive parameter within the entire methodological approach. Relationships between magnitude scales are, however, far from being trivial, and relations for transforming one magnitude type into another are usually

empiric and only valid for specific regions (e.g., [Utsu, 2002b; Das et al., 2018; Gasperini et al., 2015]) although some authors (e.g., [Lolli et al., 2014]) have proposed globally applicable laws.

Funding. This study did not receive funding.

Acknowledgments. The presented study is a refined and extended work of the Bachelor Thesis of the author [Domej, 2009], which was completed at the Institute of Earth Sciences of Karl-Franzens University, Graz, Austria, and partly published as conference contributions [Domej, 2014, 2016]. Acknowledgments are expressed for supervision.

Conflicts of interest. The author declares no conflict of interest; the work should be regarded as «food for thought» to be driven further by anyone who may find it useful.

It is noted that the Turkish State changed its name from Turkey to Türkiye in June 2022; as work for this study began before this change, the country still carries its old name.

Data sources and availability. The earthquake data for this study are freely available at the following institutions:

- ANSS ComCat/Credit: United States Geological Survey (https://earthquake.usgs.gov/data/comcat/);
- ISC Bulletin/Credit: International Seismological Centre (https://doi.org/10.31905/D808B830);
- Real Time Seismicity Catalog/Credit: European-Mediterranean Seismological Centre (https://www.emsc-csem.org/);
- SHARE European Earthquake Catalog / Credit: Collaborative project SHARE (https://www.emidius.eu/SHEEC/).

The set of the 50 Flinn-Engdahl Regions [Flinn et al., 1974], used to identify overlaps with the 16 regions of this study, is available as KMZ-file at https://www.researchgate.net/publication/370516966_Flinn-Engdahl_Regions_with_Subregions under the author's profile. Further sources are indicated in the description therein.

References

Al-Rasyid, M.B., Rusman, M.N., Hamonangan, D., Supendi, P., & Kirana, K.H. (2021). The Spatio-temporal Analysis of b-value in the Banda Arc, Indonesia. *IOP Conference Series: Earth and Environmental Science* (Vol. 873, 6 p). https://doi.org/10.1088/1755-1315/873/1/012010.

Amaro-Mellado, J.L., & Tien Bui, D. (2020). GIS-Based Mapping of Seismic Parameters for the Pyrenees. *ISPRS International Journal of Geo-Information,* 9(7), 452. https://doi.org/10.3390/ijgi9070452.

Arroyo-Solórzano, M., & Linkimer, L. (2021). Spatial variability of the b-value and seismic potential in Costa Rica. *Tectonophysics, 814*, 228951. https://doi.org/10.1016/j.tecto.2021.228951.

Bayrak, Y., Öztürk, S., Koravos, G.C., Leventakis, G.A., & Tsapanos, T.M. (2008). Seismicity assessment for the different regions in and around Turkey based on instrumental data: Gumbel first asymptotic distribution and Gutenberg-Richter cumulative frequency law. *Natural Hazards and Earth System Sciences, 8*(1), 109—122. https://doi.org/10.5194/nhess-8-109-2008.

Cheng, Q.M., & Sun, H.Y. (2018). Variation of singularity of earthquake-size distribution with respect to tectonic regime. *Geoscience Frontiers, 9*(2), 453—458. https://doi.org/10.1016/j.gsf.2017.04.006.

Clauser, C. (2014). *Einführung in die Geophysik*, ed. 2. Berlin-Heidelberg: Springer, 407 p. https://doi.org/10.1007/978-3-642-04496-0 (in German).

Das, R., Wason, H.R., Gonzalez, G., Sharma, M.L., Choudhury, D., Lindholm, C., Roy, N., & Salazar, P. (2018). Earthquake Magnitude Conversion Problem. *Bulletin of the Seismological Society of America, 108*(4), 1995—2007. https://doi.org/10.1785/0120170157.

Delvaux, D., Mulumba, J.L., Sebagenzi, M.N.S., Bondo, S.F., Kervyn, F., & Havenith, H.B. (2017). Seismic hazard assessment of the Kivu rift segment based on a new seismotectonic zonation model (western branch, East African Rift system). *Journal of African Earth Sciences, 134*, 831—855. https://doi.org/10.1016/j.jafrearsci.2016.10.004.

Domej, G. (2009). Verifizierung des Gutenberg-Richter-Gesetzes anhand ausgewählter Regionen

unterschiedlicher Seismizität. *Bachelor Thesis*. Karl-Franzens University, Graz, 51 p. https://doi.org/10.13140/RG.2.2.24670.05444 (in German).

Domej, G. (2014). Vérification de la Loi de Gutenberg-Richter dans des régions sélectionnées caractérisées par des activités sismiques différentes. *Poster & abstract, 1er Congrès de la SAG, Société Algérienne de Géophysique, Bab Ezzouar/Algiers*. https://doi.org/10.13140/RG.2.2.12087.14247 (in French).

Domej, G. (2016). Landslide Risk due to Seismic Activity in Norther Algeria: Triggers & Consequences. *Poster & abstract, 1st International Conference of the ArabGU, Arabian Geoscience Union, Bab Ezzouar/Algiers*. https://doi.org/10.13140/RG.2.2.18798.02885.

Dziewonski, A.M., Ekström, G., & Salganik, M.P. (1996). Centroid-moment tensor solutions for January—March 1995. *Physics of the Earth and Planetary Interiors, 93*(3-4), 147—157. https://doi.org/10.1016/0031-9201(95)03111-1.

EMSC (European-Mediterranean Seismological Centre). (2020). Earthquake data of the Real Time Seismicity Catalog. Retrieved from https://www.emsc-csem.org/.

Flinn, E.A., Engdahl, E.R., & Hill, A.R. (1974). Seismic and geographical regionalization. *Bulletin of the Seismological Society of America, 64*(2-3), 771—992. https://doi.org/10.1785/BSSA0643-20771.

Frohlich, C., & Davis, S.D. (1993). Teleseismic b values; Or, much ado about 1.0. *Journal of Geophysical Research: Solid Earth, 98*(B1), 631—644. https://doi.org/10.1029/92JB01891.

Gasperini, P., Lolli, B., & Castellaro, S. (2015). Comparative Analysis of Regression Methods Used for Seismic Magnitude Conversions. *Bulletin of the Seismological Society of America, 105*(3), 1787—1791. https://doi.org/10.1785/0120150018.

Godano, C., Lippiello, E., de Arcangelis, L. (2014). Variability of the b value in the Gutenberg-Richter distribution. *Geophysical Journal International, 199*(3), 1765—1771. https://doi.org/10.1093/gji/ggu359.

Gorbatov, A., Kostoglodov, V., Suárez, G., & Gordeev, E. (1997). Seismicity and structure of the Kamchatka Subduction Zone. *Journal of Geophysical Research, 102*(B8), 17883—17898. https://doi.org/10.1029/96JB03491.

Grünthal, G., & Wahlström, R. (2012). The European-Mediterranean Earthquake Catalogue (EMEC) for the last millennium. *Journal of Seismology, 16*, 535—570. https://doi.org/10.1007/s10950-012-9302-y.

Grünthal, G., Wahlström, R., & Stromeyer, D. (2013). The SHARE European Earthquake Catalogue (SHEEC) for the time period 1900—2006 and its comparison to the European-Mediterranean Earthquake Catalogue (EMEC). *Journal of Seismology, 17*, 1339—1344. https://doi.org/10.1007/s10950-013-9379-y.

Gulia, L., & Wiemer, S. (2010). The influence of tectonic regimes on the earthquake size distribution: A case study for Italy. *Geophysical Research Letters, 37*(10), L10305. https://doi.org/10.1029/2010GL043066.

Gutenberg, B. (1945). Magnitude determination for deep-focus earthquakes. *Bulletin of the Seismological Society of America, 35*(3), 117—130. https://doi.org/10.1785/BSSA0350030117.

Gutenberg, B., & Richter, C.F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America, 34*(4), 185—188. https://doi.org/10.1785/BSSA0340040185.

Gutenberg, B., & Richter, C.F. (1956). Magnitude and energy of earthquakes. *Annali di Geofisica, 9*(1), 1—15. https://doi.org/10.4401/ag-5590 (as a reprint).

Hamdache, M., Peláez, J.A., Kijko, A., & Smit, A. (2017). Energetic and spatial characterization of seismicity in the Algeria—Morocco region. *Natural Hazards, 86*, 273—293. https://doi.org/10.1007/s11069-016-2514-7.

Hanks, T.C., & Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research: Solid Earth, 84*(B5), 2348—2350. https://doi.org/10.1029/JB084iB05p02348.

Harðarson, H.B. (2015). A study of earthquakes along the Mid-Atlantic Ridge between Charlie-Gibbs and the Azores. *Bachelor Thesis*. University of Iceland, Reykjavik, 44 p.

Hergarten, S. (2002). *Self-Organized Criticality in Earth Systems*, ed. 1. Berlin-Heidelberg: Springer, 272 p. https://doi.org/10.1007/978-3-662-04390-5.

Hutton, L.K., & Jones, L.M. (1993). Local magnitudes and apparent variations in seismicity

rates in Southern California. *Bulletin of the Seismological Society of America, 83*(2), 313—329. https://doi.org/10.1785/BSSA0830020313.

ISC (International Seismological Centre). (2020). Earthquake data of the ISC Bulletin. Retrieved from https://doi.org/10.31905/D808B830.

Kagan, Y.Y. (1994). Observational evidence for earthquakes as a nonlinear dynamic process. *Physica D: Nonlinear Phenomena, 77*(1-3), 160—192. https://doi.org/10.1016/0167-2789(94)90132-5.

Kagan, Y.Y., & Knopoff, L. (1984). A stochastic model of earthquake occurrence. *Proc. of the 8th International Conference on Earthquake Engineering, San Francisco* (Vol. 1, pp. 295—302).

Kamer, Y. (2014). Minimum sample size for detection of Gutenberg-Richter's b-value. *Preprint at arXiv*, no. 1410.1815, 11 p. Retrieved from https://arxiv.org/ftp/arxiv/papers/1410/1410.1815.pdf.

Kamer, Y., & Hiemer, S. (2015). Data-driven spatial b value estimation with applications to California seismicity: To b or not to b. *Journal of Geophysical Research: Solid Earth, 120*(7), 5191—5214. https://doi.org/10.1002/2014JB011510.

Kanamori, H. (1977). The energy release in great earthquakes. *Journal of Geophysical Research, 82*(20), 2981—2987. https://doi.org/10.1029/JB082i020p02981.

Leyton, F., Ruiz, S., & Sepúlveda, S.A. (2009). Preliminary re-evaluation of probabilistic seismic hazard assessment in Chile: from Arica to Taitao Peninsula. *Advances in Geosciences, 22*, 147—153. https://doi.org/10.5194/adgeo- 22-147-2009.

Lolli, B., Gasperini, P., & Vannucci, G. (2014). Empirical conversion between teleseismic magnitudes (m_b and M_s) and moment magnitude (M_w) at the Global, Euro-Mediterranean and Italian scale. *Geophysical Journal International, 199*(2), 805—828. https://doi.org/10.1093/gji/ggu264.

Lough, A.C., Wiens, D.A., & Nyblade, A. (2018). Reactivation of ancient Antarctic rift zones by intraplate seismicity. *Nature Geoscience, 11*, 515—519. https://doi.org/10.1038/s41561-018-0140-6.

MATLAB. (2019). Software by MathWorks Inc., version 2019b.

Merz, H.A., & Cornell, C.A. (1973). Seismic risk analysis based on a quadratic magnitude-frequency law. *Bulletin of the Seismological Society of America, 63*(6-1), 1999—2006. https://doi.org/10.1785/BSSA0636-11999.

Mogi, K. (1962). Magnitude-Frequency Relationship for Elastic Shocks Accompanying Fractures of Various Materials and Some Related Problems in Earthquakes. *Bulletin of the Earthquake Research Institute, University of Tokyo, 40*, 831—853.

Mogi, K. (1967). Regional Variation in Magnitude-Frequency Relation of Earthquake. *Bulletin of the Earthquake Research Institute, University of Tokyo, 45*, 313—325.

Mori, J., & Abercrombie, R.E. (1997). Depth dependence of earthquake frequency-magnitude distributions in California: Implications for rupture initiation. *Journal of Geophysical Research: Solid Earth, 102*(B7), 15081—15090. https://doi.org/10.1029/97JB01356.

Nanjo, K.Z., & Yoshida, A. (2021). Changes in the b value in and around the focal areas of the M6.9 and M6.8 earthquakes off the coast of Miyagi prefecture, Japan. *Earth Planets Space, 73*, 176. https://doi.org/10.1186/s40623-021-01511-3.

NASA (National Aeronautics and Space Administration). (2020). Satellite image of the Earth on a Robinson Projection. Retrieved from https://commons.wikimedia.org/wiki/File:Robinson-projection.jpg (on 8th of April 2023).

Nuttli, O.W. (1983). Average seismic source-parameter relations for mid-plate earthquakes. *Bulletin of the Seismological Society of America, 73*(2), 519—535. https://doi.org/10.1785/BSSA0730020519.

Nuttli, O.W. (1973). Seismic wave attenuation and magnitude relations for eastern North America. *Journal of Geophysical Research, 78*(5), 876—885. https://doi.org/10.1029/JB078i005p00876.

Nowroozi, A.A., & Ahmadi, G. (1986). Analysis of earthquake risk in Iran based on seismotectonic provinces. *Tectonophysics, 122*(1-2), 89—114. https://doi.org/10.1016/0040-1951(86)90160-5.

Rhoades, D.A. (1996). Estimation of the Gutenberg-Richter relation allowing for individual earthquake magnitude uncertainties. *Tectonophysics, 258*(1-4), 71—83. https://doi.org/10.1016/0040-1951(95)00182-4.

Richter, C.F. (1935). An instrumental earthquake magnitude scale. *Bulletin of the Seismological Society of America, 25*(1), 1—32. https://doi.org/10.1785/BSSA0250010001.

Sandri, L., & Marzocchi, W. (2007). A technical note on the bias in the estimation of the b-value and its uncertainty through the Least Squares technique. *Annals of Geophysics, 50*(3), 329—339. https://doi.org/10.4401/ag-4422.

Schorlemmer, D., Wiemer, S., & Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes. *Nature, 437*, 539—542. https://doi.org/10.1038/nature04094.

Schwartz, D.P., & Coppersmith, K.J. (1984). Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas Fault zones. *Journal of Geophysical Research: Solid Earth, 89*(B7), 5681—5698. https://doi.org/10.1029/JB089iB07p05681.

Shah, H.C., Mortgat, C.P., Kiremidjan, A., & Zsutty, T.C. (1975). *A study of seismic risk for Nicaragua, Part I and II*. The John A. Blume Earthquake Engineering Center, Department of Civil Engineering, Standford University, Stanford, report no. 11, 369 p. Retrieved from https://searchworks.stanford.edu/view/vn596mw3936.

SHARE (Seismic Hazard Harmonization in Europe). (2020). Earthquake data of the SHARE European Earthquake Catalogue (SHEEC). Retrieved from https://www.emidius.eu/SHEEC/.

Shearer, P.M. (2009). *Introduction to Seismology, ed. 2*. Cambridge University Press, 412 p.

Singh, C. (2016). Spatial variation of seismic b-values across the NW Himalaya. *Geomatics, Natural Hazards and Risk, 7*(2), 522—530. https://doi.org/10.1080/19475705.2014.941951.

Sornette, D. (2006). *Critical Phenomena in Natural Sciences — Chaos, Fractals, Selforganization and Disorder: Concepts and Tools, ed. 2*. Springer, Berlin-Heidelberg, 528 p. https://doi.org/10.1007/3-540-33182-4.

Sornette, D., Knopoff, L., Kagan, Y.Y. & Vanneste, C. (1996). Rank-ordering statistics of extreme events: Application to the distribution of large earthquakes. *Journal of Geophysical Research: Solid Earth, 101*(B6), 13883—13893. https://doi.org/10.1029/96JB00177.

Spada, M., Tormann, T., Wiemer, S., & Enescu, B. (2013). Generic dependence of the frequency-size distribution of earthquakes on depth and its relation to the strength profile of the crust. *Geophysical Research Letters, 40*(4), 709—714. https://doi.org/10.1029/2012GL054198.

Stein, S., & Wysession, M. (2003). *An Introduction to Seismology, Earthquakes, and Earth Structure*. Blackwell Publishing, Oxford, 498 p.

Storchak, D.A., Harris, J., Brown, L., Lieser, K., Shumba, B., & Di Giacomo, D. (2020). Rebuild of the Bulletin of the International Seismological Centre (ISC) — part 2: 1980—2010. *Geoscience Letters, 7*, 18. https://doi.org/10.1186/s40562-020-00164-6.

Storchak, D.A., Harris, J., Brown, L., Lieser, K., Shumba, B., Verney, R., Di Giacomo, D., & Korger, E.I.M. (2017). Rebuild of the Bulletin of the International Seismological Centre (ISC), part 1: 1964—1979. *Geoscience Letters, 4*, 32. https://doi.org/10.1186/s40562-017-0098-z.

Stucchi, M., Rovida, A., Gomez Capera, A.A., Alexandre, P., Camelbeeck, T., Demircioglu, M.B., Gasperini, P., Kouskouna, V., Musson, R.M.W., Radulian, M., Sesetyan, K., Vilanova, S., Baumont, D., Bungum, H., Fäh, D., Lenhardt, W., Makropoulos, K., Martinez Solares, J.M., Scotti, O., Živčić, M., Albini, O., Batllo, J., Papaioannou, C., Tatevossian, R., Locati, M., Meletti, C., Viganò, D., & Giardini, D. (2013). The SHARE European Earthquake Catalogue (SHEEC) 1000—1899. *Journal of Seismology, 17*, 523—544. https://doi.org/10.1007/s10950-012-9335-2.

Turcotte, D.L. (1997). *Fractals and Chaos in Geology and Geophysics, ed. 2*. Cambridge University Press, 416 p. https://doi.org/10.1017/CBO9781139174695.

Ullah, S., Bindi, D., Pilz, M., Danciu, L., Weatherill, G., Zuccolo, E,. Ischuk, A., Mikhailova, N.N., Abdrakhmatov, K., & Parolai, S. (2015). Probabilistic seismic hazard assessment for Central Asia. *Annals of Geophysics, 58*(1), 21 p. https://doi.org/10.4401/ag-6687.

U.S. Department of Energy. (2010). Kemper County IGCC Project, Final Environmental Impact Statement. U.S. Department of Energy, Washington D.C., 671 p. Retrieved from https://www.energy.gov/nepa/articles/eis-0409-final-environmental-impact-statement.

USGS (United States Geological Survey). (2009). Earthquake data of the Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat). Retrieved from https://earthquake.usgs.gov/data/comcat/.

USGS (United States Geological Survey). (2021). 1811-1812 New Madrid, Missouri Earthquakes. Retrieved from https://www.usgs.gov/natural-hazards/earthquake-hazards/science/1811-1812-new-madrid-missouri-earthquakes?qt-science_center_objects=0#qt-science_center_objects (on 3rd of January 2024).

USGS (United States Geological Survey). (2023). Earthquake data of the Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat). Retrieved from https://earthquake.usgs.gov/data/comcat/.

Urban, P., Lasocki, S., Blascheck, P., do Nascimento, A.F., Van Giang, N., & Kwiatek, G. (2016). Violations of Gutenberg—Richter Relation in Anthropogenic Seismicity. *Pure and Applied Geophysics, 173*, 1517—1537. https://doi.org/10.1007/s00024-015-1188-5.

Utsu, T. (2002a). Relationships between Magnitude Scales. *International Geophysics, 81*(A), 733—46. https://doi.org/10.1016/S0074-6142(02)80247-9.

Utsu, T. (2002b). Statistical Features of Seismicity. *International Geophysics, 81*(A), 719—732. https://doi.org/10.1016/S0074-6142(02)80246-7.

Wang, K., Chen, Q.F., Sun, S.H., & Wang, A.D. (2006). Predicting the 1975 Haicheng Earthquake. *Bulletin of the Seismological Society of America, 96*(3), 757—795. https://doi.org/10.1785/0120050191.

Wesnousky, S.G. (1994). The Gutenberg-Richter or characteristic earthquake distribution, which is it? *Bulletin of the Seismological Society of America, 84*(6), 1940—1959. https://doi.org/10.1785/BSSA0840061940.

Wesnousky, S.G., Scholz, C.H., Shimazaki, K., & Matsuda, T. (1984). Integration of geological and seismological data for the analysis of seismic hazard: A case study of Japan. *Bulletin of the Seismological Society of America, 74*(2), 687—708. https://doi.org/10.1785/BSSA0740020687.

Wu, S.C., Cornell, C.A., & Winterstein, S.R. (1995). A hybrid recurrence model and its implication on seismic hazard results. *Bulletin of the Seismological Society of America, 85*(1), 1—16. https://doi.org/10.1785/BSSA0850010001.

Zipf, G.K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley, Cambridge, 573 p.

Ймовірності магнітуд екстремальних землетрусів по всьому світу за допомогою рангового впорядкування

Г. Домей[1,2], 2024

[1]Університет імені Адама Міцкевича, географічно-геологічний факультет, Інститут геоекології та геоінформації, кафедра геоморфології, Познань, Польща
[2]Геологічна служба Словенії, відділ регіональної геології, Любляна, Словенія

З давніх часів імовірність землетрусів хвилює людство, і оцінювання потенційних магнітуд має вирішальне значення для багатьох аспектів безпеки. У статті наведено ймовірнісний аналіз екстремальних магнітуд у 16 регіонах земної кулі, що характеризуються різною сейсмічністю, щоб перевернути традиційне запитання — «яка ймовірність пов'язана з певними величинами». Ми об'єднуємо закон Гутенберга—Ріхтера та рангову статистику в методологічному підході, щоб оцінити, які діапазони величин можна майже напевно (тобто з 95%) очікувати, а які величини стають малоймовірними за межами цих діапазонів.

Цей підхід дає можливість оцінити ймовірності максимальних величин для регіону та порівняти з ними максимальні величини (m_r), які виявилися в реальності.

Метод досліджує максимальні величини, які можуть виникнути або бути перевищені з ймовірністю 95%, якщо відповідні m_r дорівнюють або більше цих 95%-них прогнозів, і наскільки ймовірно, що ці m_r можуть бути відтворені або перевищені. З цих статистичних міркувань ми припускаємо відсутність великих величин в Альпах і надлишок по той бік Атлантичного океану.

Ключові слова: закон Гутенберга—Ріхтера, рангова статистика, амплітудно-частотний розподіл, імовірність величини, b-значення, геостатистика.

Supplement

S1. List of variables

Note that there are three lowercase «*m*». The first is used as such in this study. The italic one is a variable in the Gutenberg-Richter Law. The last refers to the undefined magnitude within the «Alps Test» (Section 3.4.) and appears as such only in the context.

M — maximal magnitudes that could occur or be exceeded with 95%;
m_r — maximal magnitudes that appeared in reality (in study time spans);
m_{ever} — maximal magnitudes that ever occurred (up to February 2023);

m — a distinct value of magnitude in the Gutenberg-Richter Law;
$N(m)$ — number of earthquakes per year with at least this magnitude (m);
a — constant referring to the seismic activity in a region;
b — (negative) gradient of regression in semi-logarithmic plot;
Δb — associated error of the gradient b;
m_0 — magnitude above which a particular earthquake record is complete;
m_{10} — greatest magnitude at which at least ten earthquakes were recorded;
$P(x)$ — cumulative probability distribution;
$p(x)$ — probability density function;

M — all magnitudes (i.e., combined set) (Table 2);
m — undefined magnitude (Table 2);
m_b — body-wave magnitude (Table 2);
m_d — duration magnitude (Table 2);
m_l — local magnitude (Table 2);
m_s — surface-wave magnitude (Table 2);
m_w — moment magnitude (Table 2);
m_{ukn} — unknown magnitude (Table 2);
m_c — coda magnitude (Fig. S1);
m_e — energy magnitude (Fig. S1);
m_h — magnitude derived by hand (Fig. S1);
m_x — velocity/amplitude/mixed magnitude (Fig. S1);

M_0 — seismic moment (i.e., not a magnitude);
R1—R16 — identifier for the 16 regions of this study;
N, S, W, E — north, south, west, east.

S2. Magnitude type fractions per region

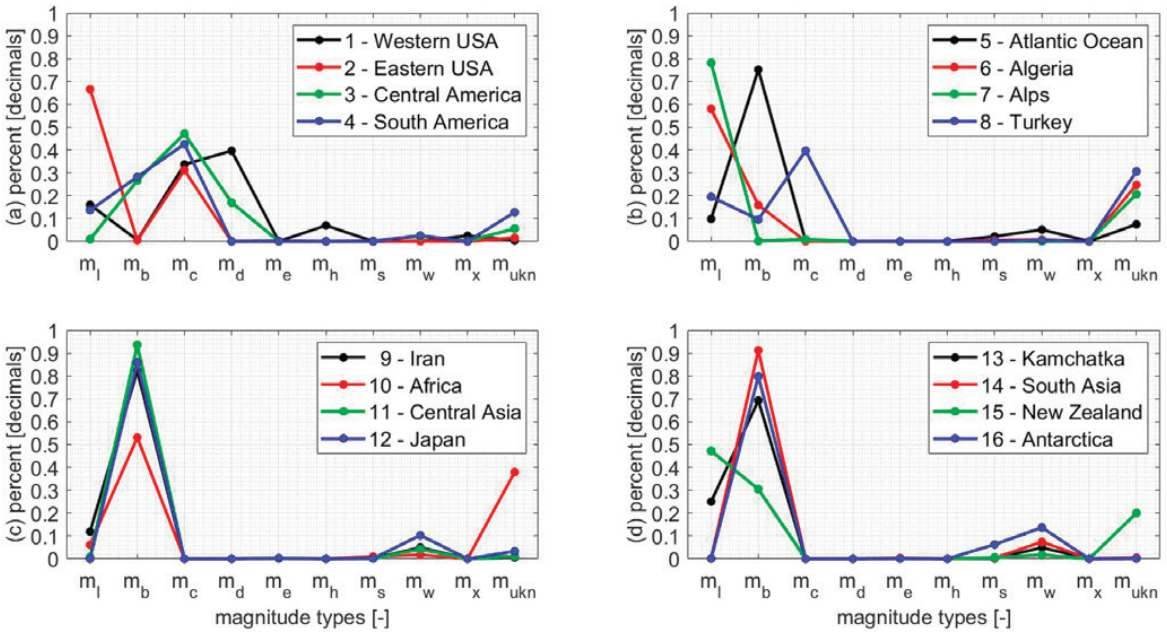

Fig. S1. Magnitude type fractions per region in percent (a—d). The connecting lines serve only for readability; they do not indicate a relation between magnitude types.

S3. Discussion on *b*-values of the «Alps Test»

By close inspection and plain statistical consideration of numbers in Table 2, it becomes apparent that different magnitude types in one dataset will never allow for ideal analysis.

The most striking oddity is that the b-value of the combined set (M; i.e., 0.8930) is smaller than the average b-value of the subsets of undefined magnitudes (m; i.e., 1.6470), body-wave magnitudes (m_b; i.e., 1.5820), duration magnitudes (m_d; i.e., 1.0340), and local magnitudes (m_l; i.e., 1.9540). At first glance, one could expect that the b-value of the combined set (M) should lie somewhere in between those other four b-values and that it also would be similar to the b-value of the local magnitudes (m_l), as their event numbers amount to 48,099 and 39,727, respectively.

This supposed artifact is explained by separate magnitude-frequency distribution (MFD) consideration per magnitude type:

– The 39,727 local magnitudes (m_l) can comfortably hide in the MFD of the combined set (M), report magnitudes only up to 4—5, and, hence, result in a high b-value by causing the regression to «plummet» before magnitude 5.5 (Fig. 3).

– The 5,761 duration magnitudes (m_d) and the 697 undefined magnitudes (m) report magnitudes up to 4—5 and 3—4, respectively; event numbers are, however, considerably small, and therefore, those two magnitude types also hide in the MFD of the combined set (M) by solely increasing the summed counts.

– The 1,403 body-wave magnitudes (m_b) brought a few magnitudes between 5—6, but still not enough to affect the b-value of the MFD of the combined set (M).

– However, the 15 surface-wave magnitudes (m_s) and particularly the 228 moment magnitudes (m_w; Fig. S2) introduced the majority of magnitudes between 5—7. As their event numbers are too small to (properly) fit regressions separately per magnitude type, no b-values could be estimated nor listed in Table 2. Nonetheless, the events exist in the

MFD of the combined set (M) and explain the low b-value of 0.8930.

– The three unknown magnitudes (m_{ukn}) do not play a significant role and hide like the duration magnitudes (m_d) and the undefined magnitudes (m) in the combined set (M) between magnitudes 4—5.

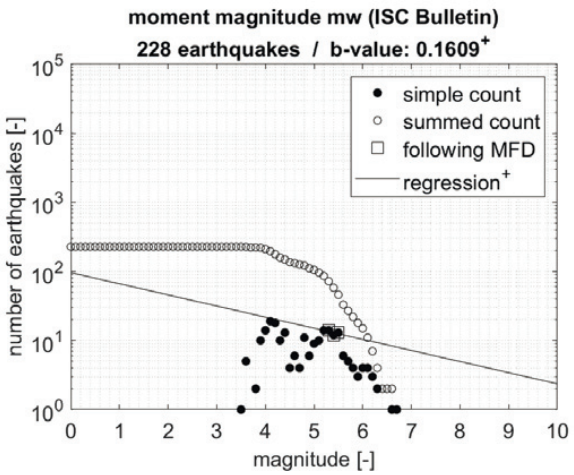

Fig. S2. Magnitude-frequency distribution (MFD) of events reported in moment magnitudes (m_w) in the «Alps Test» (Table 2). The automatically fitted regression (+) between m_0 and m_{10} with a b-value of 0.1609 (+) is not representative.

S4. Overlap between the 16 regions of this study and the Flinn-Engdahl Regions

Table S1. Ranges for b-values resulting from overlapping the 16 regions of this study with the 50 Flinn-Engdahl Regions [Flinn et al., 1974], for which Cheng & Sun [2018] computed b-values individually

Region		Flinn-Engdahl Regions				
		Main overlap	Minor overlap	Resulting b-value ranges		
1	Western USA	2, 3, 34	4*, 39	—	0.25—1.07	1.32*
2	Eastern USA	34	—	—	0.25—0.94	—
3	Central America	4, 5, 6, 7, 32, 34, 44	8, 35nd	—	0.25—1.74	—
4	South America	8, 9, 43	6, 7, 10, 32, 35nd, 39, 44	—	0.95—1.74	—
5	Atlantic Ocean	32, 40, 42	31, 34*, 36, 37	0.25*	1.21—1.51	—
6	Algeria	31	—	—	1.17—1.21	—
7	Alps	31, 36	—	—	1.17—1.32	—
8	Turkey	29, 30, 31	—	—	1.02—1.21	—
9	Iran	29	30*		1.02—1.07	1.16*
10	Africa	33, 37	29*disc	1.02*disc	1.17—1.74	—
11	Central Asia	25, 26, 27, 28, 47, 48	41		0.25—1.51	
12	Japan	17, 18, 19, 20, 22, 39, 41	21	—	0.25—1.51	—
13	Kamchatka	1, 19, 42	41*	0.25*	1.02—1.32	—
14	South Asia	15, 16, 17, 18, 22, 23, 24, 25, 38, 39, 46	20, 33*disc	—	0.25—1.51	1.70*disc
15	New Zealand	11, 12, 38	—	—	0.95—1.21	—
16	Antarctica	10, 33, 39, 43, 45, 50nd	11, 32	—	nd—1.74	—

Note: Range decrease or increase caused by minor overlaps (i.e., when Flinn-Engdahl Regions cover only a very small part of the 16 regions of this study; dotted lines in Fig. 4) are marked with asterisks (*) together with the responsible Flinn-Engdahl Region, and two of them are discarded due to an almost unnoticeable overlap and marked with «disc»; «nd» stands for «no data» in Flinn-Engdahl Regions 35 and 50.

S5. Discrepancy: Why m_{ever} is not m_r in the Atlantic Ocean (R5) and Africa (R10)?

When comparing Fig. 5, *a* and Table 4, it is apparent that for the Atlantic Ocean (R5) and Africa (R10), the maximal magnitudes (m_{ever}) that occurred since registration up to February 2023 are not the same as the maximal magnitudes (m_r) that appeared in reality in the respective periods of 1960—2009 and 1961—2009 (Table 1).

One might assume, that the reason is simply explained by the fact that earthquakes with greater magnitudes have occurred before or after the target time spans (Table 1). This is true for most of the listed m_{ever}, but for the Atlantic Ocean (R5) and Africa (R10), a different data discrepancy arose from improvements with time within the Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat) of the USGS (Table S2).

Table S2. Data discrepancy between the maximal magnitudes (m_{ever}) that occurred since registration up to February 2023 and the maximal magnitudes (m_r) that appeared in reality (Fig. 5, a, Table 4)

Region	Replacement	Magnitude	Date	ANSS ComCat	Location
Atlantic Ocean (R5) 1960—2009	initial entry: entry corrected to:	m_s=8.0 m_w=6.2	28-02-1969	retrieved in 2009 retrieved in 2023	36.181°N, 10.540°W
	replaced by:	m_w=7.9	26-05-1975	retrieved in 2023	35.997°N, 17.649°W
Africa (R10) 1961—2009	initial entry:	m_w=7.0	22-02-2006	retrieved in 2009	21.324°S, 33.583°E
	replaced by:	m_w=7.2	20-05-1990	retrieved in 2023	05.121°N, 32.145°E

In the Atlantic Ocean (R5), the earthquake with a surface-wave magnitude (m_s) of 8.0 was found to be down-corrected in 2023; it was replaced by another earthquake with a moment magnitude (m_w) of 7.9. In Africa, the earthquake with a moment magnitude (m_w) of 7.0 existed as such in 2023; however, it is still replaced by another earthquake with a moment magnitude (m_w) of 7.2, which was not listed as such in 2009 (and probably up-corrected).

Replacements of earthquakes only concern Fig. 5, a and Table 4, and not the statistical analysis of this study, which relies on the datasets per region (1—16) as downloaded from the ANSS ComCat in 2009.

S6. Magnitude-frequency distributions for R1—R4, R6, and R8—R16

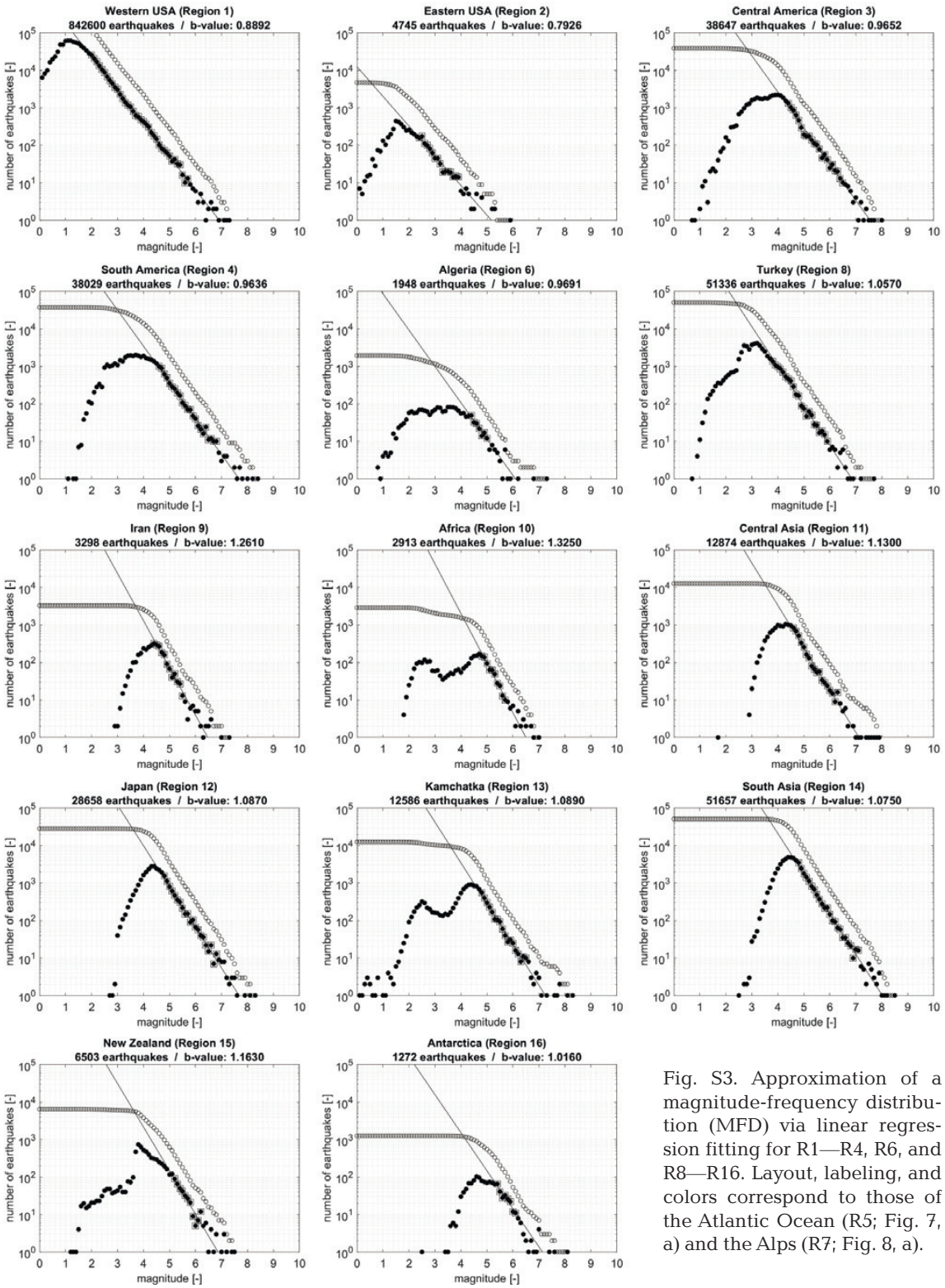

Fig. S3. Approximation of a magnitude-frequency distribution (MFD) via linear regression fitting for R1—R4, R6, and R8—R16. Layout, labeling, and colors correspond to those of the Atlantic Ocean (R5; Fig. 7, a) and the Alps (R7; Fig. 8, a).

S7. Probability estimations for R1—R4, R6, and R8—R16

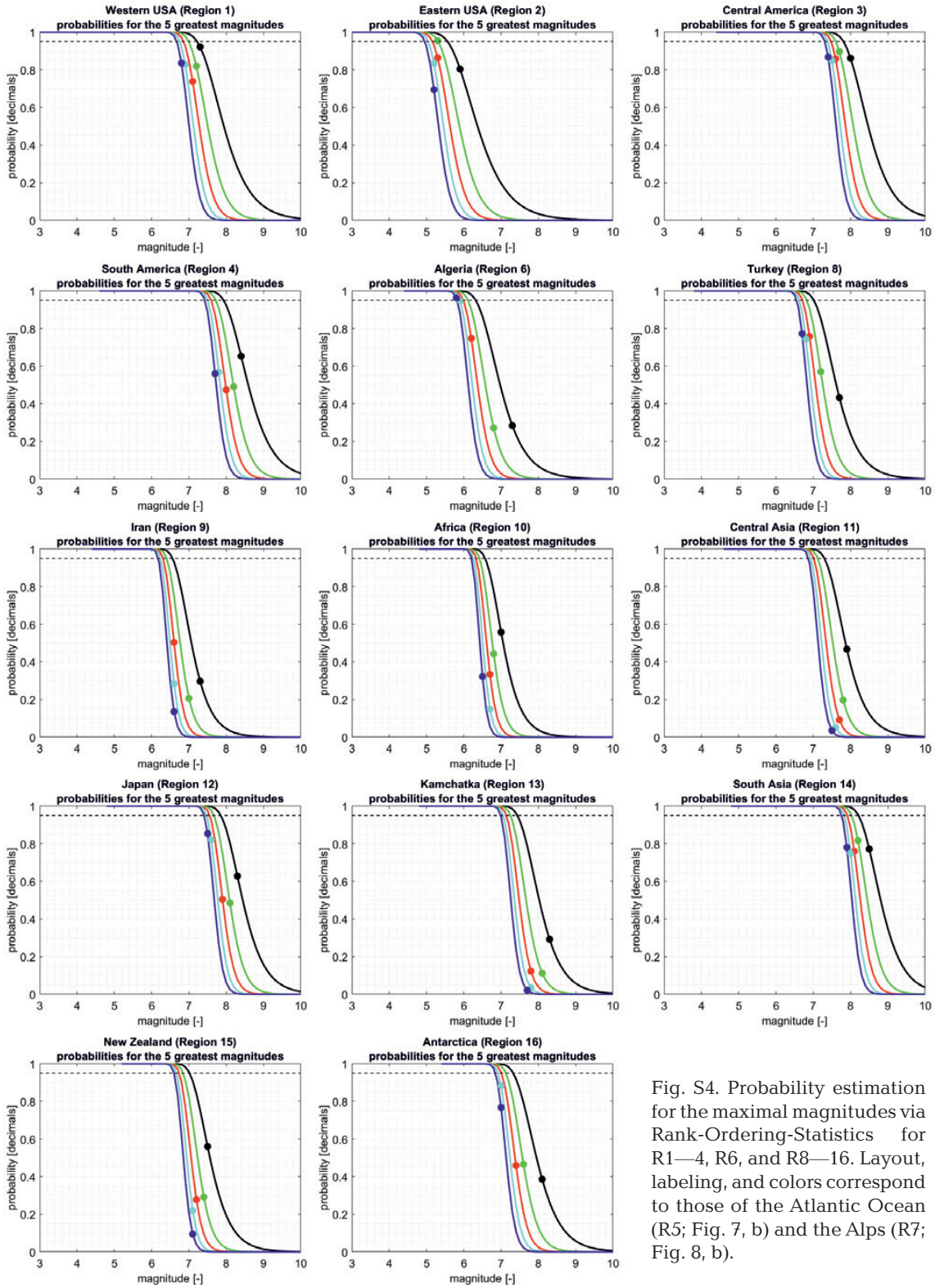

Fig. S4. Probability estimation for the maximal magnitudes via Rank-Ordering-Statistics for R1—4, R6, and R8—16. Layout, labeling, and colors correspond to those of the Atlantic Ocean (R5; Fig. 7, b) and the Alps (R7; Fig. 8, b).

S8. Type fractions of the 50 greatest magnitudes in the 16 regions of this study

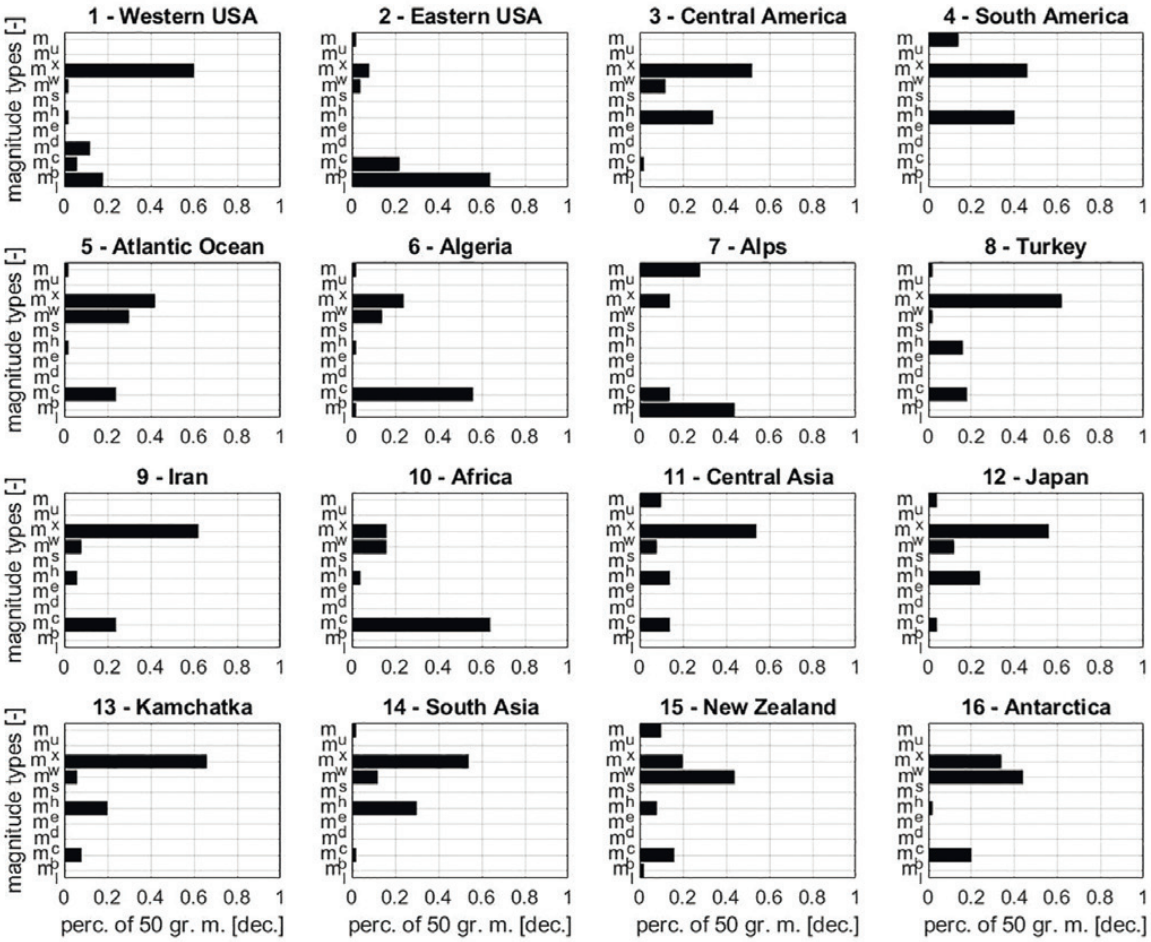

Fig. S5. Type fractions of the 50 greatest magnitudes in percent (m_u=m_{ukn}).

S9. The 50 greatest magnitudes per magnitude type in this study

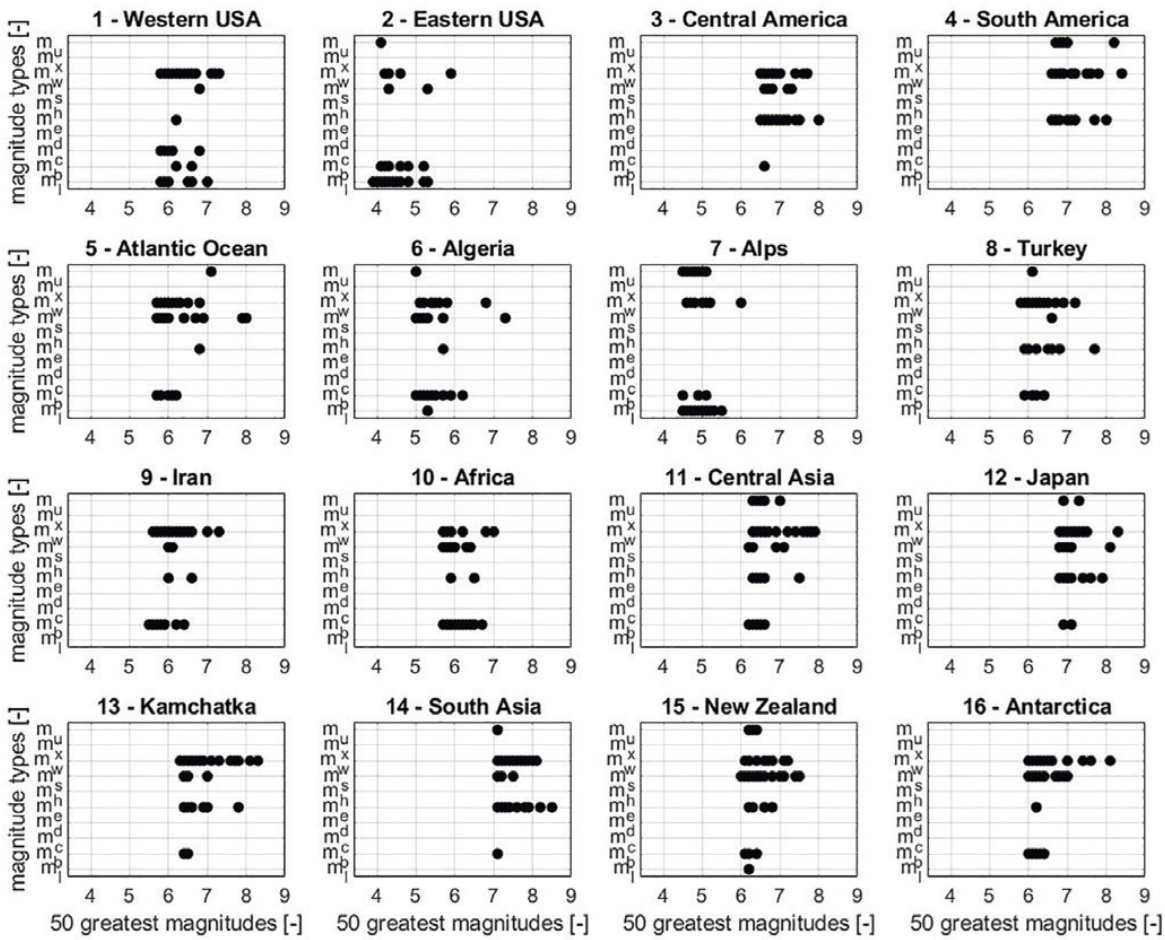

Fig. S6. The 50 greatest magnitudes per magnitude type (m_u=m_{ukn}). Points are overlapping due to earthquakes with the same magnitudes.