Grimm, C., Hainzl, S., Käser, M., Küchenhoff, H. (2024): A New Statistical Perspective on Båth's Law. - Seismological Research Letters, 95, 1, 488-498.

https://doi.org/10.1785/0220230147

A New Statistical Perspective on Båth's Law

Christian Grimm[a], **Sebastian Hainzl**[b,c], **Martin Käser**[d,e], **Helmut Küchenhoff**[a]
[a]Ludwig-Maximilians-University Munich, Department of Statistics, Ludwigstraße 33, 80539 Munich, Germany, christian.grimm@posteo.de, https://orcid.org/0000-0002-2190-2981 (CG); Kuechenhoff@stat.uni-muenchen.de (HK)
[b]GFZ German Research Centre for Geoscience, Physics of Earthquakes and Volcanoes, Helmholtzstraße 6/7, 14467 Potsdam, Germany; hainzl@gfz-potsdam.de, https://orcid.org/0000-0002-2875-0933 (SH)
[c]University of Potsdam, Institute of Geosciences, 14476 Potsdam, Germany
[d]Ludwig-Maximilians-University Munich, Department of Earth and Environmental Sciences, Geophysics, Theresienstraße 41, 80333 Munich, Germany; martin.kaeser@geophysik.uni-muenchen.de (MK)
[e]also at Munich Re, Section GeoRisks, Königinstr. 107, 80802 Munich, Germany

Abstract. The empirical Båth's law states that the average magnitude difference (ΔM) between a mainshock and its strongest aftershock is roughly 1.2, independently of the size of the mainshock. While this observation can generally be explained by a scaling of aftershock productivity with mainshock magnitude in combination with a Gutenberg-Richter frequency-magnitude distribution, estimates of ΔM may be preferable because they are directly related to the most interesting information, namely the magnitudes of the main events, without relying on assumptions. However, a major challenge in calculating this value is the bias introduced by missing data points when the strongest aftershock is below the observed cut-off magnitude. Ignoring missing values leads to a systematic error, because the data points removed are those with particularly large magnitude differences ΔM. The error can be minimized by restricting the statistics to mainshocks that are at least two magnitude units above the cut-off, but then the sample size is strongly reduced. This work provides an innovative approach for modeling ΔM by adapting methods for time-to-event data, which often suffer from incomplete observations (censoring). In doing so, we adequately account for unobserved values and estimate a fully parametric distribution of the magnitude differences ΔM for $M > 6$ mainshocks in a global earthquake catalog. Our results suggest that magnitude differences are best modeled by the Gompertz distribution, and that larger ΔM are expected at increasing depths and higher heat flows.

Abbreviated title: New Statistical Perspective on Båth's Law

Keywords: Båth's law, survival models, parametric distribution.

Main author contact information: Christian.Grimm@posteo.de

1 Introduction

As energy is released in the event of a strong earthquake, tectonic stress redistributes in the surroundings of the initial rupture, usually resulting in further earthquakes, so-called *aftershocks* (Utsu et al., 1995). The cascade of aftershocks is commonly referred to as an *earthquake sequence*, and the strongest event in the sequence is retrospectively defined as the *mainshock* (Taroni, 2023). Typically, events that occurred shortly before the mainshock, so-called *foreshocks*, are included in the sequence since they are believed to be physically related to the upcoming major earthquake (e.g. Helmstetter and Sornette, 2003).

Extensive research has been carried out to analyze and model the spatio-temporal properties of earthquake sequences, e.g. through the Epidemic Type Aftershock Sequence (ETAS) model (Ogata, 1988, 1998; Zhuang et al., 2002). Well-established empirical relationships are the Gutenberg-Richter (GR) law for the frequency-magnitude distribution and the Omori-Utsu law for the temporal aftershock decay (Omori, 1895; Utsu et al., 1995). Furthermore, the aftershock productivity is found to scale exponentially with the mainshock magnitude and the aftershock density decays with distance to the mainshock rupture, where the spatial cluster is typically elongated rather than

isotropic around the mainshock's rupture plane (e.g. Grimm et al., 2022, 2021; Hainzl et al., 2008; Ogata, 2011; Ogata and Zhuang, 2006; Zhang et al., 2018).

Aftershocks are a relevant risk driver since even moderate events can substantially increase damage in buildings and infrastructure destabilized by a prior mainshock. Similarly, foreshocks can set the stage for more severe mainshock damage (Abdelnaby, 2012; Kagermanov and Gee, 2019; Papadopoulos et al., 2020). Therefore, one of the central questions for insurance and risk management purposes is: *What is the expected magnitude difference between the two largest earthquakes in a sequence?*

To date, the literature only provides a starting point for answering this question. The often cited Båth's law states that the average magnitude difference ΔM between a mainshock and its strongest aftershock is roughly 1.2, *independently* of the size of the mainshock (Bath, 1965). Several studies have shown that this empirical observation can be derived from the two observations mentioned above, namely the GR law combined with the aftershock productivity scaling (Console et al., 2003; Felzer et al., 2002; Lombardi, 2002). Thus, ΔM does not appear to be an independent quantity (Taroni, 2023). In particular, to derive regional aftershock forecasts, Page et al. (2016) performed a sophisticated estimation of the regional frequency-magnitude and Omori-Utsu parameters and found generally a good agreement with Båth's ΔM value. However, such an analysis requires a detailed consideration of the short-time aftershock incompleteness (STAI) of earthquake catalogs and involves several fitting parameters for rather small data sets. Thus, the extrapolation to the largest magnitudes contain several assumptions and large uncertainties.

In contrast, ΔM is not affected by STAI and describes the most interesting quantity (the largest magnitudes) without assuming any specific form of the frequency-magnitude distribution and the temporal aftershock decay. In principle, it can be directly estimated from the data. However, a main challenge in calculating this value remains the bias introduced by missing data, if no aftershock was observed above the cut-off magnitude M_c of the catalog and therefore ΔM cannot be computed. We cannot simply ignore missing values, as these are the ones with particularly large magnitude differences ΔM. Therefore, leaving them out leads to a systematic bias. Several authors found that the statistic is robust, if we restrict the sample to mainshocks at least two magnitude units above M_c, but then the sample size is strongly reduced (e.g. Tahir et al., 2012; Zakharova et al., 2013). Another workaround was suggested by Zakharova et al. (2013), who modeled the seismic moment ratio between aftershocks and the mainshock, rather than ΔM, approximating the ratio by zero if no aftershocks were recorded. In any case, Båth's law only makes a statement about the average value of the ΔM, but not about their distribution (and its parameters) or any important quantiles in the lower tail of the distribution.

Another term that appears occasionally in the literature is that of an *earthquake doublet*. Doublets are generally defined as a pair of two similarly strong earthquakes, occurring temporally and spatially close to each other (e.g. Felzer et al., 2004; Grimm et al., 2021; Kagan and Jackson, 1999). A recent example of a destructive doublet is the M7.8 and M7.5 mainshocks that occurred in southeast Turkey and Syria on February 6, 2023. Kagan and Jackson (1999) found that approximately 22% of the $M > 7.5$ earthquakes worldwide occurred accompanied by another $M > 7.5$ event within a distance of one rupture length and with an inter-event time of considerably less than their recurrence time estimated from plate motion. Grimm et al. (2021) showed that roughly 17% of the global $M \geqslant 6$ mainshocks and more than 20% of the mainshocks in Japan were part of an earthquake doublet, defining them as a pair of earthquakes with no more than 0.4 magnitude units difference, occurring within 365 days and a radius of 2.5 rupture lengths.

In this paper, we propose an innovative approach that models the full, parametric distribution of ΔM by adapting so-called *survival models*, originally developed for medical applications. Survival models are a class of regression models that account for data with a censored (or truncated) response variable (see e.g. Klein and Moeschberger (2003) for a comprehensive overview). As the term "survival" suggests, these models were originally developed in applications where the response represents the non-negative lifetime of a patient in medical studies or the lifetime of a device in engineering contexts (so-called *reliability* or *failure time analysis*). The above applications have in common that the exact value of the response is unknown, if the event (e.g. death or device failure) has not occurred until the end of the study period. Replacing lifetimes by magnitude differences, we can therefore use survival models to account for the missing ΔM values where we did not observe the largest aftershock and, therefore, only have the partial information that $\Delta M > M - M_c$, given mainshock magnitude M.
For this analysis, we select earthquake sequences related to $M > 6$ mainshocks in a global catalog using a window method, and compute the (partially right-censored) ΔM between the mainshock and the second strongest event of each cluster. Note that the latter may be a foreshock or an aftershock, as both are relevant in a risk management context. Then, we enrich the cluster set by a plate boundary classification, relative plate velocities, sea floor age and heat flow data, to investigate the regression effects of these large-scale geophysical conditions on the distribution of ΔM.

The focus of this work is on the innovate approach to estimate a fully parametric distribution of ΔM, using survival models that take into account right-censored data rather than avoid it. To our knowledge, no similar approach has been pursued in the literature so far. In the ΔM-regression, covariates represent rather large-scale regional effects. Attempts to consider small-scale variations of these covariates or to include further event specific data are out of the scope of this paper.
Section 2 introduces the utilized datasets, the cluster selection approach and the compilation of the covariate datasets for the regression study. Next, Section 3 rigorously explains the methodological approach of survival models. Then, the results of the regression study are shown and discussed in Section 4. Finally, conclusions are drawn and related future research topics are recommended.

2 Data

This section summarizes the compilation of the regression dataset for the analysis of magnitude differences between the mainshock and the second strongest event in the cluster. First, we justify the choice of the underlying global earthquake catalog. Then, we outline the window method for cluster selection, followed by the definition of the response variable. Finally, the enrichment of further geophysical variables as regression covariates is explained.

2.1 Global Earthquake Catalog

The choice of an appropriate global earthquake catalog for the regression of magnitude differences raises two requirements which, however, are not fully met by any currently available catalog, and therefore necessitate a trade-off. On the one hand, the catalog should be complete down to the smallest possible cut-off magnitude M_c in order to increase the chance that the strongest aftershock of a given mainshock is included in the sample and, therefore, ΔM is known. This is particularly important since we need to account for a sufficient observable magnitude range of roughly one unit below the smallest mainshock magnitude of interest, in order to achieve enough observed ΔM samples to assure statistical robustness in the survival model fit. On the other hand,

the catalog should ideally have homogeneous magnitude scales and be reliably complete in any part of the world, including remote off-shore regions and aftershocks occurring shortly after the mainshock, in order to minimize potential biases.
For magnitude completeness reasons, we chose the U.S. Geological Survey National Earthquake Information Center (USGS-NEIC) catalog, despite not providing homogenized magnitude scales. We extracted all events from 1973 until 2021 with depths smaller than 70 km that occurred at a maximum of 300 km distance to a tectonic plate boundary according to the digital model by Bird (2003). The completeness magnitude of this dataset is $M_c = 5.0$ according to Kagan and Jackson (2010) and Tahir et al. (2012), which allows us to apply the regression model to sequences with a mainshock magnitude larger than 6.0.
To test the influence of inhomogeneous magnitude scales, we performed sensitivity analyses using the International Seismological Centre – Global Earthquake Model (ISC-GEM) instrumental catalog, which is a relocated global event set with homogenized magnitude scales (Bondár et al., 2015; Di Giacomo et al., 2015a,b, 2018; Storchak et al., 2015). Due to its higher level of magnitude completeness, $M_c = 5.6$ according to Di Giacomo et al. (2015b) and $M_c = 6.0$ according to Michael (2014) since 1964, we had to limit our statistical analysis to mainshocks with $M > 6.5$ in order to assure sufficient observed data for ΔM. The corresponding results are presented in the Supplementary Material, confirming the results presented in the following based on the USGS-NEIC catalog.

2.2 Cluster Selection

In order to obtain a set of *independent* clusters, including the information about the magnitude difference ΔM between the mainshock and the largest aftershock (or foreshock), we chose a rather simple window method (see e.g. Gardner and Knopoff, 1974; Uhrhammer, 1986; van Stiphout et al., 2012). To do so, we first sorted the catalog in descending magnitude order. Then, we consecutively searched aftershocks occurring within a time window of $T = 100$ days and a spatial radius of $R(M) = 2.5\,L(M)$, where $L(M) = 10^{-2.44+0.59M}$ is the expected rupture length of the mainshock, depending on its magnitude m, according to the empirical relation of Wells and Coppersmith (1994) for the subsurface rupture length and all mechanisms. Similar to Reasenberg (1985), we linked clusters if an event *B* is found to trigger the potential aftershock *A*, but *A* is the mainshock of an already identified cluster. In this case, due to prior re-ordering of the catalog, $m_A \geqslant m_B$, and event *B* is called a foreshock of *A*.
We conducted sensitivity studies that showed that the regression results are insensitive to varying definitions such as $T = 365$ days and $R(M)$ varying between $1.0\,L(M)$ and $2.5\,L(M)$.

2.3 Response Variable

For each cluster, the magnitude difference ΔM is computed between the mainshock (i.e., the strongest event of the cluster) and the second-strongest event, be it a foreshock or aftershock. In total, we obtain 2,933 clusters with mainshock magnitudes $M > 6.0$.
Note that 1,180 of these are *single-event* clusters, i.e., no associated foreshock or aftershock was found in the corresponding time-space window. Based on seismological reasoning, we can assume that these mainshocks actually triggered aftershocks that were too weak to be recorded in the dataset, given its cut-off magnitude M_c. Therefore, if for a mainshock i with magnitude $M_i \geqslant M_c$ no second event is listed, we have the partial information that the magnitude difference is $\Delta M_i >$

$M_i - M_c$. The single clusters are the reason why we need advanced regression models that can deal with censored data.

2.4 Covariates

We enriched the selected sequences with additional geophysical site information interpolated to the mainshock locations by a nearest-neighbor approach.
Using the digital plate boundary model of Bird (2003), we categorized each event into one of seven *plate boundary classes* continental convergence boundary (CCB), continental transform fault (CTF), continental rift boundary (CRB), oceanic spreading ridge (OSR), oceanic transform fault (OTF), oceanic convergent boundary (OCB), and subduction zone (SUB). Figure 1 shows the mainshock locations, color-coded by the corresponding plate boundary class assigned to them. Table 1 lists the number of clusters with censored ΔM out of the observed values per boundary class, respectively. Nearly half of the clusters are assigned to a subduction zone, and oceanic spreading ridges and transform faults host more censored than non-censored data points.
From the same digital model, we assigned estimates of the *relative plate velocity* and *sea floor age* from the next boundary segment point to the mainshock locations. The value of the *sea floor age* for continental sites is thus related to the age of the nearest oceanic crust. The resulting dependence on *sea floor age*, if continental regions are ignored, is found to be almost indistinguishable and is shown in the Supplementary Material. Likewise, using a nearest-neighbor approach, we interpolated values from the scattered *heat flow* dataset of Bird et al. (2008), provided to us by the author. Figure 2 illustrates the distributions of the interpolated covariate data at the mainshock locations, grouped by the assigned plate boundary class. Subduction zones show the largest relative plate velocities ranging between 0.4 and 262 mm/a (Fig. 2a), while oceanic spreading ridges and transform faults provide the youngest sea floor ages between 0 and 262 Ma (Fig. 2b) and the largest heat flows between 0.025 and 0.3 Wm^{-2} (Fig. 2c).

3 Survival Models

In this section, we introduce survival regression models that we then use in our regression study in order to account for the censored ΔM response data due to unobserved aftershocks. All statistical analyses were performed with the open source software `R` (R Core Team, 2021).

3.1 Why Using a Survival Model for Earthquakes?

The magnitude difference ΔM between the mainshock and the second-largest earthquake of a sequence is only known, if at least one foreshock or aftershock was observed and assigned to the mainshock. Indeed, roughly 40% of the global clusters consist of a stand-alone mainshock. For these clusters, we can conclude that the second strongest event must be smaller than the cut-off magnitude M_c, i.e., that $\Delta M_i > M_i - M_c$, where M_i is the magnitude of mainshock i. In statistics, data points which are capped by such an upper observable threshold are called *right-censored* (Klein and Moeschberger, 2003, section 3.2). Classical statistical models would substantially underestimate ΔM due to the relevant proportion of censored observations.
Replacing lifetimes by magnitude differences, our data meets the necessary requirements of a survival model,

- *non-negative responses* ($\Delta M \geqslant 0$)

- *independent responses* (mainshocks result from declustered catalog)
- *non-informative censoring* (i.e., conditional on covariates, censored clusters are not suspected to deviate structurally in their ΔM-distribution from non-censored clusters).

3.2 Model Formulation and Software

In order to estimate both covariate effects and the entire distribution of magnitude differences ΔM, we need a fully parametric survival model approach. As will be shown in the results section, the best model fits were achieved assuming a *Gompertz* distribution for the magnitude differences, rather than other candidates such as Weibull or Generalized Gamma. The Gompertz distribution is defined on $(0, \infty)$. Therefore data points with $\Delta M = 0$ were substituted by the value 0.01. In the `R` package *flexsurv* (Jackson, 2016), the Gompertz distribution is parameterized by its probability density function

$$f(x|a,b) = be^{ax} \exp\left(-\frac{b}{a}(e^{ax}-1)\right)$$

with shape parameter $a \in \mathbb{R}$ and scale parameter $b > 0$. Besides the categorical plate boundary class, we modeled the effects of the mainshock magnitude $\mathbf{x}_{mag}$ and depth $\mathbf{x}_{depth}$, as well as the locally interpolated relative plate velocity $\mathbf{x}_{veloc}$, heat flow $\mathbf{x}_{heat}$ and sea floor age $\mathbf{x}_{age}$. In the resulting full Gompertz survival model, we regressed the scale parameter b through all covariates for observation i by

$$\begin{aligned} log(b(\mathbf{x_i})) =& \beta_0 + \beta_1 \mathbf{x}_{class=CCB,i} + ... + \beta_6 \mathbf{x}_{class=OTF,i} + \\ & f_{mag}(\mathbf{x}_{mag,i}) + f_{depth}(\mathbf{x}_{depth,i}) + f_{veloc}(\mathbf{x}_{veloc,i}) + f_{heat}(\mathbf{x}_{heat,i}) + f_{age}(\mathbf{x}_{age,i}), \end{aligned}$$

where $\beta_0, \beta_1, ..., \beta_6$ are the coefficients related to categorical variables, where boundary class "SUB" is the reference category, represented by the intercept β_0, and the f terms denote coefficients related to categorical variables. Similarly, we modeled the shape parameter a depending on the linear effects of the plate boundary class, i.e.

$$log(a(\mathbf{x})) = \alpha_0 + \alpha_1 \mathbf{x}_{class=CCB} + ... + \alpha_6 \mathbf{x}_{class=OTF}.$$

In this work, we fitted models using the function *flexsurvreg* from the *flexsurv* package, which estimates parameters by optimizing a parametric likelihood adapted for censored data (Jackson, 2016). To allow for flexible non-linear effects, all metric variables are modeled by the penalized spline function *pspline* from the `R` package *survival* (Therneau, 2016), consistenly using $df = 2$ degrees of freedom and $n = 2.5 \times df$ basis functions (Eilers and Marx, 1996; Hurvich et al., 1998).

4 Results

In this section, we show and discuss the results of a parametric survival model fitted to the global data in order to describe the magnitude difference ΔM between the mainshock and the second strongest event in the cluster. First, we justify and validate the distribution assumption for the response variable. Then we show and interpret the effects of the modeled covariates. Finally, we assess the explanatory power of the model using a response residual plot.

4.1 *Choice of Distribution Family*

To justify the choice of the distribution family, we perform a simple synthetic experiment. We assume a mainshock of magnitude M which triggers aftershocks with magnitudes randomly chosen from the Gutenberg-Richter distribution, i.e., a probability density function (pdf)

$$f(m) = \ln(10)\, b\, 10^{-b(m-M_c)}, \qquad b > 0,\ m \geqslant M_c, \tag{1}$$

with b being the Gutenberg-Richter b-value (Gutenberg and Richter, 1944). The mean number $\langle N \rangle$ of triggered earthquakes is set that, on average, one earthquake exceeds the value of $M - 1.2$, i.e., $\langle N \rangle = 10^{b(M-1.2-M_c)}$. Using $b = 1, M_c = 5.0$, and $M = 8$, we randomly selected (i) the number N of triggered events from a Poisson distribution with mean $\langle N \rangle$, (ii) N magnitude values from Eq. (1), and (iii) determined the magnitude difference ΔM between the two largest magnitudes of the simulated sequence consisting of M and the N randomly chosen magnitude values. We repeated this procedure 1,000,000 times to analyze the ΔM distribution.

We fitted a Weibull, a Gompertz, and a Generalized Gamma distribution to the simulated magnitude differences ΔM. Fig. 3(a) shows the fits of the three distributions to the kernel density estimator of the sampled data. The Gompertz distribution clearly provides the best fit for the moderately negatively-skewed data.

In order to confirm this assumption based on the actual dataset, we fitted a Gompertz survival model with only the scale parameter depending on the categorical plate boundary class and compared the predicted survival curves to those provided by the non-parametric Kaplan-Meier estimator, which does not require a specific distribution assumption (Klein and Moeschberger, 2003, ch. 4). In Fig. 3(b), the step functions, colored according to the seven boundary classes, refer to the Kaplan-Meier estimates. The Gompertz survival model survival curves are plotted on top by black lines, generally showing good agreement.

4.2 *Synthetic Test*

To quantify the benefit of accounting for censored data in the estimation, we performed a simple test with synthetic simulations, consistent with those in the previous section. Specifically, we set the mean number of the aftershocks to $\langle N \rangle = 10^{b(M-\delta M-M_c)}$ for the three δM values (δM = 1.1, 1.3, 1.5) and mainshock magnitudes ranging from 6.0 to 7.5. For each case, we calculated 1000 aftershock sequences, where we randomly selected the number of aftershocks from a Poisson distribution and the magnitudes from a Gutenberg-Richter distribution with $b = 1$ and $M_c = 5$. We then determine the classical estimate (mean ΔM value for those sequences with at least one triggered event) and the estimate from our method, which takes into account the censored data. Both results are compared with the *true* value computed for 100,000 sequences with $M_c = 3.0$, where virtually all sequences have at least one triggered event. Note that the expectation value of the maximum of $\langle N \rangle$ events is not equal to the magnitude ($M - \delta M$), which is exceeded on average once. The results are visualized in Fig. 4 and summarized in the Supplementary Table S1. They show that our method leads to nearly unbiased results, where the true value lies within the error bars, which are related to plus/minus one standard deviation of the estimated Gompertz parameters. In contrast, the classical approach leads to a strong underestimation of ΔM for mainshock magnitudes smaller than 7.0.

4.3 *Covariate Effects*

Figure 5 shows the covariate effects for the full parametric Gompertz survival model in the case of the global earthquake catalog. The categorical effects in Fig. 5(a) represent predictions of the response ΔM given the various boundary classes, if the other covariates are held fixed at their median values (magnitude=6.4, depth=23 km, velocity=66.5 mm/a, sea floor age=220 Ma, heat flow $\approx 0.07\ Wm^{-2}$). The effects of the metric covariates in Fig. 5(b-f) are similarly predicted for a fine grid of values of the considered variable, holding the other covariates fixed and assuming a subducting environment (i.e., boundary class *"SUB"*). Gray shades represent the 95% confidence interval.

4.3.1 *Effect of Boundary Class*

Figure 5(a) reveals no structural effects of specific boundary classes. If we were fitting the same model, but leaving out sea floor age and heat flow, the boundary classes *OSR* and *OTF* would show a substantial and *OCB* a moderate increase in magnitude differences. This is demonstrated in the Supplementary Fig. S3, which shows the ΔM dependence on the boundary class, when the other covariates are ignored. In other words, mainshocks at oceanic, especially transform and divergent type boundaries, produce weaker second strongest events than those in continental zones, which fits with the generally limited magnitude sizes in these two boundary classes (Bird et al., 2002; Boettcher and Jordan, 2004). However, this effect seems to be sufficiently represented by the added metric covariates.

We also repeated the same analysis for simplified plate boundary classes, using only the categories (i) continental (combining CCB, CTF, CRB), (ii) oceanic (combining OSR, OTF, OCB), and (iii) subduction. Again, we observe no significant differences between the plate boundary types when the covariates are taken into account (see Supplementary Fig. S4).

4.3.2 *Effect of Mainshock Magnitude*

For values smaller than $M = 7.8$, the mainshock magnitude effect depicted in Fig. 5(b) confirms the well-established Båth's law hypothesis that the average magnitude difference ΔM is roughly 1.2, *independently* of the mainshock magnitude. For larger magnitudes there seems to be a tendency towards smaller ΔM, but the uncertainties are large. When all other covariates are ignored, ΔM remains constant within the uncertainty range over the whole magnitude range (see Supplementary Fig. S5), as proposed by the Båth law.

A decrease of ΔM for the largest mainshock is very uncertain for two reasons. First, the sample size of $M > 7.8$ events (41 data points) is very small compared to the lower magnitude ranges, leading to large standard errors. Second, the mainshock magnitude controls the radius of the spatial window in the declustering approach. Thus, larger mainshocks span an exponentially increasing area, in which potential aftershocks are searched. However, the larger the area, the more independent background events are included in the analysis, increasing the likelihood of a larger maximum magnitude. To test, whether the observed effect of strong mainshocks may be an artifact of a too generous choice of the spatial window radius, we repeated the study for an event set declustered with radius $R(M) = K(M)\,L(M)$, where the factor $K(M)$ gradually decreases from 2.5 to 1.0 for magnitudes between 6.0 and 9.0. This sensitivity study confirmed the shape of the effect curve, indicating that the second strongest event usually occurred relatively close to the mainshock.

4.3.3 *Effect of Mainshock Depth*

Figure 5(c) shows that the effect of the mainshock depth is almost constant for depths smaller than 40 km. Between 40 km und 50 km, ΔM increases from roughly 1.2 to a new level of approximately 1.5. This effect is consistent with the observation of Hainzl et al. (2019), who showed that aftershock productivity decreases at higher depths due to reduced seismic coupling, i.e. the energy discharges increasingly through seismic creep rather than through aftershocks. Given a constant magnitude size distribution, this would immediately translate into higher average magnitude differences ΔM. Another explanation could be that the observed increase of ΔM with depth is related to a decreasing completeness of our catalog data with depth. However, the analysis of the frequency-magnitude distribution for the deeper events shows no signs of incompleteness (see Supplementary Fig. S6), ruling out this possibility. We find that the events in the depth interval between 40 and 70 km have a larger b value of 1.18 compared to 1.04 of the shallower earthquakes. Thus, the larger b value could be another reason for the observed larger ΔM value at depth.

4.3.4 *Effect of Relative Plate Velocity*

Plate velocities play an important role for the duration of stress re-accumulation at a fault after the occurrence of a large earthquake. However, recurrence intervals of so-called *characteristic earthquakes* are typically in the range of multiple decades or even centuries. For the short-term recurrence of strong aftershocks, Fig. 5(d) reveals no clear effect of the relative plate velocity. As an alternative covariate representing the velocity of deformation in the tectonic system, we tested global strain rate data (Kreemer et al., 2014), which similarly showed no structural effect.

4.3.5 *Effect of Heat Flow*

According to Fig. 5(e), regions with heat flow larger than 0.23 W/m^{-2} show a substantial increase of magnitude differences. Warmer rock is known to be more viscous, which discharges stress through seismic creep rather than abrupt fractures, leading to the same aftershock productivity argument as for higher depths. As Fig. 2 shows, high heat flow values are typically prevalent in oceanic ridges and transform faults, which explains why the model predicts larger ΔM for the plate boundary classes *OSR* and *OTF* if heat flow is left out as a covariate.

4.3.6 *Effect of Sea Floor Age*

Figure 5(f) shows that magnitude differences are substantially larger in young compared to old oceanic crusts. The same result is obtained when the continental regions are ignored (Supplementary Fig. S7). A potential causal reason for the effect of the plate age cannot be ruled out, but is unknown to the authors. Note that young sea floor typically comes with large heat flows. Therefore, the effects of the two variables are consistent. As new oceanic crust is formed at oceanic ridges, the effect also coincides with the increased magnitude differences in the nested model without sea floor age.

If we fit the full model to the subset of subduction zone mainshocks only, both heat flow and sea floor age show no clear signal. Thus, it is likely that their effect is mainly driven by their tails at oceanic ridges.

4.4 Response Residuals

Fig. 6 shows the response residuals (i.e., observed minus predicted values) plotted against the mainshock magnitude of each cluster. Note that, as observations are censored, residuals are censored as well. Therefore, we can show only residuals for *non-censored observations* here. This explains the superiority of negative residuals, especially in the range of lower mainshock magnitudes. If a mainshock with magnitude M slightly above 6 has an observed largest aftershock, this aftershock must be above cut-off magnitude M_c by experimental design and, consequently, the observed, non-censored $\Delta M \leqslant M - M_c$ must be relatively small. The chance that the ΔM observation is smaller than the prediction is accordingly high. In contrast, mainshocks with, in reality, large ΔM and positive residuals are much more likely to be censored and absent in the residual statistic. In the larger mainshock magnitude ranges, this effect gradually disappears, and the scatter of the residuals become symmetrical around 0 for mainshocks greater than magnitude 7.0.

The overall large variation of the residuals suggests a weak predictive power of the model. Residuals of more than one magnitude unit are not rare, and can even reach up to almost two units. Small observations are typically substantially overestimated, and vice versa. The root mean square error for predictions by the full model, 0.62, is only minimally better than by a Gompertz intercept model, 0.63. However, these values only account for predictions of non-censored observations. The majority of substantial covariate effects identified above explain *increases* of the expected magnitude difference, which means that related observations (e.g. events with larger depth or heat flow, or at younger sea floors) are considerably more likely to be censored and therefore left out of the residuals statistics.

The censoring of observations and residuals hinder a rigorous diagnosis of the model. Despite the covariates showing some relevant signals, it is evident that the model misses additional high-resolution geophysical variables for local site effects or event-specific properties that can help explain a larger proportion of the variance in the data.

4.5 Sensitivity Studies

As partly mentioned above, we tested the influence of varying time windows (e.g. $T = 365$ days) and spatial windows (e.g. $R(M) = L(M)$ or $R(M) = K(M)\,L(M)$ with gradually decreasing $K(M)$ as described above) in the cluster selection approach on the regression results. The covariate effects are very insensitive, indicating that the second strongest event typically occurs close to and shortly after (or before) the mainshock. In other words, the contamination of the response variable through background seismicity is negligible.

5 Conclusions

We adapted a survival regression model approach from medical studies to estimate the parametric distribution of the magnitude difference ΔM between the mainshock and its strongest foreshock or aftershock. The highlight of this regression class is that it accounts for right-censored observations. In our case, these are mainshocks for which no aftershock or foreshock is recorded above cut-off magnitude M_c, and for which we therefore have only the partial information that $\Delta M > M - M_c$, where M is the mainshock magnitude.

We selected earthquakes clusters within a global earthquake catalog using a window method and computed ΔM for each of the independent clusters. Then, we enriched the cluster dataset with a

plate boundary classification, relative plate velocities and sea floor ages obtained from the digital plate boundary model of Bird (2003) and with heat flow from Bird et al. (2008). Using a simple simulation experiment where the magnitudes are drawn from the Gutenberg-Richter distribution, we conclude that the Gompertz distribution describes ΔM better than Weibull or Generalized Gamma, and that the approach yields nearly unbiased ΔM estimates even for small magnitude mainshocks.
The regression results show that larger ΔM values are expected at larger heat flows and depths, and in younger ocean crust. However, the residuals of ΔM are still high and an extension of the ΔM-regression model using small-scale covariate data could contribute to a better understanding of magnitude differences in different geophysical settings.

Data and Resources

The U.S. Geological Survey National Earthquake Information Center (USGS-NEIC) catalog has been downloaded from https://earthquake.usgs.gov/earthquakes/search/ (last accessed on March 30, 2022). Global covariate data has been downloaded from http://peterbird.name/publications/2003_pb2002/2003_pb2002.htm (Bird, 2003, last accessed on March 30, 2022) or has been made available by Peter Bird after personal contact (heat flow data, Bird et al., 2008). All statistical analyses were performed with the open source software *R*. Supplementary Material for this article includes the results of a sensitivity study using the homogenized ISC-GEM global earthquake catalog and other supplementary material.

Acknowledgments

We are grateful to Matteo Taroni for his comments and suggestions, which helped us to improve the manuscript. A special thanks goes also to Andreas Bender (LMU Munich) for numerous helpful discussions on the use of survival models in our study and Peter Bird for providing data.

Conflicts of Interest

The authors acknowledge that there are no relevant financial or non-financial interests to disclose.

Financial Declaration

Financial support for this work was provided by Munich Re through a scholarship granted to the first author, and by the Department of Statistics at Ludwig-Maximilians-University Munich. S.H. was supported by the Deutsche Forschungsgemeinschaft (DFG) Collaborative Research Centre 1294 (Data Assimilation – The seamless integration of data and models, project B04) and the European Unions Horizon 2020 research and innovation program under Grant Agreement Number 821115, realtime earthquake risk reduction for a resilient Europe (RISE).

Contribution of the authors

CG (first author) designed the study, compiled and declustered the datasets, conducted and evaluated the ΔM regression model, prepared figures and wrote the paper. SH performed the analysis of the synthetic data and designed the sensitivity analysis. He and MK contributed significantly to the interpretation of the model results in the geophysical context. HK contributed to the project idea and advised on data and methodology.

References

Abdelnaby, A. E. (2012). Multiple earthquake effects on degrading reinforced concrete structures. PhD thesis, University of Illinois at Urbana-Champ.

Bath, M. (1965). Largest inhomogeneities of the upper mantle. *Tectonophysics*, **2**, 483–514.

Bird, P. (2003). An updated digital model of plate boundaries. *Geochemistry, Geophysics, Geosystems*, **4**(3). doi: 10.1029/2001GC000252.

Bird, P., Kagan, Y. Y., Jackson, D. D., Stein, S., and Freymueller, J. (2002). Plate Tectonics and Earthquake Potential of Spreading Ridges and Oceanic Transform Faults. In *AGU Monograph*, pages 203–218. doi: 10.1029/gd030p0203.

Bird, P., Liu, Z., and Rucker, W. K. (2008). Stresses that drive the plates from below: Definitions, computational path, model optimization, and error analysis. *Journal of Geophysical Research: Solid Earth*, **113**, B11406. doi: 10.1029/2007JB005460.

Boettcher, M. S. and Jordan, T. H. (2004). Earthquake scaling relations for mid-ocean ridge transform faults. *Journal of Geophysical Research: Solid Earth*, **109**(12), 1–21. doi: 10.1029/2004JB003110.

Bondár, I., Engdahl, E. R., Villaseñor, A., Harris, J., and Storchak, D. (2015). ISC-GEM: Global Instrumental Earthquake Catalogue (1900-2009), II. Location and seismicity patterns. *Physics of the Earth and Planetary Interiors*, **239**, 2–13. doi: 10.1016/j.pepi.2014.06.002.

Console, R., Lombardi, A. M., Murru, M., and Rhoades, D. (2003). Båth's law and the self-similarity of earthquakes. *Journal of Geophysical Research: Solid Earth*, **108**, 2128.

Di Giacomo, D., Bondár, I., Storchak, D. A., Engdahl, E. R., Bormann, P., and Harris, J. (2015a). ISC-GEM: global instrumental earthquake catalogue (1900-2009), III. Re-computed ms and mb, proxy mw, final magnitude composition and completeness assessment. *Physics of the Earth and Planetary Interiors*, **239**, 33–47. doi: 10.1016/j.pepi.2014.06.005. URL http://dx.doi.org/10.1016/j.pepi.2014.06.005.

Di Giacomo, D., Harris, J., Villaseñor, A., Storchak, D. A., Engdahl, E. R., Lee, W. H., Verney, R., Safronova, N., Wylie, R., Baranauskaite, A., Wilson, J., and Simpson, H. (2015b). ISC-GEM: global instrumental earthquake catalogue (1900-2009), I. Data collection from early instrumental seismological bulletins. *Physics of the Earth and Planetary Interiors*, **239**, 14–24. doi: 10.1016/j.pepi.2014.06.003. URL http://dx.doi.org/10.1016/j.pepi.2014.06.003.

Di Giacomo, D., Robert Engdahl, E., and Storchak, D. A. (2018). The ISC-GEM earthquake catalogue (1904-2014): Status after the extension project. *Earth System Science Data*, **10**, 1877–1899. doi: 10.5194/essd-10-1877-2018.

Eilers, P. H. and Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, **11**, 89–121.

Felzer, K. R., Becker, T. W., Abercrombie, R. E., Ekström, G., and Rice, J. R. (2002). Triggering of the 1999 Mw 7.1 Hector Mine earthquake by aftershocks of the 1992 Mw 7.3 Landers earthquake. *Journal of Geophysical Research: Solid Earth*, **107**, ESE–6.

Felzer, K. R., Abercrombie, R. E., and Ekström, G. (2004). A common origin for aftershocks, foreshocks, and multiplets. *Bulletin of the Seismological Society of America*, **94**(1), 88–98. doi: 10.1785/0120030069.

Gardner, J. K. and Knopoff, L. (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bulletin of the Seismological Society of America*, **64**(5), 1363–1367.

Grimm, C., Hainzl, S., Käser, M., and Küchenhoff, H. (2022). Solving three major biases of the ETAS model to improve forecasts of the 2019 Ridgecrest sequence. *Stochastic Environmental Research and Risk Assessment*, **36**, 2133–2152. doi: 10.1007/s00477-022-02221-2.

Grimm, C., Käser, M., Hainzl, S., Pagani, M., and Küchenhoff, H. (2021). Improving Earthquake Doublet Frequency Predictions by Modified Spatial Trigger Kernels in the Epidemic-Type Aftershock Sequence (ETAS) Model. *Bulletin of the Seismological Society of America*, **112**(1). doi: 10.1785/0120210097.

Gutenberg, B. and Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, **34**, 185–188. doi: 10.1038/156371a0.

Hainzl, S., Christophersen, A., and Enescu, B. (2008). Impact of earthquake rupture extensions on parameter estimations of point-process models. *Bulletin of the Seismological Society of America*, **98**(4), 2066–2072. doi: 10.1785/0120070256.

Hainzl, S., Sippl, C., and Schurr, B. (2019). Linear Relationship Between Aftershock Productivity and Seismic Coupling in the Northern Chile Subduction Zone. *Journal of Geophysical Research: Solid Earth*, **124**(8), 8726–8738. doi: 10.1029/2019JB017764.

Helmstetter, A. and Sornette, D. (2003). Foreshocks explained by cascades of triggered seismicity. *Journal of Geophysical Research: Solid Earth*, **108**(B10). doi: 10.1029/2003jb002409.

Hurvich, C. M., Simonoff, J. S., and Tsal, C.-L. (1998). Smoothing parameter selection in non-parametric regression using an improved Akaike information criterion. *JRSSB*, **60**, 271–293.

Jackson, C. H. (2016). Flexsurv: A platform for parametric survival modeling in R. *Journal of Statistical Software*, **70**(8). doi: 10.18637/jss.v070.i08.

Kagan, Y. Y. and Jackson, D. D. (1999). Worldwide doublets of large shallow earthquakes. *Bulletin of the Seismological Society of America*, **89**(5), 1147–1155.

Kagan, Y. Y. and Jackson, D. D. (2010). Earthquake forecasting in diverse tectonic zones of the Globe. *Pure and Applied Geophysics*, **167**(6), 709–719. doi: 10.1007/s00024-010-0074-4.

Kagermanov, A. and Gee, R. (2019). Cyclic pushover method for seismic assessment under multiple earthquakes. *Earthquake Spectra*, **35**(4), 1541–1558. doi: 10.1193/010518EQS001M.

Klein, J. and Moeschberger, M. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. ISBN 038795399X. doi: 10.1145/390011.808243.

Kreemer, C., Blewitt, G., and Klein, E. C. (2014). A geodetic plate motion and global strain rate model. *Geochemistry, Geophysics, Geosystems*, **15**. doi: 10.1002/2014GC005407.

Lombardi, A. M. (2002). Probabilistic interpretation of Båth's law. *Annals of Geophysics*, **45**, 455–470.

Michael, A. J. (2014). How complete is the ISC-GEM global earthquake catalog? *Bulletin of the Seismological Society of America*, **104**(4), 1829–1837. doi: 10.1785/0120130227.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, **83**(401), 9–27.

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, **50**(2), 379–402.

Ogata, Y. (2011). Significant improvements of the space-time ETAS model for forecasting of accurate baseline seismicity. *Earth, Planets and Space*, **63**(3), 217–229. doi: 10.5047/eps.2010.09.001.

Ogata, Y. and Zhuang, J. (2006). Space-time ETAS models and an improved extension. *Tectonophysics*, **413**(1-2), 13–23. doi: 10.1016/j.tecto.2005.10.016.

Omori, F. (1895). On the aftershocks of earthquakes. *Journal of the College of Science, Imperial University of Tokyo*, **7**, 111–200.

Page, M., van der Elst, N., Hardebeck, J., Felzer, K., and Michael, A. (2016). Three ingredients for improved global aftershock forecasts: Tectonic region, time-dependent catalog incompleteness, and intersequence variability. *Bulletin of the Seismological Society of America*, **106**, 2290–2301.

Papadopoulos, A. N., Bazzurro, P., and Marzocchi, W. (2020). Exploring probabilistic seismic risk assessment accounting for seismicity clustering and damage accumulation: Part I. Hazard analysis. *Earthquake Spectra*. doi: 10.1177/8755293020957338.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reasenberg, P. (1985). Second-order moment of central California seismicity, 1969–1982. *Journal of Geophysical Research: Solid Earth*, **90**, 5479–5495.

Storchak, D. A., Di Giacomo, D., Engdahl, E. R., Harris, J., Bondár, I., Lee, W. H., Bormann, P., and Villaseñor, A. (2015). The ISC-GEM global instrumental earthquake catalogue (1900-2009): Introduction. *Physics of the Earth and Planetary Interiors*, **239**, 48–63. doi: 10.1016/j.pepi.2014.06.009. URL http://dx.doi.org/10.1016/j.pepi.2014.06.009.

Tahir, M., Grasso, J. R., and Amorse, D. (2012). The largest aftershock: How strong, how far away, how delayed? *Geophysical Research Letters*, **39**(4), 1–5. doi: 10.1029/2011GL050604.

Taroni, M. (2023). Against Båth's law: When aftershocks became mainshocks - implications for earthquake forecasting communication. *Seismological Research Letters*, **XX**, XX.

Therneau, T. (2016). survival: Survival Analysis, R package version 2.39-3.

Uhrhammer, R. A. (1986). Characteristics of northern and central California seismicity. *Earthquake Notes*, **57**(1), 21.

Utsu, T., Ogata, Y., and Matsu'ura, R. S. (1995). The centenary of the Omori formula for a decay law of aftershock activity. *J. Phys. Earth*, **43**, 1–33.

van Stiphout, T., Zhuang, J., and Marsan, D. (2012). Seismicity declustering. *Community Online Resource for Statistical Seismicity Analysis*. doi: 10.5078/corssa-52382934. URL http://www.corssa.org/articles/themev/van{_}stiphout{_}et{_}al/vanstiphoutetal2012.pdf.

Wells, D. L. and Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacements. *Bulletin of the Seismological Society of America*, **84**(4), 974–1002.

Zakharova, O., Hainzl, S., and Bach, C. (2013). Seismic moment ratio of aftershocks with respect to main shocks. *Journal of Geophysical Research: Solid Earth*, **118**(11), 5856–5864. doi: 10.1002/2013JB010191.

Zhang, L., Werner, M. J., and Goda, K. (2018). Spatiotemporal seismic hazard and risk assessment of aftershocks of M 9 megathrust earthquakes. *Bulletin of the Seismological Society of America*, **108**(6), 3313–3335. doi: 10.1785/0120180126.

Zhuang, J., Ogata, Y., and Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association*, **97**(458), 369–380. doi: 10.1198/016214502760046925.

Contact to authors

Christian Grimm (main author), *Ludwig-Maximilians-University Munich, Department of Statistics, Ludwigstraße 33, 80539 Munich, Germany; Christian.Grimm@posteo.de*,
Sebastian Hainzl, *GFZ German Research Centre for Geoscience, Physics of Earthquakes and Volcanoes, Helmholtzstraße 6/7, 14467 Potsdam, Germany*,
Martin Käser, *Ludwig-Maximilians-University Munich, Department of Earth and Environmental Sciences, Geophysics, Theresienstraße 41, 80333 Munich, Germany*, also at *Munich Re, Section GeoRisks, Königinstr. 107, 80802 Munich, Germany*,
Helmut Küchenhoff, *Ludwig-Maximilians-University Munich, Department of Statistics, Ludwigstraße 33, 80539 Munich, Germany*

List of Tables

Table 1: Number of all, censored, and observed ΔM data points, grouped by plate boundary classes according to the digital plate model of Bird (2003).

Plate Boundary	**Number of Clusters**		
Class	**#total**	**#censored**	**#observed**
CCB (continental convergence boundary)	219	71	148
CTF (continental transform fault)	238	81	157
CRB (continental rift boundary)	183	69	114
OSR (oceanic spreading ridge)	181	108	73
OTF (oceanic transform fault)	513	319	194
OCB (oceanic convergent boundary)	142	59	83
SUB (subduction zone)	1457	473	984

List of Figures

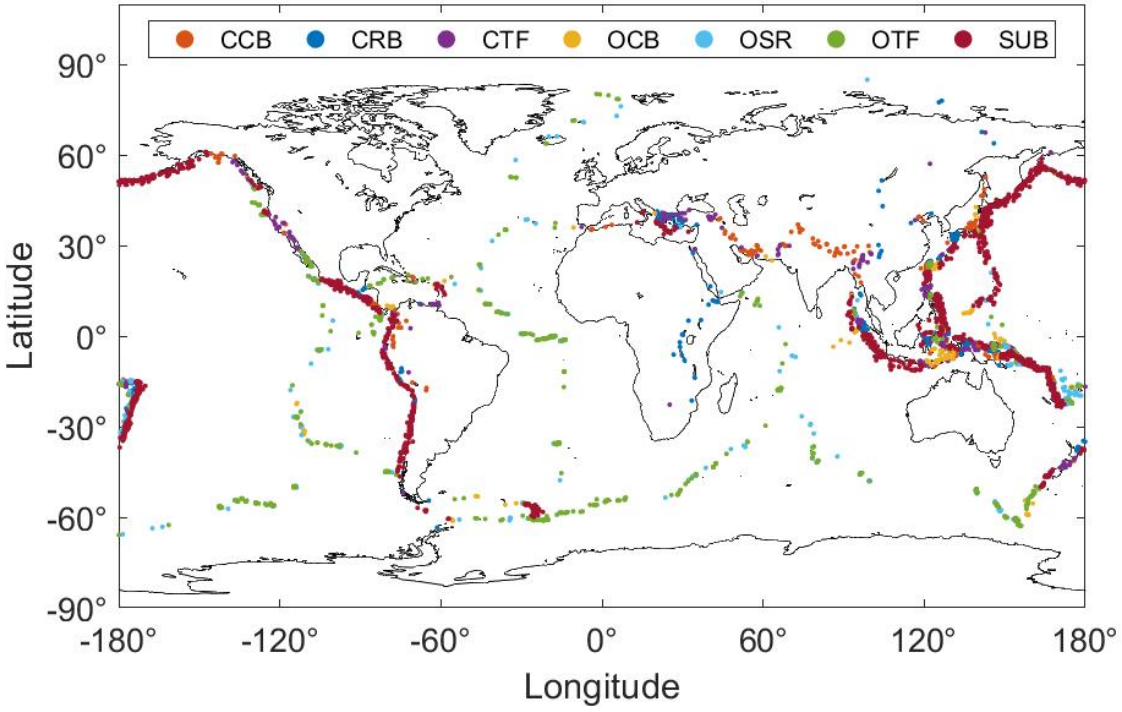

Fig 1: Locations of 2,933 global $M > 6$ mainshocks between 1973 and 2020, selected from the USGS-NEIC catalog. Mainshocks are color-coded according to their assignment to the plate boundary classes, listed in Tab. 1, which are introduced in the digital plate model of Bird (2003).

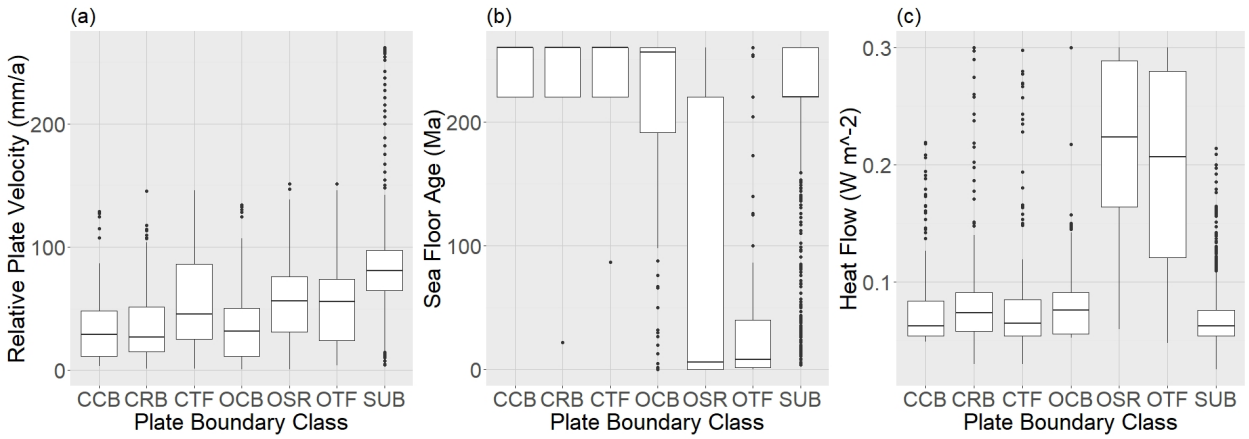

Fig 2: Boxplots of (a) relative plate velocity, (b) sea floor age, and (c) heat flow values assigned to cluster mainshocks by a nearest-neighbor approach from original scatter data (Bird, 2003; Bird et al., 2008), grouped by the plate boundary class which are listed in Tab. 1.

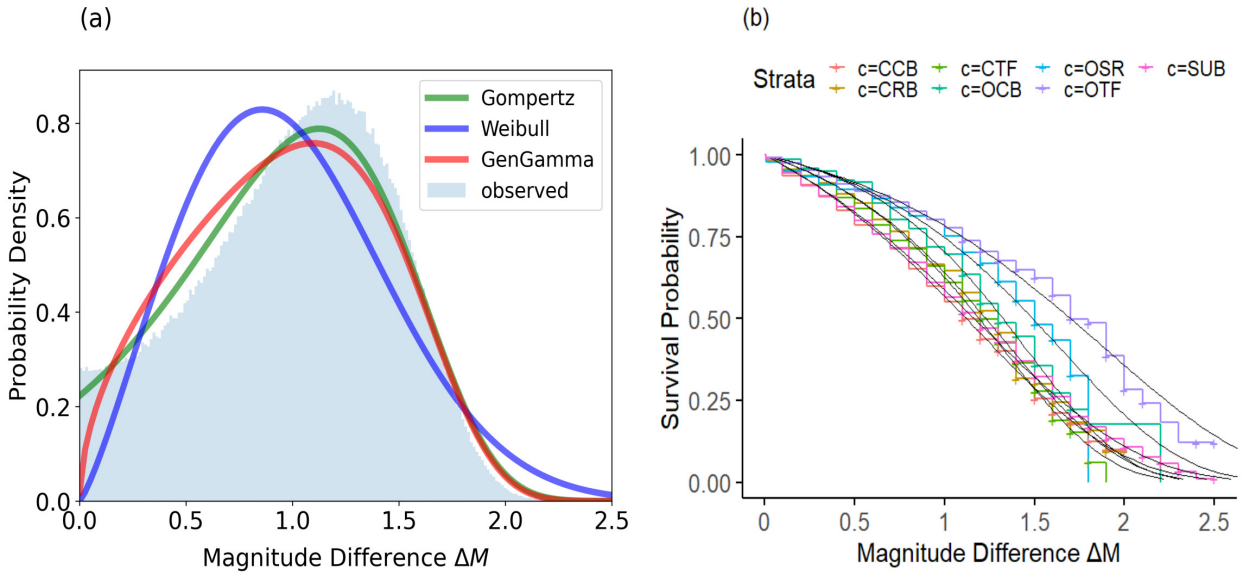

Fig 3: (a) Fits of the Gompertz, Weibull and Generalized Gamma distribution to simulated magnitude differences ΔM, represented by the kernel density estimator (black curve). (b) Comparison of survival curves estimated from a Gompertz model (continuous black lines) and a non-parametric Kaplan-Meier estimator (colored step functions), stratified for plate boundary classes (c, see legend).

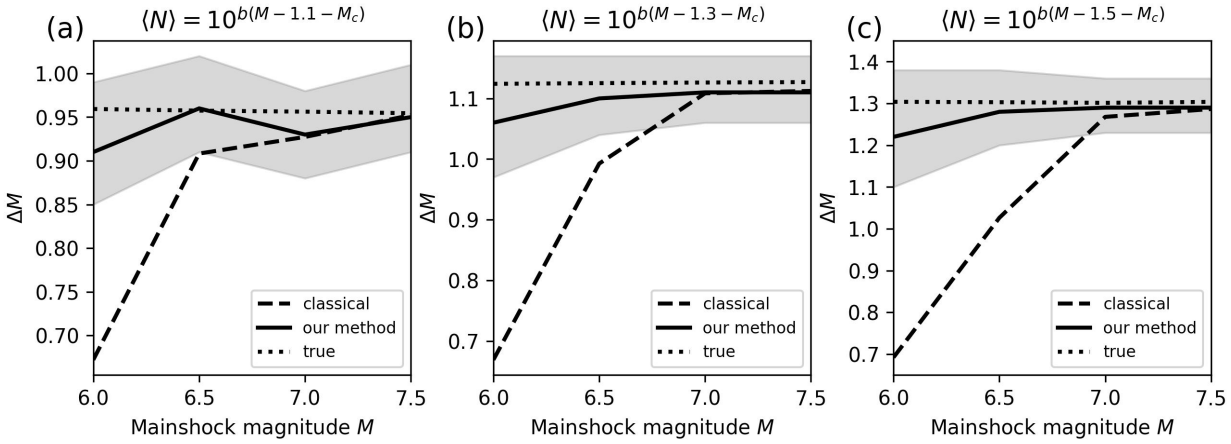

Fig 4: ΔM estimates for synthetic aftershock sequences as a function of the mainshock magnitude M and the parameter δM, determining the mean aftershock number $\langle N \rangle$ (see title lines) with $b = 1, M_c = 5$. Each ΔM value was estimated by 1000 synthetic sequences, where ΔM is estimated in the case of the classical method by the average of the observed value for those sequences with at least one aftershock, while our method takes into account the censored sequences where no aftershock occurred. In the latter case, the error bars refer to plus/minus one standard deviation of the estimated Gompertz parameters. The true value was estimated from the mean of 10^5 simulations with $M_c = 3$.

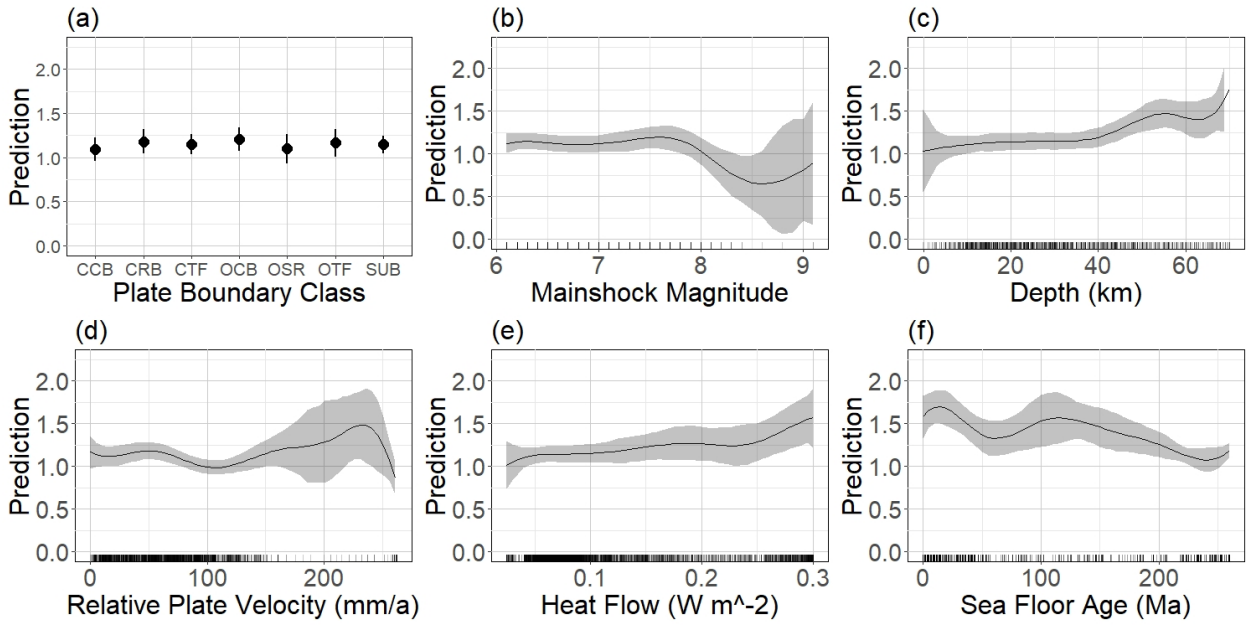

Fig 5: Covariate effects of the ΔM-Regression, by (a) plate boundary type (categorical), (b) mainshock magnitude, (c) mainshock depth, (d) relative plate velocity, (e) heat flow and (f) sea floor age on the magnitude difference between a mainshock and the second largest event of the cluster. For linear effects (a), 95% confidence intervals are represented by bars. For smooth effects (b-f), 95% confidence intervals are depicted by gray shades. The effects are computed as predictions of the response variable, fixing the other variables at their median values. Rug lines on the x axis visualize the marginal distributions of the corresponding metric covariate.

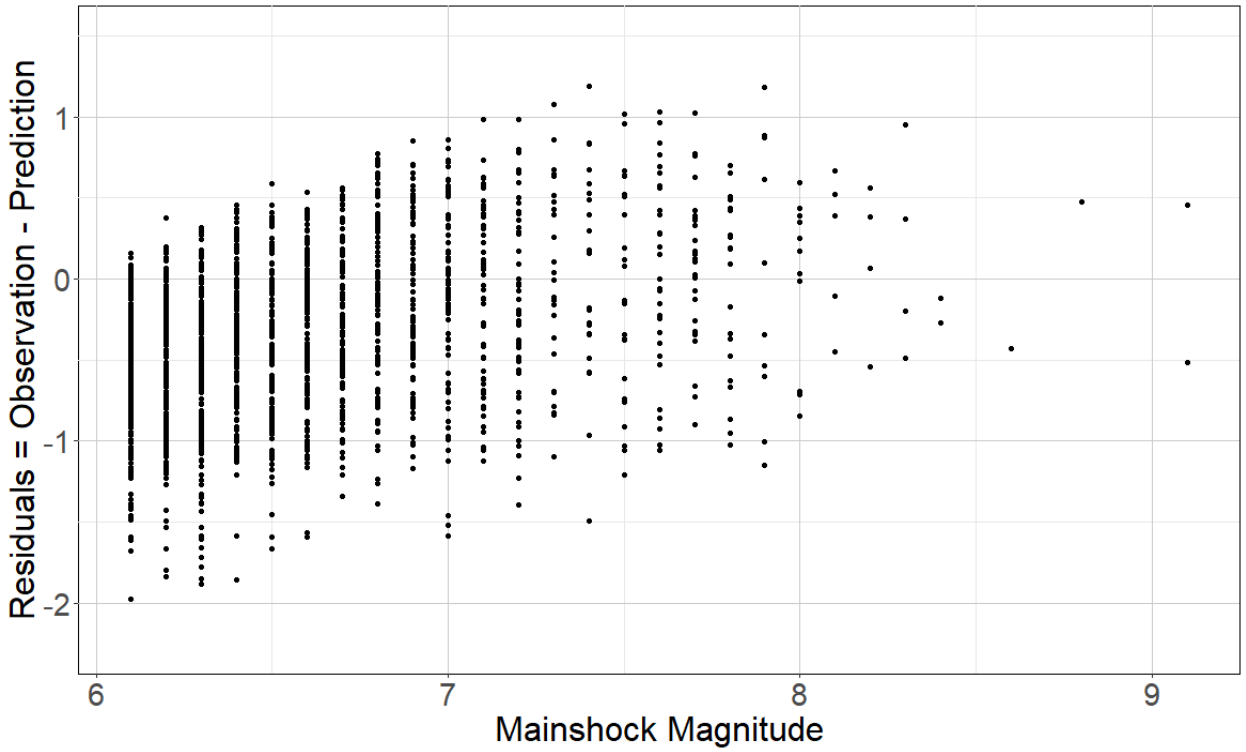

Fig 6: Response Residuals of the ΔM-regression for non-censored observations only, plotted against the mainshock magnitudes. The row arrangement of the points is due to the rounding of the observed data to one decimal place.