New Statistical Perspectives on Bath's Law and Aftershock Productivity

Christian Grimm (✉ Christian.Grimm@stat.uni-muenchen.de)

Ludwig Maximilians University Munich: Ludwig-Maximilians-Universitat Munchen

https://orcid.org/0000-0002-2190-2981

Teresa Rupprecht

Ludwig Maximilians University Munich: Ludwig-Maximilians-Universitat Munchen

Kendra Johnson

Global Earthquake Model

Sebastian Hainzl

GFZ: Deutsches Geoforschungszentrum Potsdam

Helmut Küchenhoff

Ludwig Maximilians University Munich: Ludwig-Maximilians-Universitat Munchen

Martin Käser

Ludwig Maximilians University Munich: Ludwig-Maximilians-Universitat Munchen

Marco Pagani

Global Earthquake Model

New Statistical Perspectives on Bath's Law and Aftershock Productivity

Christian Grimm[a], Teresa Rupprecht[a], Kendra Johnson[b], Sebastian Hainzl[c], Helmut Küchenhoff[a], Martin Käser[d,e], Marco Pagani[b]

[a]Ludwig-Maximilians-University Munich, Department of Statistics, Ludwigstraße 33, 80539 Munich, Germany,

christian.grimm@stat.uni-muenchen.de, https://orcid.org/0000-0002-2190-2981 (CG);

Teresa.Rupprecht@campus.lmu.de (TR); Kuechenhoff@stat.uni-muenchen.de (HK)

[b]Global Earthquake Model Foundation, Via Ferrata 1, 27100 Pavia, Italy, kendra.johnson@globalquakemodel.org

(KJ); marco.pagani@globalquakemodel.org (MP)

[c]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)

[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 well-established Bath's law states that the average magnitude difference between a mainshock and its strongest aftershock is roughly 1.2, independently of the size of the mainshock. The main 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 is minimized, if we restrict the statistics to mainshocks at least two magnitude units above the cut-off, but then the sample size is strongly reduced. This work provides an innovative approach for modelling ΔM by adapting methods for time-to-event data, which often suffers from incomplete observation (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. Results show that magnitude differences are best modeled by the Gompertz distribution, and that larger ΔM are expected at increasing depths and higher heat flows. A simulation experiment suggests that ΔM is mainly driven by the number and the magnitude distribution of aftershocks. Therefore, in a second study, we modelled the variation of aftershock productivity in a stochastically declustered local catalog for New Zealand, using a generalized additive model approach. Results confirm that aftershock counts can be better modelled by a Negative Binomial than a Poisson distribution. Interestingly, there is indication that triggered earthquakes trigger themselves two to three times more aftershocks than comparable

background events. This effect can either be an indicator for incorrect trigger pair assignments as a result of the declustering approach, or it may represent an actual effect due to a higher prevalent energy level in the tectonic system during on-going earthquake sequences. If confirmed, this effect must be carefully considered in forward simulations of earthquake sequences, as otherwise there is a risk of substantially underestimating cluster sizes and consequently overestimating ΔM.

Keywords: Bath's law, survival models, aftershock productivity, generalized additive models.

Main author contact information: Christian.Grimm@stat.uni-muenchen.de

1 Introduction

As energy is released in the event of a strong earthquake, tectonic stress redistributes in the surroundings of the initial rupture and typically results 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 of the sequence is called the *mainshock*. 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). Studies found a well-established power-law decay of aftershock rates as a function of the time after the mainshock (Omori, 1895), while 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 emergency and risk management purposes is: *What is the second strongest magnitude to be expected in a sequence?*

To date, the literature only provides a starting point for answers to this question. The well-established Bath'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). The main challenge in calculating this value is 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 would lead to a systematic bias. Several authors found that the statistics 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 work around 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, Bath'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 temporarly and spatially close to each other (e.g. Felzer et al., 2004; Grimm et al., 2021; Kagan and Jackson, 1999). 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.

3

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.

1.1 Survival Model Regression of ΔM

In the first part of this work, 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 comprehensive overviews). 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 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 only have the partial information that $\Delta M > M - M_c$, given mainshock magnitude M.

To do so, we decluster 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

4

98 conditions on the distribution of ΔM.

1.2 Drivers of ΔM

100 The magnitude difference ΔM is controlled by the two drivers (Grimm et al., 2021)

101   1. number of aftershocks (hereafter called *aftershock productivity*) and

102   2. frequency-magnitude distribution (FMD) of triggered events.

103 A simple simulation shall demonstrate the effect of both factors. Assume an initial mainshock of
104 magnitude $M = 8$. Let the expected aftershock productivity of an earthquake with magnitude M
105 be

$$k(M) = A \, \exp(\alpha \, (M - M_c)), \qquad M \geqslant M_c, \tag{1}$$

106 and the FMD be the exponential distribution with probability density function (pdf)

$$f(M) = \beta \, e^{-\beta(M - M_c)}, \qquad \beta > 0, \ M \geqslant M_c, \tag{2}$$

107 where β is related to the Gutenberg-Richter b-value by $b = \beta/ln(10)$ (Gutenberg and Richter,
108 1944). If we assume the realistic parameters $A = 0.13$, $\alpha = 2.0$ and $b = 1.0$ and, for simplicity, a
109 Poisson distributed number of aftershocks with trigger-magnitude dependent parameter $\lambda = k(M)$,
110 we can simulate a distribution of ΔM with a mean of 1.2, consistent with Bath's Law.

111 A doubled aftershock productivity (i.e. $A = 0.26$) would lead to a pronounced drop of the average
112 magnitude difference down to 0.66. If the increase of A does only apply to secondary triggering,
113 but the number of direct aftershocks of the $M8$ mainshock remains unchanged, the mean of ΔM

114 still decreases to below 0.9. A similar reduction of the average ΔM is achieved, if instead we

115 modify the FMD, sampling magnitudes according to $b = 0.85$.

1.3 Regression of Aftershock Productivity

117 The simplified simulation experiment above illustrates the leverage of both aftershock productivity

118 and FMD on the magnitude differences ΔM. This gave motivation for a more in-depth analysis

119 of the variation in aftershock productivity in the second part of this study. To do so, we declus-

120 ter a local earthquake catalog for New Zealand and fit a generalized additive model (GAM) to

121 the estimated number of direct aftershocks per event, comparing a Poisson with a Negative Bi-

122 nomial distribution. Here, we use the stochastic declustering method introduced by Zhuang et al.

123 (2002), based on the ETAS-Incomplete model proposed by Grimm et al. (2022) that accounts for

124 anisotropic ruptures as well as short-term aftershock incompleteness to reduce estimation biases.

125 We enrich the local catalog by classifying the events to the main tectonic contexts in the surround-

126 ing of the Hikurangi fault, and by concluding the slip type from additional focal mechanism data.

127 Extensive research has been done on the variation of aftershock productivity. Kagan (2017) and

128 Shebalin et al. (2018) showed that aftershock counts are best modeled by the Negative Binomial

129 distribution due to their large variance. Page et al. (2016) found that aftershock productivity may

130 regionally vary by a factor of almost 10, which would explain the variation in ΔM to large extent.

131 Marsan and Helmstetter (2017) found that 40-80% of the aftershock variability may be related

132 to variation in the mainshock stress drop. Dascher-Cousineau et al. (2020) investigated a large

133 number of source and site effects on aftershock productivity and showed individual correlations of

134 stress drop and rupture dimension with the number of aftershocks. Wetzler et al. (2016) suggest

135 a larger productivity in subduction zones of the western circum-Pazific, compared to the eastern

6

side. Zhuang et al. (2004) proposed that triggered events produce more aftershocks than comparable background events, which would provide an additional boost in clustering and decrease expected ΔM's.

Potential correlations of the FMD of triggered events with the magnitude of the direct ancestor or the cluster mainshock were found by Zhuang et al. (2004), Gulia et al. (2018) and Nandan et al. (2019), and may considerably increase the chance of small ΔM. However, they are out of the scope of this work.

1.4 Scope and Outline

This work consists of two regression studies, the analysis of magnitude differences (hereafter called ΔM-*regression*), and the analysis of the aftershock productivity (referred to as *productivity regression*). The focus 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. Especially 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.

Sections 2 and 3 introduce the utilized datasets, the declustering approaches and the compilation of the covariate datasets for both regression studies, respectively. Next, Section 4 rigorously explains the survival model and GAM methodological approaches. Then, the results of the regression studies are shown and discussed in Sections 5 (ΔM-regression) and 6 (productivity regression). Finally, conclusions are drawn from a joint interpretation and related future research topics are recommended.

2 Data for ΔM-Regression

This section summarizes the compilation of the regression dataset for the analysis of global 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 to decluster the catalog, 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 requires a trade-off. On the one hand, the catalog should ideally have homogeneous magnitude scales, and be reliably complete in any part of the world, including far off-shore regions and aftershocks occurring shortly after the mainshock. On the other hand, it should be complete to the smallest possible magnitude level to assure a sufficient observable magnitude range of at least one unit below the smallest mainshock magnitude of interest, $M > 6.0$.

Despite not providing homogenized magnitude scales, we chose the U.S. Geological Survey National Earthquake Information Center (USGS-NEIC) catalog. 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 clusters with mainshock magnitude larger than 6.0.

To test the influence of inhomogeneous magnitude scales, we performed sensitivity analyzes using the International Seismological Centre – Global Earthquake Model (ISC-GEM) instrumental cat-

8

alogue, 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 have to limit our statistical analysis to mainshocks with $M \geqslant 6.6$.

2.2 Declustering of Global Catalog

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 declustered the global earthquake catalog using 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 Wells and Coppersmith (1994). 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 and therefore, due to prior re-ordering of the catalog, has the larger magnitude, $m_A \geqslant m_B$. In this case, 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, but that these were simply 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 declustered catalog by 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). Fig. 1 shows the mainshock locations of the declustered catalog, color-coded by the corresponding plate boundary class assigned to them. Table 1 lists the number of clusters with censored and observed ΔM value per boundary class, respectively. Note that almost half of the clusters are assigned to a subduction zone, and that oceanic spreading ridges and transform faults host more censored than observed 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. Likewise, using a near-

10

est neighbor approach, we interpolated values from the scattered *heat flow* dataset of Bird et al. (2008), provided to us by the author. Fig. 2 illustrates the marginal 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 (values range between 0.4 and 262 mm/a, see Fig. 2(a)), while oceanic spreading ridges and transform faults provide the youngest sea floor ages (between 0 and 262 Ma, see Fig. 2(b)) and the largest heat flows (between 0.025 and 0.3 Wm^{-2}, see Fig. 2(b)).

3 Data for Productivity-Regression

This section summarizes the dataset compilation for the regression of aftershock productivity. First, we briefly introduce the chosen local event set for New Zealand. Then, we rigorously describe the stochastic declustering method which is applied in order to estimate the number of aftershocks as the response variable. Finally, we describe the enrichment of the local event set by further geophysical properties.

3.1 Local New Zealand Catalog

We limited this study to the Hikurangi subduction zone in New Zealand. A local event set was provided by GNS Science as an input to the ongoing 2022 revision of the New Zealand National Seismic Hazard Model. Using the algorithm of Stepp (1972), we computed that the catalog is complete down to $M_c = 3.5$ from 1982; however, to be conservative, we assumed $M_c = 3.5$ from 1987, concurrent with an improvement to the seismic network. Fig. 3 shows the chosen extract of 11,091 events surrounding the Hikurangi fault, between 1987 and end of 2020, at depths down to 80 km.

3.2 Declustering of Local Catalog

For the regression of aftershock counts, we cannot use the window declustering method, as it does not distinguish *direct* from *secondary* aftershocks. Instead, we used the stochastic declustering approach based on the Epidemic Type Aftershock Sequence (ETAS) model, as introduced by Zhuang et al. (2002).

The ETAS model describes the spatio-temporal clustering behavior of the entire catalog, and models a dynamic event rate at time t and location (x, y), given the prior event history H_t, through two overlapping processes,

$$R(t, x, y | H_t) = u(x, y) + \sum_{i:t_i < t} R^{trig}(t, x, y, i), \qquad (3)$$

where $u(x, y)$ denotes the time-invariant, aftershock-independent *seismic background* rate, and $\sum_{i:t_i < t} R^{trig}(t, x, y, i)$ is the sum of the *trigger rate* contributions by all events i that occurred prior to time t. For more details about the ETAS model, see e.g. Jalilian (2019); Ogata (1988, 1998); Zhuang et al. (2002).

From Equation (3), Zhuang et al. (2002) concluded that the probability, that the event j at time t_j and location (x_j, y_j) was an aftershock of the prior event i, is

$$P_{j,i} = \frac{R^{trig}(t_j, x_j, y_j, i)}{R(t_j, x_j, y_j | H_t)}.$$

Similarly, the probability that event j is a seismic background event and therefore independent of any prior trigger, is

$$P_{j,backgr} = \frac{u(x, y)}{R(t_j, x_j, y_j | H_t)}.$$

Thus, unlike the window method, the ETAS model provides probabilistic trigger associations between event pairs, and does not require the arbitrary choice of a fixed space-time window to search aftershocks. Instead, it optimizes built-in parametric functions that model the expected number of aftershocks (for trigger magnitude m_i),

$$K(m_i) = A\,e^{\alpha\,(m_i - M_c)}, \qquad A > 0, \quad \alpha > 0, \tag{4}$$

the temporal decay of aftershock rates (e.g. Omori-Utsu power law, see Omori, 1895), and a typically isotropic spatial distribution of aftershocks.

In this work, we used the *ETAS-Incomplete* model version of Grimm et al. (2022), who introduced a novel, anisotropic and locally restricted spatial kernel and accounted for incomplete records of aftershock sequences. The estimation source code is available in a public github repository (see Data and Resources). The new features solve the estimation biases due to the misfit of mostly elongated aftershock clouds by an isotropic kernel, and an underestimation of the trigger potential of strong mainshocks as a consequence of missing aftershock data. Both biases were shown to heavily affect our response variable, the aftershock productivity (Grimm et al., 2022, 2021; Hainzl, 2021; Hainzl et al., 2013; Page et al., 2016; Seif et al., 2017).

3.3 Response Variable

As the response variable of the regression model, we defined the estimated *number of direct aftershocks* for each event i in the catalog. We did this by counting the number of subsequent events j, for which i is the most probable trigger event, i.e. $P_{j,i} > P_{j,k} \ \ \forall \ k \neq i$, and that are more likely triggered by i than being a background event, i.e. $P_{j,i} > P_{j,backgr}$.

13

Note that the response is inevitably affected by the short-term incomplete records of aftershock sequences, as we can only count aftershocks that are recorded in the dataset. Nevertheless, the ETAS-Incomplete model approach avoids a biased parameter estimation, that would lead to manipulated declustering probabilities.

3.4 Covariates

The local catalog provides us the magnitude and depth for each event as immediate covariates for the regression model. Additionally, if the triggering event i was itself already triggered by a previous event, we traced back the trigger sequence and identified the largest magnitude in the cluster, that occurred before event i. This covariate tests whether an event, that is a member of a cluster, is more or less productive than an independent background event, and whether its aftershock productivity is influenced by the previous mainshock.

A focal mechanism data set comprising 1,581 events of the chosen catalog extract was provided by GNS Science. We used the *nodal-plane.py* function from the public *GEMScienceTools/oq-mbtk* repository, based on the algorithm in Álvarez-Gómez (2019), to translate the given focal mechanisms into the slip type categories *normal (N)*, *strike-slip (SS)* and *reverse (R)* and mixed categories.

We classified each earthquake in the catalogue to the main tectonic regions. These regions are the *shallow crust*, *subduction interface*, *subduction intraslab deep* (within the subducting plate, but deeper than the zone of contact between the subducting and overriding plate), and *subduction intraslab shallow* (within the subducting plate, but in the shallow part of the plate beneath the interface, e.g. Reyners et al., 2010). To do so, we used the methodology described by Pagani et al. (2020), which classifies each earthquake based on its hypocentral position relative to surfaces

14

(with buffers) that demarcate the boundaries of these regions. The surface used to represent the Hikurangi subduction interface and the top of the subducting plate was derived from a 3D model provided by GNS Science. Earthquakes shallower than 40 km and within 5 km of this surface were classified as interface; those shallower than 40 km and more than 5 km below the surface were classified as shallow slab; and those deeper than 40 km and within 5 km above or 60 km below this surface were classified as deep slab (the large below-slab buffer helps to capture earthquakes with large depth errors). Earthquakes shallower than the Moho (depths defined by LITHO1.0, Pasyanos et al., 2014) with a 10 km buffer were classified as crustal; crustal earthquakes were then sub-classified as occurring within or outside of the surface projection of the subduction zone. If an earthquake was classified into more than one tectonic region, then the following hierarchy was applied: interface is more likely than shallow slab, which is more likely than deep slab, which is more likely than shallow crustal. All other earthquakes were labelled as "unclassified" and not used in further analyses.

4 Regression Methods

In this section, we summarize the statistical models used in the two regression studies. Subsection 4.1 introduces survival regression models that can account for the censored ΔM response data due to unobserved aftershocks. Subsection 4.2 explains the use of a generalized additive model (GAM) for modeling aftershock counts in the local New Zealand catalog. All statistical analyses were performed with the open source software R (R Core Team, 2021).

4.1 Survival Models

4.1.1 Why Using a Survival Model for Earthquakes?

The magnitude difference ΔM between the mainshock and the second-largest earthquake of a se-

quence 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 statis-

tics, 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 un-

derestimate ΔM due to the relevant proportion of censored observations.

Replacing lifetimes by magnitude differences, our data meets the necessary requirements of a sur-

vival 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 sus-

pected to deviate structurally in their ΔM-distribution from non-censored clusters).

4.1.2 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

$$log(b(\mathbf{x_i})) = \beta_0 + \beta_1 \mathbf{x}_{class=CCB,i} + ... + \beta_6 \mathbf{x}_{class=OTF,i} +$$

$$f_{mag}(\mathbf{x}_{mag,i}) + f_{depth}(\mathbf{x}_{depth,i}) + f_{veloc}(\mathbf{x}_{veloc,i}) + f_{heat}(\mathbf{x}_{heat,i}) + f_{age}(\mathbf{x}_{age,i}),$$

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

17

₃₅₆ spline function *pspline* from the R package *survival* (Therneau, 2016), consistently using $df = 2$

₃₅₇ degrees of freedom and $n = 2.5 \times df$ basis functions (Eilers and Marx, 1996; Hurvich et al., 1998).

4.2 Generalized Additive Count Models

₃₅₉ In the second part of this study, we model the number of aftershocks N_i of each event i. The

₃₆₀ starting point for modeling count data response variables are so-called generalized linear models.

₃₆₁ Given covariate values $\mathbf{x}_{i1}, ..., \mathbf{x}_{ik}$, the expected aftershock productivity of event i is modeled by

₃₆₂ the log-linear relationship

$$\mathbb{E}[N_i] = \exp(\eta_i) \tag{5}$$

₃₆₃ where $\eta_i = \beta_0 + \beta_1 \mathbf{x}_{i1} + ... + \beta_k \mathbf{x}_{ik}$ is the linear predictor and $\beta_0, \beta_1, ..., \beta_k$ are the estimated

₃₆₄ coefficients. Note that the covariates have an *exponentially multiplicative effect* on the expected

₃₆₅ number of aftershocks (see e.g. Fahrmeir et al., 2013, section 5.2).

₃₆₆ In this work, we used a GAM approach and replaced the linear effects of all metric covariates

₃₆₇ by potentially smooth functions that can more flexibly represent varying effects in different value

₃₆₈ ranges of the covariates (e.g., see Fahrmeir et al., 2013, section 9.1). The full model is then

₃₆₉ specified by the predictor

$$\eta_i = \beta_0 + \sum_{k=1}^{5} \beta_k \, I(\mathbf{x}_{TR,i} = k) + \sum_{h=6}^{12} \beta_h \, I(\mathbf{x}_{SL,i} = h) + ...$$
$$f_{mag}(\mathbf{x}_{mag,i}) + f_{depth}(\mathbf{x}_{depth,i}) + ... \tag{6}$$
$$f_{mainshMag}(\mathbf{x}_{mainshMag,i}) \, I(\mathbf{x}_{isBackground,i} = false),$$

₃₇₀ where the β's are the estimated coefficients for the linear categorical effects of the tectonic region

₃₇₁ (TR, $k = 1, ..., 5$) and slip type (SL, $h = 6, ..., 12$), and β_0 is the intercept representing the refer-

372 ence categories *TR=crustal outside* and *SL=unknown*. The functions f_{mag}, f_{depth} and $f_{mainshMag}$

373 represent the smooth effects of the magnitude and depth of the triggering event as well as of the

374 prior mainshock magnitude, and $I(...)$ is the indicator function that is 1 if the inside condition is

375 fulfilled, and 0 otherwise.

376 To fit the model, we used the function *gam* from the R package *mgcv* (Wood, 2017), using a loga-

377 rithmic link function and the restricted maximum likelihood estimator (REML) as the smoothing

378 parameter estimation method. Penalized splines based on a basic spline basis (P-Splines) were

379 used to model the unspecified smooth functions (e.g., see Fahrmeir et al., 2013, section 8.1). We

380 used the standard smooth term function of *mgcv*, choosing $k = 5$ and $k = 8$ (for depth) as the

381 dimensions of the basis.

5 Results of the ΔM-Regression

383 In this section, we show and discuss the results of a parametric survival model fitted to the global

384 declustered earthquake catalog in order to describe the magnitude difference ΔM between the

385 mainshock and the second strongest event in the cluster. First, we justify and validate the distribu-

386 tion assumption for the response variable. Then we show and interpret the effects of the modeled

387 covariates. Finally, we assess the explanatory power of the model using a response residual plot.

5.1 Choice of Distribution Family

389 Following the simple simulation model outlined in the section *Introduction*, with parameters $b = 1$,

390 $A = 0.13$ and $\alpha = 2$, we fitted a Weibull, a Gompertz and a Generalized Gamma distribution to

391 the simulated magnitude differences ΔM. Fig. 4(a) shows the fits of the three distributions to the

392 kernel density estimator of the sampled data. The Gompertz distribution clearly provides the best

fit to 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. 4(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, showing generally good agreement.

5.2 Covariate Effects

Fig. 5 shows the covariate effects for the full parametric Gompertz survival model. 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.

5.2.1 Effect of Boundary Class

Fig. 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. In other words, mainshocks at oceanic, especially transform and divergent type boundaries, produce weaker second strongest

20

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.

5.2.2 *Effect of Mainshock Magnitude*

For values smaller than $M = 7.8$, the mainshock magnitude effect depicted in Fig. 5(b) confirms the well-established Bath'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 toward smaller ΔM.

However, this effect 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. 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.

5.2.3 *Effect of Mainshock Depth*

Fig. 5(c) shows that the effect of the mainshock depth is almost constant for depths smaller than 40 km. Between 40km und 50km, ΔM increases from roughly 1.2 to a new level of approximately

21

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 magnitude differences ΔM. A connection with missing data at greater depths is unlikely, as we are only interested in the largest aftershock rather than the entire sequence.

5.2.4 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.

5.3 Effect of Heat Flow

According to Fig. 5(e), regions with heat flow larger than $0.23W/m^{-2}$ 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.

22

5.3.1 Effect of Sea Floor Age

Fig. 5(f) shows that magnitude differences are substantially larger in young compared to old oceanic crusts. 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.

5.4 Response Residuals

Fig. 6 shows the response residuals (i.e., observed minus predicted values) plotted against predicted magnitude differences. Note that, as observations are censored, residuals are censored as well. Therefore, we show only residuals for *non-censored observations* here. This explains the superiority of negative residuals. The blue line represents the linear trend of the residuals.

The 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

23

at younger sea floors) are considerably more likely to be censored and therefore left out of the residuals statistics.

A similar argument holds for the negative linear trend of the residuals. For instance, the lowest mainshock magnitude, $M = 6.1$, can only have observed magnitude differences up to $\Delta M_{maxObs} = M - M_c = 1.1$. Therefore, if the model predicts $\Delta M > 1.1$, only negative residuals will occur in the statistic. As we move to larger predicted values on the x-axis, the selection bias affects even larger mainshock magnitudes.

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.

5.5 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 declustering 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.

6 Results of Productivity-Regression

In this section, we present the results of the GAM regression of aftershock count data in New Zealand. First, we justify and validate our choice of the Negative Binomial distribution instead

of the commonly used Poisson distribution for aftershock counts. Then we show and interpret the effects of the modeled covariates. Finally, we illustrate the impact of the results in a simulation experiment.

6.1 Choice of Distribution Family

Fig. 7(a) and (b) show the quantile-quantile (Q-Q) plots of the deviance residuals for the Poisson distribution (a) and the Negative Binomial distribution (b). The latter fits the data better, as it adjusts the variance independently of the mean and therefore allows for larger variation than the Poisson distribution, resulting in a more adequate representation of the upper and lower tails of the distribution. The dispersion parameter of the Negative Binomial fit is 2.3.

In Fig. 7(c) and (d), the corresponding model residuals are plotted against the linear predictors η_i (see Equation 6) for the Poisson (c) and Negative Binomial distribution (d). The Negative Binomial fit shows substantially less spread in the residuals than the Poisson fit, confirming that aftershock count data is rather Negative Binomial distributed.

Note that, according to our simplified simulation model, a larger variance of aftershock counts directly translates into a larger variance of ΔM. In other words, the Negative Binomial distribution increases the likelihood of particularly small ΔM.

Alternative approaches such as a Quasi Poisson or a zero-inflated model were tested, but did not stand out substantially from the respective basic models. Additionally, from a substantive point of view, there seems to be no causal reason for "excess zeros" that would suggest the use of zero-inflated approaches.

25

6.2 Covariate Effects

Fig. 8 and 9 present the exponential, multiplicative effects of the categorical and metrical covariates on the expected number of aftershocks according to the relationship in Equation (5). That is, if the exponential effect of a category is larger than 1, it has a positive impact on aftershock counts, and vice-versa. If the exponential effect is equal to 1, the model shows no effect of the respective category.

6.2.1 Effects of Categorical Variables

Fig. 8 shows the effects of the categorical covariates *Tectonic Region* and *Slip Type*. The effects of the tectonic region are presented relative to their reference category *crustal outside*. Crustal events in the subduction zone show a substantially increased aftershock productivity. The expected number of aftershocks is approximately 1.8 times larger than for crustal events outside of the subduction zone. For interface and slab events, both shallow and deep, no clear signal is found, as their confidence intervals overlap with the reference line at $\exp(0) = 1$. Unclassified events appear to have a slightly positive effect, but the uncertainty is large. Sensitivity tests with different buffer sizes of the slab do not consistently confirm the effect for unclassified events. The positive effect of crustal events on the aftershock productivity, compared to interface, might be explained by reduced seismic coupling in subduction zones (Hainzl et al., 2019). On the interface, a substantial part of the deformation is often aseismic (Lay et al., 2012). A second possibility is that, in the proximity of the study region, the crust contains a dense network of faults with a wide range of orientations, and therefore more structures that could be brought closer to failure by a change in stress conditions due to mainshock earthquakes. Slip type effects

26

540 are depicted relatively to their reference category *unknown*. None of the focal mechanisms appears

541 to be substantially more or less productive. Additionally, there seems to be no *selection effect* in

542 the sense that events where the focal mechanism is known have a common effect on the aftershock

543 productivity.

6.2.2 Effects of Magnitude and Depth

545 Fig. 9(a) shows an exponential effect of the triggering magnitude on aftershock productivity. How-

546 ever, this effect is enforced by the declustering approach, as the ETAS model fits the exponential

547 aftershock productivity function (4) to optimize the aftershock trigger rates $R^{trig}(t, x, y, i)$. There-

548 fore, the effect only has control character.

549 The effect of the depth, shown in Fig. 9(b), confirms the argumentation in the discussion of the

550 ΔM regression results, that increasing magnitude differences may coincide with reduced after-

551 shock productivity at higher depths. The physical reason may again be reduced seismic coupling

552 along the subduction interface relative to the shallow crust (Hainzl et al., 2019; Lay et al., 2012).

6.2.3 Effect of Mainshock Magnitude

554 Fig. 9(c) shows no clear trend in the effect of varying mainshock magnitudes. However, indepen-

555 dently of the size of the mainshock, triggered events in general appear to be two to three times

556 more productive than a comparable background event.

557 This finding has two possible explanations. On the one hand, it may be an indicator that the ETAS

558 model based declustering approach does not adequately disentangle spatio-temporal clusters in

559 the catalog, and incorrectly assigns too much aftershock productivity to the smaller events, at the

560 cost of the mainshock productivity. In other words, in seismically active periods, small events

27

may be simply more likely to be assigned offsprings than in seismically quiet times, but in reality, many of these aftershocks should perhaps be direct rather than secondary ones of the mainshock. Such a rearrangement of trigger relationships would have a strongly distorting effect on our model, in which we consider the number of *direct* aftershocks. Note, however, that we used here an ETAS model approach that accounts for short-term aftershock incompleteness as well as locally restricted, anisotropic spatial kernels and therefore already improves some of the major biases of common ETAS models (de Arcangelis et al., 2018; Grimm et al., 2022, 2021; Hainzl et al., 2008; Hainzl, 2021; Seif et al., 2017).

On the other hand, Zhuang et al. (2004) proposed that triggered events are more productive than background events, based on a similar study. It seems reasonable that during an on-going sequence the aftershock productivity could temporarily increase due to a higher level of energy prevalent in the tectonic system, compared to seismically quiet periods with occasional background activity. A doubling of the productivity parameter A in the simulation model (see Introduction), applied only to secondary triggering, led to a reduction of the expected magnitude difference ΔM from 1.2 to below 0.9 due to the increasing cluster sizes. This additional "boost" in triggering illustrates the relevance of the observed effect. The finding may also contribute to an explanation as to why the ETAS model tends to underestimate cluster sizes and doublet probabilities in forward simulations, as observed in my first contribution (Grimm et al., 2021). Further research is recommended to evaluate this finding.

7 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

28

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 declustered 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). From a simplified simulation model, assuming an exponential aftershock productivity law and the Gutenberg-Richter type magnitude size distribution, we concluded that the Gompertz distribution may be the better choice than Weibull or Generalized Gamma.

The regression results show that larger ΔM are expected at higher depths and in younger ocean crust. This may be an indication, that aftershock productivity is a relevant driver of ΔM, as in these conditions lower aftershock productivity is expected due to reduced seismic coupling (Hainzl et al., 2019).

In the second part of this study, we used the stochastic declustering method of Zhuang et al. (2002) to estimate the aftershock productivity per event in a local catalog for New Zealand. To do so, we used the anisotropic ETAS-Incomplete model (Grimm et al., 2022) to disentangle the trigger relations between events. We further enriched the event set by a categorization of events in tectonic regions and slip types and used a generalized additive regression approach to model the aftershock productivity.

The results clearly confirm that aftershock counts follow a Negative Binomial rather than a Poisson distribution (Kagan, 2017; Shebalin et al., 2018). Also, aftershock productivity decreases with

29

increasing depth, supporting the reasoning regarding the depth effect on ΔM above. Furthermore, results indicate that triggered earthquakes trigger themselves two to three times more aftershocks than non-triggered ones. In other words, secondary aftershock triggering is substantially stronger than direct triggering by background events. This effect may either be an indicator for wrongly disentangled sequences in the sense that secondary triggering is overestimated at the cost of the productivity of the mainshock, or it may represent an actual effect due to the temporarily higher energy level after the occurrence of a strong mainshock. A causal effect, if confirmed, would have an enormous impact on the expected cluster sizes (compare with Grimm et al., 2021) and could explain some of the rather small ΔM observations in the first study.

Future research should identify whether small magnitude differences ΔM are typically character-ized rather by above-average aftershock productivity or by magnitude size distributions favoring large aftershocks. To do so, one could compile a sufficiently large set of earthquake sequences and analyze the correlation of their ΔM with estimates of the aftershock productivity (Equation 4) and frequency-magnitude distribution (Equation 2). Additionally, it should be verified whether triggered events indeed have a larger aftershock productivity, and how this effect impacts ΔM. Similarly, potential correlations of aftershock magnitudes with their ancestors should be evaluated. Finally, an extension of the ΔM-regression model using small-scale covariate data could certainly 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://

30

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). The New Zealand event set and focal mechanism data was provided by GNS Science as an input to the ongoing 2022 revision of the New Zealand National Seismic Hazard Model. For the stochastic declustering, we used the ETAS-Incomplete model source code available in the Github repository https://github.com/ChrGrimm/ETASanisotropic (Grimm et al., 2022), implemented using the software *Matlab*. All statistical analyses were performed with the open source software *R*.

31

that there are no relevant financial or non-financial interests to disclose. This article does not contain any studies involving human participants or animals performed by any of the authors.

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. TR conducted and evaluated the productivity regression model. SH, KJ and MK contributed significantly to the interpretation of the model results in the geophysical context. Additionally, KJ categorized New Zealand events into tectonic regions. MP und 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.

Álvarez-Gómez, J. A. (2019). FCM - Earthquake focal mechanisms data management, cluster and classification. *SoftwareC*, **9**, 299–307.

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.

Dascher-Cousineau, K., Brodsky, E. E., Lay, T., and Goebel, T. H. (2020). What controls variations in aftershock productivity? *Journal of Geophysical Research: Solid Earth*, **125**. doi: 10.1029/2019JB018111.

de Arcangelis, L., Godano, C., and Lippiello, E. (2018). The overlap of aftershock coda waves and short-term postseismic forecasting. *Journal of Geophysical Research: Solid Earth*, **123**, 5661–5674. doi: 10.1029/2018JB015518.

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., Ver-
ney, 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.

Fahrmeir, L., Kneib, T., Lang, S., and Marx, B. (2013). *Regression - Models, Methods and Appli-
cations*. ISBN 9783642343322. doi: 10.1007/978-3-642-34333-9.

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*, **2**. 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*, (Xx). doi: 10.1785/0120210097.

Gulia, L., Rinaldi, A. P., Tormann, T., Vannucci, G., Enescu, B., and Wiemer, S. (2018). The effect of a mainshock on the size distribution of the aftershocks. *Geophysical Research Letters*, **45**(24), 13,277–13,287. doi: 10.1029/2018GL080619.

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.

Hainzl, S. (2021). ETAS-Approach Accounting for Short-Term Incompleteness of Earthquake Catalogs. *Bulletin of the Seismological Society of America*. doi: 10.1785/0120210146.

Hainzl, S., Zakharova, O., and Marsan, D. (2013). Impact of aseismic transients on the estimation of aftershock productivity parameters. *Bulletin of the Seismological Society of America*, **103**(3), 1723–1732. doi: 10.1785/0120120247.

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.

Jalilian, A. (2019). ETAS: An R package for fitting the space-time ETAS model to earthquake data. *Journal of Statistical Software*, **88**(1). doi: 10.18637/jss.v088.c01.

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.

Kagan, Y. Y. (2017). Earthquake number forecasts testing. *Geophysical Journal International*, **211**(1), 335–345. doi: 10.1093/gji/ggx300.

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.

Lay, T., Kanamori, H., Ammon, C. J., Koper, K. D., Hutko, A. R., Ye, L., Yue, H., and Rushing, T. M. (2012). Depth-varying rupture properties of subduction zone megathrust faults. *Journal of Geophysical Research: Solid Earth*, **117**(4), 1–21. doi: 10.1029/2011JB009133.

Marsan, D. and Helmstetter, A. (2017). How variable is the number of triggered aftershocks? *Journal of Geophysical Research: Solid Earth*, **122**(7), 5544–5560. doi: 10.1002/2016JB013807.

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.

Nandan, S., Ouillon, G., and Sornette, D. (2019). Magnitude of earthquakes controls the size distribution of their triggered events. *Journal of Geophysical Research: Solid Earth*, **124**(3), 2762–2780. doi: 10.1029/2018JB017118.

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.

37

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. *Tectono-physics*, **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.

Pagani, M., Johnson, K., and Garcia-Pelaez, J. (2020). Modelling subduction sources for proba-bilistic seismic hazard analysis. *Geological Society, London, Special Publications 501*.

Page, M. T., van Der Elst, N., Hardebeck, J., Felzer, K., and Michael, A. J. (2016). Three in-gredients for improved global aftershock forecasts: Tectonic region, time-dependent catalog incompleteness, and intersequence variability. *Bulletin of the Seismological Society of America*, **106**(5), 2290–2301. doi: 10.1785/0120160073.

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.

Pasyanos, M., Masters, T., Laske, G., and Ma, Z. (2014). LITHO1. 0: An updated crust and lithospheric model of the Earth. *Journal of Geophysical Research: Solid Earth*, **119**(3), 2153–2173.

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.

Reyners, M., Eberhart-Phillips, D., and Ristau, J. (2010). Final report on EQC Project BIE 08/549–Quantifying the hazard from New Zealand's shallow intraslab earthquakes. Technical report.

Seif, S., Mignan, A., Zechar, J. D., Werner, M. J., and Wiemer, S. (2017). Estimating ETAS: The effects of truncation, missing data, and model assumptions. *Journal of Geophysical Research: Solid Earth*, **122**(1), 449–469. doi: 10.1002/2016JB012809.

Shebalin, P. N., Baranov, S. V., and Dzeboev, B. A. (2018). The law of the repeatability of the number of aftershocks. *Doklady Earth Sciences*, **481**(1), 963–966. doi: 10.1134/S1028334X18070280.

Stepp, J. C. (1972). Analysis of completeness of the earthquake sample in the Puget Sound area and its effect on statistical estimates of earthquake hazard. In *Proc. of the 1st Int. Conf. on Microzonazion*, volume 2, pages 897–910.

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.

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.

Wetzler, N., Brodsky, E. E., and Lay, T. (2016). Regional and stress drop effects on aftershock productivity of large megathrust earthquakes. *Geophysical Research Letters*, **43**(23), 12,012–12,020. doi: 10.1002/2016GL071104.

Wood, S. (2017). Generalized Additive Models: An Introduction with R. *Boca Raton, FL: CRC Press.*

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.

826 Zhang, L., Werner, M. J., and Goda, K. (2018). Spatiotemporal seismic hazard and risk assessment

827 of aftershocks of M 9 megathrust earthquakes. *Bulletin of the Seismological Society of America*,

828 **108**(6), 3313–3335. doi: 10.1785/0120180126.

829 Zhuang, J., Ogata, Y., and Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake

830 occurrences. *Journal of the American Statistical Association*, **97**(458), 369–380. doi: 10.1198/

831 016214502760046925.

832 Zhuang, J., Ogata, Y., and Vere-Jones, D. (2004). Analyzing earthquake clustering features by

833 using stochastic reconstruction. *Journal of Geophysical Research: Solid Earth*, **109**, B05301.

834 doi: 10.1029/2003JB002879.

Contact to authors

Christian Grimm (main author), *Ludwig-Maximilians-University Munich, Department of Statistics, Ludwigstraße 33, 80539 Munich, Germany; Christian.Grimm@stat.uni-muenchen.de*,

Teresa Rupprecht, *Ludwig-Maximilians-University Munich, Department of Statistics, Ludwigstraße 33, 80539 Munich, Germany*,

Kendra Johnson, *Global Earthquake Model Foundation, Via Ferrata 1, 27100 Pavia, Italy*,

Sebastian Hainzl, *GFZ German Research Centre for Geoscience, Physics of Earthquakes and Volcanoes, Helmholtzstraße 6/7, 14467 Potsdam, Germany*,

Helmut Küchenhoff, *Ludwig-Maximilians-University Munich, Department of Statistics, Ludwigstraße 33, 80539 Munich, 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*,

Marco Pagani, *Global Earthquake Model Foundation, Via Ferrata 1, 27100 Pavia, Italy*

List of Tables

43

Table 1: Number of censored and observed ΔM data points, grouped by plate boundary class.

Plate Boundary Class	Number of Clusters	
	#censored	#observed
CCB	71	148
CTF	82	157
CRB	69	114
OSR	108	73
OTF	322	194
OCB	59	83
SUB	474	989

List of Figures

45

9   Exponential, multiplicative effects of the metric covariates *magnitude* and *depth* of the triggering event as well as *mainshock magnitude*, given that the triggering event was already part of a triggered sequence. Exponential effects larger than one signify a positive effect on aftershock productivity, and vice versa. Rug lines on the x axis visualize the marginal distributions of the corresponding covariate.

Fig 1: Locations of 2,933 global $M > 6$ mainshocks between 1973 and 2020 after declustering the USGS-NEIC catalog. Mainshocks are colour-coded according to their assignment to the 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)*, introduced in the digital plate model of Bird (2003).
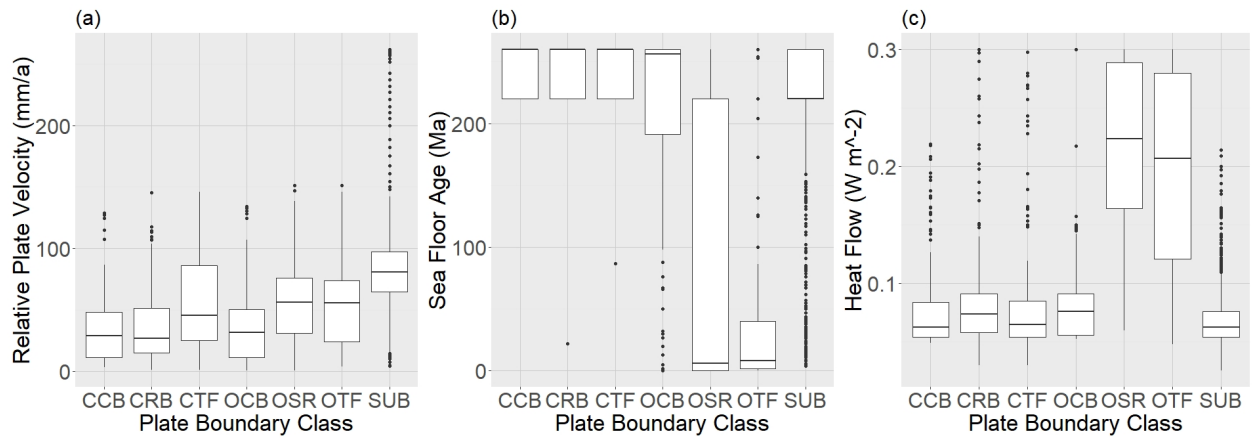
Fig 2: Boxplots of (a) relative plate velocity, (b) sea floor age, and (c) heat flow values assigned to cluster mainshocks by a nearest approach from original scatter data (Bird, 2003; Bird et al., 2008), grouped by the plate boundary class. Acronyms of boundary classes are spelled out in the caption of Fig. 1.
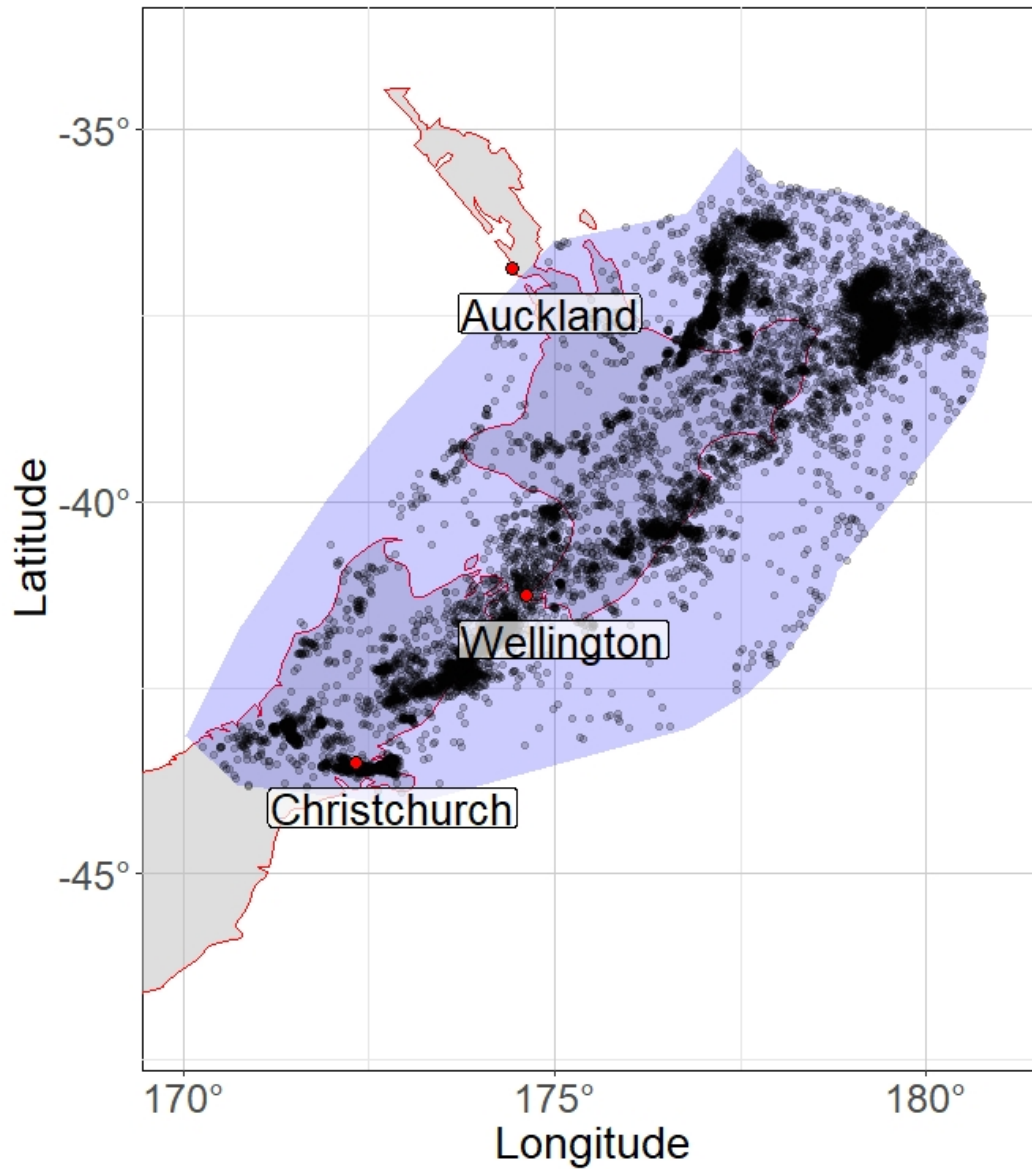
Fig 3: Spatial extract (blue polygon) of the Hikurangi subduction region in New Zealand, chosen for the aftershock productivity regression model. Black scatter points represent event locations of earthquakes with magnitude $M \geqslant 3.5$, depths $\leqslant 80$ km, that occurred between 1987 and 2020. The local event set was provided by GNS Science as an input to the ongoing 2022 revision of the New Zealand National Seismic Hazard Model.
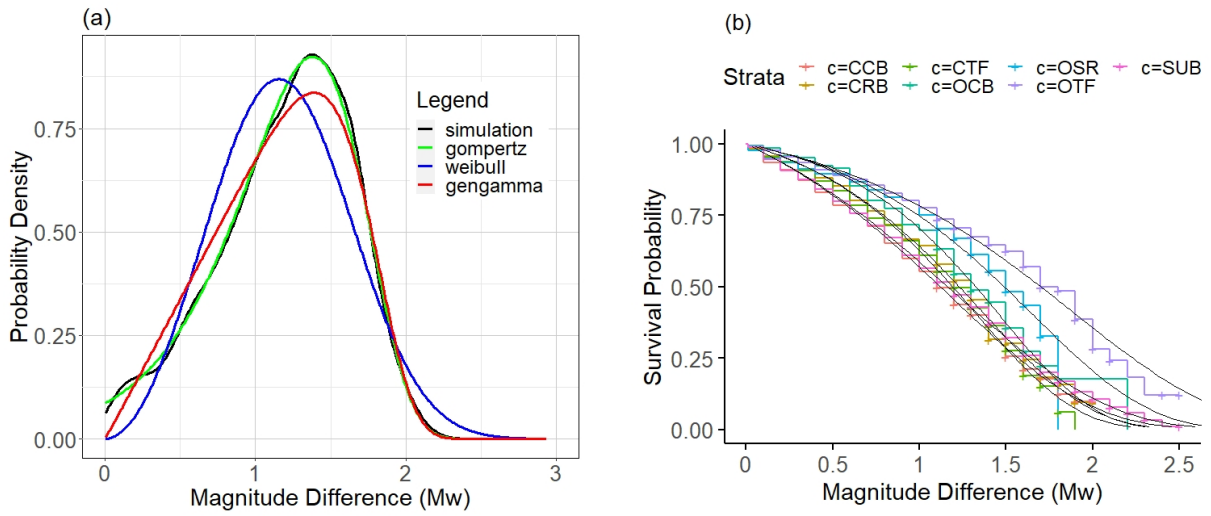
Fig 4: (a) Fits of a Gompertz, Weibull and Generalized Gamma distribution to simulated magnitude differences ΔM, represented by the kernel density estimator (black curve). (b) Comparison of survival curves estimated from a Gompertz model and a non-parametric Kaplan-Meier estimator, stratified for plate boundary classes (c).
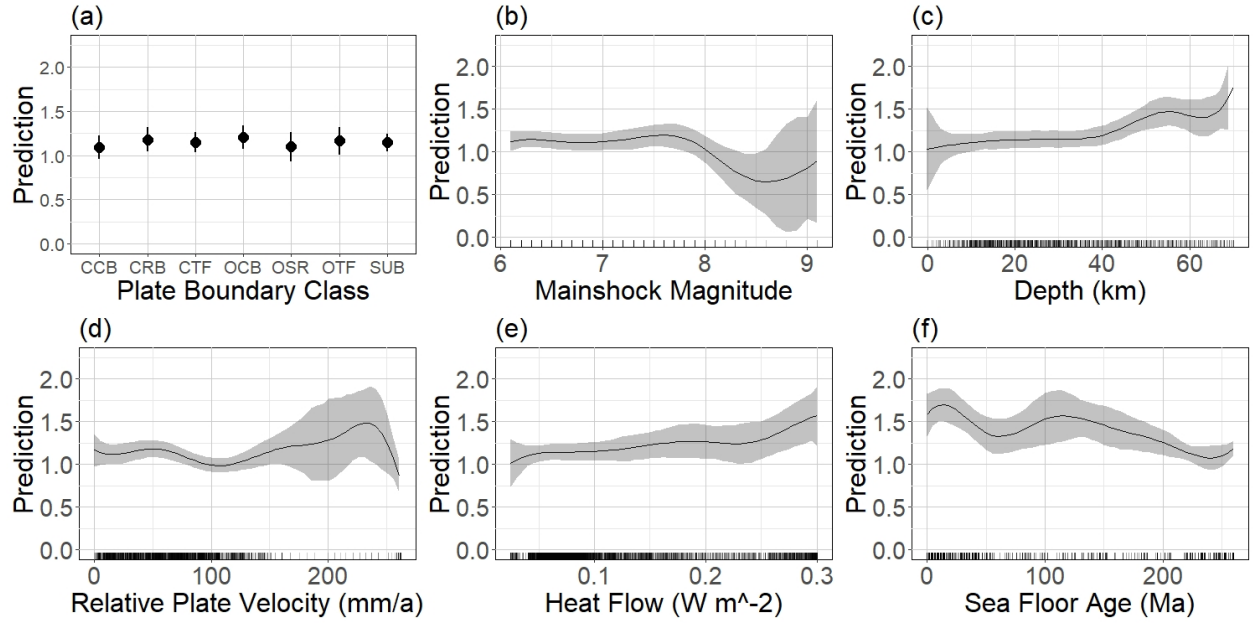
Fig 5: Covariate effects of the ΔM-Regression, by (a) plate boundary type (categorical), (b) mainshock magnitude, (c) mainshock depth, (d) relative plate velocity, (e) heat flow and (f) sea floor age on the magnitude difference between a mainshock and the second largest event of the cluster. For linear 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.
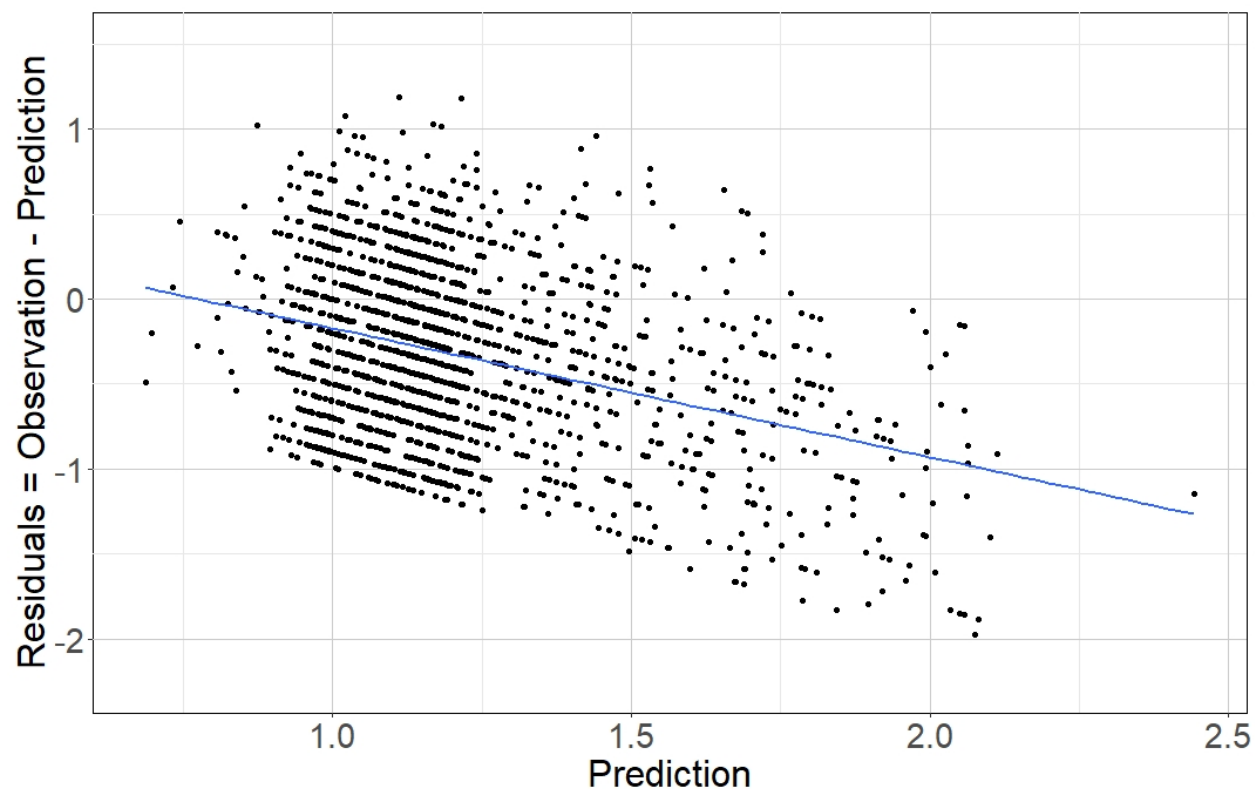
Fig 6: Response Residuals of the ΔM-regression for non-censored observations only, plotted against predicted values. The blue line represents the linear trend of the residuals. The row arrangement of the points is due to the rounding of the observed data to one decimal place. For instance, the bottom row represents observations where $\Delta M = 0$.
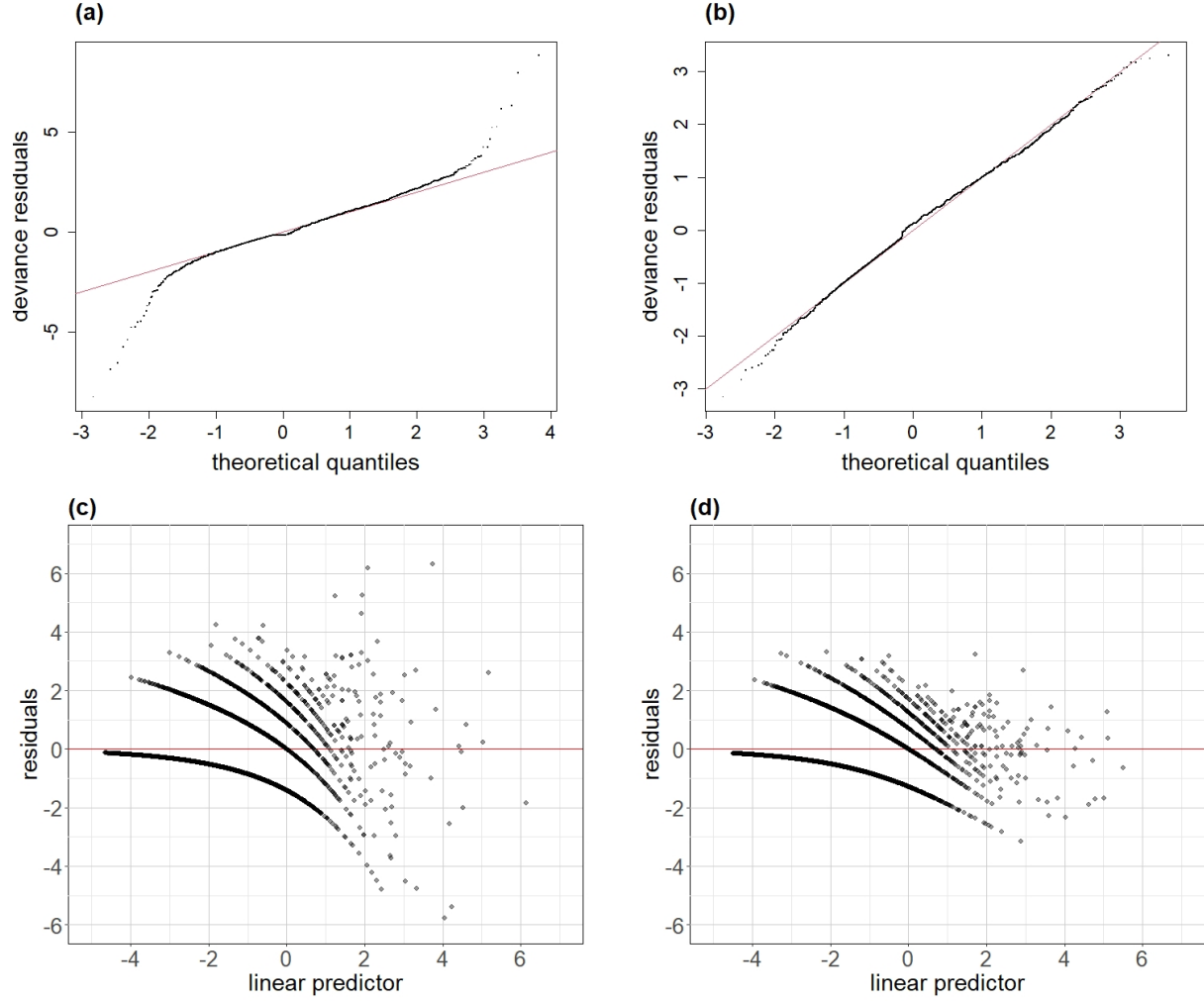
Fig 7: *Top row:* Quantile-Quantile plots of the deviance residuals for the (a) Poisson and the (b) Negative Binomial regression of the aftershock productivity. *Bottom row:* Corresponding model residuals plotted against the linear predictors η_i (see Equation 6) for the (c) Poisson and (d) Negative Binomial regression. The row arrangement of points is due to the count data structure of the response.
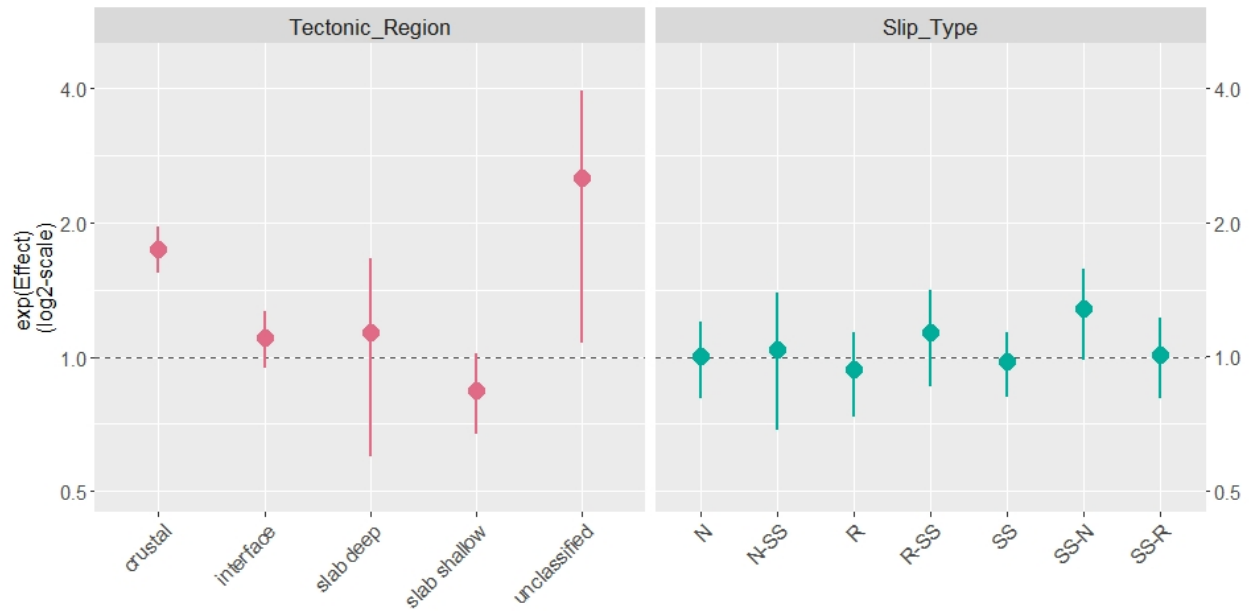
Fig 8: Exponential, multiplicative effects of the categorical covariates *tectonic region* and *slip type* relative to their reference categories "crustal outside" and "unkown", respectively, according to Equation (5). Exponential effects larger than one signify a positive effect on aftershock productivity, and vice versa.
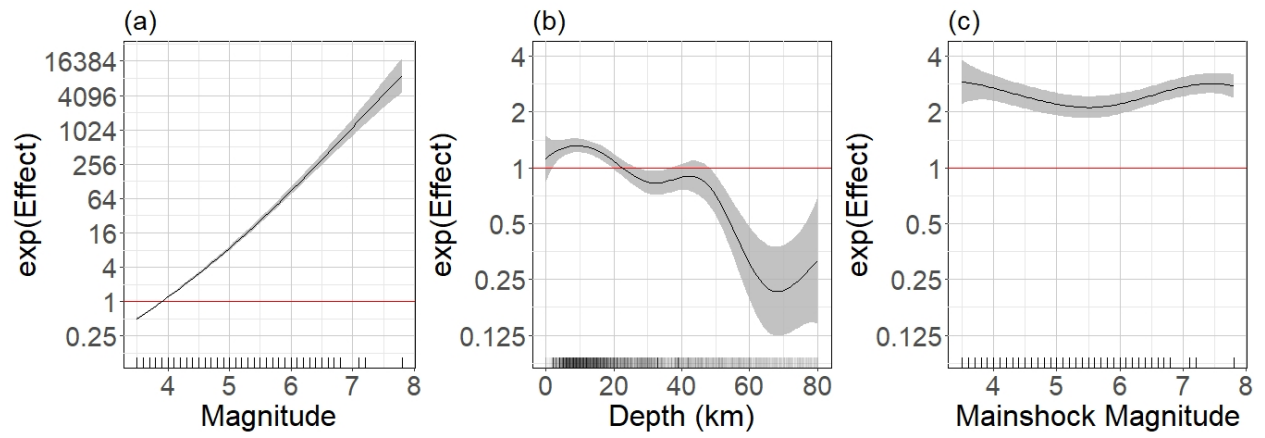
Fig 9: Exponential, multiplicative effects of the metric covariates *magnitude* and *depth* of the triggering event as well as *mainshock magnitude*, given that the triggering event was already part of a triggered sequence. Exponential effects larger than one signify a positive effect on aftershock productivity, and vice versa. Rug lines on the x axis visualize the marginal distributions of the corresponding covariate.