Surveys in Geophysics (2022) 43:437–481
https://doi.org/10.1007/s10712-022-09698-0

Modeling and Prediction of Aftershock Activity

Sergey Baranov[1,3] · Clement Narteau[2] · Peter Shebalin[3]

Received: 7 March 2021 / Accepted: 19 February 2022 / Published online: 13 April 2022

Abstract

We provide an overview of the basic models of the aftershock processes and advanced methods used to predict postseismic hazard. We consider both the physical mechanisms for aftershock generation and models of aftershocks and time-dependent models of aftershock processes. In particular, we provide a validation of the aftershock process using a superposition of the Gutenberg–Richter and Omori–Utsu laws. We show that the key role in assessment of postseismic hazard is earthquake productivity, which characterizes the ability of earthquakes to produce subsequent shocks. We discuss the recently established exponential law of earthquake productivity and show that the exponential form is invariant under variations in magnitude and focus depth. Being in discordance with the popular epidemic type aftershock sequence (ETAS) model, the law makes it possible to build a corrected model. We study versions of theoretical validation for the Båth law, which specifies the mean difference between the magnitudes of the main shock and the largest aftershock. We consider also the time-dependent Båth law. We provide a detailed review of modern approaches and methods for dealing with the estimation of the magnitude of the largest aftershock. As well, we review the problem of estimating the duration of aftershocks with magnitudes equal to or greater than a specified value, the hazardous period.

Keywords Aftershocks · Models · Earthquake productivity · Båth law · Aftershock hazard

✉ Sergey Baranov
bars.vl@gmail.com

Clement Narteau
narteau@ipgp.fr

Peter Shebalin
p.n.shebalin@gmail.com

[1] Federal Research Center Geophysical Survey, Russian Academy of Sciences, 14 Fersmana, Apatity 184209, Russia

[2] Institut de Physique du Globe de Paris, Universite de Paris, CNRS, 75005 Paris, France

[3] Institute of Earthquake Prediction Theory and Mathematical Geophysics, 84/32 Profsouznaya, Moscow 117997, Russia

Article Highlights

- Earthquake productivity plays the key role for evaluation of postseismic hazard
- The productivity of direct aftershocks varies widely from case to case and statistically obeys an exponential law
- The combination of the exponential productivity law with the Gutenberg-Richter and Omori laws makes it possible to estimate future aftershock activity

1 Introduction

It frequently happens that a significant earthquake is followed by subsequent shocks along the earthquake rupture plane (aftershocks). The earthquake that is followed by aftershocks is called a main shock. Although the bulk of the energy release commonly comes from the main shock, aftershocks may turn out at least as destructive as the main shock. A main shock causes considerable weakening of a building structure to be subsequently destroyed by a large aftershock. It sometimes happens that the epicenter of a large aftershock turns out to be nearer a population center, so that its impact may happen to be greater, despite the magnitude being smaller. A phenomenon of this kind occurred, e.g., in New Zealand in 2010–2011. The main shock of September 4, 2010, near the town of Darfield (magnitude 7.1) did not produce much damage, because the epicenter was far from the main residential blocks. However, the aftershock of February 22, 2011 (magnitude 6.3) at the city of Christchurch was much more destructive, causing loss of 185 lives. Another aftershock took place on June 13, 2011, and caused more serious damage and killed a person.

This paper reviews the most popular models of aftershock processes, modern methods, and approaches to the assessment of aftershock hazard following large earthquakes.

2 Models of Aftershock Activity

Aftershocks are caused by relaxation of stress concentrations due to a main shock rupture. Probably, the first theory of aftershock occurrence was put forward by Benioff (1951), who explained the retardation in elastic energy release in an aftershock sequence by a successive diminution of static friction. These ideas were subsequently developed by Dieterich (1992, 1994, 2007; Scholz 1998—review article; Heimisson and Segall 2018), who used the rate-and-state model of aftershock generation based on the theory of nonlinear dry friction that depends on rupture velocity and the condition of the fault (rate-and-state friction model). According to the rate-and-state model, aftershocks are generated via a stepwise variation of stress occurring at the time of the main shock, while the decay follows the Omori law (see later in this paper).

According to the rate-and-state model, the intensity of seismicity after the stress drop ΔS exceeds the background intensity before the stress drop. Thus, the main shock gives rise to a stress change, which leads to increased rupture velocity on the fault and decreased time until failure (clock advance). In other words, aftershocks are just earthquakes whose times of occurrence have been accelerated compared with the background events owing to a stress change due to the main shock.

A good example of the rate-and-state model is an increase (by more than two times) in aftershock rate during sea tides near the shores of Kamchatka and New Zealand (Baranov et al. 2019a; Shebalin and Baranov 2020). This provides evidence of two main mechanisms responsible for ocean tides as a factor influencing the rate of seismicity, namely increased pore pressure during high water or decreased normal stress during low water, which leads to a decrease in effective friction on faults.

At present, there are several physical models of aftershock processes and mechanisms of aftershock generation. The most complete classification accompanied by a description and an exhaustive review of the literature can be found in (Lindman 2009; Hainzl et al. 2010). A general characterization of aftershock processes from the viewpoint of rock failure is given in the well-known monograph of Scholz (2019). During the last 30 years, it has been generally thought that the time and location of aftershock occurrence are largely controlled by stress changes due to the main shock rupture. A connection between static stress changes and the spatial distribution of aftershocks was proposed by Kostrov and Das (1982) based on the results of a study in the aftershocks of several earthquakes (Das and Scholz 1981). Stein et al. (1992) showed that small events that occurred near the town of Landers, California, 17 years before the M 7.3 earthquake of 1992 led to an increase in the stresses at the epicenter and along the future rupture. As well, most $M_L > 1$ aftershocks of the Landers earthquake occurred (Fig. 1) in the area where the failure stress increased by more than 0.1 bar (King et al. 1994). Few events are located where the stress has dropped.

Later, numerous studies revealed a quantitative correspondence between positive static stress changes and the location of subsequent shocks and aftershocks (Steacy et al. 2005; King and Devès 2015; Hainzl et al. 2010; Hardebeck and Okada 2018; Lasocki et al. 2020).

Stress transfer models are based on a straightforward application of the Coulomb criterion

$$\Delta CFF = \Delta\tau + \mu + \left(\Delta\sigma_n + \Delta p\right) \approx \Delta\tau + \mu'\Delta\sigma_n$$

where $\Delta\tau$ is the shear stress in the direction of slip, $\Delta\sigma_n$ is the normal stress changes (positive for unclamping or extension), μ is the friction coefficient, Δp is the pore pressure change, $\mu' = \mu(1 - B)$ is usually called the effective friction coefficient, and $0 \le B \le 1$ is the Skempton coefficient (Beeler et al. 2000; Cocco and Rice 2002).

A good example of the Coulomb criterion is induced seismicity in oil production zones in Alberta (Canada) where earthquake clustering is strongly related to pore pressure increase due to the nearby hydraulic fracturing operations (Schultz et al. 2017). In 2019, in the same area, fluid injection reactivated the preexisting faults by pore pressure diffusion and triggered earthquake sequence following the Richter local magnitude (ML) 4.18 main shock (Wang et al. 2020).

The main unresolved difficulties of straightforward application of the Coulomb criterion with regard to the assessment of the hazard due to subsequent shocks are (Hainzl et al. 2010): (1) the unknown distribution and orientation of the faults affected by the rupture; (2) nonuniqueness associated with the inversion of rupture models and fault geometry for the main shock; (3) small-scale variability of the rupture, which can give rise to large inhomogeneities in the stresses near the source fault, and (4) the spatial inhomogeneity of material and initial stress values. Besides, some scientists note significant inconsistency of the stress transfer mechanism with observations: aftershocks do not fill in all stress triggering lobes, the seismicity rate declines are not always apparent at negative Coulomb stress change areas (Marsan 2003; Felzer and Brodsky 2005), aftershocks are commonly abundant not only near the tips of the fault rupture, where a highest positive stress change is predicted, but also on

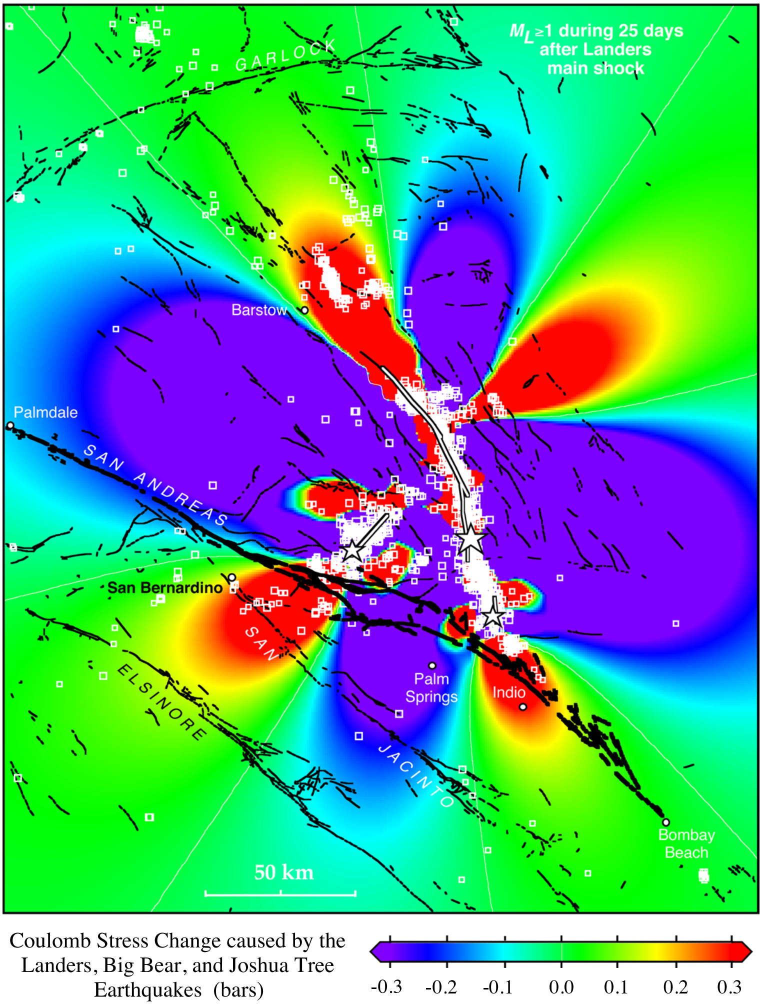


Fig. 1 Coulomb stress changes at a depth of 6.25 km caused by the Landers, Big Bear, and Joshua Tree earthquakes (King et al. 1994, Fig. 12)

the sides of the rupture zone (Hu et al. 2013). Some authors find that static Coulomb failure function (CFF) triggering is not statistically significant (Hardebeck et al. 1998; Kato 2006).

Although most researchers are unanimous concerning the point of view stating that aftershocks are triggered by stress steps due to a main shock, the time-dependent changes can also be important. In particular, viscoelastic relaxation in the lower crust (Freed and Lin 2002) and aseismic slip after the earthquake along the fault plane (afterslip) can affect the stress field and rate of events on scales ranging between a few days and a few years (Chan and Stein 2009; Helmstetter and Shaw 2009; Wang et al. 2010a; Sun et al. 2014), which decays according to the Omori–Utsu law.

Several researchers believe that seismic waves can trigger aftershocks (Freed, 2005; Felzer, Brodsky 2006; Sobolev and Zakrzhevskaya 2013; Kocharyan 2016; Zavyalov et al. 2017; Zotov et al. 2018), especially in geothermal fields (Hill et al. 1993; Brodsky 2006).

Baranov et al. (2019a) also showed that the dynamic transfer of stress due to sea tides enhances the local rate of aftershocks in the ocean during high rates of tide decrease.

Gomberg and Johnson (2005) supposed that the rather rarely observed effect of dynamic triggering of aftershocks is apparently associated with a pronounced directionality of seismic radiation in such cases. It is usually assumed that dynamic stresses can promote failure at remote distances, while at short distances the influence of static stress transfer dominates (Gomberg et al. 2001); however, Kilb et al. (2000) showed that dynamic stresses can also trigger aftershocks nearby. Dynamic triggering of earthquakes may also explain that the significant fraction of aftershocks often is found to occur in main shock stress shadows in static Coulomb stress triggering studies (Felzer et al., 2002).

The idea of direct aftershock triggering by the dynamic transfer of stress was also based on an interesting feature of the spatial localization of aftershocks: The aftershock rate decays following a power law as we move away from the main shock (Huc and Main 2003; Felzer and Brodsky 2006; Richards-Dinger et al. 2010). The last two papers reflect the well-known discussion concerning the dynamic transfer of stress to aftershock triggering. Felzer and Brodsky (2006) came to the conclusion that the observed power law for the distribution of the main shock–aftershock distance (Fig. 2) is consistent with the fact that the probability of aftershock occurrence is practically proportional to seismic wave amplitude. These authors also showed that this distribution is in poor agreement with rate-and-state models, which describe movements on a fault involving a friction that is a function of static stress change, velocity, and state (Dieterich 1994; Scholz 1998). Bearing in mind this observation, as well as the fact that changes in static stress for more remote aftershocks are small, the authors hypothesized that aftershocks can be triggered by the dynamic transfer of stress from the main shock. Richards-Dinger et al. (2010) used an algorithm from Felzer and Brodsky (2006) to identify main shocks and aftershocks, showing that the power law decay in the rate of subsequent shocks over distance also applies to those aftershocks which had occurred before the seismic wave excited by the main shock arrived, and this violates the principle of causality when we deal with the dynamic transfer of stress. It thus appears that the power law decay in

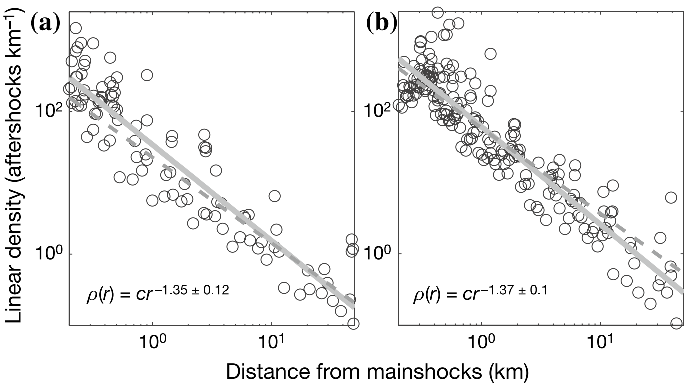


Fig. 2 Distribution of distance from main shock hypocenters to its aftershock (Felzer and Brodsky 2006, Fig. 2). Aftershocks are $M_m > 2$ and occur in the first 5 min. **(a)** $2 \le M_m < 3$ main shock; **(b)** $3 \le M_m < 4$ main shocks. The data are fitted with an inverse power law (gray solid line). The fit is made from 0.2 to 50 km for both plots. Dashed lines give the decay of maximum seismic wave amplitude, a proxy for dynamic stress, as derived from the standard Richter relationship

aftershock rate over distance does not support the hypothesis of the dynamic transfer of stress-producing aftershocks.

Other natural explanation of the specific spatial distribution of the aftershocks is fractal structure of the preexisting fracture systems where aftershocks occur. It was found that aftershock distributions become less clustered with increasing fractal dimensions of the active fault system (Nanjo and Nagahama 2000).

2.1 The Omori Law

Another approach to modeling aftershock processes is offered by statistical seismology. The law due to Omori (1984) is the first empirical model which states that the rate of aftershocks decays over time t as a hyperbolic function, $1/t$. The history of the discovery of the law is briefly provided by Guglielmi and Zavyalov (2018) whose work is devoted to the 150th anniversary of Fusakichi Omori.

An analysis of aftershock activity for the Ms 8.0, 1891 earthquake of Nobi, Japan, showed that the hyperbolic decay has been preserved even after 100 years since the main shock (Utsu et al. 1995). This long-continued aftershock activity is explained by the fact that Nobi lies in a relatively quiet seismic region of Japan. Where more active seismic regions are concerned, the rate of aftershocks reaches the background level faster.

Utsu (1961) remarked that there are several cases in which the aftershock decay occurs faster than is prescribed by the function $1/t$ and accordingly proposed a generalization of the Omori law commonly termed the Omori–Utsu law or the modified Omori law:

$$\lambda(t) = \frac{K}{(t+c)^p} \tag{1}$$

where t is the time elapsed since the main shock; $\lambda(t)$ is the rate of aftershocks at time t; and K, c, and p are positive parameters in the model. (If $p=1$, then Eq. (1) becomes the original Omori law.)

According to the results from at least 50 studies reviewed by Utsu et al. (1995), the parameter p has values in the range between 0.6 and 2.5, with the median being 1.1. Worldwide statistics demonstrates distribution of both c and p values close to normal (Fig. 3).

The parameter K in Eq. (1), which is commonly called productivity, is a function of the lower magnitude level M_0 for the aftershocks considered. The parameters p and c determine the shape of the cumulative curve (total number of aftershocks versus time). If the data are complete in the time span of interest, then p is independent of M_0, while c is decreasing with increasing M_0 (Utsu et al. 1995; Nanjo et al. 2007).

The physical meaning of K and c in Eq. (1) can be gathered from the rate-and-state model (Dieterich 1994; Cocco et al. 2010). The productivity K and the delay of the start of hyperbolic decay in aftershock activity c depend on the rate of tectonic deformation, the stress drop at the main shock, characteristics of the fault zone, and the seismicity rate before the main shock (Fig. 4). The parameter p referred to as the relaxation parameter depends on the velocity of aseismic slip after the earthquake on the fault plane (afterslip) (Helmstetter and Shaw 2009).

Narteau et al. (2009) found the relationship between c and the difference of stress and strength in the aftershock area. Shebalin and Narteau (2017) discovered a persistent tendency for c to increase with increasing depth. It was found in laboratory experiments

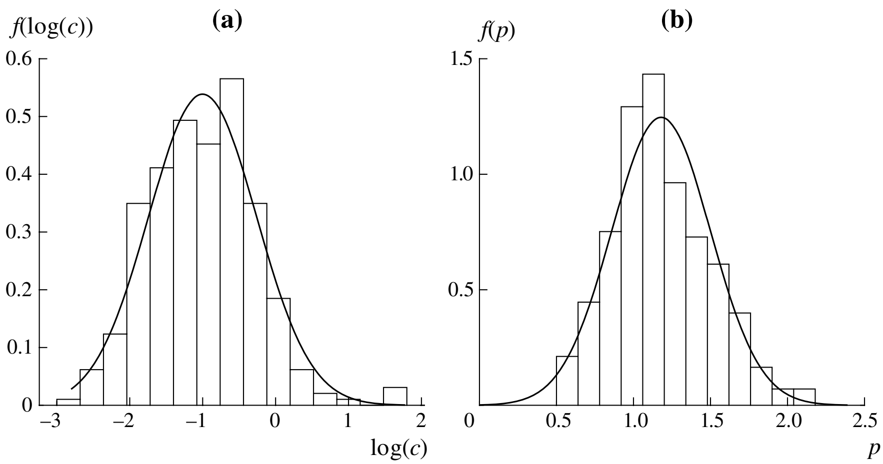
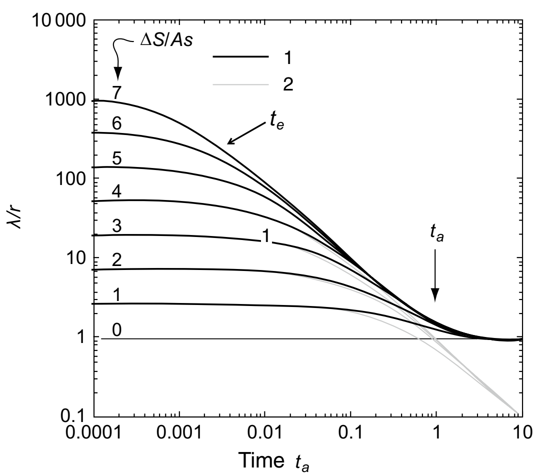


Fig. 3 Empirical distributions of estimates of parameters of Eq. (1) log(c) **(a)** and p **(b)** (Baranov et al. 2019c, Fig. 2). Histograms are based on 334 aftershock series from $M_m \geq 6.5$ world earthquakes taken from ANSS ComCat for 1980–2016. Solid line is a normal approximation with means $E[\log(c)] = -1$ and $E[p] = 1.05$ and standard deviations $\sigma[\log(c)] = 0.74$ and $\sigma[p] = 0.25$

on two identical sandstone specimens that, when failure occurs on a preexisting fault, p increases with increasing axial stresses, c decreases with increasing axial stresses and increases with increasing confining pressure (Smirnov et al. 2019). These authors also showed that the b value in the Gutenberg–Richter relation varies inversely to stress variation. Negative correlation between the values of b and p for intraplate seismicity is also observed in actual data (Wang 1994; Ogata and Guo 1997; Gasperini and Lolli 2006).

The Omori–Utsu law is the basic model for the aftershock process. It describes the decay of aftershock activity due to physical properties of the aftershock processes, including afterslip. The parameters that enter the law can be clearly interpreted in physical terms following the rate-and-state model and depend on material properties. The

Fig. 4 The rate of aftershocks after the stress drop caused by the main shock for cases where the rate of tectonic deformation is $\dot{\tau} \neq 0$ (curve 1) and $\dot{\tau} = 0$ (curve 2) (Dieterich 2007, Fig. 6)



parameters in Eq. (1) for an aftershock sequence can be estimated either by the method of maximum likelihood (Utsu et al. 1995) or by the Bayesian method (Holschneider et al. 2012). The latter approach is advantageous in that it can incorporate prior information on the distribution of the parameters and can obtain posterior estimates of the parameter values.

Due to the latter, surfaces of constant gravitational potential are no longer spherical. Its topography at the Earth's surface is referred to as the geoid. Dynamical core processes contribute to a quasi-steady part of the non-spherical gravity field.

2.2 Exponential Models

Power law decay describes a "long tail" of the aftershock sequences. But sometimes a faster decay is observed (Utsu 1957, Mogi 1962, Watanabe and Curoiso 1970, Otsuka 1985). Various models have been suggested to describe this transition power law and exponential behavior (Mignan 2015). Otsuka (1985) introduced a model which is a product of exponential and Omori–Utsu laws:

$$\lambda(t) = Ke^{-\alpha t}(t + c)^{-p} \tag{2}$$

where t means time after main shock, K, α, c, p are model parameters. Souriau et al. (1982) and Kisslinger (1993) suggested the stretched exponential function, often describing relaxation processes:

$$n(t) = \frac{qN_a}{t\left(\frac{t}{t_0}\right)^q \exp\left[\left(\frac{t}{t_0}\right)^q\right]} \tag{3}$$

where N_a is the number of "potential aftershock sites" at $t=0$ (corresponding to the total number of aftershocks for $t \to \infty$, t_0 (relaxation time) and q are parameters. For $q=1$, the model becomes the classical exponential decay, and for $q=0$ it corresponds to the simple Omori model.

Using three declustering techniques, Mignan (2015) has compared power law, pure exponential, and stretched exponent models. As it turned out, the stretched exponential provides better fit than the power law. He concluded that aftershock occurrence is due to a simple relaxation process, in accordance with most other relaxation processes observed in the nature.

Narteau et al. (2002) used the idea of Scholz (1968) about an explanation of the Omori–Utsu law by superposition of relaxation processes with different times to develop the limited power law (LPL) model:

$$\lambda(t) = \frac{A\left[\gamma(q, \lambda_b t) - \gamma(q, \lambda_a t)\right]}{t^q} \tag{4}$$

where

$$\gamma(\rho, x) = \int_0^x \tau^{\rho-1} \exp(-\tau) d\tau$$

Here, $\lambda(t)$ is the rate of aftershocks at time t following the main shock; $\gamma(\rho,x)$ is the incomplete gamma function; and A, q, λ_a, λ_b are positive model parameters.

Equation (4) is derived analytically (Narteau et al. 2002) based on the following idea. At once after a main shock, the aftershock area is modeled by a finite number of isolated and independent undamaged domains. Each domain is triggered by a local overload that is a local combination of stress and strength. When the stress exceeds the strength, the overload is positive, the excess is removed by fracture, and the domain generates an aftershock. Accordingly, under the conditions of a constant stress, the response to the overload σ_0 involving the triggering of a fracture is a final manifestation of fracture organization on small scales. The dependence on the time can come from the interaction between the growing number of microfractures and/or rates of chemical reaction, which control the growth of the fracture on the atomic scale (Das and Scholz 1981). The aftershocks are distributed over time in conformance with the Markov process with a stationary rate of transition (Scholz 1968).

The parameters of the limited power law (LPL) model, λ_a and λ_b, are characteristic rates of aftershock generation where the linear type of aftershock decay gives way to a power law and then to an exponential type (Narteau et al. 2002). The variable $1/\lambda_b$ is an estimate of c in the Omori–Utsu law (1) and hence has a physical meaning that characterizes the delay in the power law decay of the aftershock process (Shebalin 2004; Narteau et al. 2002).

Using the estimates of q, λ_a, λ_b in the LPL model (4), one can estimate the characteristic times t_e, t_a such that the aftershock decay changes in a linear manner as $0<t<t_e$ and decreases according to the power law $1/t^p$ as $t_a<t<t_e$ and exponentially as $t>t_e$ (Narteau et al. 2002). We note that the existence of such times is in accordance with the rate-and-state model. We thus see that the decay of the aftershock process during the time interval $(0, t_a)$ as described by the LPL model is close to that according to the Omori–Utsu law (1). At late times, as $t>t_e$, the LPL model demonstrates faster (exponential) decay on a way similar to various exponential models described above.

An exponential decay rate is a major ingredient of the aftershock decay of the nuclear tests (Gross, 1996). The quantity t_a for the aftershock sequences due to major explosions can be a few days or even a few hours (Narteau et al. 2002). Narteau et al. (2003) used the example of synthetical aftershock sequences to show that the parameter λ_a (hence the time t_a) carries information on the maturity of a fault zone. The values of λ_a around 0 tell us that the fault zone of interest is smooth and mature. Large values of λ_a (a rapid transition to the exponential decay in aftershock rate) correspond to a fault zone of some roughness.

Large earthquakes generally occur on mature faults; accordingly, the power law decay phase for their aftershock sequences generally lasts until the aftershock activity has reached the seismic background level and thus cannot be distinguished. In connection with this fact and considering that the Omori–Utsu model is in general use, it should preferably be employed for hazard assessment.

We agree with Richards-Dinger et al. (2010) that the principal difference of the aftershock occurrence from other earthquakes is that their rate decays according to the Omori–Utsu law.

2.3 The Epidemic Type Aftershock Sequence (ETAS) Model

The increasing density of seismograph networks and enhanced sensitivity of the recording instruments resulted in accumulation of seismological data and helped detect aftershock processes in which the events were themselves followed by sequences of subsequent shocks. In that case, each earthquake triggers an aftershock sequence of its own depending on its magnitude. When these aftershocks are included in the total sequence, this leads to departures of the total rate of aftershocks from the steady decay described by the Omori–Utsu law (1). This branching process can be described by the epidemic-type aftershock sequence (ETAS) model proposed by the noted Japanese mathematician Ogata (1988). The ETAS model is a superposition of independent sequences that obey the Omori–Utsu law, with a new event "generating" (triggering) a sequence of its own whose rate depends on its magnitude:

$$\lambda(t) = r + \sum_{i,t_i<t} \frac{K_1 10^{\alpha(M_i - M_{\min})}}{(t-t_i+c)^p} \tag{5}$$

where t_i and M_i are the time and magnitude of event i; r (events per unit time) is the rate of the so-called background seismicity; K_1 describes the productivity of an event with $M \geq M_{min}$; α measures the efficiency of an earthquake to trigger aftershocks relative to its magnitude; c, p are the parameters that determine aftershock decay rate according to the Omori–Utsu law (1). The model parameters r, K_1, α, c, p are the same for all events in the catalog. The summation in Eq. (5) is over all aftershocks that have occurred until time t.

After the temporal ETAS model had been published, some spatiotemporal generalizations were developed. The generally accepted form of the space–time ETAS model was suggested by Ogata (1998, 2004):

$$\lambda(t) = r(x,y) + \sum_{i,t_i<t} \frac{10^{\alpha(M_i - M_{\min})}}{(t-t_i+c)^p} f\left(x-x_i, y-y_i; M_i\right) \tag{6}$$

This equation, in comparison with Eq. (5), introduces a dependency of background seismicity r and coefficient K_1 on the spatial coordinates (x, y). Function f in addition may depend on the magnitude of the triggering event M_i. Several variants of the function have been suggested (see (Zhuang et al. 2011) and references therein). Among them, two variants are most common. Function suggested by Ogata (1998) is independent of magnitude:

$$f(x,y) = \frac{K}{\left(x^2+y^2+d^2\right)^q} \tag{7}$$

Its modification (Falcone et al. 2010) introduces a magnitude scaling of the triggering distance d_i of the aftershock zone:

$$f\left(x-x_i, y-y_i; M_i\right) = \frac{K d_i^2}{\left(\left(x-x_i\right)^2+\left(y-y_i\right)^2+d_i^2\right)^q}, d_i = d_0 10^{\alpha\left(M_i - M_{\min}\right)} \tag{8}$$

In various versions of the ETAS model, both the background seismicity and the productivity K (or K_1 in Eq. (5)) are supposed to follow the Gutenberg–Richter law (Gutenberg and Richter 1956) with the same b value. Another commonly used simplification concerns the parameter α in Eqs. (5) and (6). In case $\alpha < b$, the ETAS model is stable and stationary; otherwise, there is a finite probability that the number of events in a unit time interval becomes infinite as t increases to infinity (Helmstetter, Sornette 2002; Zhuang and Ogata 2006; Saichev and Sornette 2007). Accordingly, it is often supposed that $\alpha = b$; this simultaneously contributes to a decrease in the number of model parameters (Felzer et al. 2004; Helmstetter et al. 2005; Falcone et al. 2010; Page et al. 2016).

The ETAS models are widely used to assess the seismic hazard due to aftershocks (Hainzl 2016; Helmstetter and Sornette 2003; Helmstetter et al. 2006; Omi et al. 2014, 2015; Ebrahimian and Jalayer 2017; Harte 2017; Shcherbakov et al. 2019; Trevlopoulos et al. 2020).

3 The Reasenberg and Jones Approach

The approach of Reasenberg and Jones (1989) consists in representing the aftershock process by superposition of a time-dependent model and the Gutenberg–Richter relation (Gutenberg and Richter 1956). One important assumption for this representation consists in the independence of times and magnitudes of aftershocks.

It is important to confirm this assumption due to the fact that the approach of Reasenberg and Jones underlies the methods of aftershock activity prediction (Gerstenberger et al. 2005; Baranov et al. 2019c, Shcherbakov et al. 2018, 2021). At the same time, many studies indicate a variation of the *b* value before an earthquake, see, e.g., (Molchan et al. 1999; Papadopoulos et al. 2006; Nanjo et al. 2012; Rodkin and Tikhonov 2016). Accordingly, it would be natural to expect that the aftershock process returns to the common values. Within a concept of aftershock sequences as direct failure cascades (Marsan and Lengliné 2008; Smirnov et al. 2010; Gu et al. 2013), the *b* value is expected to gradually increase. This behavior was observed in laboratory experiments (Smirnov et al. 2010, 2019; Sobolev et al. 1996), as well as in real and synthetic aftershock sequences (Knopoff et al. 1982; Helmstetter and Sornette 2002; Ogata and Katsura 2014; Rodkin and Tikhonov 2016; Tamaribuchi et al. 2018). However, the effect referred to may be due to catalog incompleteness, which depends both on magnitude and on time (Helmstetter et al. 2006; Hainzl 2016; Shebalin and Baranov 2017; Baranov et al. 2019c). Gulia et al. (2018) in contrary observed a positive step of the *b* value after the main shock and its subsequent decrease. In recent studies (Gulia and Wiemer 2019; Dascher-Cousineau et al. 2019), characteristic *b* value changes before and after large earthquakes are considered to determine in real time whether an earthquake would become a foreshock and be followed by an even larger earthquake. The instability of *b* value in time is also observed in the induced seismicity that occurs due to fluid injection for oil production in Oklahoma (Vorobieva et al. 2020). This observation, however, may reflect an effect of a break of slope of the magnitude–frequency distribution at early stages of the aftershock sequences (Shebalin et al. 2021).

To eliminate the influence of catalog incompleteness after large earthquakes, Baranov and Shebalin (2019) investigated the global distribution of the larger aftershocks in each aftershock sequence. Larger aftershocks are usually not missed even at short times after the main shock (Helmstetter et al. 2006). They used the ANSS ComCat catalog for 1975–2017 to which the algorithm of Molchan and Dmitrieva (1992) was applied to extract 526 sequences after $M \geq 6.5$ earthquakes.

From the hypothesis about the independence of times and magnitudes and thus the temporal constancy of the *b* value, it follows that the times of aftershocks, which are the largest in their sequence, should follow the same distribution as the distribution of aftershocks of arbitrary magnitude in each sequence. If we adopt the concept of aftershock sequences as direct failure cascades with growing *b* value, we can expect the times of the largest aftershocks to be shifted toward smaller values; therefore, the distribution function is biased toward larger values. The authors found no bias of the distribution function for the times of the largest aftershocks relative to the theoretic distribution obtained for the stack of aftershocks in 526 sequences (Fig. 5). Next, they compared distributions of times of k-largest aftershock from each sequence for $1 < k \leq 15$ (Fig. 6). Again, no significant difference of the distributions has been found. Both results were obtained using Smirnov–Kolmogorov test at a significance level 5%.

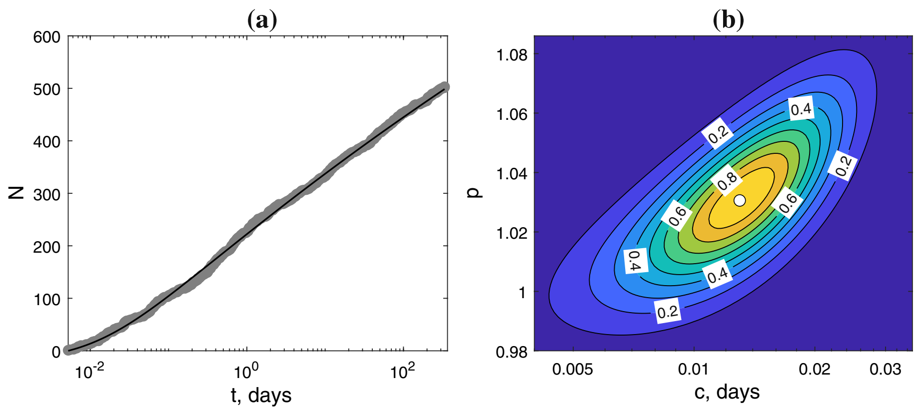


Fig. 5 The distribution of times of largest aftershocks in the 526 sequences (Baranov and Shebalin 2019, Fig. 2). **(a)** The number of sequences with $t_{j,1} < t$ (gray circles) and the Omori–Utsu fit (line); **(b)** posterior probabilities from the joint distribution of estimated parameters c and p in the Omori–Utsu fit (Holschneider et al. 2012). The marked contours are level lines; the white circle marks the likelihood maximum. To avoid distortions due to catalog incompleteness immediately after main shocks, the times in the interval ($t_{start} = 0.005$, $t_{stop} = 365$) days were considered

We note that the issue remains open as to possible variation in the b value in aftershock sequences, since the variation can take place owing to a changed fraction of smaller aftershocks. It seems impossible to test this hypothesis, because it would be hard to distinguish a real deficit in smaller events from that due to catalog incompleteness.

In some individual cases, one can well encounter magnitudes showing some orderliness over time. Since the results above were obtained by a joint analysis of many aftershock sequences in various regions, the independence of magnitudes and times of aftershocks may be considered true in average. This signifies that if in some regions actually magnitudes of aftershocks tend to gradually decrease, then in other regions this should be compensated by a gradual increase in the magnitudes. The dependence of aftershock times and magnitudes under a specific tectonic setting calls for a special study.

4 Earthquake Productivity

As pointed out above, productivity is a key parameter in statistical seismology, and is of critical importance for assessing the hazard posed by aftershocks, because it controls the rate of events within the space–time interval of observation. Viewed from the standpoint of earthquake magnitude distribution due to Gutenberg and Richter (1956), productivity is a constant quantity that can be treated as independent of the scale, the b value, which is the slope of the recurrence curve. Bearing in mind the whole process of earthquakes following each other, productivity is also the total number of events caused by the disturbance in the state of stress due to another, earlier earthquake. This concept of productivity was first used to develop a suitable model of aftershock generation (Utsu 1970). Knowing the productivity and using the Omori–Utsu law (1), one can find the number of subsequent shocks to be expected in a specified time interval from the magnitude of the main shock

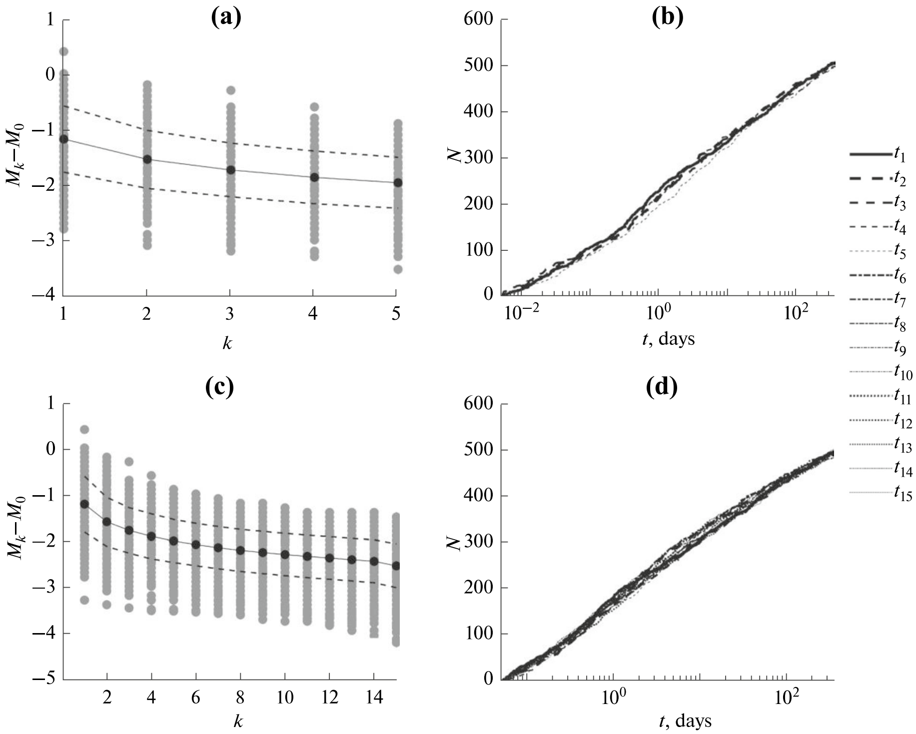


Fig. 6 The distribution of k-largest magnitudes M_k **(a, c)** and the cumulative distribution of the corresponding aftershock times t_k**(b, d)** for $k=1$, K: $K=5$(a, b); $K=15$ **(c, d)** (Baranov and Shebalin 2019, Fig. 4). Index k means the serial number of the aftershock in each sequence, sorted in descending order of magnitude. M_0—magnitude of the main shock. In plots *a* and *c*, gray circles mark relative magnitudes of the k-largest aftershock in individual sequence, black circles and the solid line represent mean values, and dashed lines show the standard deviations

M_m, the slope of decay rate, p, in accordance with a power law function, and the delay in time before the aftershock process would arrive to this decay behavior (Holschneider et al. 2012; Shcherbakov et al 2018; Davidsen et al. 2015; Baranov et al. 2019c).

Since the first epidemic seismicity models appeared (i.e., ETAS), the study of earthquake productivity became the main subject of study, because it is the main parameter that controls the increase in seismicity rate after an earthquake (Kagan and Knopoff 1981; Ogata 1989; Helmstetter and Sornette 2002). All these models assume the rate of events due to an earthquake of magnitude m to vary as a Poisson process with rate

$$N(m) = Ke^{\alpha m} \tag{9}$$

The values of α are between 0.5 and 2.3 (Felzer et al. 2004; Hainzl and Marsan 2008; Hainzl et al. 2013; Marsan and Helmstetter 2017; Wang et al. 2010b; Werner and Sornette 2008; Zhuang et al. 2004, 2005), but they are invariably close to the observed b value. Nevertheless, these estimates remain questionable because it is difficult to identify the relative contribution of successive events in a sequence.

Even though the declustering methods in use are various and diverse, see the reviews of Van Stiphout et al. (2012); Gulia and Wiemer (2019); Pisarenko and Rodkin (2019), the study of cause-and-effect relationships within cascades of triggered seismicity is still in its infancy, no definitive classification has been developed yet. One approach to the issue consists in separating the branching structure of earthquake chains from background seismicity using an iteration algorithm related to maximum likelihood estimates in the epidemic model (Zhuang et al. 2002, 2004). There is another method, consisting of identifying events that were directly and indirectly triggered, assuming a linear contribution of each earthquake into the net level of seismicity without having recourse to any a priori model (Marsan and Lengline, 2008). Finally, there is an alternative approach which seeks to identify earthquake clusters using nearest neighbor method (Zaliapin et al. 2008; Zaliapin and Ben-Zion 2013, 2016) with a proximity function in time–space–magnitude domains (Baiesi and Paczuski 2004).

The nearest neighbor method stands out for its logical simplicity and convenience in calculating productivity. With this method, each seismic event can be a trigger (parent) for several events (offsprings), but each offspring can only have one parent triggering event (Fig. 7). Due to this, when summing (or averaging) the productivity values, there is no problem of multiple counting of the same offspring.

Practically, all declustering methods exploit the fact that productivity depends on the magnitude of the triggering event. Nevertheless, less attention is paid to the general variability in the rate of triggered events (Marsan and Helmstetter 2017) in earthquake catalogs.

Shebalin et al. (2020) examined global and regional earthquake statistics to find that the rate of seismic events triggered by an earlier earthquake obeys the exponential distribution rather than the Poisson distribution as commonly assumed (Fig. 8):

$$F(x) = P(k < x) = 1 - e^{-x/\Lambda_{\Delta M}} \tag{10}$$

with density

$$f(x) = \frac{1}{\Lambda_{\Delta M}} e^{-x/\Lambda_{\Delta M}} \tag{11}$$

Equations (10) and (11) constitute the *law of earthquake productivity*. Here, $\Lambda_{\Delta M}$, called clustering factor, is the mean rate of the events triggered by an earthquake with magnitude $M \geq M_m$-ΔM (M_m is the magnitude of the triggering event). To distinguish the productivity calculated in the ΔM-band of magnitudes from the productivity calculated with the general magnitude threshold, for example, the completeness magnitude, we will call it the ΔM-productivity or simply productivity when it is clear from the context.

Each hierarchical clustering tree was based on the main triggering event. At lower hierarchical levels, the triggered earthquakes are themselves secondary triggering events, and so forth as far as the last branches of the cluster. The inset in Fig. 8 shows that the productivity distribution is invariant with respect to the hierarchical level of the triggering event. A similar behavior is observed for triggering events at all hierarchical levels, and the exponential function seems to control the productivity of all $M \geq 6.5$ earthquakes. The slope of the graphs is equal to the clustering factor in Eq. (10). Almost equal slopes indicate that the clustering factor remains constant from primary to secondary triggered events. This important property allows productivity to be summed or averaged regardless of the hierarchy level in the clustering tree.

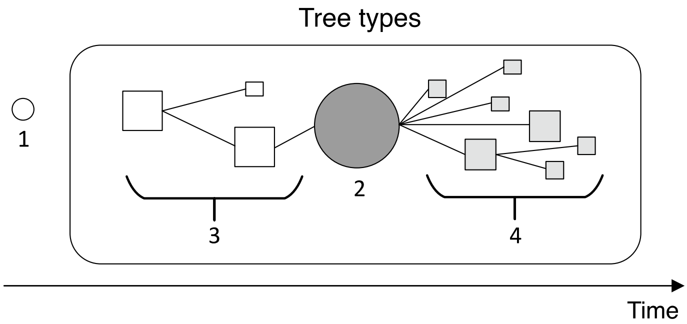
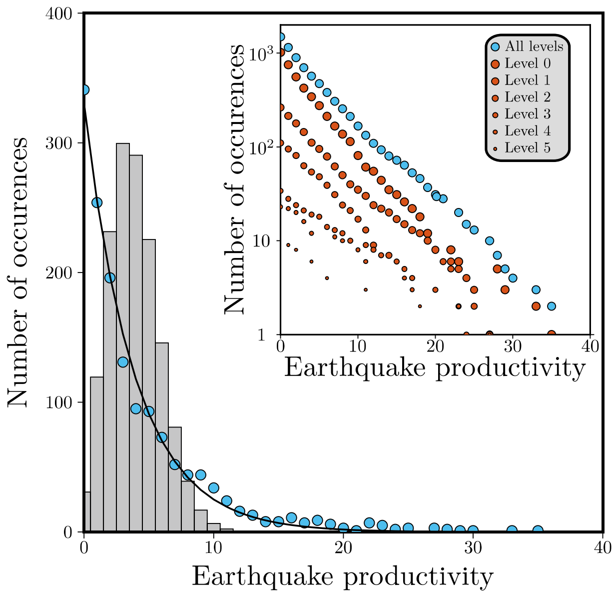


Fig. 7 A scheme of classification of events in the nearest neighbor method (Zaliapin and Ben-Zion 2013, Fig. 6). (1) Single event; (2) main shock; (3) foreshocks; (4) aftershocks. The aftershocks whose parent is the main shock are referred to as primary aftershocks; secondary aftershocks are produced by another aftershock as parent

Fig. 8 Earthquake productivity in the worldwide ANSS Comprehensive Catalog (Shebalin et al. 2020, Fig. 1). Dots show the distribution of the number of triggered events for $M_m \geq 6.5$ earthquakes using a relative magnitude threshold $\Delta M = 2$. The solid line is the exponential law with parameter Λ_2, the mean number of triggered events derived from the data. The histogram shows the Poisson distribution with parameter Λ_2. Inset shows the cumulative productivity distributions for primary and secondary triggering events



The productivity distribution remains an exponential one, when the minimum relative magnitude ΔM is increased from 1 to 2.6 (Fig. 9a). The means of $\Lambda_{\Delta M}$ decrease in accordance with the b value in the distribution of earthquake magnitude (Fig. 9b) and show no dependence on the magnitude of triggering earthquakes M_m (Fig. 9c, d). As well, keeping the minimum relative cutoff value constant, $\Delta M = 2$, we obtained the result that the distribution of triggered events and its mean Λ_2 are practically identical, whatever the triggering event magnitude M_m(Fig. 9c, d). It follows that the exponential distribution of productivity can be treated as a general property of all earthquakes, whatever the size.

Exponential shape of the ΔM-productivity was confirmed for lower magnitudes of triggering earthquakes using seven regional earthquake catalogs (Fig. 10): Baikal, Italy, Japan, Kamchatka and the Kurile, northern California, southern California, and New Zealand, while the clustering factor Λ_2 was found to vary from one region to another (Shebalin

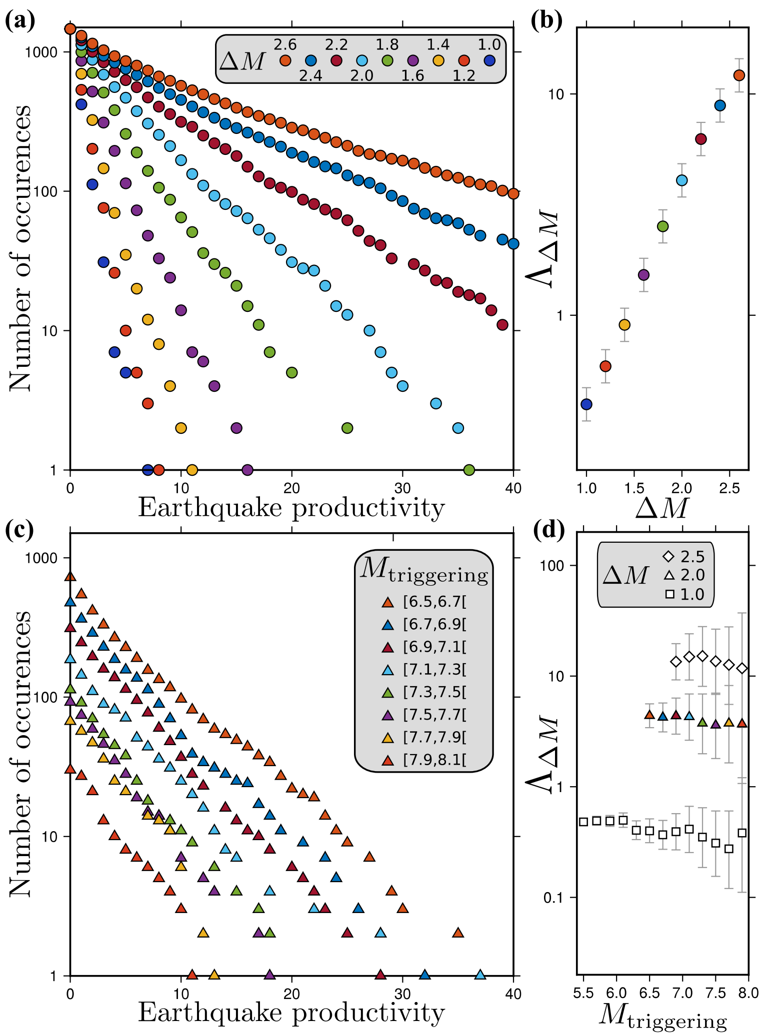


Fig. 9 Dependence of earthquake productivity on magnitude ranges in the worldwide catalog ANSS (Shebalin et al. 2020, Fig. 2). **(a)** Distribution of the number of triggered events of $M \geq 6.5$ earthquakes using a relative magnitude threshold $\Delta M \in \{1, 1.2, \ldots, 2.6\}$. **(b)** The mean number of events $\Lambda_{\Delta M}$ with respect to the relative magnitude threshold ΔM. **(c)** Distribution of the number of triggered events with respect to the magnitude Mtriggering of the triggering event using a relative magnitude threshold $\Delta M = 2$. We take $M_{\min} \leq M_{\text{triggering}} < M_{\max}$ with $M_{\max} = M_{\min} + 0.2$ and $M_{\max} \in \{6.5, 6.7, \ldots, 8.1\}$. **(d)** The average number $\Lambda_{\Delta M}$ of triggered events with respect to the magnitude of the triggering event for $\Delta M \in \{1, 2, 2.5\}$

et al., 2020). The similar properties of the ΔM-productivity were also observed for low-magnitude mining-induced seismicity (the lower cutoff magnitude $M_c \geq 0$) in the Khibiny Massif (Baranov et al. 2020).

Physical reasons of the variation of the earthquake productivity remain a subject of great interest. It was assumed that the high background stress characterizing the thrust faulting style is associated with increased productivity (Tahir et al. 2012; Tahir and Grasso 2015). Zaliapin and Ben-Zion (2019) compared the productivity and other characteristics

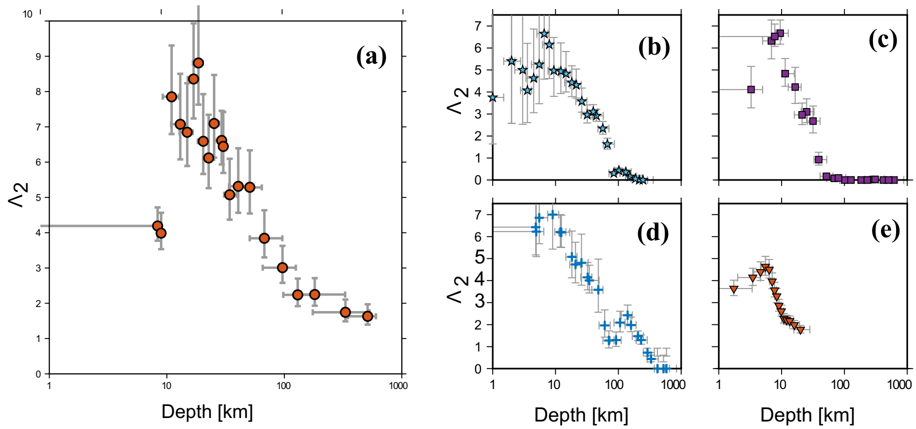


Fig. 10 Earthquake productivity (clustering factor) versus depth (Shebalin et al. 2020, Fig. 3, Fig. 4). Mean rate of triggered earthquakes with minimum relative magnitude $\Delta M = 2$ or greater plotted against the depth of triggering events. **(a)** Worldwide data, ANSS ComCat catalog, $M \geq 6.5$, **(b)** Kamchatka, $M \geq 5$, **(c)** California, $M \geq 3.3$, **(d)** New Zealand, $M \geq 4.5$, **(e)** Italy, $M \geq 3.0$. The horizontal bars represent the depth ranges around the mean, while the vertical bars show the standard deviation Λ_2

of the earthquake clusters with the heat flow level and type of deformation. They found that the dominant type of seismicity clusters depends on the heat flow and much less on the type and intensity of deformation. This confirmed earlier results obtained for California and showing positive correlation of the productivity and heat flow (Yang and Ben-Zion 2009). Zaliapin and Ben-Zion (2019) suggested two dominant types of global clustering: burst-like and swarm-like clusters. The former are characterized by a brittle type of fracturing in a relatively cold lithosphere (for example, surface earthquakes in subduction zones), and the latter are characterized by brittle–ductile deformation in hotter areas (for example, mid-ocean ridges). Those results are in accordance with the observation that oceanic transform faults are characterized by relatively weak aftershock sequences (Boettcher and Jordan 2004).

Hainzl et al. (2019) using a detailed earthquake catalog for the subduction zone of northern Chile, 2007–2014, found that the earthquake productivity decays systematically with depth. Similar result for worldwide catalog and for regional data in Japan, Kamchatka, New Zealand, and Italy (Fig. 10) was obtained by Shebalin et al. (2020). Hainzl et al. (2019) compared the analysis of the earthquake catalog with three different coupling maps inferred from interseismic geodetic deformation data. They show that the observed depth dependence can be explained by a linear relationship between the productivity and the coupling coefficient.

5 Productivity and the ETAS Model

As mentioned, the ETAS model is based on the hypothesis that the rate of generated aftershocks is constant for a given magnitude. The Poisson distribution is a natural means thereby to model deviations from the mean in specific samples. As a matter of fact, however, the rate of generated events follows an exponential distribution, hence is below the mean in the bulk of the cases. It is for this reason that the ETAS model can return overestimated aftershock rates when used for forecasting (Baranov et al. 2019b).

On the other hand, it is of utmost importance to see that the observed exponential distribution for the ΔM-productivity is not a consequence of the hierarchical declustering procedure based on the nearest neighbor method (Zaliapin and Ben-Zion 2013). With this end in view, Shebalin et al. (2020) applied the declustering procedure to catalogs that have been synthesized following the space–time ETAS model (Zhuang et al. 2002; Helmstetter and Sornette 2002; Felzer et al. 2004). The model consists of background events that obey a homogeneous stationary Poisson process of rate μ. Each earthquake in the catalog triggers the first generation; these events trigger sequences of their own in turn, and so on. The magnitudes are assumed to be independent and to obey the Gutenberg–Richter distribution with a constant b value, while the times of the events obey the Omori–Utsu law (1) with $p > 1$. The total seismic rate consisting of the background and triggered events of all generation is given by the rate (Ogata 1998):

$$\lambda(t,x,y,z) = \mu + \sum_{i:t_i<t} \frac{K}{(t-t_i+c)^p} \times \frac{(p-1)c^{p-1}10^{-b\Delta M}\Lambda_{\Delta M}}{\left(\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2+d}\right)^q} \tag{12}$$

where M_0 is a magnitude threshold and t_i, x_i, y_i, z_i, and m_i are time, coordinates, and magnitude of the ith event. The temporal component corresponds to the Omori–Utsu law (1). Its spatial counterpart follows a similar power law distribution.

The model (12) is specified by eight scalar parameters $\{\mu, b, K, c, p, \alpha, d, q\}$. The parameter values were chosen with respect to the properties of seismicity in the global catalog: $M_0 = M_c = 4.5$, $b = 1.15$, $c = 0.0356$ day and $p = 1.08$, $\Delta M = 2$, $\Lambda_2 = 4.32$. To obtain approximately the same spatial and temporal density of earthquakes as in the global catalog, it was chosen $\mu = 10$ events per day in an area of 5000×5000 km^2. The catalog was simulated with total duration of 20,000 days for different q values in a range of 1.5 to 3 and different d values in a range of 1 to 30.

For all versions of the synthetic catalog, distributions of the ΔM-productivity are in good agreement with Poisson distributions with modal values that are significantly different from zero. These modal values are close to the $\Lambda_{\Delta M}$ value. Figure 11a shows an example of the productivity distribution for $\Delta M = 2$, $q = 2$ and $d = 1$ km and compared it to the Poisson and exponential distributions with parameter Λ_2.

To incorporate earthquake productivity law into the ETAS model, Shebalin et al. (2020) modified a single ingredient of the spatiotemporal ETAS model (12). The productivity $\Lambda_{\Delta M}$ is the mean value of a random variable following an exponential distribution. When it comes to generating the synthetic catalog, this new ingredient takes the form of an additional random draw associated with each event. This modified ETAS model was called the ETAS$^{(e)}$ model. It does not have an explicit form of the conditional intensity function, but one can still numerically simulate this process. The declustering procedure was applied to synthetic catalogs produced by the ETAS$^{(e)}$ model; the ΔM-productivities are in good agreement with the exponential distribution and have maximum values at zero (Fig. 11b).

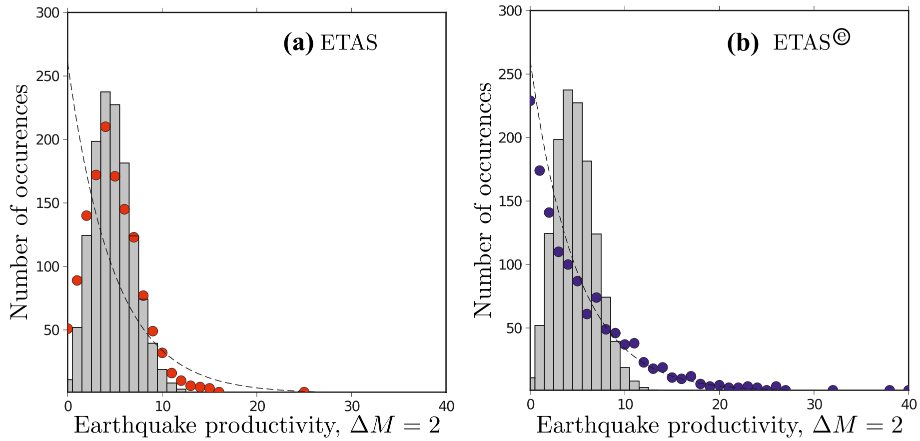


Fig. 11 Distribution of the earthquake ΔM-productivity in synthetic catalogs using **(a)** ETAS and **(b)** ETAS$^{(e)}$ models (Shebalin et al. 2020, Fig. S8). Dots show the distribution of the number of triggered events for $M \geq 6.5$ earthquakes using a relative magnitude threshold $\Delta M = 2$. The dashed line is the exponential law with parameter Λ_2, the mean number of triggered events derived from the data. The histogram shows the Poisson distribution with parameter Λ_2

Testing with synthetic catalogs demonstrates that hierarchical declustering procedure is able to recover both the predefined Poisson and exponential distributions of the productivity in synthetic catalogs produced by epidemic models of seismicity. Hence, the hypothesis that the exponential behavior observed in real catalogs is an artifact of the declustering procedure should be rejected.

6 Earthquake Productivity and the Båth Law

Seismology has the well-known Båth law (Båth 1965) stating that the difference between the magnitude of the main shock (M_m) and that of the largest aftershock (M_1) is independent of main shock magnitude, being approximately equal to 1. At the same time, the bulk of larger aftershocks generally occurs during the few first hours. In practice, this activity is naturally perceived as a direct continuation of the main earthquake. The later larger aftershocks occur upon the background of less frequent shocks, are less expected, and for this reason pose a hazard of their own.

The modeling of the distribution for the difference in magnitude between the main shock and the largest aftershock for $t = 0$ with a view to a theoretical validation of the Båth law was the subject of extensive research (Vere-Jones 1969, 2008; Lombardi 2002; Console et al. 2003; Helmstetter and Sornette 2003; Baranov and Shebalin 2018; Baranov et al. 2019c, Shebalin, Baranov 2021). Early work focused mainly on confirming a mean of about 1.2 regardless of the shape of the distribution, although the distribution was known to be quite large, with a standard deviation of about one unit of magnitude.

Vere-Jones (1969), based on the Gutenberg–Richter law and the assumption that the magnitudes of different events are independent, obtained a mean value equal to $\sim \frac{1}{\beta} = \frac{1}{ln10b}$. This gives about 0.5 which is significantly different from usually observed value 1.2. In addition, the author found that M_m-M_1 depends on M_m. He suggested that the discrepancies with Båth law arise from hidden bias from the use of different lower thresholds for M_m and M_1.

Following mainly the ideas of Vere-Jones (1969), Lombardi et al. (2002) and Console et al. (2003) noted that the quantities M_m and M_1 are random variables with different distributions, not because the main shocks are of a different nature than aftershocks, but because the ordinal statistics are not independent and equally distributed random variables, in contrast to a sample of random variables. They also concluded that the choice of thresholds for M_m and M_1 is crucial for the distribution of M_m-M_1.

Drawing on the work of Vere-Jones (1969) and Console et al, (2003), Helmstetter and Sornette (2003) showed that the origin of Båth law should be sought in the selection procedure used to determine the main shocks and aftershocks, and not in any differences in the mechanisms controlling the main shock and aftershocks. Using the ETAS model, the authors showed that this model agrees well with Båth law in a certain range of parameters (Fig. 12).

It has also been shown that the standard interpretation of Båth law is incorrect in terms of the two strongest events from a self-similar set of independent events. The authors believe that mean difference M_m-M_1 is determined not only by the distribution of magnitudes, but also by the productivity of aftershocks. The large difference in magnitudes can be explained by the low productivity of aftershocks.

Vere-Jones (2008) proved a limit theorem that establishes conditions under which the distribution of the difference M_m-M_1 approaches a limit form independent of M_m. He assumed that the magnitudes of individual events obey an exponential distribution and that the structure of the sequence approaches the structure of a Poisson process. It is shown that in these cases the form of the limiting distribution is a double exponent. If additionally, to take into account that the expected total number of aftershocks increases exponentially with increasing M_m, then this gives both an explanation of Båth law and its features in a wide range of conditions. We note that the research of Vere-Jones (2008) gives a good explanation of the mean observed difference M_m-M_1, but gives significantly asymmetric distribution, while empiric distribution is rather symmetric.

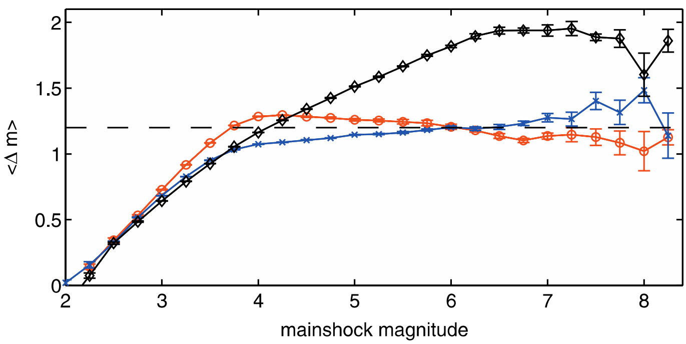


Fig. 12 Average magnitude difference $<\Delta \mathrm{m}> = <M_m\text{-}M_1>$ between a main shock and its largest aftershock (Helmstetter and Sornette 2003, Fig. 2), for numerical simulations of the ETAS model with parameters $b=1$, $c=0.001$ day, $p=1.2$, a minimum magnitude $m_0=2$, a maximum magnitude $m_{max}=8.5$ and a constant loading $m=300$ events per day. Each curve corresponds to a different value of the ETAS parameters: a $=0.8$ and $n=0.76$ (crosses), $\alpha=0.5$ and $n=0.8$ (diamonds) and $\alpha=1$ and $n=0.6$ (circles). The error bars give the uncertainty of $<\Delta \mathrm{m}>$ (1 standard deviation). The horizontal dashed line is the empirical value $<\Delta \mathrm{m}> = 1.2$

Baranov and Shebalin (2018) observed that the average difference between the magnitudes of the main shock and the largest aftershock, which occurs in the interval beginning sometime after the main shock, gradually increases. They used worldwide data to show that the distribution of M_m-M_1 can be fitted with the normal distribution. As time goes on, the distribution is displaced toward lower magnitudes while retaining its shape and width. The authors called this feature the dynamic Båth law. The amount of the displacement depends on the b value in the Gutenberg–Richter law, on the parameters c, p in the Omori–Utsu law, and the time span of interest.

A little later, Baranov et al. (2019c) found a theoretical explanation of the Gaussian-like form of the Båth law distribution taking into account the exponential earthquake productivity law. Combining for aftershocks the Omori–Utsu law (1) for time distribution, Gutenberg–Richter law (Gutenberg and Richter 1956) for magnitude distribution (times and magnitudes are assumed to be independent, see Sect. 3), the exponential earthquake productivity law (see Sect. 4), and finally the Poisson distribution for the actual realization of the number of aftershocks at a given rate, they came to a logistic distribution of the M_1-M_m quantity. The shape of the logistic distribution is very close to Gaussian distribution.

The Gutenberg–Richter law (Gutenberg and Richter 1956) in the form of a probability distribution is written as follows:

$$F(M) = 1 - e^{-\beta\left(M - M_c\right)} = 1 - 10^{-b\left(M - M_c\right)}, M \geq M_c \tag{13}$$

where $\beta/\ln 10 = b$ is the slope of the recurrence curve for the events in the sequence; M_c—is the lower cutoff magnitude.

Equation (13) determines the probability that an arbitrary aftershock has a magnitude less than M. Then, the probability that N aftershocks have magnitudes less than M is equal to $F(M)^N$. Accordingly, under assumption that N obeys the Poisson distribution with rate Λ and following the formula of complete probability, we can obtain that the probability distribution for magnitude M_1 in an individual aftershock sequence is a double exponent (Vere-Jones 2008; Zöller et al. 2013):

$$P\left(M_1 < M\right) = F_s(M) = \sum_{N=0}^{\infty} \frac{\Lambda^N}{N!} e^{-\Lambda} F(M)^N = e^{-\Lambda[1-F(M)]} = e^{-\Lambda e^{-\beta\left(M-M_c\right)}} = e^{-\Lambda 10^{-b\left(M-M_c\right)}} \tag{14}$$

In order to pass to the Båth law and to derive the distribution of M_1 based on all sequences, we need to recall that Λ obeys the exponential distribution, in accordance with the earthquake productivity law Eqs. (10) and (11). Taking account Eqs. (10) and (14), Baranov et al. (2019c) found that the distribution of M_1 based on all sequences has the form

$$G(M) = \int_0^{\infty} F_s(M) f_{\text{ex}}(\Lambda) d\Lambda = \frac{1}{\Lambda_{\Delta M}} \int_0^{\infty} \exp\left[-\Lambda\left(10^{-b\left(M-M_c\right)+\frac{1}{\Lambda_{\Delta M}}}\right)\right] d\Lambda.$$

On integration they found (here, as in Sect. 4, we use positive value ΔM):

$$G\left(M_1 - M_m < m\right) = \frac{1}{1+\Lambda_{\Delta M} 10^{-b(m+\Delta M)}}. \tag{15}$$

The distribution (15) is known as logistic distribution. Baranov et al. (2019c) found that this distribution also held for an arbitrary time interval (t, T). In that case, the quantity $\Lambda_{\Delta M}$ in Eq. (15) must be replaced with the time-dependent $\Lambda_{\Delta M}$ (t,T). The dependence of the value of $\Lambda_{\Delta M}$ (t,T) on time is given by the Omori–Utsu law (1):

$$\Lambda_{\Delta M}(t,T) = \Lambda_{\Delta M}\frac{\int_t^T (t+c)^{-p}dt}{\int_0^T (t+c)^{-p}dt} = \Lambda_{\Delta M}\frac{D(0,T;c,p)}{D(t,T;c,p)} \tag{16}$$

where c, p are the parameters in the Omori–Utsu law; the function $D(t_1,t_2;c,p)$ is specified by the relations

$$D\left(t_1, t_2;c,p,\right) = \begin{cases} \frac{\frac{1-p}{c}}{(1+t_2/c)^{1-p}-(1+t_1/c)^{1-p}}, \ p \neq 1 \\ \frac{\frac{1}{c}}{\ln\left(1+\frac{t_2}{c}\right)-\ln\left(1+\frac{t_1}{c}\right)}, \ p = 1 \end{cases} \tag{17}$$

Thus, the quantity $M_1(t,T)$-M_m in the dynamic Båth law has a logistic-type distribution with a shape similar to the Gaussian distribution:

$$P\left(M_1(t,T) - M_m < m\right) = F_{db}(m;t,T) = \frac{1}{1+\Lambda_{\Delta M}(t,T)10^{-b(m+\Delta M)}} \tag{18}$$

and its density is

$$f_{db}(m,t) = \frac{\Lambda_{\Delta M}(t,T)b\ln(10)10^{-b(m+\Delta M)}}{\left[1+\Lambda_{\Delta M}(t,T)10^{-b(m+\Delta M)}\right]^2} \tag{19}$$

The mean, the median, and the mode are identical and have the form

$$\mathrm{E}_{db}\left[M_1(t,T) - M_m\right] = Mode\left[M_1(t,T) - M_m\right] = -\Delta M + \frac{1}{b}lg\left(\Lambda_{\Delta M}\right) + \frac{1}{b}lg\left(\frac{D(t,T;c,p)}{D(0,T;c,p)}\right) \tag{20}$$

The variance is

$$Var\left[M_1(t,T) - M_m\right] = \frac{\pi^2}{3\beta^2} = \frac{\pi^2}{3b^2\ln^2 10} \tag{21}$$

The quantity $-\Delta M+\lg(\Lambda_{\Delta M})/b$ in Eq. (20) specifies the mean difference in magnitude between the largest aftershock and the main shock $\mathrm{E}[M_1(0,T)$-$M_m]$. According to the empirical Båth law, this value is equal to about −1.2. The third term in Eq. (20) determines the temporal decay of this mean value.

To test Eqs. (18) and (19) on real data, Baranov et al. (2019c) and Shebalin and Baranov (2021) used the technique of aftershock stacking (Shebalin and Narteau 2017). All aftershock sequences are stacked together, and the magnitudes M in each sequence are replaced with differences M-M_m. Times are calculated relative to the time of corresponding main shock. This data set is then rearranged in increasing time.

The estimate of the b value in the Gutenberg–Richter law for relative magnitudes has the following advantage. It is known that completeness magnitude for early aftershocks depends on the magnitude of the main shock, while the relative completeness magnitude does not (Helmstetter et al. 2006; Hainzl 2016; Shebalin and Baranov 2017).

Correct estimation of b value from an earthquake catalog is a critical problem of statistical seismology. In practice, the magnitudes are binned in δM intervals, typically $\delta M = 0.1$. The binned magnitudes led to biased estimate of b value. Marzocchi et al. (2003, 2020) reviewed many methods for estimating b value, studied some potential source of bias, and also provided some recipes to minimize the impact of these potential sources of bias. In the spirit of this work, Baranov and Shebalin (2018) estimated the b value for the stack of 777 sequences with main shocks of magnitude $M \geq 6.5$ identified in the ANSS ComCat catalog, 1980–2016, using the method of Molchan and Dmitrieva (1992) (Fig. 13). They estimated $b = 1.0$ using the procedure described by Bender (1983) in the interval [−2, −0.5]. Limiting the estimate on the right at −0.5 was introduced to exclude a possible effect of finite volumes (Romanowicz 1992) or break of slope in magnitude–frequency distribution caused by postseismic creep (Shebalin, Baranov 2017; Shebalin et al., 2021). Both effects are expressed in a deficit of larger events, which may lead to significant overestimation of the b value (Marzocchi et al. 2020) when the completeness magnitude is 2 units below the maximum magnitude. The parameters c and p in the Omori–Utsu law (1) were estimated (Fig. 14) by the Bayesian method (Holschneider et al. 2012) for the interval (0.01, 365) days using uniform prior distributions for c in the interval [0.005*,* 0.02] and for p in [0.5, 1.5].

Parameter $\Lambda_{\Delta M}$ was estimated as the mean rate of aftershocks in the sequences during the time (0, 365) days after the main shock. The list of estimated parameters for the worldwide catalog and three regional catalogs is given in Table 1 (Shebalin, Baranov, 2021).

Those estimates were used to calculate using Eqs. (20) and (21) the means and the standard deviations for the $M_1(t{,}365)$-M_m quantity at a set of times t and compare them with the observed values (Fig. 15). For all four considered catalogs, the dynamic Båth law is in good agreement with observations.

The above combination of the Gutenberg–Richter law and earthquake productivity enables one to explain the shape of the associated distribution in addition to validating the empirical Båth law. Indeed, incorporation of the decay for the aftershock process using the Omori–Utsu law has enabled the dynamic Båth law to be derived, this law being an extension of the Båth law to include the time factor.

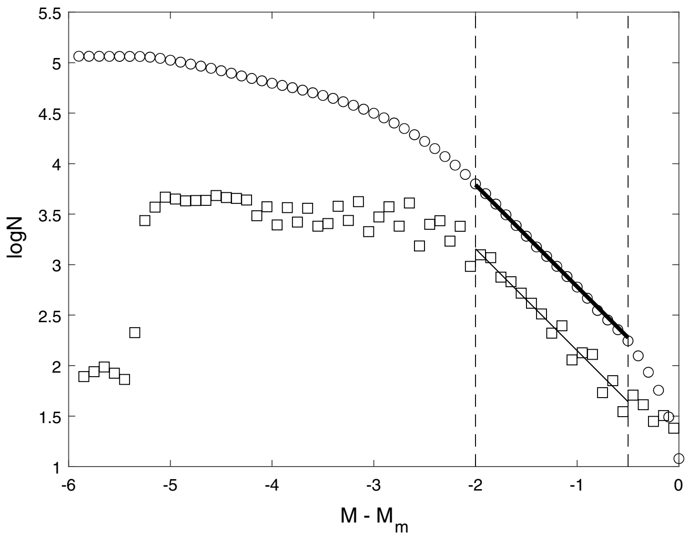


Fig. 13 Estimating Gutenberg–Richter b value for the stack of 777 aftershock series (Baranov and Shebalin 2018, Fig. 2). Cumulative (circles, thick line) and differential graphs (squares, thin line) of frequency–magnitude relationship for $M - M_m$. Dashed lines mark the interval (− 2, − 0.5) for estimating b value. Estimated b value is 1.0

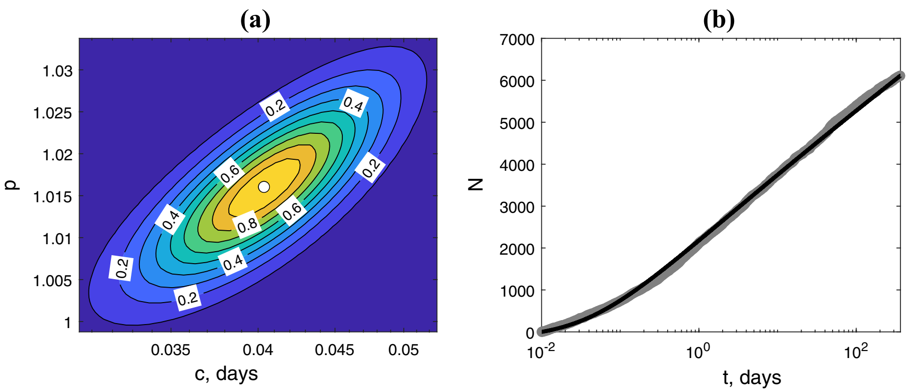


Fig. 14 Estimating parameters c and p of Omori–Utsu law (1) for the stack of 777 aftershock series (Baranov and Shebalin 2018, Fig. 3): **(a)** a posteriori distribution of Bayesian estimates of c and p. Contours with markers show lines of levels of quantiles; white circle marks position of maximum likelihood; **(b)** distribution of times of aftershocks. Thick gray line shows empirical distribution over the stack; thin black line shows distribution by Omori–Utsu law with estimated parameter values $c = 0.04$ day and $p = 1.016$

7 Predicting the Magnitude of the Largest Aftershock and the Hazardous Period of Aftershock Activity

Estimating the future activity of aftershocks after a main shock has occurred is an important practical problem facing statistical seismology, because the larger aftershocks may constitute severe threat for infrastructure weakened by the impact of the main shock. Here, we touch upon two most important aspects of forecasting aftershocks: what maximum magnitude should be expected and how long the danger of destructive aftershocks will persist.

The forecasting aftershocks is based on modeling their activity. Although different models can be used, the most common are the ETAS model and the Reasenberg–Jones (1989) approach discussed above. When using models, the question arises about the estimation of the parameters. Verification methods are needed to test how well the models predict the activity of aftershocks. Below, before proceeding to the description of the predictive methods used, we consider separately these two most important issues.

If the Reasenberg–Jones approach is applied, the forecasts can be derived directly from the model. In the case of ETAS model, it is necessary to simulate a set of forecasts and calculate the required probabilities as frequencies of occurrence.

Table 1 The estimated parameters to test the dynamic Båth law (18)

Catalog	Period	ΔM	b	c	p	$\Lambda_{\Delta M}$
ANSS ComCat	1975–2017	2	1.0	0.04	1.016	7.5
Kamchatka	1969–2017	2	1.2	0.08	0.9	16.3
Baikal	1960–2014	2.5	0.84	0.02	0.87	7.1
Caucasus	1962–2015	2	0.81	0.123	1.014	5.2

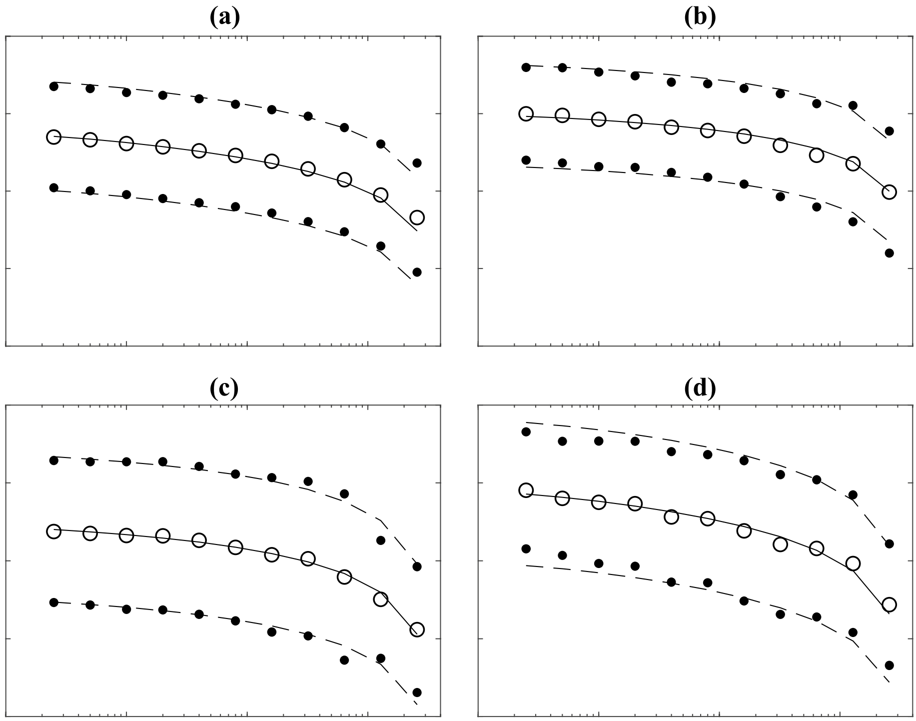


Fig. 15 Comparing observed and estimated using Eqs. (20) and (21) means (circles and solid lines) and standard deviations (dots and dashed lines) for the magnitude difference between the largest aftershock and the main shock $M_1(t,365)$—M_m for the worldwide ANSS ComCat catalog **(a)**, Kamchatka **(b)**, Baikal **(c)**, and the Caucasus **(d)** catalogs with the parameter values from Table 1

7.1 Estimating Model Parameters

Basic mathematical concepts of the ETAS model are summarized in Zhouang et al. (2012). The ETAS parameters are estimated by the method of maximum likelihood (Ogata et al. 1993). This method gives point estimates for the parameters ignoring the inherent uncertainty that arises from historical earthquake catalogs (Ross 2021) or incomplete detection after large earthquake when many aftershocks are skipped during processing due to overlapping seismic waves.

Another way to estimate the parameters is Bayesian approach which is more preferred for incomplete data. Omi et al. (2014) suggested a method based on Bayesian approach for estimating the ETAS model from data of incompletely detected early aftershocks. The idea is to jointly estimate the b value of Gutenberg–Richter distribution and the time-varying detection rate function, which describes the lack of aftershocks, using Bayesian procedure applied, e.g., to data for 24 h after the main shock. Then, these values were used as a prior information for estimating the α value of the ETAS (5). Later, the same authors (Omi et al. 2015) tested this method using data of 38 aftershock series selected by window procedure from JMA catalog. The ETAS parameters were estimated from 1-day data after the main shock and an expected aftershock number for 1–30 day was estimated using the so-called scenario approach.

Strictly speaking, an empirical procedure by Omi et al. (2014, 2015) is not completely Bayesian. Applying full Bayesian statistics to the ETAS model is difficult because of complex nature of the resulting posterior distribution, which makes it infeasible to apply to catalogs containing more than a few hundred earthquakes. Ross (2021) developed a procedure to estimate parameters of the space–time ETAS model in fully Bayesian way, which can be efficiently scaled up to large catalogs containing thousands of earthquakes and provided an appropriate software.

The "scenario" approach is commonly employed to predict aftershock activity following large earthquakes (Kagan and Jackson 2000). The parameters are estimated from data for a period of time (it may the initial part of the aftershock sequence of a large earthquake) in an area (Omi et al. 2014, 2015; Ebrahimian and Jalayer 2017). Afterward, the model is simulated multiple times with these parameter values, and the distribution of the desired quantity (as an example, the last aftershock of a specified magnitude) is estimated from the frequency observed in these model samples. However, this gives rise to several problems that call for extra conditions and restrictions to be introduced in the model that are frequently little justified. The problems that are the most important for the present study are the following.

First, as follows from Eq. (5), the number of primary aftershocks (that is, those which are not aftershocks of aftershocks) of any event only depends on the difference in magnitudes between that event and the minimum magnitude used, while the parameters K_1 and α are assumed to be constants. This may make the estimates based on the ETAS model unreliable (Marsan and Helmstetter 2017). It is also known that the number of aftershocks can vary within a few orders (Marsan and Helmstetter 2017); indeed, according to the law of earthquake productivity, this quantity obeys an exponential distribution whose mode is at zero, when the difference between the main shock magnitude and the minimum magnitude used is fixed (Shebalin et al. 2020). At the same time, the ETAS model assumes the aftershock rate to obey the Poisson distribution. That assumption and the nonlinear form of Eq. (5) leads to overestimation of expected rates in aftershock sequences even if the condition $\alpha < b$ is true (Baranov et al. 2019b).

The Reasenberg–Jones (1989) approach assumes that the magnitudes and times of aftershocks are independent, and the model is a direct product of the Gutenberg–Richter law and the Omori–Utsu model. The total set of parameters is: b value of the Gutenberg–Richter law (13), parameters c, p and K of the Omori–Utsu law (1). In recent studies, the most common way to estimate parameter values is based on likelihood: either a point maximum likelihood, or an interval Bayesian estimate. The Bayesian method additionally makes it possible to regularize estimates.

A standard way to apply the Bayesian approach consists in simultaneous optimization of the whole set of parameters and constructing posterior distribution for the target value. For example, Shcherbakov et al. (2018) obtained Bayesian estimates of the parameters by multiple integration of Bayes' formula to derive a posterior distribution for the magnitude of the largest aftershock $M_1(t,T)$. Usually, a "flat" prior (uniform distribution in a given range of values) is applied to each parameter.

At the same time, retrospective testing based on worldwide data showed (Baranov et al. 2019c) that some aftershock sequences frequently displaced the maximum of the posterior distribution for the parameter so that it was near the boundary of specified prior limits, and this ultimately led to erroneous estimates of $M_1(t,T)$. Regularization of estimates requires a more detailed prior information on the parameters. Thus, it is technically more convenient to separate the estimate of the b value and the parameters of the Omori law.

Correct estimating *b* value from an earthquake catalog is a critical problem of statistical seismology. In practice, the magnitudes are binned in δM intervals, typically $\delta M = 0.1$. The binned magnitudes led to biased estimate of *b* value. Marzocchi et al. (2003, 2020) reviewed many methods for estimating *b* value, studied some potential source of bias, and also provided some recipes to minimize the impact of these potential sources of bias. The Bayesian procedure of Bender (1983) allows, among other advantages, to estimate the b value for a two-sided limited magnitude range. A "cookbook" on the application of this procedure can be found in (Vorobieva et al. 2013).

Bender's standard procedure assumes a flat prior. Baranov et al. (2019c) modified this procedure using Gaussian prior for the *b* value. They have considered 334 aftershock sequences and estimated the *b* value using Bender's procedure with a uniform prior distribution in [0.5, 1.5] (Fig. 16). The distribution of the estimates then was fitted using a normal function, which was then used as a prior for other b value estimates.

Correct estimation of the *b* value depends on a correct choice of the completeness magnitude M_c. This problem is well reflected in the literature; there are many methods for assessing M_c, for example, the method of Woessner and Wiemer (2005). For aftershock analysis, however, this problem has an essential feature: M_c depends on time from the main shock. We discuss this in more detail below.

When estimating the *b* value in an aftershock sequence, the result can significantly depend on the declustering method and the parameters of the declustering procedure. Using the catalog of earthquakes in California since 1980, Mizrahi et al (2021) showed that the *b* value after declustering procedure (aftershocks removed) decreases by up to 30% due to declustering. However, we think, this may be a result of usual assumption that magnitude of aftershock does not exceed the magnitude of its main shock. This automatically forms a general deficit of larger aftershocks and results in an increase in the *b* value in aftershocks and corresponding decrease in the declustered catalog.

Bayesian estimates for *c* and *p* in the Omori–Utsu law can be obtained using the method of Holschneider et al. (2012). The input information are times of aftershocks in an interval $[t_{start}, t_{stop}]$. The recommendations in this publication are to search for the optimal parameter value for the quantity lg(*c*). Flat priors for lg(*c*) and *p* are the default option in the program provided in the public domain by the authors of the paper. Baranov et al. (2019c) modified the procedure for the use of the normal prior. In a similar manner to the estimation of *b*, using 334 aftershock sequences, they obtained normal approximations

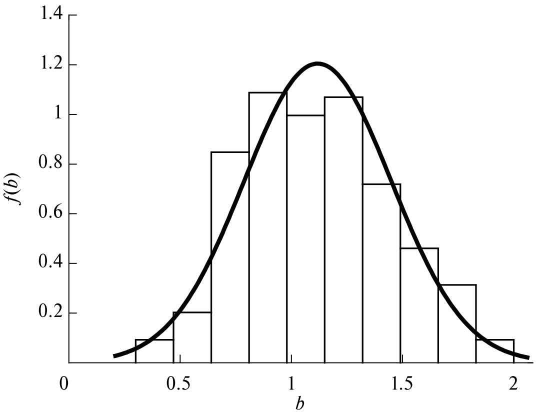


Fig. 16 Empirical distribution of b value estimates (Baranov et al. 2019c, Fig. 1). Density *f*(*b*) (histogram) is based on 334 aftershock series from world's earthquakes with $M_m \geq 6.5$. Heavy line is normal approximation with mean $Eb = 1.12$ and standard deviation $\sigma b = 0.3$

for the distributions of $\lg(c)$ and p with $E\lg(c) = -1$ and $Ep = 1.05$ and standard deviations $\sigma\lg(c) = 0.74$ and $\sigma p = 0.25$.

The choice of t_{start} is an important element when applying the Holschneider's procedure. Obviously, some of the aftershocks that occur immediately after the main shock are not included in the catalog, primarily because their waveforms are overlapped by the main shock, and also because of the overlapping the aftershock waveforms. Helmstetter et al. (2006) analyzed the Californian earthquake catalog to detect an approximate relationship between main shock magnitude M_m, the completeness magnitude M_c, and time t_{start} after which the catalog is complete for $M \geq M_c$ They obtained:

$$M_c = M_m - 4.5 - 0.761\,\lg\left(t_{start}\right)$$

where time is in days relative to the main shock. Baranov et al. (2019a, b, c) found for worldwide seismicity:

$$t_{start} = 10^{\left(M_m - M_c - 3.5\right)/0.7} \tag{23}$$

Hainzl (2016) proposed an approach for estimating the pair (M_c, t_{start}), based on the idea that a catalog becomes incomplete, when the rate of events is too high. This approach requires adjustment of the parameters, but there is no universal criterion for such an adjustment.

Another approach, proposed by Shebalin and Baranov (2017), is to find a t_{start} value that stabilizes the b value estimates (or a mean magnitude that has a one-to-one correspondence with b value according to (Aki 1965).

There is another version (Baranov et al. 2019c) which supposes that the parameter c is really small for aftershocks of a large earthquake and considerably below the estimate based on values of M_c by Eq. (23). The actual estimate of c can thus serve to determine t_{start} for the associated value of M_c. At the first step the parameters in the Omori–Utsu law are estimated for the interval (0.0001, t) days, say, for $M \geq M_c$. The next step would be to use [$c = t_{start}$, t] with $M \geq M_c$ to find new estimates of c, p.

Baranov et al. (2019c) tested all the four methods for estimating the {t_{start}, M_c} pair based on the world data for 777 aftershock sequences extracted from epy ANSS ComCat catalog. They found a method that used Eq. (23) as the most suitable for such a worldwide analysis.

7.2 Testing Forecasts

Evaluation of forecasts and comparison of different forecasts is an issue in its own right in statistical seismology. There are two generally different approaches for this. Alarm-based forecasts, usually also called predictions specify areas in time, space, or both in time and space and magnitude interval (Zechar 2010). Another common format for earthquake forecasts is a gridded rate forecast, in which the geographic region of interest is divided into sections and the forecast indicates the expected number of earthquakes in each section. This format is widely used by the Collaboration for Earthquake Predictability Research (CSEP) testing centers (Jordan 2006; Zechar et al. 2010).

Two types of error analysis are used to evaluate alarm-based predictions: target misses and false alarms. The most commonly used are receiver operating characteristic (ROC) diagram (Mason 2003, and references therein) and Molchan diagram (Molchan 1991). Molchan diagram, which is a plot of the miss rate versus the fraction of space–time volume occupied by "alarms". An important feature of the Molchan diagram which makes it preferable in comparison to ROC diagram is that the time–space may be defined using a reference model. The reference model specifies what is the probability that an event occurs within the time–space of alarms. Alarm-based forecasts usually imply a control parameter. “Alarms” are “switched on” in case the control parameter exceeds a given threshold. Increasing threshold leads to an increasing number of the target misses but decreasing the time–space of alarms. The whole range of the control parameter forms a trajectory in Molchan diagram. If the trajectory goes below the diagonal (0.1; 1.0), then the tested predictions are more performant than the reference model.

The reference model can define the distribution of a control parameter, not necessarily related to space or time—for example, to predict only the magnitude of an earthquake. In this case, the time–space of alarms on the Molchan diagram can be replaced by the function of the distribution of the control parameter (Baranov et al. 2019c).

For the format of the gridded rate forecasts a set of tests have been developed based on the joint likelihood function (Zechar 2010). The forecasts are done in a form:

$$\Lambda = \left\{ f_{ij} | i \in M, j \in S \right\} \tag{24}$$

where M is the binned magnitude range of interest, $\boldsymbol{S}$ is the binned spatial domain of interest, and f_{ij} specifies the probability of observing zero earthquakes, one earthquake, etc.

The locations of the observed earthquakes are binned using the same discretization:

$$\Omega = \{ \omega(i,j) | i \in M, j \in S \} \tag{25}$$

where $\omega(i,j)$ specifies the number of observed earthquakes within a bin (i, j). The joint likelihood is the probability to observe $\omega(i,j)$ earthquakes in all bins according to the forecast model. It is assumed that the values in different cells are independent. In this case, the joint likelihood is the product of the probabilities calculated in each bin, and the log-likelihood is convenient for replacing the product with the sum:

$$L(\Omega|\Lambda) = \sum \log \left[f_{ij}(\omega(i,j) \right] \tag{26}$$

Assuming that the number of earthquakes in each bin has a Poisson distribution, the probabilities f_{ij} can be calculated using the modeled rates in the bins. In that case Eq. (26) is:

$$L(\Omega|\Lambda) = \sum \left[-\lambda(i,j) + \omega(i,j) \log \lambda(i,j) - \log(\omega(i,j)! \right] \tag{27}$$

where $\lambda(i, j)$ is Poisson rate in bin (i, j). In more general case the probabilities can be calculated using numerical simulations.

Several likelihood-based tests have been developed for gridded-type forecasts: L-test, N-test, S-test, and R-test (see Zechar 2010 and references therein).

The Likelihood test (L-test) answers a question: is the observed catalog of earthquakes consistent with the forecast? The test is based on multiple simulations of the earthquake catalog. For each catalog, the joint log-likelihood is calculated by Eq. (26) or (27) and compared to the log-likelihood value calculated using the real catalog. The proportion γ of cases in which the likelihood value according to the simulated catalog is less than according to the real one is calculated. A very small value of γ indicates that the observation is inconsistent with the forecast (at a 100(1-γ)% confidence level).

The Number test (N test) is aimed to verify whether the number of observed target earthquakes is consistent with the forecast. Using catalog simulations or expected Poisson rates if appropriate, one may calculate the distribution function for the total number of earthquakes according to the forecast model. At a given significance level α, the total observed number of earthquakes is consistent with the model if this value lies between $\alpha/2$ and 1-$\alpha/2$ quantiles of the obtained distribution.

The Magnitude test (M-test) and Space test (S-test) are designed to consider only the magnitude or spatial distributions of the model and the observation. The forecast is reduced to a set of magnitude bins by summing expected rates over spatial bins in M-test and to a set of spatial bins in S test. In multiple simulations of catalogs, the "reduced" log-likelihood is compared between the simulated and real catalogs, and a proportion g analogous to that calculated in L test is determined. Again, a very small value of γ indicates that the observation is inconsistent with the forecast at a 100(1-γ)% confidence level.

The Likelihood Ratio test (R-test) is designed to compare two forecast models. It is based on the natural idea that a forecast with a higher joint log-likelihood is better. The likelihood ratio for two forecasts Λ_A and Λ_B is the difference of the joint log-likelihoods:

$$R = L\left(\Omega|\Lambda_A\right) - L\left(\Omega|\Lambda_B\right) \tag{28}$$

The likelihood scores and the Molchan (2010) diagram may complement each other. Shebalin et al. (2014) demonstrated that in the error diagram, relatively more weight is given to successful forecasts and false alarms in bins with high forecasted rates. In contrast, likelihood criteria give relatively more weight to bins of low probability of events (including target misses). For likelihood tests, moreover, there is often a problem of zero or very low expected rates. The Molchan diagram, in turn, has the disadvantage that it is not applicable for assessing the forecast of the overall expected rate, but only its distribution in time and / or space.

7.3 Magnitude of the Largest Aftershock

Very popular Båth law forecasts the magnitude of the largest aftershock as the magnitude of the main shock minus about 1. This is probably the reason that little attention has been paid to the problem of forecasting this value, although Båth law gives a forecast only in general, and the actual values vary greatly. In addition, as shown above, over time, the magnitude of the subsequent strongest aftershocks decreases, and often the strongest aftershocks occur in the first hours after the main shock.

The procedure to use for estimating the maximum magnitudes of future aftershocks based on observations of past ones was developed by a joint team of researchers from Canada and Japan (Shcherbakov et al. 2018; Shcherbakov 2021) and by a Russian team (Baranov and Shebalin 2018; Baranov et al. 2019c). The approaches of both teams are based on using a straightforward superposition of the Gutenberg–Richter and Omori–Utsu laws with a view to simulating aftershock processes (see above). Both approaches make use of the Gutenberg–Richter law (13), the Omori–Utsu law (1), and Vere Jones' Eq. (14) in the form of a double exponent.

The main differences between the two approaches are as follows. First, Baranov et al. (2019) estimate the parameters with due account of catalog incompleteness at the beginning of a sequence. To do this, these authors estimate, not only the completeness magnitude, but also the time since the main shock when the magnitude actually starts to be complete. Second, they use the Bayesian method merely to regularize the estimates of the parameters of the Gutenberg–Richter and Omori–Utsu laws rather than for estimating the magnitude distribution of the largest aftershock, as was done by Shcherbakov et al. (2018).

Third, they regularize the estimates of the parameters by using their distributions based on worldwide data as Bayesian priors. Shcherbakov et al. (2018) use 17 aftershock sequences from several different regions worldwide. Baranov et al. (2019c) use global data for 777 aftershock sequences as identified by the algorithm of Molchan and Dmitrieva (1992) from the ANSS ComCat catalog for the period 1980–2016.

Baranov et al. (2019c) compare these two approaches by integral collation of results from a retrospective estimate for the magnitude of the largest aftershock for the time interval (t, $T=365$) days, i.e., the quantity $M_1(t, T)$. Two independent tests are used for the comparison. One test, R test (see Sect. 7.2), is based on the ratio between the likelihood function of the estimate under analysis and the estimate using the dynamic Båth law (15) using the actual maximum magnitudes (Schorlemmer et al. 2007). The other test uses of the error diagram due to Molchan (1991) relative to the dynamic Båth law which is the reference model.

Using the R test Eq. (28), the authors calculate the geometric mean $LG(N)$ of the value exp(R) in the N terminated forecasts. The $LG(N)$ value is the greater, the more often the values of $M_{1,i}(t,T)$ fall into the region of higher probability density in the tested model compared to the reference model, and the less often they fall into the region of low probability density.

In the Molchan diagram the control parameter is the distance to the maximum density (mode) of the forecast distribution, δM, since the forecast is considered the better the nearer is the sample to the mode. Time–space of alarms in this case is replaced by the probability of $M_1(t,T)$ falling in $[\mathrm{E}_{db}(t) - \delta M, \mathrm{E}_{db}(t) + \delta M]$, and the fraction of failures-to-predict is the number of cases in which $M_{i,1}(t,T) < E(t) - \delta M$ or $M_{i,1}(t,T) > E(t) + \delta M$, where $E(t)$ is the mode of the distribution for the tested model, while the mode of the distribution for the dynamic Båth law, $E_{db}(t)$, is given by Eq. (20).

The tangent of the angle in the straight line passing through the point (0,1) and a point at the trajectory in the Molchan diagram is called probability gain (Molchan 1991; Shebalin et al. 2014). The better the model, the greater the probability gain. For convenience in the comparison for different values of t that quantity was considered at $\nu = 0.5$; we denote it as $PG_{0.5}$. The value 1 for both tests means that the performance in predicting $M_1(t, T)$ using the method being tested is the same as the estimate based on the dynamic Båth law, that is, without appealing to the information on past aftershocks. Values greater than 1 mean that the method is preferable compared with the reference model, while values below 1 make the dynamic Båth law the method of choice.

Estimates were made when the number of aftershocks above the completeness magnitude in the interval (t_{start}, t) was at least N_0. Otherwise, the estimate based on the dynamic Båth law was used. Since the evaluation of performance assumed the dynamic Båth law as the reference model, such cases were excluded from the evaluation. The quantity N_0 is the only free parameter in this method due to Baranov et al. (2019c). Tests with N_0 varying between 4 and 20 resulted in the choice of the optimal value, $N_0 = 5$. That value was afterward used in all tests.

The testing results are listed in Table 2. For all times t tested, the values of the LG and $PG_{0.5}$ tests were greater than 1 and were 1.22 on average. It thus appears that the procedure described here incorporates information on past aftershocks to yield considerably better time-dependent estimates of the future largest aftershock compared with the estimates based on the dynamic Båth method. The mean gain in probability was approximately 22%.

Comparison of these results (Table 2) with the results obtained by the basic method showed a substantial advantage of the basic option where point parameter estimates are employed. The same retrospective data were used to compare the estimates using our

Table 2 The results of retrospective testing of the procedure for forecasting $M_1(t,T)$ (Baranov et al. 2019c, Table 2)

t, days	$N^{(1)}$	Basic method (from point parameter estimates)		Method based on the Bayesian approach for estimating $M_1(t,T)M_1(t,T)$		After (Scherbakov et al. 2017)	
		LG	$PG_{0.5}$	LG	$PG_{0.5}$	LG	$PG_{0.5}$
0.25	107	1.392	1.337	1.222	1.215	1.081	0.743
0.5	180	1.354	1.292	1.224	1.006	1.078	0.756
1	257	1.243	1.218	1.136	1.007	1.036	0.812
2	318	1.273	1.146	1.176	1.034	1.084	0.839
4	362	1.266	1.155	1.201	1.103	1.121	0.929
8	407	1.302	1.093	1.203	1.056	1.116	0.923
16	447	1.148	1.096	1.139	1.045	1.096	0.961
32	486	1.216	1.055	1.178	1.033	1.155	1.014
64	523	1.265	1.159	1.220	1.147	1.206	1.062
Mean	343	1.273	1.172	1.189	1.072	1.108	0.893
		1.223		1.130		1.000	

N is the number of sequences for which the number of completely reported aftershocks in the interval (t_{start}, t) is 5 or greater.

method with those from (Scherbakov et al. 2018). These results are also listed in Table 2. It has turned out that the estimates following (Shcherbakov et al. 2018) were much worse, not better than those based on the dynamic Båth law.

7.4 Hazardous Period Duration

Many researchers considered the problem of estimating the time of the largest aftershock (Saichev and Sornette 2005; Tahir et al. 2002; Shcherbakov et al. 2018). The problem can be dealt simultaneously with that of estimating the magnitude of the largest aftershock (Shcherbakov et al. 2018). In our opinion, there is a problem that is more urgent in practice, namely, the estimation of the time of the last aftershock with a magnitude above a specified value, since it is not only the largest aftershock may pose serious threat, but other large aftershocks too.

A natural way to solve this problem would be applying the ETAS model with multiple simulations to find the distribution of the time of the last aftershock with magnitude above a given threshold. The ETAS model implies that large aftershocks may significantly change the duration of the hazardous period by breaking an Omori-like behavior of the sequence. Spassiani et al. (2018) studied theoretical probabilities of such a break. They found that for some sequences conform to the Omori law, some do not. In our opinion, this is a consequence of the productivity law (Sect. 4), from which it follows that the number of aftershocks in an event of a given magnitude can be both large and, most likely, very small.

Shebalin and Baranov (2019) constructed a solution to this problem based on the model of the aftershock process in accordance with the Omori–Utsu law (1) and taking into account the exponential productivity law (Eq. (11)). The magnitude threshold can be specified with regard to how dangerous are events of this magnitude for the region in question.

The density and distribution function for the time t of an arbitrary aftershock for an assumed case in which times are considered within some interval T (everywhere below we will be using the value $T=365$ days) after the main shock and have the form (Holschneider et al. 2012):

$$f(t) = \frac{D_{c,p}(0,T)}{(1+t/c)^p}, \quad F(t) = \frac{D_{c,p}(0,T)}{D_{c,p}(0,t)} \tag{29}$$

Here, c and p are the parameters in the Omori–Utsu law (1), and function $D_{c,p}\left(t_1, t_2\right)$ is defined by Eq. (17).

The probability that all k aftershocks will occur before a time τ, given $\tau < T$, is $F(\tau)^k$. We will assume the Poisson distribution with mean Λ to hold for the rate of aftershocks in the interval $(0,T)$ in each sequence. In that case, according to (8), the probability distribution for the time τ when the last aftershock of a specified magnitude will occur (given $\tau < T$) is

$$F_\Lambda(\tau) = e^{-\Lambda[1-F(\tau)]} \tag{30}$$

and its density is

$$f_\Lambda(\tau) = \Lambda f(\tau) e^{-\Lambda[1-F(\tau)]} \tag{31}$$

Note that Eq. (30) does not vanish at $\tau=0$, since there is always a certain probability that the earthquake will not be followed by aftershocks of the indicated magnitude, which the larger, the smaller Λ. All other things being equal, that quantity can vary from earthquake to earthquake by an order or greater (Marsan and Helmstetter 2017). As was shown above, when there is a fixed threshold for the magnitudes under consideration relative to the main shock magnitude, the global and regional distributions of Λ have the exponential form, Eq. (10). As well, the parameter $\Lambda_{\Delta M}$ (the mean rate of $M \geq M_m$-ΔM aftershocks) decreases with increasing main shock depth on the global (Fig. 10a) and on the regional level (Fig. 10b–c). This tells us that depth should be incorporated in the calculation.

The quantity $\Lambda_{\Delta M}$ frequently has values near 0 in some sequences. For such sequences, the probability of large subsequent shocks is low in a short time after the main shock. Average estimates obviously require the distribution (10) to be included. Based on Eqs. (10) and (30), we derive an average distribution function $\overline{F_\Lambda}(\tau)$ and the density $\overline{f_\Lambda}(\tau)$ for τ under the condition $\tau < T$ (Shebalin and Baranov 2019):

$$\overline{F_\Lambda}(\tau) = \int_0^\infty F_\Lambda(\tau) f(\Lambda) d\Lambda = \frac{1}{\Lambda_{\Delta M}} \int_0^\infty e^{-\Lambda/\Lambda_{\Delta M}} e^{-\Lambda[1-F(\tau)]} d\Lambda = \frac{1}{1+\Lambda_{\Delta M}[1-F(\tau)]} \tag{32}$$

$$\overline{f_\Lambda}(\tau) = \frac{\Lambda_{\Delta M} f(\tau)}{\{1+\Lambda_{\Delta M}\left[1-f(\tau)\right]\}^2} \tag{32a}$$

The key parameter in Eq. (32) is $\Lambda_{\Delta M}$, which determines the probability function at $\tau=0$. The smaller $\Lambda_{\Delta M}$, the higher is the probability of no aftershocks of the specified magnitude at all. The quantity $\Lambda_{\Delta M}$ depends on the magnitude threshold selected relative to main shock magnitude. The value $\Delta M=2.0$ is a convenient threshold to use. This is considerably below the mean magnitude difference between the largest aftershock and the main shock as given by the Båth law, but high enough for the events to be above the completeness magnitude in most cases.

Shebalin and Baranov (2019) estimated of the parameters in Eq. (32) based on worldwide and regional data and plotted against depth. Figure 17 shows the parameter Λ_2 as

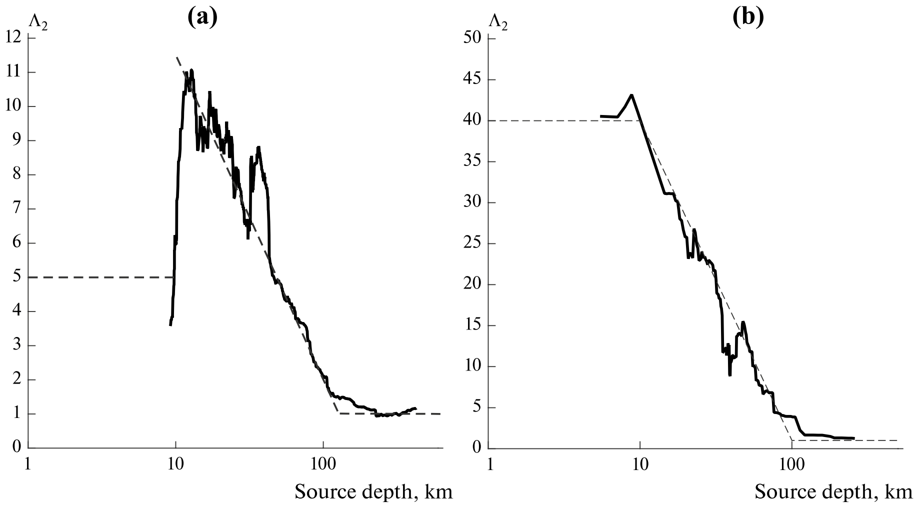


Fig. 17 Dependence of parameter Λ_2 on source depth for aftershock sequences (Shebalin and Baranov 2019, Fig. 2) **(a)** from world earthquakes with Mm ≥ 6.5, 1980–2018; **(b)** from earthquakes with Mm ≥ 6.0 in Kuril–Kamchatka region. Solid line estimates of parameter Λ_2 from main shocks ordered by increasing source depth in moving window of 50 events with step of 5 events (average over 50 events is assumed as depth value); dashed line, piecewise linear approximation on logarithmic scale (see text)

a function of depth of focus for aftershock sequence due to large ($M_m \geq 6.5$) earthquakes worldwide based on the ANSS ComCat catalog and for Kamchatka and the Kurile ($M_m \geq$ 6.0). It has turned out that Λ_2 can vary by factors of a few tens over depth. For both of these cases it has turned out that the dependence is exponential for depths between 10 and 100 km.

Model (32) was examined for consistency with empirical data based on the global catalog (Shebalin, Baranov, 2019) for four ranges of depth, resulting in an estimate of Λ_2. There is not enough data for the Baikal and Caucasian regions to estimate the dependence of Λ_2 on depth, consequently, the estimates used for the two regions were common values for all depths.

It seems to be impossible to estimate real values of c for the aftershock sequences of large earthquakes, but one can obtain indirect estimates of the parameter based on the aftershocks of smaller earthquakes, doing a joint analysis of many sequences (Narteau et al. 2009). Shebalin and Narteau (2017) have identified a dependence of c on main shock depth. Indeed, it has turned out that c is little dependent on the magnitude of both the main shock and aftershocks. The values of c vary within $10^{-4}-10^{-1}$ days for strike slip faults in California, demonstrating a persistent tendency of decreasing with increasing depth in the interval 2—15 km, with the bulk of Californian earthquakes occurring in this layer. Shebalin and Baranov (2019) obtained a similar tendency for the Japan subduction zone (Fig. 18a) in the range down to 40 km depth where c decreases from 10^{-2} to 10^{-3} days.

An analogous analysis has also been done for the earthquakes in the Kuril–Kamchatka region (Fig. 18b). It has turned out that the parameter decreases similarly in the depth range down to 40 km from 10^{-2} days to 10^{-3} days, while increasing somewhat again from 40 to 100 km. It can be surmised that the parameter c varies over depth in a similar manner in about the same limits at other subduction zones. The subduction earthquakes make the bulk of all world earthquakes; hence c must have the same order of magnitude over all large earthquakes on average. With this consideration in mind, Shebalin and Baranov (2019) adopted values of c for the four depth ranges used for testing the model.

It is quite legitimate to use aftershock sequences of large earthquakes in order to estimate the parameter p in the Omori–Utsu law (1), provided the initial segment of each sequence has been excluded from estimation. Shebalin and Baranov (2019) used the worldwide catalog and the Kuril–Kamchatka data to obtain dependences of p on depth of focus (Fig. 19). For the worldwide catalog they also obtained analogous estimates for the four depth ranges considered here.

The model for the distribution of τ was tested (Shebalin and Baranov 2019) by comparing theoretical and empirical distribution functions for this quantity in the four depth ranges using the worldwide catalog and three regional catalogs (Fig. 20). The theoretical distributions were constructed in accordance with Eq. (32). The empirical

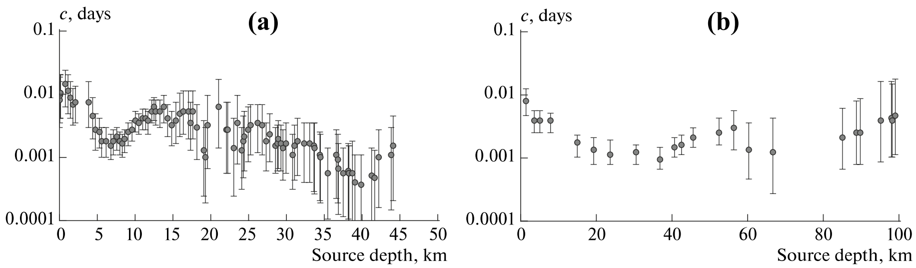


Fig. 18 Dependence of estimates of parameter c of Omori–Utsu law on source depth of main shock (Shebalin and Baranov 2019, Fig. 3) obtained by method of Shebalin and Narteau (2017): **(a)** earthquakes in subduction zones near Japan; **(b)** Kuril–Kamchatka earthquakes. Maximum likelihood estimates (circles) and confidence intervals at level of 95% are indicated in graphs

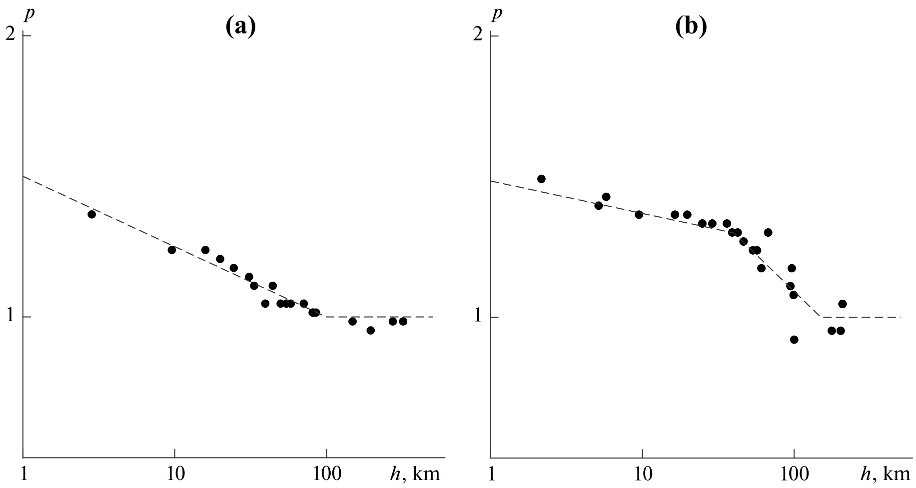


Fig. 19 Dependence of parameter p on source depth for aftershock sequences **(a)** from world earthquakes with $M_m \geq 6.5$, 1980– 2018; **(b)** from earthquakes with $M_m \geq 6.0$ in Kuril–Kamchatka region (Shebalin and Baranov 2019, Fig. 4). Circles mark estimates of parameter p from main shocks ordered by increasing source depth in moving window of 50 events with step of 5 events (average over 50 events is assumed as depth value); dashed line is piecewise linear approximation on logarithmic scale

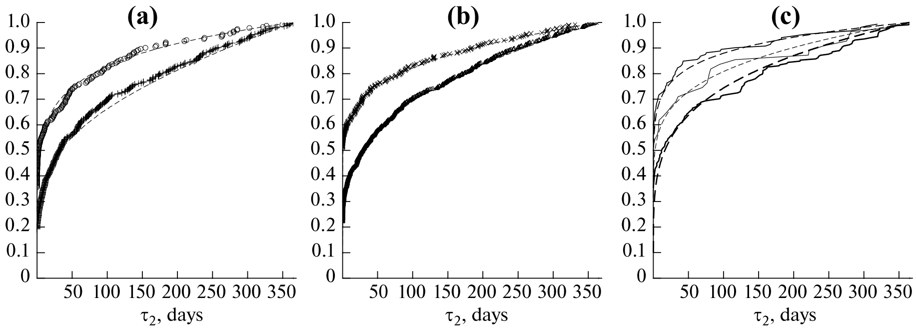


Fig. 20 Empirical and theoretical distribution functions of duration of aftershock hazardous period τ_2 (a), (b) for global catalog $M_m \geq 6.5$, depth intervals of main shock **(a)** $0 \leq h < 10$ km (circles), $10 \leq h < 30$ km (triangles), **(b)** $30 \leq h < 50$ km (plus signs), 50 km $\leq h$ (crosses)) and **(c)** for regional catalogs: Kuril–Kamchatka region, $M_m \geq 6.0$ (heavy line), Baikal region, $M_m \geq 5.5$ (thin line), Caucasus region, $M_m \geq 5.0$ (medium heavy line). Theoretical distribution functions Eq. (32) are shown by dashed lines. (Shebalin and Baranov 2019, Fig. 5.)

distributions differ considerably among themselves, but the theoretical ones differ from the respective empirical distributions much less. The best agreement is achieved for $h \geq 50$ km when using the worldwide catalog (at confidence level 0.99 for the Kolmogorov test) and the worse agreement for $h = 30 - 50$ km (confidence level 0.52). Considering that each aftershock sequence can involve some local fluctuations over time from the Omori–Utsu law and that the model Eq. (32) is a strongly nonlinear function of c and p, the results can be considered a good confirmation of the model. This good agreement between theoretical and empirical distributions also corroborates that the use of indirect (rather than impossible direct) estimates of c whose values strongly affect the distribution shape is legitimate.

Shebalin and Baranov (2019) suggested also to incorporate the information on the first aftershocks to obtain more accurate estimates of the duration of the hazardous period. But the interval before the time tstart from which the catalog can be considered complete, must be excluded. Equation 30 is again the basis for such forecasts. The total rate Λ of aftershocks with magnitude above the relative threshold M_m—2 in the interval [0, T] is estimated using Omori–Utsu model:

$$\Lambda = n_c\left(t_{start}, t_0\right) 10^{b(M_c - M_m + 2)} \frac{D_{c,p}\left(t_{start}, t_0\right)}{D_{c,p}(0,T)} \tag{33}$$

where the function $D_{c,p}$ is given by Eq. (17).

For the forecasts the authors used the predefined values of parameters c, p, b in the same way as without using data on the first aftershocks. The results were evaluated based on data on the first aftershocks using the information gain test LG (LG is defined as a geometric mean of the value exp(R) in the N forecasts, and R is the value of R test, Eq. (28), for each forecast). Equation (32) was used as the reference model for the R test. The results are listed in Table 3. Forecasts using information on the first aftershocks are significantly more performant.

Duration of the hazardous period obviously depends on the total duration of the aftershock sequences, a problem that has received much more attention. In particular, it was found that the magnitude of the main shock is not a determining factor for the duration

Table 3 The results of a retrospective test for estimates of the duration of the hazardous period due to the $M \geq M_m - 2$ aftershocks using the information on the aftershocks in the interval [0, 0.5] days (Shebalin and Baranov 2019, Table 2.)

Catalog	Number of sequences N with $n_c(t_{start}, t_0) \geq 5$ $n_c(t_{start}, t_0) \geq 5$	$LG(N)$
Worldwide ANSS catalog, $M_m \geq 6.5$	181	1.48
Kuril–Kamchatka region, $M_m \geq 6.0$	32	1.59
Caucasus, $M_m \geq 5.0$	10	1.39
Baikal, $M_m \geq 5.5$	4	1.64

of the series of aftershocks (Ziv 2006; Toda and Stein 2018). Important here finding is that duration depends on the faulting style: longer aftershock duration is characteristic to extensional tectonic settings or normal faulting (Tahir and Grasso 2015; Valerio et al. 2017). Accordingly, not only depth, but also faulting style of main shock should be taken into account in forecasts of the duration of the hazardous period.

8 Conclusions

We have reviewed the main models and contemporary approaches and methods used for predicting aftershock activity in application to evaluation of postseismic hazard. We considered both the physical mechanisms and models of aftershock generation on the one hand and time-dependent models on the other.

We present validation of the approach to the representation of the aftershock process as a direct superposition of the Omori–Utsu and Gutenberg–Richter laws as suggested by Reasenberg and Jones (1989). Among other things, worldwide data were used to corroborate the hypothesis that aftershock times and magnitudes are independent.

We noted that earthquake productivity plays the key role for evaluation of postseismic hazard; the productivity characterizes the ability of different earthquakes to generate subsequent shocks. We review the main approaches to the study of productivity. The Δ-analysis as applied to worldwide and regional data showed that productivity (the number of events triggered by an earlier earthquake) obeys the exponential whose only parameter is independent of main shock magnitude and can be viewed as a clustering factor. Productivity decreases with increasing depths of events, while the exponential shape of the distribution persists for different depths and magnitudes of the triggered events. That result is a recent achievement in statistical seismology and bears the name of earthquake productivity law.

An analysis of data synthesized for the spatial ETAS model showed that the earthquake productivity law is violated in this case, because the model assumes productivity to obey the Poisson distribution. This contradiction was eliminated by incorporating the productivity law in the ETAS model, with the result being called ETAS(e).

The empirical Båth law is another important pattern in statistical seismology. We cite the main publications that contained attempts at theoretical validation of the Båth law. A recent achievement in this area is validation of the Båth law using the Gutenberg–Richter relation and the earthquake productivity law. Modeling the decay of aftershocks using the Omori–Utsu law has enabled the Båth law to be generalized, resulting in the distribution of the difference between the magnitudes of the largest aftershock and main shock over time. The distribution was called the dynamic Båth law. This law has been shown to be consistent with worldwide and regional data.

Estimating the magnitude of the largest aftershock is an important problem in statistical seismology. The general approach to such estimation consists in representing the aftershock process by a superposition of the Gutenberg–Richter and Omori–Utsu laws (the model of P. Reasenberg and L. Jones) followed by using the Bayesian method either to estimate the model parameters or to regularize the estimates based on worldwide data. The modern studies based on worldwide data showed that the latter approach is to be preferred, while a straightforward application of the Bayesian method to parameter estimation gains nothing compared with the dynamic Båth law.

There is another important problem, namely estimating the duration of the hazardous period in aftershock activity i.e., the time span where aftershocks with magnitudes equal to or greater than a specified value are to be expected. We cite the main publications where the problem was examined. The most significant results here include the derivation of the distribution for the duration of the hazardous period in an aftershock sequence and of an averaged distribution for the duration of the hazardous period based on the totality of all sequences. This last result was derived using the law of earthquake productivity. We quote estimates for the parameters in the averaged distribution based on worldwide and regional data. It is shown that the estimates using the averaged model can be improved using the information on the first aftershocks.

It is also necessary to point out that the practical application of these procedures consists in the Aftershock Hazard Assessment System, AFCAST (https://itpz-ran.ru/afcast/). The system uses the worldwide ANSS catalog and performs near-real-time estimation of areas of aftershock activity, the magnitude of the largest aftershock, and the duration of the hazardous period for the $M \geq 6.5$ earthquakes.

In our opinion, one promising line of research in the estimation of largest aftershock magnitudes and hazardous period duration is a scenario approach based on the generation of synthetic catalogs using the ETAS(e) model.

Acknowledgements The authors thank an anonymous reviewer for his insightful comments and suggestions that helped to improve the manuscript. We acknowledge the financial support from Russian Science Foundation, project no. 20-17-00180.

References

Aki K (1965) Maximum likelihood estimate of b in the formula log $N = a{-}bM$ and its confidence level. Bull Earthquake Res Inst 43:237–239

Baiesi M, Paczuski M (2004) Scale-free networks of earthquakes and aftershocks. Phys Rev E 69(6):066106. https://doi.org/10.1103/PhysRevE.69.066106

Baranov SV, Shebalin PN (2018) Forecasting aftershock activity: 3. båth's dynamic law. Izv. Phys Solid Earth 54:926–932. https://doi.org/10.1134/S1069351318060022

Baranov SV, Shebalin PN (2019) Global statistics of aftershocks following large earthquakes: independence of times and magnitudes. J Volcanolog Seismol 13:124–130. https://doi.org/10.1134/S0742046319020027

Baranov AA, Baranov SV, Shebalin PN (2019a) A Quantitative estimate of the effects of sea tides on aftershock activity: Kamchatka. J Volcanolog Seismol 13:56–69. https://doi.org/10.1134/S0742046319010020

Baranov SV, Gvishiani AD, Narteau C, Shebalin PN (2019b) Epidemic type aftershock sequence exponential productivity. Russ J Earth Sci. https://doi.org/10.2205/2019ES000695.s

Baranov SV, Pavlenko VA, Shebalin PN (2019) Forecasting aftershock activity: 4. estimating the maximum magnitude of future aftershocks. Izv. Phys Solid Earth 55:548–562. https://doi.org/10.1134/S1069351319040013

Baranov SV, Zhukova SA, Shebalin KPA, PN, (2020) Productivity of mining-induced seismicity Izv. Phys Solid Earth 56:326–336. https://doi.org/10.1134/S1069351320030015

Båth M (1965) Lateral in homogeneities in the upper mantle. Tectonophysics 2:483–514. https://doi.org/10.1016/0040-1951(65)90003-X

Beeler N, Simpson RW, Hickman SH, Lockner DA (2000) Pore fluid pressure, apparent friction, and Coulomb failure. J Geophys Res 105(B11):25533–25542. https://doi.org/10.1029/2000JB900119

Bender B (1983) Maximum likelihood estimation of b values for magnitude grouped data. Bull Seism Soc Am 73:831–851. https://doi.org/10.1785/BSSA0730030831

Benioff H (1951) Earthquakes and rock creep: (part I: creep characteristics of rocks and the origin of aftershocks). Bull Seism Soc Am 41(1):31–62. https://doi.org/10.1785/BSSA0410010031

Boettcher MS, Jordan TH (2004) Earthquake scaling relations for mid-ocean ridge transform faults. J Geoph Res 109:B12302. https://doi.org/10.1029/2004JB003110

Brodsky EE (2006) Long-range triggered earthquakes that continue after the wave train passes. Geophys Res Lett. https://doi.org/10.1029/2006GL026605

Chan CH, Stein RS (2009) Stress evolution following the 1999 Chi-Chi, Taiwan, earthquake: consequences for afterslip, relaxation, aftershocks and departures from Omori decay. Geophys J Int 177(1):179–192. https://doi.org/10.1111/j.1365-246X.2008.04069.x

Cocco M, Rice JR (2002) Pore pressure and poroelasticity effects in Coulomb stress analysis of earthquake interactions. J Geophys Res 107(B2):2030. https://doi.org/10.1029/2000JB000138

Cocco M, Hainzl S, Catalli F, Enescu B, Lombardi AM, Woessner J (2010) Sensitivity study of forecasted aftershock seismicity based on Coulomb stress calculation and rate- and state-dependent frictional response. J Geophys Res. https://doi.org/10.1029/2009JB006838

Console R, Lombardi AM, Murru M, Rhoades D (2003) Båth's law and the self-similarity of earthquakes. J Geophys Res. https://doi.org/10.1029/2001JB001651

Das S, Scholz CH (1981) Off-fault aftershock clusters caused by shear-stress increase? Bull Seis Soc Am 71:1669–1675

Dascher-Cousineau K, Lay T, Brodsky EE (2019) Two foreshock sequences post gulia and wiemer. Seismol Res Lett 91(5):2843–2850. https://doi.org/10.1785/0220200082

Davidsen J, Gu C, Baiesi M (2015) Generalized Omori-Utsu law for aftershock sequences in southern California. Geophys J Inter 201:965–978. https://doi.org/10.1093/gji/ggv061

Dieterich JH (1992) Earthquake nucleation on faults with rate- and state- dependent strength. Tectonophysics 211:115–134. https://doi.org/10.1016/0040-1951(92)90055-B

Dieterich JH (1994) A constitutive law for rate of earthquake production and its application to earthquake clustering. J Geophys Res V. https://doi.org/10.1029/93JB02581

Dieterich JH (2007) Applications of rate- and state-dependent friction to models of fault slip and earthquake occurrence. Treatise on Geophysics 4:107–129. https://doi.org/10.1016/B978-0-444-53802-4.00075-0

Ebrahimian H, Jalayer F (2017) Robust seismicity forecasting based on Bayesian parameter estimation for epidemiological spatio-temporal aftershock clustering models. Sci Rep 7:9803. https://doi.org/10.1038/s41598-017-09962-z

Falcone G, Console R, Murru M (2010) Short-term and long-term earthquake occurrence models for Italy: ETES, ERS and LTST. Ann Geophys 53:41–50. https://doi.org/10.4401/ag-4760

Felzer K, Brodsky EE (2005) Testing the stress shadow hypothesis. J Geophys Res. https://doi.org/10.1029/2004JB003277

Felzer KR, Brodsky EE (2006) Decay of aftershock density with distance indicates triggering by dynamic stress. Nature 441:735–738. https://doi.org/10.1038/nature04799

Felzer K, Becker TW, Abercrombie EE, Ekström G, Rice JR (2002) Triggering of the 1999 Mw 7.1 hector mine earthquake by aftershocks of the 1992 Mw 7.3 landers earthquake. J Geophys Res. https://doi.org/10.1029/2001JB000911

Felzer KR, Abercrombie RE, Ekström G (2004) A common origin for aftershocks, foreshocks, and multiplets. Bull Seismol Soc Am 94:88–98. https://doi.org/10.1785/0120030069

Freed AM (2005) Earthquake triggering by static, dynamic, and postseismic stress transfer. Annual Rev Earth and Planet Sci 33:335–367. https://doi.org/10.1146/annurev.earth.33.092203.122505

Freed AM, Lin J (2002) Accelerated stress buildup on the southern San Andreas fault and surrounding regions caused by Mojave desert earthquakes. Geology 30:571–574. https://doi.org/10.1130/0091-7613(2002)030%3c0571:ASBOTS%3e2.0.CO;2

Gasperini P, Lolli B (2006) Correlation between the parameters of the aftershock rate equation: implications for the forecasting of future sequences. Phys Earth Planet Inter 156:41–58. https://doi.org/10.1016/j.pepi.2006.01.005
Gerstenberger MC, Wiemer S, Jones LM, Reasenberg PA (2005) Real-time forecasts of tomorrow's earthquakes in California. Nature 435(7040):328–331. https://doi.org/10.1038/nature03622
Gomberg J, Johnson P (2005) Dynamic triggering of earthquakes. Nature 437:830. https://doi.org/10.1038/nature04167
Gomberg J, Reasenberg PA, Bodin P, Harris RA (2001) Earthquake triggering by seismic waves following the landers and hector mine earthquakes. Nature 411(6836):462–466. https://doi.org/10.1038/35078053
Gross S (1996) Aftershocks of nuclear explosions compared to natural aftershocks. Bull Seismol Soc Am 86:1054–1060. https://doi.org/10.1785/BSSA0860041054
Gu C, Schumann AY, Baiesi M, Davidsen J (2013) Triggering cascades and statistical properties of aftershocks. J Geophys Res Solid Earth 118:4278–4295. https://doi.org/10.1002/jgrb.50306
Guglielmi AV, Zavyalov AD (2018) The Omori law: the 150-year birthday jubilee of Fusakichi Omori. J Volcanolog Seismol 12:353–358. https://doi.org/10.1134/S0742046318050044
Gulia L, Wiemer S (2019) Real-time discrimination of earthquake foreshocks and aftershocks. Nature 574:193–199. https://doi.org/10.1038/s41586-019-1606-4
Gulia L, Rinaldi AP, Tormann T, Vannucci G, Enescu B, Wiemer S (2018) The effect of a mainshock on the size distribution of the aftershocks. Geophys Res Lett 45:13277–13287. https://doi.org/10.1029/2018GL080619
Gutenberg B, Richter CF (1956) Earthquake magnitude, intensity, energy, and acceleration. Bull Seismol Soc Am 46(2):105–145. https://doi.org/10.1785/BSSA0460020105
Hainzl S (2016) Rate-dependent incompleteness of earthquake catalogs. Seismol Res Lett 87(2A):337–344. https://doi.org/10.1785/0220150211
Hainzl S, Marsan D (2008) Dependence of the Omori-Utsu law parameters on main shock magnitude: observations and modeling. J Geophys Res. https://doi.org/10.1029/2007JB005492
Hainzl S, Steacy S, Marsan D (2010) Seismicity models based on coulomb stress calculations, community online resource for statistical seismicity. Analysis. https://doi.org/10.5078/corssa-32035809
Hainzl S, Zakharova O, Marsan D (2013) Impact of aseismic transients on the estimation of aftershock productivity parameters. Bull Seismol Soc Am 103:1723–1732. https://doi.org/10.1785/0120120247
Hainzl S, Sippl C, Schurr B (2019) Linear relationship between aftershock productivity and seismic coupling in the Northern Chile subduction zone. J Geophys Res: Solid Earth 124(8):8726–8738. https://doi.org/10.1029/2019JB017764
Hardebeck JL, Okada T (2018) Temporal stress changes caused by earthquakes: a review. J Geophys Res Solid Earth 123(2):1350–1365. https://doi.org/10.1002/2017JB014617
Hardebeck JL, Nazareth JJ, Hauksson E (1998) The static stress change triggering model: constraints from two southern California aftershock sequence. J Geophys Res 103(B10):24427–24437. https://doi.org/10.1029/98JB00573
Harte DS (2017) Probability distribution of forecasts based on the ETAS model. Geophys J Int 210(1):90–104. https://doi.org/10.1093/gji/ggx146
Heimisson ER, Segall P (2018) Constitutive law for earthquake production based on rate-and-state friction: Dieterich 1994 revisited. J Geophys Res Solid Earth 123:4141–4156. https://doi.org/10.1029/2018JB015656
Helmstetter A, Shaw BE (2009) Afterslip and aftershocks in the rate-and-state friction law. J Geophys Res. https://doi.org/10.1029/2007JB005077
Helmstetter A, Sornette D (2002) Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. J Geophys Res. https://doi.org/10.1029/2001JB001580
Helmstetter A, Sornette D (2003) Båth's law derived from the Gutenberg-Richter law and from aftershock properties. Geophys Res Lett. https://doi.org/10.1029/2003GL018186
Helmstetter AY, Kagan Y, Jackson DD (2005) Importance of small earthquakes for stress transfers and earthquake triggering. J Geophys Res. https://doi.org/10.1029/2004JB003286
Helmstetter A, Kagan YY, Jackson DD (2006) Comparison of short-term and time- independent earthquake forecast models for southern California. Bull Seism Soc Am 96(1):90–106. https://doi.org/10.1785/0120050067
Hill DP, Reasenberg PA, Michael A, Arabaz WJ, Beroza G, Brumbaugh D, Brune JN, Castro R, Davis S, Depolo D, Ellsworth WL, Gomberg J, Harmsen S, House L, Jackson SM, Johnston MJS, Jones L, Keller R, Malone S, Munguia L, Nava S, Pechmann JC, Sanford A, Simpson RW, Smith RB, Stark M, Stickney M, Vidal A, Walter S, Wong V, Zollweg J (1993) Seismicity remotely triggered by the magnitude 7.3 Landers. Science. 260:1617–1623. https://doi.org/10.1126/science.260.5114.1617

Holschneider M, Narteau C, Shebalin P, Peng Z, Schorlemmer D (2012) Bayesian analysis of the modified Omori law. J Geophys Res 117:B06317. https://doi.org/10.1029/2011JB009054

Hu C, Cai Y, Liu M, Wang Z (2013) Aftershocks due to property variations in the fault zone: a mechanical model. Tectonophysics 588:179–188. https://doi.org/10.1016/j.tecto.2012.12.013

Huc M, Main IG (2003) Anomalous stress diffusion in earthquake triggering: correlation length, time dependence, and directionality. J Geophys Res. https://doi.org/10.1029/2001JB001645

Jordan TH (2006) Earthquake predictability, brick by brick. Seismol Res Lett 77:3–6. https://doi.org/10.1785/gssrl.77.1.3

Kagan YY, Jackson D (2000) Probabilistic forecasting of earthquakes. Geophys J Int 143(2):438–453. https://doi.org/10.1046/j.1365-246X.2000.01267.x

Kagan YY, Knopoff L (1981) Stochastic synthesis of earthquake catalogs. J Geophys Res 86:2853–2862. https://doi.org/10.1029/JB086iB04p02853

Kato M (2006) Static coulomb failure function and aftershocks of 1995 Kobe earthquake: a statistical test. Geophys Res Lett 33:L17313. https://doi.org/10.1029/2006GL026970

Kilb D, Gomberg J, Bodin P (2000) Triggering of earthquake aftershocks by dynamic stresses. Nature 408:570–574. https://doi.org/10.1038/35046046

King GCP, Devès MH (2015) Fault interaction, earthquake stress changes, and the evolution of seismicity. Treatise on Geophysics: Second Edition 4:243–271

King GCP, Stein RS, Lin J (1994) Static stress changes and the triggering of earthquakes. Bull Seimol Soc Am 84(3):935–953. https://doi.org/10.1785/BSSA0840030935

Kisslinger C (1993) The stretched exponential function as an alternative model for aftershock decay rate. J Geophys Res 98:1913–1921. https://doi.org/10.1029/92JB01852

Knopoff L, Kagan YY, Knopoff R (1982) Values for foreshocks and aftershocks in real and simulated earthquake sequences. Bull Seimol Soc Am 72(5):1663–1676

Kostrov BV, Das S (1982) Idealized models of fault behavior prior to dynamic rupture. Bull Seimol Soc of Am 2(3):679–703. https://doi.org/10.1785/BSSA0720051663

Lasocki S, Karakostas VG, Papadimitriou EE (2009) Assessing the role of stress transfer on aftershock locations. J Geophys Res 114:B11304. https://doi.org/10.1029/2008JB006022

Lombardi AM (2002) Probabilistic interpretation of Båth's law. Ann Geophys 45:455–472. https://doi.org/10.4401/ag-3520

Marsan D (2003) Triggering of seismicity at short timescales following Californian earthquakes. J Geophys Res 108(B5):2266. https://doi.org/10.1029/2002JB001946

Marsan D, Helmstetter A (2017) How variable is the number of triggered aftershocks? J Geophys Res Solid Earth 122:5544–5560. https://doi.org/10.1002/2016JB013807

Marsan D, Lengline J (2008) Extending earthquakes' reach through cascading. Science 319:1076–1079. https://doi.org/10.1126/science.1148783

Marzocchi W, Sandri L (2003) A review and new insights on the estimation of the b-value and its uncertainty. Ann Geophys 46(6):1271–1282. https://doi.org/10.4401/ag-3472

Marzocchi W, Spassiani I, Stallone A, Taroni M (2020) How to be fooled searching for significant variations of the b-value. Geophys J Int 220(3):1845–1856. https://doi.org/10.1093/gji/ggz541

Mignan A (2015) Modeling aftershocks as a stretched exponential relaxation. Geophys Res Lett 42(22):9726–9732. https://doi.org/10.1002/2015GL066232

Mizrahi L, Nandan S, Wiemer S (2021) The effect of declustering on the size distribution of mainshocks. Seismol Res Lett 92(4):2333–2342. https://doi.org/10.1785/0220200231

Mogi K (1962) On the time distribution of aftershocks accompanying the recent major earthquakes in and near Japan. Bull Earthq Res. Inst Univ Tokyo. 40:175–185

Molchan G (1991) Structure of optimal strategies in earthquake prediction. Tectonophysics 193:267–276. https://doi.org/10.1016/0040-1951(91)90336-Q

Molchan G (2010) Space-time earthquake prediction: the error diagrams. Pure Appl Geophys 167:907–917. https://doi.org/10.1007/s00024-010-0087-z

Molchan GM, Dmitrieva EO (1992) Aftershock identification: methods and new approaches. Geophys J Int 109(3):501–516. https://doi.org/10.1111/j.1365-246X.1992.tb00113.x

Molchan G, Kronrod T, Nekrasova A (1999) Immediate foreshocks: time variation of the b- value. Phys Earth Planet Int 111:129–140. https://doi.org/10.1016/S0031-9201(98)00163-0

Nanjo K, Nagahama H (2000) Spatial distribution of aftershocks and the fractal structure of active fault systems. Pure Appl Geophys 157:575–588. https://doi.org/10.1007/PL00001108

Nanjo KZ, Enescu B, Shcherbakov R, Turcotte DL, Iwata T, Ogata Y (2007) Decay of aftershock activity for Japanese earthquakes. J Geophys Res. https://doi.org/10.1029/2006JB004754

Nanjo KZ, Hirata N, Obara K, Kasahara K (2012) Decade-scale decrease in b value prior to the M9-class 2011 Tohoku and 2004 Sumatra quakes. Geophys Res Lett. https://doi.org/10.1029/2012GL052997

Narteau C, Shebalin P, Holschneider M (2002) Temporal limits of the power law aftershock decay rate. J Geophys Res. https://doi.org/10.1029/2002JB001868

Narteau C, Shebalin P, Hainzl S, Zoller G, Holschneider M (2003) Emergence of a band- limited power law in the aftershock decay rate of a slider-block model. Geophys Res Lett. https://doi.org/10.1029/2003GL017110

Narteau C, Byrdina S, Shebalin P, Schorlemmer D (2009) Common dependence on stress for the two fundamental laws of statistical seismology. Nature 462(2):642–645. https://doi.org/10.1038/nature08553

Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. J Am Stat Assoc 83:9–27

Ogata Y (1989) Statistical models for standard seismicity and detection of anomalies by residual analysis. Tectonophysics 69:159–174. https://doi.org/10.1016/0040-1951(89)90191-1

Ogata Y (1998) Space-time point-process models for earthquake occurrences. Ann Inst Stat Math 50:379–402. https://doi.org/10.1023/A:1003403601725

Ogata Y (2004) Space-time model for regional seismicity and detection of crustal stress changes. J Geophys Res 109:B03308. https://doi.org/10.1029/2003JB002621

Ogata Y, Guo Z (1997) Statistical relations between the parameters of aftershocks in time, space, and magnitude. J Geophys Res Solid Earth 102:2857–2873. https://doi.org/10.1029/96JB02946

Ogata Y, Katsura K (2014) Comparing foreshock characteristics and foreshock forecasting in observed and simulated earthquake catalogs. J Geophys Res Solid Earth 119:8457–8477. https://doi.org/10.1002/2014JB011250

Ogata Y, Matsuura RS, Katsura K (1993) Fast likelihood computation of epidemic type aftershock-sequence model. Geophys. Res. Lett. 20(19):2143–2146. https://doi.org/10.1029/93GL02142

Omi T, Ogata Y, Hirata Y, Aihara K (2014) Estimating the ETAS model from an early aftershock sequence. Geophys Res Lett 41:850–857. https://doi.org/10.1002/2013GL058958

Omi T, Ogata Y, Hirata Y, Kazuyuki Aihara K (2015) Intermediate-term forecasting of aftershocks from an early aftershock sequence: Bayesian and ensemble forecasting approaches. J Geophys Res Solid Earth 120:2561–2578. https://doi.org/10.1002/2014JB011456

Omori F (1894) On the aftershocks of earthquake. J Coll Sci Imp Univ Tokyo 7:111–200

Otsuka M (1985) Studies on aftershock sequences - part 1. physical interpretations of Omori's formula (in Japanese). Sci Rep Shimabara Earthquake Volcano Obs 13:11–20

Page MT, van der Elst N, Hardebeck J, Felzer K, Michael AJ (2016) Three ingredients for improved global aftershock forecasts: tectonic region, time-dependent catalog incompleteness, and intersequence variability. Bull Seismol Soc Am 106(5):2290–2301. https://doi.org/10.1785/0120160073

Papadopoulos GA, Latoussakis I, Daskalaki E, Diakogianni G, Fokaefs A, Kolligri M, Liadopoulou K, Orfanogiannaki K, Pirentis A (2006) The east aegean sea strong earthquake sequence of october–november 2005: lessons learned for earthquake prediction from foreshocks. Nat Hazards Earth Syst 6:895–901. https://doi.org/10.5194/nhess-6-895-2006

Pisarenko VF, Rodkin MV (2019) Declustering of seismicity flow: statistical analysis. Izv. Phys Solid Earth 55:733–745. https://doi.org/10.1134/S1069351319050082

Reasenberg PA, Jones LM (1989) Earthquake hazard after a main shock in California. Science 242(4895):1173–1176. https://doi.org/10.1126/science.243.4895.1173

Richards-Dinger K, Stein RS, Toda S (2010) Decay of aftershock density with distance does not indicate triggering by dynamic stress. Nature 467(7315):583–586. https://doi.org/10.1038/nature09402

Rodkin MV, Tikhonov IN (2016) The typical seismic behavior in the vicinity of a large earthquake. Phys Chem Earth 95:73–84. https://doi.org/10.1016/j.pce.2016.04.001

Romanowicz B (1992) Strike-slip earthquakes on quasi-vertical transcurrent faults: inferences for general scaling relations. Geophys Res Lett 19(5):481–484. https://doi.org/10.1029/92GL00265

Ross GJ (2021) Bayesian estimation of the etas model for earthquake occurrences. Bull Seimol Soc of Am 111(3):1473–1480. https://doi.org/10.1785/0120200198

Saichev A, Sornette D (2005) Distribution of the largest aftershocks in branching models of triggered seismicity: theory of the universal Båth law. Phys Rev. https://doi.org/10.1103/PhysRevE.71.056127

Saichev A, Sornette D (2007) Theory of earthquake recurrence times. J Geophys Res. https://doi.org/10.1029/2006JB004536.33

Scholz CH (1968) Microfractures, aftershocks and seismicity. Bull Seismol Soc Am 58:1117–1130

Scholz CH (1998) Earthquakes and friction laws. Nature 391:37–42. https://doi.org/10.1038/34097

Scholz CH (2019) The mechanics of earthquakes and faulting. Cambridg Univ. Press, Cambridge. https://doi.org/10.1017/9781316681473
Schorlemmer D, Gerstenberger MC, Wiemer S, Jackson DD, Rhoades DA (2007) Earthquake likelihood model testing. Seismol Res Lett 78(1):17–29. https://doi.org/10.1785/gssrl.78.1.17
Schultz R, Wang R, Gu YJ, Haug K, Atkinson G (2017) A seismological overview of the induced earthquakes in the Duvernay play near Fox Creek, Alberta. J Geophys Res Solid Earth 122:492–505. https://doi.org/10.1002/2016JB013570
Shcherbakov R (2021) Statistics and forecasting of aftershocks during the 2019 Ridgecrest, California, earthquake sequence. J Geophys Res Solid Earth. https://doi.org/10.1029/2020JB020887
Shcherbakov R, Zhuang J, Ogata Y (2018) Constraining the magnitude of the largest event in a foreshock-mainshock-aftershock sequence. Geophys J Int. https://doi.org/10.1093/gji/ggx407
Shcherbakov R, Zhuang J, Zöller G, Ogata Y et al (2019) Forecasting the magnitude of the largest expected earthquake. Nat Commun 10:4051. https://doi.org/10.1038/s41467-019-11958-4
Shebalin PN (2004) Aftershocks as indicators of the state of stress in a fault system. Dokl Earth Sci 398(7):978–982
Shebalin P, Baranov S (2017) Long-delayed aftershocks in New Zealand and the 2016 M7.8 Kaikoura earthquake. Pure Appl Geophys 174:3751–3764. https://doi.org/10.1007/s00024-017-1608-9
Shebalin PN, Baranov SV (2019) Forecasting aftershock activity: 5. estimating the duration of a hazardous period. Izv. Phys Solid Earth 55:719–732. https://doi.org/10.1134/S1069351319050112
Shebalin PN, Baranov AA (2020) Aftershock rate changes at different ocean tide heights. Front Earth. https://doi.org/10.3389/feart.2020.559624
Shebalin P, Baranov S (2021) Statistical Laws of Post-seismic Activity. In: Limnios N, Papadimitriou E, Tsaklidis WG (eds) Statistical methods and modeling of seismogenesis. Wiley, Hoboken, pp 64–103. https://doi.org/10.1002/9781119825050.ch3
Shebalin P, Narteau C (2017) Depth dependent stress revealed by aftershocks. Nat Commun. https://doi.org/10.1038/s41467-017-01446-y
Shebalin PN, Narteau C, Zechar JD, Holschneider M (2014) Combining earthquake forecasts using differential probability gains. Earth Planet Sp. https://doi.org/10.1186/1880-5981-66-37
Shebalin PN, Narteau C, Baranov SV (2020) Earthquake productivity law. Geophys. J. Int. 222(2):1264–1269. https://doi.org/10.1093/gji/ggaa252
Shebalin PN, Vorobieva IA, Baranov SV, Mikhailov VO (2021) Deficit of large aftershocks as an indicator of afterslip at the sources of earthquakes in subduction zones. Dokl. Earth Sc. 498:423–426. https://doi.org/10.1134/S1028334X21050172
Smirnov N (1948) Table for estimating the goodness of fit of empirical distributions. Ann Math Stat 19:279–281. https://doi.org/10.1214/aoms/1177730256
Smirnov VB, Ponomarev AV, Benard P, Patonin AV (2010) Regularities in transient modes in the seismic process according to the laboratory and natural modeling. Izv. Phys Solid Earth 46:104–135. https://doi.org/10.1134/S1069351310020023
Smirnov VB, Ponomarev AV, Stanchits SA, Potanina MG, Patonin AV, Dresen G, Narteau C, Bernard P, Stroganova SM (2019) Laboratory modeling of aftershock sequences: stress dependences of the omori and gutenberg-richter parameters. Izv. Phys Solid Earth 55:124–137. https://doi.org/10.1134/S1069351319010105
Sobolev GA, Zakrzhevskaya NA (2013) On the question of influence of remote earthquakes on seismicity. Izv. Phys Solid Earth 49:474–487. https://doi.org/10.1134/S1069351313030166
Sobolev GA, Ponomarev AV, Koltsov AV, Smirnov VB (1996) Simulation of triggered earthquakes in the laboratory. Pure Appl Geophys 147:345–355. https://doi.org/10.1007/BF00877487
Spassiani I, Marzocchi W (2018) How likely does an aftershock sequence conform to a single omori law behavior? Seismol Res Lett 89(3):1118–1128. https://doi.org/10.1785/0220170224
Steacy S, Gomberg J, Cocco M (2005) Introduction to special section: stress transfer, earthquake triggering, and time-dependent seismic hazard. J Geophys Res. https://doi.org/10.1029/2005JB003692
Stein RS, King GCP, Lin J (1992) Change in failure stress on the southern San Andreas fault system caused by the 1992 Magnitude = 7.4 Landers earthquake. Science 258:1328–1332. https://doi.org/10.1126/science.258.5086.1328
Sun T, Wang K, Iinuma T, Hino R, He J, Fujimoto H, Kido M, Osada Y, Miura S, Ohta Y, Hu Y (2014) Prevalence of viscoelastic relaxation after the 2011 Tohoku-oki earthquake. Nature 514:84–89. https://doi.org/10.1038/nature13778
Tahir M, Grasso JR (2015) Faulting style controls for the space-time aftershock patterns. Bull Seismol Soc of Am 105:2480–2497. https://doi.org/10.1785/0120140336
Tahir M, Grasso JR, Amorèse D (2012) The largest aftershock: how strong, how far away, how delayed? Geophys Res Let 39:L04301. https://doi.org/10.1029/2011GL050604

Tamaribuchi K, Yagi Y, Enescu B, Hirano S (2018) Characteristics of foreshock activity inferred from the JMA earthquake catalog. Earth Planets Space 70:90. https://doi.org/10.1186/s40623-018-0866-9

Toda S, Stein RS (2018) Why aftershock duration matters for probabilistic seismic hazard assessment. Bull Seism Soc Am 108(3A):141–1426. https://doi.org/10.1785/0120170270

Trevlopoulos K, Guéguen P, Helmstetter A, Cotton F (2020) Earthquake risk in reinforced concrete buildings during aftershock sequences based on period elongation and operational earthquake forecasting. Struct Saf 84:101922. https://doi.org/10.1016/j.strusafe.2020.101922

Utsu T (1957) Magnitudes of earthquakes and occurrence of their aftershocks. Zisin, Ser 2(10):35–45 ((in Japanese))

Utsu TA (1961) Statistical study on the occurrence of aftershocks. Geophys Mag 30:521–605

Utsu T (1970) Aftershocks and earthquake statistics: some parameters which characterize an aftershock sequence and their interrelations. J Faculty Sci, Hokkaido Univ Series Geophys. 3(3):129–195

Utsu T, Ogata Y, Matsuura R (1995) The centenary of the Omori formula for a decay law of aftershocks activity. J Phys Earth. 43:1–33. https://doi.org/10.4294/jpe1952.43.1

Valerio E, Tizzani P, Carminati E, Doglioni C (2017) Longer aftershocks duration in extensional tectonic settings. Sci Rep 7:16403. https://doi.org/10.1038/s41598-017-14550-2

Van Stiphout T, Zhuang J, Marsan D (2012). Seismicity Declustering. https://doi.org/10.5078/corssa-52382934

Vere-Jones D (1969) A note on the statistical interpretation of Båth's law. Bull Seism Soc Am. 59:1535–1541. https://doi.org/10.1785/BSSA0590041535

Vere-Jones D (2008) A limit theorem with application to båth's law in seismology. Adv Appl Prob 40:882–896

Vorobieva I, Narteau C, Shebalin P, Beauducel F, Nercessian A, Clouard V, Bouin MP (2013) Multiscale mapping of completeness magnitude of earthquake catalogs. Seismol Soc Am Bull. 103(4):2188–2202. https://doi.org/10.1785/0120120132

Vorobieva I, Shebalin P, Narteau C (2016) Break of slope in earthquake size distribution and creep rate along the San Andreas fault system. Geophys Res Lett 43:6869–6875. https://doi.org/10.1002/2016GL069636

Vorobieva I, Shebalin P, Narteau C (2020) Condition of occurrence of large man-made earthquakes in the zone of oil production. Oklahoma Izv Phys Solid Earth 56(6):911–919. https://doi.org/10.1134/S1069351320060130

Wang J-H (1994) On the correlation of observed Gutenberg-Richter's b value and Omori's p value for aftershocks. Bull Seis Soc Am 84(6):2008–2011

Wang L, Hainzl S, Sinan Özeren M, Ben-Zion Y (2010a) Postseismic deformation induced by brittle rock damage of aftershocks. J Geophys Res. https://doi.org/10.1029/2010JB007532

Wang Q, Schoenberg FP, Jackson DD (2010b) Standard errors of parameter estimates in the ETAS model. Bull Seis Soc Am 100:1989–2001. https://doi.org/10.1785/0120100001

Wang J, Li T, Gu YJ, Schultz R, Yusifbayov J, Zhang M (2020) Sequential fault reactivation and secondary triggering in the March 2019 red deer induced earthquake swarm. Geophys Res Lett. https://doi.org/10.1029/2020GL090219

Watanabe H, Kuroiso A (1970) Aftershocks of the earthquake of the central part of Gifu Prefecture. Bull Earthq Res Inst Univ. Tokyo 48:1195–1208 ((**in Japanese**))

Werner MJ, Sornette D (2008) Magnitude uncertainties impact seismic rate estimates, forecasts, and predictability experiments. J Geophys Res. https://doi.org/10.1029/2007JB005427

Woessner J, Wiemer S (2005) Assessing the quality of earthquake catalogues: estimating the magnitude of completeness and its uncertainty // Bull. Seis Soc Am 95:2684–2698. https://doi.org/10.1785/0120040007

Yang W, Ben-Zion Y (2009) Observational analysis of correlations between aftershock productivities and regional conditions in the context of a damage rheology model. Geophys J Int 177(2):481–490. https://doi.org/10.1111/j.1365-246X.2009.04145.x

Zaliapin I, Ben-Zion Y (2013) Earthquake clusters in southern California I: identification and stability. J Geophys Res 118:2847–2864. https://doi.org/10.1002/jgrb.50179

Zaliapin I, Ben-Zion Y (2016) A global classification and characterization of earthquake clusters. Geophys J Int 207:608–634. https://doi.org/10.1093/gji/ggw300

Zaliapin I, Gabrielov A, Keilis-Borok VI, Wong H (2008) Clustering analysis of seismicity and aftershock identification. Phys Rev Lett. https://doi.org/10.1103/PhysRevLett.101.018501

Zechar JD (2010) Evaluating earthquake predictions and earthquake forecasts: a guide for students and new researchers. Comm Online Res Stat Seism Anal. https://doi.org/10.5078/corssa-77337879

Zechar JD, Gerstenberger MC, Rhoades DA (2010) Likelihood-based tests for evaluating space-rate-magnitude forecasts. Bull Seismol Soc Am 100(3):1184–1195. https://doi.org/10.1785/0120090192

Zhuang J, Ogata Y, Vere-Jones D (2002) Stochastic declustering of space-time earthquake occurrences. J. Am. Stat. Ass. 97(458):369–380. https://doi.org/10.1198/016214502760046925
Zhuang J, Ogata Y, Vere-Jones D (2004) Analyzing earthquake clustering features by using stochastic reconstruction. J Geophys Res. https://doi.org/10.1029/2003JB002879
Zhuang J, Chang C-P, Ogata Y, Chen Y-I (2005) A study on the background and clustering seismicity in the Taiwan region by using point process models. J geophys Res. https://doi.org/10.1029/2004JB003157157
Zhuang J, Werner MJ, Hainzl S, Harte DS, Zhou S (2011) Basics models of seismicity: spatiotemporal models. Comm Online Res Stat Seism Anal. https://doi.org/10.5078/corssa-07487583
Zhuang J, Harte DS, Werner MJ, Hainzl S, Zhou S (2012) Basic models of seismicity: temporal models. Comm Online Res Stat Seism Anal. https://doi.org/10.5078/corssa-79905851
Ziv A (2006) Does aftershock duration scale with mainshock size? Geophys Res Lett 33:L17317. https://doi.org/10.1029/2006GL027141
Zöller G, Holschneider M, Hainzl S (2013) The maximum earthquake magnitude in a time horizon: theory and case studies. Seismol. Soc. Am Bull. 103(2A):860–875. https://doi.org/10.1785/0120120013
Zotov OD, Zavyalov AD, Guglielmi AV, Lavrov IP (2018) On the possible effect of round-the-world surface seismic waves in the dynamics of repeated shocks after strong earthquakes. Izv. Phys Solid Earth 54:178–191. https://doi.org/10.1134/S1069351318010159