





OPEN

Statistical bounds on how induced seismicity stops

Ryan Schultz  , William L. Ellsworth & Gregory C. Beroza

Earthquakes caused by human activities receive scrutiny due to the risks and hazards they pose. Seismicity that occurs after the causative anthropogenic operation stops has been particularly problematic—both because of high-profile cases of damage caused by this trailing seismicity and due to the loss of control for risk management. With this motivation, we undertake a statistical examination of how induced seismicity stops. We borrow the concept of Båth's law from tectonic aftershock sequences. Båth's law anticipates the difference between magnitudes in two subsets of seismicity as dependent on their population count ratio. We test this concept for its applicability to induced seismicity, including ~80 cases of earthquakes caused by hydraulic fracturing, enhanced geothermal systems, and other fluid-injections with clear operational end points. We find that induced seismicity obeys Båth's law: both in terms of the magnitude-count-ratio relationship and the power law distribution of residuals. Furthermore, the distribution of count ratios is skewed and heavy-tailed, with most earthquakes occurring during stimulation/injection. We discuss potential models to improve the characterization of these count ratios and propose a Seismogenic Fault Injection Test to measure their parameters in situ. We conclude that Båth's law quantifies the occurrence of earthquake magnitudes trailing anthropogenic operations.

The injection of fluids in the subsurface has the potential to cause faults to reactivate with earthquake slip^{1,2}. Numerous high-magnitude cases of induced seismicity causing significant impacts^{3–6} have led to increased attention. Study topics have covered a broad range of aspects including triggering mechanisms, maximum magnitude, potential risks/hazards, and managements strategies^{7,8}. Despite the proliferation of scientific studies, relatively little attention has been devoted to understanding how induced earthquake sequences stop. This is surprising because earthquakes trailing well shut-in have been identified as one of the most significant sources of risk for induced seismicity management⁹. In practice, there are many cases where the largest magnitude earthquake trails the shut-in of the responsible well^{5,10–12} or where abrupt 'jumps' in magnitude have occurred¹³.

In this study we build a conceptual framework for statistical bounds on how induced earthquake sequences stop. We start by following the development of Båth's law for aftershocks of natural earthquakes and adapt it through a set of assumptions and justifications for induced earthquakes. We test the validity of this framework on sequences of induced seismicity and find that in most cases the majority of earthquakes occur during the stimulation/injection periods, rather than trailing well shut-in. As well, the largest events occurring during the stimulation/trailing periods follow Båth's law. We also discuss the implications of our results and suggest recommendations for in situ tests that could be useful for constraining induced earthquake risks. Finally, we speculate on how improved physical models could be used to build upon our empirical analysis of how induced earthquakes stop.

Background information on tectonic aftershocks and Båth's law

Båth's law is an empirical observation that the average magnitude difference between a mainshock earthquake and its largest aftershock is approximately 1.2 magnitude units¹⁴. Statistical models of earthquakes^{15,16} have come to interpret Båth's law as a probabilistic consequence of two more fundamental observations^{17–20}. The first being an aftershock law (i.e., Omori's law), which describes the temporal decay of earthquake rate following a main shock^{21–23}.

$$n^o(t) = K(t + c)^{-p} \quad (1)$$

In Omori's law $n(t)$ is the time-dependent rate of aftershocks, K describes the productivity of the aftershock sequence and strongly influences the total number of aftershocks, p is the decay exponent, and c is an ad hoc parameter added to ensure the equation doesn't diverge at the time of the mainshock (i.e., $t=0$). In this formulation, the value of p is bounded such that $p > 1$ for a finite number of earthquakes to occur²⁴. The second

Department of Geophysics, Stanford University, Stanford, CA, USA. ✉email: rjs10@stanford.edu

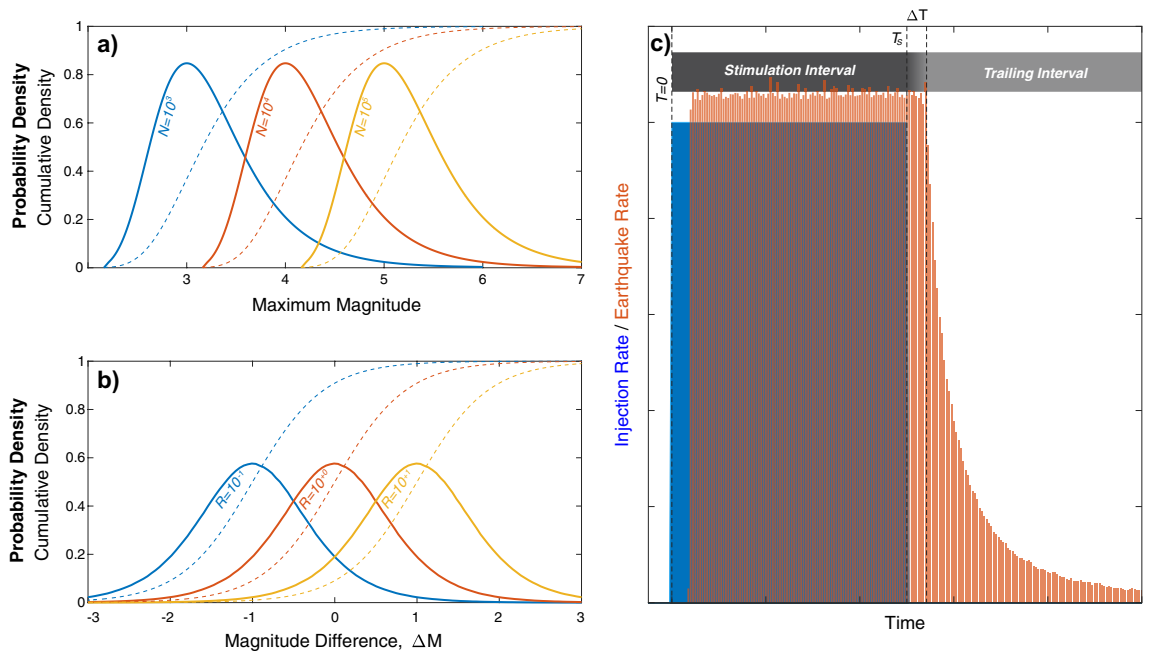


Figure 1. Concepts of Båth’s law and a simple adaptation for trailing seismicity. (a) Probability density functions (solid lines) and cumulative density functions (dashed lines) for the expected maximum magnitude for various population sizes N , based on numerically sampling Eq. (4). (b) Probability density functions (solid lines) and cumulative density functions (dashed lines) for the magnitude difference ΔM for various population ratios R , based on numerically sampling Eq. (5). (c) Simple model of induced seismicity, where earthquakes lag injection/shut-in by time ΔT and trailing seismicity decays with time.

fundamental observation is the Gutenberg Richter magnitude frequency distribution (GR-MFD), which describes the relative number of small and large earthquakes in a sample^{25,26}.

$$N = 10^a 10^{-bM} \tag{2}$$

In the GR-MFD N is the sample size of the catalogue, a is the a -value which is related to the productivity of the sequence, and b is the b -value, which controls the proportionality of big-to-small magnitude earthquakes.

Omori’s law governs the total number of earthquakes that follow a mainshock N_A and the GR-MFD controls the distribution of earthquake magnitudes ‘randomly’ sampled in that population. The combination of these two observations gives the expression for the average magnitude difference (where $N > N_A > 1$).

$$\langle \Delta M \rangle \approx \frac{1}{b} \log_{10}(N/N_A) \tag{3}$$

where the total number of earthquakes in a catalogue N is an observable and the number of aftershocks N_A can be forecast/estimated from the Omori’s law parameters. Typical estimates of Omori and GR-MFD parameters yields estimates of the average magnitude difference (ΔM) that are similar to the empirical observation of $\sim 1.2 M$ ^{27,28}.

Deriving a trailing earthquake model for induced seismicity

Following a similar rationale, we derive a trailing earthquake model from the statistical viewpoint of induced earthquakes as a point-process. We first assume that induced earthquake magnitudes are independent and identically distributed random variables that follow the GR-MFD. This assumption is supported by numerous cases that obeying this distribution^{2,29,30}, similar to naturally occurring tectonic sequences. From this assumption an expression for the distribution of the maximum magnitude earthquake follows (in the limit of large N).

$$M_{\max} \approx M_C + \frac{1}{b} \log_{10}(N) - \frac{1}{b} \log_{10}(-\ln(u)) \tag{4}$$

This is often referred to as the sample size hypothesis²⁹ for the distribution of maximum magnitudes M_{\max} , where the newly introduced parameters are the magnitude of completeness M_C ³¹ and a confidence level u that can be utilized to sample the M_{\max} distribution. The distribution of expected maximum magnitudes for different population sizes is shown in Fig. 1a.

Based on this sample size hypothesis, we can consider an archetype of an induced earthquake sequence that is broken into two segments: the injection/stimulation phase (denoted by a subscript S) and the trailing shut-in phase (denoted by a subscript T) that follows (Fig. 1c). The stimulation phase would entail a population count of N_S stimulation earthquakes that occur until the anthropogenic operation is halted at time T_S ,

Model name [References]	Aftershock rate decay	Cumulative aftershock count	Total aftershock count
Modified Omori ³⁹	$n^o(t) = K(t + c)^{-p}$	$N^o(t) = \left(\frac{K}{p-1}\right)(c^{1-p} - (t + c)^{1-p})$	$N^o(\infty) = K \frac{c^{1-p}}{p-1}$
Exponential ⁴⁰	$n^e(t) = n^e(0)e^{-t/\tau}$	$N^e(t) = \tau n^e(0)(1 - e^{-t/\tau})$	$N^e(\infty) = \tau n^e(0)$
Stretched exponential ⁴¹	$n^s(t) = qN^* e^{\left(\frac{d}{t_0}\right)^q} \left(\frac{1}{t+d}\right)^q e^{-\left(\frac{t+d}{t_0}\right)^q}$	$N^s(t) = N^* \left(1 - e^{\left(\frac{d}{t_0}\right)^q} e^{-\left(\frac{t+d}{t_0}\right)^q}\right)$	$N^s(\infty) = N^*$
Cut-off Power Law ⁴²	$n^c(t) = K(t + c)^{-p} e^{-t/\tau}$	$N^c(t) = K\tau^{1-p} e^{c/\tau} (\Gamma(1 - p, \frac{c}{\tau}) - \Gamma(1 - p, \frac{t+c}{\tau}))$	$N^c(\infty) = K\tau^{1-p} e^{c/\tau} \Gamma(1 - p, \frac{c}{\tau})$
Gamma ⁴³	$n^\gamma(t) = A t^{-q} (\gamma(q, \lambda_b t) - \gamma(q, \lambda_a t))$	$N^\gamma(t) = \left(\frac{A}{q-1}\right) \left[\lambda_b^{q-1} - \lambda_a^{q-1}\right] - \left[-\lambda_a^{q-1} e^{-\lambda_a t} + t^{1-p} \Gamma(q, \lambda_a t) + \lambda_b^{q-1} e^{-\lambda_b t} - t^{1-p} \Gamma(q, \lambda_b t)\right]$	$N^\gamma(\infty) = \left(\frac{A}{q-1}\right) (\lambda_b^{q-1} - \lambda_a^{q-1})$
Epidemic type ⁴⁴	$n^E(t) = \sum_i K_i (t - t_i + c)^{-p}$	$N^E(t) = \sum_i \left(\frac{K_i}{p-1}\right) (c^{1-p} - (t - t_i + c)^{1-p})$	$N^E(\infty) = \sum_i K_i \frac{c^{1-p}}{p-1}$

Table 1. Aftershock models proposed for tectonic earthquakes. For each model we provide an expression for the aftershock rate function $n(t)$, the cumulative aftershock count function $N(t)$, and the total number of aftershocks $N(\infty)$. Corresponding references are also provided for each model. Note that $\gamma(a, x)$ and $\Gamma(a, x)$ are the lower and upper incomplete gamma functions, respectively. Superscripts are used to denote the aftershock decay model used.

lagged by stimulation events for time ΔT , and followed by aftershock-like events (starting at an earthquake rate approximately equal to the stimulation rate). We consider trailing seismicity to be those N_T events that occur after time T_S i.e., the sum of lagged stimulation events and the aftershock-like events. Taking the difference of the maximum magnitudes (Eq. 4) for both segments and the stimulation segment yields the relationship for the trailing magnitude difference below.

$$\Delta M = M_{\max} - M_{\max,S} \approx \frac{1}{b} \log_{10} \left(\frac{1}{R_S}\right) + \frac{1}{b} \log_{10} \left(\frac{\ln(u_S)}{\ln(u)}\right) \tag{5}$$

This produces a form for the differential maximum magnitude that is reminiscent of Båth’s law (Eq. 3), where $1/R_S = N/N_S = 1 + N_T/N_S = 1 + R_{TS}$. Here we note that ΔM and R_S are physically observable quantities that may be subjected to hypothesis testing.

In continuing to develop our model, we assume a further constraint on the stimulated and trailing earthquake counts N_S and N_T . Similar to the prior logic in Båth’s law, we bring additional quantification to the value of R_S . First, the stimulated earthquake count N_S has been suggested to have a linear relationship with injection volume³².

$$N_S = V(t) 10^\Sigma 10^{-bM} \tag{6}$$

This expression effectively modifies the a -value of the GR-MFD to have an explicit dependence on the volume injected $V(t)$ (in m^3) and the background potential for earthquakes i.e., the seismogenic index Σ . This relationship assumes a non-homogenous Poisson process for the stimulation events³³ and has been well fit to numerous induced earthquake sequences^{29,34,35}. Second, the total aftershock count is the integration of an aftershock rate decay function from the time of the mainshock to infinite time. In practice, induced sequences have often been observed to fit aftershock-like rate decay sequences following shut-in³⁶. While there is evidence to suggest a linear dependence between injected volume and cumulative earthquake counts²⁹, we note that alternative models^{7,37} could be utilized without changing the results of our study. Because there is an abundance in the choice of aftershock rate decay model, we summarize the prominent ones in Table 1. For each of the aftershock rate decay models we provide the rate decay function $n(t)$, the cumulative count function $N(t)$, and the total aftershock count $N(\infty)$. We also refer the reader to prior summaries of tectonic aftershock models³⁸. In this study we included the most general version of aftershock models that (with parameter constraints) lead to finite cumulative event counts $N(\infty)$ and are widely adopted by the seismological community.

We can expand upon Eq. (5) by imposing our archetypal induced earthquake sequence (Fig. 1c) in combination with the aftershock and stimulation models. Combining these models with the archetypal sequence allows for closed form expression of R_{TS} , and thus the distribution of the expected maximum earthquake magnitude.

$$R_{TS}^o = \frac{N_T}{N_S} = \frac{\dot{V}(t) \Delta T 10^\Sigma + K \frac{c^{1-p}}{p-1}}{\dot{V}(t) T_S 10^\Sigma} = \frac{\dot{V}(t) \Delta T 10^\Sigma + f \dot{V}(t) 10^\Sigma \frac{c}{p-1}}{\dot{V}(t) T_S 10^\Sigma} = \frac{\Delta T + f \frac{c}{p-1}}{T_S} \tag{7}$$

$$R_{TS}^e = \frac{N_T}{N_S} = \frac{\dot{V}(t) \Delta T 10^\Sigma + \tau n^e(0)}{\dot{V}(t) T_S 10^\Sigma} = \frac{\dot{V}(t) \Delta T 10^\Sigma + f \dot{V}(t) 10^\Sigma \tau}{\dot{V}(t) T_S 10^\Sigma} = \frac{\Delta T + f \tau}{T_S} \tag{8}$$

Here, R_{TS} has been described for both the Modified Omori (R_{TS}^o , Eq. 7) and Exponential (R_{TS}^e , Eq. 8) aftershock rate decay models. Note that we’ve assumed the initial aftershock rate is some fraction f of the ending stimulation rate—to account for slight increases or decreases in earthquake rate after the lagged shut-in response.

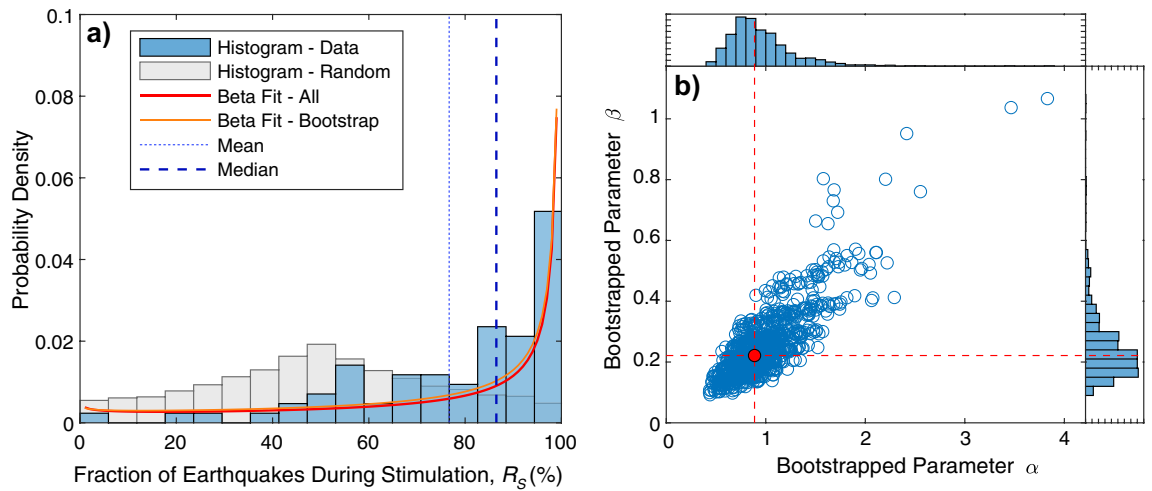


Figure 2. The empirical distribution of R_S and bootstrapped beta distribution fit. **(a)** Histogram of R_S for all the induced seismicity cases (blue bars) with its median and mean values (vertical lines), best fit to the beta distribution (crimson lines), compared against a random distribution (grey bars). **(b)** Scatterplot of bootstrapped beta distribution parameters (blue circles) with the best fit parameter (red circle). Inset histograms show the univariate distribution of beta parameters.

Testing the model's validity

In adapting Båth's law to induced seismicity, we've made several assumptions (Fig. 1): that the sequence is described by a nonhomogenous Poisson process, that the extended fault system can effectively be modelled as a point-process, that there's a single lag time between operations and seismogenic response, that statistical parameters are stationary during the observation interval, and that there are no injection-volume or tectonic limits on the maximum magnitude. To validate (or refute) these assumptions, we begin by examining induced seismicity cases.

We focus only on cases where there has been a clearly delineated end of operations (Fig. S1). This criterion dominantly encompasses cases related to hydraulic fracturing² or enhanced geothermal systems⁴⁵. We select cases of induced seismicity where both catalogue and injection time series information is available. Cases include the Duvernay^{11,13,46,47} and Montney⁴⁸ in Canada, the Bowland Shale in the UK^{49,50}, the Appalachian Basin^{51,52}, SCOOP & STACK play^{53–55}, Eagle Ford⁵⁶, and Arkoma Basin⁵⁷ in the United States, the Sichuan Basin in China^{6,58,59}, Soultz-sous-Forêts in France^{60,61}, Cooper Basin⁶² and Paralana⁶³ in Australia, KTB⁶⁴ and Groß Schönebeck⁶⁵ in Germany, Basel⁶⁶ and St Gallen⁶⁷ in Switzerland, Helsinki in Finland⁶⁸, Pohang in South Korea^{5,69}, and Berlin in El Salvador⁷⁰. Many of these cases have data available in online repositories⁷¹, within aggregate comparison studies^{12,29,72}, or as supplements to the individual case studies.

Briefly, individual cases are truncated at their magnitude of completeness (M_C) and then subdivided into two populations: those occurring during stimulation or those trailing stimulation. We report statistics such as earthquake counts, largest magnitudes, b -value, and M_C , which we summarize in Table S1. For additional details on the data acquisition and preprocessing, we direct the reader to the Supplementary Information (Section S1).

Examining the Statistics of R_S . From our compiled database (Table S1), we examine the parameter R_S observed during numerous induced seismicity sequences (Fig. 2a). R_S shows a highly skewed distribution—tending towards values of 77% on average and 86% as a median (Fig. S5). This is comparable to observations in prior studies on hydraulic fracturing that found ~90% of events occurred during stage stimulation in the Duvernay³⁵ and ~87% of events in the Sichuan Basin were 'forced' under a modified Epidemic Type Aftershock Model^{6,58}. While the distribution is skewed towards high R_S values, we also note that it is heavy-tailed with a handful of cases with low R_S values. For example, the enhanced geothermal stimulation at Pohang shows a large trailing aftershock sequence^{5,69}.

To further examine the skewed nature of the R_S distribution, we compare our empirical observations to the distribution expected from a random process (Fig. 2a). To model the random process, we consider a uniform random variable for both N_T and N_S . We then perform the two-sample Kolmogorov–Smirnov test between our empirically observed R_S and the random process model. The two-sample Kolmogorov–Smirnov test is a non-parametric hypothesis test that assesses the probability that two samples are drawn from the same (arbitrary) distribution⁷³. In performing this test, we find extremely small p values that are orders of magnitude less than 5% (Fig. S6)—which rejects the null hypothesis that the two samples are from the same distribution. The failure of R_S to be modelled as a random process suggests that there is a physical process underlying this distribution (e.g., Eqs. 7 and 8).

A better representation of the statistics of the R_S would be useful for estimation of differential magnitudes ΔM (e.g., Eq. 5). Towards this end, we fit our histogram results to a beta distribution—a two-parameter (α and β) family of distributions defined over the interval 0–1⁷⁴. We select the beta distribution due to its flexibility in fitting probability density functions over this interval. Further, to capture the variability in the fitted parameters

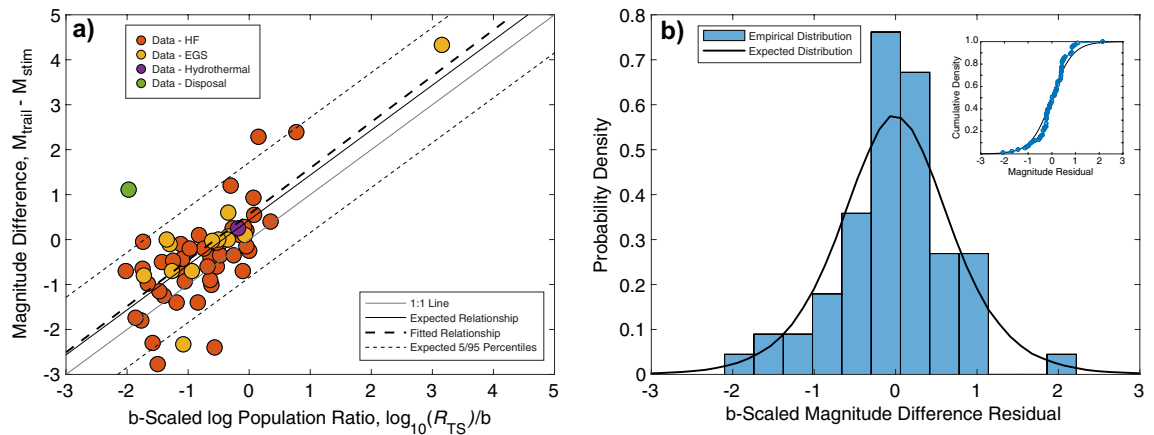


Figure 3. Testing Båth's law for induced seismicity. (a) Linear regression of magnitude differences ΔM versus logarithmic population ratios $\frac{1}{b} \log_{10}(R_{TS})$: data is separated by anthropogenic type (circles) and shown alongside the 1:1 line (solid gray line), the relationship expected by Båth's law (solid black line), the regression fit to data (dashed line), and the corresponding 5/95 percentiles (dotted line). (b) The residuals to the expected Båth's law relationship are shown as probability density function for the expected (solid black line) and empirically derived from data (blue bars). Inset panel shows the cumulative density function for the expected (solid black line) and empirically derived residuals (blue circles/line).

we performed a 1000-trial weighted bootstrap analysis (see Supplementary Information Section S3 for more details). We find that the mean best-fit parameters for α and β are 1.016 and 0.271, respectively. Similarly, the median best-fit parameters for α and β are 0.940 and 0.238, respectively. The covariances ($\sigma_{\alpha\alpha}$, $\sigma_{\beta\beta}$, $\sigma_{\alpha\beta}$) for these parameter fits are 0.1151, 0.0130, and 0.0317 (Fig. 2b). In performing sensitivity tests, we do not find that fitted parameters differ appreciably, when comparing hydraulic fracturing or enhanced geothermal system subsets. However, larger datasets could be useful to discern systematic differences in R_S according to operation and setting. These fitted R_S distribution results could constrain trailing seismicity differential magnitudes ΔM , provided the validity of using R_S in Båth's law for induced seismicity was established.

Testing Båth's law for induced seismicity. With constraints on the empirical distribution of R_S , we examine the statistical validity of Båth's law for induced seismicity. Equation (5) predicts that there is a linear relationship between the trailing-stimulation maximum magnitude difference ΔM (left hand side of equation) and the b -value scaled logarithm of the population ratio $\frac{1}{b} \log_{10}(R_{TS})$ (first term of the right-hand side), where the remaining term describes the statistical distribution of the residuals (second term of right-hand side).

Based on this, we first consider a weighted linear regression applied to our dataset (Fig. 3a). We observe a reasonable agreement between observation and expectation. The fitted parameters for the linear regression are 1.06 ± 0.16 for the slope and 0.58 ± 0.18 for the y-intercept. The theoretically expected slope of 1.0 agrees within error of our regression-fitted slope (see Supplementary Information Section S4 for more details). The observation of a y-intercept term is the result of a correction term that conditions on the sample count, when there is an unknown true sample count (see van der Elst et al.²⁹ Sections 2.2 & A4 for additional details). The value of this correction term approaches $\frac{1}{b} \log_{10}(e)$ (i.e., ~ 0.43) as the population count N becomes large, which agrees within error of our regression fitted parameter. While both the empirically derived slope and y-intercept agree within error to their expected values, we note that the best estimate of both values are higher than those expected; this larger value bias could be the result of a publication bias where scientific studies preferentially describe cases with 'exceptional behavior', like large magnitudes, red-light events, or lingering sequences of trailing seismicity.

The second test we consider is an examination of the distribution of the Båth's law residuals (Fig. 3b). In this test, we rearrange the terms in Eq. (5) such that just the power law distribution remains: i.e., $\log_{10} \left(\frac{\ln(us)}{\ln(ur)} \right)$. Note that individual correction terms are accounted for using the true population size, rather than the large N limit. We find a correspondence between the histogram of data residuals and the expected probability density function for Båth's law. More quantitatively, we compare the expected and observed residual samples using the two-sample Kolmogorov–Smirnov test. We find a p value ($\sim 45\%$) that is larger than 5%, meaning we are unable to reject the null hypothesis that the two samples are from the same distribution. Most bootstrapped perturbations ($\sim 61\%$) support this result (Fig. S7). Of course, additional high-resolution cases of induced seismicity would be helpful for further scrutiny. The observation that induced earthquakes trailing the stimulation follow Båth's law, both for expected magnitudes versus population counts, and for their residual distribution suggests that this approach should be useful for evaluating trailing seismicity.

Discussion

In this section we discuss the interpretations and implications of our results. We divide the discussion into three parts: (1) the applicability of Båth's law for induced seismicity, (2) the relative importance of parameters in Båth's law, and (3) a diagnostic test for operators to use in situ when encountering seismogenic faults.

The applicability of Båth's law for induced seismicity. To date, many models have been introduced to capture the anticipated maximum magnitude of induced earthquakes from either physical or statistical perspectives^{7,29,37}. Typically, these models have focused on the earthquakes occurring during the stimulation process. In this paper, we have described a statistical framework to extend these considerations to those earthquakes that trail after well shut-in based on Båth's law for tectonic earthquakes. This conceptual extension is partly driven by the observation of large earthquake magnitude jumps¹³, high-profile cases of large trailing events^{5,11,12,48}, and the contribution of trailing events to earthquake risk management⁹.

Our results suggest that Båth's law is a viable foundation for anticipating the magnitude of earthquakes trailing an operation that produces anthropogenic seismicity. This assertion is supported by the observation that: (1) induced seismicity cases are well fit by the linear relationship between magnitude differences ΔM and logarithmic population ratios $\frac{1}{b} \log_{10}(R_{TS})$ that is predicted by Båth's law (Fig. 3a), and (2) the residuals from this relationship are approximately consistent with the power law distribution anticipated by Båth's law (Fig. 3b). Based on this, we further suggest that the description of Båth's law (Eq. 5) will have predictive power for quantitative forecasts of trailing seismicity. Of course, this is predicated on knowing/estimating the population ratios R_{TS} —either a priori or updating in real-time during an operation. In this sense, the empirical distribution of R_S and its best fit to a beta distribution (Fig. 2) provides some utility—possibly as prior information in a Bayesian updating scheme. Much of the literature that has been described for tectonic aftershock forecasts through extreme value statistics and Bayesian approaches^{22,75}, will likely be adaptable for the purpose of constraining N_T . Additionally, models of stimulation seismicity that consider the full injection history should provide better predictability to the anticipated stimulation events N_S ⁵⁰.

This interpretation is in line with recent work that has found that induced earthquake sequences have statistical properties that are similar to (or subtle modifications of) their tectonic counterparts^{29,36}. These results could be indicative of the importance of fluid-flow along faults for tectonic earthquakes, which has been suggested in some studies as important factor for the aftershock process^{76,77}. General links between statistical models and physical processes have been made for earthquakes and aftershocks⁷⁸. Thus, incorporating and testing physical models of aftershocks that consider realistic fault geometries, fault rheology, stress, material heterogeneity, and fluid flow will likely be important for better understanding post shut-in induced earthquakes.

Relative importance of Båth's law parameters. A better understanding the physical processes that underlie the observed statistics would allow for better understanding of trailing seismicity. While the collected dataset (Table S1) is too small to meaningfully discern systematic trends in R_S that vary with anthropogenic type, operational parameters, fault architecture, or geological setting, we may still apply our simple model to better understand the relative importance of parameters that contribute to a skewed distribution of R_S . We focus first on the exponential decay model (Eq. 8) both for its simplicity and the breadth of induced seismicity studies it has been applied to^{79,80}. If we assume representative values for a hydraulic fracturing operation ($\Delta T = 1$ h, $\tau = 3$ days, $f = 1.0$, $T_S = 2$ weeks), we find that $R_S = 82\%$; a value between those reported for the mean/median of the empirical R_S . To build on this further, we assume that the parameters ΔT and τ follow a lognormal distribution while f follows a Gaussian distribution so that we may compute a synthetic R_S distribution. From this, we perform a Markov Chain Monte Carlo (MCMC) optimization to maximize the Kolomogorov–Smirnov p value between the empirical and synthetic R_S . The MCMC approach randomly samples the parameter space to find optimal mean and standard deviations of these parameters (ΔT , τ , f) that are bounded within those reasonably observed in the literature^{2,79}.

The MCMC approach produces synthetic R_S distributions that fit the empirical distribution well (likelihoods of up to 90%). The input parameters distributions needed to fit R_S well (Fig. 4) suggests that a skewed distribution of input parameters (ΔT or τ) is required to match the skewed R_S distribution. Such a skewed distribution includes chances for black swan occurrences of extreme value events. In practice, this is what is observed in cases of induced seismicity: relatively rare events of large consequence. For example, the earthquakes at Pohang were exceptional in that a dormant fault system was reactivated by the geothermal stimulation program⁵. Similarly, most events in a 2018 episode in British Columbia occurred as part of a delayed trigger after hydraulic fracturing stage completion⁴⁸. The largest event during the second well stimulation at Preston New Road occurred 60 h after the end of the last stimulation⁵⁰. Earthquakes were apparently induced at distances of up to ~ 5 km away from the enhanced geothermal stimulation at Strasbourg–Vendenheim⁸¹. These consequential trailing seismicity events suggest the importance of better modelling and constraining parameters relevant to Båth's law. We note, however, that there are significant limitations to our MCMC approach, as the distributions of ΔT and τ trade-off with one another. Thus, empirical observations of aftershock parameters (τ or c and p) and injection responses (Σ , b , ΔT) would be needed to better constrain these values/distributions.

The Seismogenic fault injection test (SFIT). Previously we showed that skewed values of ΔT or τ are required for a skewed R_S —and represent an increased potential for large trailing magnitude differences ΔM . Given the possibility of consequential impacts from large magnitude trailing events, it would be prudent to measure these parameters in situ to reduce the uncertainties associated with fault reactivation.

Similar to Diagnostic Fracture Injection Tests⁸² where fluids are intentionally injected into a reservoir to constrain its hydrological properties and stress state, we suggest an analogous Seismogenic Fault Injection Test (SFIT) to calibrate the seismic response to fluid injection (Fig. 5). Ideally, an inventory of faults would be available from a pre-assessment to injection. Injection would be performed with the intention of measuring both the time-delayed response to injection (Σ , b , ΔT) and the aftershock-like decay of seismicity (τ or c and p) for each of the known faults⁷⁹. The SFIT should be designed to constrain parameters of greatest consequence (ΔT and τ). In the context of our study, unknown faults activated at distance from the stimulation interval (e.g., large

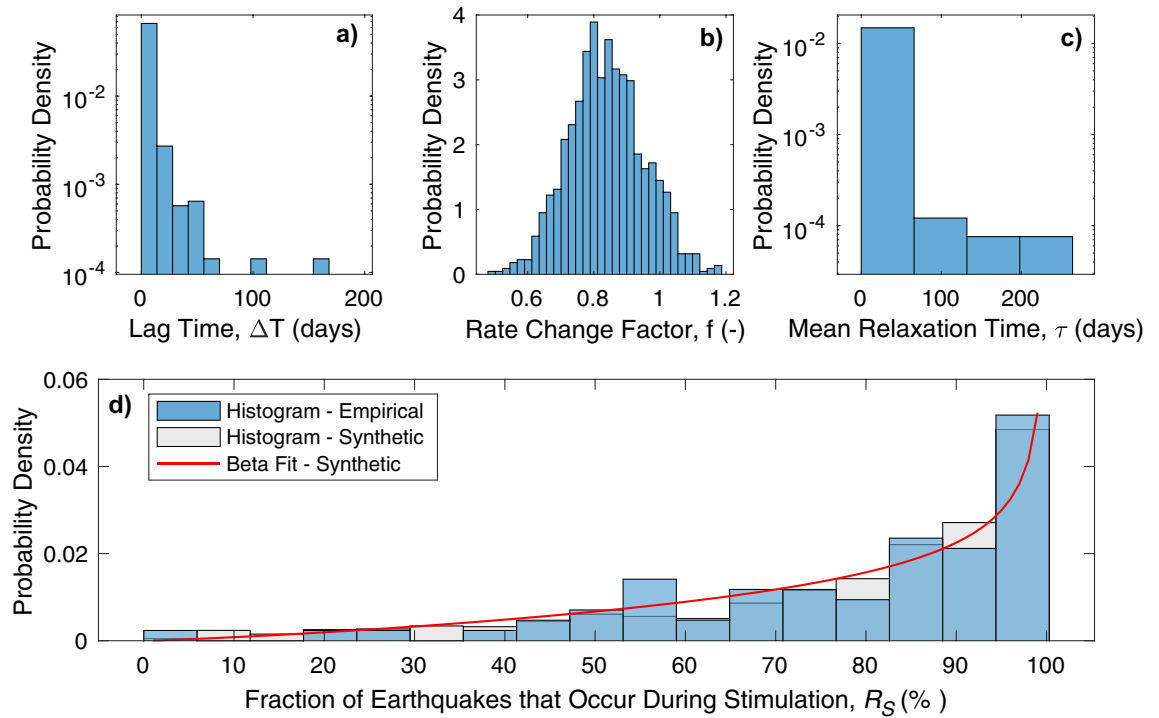


Figure 4. Synthetically fitting the R_S distribution to model parameters. (d) The archetypal model for trailing seismicity ratios (Eq. 8) is fit (grey bars) to the empirical distribution of R_S (blue bars). The distribution of model parameters (a) ΔT , (b) f , and (c) τ are also shown.

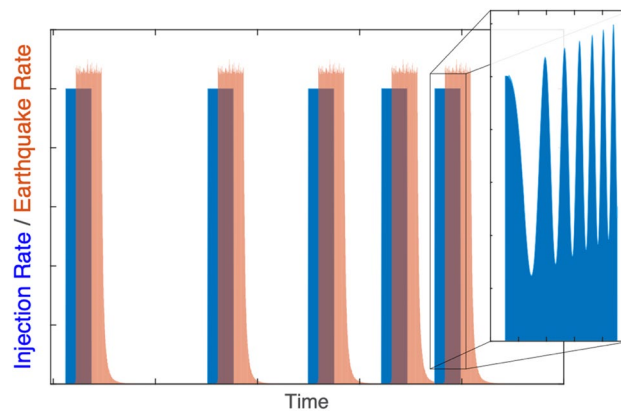


Figure 5. The Seismogenic Fault Injection Test (SFIT). Schematic diagram of a diagnostic test to measure a fault's seismogenic response through parameters such as Σ , b , ΔT , and τ (or c and p). Repeated injection tests (at non-periodic intervals) are performed to verify the accuracy of measurements. Inset diagram shows how individual tests could be 'chirped' in their injection rate to better constrain parameters.

ΔT) or with prolonged trailing sequences (e.g., large τ) pose the greatest risks. With this in mind, an individual injection phase would span a range of injection rates, potentially sweeping through them cyclically in a 'chirped' manner (Fig. 5); the intent would be to both discern any rate-dependent features and better highlight potential time lags ΔT . As well, individual injection phases should be repeated to bolster confidence in the measurements, in particular the decay rate τ which only is well constrained after shut-in⁷⁹. Ideally, these repeated injections phases would be spaced at non-periodic intervals to unambiguously measure delay times ΔT ^{83,84}. Finally, if we are making suggestions outside of what the simple models predict, varying the shut-in procedure for individual injection phases would be useful to discern safer responses to unwanted seismicity.

Variable	Description
a	GR-MFD productivity parameter, a -value
b	GR-MFD slope parameter, b -value
$M, \Delta M$	Earthquake magnitude and magnitude difference, respectively
M_C	Magnitude of completeness
Σ	Seismogenic Index
$V(t), \dot{V}(t)$	Cumulative injection volume and injection rate, respectively
$n(t)$	Earthquake rate
N, N_A, N_T, N_S	Cumulative earthquake counts: total, aftershocks, trailing, and stimulation
R_S, R_{TS}	Earthquake count ratios: related by $1/R_S = N/N_S = 1 + N_T/N_S = 1 + R_{TS}$
u, u_T, u_S	Confidence level variables
ΔT	Simple archetype model, stimulation-earthquake response delay
T_S	Stimulation/injection time interval
f	Stimulation-trailing rate change factor
K	Omori aftershock productivity parameter
p	Omori aftershock decay exponent
c	Omori aftershock singularity parameter
τ	Exponential aftershock mean decay time

Table 2. Summary of equation parameters and variables. Here we list the relevant parameters used in equations throughout our study, for convenience to the reader.

Conclusions

In summary, we have tested the applicability of Båth's law to cases of induced seismicity with clear ends to fluid injection. Båth's law fits the expected relationship, as confirmed by a linear regression of the collected data. Additionally, the residuals to the data follow the expected power law distribution, as confirmed by Kolomogorov–Smirnov tests. Thus, our compilation of population ratio R_S statistics, including a fit to the beta distribution, could be useful for forecasting trailing seismicity. Finally, simple point-like models describing R_S highlighted parameters of interest (ΔT and τ); we propose in situ tests to help quantify the risks associated with seismicity trailing an anthropogenic operation. Variables used throughout the paper are summarized in Table 2.

Data availability

Data used to produce the results of this study are available as an electronic supplement (Table S1). In addition, the codes used to derive these results are available online, through GitHub (<https://github.com/RyanJamesSchultz/IS-Bath>). Catalogue data and injection data are available in online repositories⁷¹, within aggregate comparison studies^{12,29,72}, or as supplements to the individual case studies.

Received: 21 October 2021; Accepted: 7 January 2022

Published online: 24 January 2022

References

- Ellsworth, W. L. Injection-induced earthquakes. *Science* <https://doi.org/10.1126/science.1225942> (2013).
- Schultz, R. *et al.* Hydraulic fracturing-induced seismicity. *Rev. Geophys.* **58**(3), e2019RG000695. <https://doi.org/10.1029/2019RG000695> (2020).
- Keranen, K. M., Savage, H. M., Abers, G. A. & Cochran, E. S. Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 Mw 5.7 earthquake sequence. *Geology* **41**(6), 699–702. <https://doi.org/10.1130/G34045.1> (2013).
- Van der Voort, N. & Vanclay, F. Social impacts of earthquakes caused by gas extraction in the Province of Groningen, The Netherlands. *Environ. Impact Assess. Rev.* **50**, 1–15. <https://doi.org/10.1016/j.eiar.2014.08.008> (2015).
- Ellsworth, W. L., Giardini, D., Townend, J., Ge, S. & Shimamoto, T. Triggering of the Pohang, Korea, earthquake (M w 5.5) by enhanced geothermal system stimulation. *Seismol. Res. Lett.* **90**(5), 1844–1858. <https://doi.org/10.1785/0220190102> (2019).
- Lei, X., Wang, Z. & Su, J. The December 2018 ML 5.7 and January 2019 ML 5.3 earthquakes in South Sichuan Basin induced by shale gas hydraulic fracturing. *Seismol. Res. Lett.* **90**(3), 1099–1110. <https://doi.org/10.1785/0220190029> (2019).
- McGarr, A. Maximum magnitude earthquakes induced by fluid injection. *J. Geophys. Res. Solid Earth* **119**(2), 1008–1019. <https://doi.org/10.1002/2013JB010597> (2014).
- Lee, K. K. *et al.* Managing injection-induced seismic risks. *Science* **364**(6442), 730–732. <https://doi.org/10.1126/science.aax1878> (2019).
- Schultz, R., Beroza, G. C. & Ellsworth, W. L. A risk-based approach for managing hydraulic fracturing-induced seismicity. *Science* **372**(6541), 504–507. <https://doi.org/10.1126/science.abg5451> (2021).
- Segall, P. & Lu, S. Injection-induced seismicity: Poroelastic and earthquake nucleation effects. *J. Geophys. Res. Solid Earth* **120**(7), 5082–5103. <https://doi.org/10.1002/2015JB012060> (2015).
- Schultz, R., Wang, R., Gu, Y. J., Haug, K. & Atkinson, G. A seismological overview of the induced earthquakes in the Duvernay play near Fox Creek, Alberta. *J. Geophys. Res. Solid Earth* **122**(1), 492–505. <https://doi.org/10.1002/2016JB013570> (2017).
- Verdon, J. P. & Bommer, J. J. Green, yellow, red, or out of the blue? An assessment of traffic light schemes to mitigate the impact of hydraulic fracturing-induced seismicity. *J. Seismol.* <https://doi.org/10.1007/s10950-020-09966-9> (2020).
- Schultz, R. & Wang, R. Newly emerging cases of hydraulic fracturing induced seismicity in the Duvernay East Shale Basin. *Tectonophysics* **779**, 228393. <https://doi.org/10.1016/j.tecto.2020.228393> (2020).

14. Båth, M. Lateral inhomogeneities of the upper mantle. *Tectonophysics* **2**(6), 483–514. [https://doi.org/10.1016/0040-1951\(65\)90003-X](https://doi.org/10.1016/0040-1951(65)90003-X) (1965).
15. Zhuang, J. *et al.* Basic models of seismicity: spatiotemporal models. *Community Online Resour. Stat. Seism. Anal.* **1**, 20. <https://doi.org/10.5078/corssa-07487583> (2011).
16. Zhuang, J., Harte, D., Werner, M. J., Hainzl, S. & Zhou, S. Basic models of seismicity: temporal models. *Community Online Resour. Stat. Seism. Anal.* **1**, 42. <https://doi.org/10.5078/corssa-79905851> (2012).
17. Vere-Jones, D. A note on the statistical interpretation of Båth's law. *Bull. Seismol. Soc. Am.* **59**(4), 1535–1541 (1969).
18. Lombardi, A. M. Probabilistic interpretation of Båth's law. *An. Geophys.* <https://doi.org/10.4401/ag-3520> (2002).
19. Helmstetter, A. & Sornette, D. Båth's law derived from the Gutenberg–Richter law and from aftershock properties. *Geophys. Res. Lett.* <https://doi.org/10.1029/2003GL018186> (2003).
20. Sornette, D. & Werner, M. J. Constraints on the size of the smallest triggering earthquake from the epidemic-type aftershock sequence model, Båth's law, and observed aftershock sequences. *J. Geophys. Res. Solid Earth* <https://doi.org/10.1029/2005JB003621> (2005).
21. Omori, F. On After-shocks of Earthquakes. *J. Coll. Sci. Imp. Univ. Tokyo* **7–2**, 113–200 (1894).
22. Reasenber, P. A. & Jones, L. M. Earthquake hazard after a mainshock in California. *Science* **243**(4895), 1173–1176. <https://doi.org/10.1126/science.243.4895.1173> (1989).
23. Utsu, T. & Ogata, Y. The centenary of the Omori formula for a decay law of aftershock activity. *J. Phys. Earth* **43**(1), 1–33. <https://doi.org/10.4294/jpe1952.43.1> (1995).
24. Kisslinger, C. The stretched exponential function as an alternative model for aftershock decay rate. *J. Geophys. Res. Solid Earth* **98**(B2), 1913–1921. <https://doi.org/10.1029/92JB01852> (1993).
25. Ishimoto, M. & Iida, K. Observations of earthquakes registered with the microseismograph constructed recently. *Bull. Earthq. Res. Inst. Univ. Tokyo* **17**, 443–478 (1939).
26. Gutenberg, B. & Richter, C. F. Earthquake magnitude, intensity, energy, and acceleration. *Bull. Seismol. Soc. Am.* **46**(2), 105–145 (1956).
27. Shcherbakov, R. & Turcotte, D. L. A modified form of Bath's law. *Bull. Seismol. Soc. Am.* **94**(5), 1968–1975. <https://doi.org/10.1785/012003162> (2004).
28. Shcherbakov, R., Turcotte, D. L. & Rundle, J. B. Aftershock statistics. *Pure Appl. Geophys.* **162**(6–7), 1051–1076. <https://doi.org/10.1007/s00024-004-2661-8> (2005).
29. Van der Elst, N. J., Page, M. T., Weiser, D. A., Goebel, T. H. & Hosseini, S. M. Induced earthquake magnitudes are as large as (statistically) expected. *J. Geophys. Res. Solid Earth* **121**(6), 4575–4590. <https://doi.org/10.1002/2016JB012818> (2016).
30. Wilson, M. P., Foulger, G. R., Gluyas, J. G., Davies, R. J. & Julian, B. R. HiQuake: The human-induced earthquake database. *Seismol. Res. Lett.* **88**(6), 1560–1565. <https://doi.org/10.1785/0220170112> (2017).
31. Mignan, A. & Woessner, J. Estimating the magnitude of completeness for earthquake catalogs. *Community Online Resour. Stat. Seism. Anal.* <https://doi.org/10.5078/corssa-00180805> (2012).
32. Shapiro, S. A., Dinske, C., Langenbruch, C. & Wenzel, F. Seismogenic index and magnitude probability of earthquakes induced during reservoir fluid stimulations. *Lead. Edge* **29**(3), 304–309. <https://doi.org/10.1190/1.3353727> (2010).
33. Langenbruch, C., Dinske, C. & Shapiro, S. A. Inter event times of fluid induced earthquakes suggest their Poisson nature. *Geophys. Res. Lett.* <https://doi.org/10.1029/2011GL049474> (2011).
34. Langenbruch, C. & Zoback, M. D. How will induced seismicity in Oklahoma respond to decreased saltwater injection rates?. *Sci. Adv.* **2**(11), e1601542. <https://doi.org/10.1126/sciadv.1601542> (2016).
35. Schultz, R., Atkinson, G., Eaton, D. W., Gu, Y. J. & Kao, H. Hydraulic fracturing volume is associated with induced earthquake productivity in the Duvernay play. *Science* **359**(6373), 304–308. <https://doi.org/10.1126/science.aao0159> (2018).
36. Langenbruch, C. & Shapiro, S. A. Decay rate of fluid-induced seismicity after termination of reservoir stimulations post injection seismicity. *Geophysics* **75**(6), MA53–MA62. <https://doi.org/10.1190/1.3506005> (2010).
37. Galis, M., Ampuero, J. P., Mai, P. M. & Cappa, F. Induced seismicity provides insight into why earthquake ruptures stop. *Sci. Adv.* <https://doi.org/10.1126/sciadv.aap7528> (2017).
38. Mignan, A. Modeling aftershocks as a stretched exponential relaxation. *Geophys. Res. Lett.* **42**(22), 9726–9732. <https://doi.org/10.1002/2015GL066232> (2015).
39. Utsu, T. A statistical study on the occurrence of aftershocks. *Geophys. Mag.* **30**, 521–605 (1961).
40. Burridge, R. & Knopoff, L. Model and theoretical seismicity. *Bull. Seismol. Soc. Am.* **57**(3), 341–371 (1967).
41. Gross, S. J. & Kisslinger, C. Tests of models of aftershock rate decay. *Bull. Seismol. Soc. Am.* **84**(5), 1571–1579 (1994).
42. Otsuka, M. Studies on aftershock sequences—Part 1. Physical interpretation of Omori's formula. *Sci. Rep. Shimabara Earthq. Volcano Obs.* **13**, 11–20 (1985) (In Japanese).
43. Narteau, C., Shebalin, P. & Holschneider, M. Temporal limits of the power law aftershock decay rate. *J. Geophys. Res. Solid Earth* **107**(B12), ESE-12. <https://doi.org/10.1029/2002JB001868> (2002).
44. Ogata, Y. Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.* **83**(401), 9–27. <https://doi.org/10.1080/01621459.1988.10478560> (1988).
45. Olasolo, P., Juárez, M. C., Morales, M. P. & Liarte, I. A. Enhanced geothermal systems (EGS): A review. *Renew. Sust. Energ.* **56**, 133–144. <https://doi.org/10.1016/j.rser.2015.11.031> (2016).
46. Bao, X. & Eaton, D. W. Fault activation by hydraulic fracturing in western Canada. *Science* **354**(6318), 1406–1409. <https://doi.org/10.1126/science.aag2583> (2016).
47. Schultz, R. *et al.* The Cardston earthquake swarm and hydraulic fracturing of the Exshaw Formation (Alberta Bakken play). *Bull. Seismol. Soc. Am.* **105**(6), 2871–2884. <https://doi.org/10.1785/0120150131> (2015).
48. Castro, P. *et al.* Stress chatter via fluid flow and fault slip in a hydraulic fracturing-induced earthquake sequence in the Montney Formation, British Columbia. *Geophys. Res. Lett.* **47**(14), e2020GL087254. <https://doi.org/10.1029/2020GL087254> (2020).
49. Clarke, H., Eisner, L., Styles, P. & Turner, P. Felt seismicity associated with shale gas hydraulic fracturing: The first documented example in Europe. *Geophys. Res. Lett.* **41**(23), 8308–8314. <https://doi.org/10.1002/2014GL062047> (2014).
50. Mancini, S., Werner, M. J., Segou, M. & Baptie, B. Probabilistic forecasting of hydraulic fracturing-induced seismicity using an injection-rate driven ETAS model. *Seismol. Res. Lett.* **92**(6), 3471–3481. <https://doi.org/10.1785/0220200454> (2021).
51. Kim, W. Y. Induced seismicity associated with fluid injection into a deep well in Youngstown, Ohio. *J. Geophys. Res. Solid Earth* **118**(7), 3506–3518. <https://doi.org/10.1002/jgrb.50247> (2013).
52. Friberg, P. A., Besana-Ostman, G. M. & Dricker, I. Characterization of an earthquake sequence triggered by hydraulic fracturing in Harrison County, Ohio. *Seismol. Res. Lett.* **85**(6), 1295–1307. <https://doi.org/10.1785/0220140127> (2014).
53. Holland, A. A. Earthquakes triggered by hydraulic fracturing in south-central Oklahoma. *Bull. Seismol. Soc. Am.* **103**(3), 1784–1792. <https://doi.org/10.1785/0120120109> (2013).
54. Darold, A., Holland, A. A., Chen, C., & Youngblood, A. Preliminary analysis of seismicity near Eagleton 1–29, Carter County, July 2014. *Oklahoma Geol. Soc. Open File Rept. OF2-2014* (2014).
55. Skoumal, R. J., Ries, R., Brudzinski, M. R., Barbour, A. J. & Currie, B. S. Earthquakes induced by hydraulic fracturing are pervasive in Oklahoma. *J. Geophys. Res. Solid Earth* **123**(12), 10–918. <https://doi.org/10.1029/2018JB016790> (2018).
56. Fasola, S. L. *et al.* Hydraulic fracture injection strategy influences the probability of earthquakes in the Eagle Ford shale play of South Texas. *Geophys. Res. Lett.* **46**(22), 12958–12967. <https://doi.org/10.1029/2019GL085167> (2019).

57. Yoon, C. E., Huang, Y., Ellsworth, W. L. & Beroza, G. C. Seismicity during the initial stages of the Guy-Greenbrier, Arkansas, earthquake sequence. *J. Geophys. Res. Solid Earth* **122**(11), 9253–9274. <https://doi.org/10.1002/2017JB014946> (2017).
58. Lei, X. *et al.* Fault reactivation and earthquakes with magnitudes of up to Mw4.7 induced by shale-gas hydraulic fracturing in Sichuan Basin, China. *Sci. Rep.* **7**(1), 1–12. <https://doi.org/10.1038/s41598-017-08557-y> (2017).
59. Tan, Y. *et al.* Hydraulic fracturing induced seismicity in the southern Sichuan Basin due to fluid diffusion inferred from seismic and injection data analysis. *Geophys. Res. Lett.* **47**(4), e2019GL084885. <https://doi.org/10.1029/2019GL084885> (2020).
60. Charléty, J. *et al.* Large earthquakes during hydraulic stimulations at the geothermal site of Soultz-sous-Forêts. *Int. J. Rock Mech. Min. Sci.* **44**(8), 1091–1105. <https://doi.org/10.1016/j.ijrmms.2007.06.003> (2007).
61. Dorbath, L., Cuenot, N., Genter, A. & Frogneux, M. Seismic response of the fractured and faulted granite of Soultz-sous-Forêts (France) to 5 km deep massive water injections. *Geophys. J. Int.* **177**(2), 653–675. <https://doi.org/10.1111/j.1365-246X.2009.04030.x> (2009).
62. Baisch, S., Weidler, R., Vörös, R., Wyborn, D. & de Graaf, L. Induced seismicity during the stimulation of a geothermal HFR reservoir in the Cooper Basin, Australia. *Bull. Seismol. Soc. Am.* **96**(6), 2242–2256. <https://doi.org/10.1785/0120050255> (2006).
63. Albaric, J. *et al.* Monitoring of induced seismicity during the first geothermal reservoir stimulation at Paralana, Australia. *Geothermics* **52**, 120–131. <https://doi.org/10.1016/j.geothermics.2013.10.013> (2014).
64. Baisch, S., Bohnhoff, M., Ceranna, L., Tu, Y. & Harjes, H. P. Probing the crust to 9-km depth: fluid-injection experiments and induced seismicity at the KTB superdeep drilling hole, Germany. *Bull. Seismol. Soc. Am.* **92**(6), 2369–2380. <https://doi.org/10.1785/0120010236> (2002).
65. Blöcher, G. *et al.* Evaluating micro-seismic events triggered by reservoir operations at the geothermal site of Groß Schönebeck (Germany). *Rock Mech. Rock Eng.* **51**(10), 3265–3279. <https://doi.org/10.1007/s00603-018-1521-2> (2018).
66. Kraft, T. & Deichmann, N. High-precision relocation and focal mechanism of the injection-induced seismicity at the Basel EGS. *Geothermics* **52**, 59–73. <https://doi.org/10.1016/j.geothermics.2014.05.014> (2014).
67. Diehl, T., Kraft, T., Kissling, E. & Wiemer, S. The induced earthquake sequence related to the St. Gallen deep geothermal project (Switzerland): Fault reactivation and fluid interactions imaged by microseismicity. *J. Geophys. Res. Solid Earth* **122**(9), 7272–7290. <https://doi.org/10.1002/2017JB014473> (2017).
68. Kwiatek, G. *et al.* Controlling fluid-induced seismicity during a 6.1-km-deep geothermal stimulation in Finland. *Sci. Adv.* **5**(5), eaav7224. <https://doi.org/10.1126/sciadv.aav7224> (2019).
69. Woo, J. U., Kim, M., Rhie, J. & Kang, T. S. Aftershock sequence and statistics of the 2017 M w 5.5 Pohang, South Korea, earthquake: Implications of fault heterogeneity and postseismic relaxation. *Bull. Seismol. Soc. Am.* **110**(5), 2031–2046. <https://doi.org/10.1785/0120200059> (2020).
70. Kwiatek, G., Bulut, F., Bohnhoff, M. & Dresen, G. High-resolution analysis of seismicity induced at Berlin geothermal field, El Salvador. *Geothermics* **52**, 98–111. <https://doi.org/10.1016/j.geothermics.2013.09.008> (2014).
71. Leptokaropoulos, *et al.* IS-EPOS: A platform for anthropogenic seismicity research. *Acta Geophys.* **67**(1), 299–310. <https://doi.org/10.1007/s11600-018-0209-z> (2019).
72. Bentz, S., Kwiatek, G., Martínez-Garzón, P., Bohnhoff, M. & Dresen, G. Seismic moment evolution during hydraulic stimulations. *Geophys. Res. Lett.* **47**(5), e2019GL086185. <https://doi.org/10.1029/2019GL086185> (2020).
73. Berger, V. W. & Zhou, Y. Kolmogorov–Smirnov test: Overview. *Wiley Statsref Stat. Ref. online* <https://doi.org/10.1002/9781118445112.stat06558> (2014).
74. Gupta, A. K. & Nadarajah, S. *Handbook of Beta Distribution and Its Applications* 600 (CRC Press, 2004).
75. Shcherbakov, R., Zhuang, J., Zöller, G. & Ogata, Y. Forecasting the magnitude of the largest expected earthquake. *Nat. Commun.* **10**(1), 1–11. <https://doi.org/10.1038/s41467-019-11958-4> (2019).
76. Nur, A. & Booker, J. R. Aftershocks caused by pore fluid flow? *Science* **175**, 885–887. <https://doi.org/10.1126/science.175.4024.885> (1972).
77. Miller, S. A. Aftershocks are fluid-driven and decay rates controlled by permeability dynamics. *Nat. Commun.* **11**(1), 1–11. <https://doi.org/10.1038/s41467-020-19590-3> (2020).
78. Rundle, J. B., Gross, S., Klein, W., Ferguson, C. & Turcotte, D. L. The statistical mechanics of earthquakes. *Tectonophysics* **277**(1–3), 147–164. [https://doi.org/10.1016/S0040-1951\(97\)00083-8](https://doi.org/10.1016/S0040-1951(97)00083-8) (1997).
79. Broccardo, M., Mignan, A., Wiemer, S., Stojadinovic, B. & Giardini, D. Hierarchical Bayesian modeling of fluid-induced seismicity. *Geophys. Res. Lett.* **44**(22), 11–357. <https://doi.org/10.1002/2017GL075251> (2017).
80. Mignan, A., Broccardo, M., Wiemer, S. & Giardini, D. Induced seismicity closed-form traffic light system for actuarial decision-making during deep fluid injections. *Sci. Rep.* **7**(1), 1–10. <https://doi.org/10.1038/s41598-017-13585-9> (2017).
81. Schmittbuhl, J. *et al.* Induced and triggered seismicity below the city of Strasbourg, France from November 2019 to January 2021. *C. R. Géosci.* **353**(S1), 1–24. <https://doi.org/10.5802/crgeos.71> (2021).
82. Barree, R. D., Miskimins, J. L. & Gilbert, J. V. Diagnostic fracture injection tests: common mistakes, misfires, and misdiagnoses. *SPE Prod. Oper.* **30**(02), 84–98. <https://doi.org/10.2118/169539-PA> (2015).
83. Orpsal, I. & Eisner, L. Cross-correlation—An objective tool to indicate induced seismicity. *Geophys. J. Int.* **196**(3), 1536–1543. <https://doi.org/10.1093/gji/ggt501> (2014).
84. Schultz, R. & Telesca, L. The cross-correlation and reshuffling tests in discerning induced seismicity. *Pure Appl. Geophys.* **175**(10), 3395–3401. <https://doi.org/10.1007/s00024-018-1890-1> (2018).

Acknowledgements

Funding for this work came from the Stanford Center for Induced and Triggered Seismicity. We would like to thank Grzegorz Kwiatek, Simone Mancini, Volker Oye, Julie Albaric, Nicholas van der Elst, James Verdon, Toni Kraft, and Jeongung Woo for pointing us towards relevant catalogue/injection datasets. We would also like to thank James Verdon and an anonymous reviewer for their comments that helped improve this manuscript.

Author contributions

R.S. conceived the project, collected data, carried out the analysis, and wrote the manuscript with guidance, comments, and revisions from all the authors. W.L.E. assisted in data collection. G.C.B. and W.L.E. supervised the project.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-05216-9>.

Correspondence and requests for materials should be addressed to R.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022