Probabilistic Forecasting of Hydraulic Fracturing Induced Seismicity Using an Injection-Rate Driven ETAS Model

S. Mancini[1,2], M. J. Werner[2], M. Segou[1], and B. Baptie[1]

[1] British Geological Survey, Lyell Centre, EH10 4AP, Edinburgh, UK
[2] School of Earth Sciences, University of Bristol, BS8 1RL Bristol, UK

Corresponding author: Simone Mancini (simone@bgs.ac.uk)

Declaration of Competing interests:
The authors acknowledge there are no conflicts of interest recorded.

Abstract

The development of robust forecasts of human-induced seismicity is highly desirable to mitigate the effects of disturbing or damaging earthquakes. We assess the performance of a well-established statistical model, the Epidemic-Type Aftershock Sequence (ETAS) model, with a catalog of ~93,000 microearthquakes observed at the Preston New Road (UK) unconventional shale gas site during and after hydraulic fracturing of the PNR-1z and PNR-2 wells. Because ETAS was developed for slower loading rate tectonic seismicity, in order to account for seismicity caused by pressurized fluid we also generate three modified ETAS with background rates proportional to injection rates. We find that (1) the standard ETAS captures low seismicity between and after injections but is outperformed by the modified model during high seismicity periods, and (2) the injection-rate driven ETAS substantially improves when the forecast is calibrated on sleeve-specific pumping data. We finally forecast out-of-sample the PNR-2 seismicity using the average response to injection observed at PNR-1z, achieving better predictive skills than the in-sample standard ETAS. The insights from this study contribute towards producing informative seismicity forecasts for real-time decision making and risk mitigation techniques during unconventional shale gas development.

Introduction

Seismicity induced by fluid injections is a growing concern (*Schultz et al.*, 2020, and references therein). Many countries are witnessing an increased development of subsurface geo-energy reservoirs, including unconventional shale gas development, enhanced geothermal energy systems, fluid injection in salt mine fields, wastewater injection, and underground storage of liquid carbon (*Ellsworth*, 2013). These activities promote seismicity in previously low seismic hazard regions or further increase high seismic rates. In recent years, induced seismicity with moderate magnitudes (M5-5.7) in regions such as the central United States, South Korea and southwestern China has led to significant damages and losses (*Keranen et al.*, 2013; *Ellsworth et al.*, 2019; *Lee et al.*, 2019; *Lei et al.*, 2020). While several hypotheses about the interplay of deterministic physical mechanisms controlling the seismic response to subsurface fluid injection are currently under investigation (*Atkinson et al.*, 2020), probabilistic methods provide a framework for current epistemic and aleatory uncertainties. Indeed, statistical models of injection-induced seismicity have shown some skill in capturing the complex range of seismic responses to fluid injections (*e.g., Shapiro et al.*, 2007; 2010; *Kiraly-Proag et al.*, 2016; *Verdon and Budge*, 2018). In particular, a popular statistical method, the Epidemic-Type Aftershock Sequence (ETAS) model (*Ogata*, 1988), originally developed to reproduce the short-term clustering of tectonic earthquakes, was tested under different fluid-induced seismicity scenarios including natural circulation of fluids at depth (*Hainzl and Ogata*, 2005) as well as human-related activities, such as natural gas extraction (*Bourne and Oates*, 2017), Enhanced Geothermal Systems (EGS - e.g. *Bachmann et al.*, 2011; *Mena et al.*, 2013; *Asanuma et al.*, 2014), hydraulic fracturing for unconventional shale gas development (HF - *e.g. Lei et al.*, 2017; 2019; *Jia et al.*, 2020), and wastewater disposal (*Llenos and Michael*, 2013). These studies concluded that fluid-driven seismicity has distinctive spatiotemporal characteristics, some of which are different from the 'regular' tectonic seismicity dominated by earthquake-to-earthquake triggering mechanisms. While the standard ETAS features a stationary background rate due to slower tectonic loadings, *Bachmann et al.* (2011) introduced an ETAS model with a background rate linearly proportional to the injection rate and found that this model performed best in forecasting the seismicity induced in Basel (Switzerland) due to the stimulation of a deep geothermal energy reservoir.

In its limited number of applications to HF environments, the ETAS model was mostly used to explore the behavior of HF-induced seismicity and to show that time-varying background rates positively correlate with injection operations (*Lei et al.*, 2019; *Jia et al.*, 2020). *Lei et al.* (2017)

showed that an ETAS model featuring a non-stationary background rate better reproduces the observed features of seismicity when an external forcing is applied (*e.g.,* fluid flow or aseismic slip in cases of induced and natural seismicity, respectively), but their primary scope was not to assess ETAS performance in a formal forecasting experiment.

In this study, we probe the suitability of the ETAS model as a statistical tool for near real-time forecasts of the seismic rates during and after HF operations. We expand on previous applications of the ETAS model to HF by quantitatively assessing the predictive skills of a suite of temporal ETAS models that (1) are calibrated and tested on a much richer microseismicity dataset, (2) seek to reproduce seismic rates from a wider magnitude range (from M~3 down to M=-1.5), (3) explore how the forecast performance changes under different modelling assumptions (standard vs. modified model formulations) and parameterizations (in-sample vs. out-of-sample forecasts), and (4) test the influence of expressing the non-stationary background rates by using either averaged or sleeve-specific fluid pumping parameters.

We take advantage of a rich microseismicity dataset recorded at Preston New Road, Lancashire (UK), during unconventional shale gas development by Cuadrilla Ltd in two wells, PNR-1z in 2018 (*Clarke et al.*, 2019) and PNR-2 in 2019. First, we implement the ETAS model in its original tectonic formulation and assess whether (1) it captures the temporal evolution of the microseismicity, and (2) parameters optimized using the available data improve model performance. Second, we implement a modified ETAS model featuring a background seismicity rate proportional to the injection rate following *Bachmann et al.* (2011) but here applied in the context of HF. This presents a particular challenge as HF operations feature short injection episodes along different sleeves, while EGS injections are continuous with gradually changing flow rates at a single injection point. Within the modified ETAS class, we (1) assess model performance against the standard ETAS model, and (2) quantify the influence of using an average (bulk) constant of proportionality between seismicity and injection rates calculated over the entire period of operations at each well versus constants specifically calibrated on individual injection periods. For both ETAS classes we also perform an out-of-sample experiment where we calibrate the ETAS model on PNR-1z data and then use it to independently forecast microseismicity during PNR-2. We rank the forecasts by means of likelihood scores, a well-established metric (*e.g.,* used within the Collaboratory for the Study of Earthquake Predictability, CSEP, *Michael and Werner*, 2018). The comparative performance evaluation illustrates the predictive skills of injection-rate driven ETAS models and how these may inform real-time decision-making by operators and regulators during HF operations.

Operations and seismicity at Preston New Road, UK

Hydraulic fracturing operations at the PNR-1z well occurred between 15 October and 17 December 2018. The well ran for 700 m horizontally through the natural gas-bearing Carboniferous formation of the Lower Bowland Shale at a depth of ~2.3 km (*Clarke et al.*, 2018). A total of 17 sleeves were hydraulically fractured (Figure 1a) with mini fracs at 18 additional sleeves, consisting of a few tens of m^3 of fluid pumped. Overall, a total of ~4600 m^3 of slick water fluid was injected (Figure 2a) with an average volume per sleeve of 234 m^3 (and a maximum V_{MAX} = 431 m^3). Hydraulic fracturing was paused between 3 November and 4 December 2018 as flow-back from the well took place. The microseismicity at PNR-1z was recorded by a downhole array in the adjacent PNR-2 well consisting of 12 three-component geophones that detected over 38,000 events. Although local 3D reflection seismic surveys acquired before the start of operations revealed the presence of pre-existing seismic discontinuities, these were located far from the well and did not present any clear correlation with the initial microseismicity (*Clarke et al.*, 2019). As injection proceeded, hydraulic fractures started intersecting another pre-existing (but not previously identified) subvertical

NE-striking seismogenic feature, located NE of the well. The largest magnitude event that occurred on 11 December 2018 (M_L = 1.5) activated a section of such structure. However, as reported by *Kettlety et al.* (2020a), it is not clear whether this was a single contiguous fault or a dense zone of fractures.

Here, we use the available earthquake catalog that includes origin times and moment magnitudes (M_w) as determined by Schlumberger Ltd., the geophysical processing contractor. The limited dynamic range of the downhole geophones leads to problems in magnitude estimation for $M_w \geq 0.0$ events due to clipping. To avoid a potential bias, we matched these with events in the catalog obtained from broadband surface stations operated by the British Geological Survey (BGS) that reported 172 events with local magnitudes (M_L). We then replaced the moment magnitudes for all $M_w \geq 0.0$ events in the downhole catalog with the corresponding local magnitude estimate, following *Clarke et al.* (2019) for the same dataset. This *ad hoc* solution to the problem of PNR-1z magnitude conversions remains the subject of ongoing research (*Baptie et al.*, 2020). *Clarke et al.* (2019) argued that assuming $M_L = M_w$ for all $M_w \geq 0.0$ events does not produce anomalies in the frequency-magnitude distribution, suggesting that this simple approach is reasonable.

Figure 2a shows a histogram of the hourly number of events during operations along with the cumulative volume of injected fluid. The observed seismicity at PNR-1z shows multiple peaks that visually correlate well with the pumping periods and then decay rapidly with time after injection stops. We find evidence of considerable variations in seismic responses despite comparable injection rates across sleeves (*e.g.*, Figure 2c-d). For instance, at sleeve #2 (injection stage S02) event rates increase as soon as injection starts and remain relatively stable (Figure 2c), while at sleeve #40 (injection stage S17) there is a delayed onset of seismicity followed by substantially higher rates (Figure 2d).

The horizontal PNR-2 well runs roughly parallel to PNR-1z offset by approximately 200 m and was drilled through the upper part of the Lower Bowland Shale formation at a depth of ~2.1 km. Operations started on 15 August 2019 but were suspended on 26 August following a M_L = 2.9 earthquake that was felt up to a few kilometers from the epicenter (*Cremen and Werner*, 2020). Aftershocks of this event illuminated a SE-striking fault, a clearly different feature than the one activated during hydraulic stimulation at PNR-1z. Furthermore, the latter did not show any seismicity during operations at PNR-2; it is likely that a barrier blocking any interaction between the two zones was created by lateral lithological variabilities as well as by the notable vertical and lateral separation between the two wells (*Kettlety et al.*, 2020b).

PNR-2 seismicity was recorded by a downhole array of 12 geophones in the adjacent PNR-1z well, and the final catalog, extending up to 2 October 2019, consists of over 55,000 microseismic events (Figure 1b) with magnitudes reported as Mw. We added a correction of 0.15 magnitude units to the downhole moment magnitudes following *Baptie et al.* (2020). Furthermore, the PNR-2 catalog suffers from brief but critical data gaps that result in a loss of otherwise recorded seismic events, including the largest event in the sequence (M_L = 2.9) and presumably its early aftershocks. We filled these gaps with events recorded by the combined surface network of the BGS and the operator (*Baptie and Luckett*, 2019).

The early earthquake productivity at PNR-2 appears an order of magnitude larger than that observed during the initial injection stages at PNR-1z, even under similar injected volumes (Figure 2b). The complexity of the seismic response to injection is similar to PNR-1z (Figure 2e). As at PNR-1z, we observe a general positive co-dependency between seismicity and fluid injection at PNR-2.

Methods

The standard ETAS model

The Epidemic-Type Aftershock Sequence (ETAS) model (*Ogata*, 1988) is a statistical model of the time-magnitude characteristics of triggered tectonic seismicity. The model treats seismicity as a self-exciting stochastic point process, in which each earthquake produces offspring with magnitudes independently sampled from the Gutenberg-Richter distribution (that is, parent earthquakes can trigger larger events with some probability). The seismic rate $\lambda(\mathrm{t})$ at time t is given by a time-independent background rate (μ) plus a function accounting for the history (H_t) of triggering contributions from all previous events at time t_i and with magnitude M_i prior to t:

$$\lambda(t \mid H_t) = \mu + \sum_{i:t_i<t} K e^{\alpha(M_i - M_{cut})} \cdot c^{p-1}(t - t_i + c)^{-p}(p-1)\,, \tag{1}$$

where the sum includes empirically observed relations that (1) describe the short-term aftershock productivity of events above a minimum triggering threshold ($\mathrm{M_{cut}}$) with parameters K and α, (2) determine an Omori-Utsu temporal decay of the triggered rate with exponent *p* and a constant *c* (*Utsu*, 1961). We estimate the parameters (μ, K, α, c, p), by maximizing the log-likelihood function (*Zhuang et al.*, 2012) on a seismic catalog with N events and over a period from T_0 to T_1:

$$log\, L(\mu, K, \alpha, c, p) = \sum_{i=1}^{N} log\, \lambda\,(t_i \mid H_t) - \int_{T_0}^{T_1} \lambda(t)\, dt\,. \tag{2}$$

Forecasts of the ETAS model require simulations because the rate is conditional on the history (e.g., *Zhuang and Touati,* 2015; *Seif et al.*, 2017).

We create three versions of the standard ETAS model (the "ETAS1" class). In *ETAS1-optimized* we estimate ETAS parameters from the target catalog (either PNR-1z or PNR-2) and thus perform an in-sample (best-case) forecast evaluation. In *ETAS1-unoptimized* we use the parameters estimated from PNR-1z data to forecast the PNR-2 seismicity out-of-sample. *ETAS1-global* serves as an alternative benchmark model with the most recently estimated ETAS parameters from global subduction zones (except for the background rate) by *Zhang et al.* (2020). We select parameters from interplate settings because these might represent the tectonic counterpart that most closely matches the forcing and boundary conditions of in-situ fluid-induced seismicity environments, that is, high stressing rates and relatively short-lived aftershock sequences.

The modified ETAS model for injection-induced seismicity

In the second forecast class ("ETAS2"), we modify the ETAS model to account for events forced by an external driver. We couple the background rate to the time-dependent fluid injection rate $\mathrm{I_r}(\mathrm{t})$:

$$\lambda_m(t \mid H_t) = \mu(I_r) + \sum_{i:t_i<t} K e^{\alpha(M_i - M_{cut})} \cdot c^{p-1}(t - t_i + c)^{-p}(p-1)\,, \qquad (3)$$

with λ_m a "modified" seismic rate and the background rate $\mu(\mathrm{I_r})$ now assumed to be linearly related to the injection rate via a constant of proportionality $\mathrm{c_f}$ (*Bachmann et al.*, 2011):

$$\mu(I_r) = c_f\, I_r(t). \qquad (4)$$

To estimate $\mathrm{c_f}$, we maximize:

$$log\, L(c_f, K, \alpha, c, p) = \sum_{i=1}^{N} log\, \lambda_m(t_i \mid H_t) - \int_{T_0}^{T_1} \lambda_m(t)\, dt\,. \qquad (5)$$

Within the ETAS2 class, we develop three forecast versions. In *ETAS2-bulk* we estimate and use only a single value of $\mathrm{c_f}$ for each well, fit over the entire period of operations. *ETAS2-specific* implements specific values of $\mathrm{c_f}$ for each sleeve, calibrated within the individual injection periods; in this model, we fix the triggering parameters (K, c, p, α) to the respective values previously obtained for *ETAS2-bulk* assuming that the contribution of event-to-event interactions does not change in different injection periods, when the external forcing is likely to be the dominant mechanism of earthquake production. Finally, *ETAS2-unoptimized* uses the ETAS parameters estimated on the PNR-1z catalog (including its bulk proportionality constant) to forecast out-of-sample the expected seismic response at PNR-2.

Simulating ETAS2 models requires a different method for background events during injection periods. We apply the thinning algorithm (e.g. *Zhuang and Touati*, 2015): (i) estimate a mean expected number of forced events ($\overline{N_f}$) by multiplying c_f by the injection rate integrated over the duration of either the injection period or the forecast window (whichever is shorter); (ii) draw a random variable (N_f) from a Poisson distribution with mean equal to $\overline{N_f}$; (iii) distribute the N_f events in time according to a piece-wise linear, non-homogeneous Poisson process with rate $\mu(I_r)$ driven by the injection rate (smoothed using 1-minute moving windows); (iv) simulate all aftershock generations triggered by the directly forced events by means of the standard procedure.

For consistency, all six ETAS versions are updated hourly (or when an injection period starts, whichever comes sooner), and estimated by 1,000 stochastic ETAS simulations with fixed M_{max}= 6.5 (the most likely regional maximum expected tectonic magnitude; *Woessner et al.*, 2015). It is worth noting that incomplete datasets can bias the estimation of the ETAS parameters and potentially lead to seismic rate underpredictions (*Seif et al.*, 2017). For the PNR-1z microseismicity catalog, we estimate a magnitude of completeness (M_c) between -1.2 and -1.5 (Figure S1a), while our M_c estimate for PNR-2 is below -1.5 (Figure S1b). However, $M \geq -1.2$ events represent only ~7% of earthquakes recorded at PNR-1z. Furthermore, here we are interested in producing earthquake models that can forecast events also during periods of intense injection-induced seismicity, which instead consist primarily of very small magnitude earthquakes. Therefore, to find a pragmatic compromise and to increase the number of events to around 20% of the entire PNR-1z dataset, we conduct our analyses using the lower bound of the estimated PNR-1z catalog completeness range (M_c = -1.5). For comparability, we use the same magnitude threshold for PNR-2. Accordingly, all our ETAS models seek to forecast the number of M ≥ -1.5 events at the two wells.

In the electronic supplement, we report a summary of the tested ETAS versions (Table S1) and the values of the ETAS parameters (Table S2), including the bulk and sleeve-specific values of c_f (Tables S3 and S4 for PNR-1z and PNR-2, respectively).

Evaluation of model performance

Because each forecast consists of a probability distribution of earthquake numbers over the forecast period, we evaluate and rank forecast models using a probabilistic score, namely the log-likelihood values. The score quantifies the likelihood of the observed number if the models were the data-generator, specifically the logarithm of the probability $Pr(\omega|model)$ of observing ω earthquakes given the ETAS forecasts (*Zechar*, 2010):

$$LL(\omega|model) = \log\big(Pr(\omega|model)\big). \tag{6}$$

To compensate for the limited number of simulations, which is likely to under-sample the range of possible simulated ETAS rates, we approximate the simulation histogram of each forecast window with a Negative Binomial Distribution (NBD; *Harte*, 2015) (Figures S2 and S3). We choose the two-parameter NBD because it characterizes earthquake clustering and process overdispersion much better than the Poisson distribution (*Kagan*, 2010). We calculate the likelihood scores from the fitted NBD. Greater log-likelihood scores indicate greater predictive skill.

Results

Forecast timeseries

In Figures 3a and 4a, we present the incremental hourly timeseries of the three in-sample ETAS forecasts for PNR-1z and PNR-2. We select illustrative subperiods characterized by (1) weak and strong seismic responses to injection, and (2) seismicity without injection. The panels compare the observed number of $M \geq -1.5$ events per hour with the mean and 95% predictive interval of the ETAS model. Firstly, we find that the ETAS1 class projects the onset of increased rates with a 1-hour delay compared to observations. This is not an unexpected effect due to the scarcity of $M \geq -1.5$ parent earthquakes prior to each injection period and the fact that ETAS1 does not account for external seismicity forcing. Secondly, the standard *ETAS1-optimized* severely underestimates the observed rates by an order of magnitude during the higher seismicity periods, whether the seismic response to injection is weak or strong. The reason for this underprediction is the fact that *ETAS1-optimized* lacks information about impending active fluid injections. In contrast, other forecast time windows characterized by underpredictions, such as those immediately following the stop of injections, may suffer from the possible temporary incompleteness of the catalog. Although the estimated ETAS parameters may compensate for this effect, the time-varying incompleteness results in some target periods with fewer small events that would have otherwise increased the chances of triggering additional events. Therefore, the early post-injection model performance might improve with a more complete catalog. However, in post-injection conditions (*i.e.*, a few hours after the end of pumping), when any earthquake clustering is likely driven by event-to-event triggering, *ETAS1-optimized* generally reproduces well the hourly seismicity within the model's 95% ranges at PNR-1z (Figure 3a) and PNR-2 (Figure 4a). Interestingly, during periods of no injection and low seismicity at PNR-1z, the 95% forecast range often encompasses the critical value of zero events, reflecting the intrinsic stochasticity of the ETAS model.

The ETAS2 class, featuring an injection-rate-driven background rate, substantially reduces the discrepancies with the observed rates. *ETAS2-bulk*, which captures the average seismic response to injection, both under- and over-predicts during injection periods. This mixed performance is a result of the single proportionality constant for each dataset that does not sufficiently capture the complex relationships between injection rate and seismicity. *ETAS2-specific*, which describes the seismicity response with sleeve-specific injection data, presents the best match during the periods of high seismicity rate due to pressurized fluid forcing. Here, the visual comparison is very encouraging, but hinges on in-sample, sleeve-specific proportionality constants between seismic rates and injection rates.

We next analyze the performance of all ETAS models, including the out-of-sample versions, over the entire testing periods at PNR-1z (Figure 3b-d) and at PNR-2 (Figure 4b-d). Using a simple acceptance/rejection criterion, we consider a forecast *accepted* (green symbols) if the observations fall within the 95% model range, otherwise we mark it as *rejected* (red symbols). An ideal forecast, which predicts the observations perfectly, aligns along the diagonal lines of Figures 3b-d and 4b-d. While the observations fall into the 95% forecast range of the ETAS1 models about 80% of the time, these matches correspond to periods of low seismicity: accepted forecasts occur only when less than 40 events are observed at PNR-1z (Figure 3b,c) and less than 150 events are observed at PNR-2 (Figure 4b,c). We also note that (1) at both PNR-1z and PNR-2 *ETAS1-global* overpredicts less frequently than models parameterized on well-specific seismicity when the seismicity rate is extremely low (Figure 3b,c and Figure 4b,c) but underpredicts more during high-rate windows, and (2) in PNR-2 the differences between

ETAS1-optimized and *ETAS1-unoptimized* are negligible (Figure 4b), a result of the similar parameters estimated from the two wells (Table S2).

The performance of the ETAS2 class (Figures 3d and 4d) differs from ETAS1 mostly during injection periods, and the improvement is appreciable. *ETAS2-specific* performs strikingly well, as the only model to forecast very productive periods with more than 300 events at PNR-1z (Figure 3d) and more than 1,000 events at PNR-2 (Figure 4d). Finally, the out-of-sample *ETAS2-unoptimized* model, which uses the bulk seismic response to injection at PNR-1z to forecast seismicity at PNR-2, persistently underpredicts injection-induced high rates (Figure 4d), but its underprediction is less severe than that of the ETAS1 class.

Likelihood scores

The cumulative log-likelihood scores of the models over the entire duration of the PNR catalogs show that the injection-rate driven ETAS2 realizations considerably outperform models belonging to the standard ETAS1 class (Figure 5). In particular, *ETAS2-specific* has the highest likelihood scores at both wells and thus ranks as the best performing model, followed by *ETAS2-bulk* as second-best. The latter performs unevenly in the two wells, with better predictive skill in PNR-1z (Figure 5a) than in PNR-2 (Figure 5b) during the first few days of operations. Encouragingly, the out-of-sample *ETAS2-unoptimized* model scores better than all ETAS1 models and performs similarly to *ETAS2-bulk* during the first week of treatment of PNR-2. In other words, a model calibrated on PNR-1z data could have provided informative forecasts for PNR-2.

ETAS1-global performs worse than the injection-rate driven ETAS2 class but compares well with the other ETAS1 models and even with the *ETAS2-unoptimized* and *ETAS2-bulk* models in the early stages of PNR-2 (inset of Figure 5b); this is a priori surprising for a model calibrated on moderate to large subduction zone earthquakes.

Conclusions

The PNR microseismic datasets present a unique opportunity to develop and evaluate statistical forecasting models of hydraulic fracturing induced seismicity. Notwithstanding the variability and uncertainties linking pumping data to the induced seismicity response at both PNR wells, we observe a generally positive co-dependency between seismicity and injection rates that supports the incorporation of operational parameters into the standard tectonic ETAS model.

In comparing the performance of the standard and injection-rate driven ETAS forecasts, we find that the seismicity decay after the operations, or between stages, is satisfactorily captured by the standard ETAS. We interpret this result as follows. During operations we witness the complex interplay of rapid pore pressure effects and earthquake clustering, expressing a variety of possible mechanisms (*e.g.*, elastostatic stress transfer, poroelastic effects, aseismic creep) (*Schultz et al.*, 2020), while external forcing ceases in inter- and post-injection periods and seismicity shows a more typical tectonic behavior.

However, the log-likelihood scores of the ETAS models demonstrate that a non-stationary background rate tied to the injection rate is necessary to avoid severe underpredictions during injection periods, when the seismic productivity is high. Thus, even a simplistic linear relationship between injection rate and induced seismicity leads to informative ETAS forecasts in HF environments. From the model comparison, we conclude that (1) bulk constants of proportionality do not accurately describe the variable seismic response to fluid injection, and (2) a sleeve-specific modulation of the seismic response to injection is the most critical element for producing reliable forecasts.

In our study, the best-performing ETAS model is an in-sample forecast that represents a best-case scenario. This performance may be difficult to attain out-of-sample. However, the sleeve-specific constants of proportionality could be estimated and fine-tuned in near real-time conditions from the initial seismic response at the sleeve, similarly to real-time attempts to estimate parameters of other models (e.g., *Clarke et al.*, 2019). Given the temporal variability of the seismic response to constant injection and the time-varying catalog completeness thresholds, the parameters will doubtlessly be more uncertain, and this additional uncertainty should be propagated into the forecasts. In this regard, the operator would have to assume that (i) the injection rate at each sleeve is known in advance and (ii) the evolving sleeve-specific seismic response is continuously acquired and adequately detected to support frequent model calibrations.

To mimic real-time conditions (*i.e.,* before data are available for parameter estimation), we also evaluate forecasts from three out-of-sample models. Although their performance is worse than the in-sample models, we also see encouraging results. The models present low log-likelihood scores in the longer term (*i.e.,* more than 3-5 days after the start of operations), but they perform comparably to some in-sample models during the first few days of operations. This is true even for the ETAS model calibrated on data from global subduction zones. This is promising for operational conditions: operators could provide forecasts during the very early stages of operations using parameters that are either generic or previously calibrated on adjacent wells. As well-specific and stage-specific data become available, forecasts can be improved with re-estimated parameters and the operational injection data, similarly to an ETAS approach proposed for other time-varying fluid-driven processes such as natural seismic swarms (*Llenos & Michael*, 2019). To further assess the robustness of the model parameterization and performance, future tests should involve datasets with a coherent magnitude scale and a less time-variant magnitude completeness level.

In light of the results from the PNR experiments, we conclude that injection-rate driven ETAS models produce informative time-dependent probabilistic seismic rate forecasts. The seismicity forecasts, when convolved with models of ground motion, exposure and vulnerability, can support time-dependent probabilistic seismic hazard and risk assessment. These forecast models may provide useful information for operators, regulators, residents and other stakeholders in HF environments.

Data and resources

The PNR-1z and PNR-2 microseismicity catalogs as well as the fluid injection rate data used in this study can be acquired through access to the UK Oil and Gas Authority website at https://www.ogauthority.co.uk/exploration-production/onshore/onshore-reports-and-data/. The supplemental material attached to this manuscript illustrates examples of histograms from the ETAS simulations performed for PNR-1z and PNR-2; it also provides a summary of the developed ETAS models along with their parameterizations.

Acknowledgements

The authors thank the editor and two anonymous reviewers for their constructive comments. We would also like to thank the UK Oil and Gas Authority (OGA) for providing the datasets. SM was supported by a Great Western Four+ Doctoral Training Partnership (GW4+ DTP) studentship from the Natural Environment Research Council (NERC) (NE/L002434/1) and by a studentship from the British Geological Survey University Funding Initiative (BUFI) (S350). MJW and BB were supported by NERC (NE/R017956/1, “EQUIPT4RISK”). MJW and MS were supported by the European Union H2020 program (No 821115, “RISE”). BB was also supported by the NERC grant NE/R01809X/1. This work was also supported by the Bristol University Microseismic ProjectS (“BUMPS”) and by the Southern California Earthquake Center (SCEC) (Contribution No. 10149). SCEC is funded by the National Science foundation Cooperative Agreement EAR-1600087 & US Geological Survey Cooperative Agreement G17AC00047.

References

Atkinson, G. M., D. W. Eton, and N. Igonin (2020). Developments in understanding seismicity triggered by hydraulic fracturing, *Nat. Rev. Earth Environ.*, 1, 264-277. https://doi.org/10.1038/s43017-020-0049-7.

Asanuma, H., T. Eto, M. Adachi, K. Saeki, K. Aoyama, H. Ozeki, and M. Häring (2014). Seismostatistical Characterization of Earthquakes from Geothermal Reservoirs. *Proceedings of the Thirty-Ninth Workshop on Geothermal Reservoir engineering Stanford University, Stanford, California, February 2014, SGP-TR-202.*

Baptie, B. and R. Luckett (2019). Seismicity Induced by Hydraulic Fracturing Operations at Preston New Road, Lancashire, 2018. *Proceedings of the Society of Earthquake and Civil Engineering Dynamics Conference, September 2019, Greenwich, London.*

Baptie, B., R. Luckett, A. Butcher, and M. J. Werner (2020). Robust relationships for magnitude conversion of PNR seismicity catalogues. *British Geological Survey Open Report OR/20/042, British Geological Survey for Oil and Gas Authority, London, United Kingdom, 32 pp.*

Bachmann, C., S. Wiemer, J. Woessner, and S. Hainzl (2011). Statistical analysis of the induced Basel 2006 earthquake sequence: Introducing a probability-based monitoring approach for Enhanced Geothermal Systems, *Geophys. J. Int.*, 186, 793-807.

Bourne, S., and S. Oates (2017). Development of statistical geomechanical models for forecasting seismicity induced by gas production from the Groningen field, *Netherlands Journal of geosciences*, 96(5), S175-S182. http://doi.org.10.1017/njg.2017.35.

Clarke, H., P. Turner, R. M. Bustin, N. Riley, and B. Besly (2018). Shale Gas Resources of the Bowland Basin, NW England: A Holistic Study, *Petrol. Geosci.*, 24(3), 287-322, https://doi.org/10.1144/petgeo2017-066.

Clarke, H., J. P. Verdon, T. Kettlety, A. F. Baird, and J. M. Kendall (2019). Real-Time Imaging, Forecasting, and Management of Human-Induced Seismicity at Preston New Road, Lancashire, England, *Seismol. Res. Lett.*, 90(5), 1902-1915.

Cao, A. M., and S. S. Gao (2002). Temporal variation of seismic b-values beneath northeastern Japan island arc, *Geophys. Res. Lett.*, 29(9), 1334. https://doi.org/10.1029/2001GL013775.

Cremen, G. and M. J. Werner (2020). A Novel Approach to Assessing Nuisance Risk from Seismicity Induced by UK Shale Gas Development, with Implications for Future Policy Design, *Nat. Hazards Earth Syst. Sci. Discuss.*, https://doi.org/10.5194/nhess-2020-95.

Cuadrilla Resources Inc. (2019). Hydraulic Fracture Plan PNR 2. Cuadrilla Resources Inc. Report CORP-HSE-RPT-003.

Ellsworth, W. L. (2013). Injection-induced earthquakes, *Science*, 341(6142), 1225942. https://doi.org/10.1126/science.1225942

Ellsworth, W. L., D. Giardini, J. Townend, S. Ge, and T. Shimamoto (2019). Triggering of the Pohang, Korea, Earthquake (Mw 5.5) by Enhanced Geothermal System Stimulation, *Seismol. Res. Lett.*, 90(5), 1844-1858.

Hainzl, S., and Y. Ogata (2005). Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. Geophys. Res. Solid Earth*, 110(B5).

Harte, D. (2015). Log-likelihood of earthquake models: evaluation of models and forecasts, *Geophys. J. Int.*, 201, 711-723, https://doi.org/10.1093/gji/ggu442.

Jia, K., S. Zhou, J. Zhuang, C. Jiang, Y. Guo, Z. Gao, S. Gao, Y. Ogata, and X. Song (2020). Nonstationary Background Seismicity Rate and Evolution of Stress Changes in the Changning Salt Mining and Shale-Gas Hydraulic Fracturing Region, Sichuan Basin, China, *Seismol. Res. Lett.* 91, 2170–2181, https://doi.org/10.1785/0220200092.

Kagan, Y. Y. (2010). Statistical distribution of earthquake numbers: consequence of branching process, *Geophys. J. Int.*, 180, 1313-1328, https://doi.org/10.1111/j.1365-246X.2009.04487.x.

Keranen, K. M., H. M. Savage, G. A. Abers, and E. S. Cochran (2013). Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 Mw 5.7 earthquake sequence, *Geology*, https://doi.org/10.1130/G34045.1.

Kettlety, T., J. P. Verdon, M. J. Werner, and J. M. Kendall (2020a). Stress transfer from opening hydraulic fractures controls the distribution of induced seismicity, *J. Geophys. Res. Solid Earth*, 125, e2019JB018794, https://doi.org/10.1029/2019JB018794.

Kettlety, T., J. P. Verdon, A. Butcher, M. Hampson, and L. Craddock (2020b). High-resolution imaging of the ML 2.9 August 2019 earthquake in Lancashire, United Kingdom, induced by hydraulic fracturing during Preston New Road PNR-2 operations, *Seismol. Res. Lett.*, 92(1), 151-169, https://doi.org/10.1785/0220200187.

Kiraly-Proag, E., J. D. Zechar, V. Gischig, S. Wiemer, D. Karvounis, and J. Doetsch (2016). Validating induced seismicity forecast models—Induced Seismicity Test Bench, *J. Geophys. Res. Solid Earth*, 121, 6009–6029, https://doi.org/10.1002/2016JB013236.

Lee, K., W. L. Ellsworth, D. Giardini, J. Townend, S. Ge, T. Shimamoto, I. Yeo, T. Kang, J. Rhie, D. Sheen, C. Chang, J. Woo, and C. Langenbruch (2019). Managing injection-induced seismic risks, *Science*, 364 (6442), 730–32.

Lei, X., D. Huang, J. Su, G. Jiang, X. Wang, H. Wang, X. Guo, and H. Fu (2017). Fault reactivation and earthquakes with magnitudes of up to Mw4.7 induced by shake-gas hydraulic fracturing in Sichuan Basin, China, *Sci. Rep.*, 7 (7971). https://doi.org/10.1038/s41598-017-08557-y.

Lei, X., Z. Wang, and J. Su (2019). 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.

Llenos, A. L., and A. J. Michael (2013). Modeling earthquake rate change in Oklahoma and Arkansas: possible signatures of induced seismicity, *Bull. Seismol. Soc. Am.*, 103(5), 2850-2861. https://doi.org/10.1785/0120130017.

Llenos, A. L., and A. J. Michael (2019), Ensembles of ETAS models provide optimal operational earthquake forecasting during swarms: Insights from the 2015 San Ramon, California swarm, *Bull. Seismol. Soc. Am.*, 109, 2145-2158. https://doi.org/10.1785/0120190020.

Mena, B., S. Wiemer and C. Bachmann (2013). Building robust models to forecast the induced seismicity related to geothermal reservoir enhancement, Bull. Seismol. Soc. Am., 103(1), 383-392. https://doi.org/10.1785/0120120102.

Michael, A. J., and M. J. Werner (2018). Preface to the Focus Section on the Collaboratory for the Study of Earthquake Predictability (CSEP): New Results and Future Directions, *Seismol. Res. Lett.*, 89(4), 1226-1228. https://doi.org/10.1785/0220180161.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.*, 83(401), 9–27. 2861–2864.

Seif, S., A. Mignan, J. D. Zechar, M. J. Werner, and S. Wiemer (2017). Estimating ETAS: The effects of truncation, missing data, and model assumptions, *J. Geophys. Res. Solid Earth* 121, 449–469. https://doi.org/10.1002/2016JB012809.

Shapiro, S. A., C. Dinske, and J. Kummerow (2007). Probability of a given-magnitude earthquake induced by a fluid injection, *Geophys. Res. Lett.*, 34, L22, 314, https://doi.org/10.1029/2007GL031615.

Schultz, R., R. J. Skoumal, M. R. Brudzinski, D. Eaton, B. Baptie, and W. Ellsworth (2020). Hydraulic fracturing-induced seismicity, *Rev. Geophys.*, 58, e2019RG000695. https://doi.org/10.1029/2019RG000695.

Utsu, T. (1961). A statistical study on the occurrence of aftershocks, *Geophys. Mag.*, 30, 521-605.

Verdon, J., and J. Budge (2018). Examining the Capability of Statistical Models to Mitigate Induced Seismicity during Hydraulic Fracturing of Shale Gas Reservoirs, *Bull. Seismol. Soc. Am.*, *108*(2), 690-701. https://doi.org/10.1785/0120170207

Woessner, J., L. Danciu, D. Giardini, H. Crowley, F. Cotton, G. Grünthal, G. Valensise, R. Arvidsson, R. Basili, M. Betül Demircioglu, S. Hiemer, C. Meletti, R. Musson, A. Rovida, K. Sesetyan, M. Stucchi, and the Seismic Hazard Harmonization in Europe (SHARE) Consortium (2015). The 2013 European Seismic hazard model: key components and results, *Bull. Earth. Eng.* https://doi.org/10.1007/s10518-015-9795-1.

Zechar, J.D. (2010). Evaluating earthquake predictions and earthquake forecasts: a guide for students and new researchers, *Community Online Resource for Statistical Seismicity Analysis*, https://doi.org/10.5078/corssa-77337879.

Zhang, L., M. J. Werner, and K. Goda (2020). Variability of ETAS parameters in global subduction zones and applications to mainshock-aftershock hazard assessment, *Bull. Seismol. Soc. Am.*, *110*(1), 191-212. https://doi.org/10.1785/0120190121.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2002). Stochastic declustering of space-time earthquake occurrences, *J. Am. Stat. Assoc.*, 97(458), 369–380. https://doi.org/10.1198/016214502760046925.

Zhuang, J., D. Harte, M. J. Werner, S. Hainzl, and S. Zhou (2012). Basic models of seismicity: temporal models, *Community Online Resource for Statistical Seismicity Analysis*. https://doi.org/10.5078/corssa-79905851.

Zhuang, J., and S. Touati (2015). Stochastic simulation of earthquake catalogs, *Community Online Resource for Statistical Seismicity Analysis*. https://doi.org/10.5078/corssa-43806322.

Addresses of authors

Simone Mancini (simone@bgs.ac.uk), British Geological Survey, The Lyell Center, Research Avenue South, EH14 4AP, Edinburgh, UK

Maximilian Jonas Werner (max.werner@bristol.ac.uk), School of Earth Sciences, University of Bristol, BS8 1RL Bristol, UK

Margarita Segou (msegou@bgs.ac.uk), British Geological Survey, The Lyell Center, Research Avenue South, EH14 4AP, Edinburgh, UK

Brian Baptie (bbap@bgs.ac.uk), British Geological Survey, The Lyell Center, Research Avenue South, EH14 4AP, Edinburgh, UK

List of Figure Captions

Figure 1. Map view of earthquakes recorded during hydraulic fracturing at the Preston New Road unconventional shale gas site. Events are color-coded by the associated injection stage and their size scales with magnitude. (a) Seismicity between 15 October and 17 December 2018 during and after injection at the PNR-1z well. (b) Seismicity between 15 August and 2 October 2019 during and after injection at the PNR-2 well; grey dots indicate the epicenters of events occurred during operations at PNR-1z. The black lines represent the surface projection of the two wellpaths. Diamonds illustrate the position of the main sleeves worked during the operations at the two wells and are colored by the corresponding injection stages.

Figure 2. Seismicity response to hydraulic fracturing at the Preston New Road site. (a-b) Histograms of the number of $M \geq -1.5$ events per hour (black bars) as a function of time during operations along with the cumulative volume of injected fluid (light blue line) at PNR-1z and

PNR-2, respectively. For illustration purposes, we inserted a time gap during the pause of operations at PNR-1z, which is indicated by the grey area. (c-e) Examples of seismic productivity and earthquake magnitudes vs. time (red circles) in response to the injection history (light blue line) at selected sleeves.

Figure 3. Observed vs. forecasted number of M ≥ -1.5 events at PNR-1z. (a) Illustration of incremental 1-hour forecast timeseries vs. observations at PNR-1z injection sleeves characterized by weak and strong seismicity response as well as during the pause of operations. ETAS2-*bulk* model predictions are shown only during injection periods indicated by the "Inj." label (otherwise its forecasts are identical to ETAS1-*optimized* and ETAS2-*specific*). Black circles indicate the number of observed events in each forecast window. Other symbols represent the mean expected number from the simulations. Bars denote 95% ETAS model simulation ranges. For illustration purposes, during periods of suspended/paused injection data are plotted at 12-hour intervals. (b-d) Observed vs. expected number of events per forecast period over all injection stages. Each symbol indicates one forecast window, which is accepted if the 95% model range (black vertical bars) intersect the diagonal black line. Red symbols denote rejected forecasts (data outside model range); green symbols denote accepted forecasts.

Figure 4. Observed vs. forecasted number of M ≥ -1.5 events at PNR-2. (a) Illustration of incremental 1-hour forecast timeseries vs. observations at PNR-2 injection sleeves characterized by weak and strong seismicity response as well as during the pause of operations. ETAS2-*bulk* model predictions are shown only during injection periods indicated by the "Inj." label (otherwise its forecasts are identical to ETAS1-*optimized* and ETAS2-*specific*). Black circles indicate the number of observed events in each forecast window. Other symbols represent the mean expected number from the simulations. Bars denote 95% ETAS model simulation ranges. For illustration purposes, during periods of suspended/paused injection data are plotted at 2-hour intervals. (b-d) Observed vs. expected number of events per forecast period over all injection stages. Each symbol indicates one forecast window, which is accepted if the 95% model range (black vertical bars) intersect the diagonal black line. Red symbols denote rejected forecasts (data outside model range); green symbols denote accepted forecasts.

Figure 5. Cumulative log-likelihood timeseries. ETAS models tested on (a) PNR-1z and (b) PNR-2.

Figures with captions

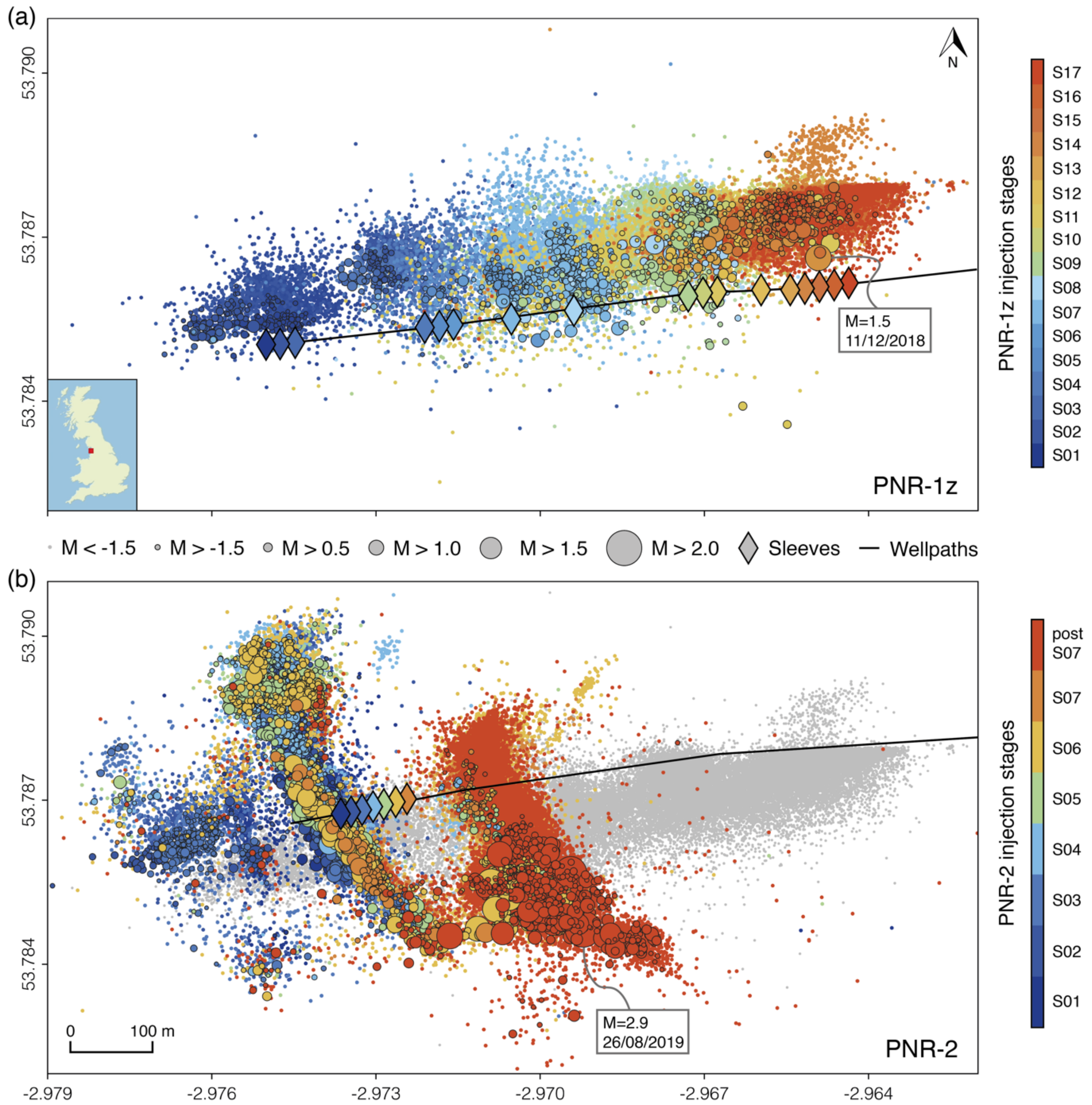

Figure 1. Map view of earthquakes recorded during hydraulic fracturing at the Preston New Road unconventional shale gas site. Events are color-coded by the associated injection stage and their size scales with magnitude. (a) Seismicity between 15 October and 17 December 2018 during and after injection at the PNR-1z well. (b) Seismicity between 15 August and 2 October 2019 during and after injection at the PNR-2 well; grey dots indicate the epicenters of events occurred during operations at PNR-1z. The black lines represent the surface projection of the two wellpaths. Diamonds illustrate the position of the main sleeves worked during the operations at the two wells and are colored by the corresponding injection stages.

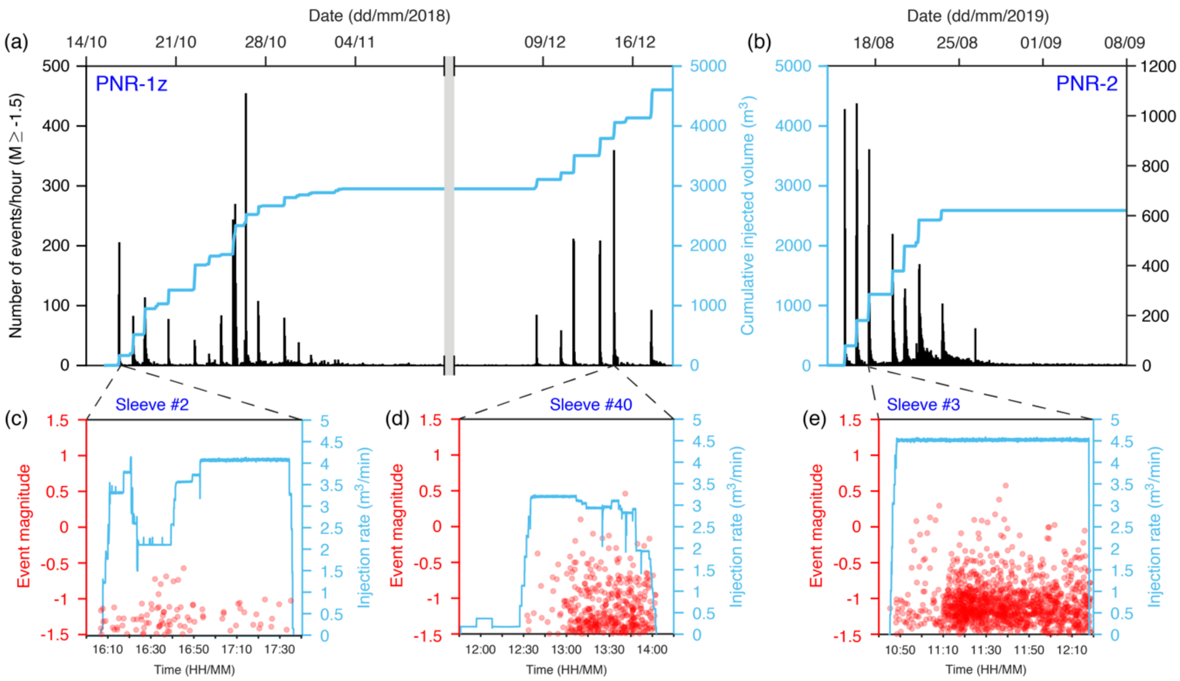

Figure 2. Seismicity response to hydraulic fracturing at the Preston New Road site. (a-b) Histograms of the number of M ≥ -1.5 events per hour (black bars) as a function of time during operations along with the cumulative volume of injected fluid (light blue line) at PNR-1z and PNR-2, respectively. For illustration purposes, we inserted a time gap during the pause of operations at PNR-1z, which is indicated by the grey area. (c-e) Examples of seismic productivity and earthquake magnitudes vs. time (red circles) in response to the injection history (light blue line) at selected sleeves.

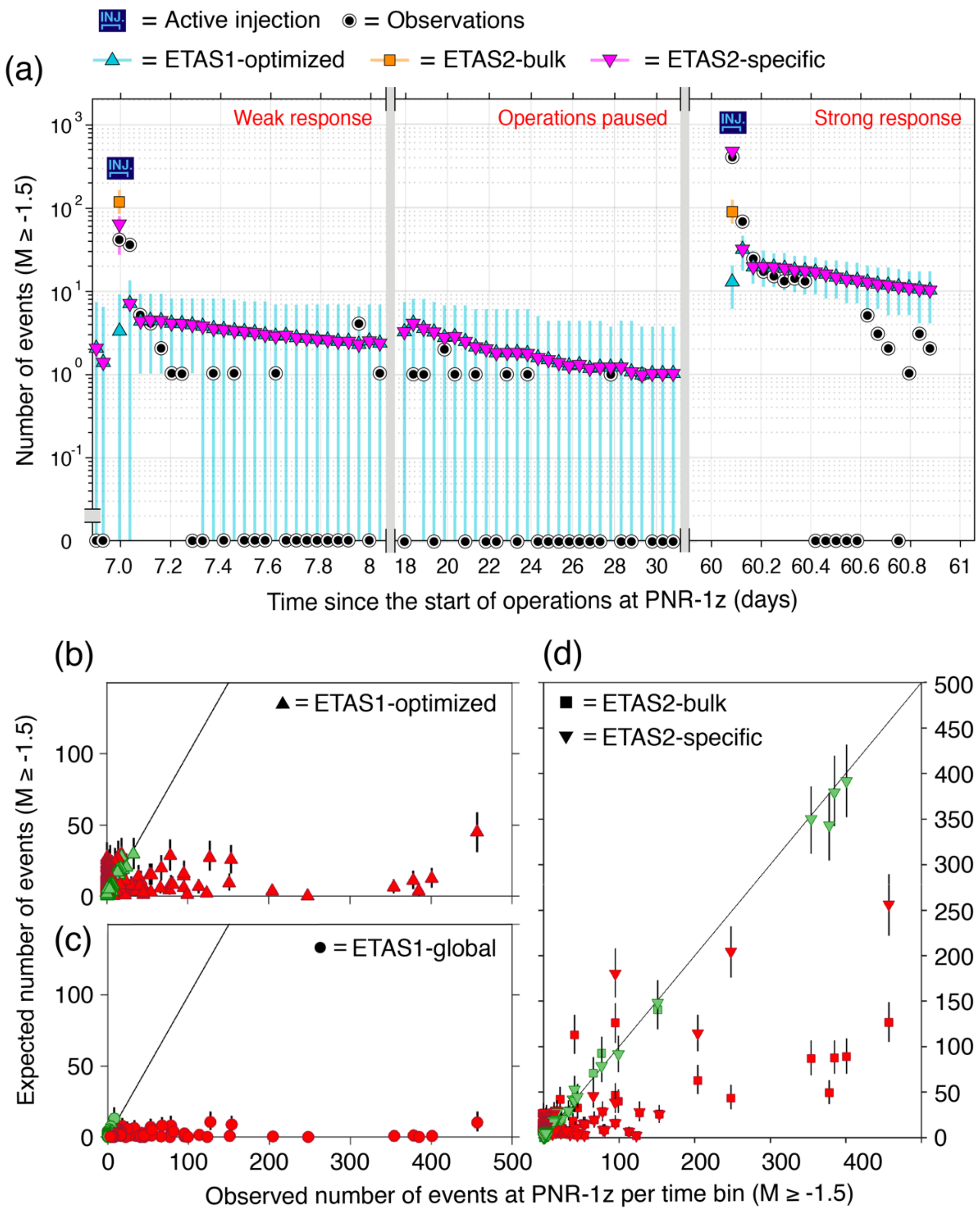

Figure 3. Observed vs. forecasted number of M ≥ -1.5 events at PNR-1z. (a) Illustration of incremental 1-hour forecast timeseries vs. observations at PNR-1z injection sleeves characterized by weak and strong seismicity response as well as during the pause of operations. ETAS2-*bulk* model predictions are shown only during injection periods indicated by the “Inj.” label (otherwise its forecasts are identical to ETAS1-*optimized* and ETAS2-*specific*). Black circles indicate the number of observed events in each forecast window. Other symbols represent the mean expected number from the simulations. Bars denote 95% ETAS model simulation ranges. For illustration purposes, during periods of suspended/paused injection data are plotted at 12-hour intervals. (b-d) Observed vs. expected number of events per forecast period over all injection stages. Each symbol indicates one forecast window, which is accepted if the 95% model range (black vertical bars) intersect the diagonal black line. Red symbols denote rejected forecasts (data outside model range); green symbols denote accepted forecasts.

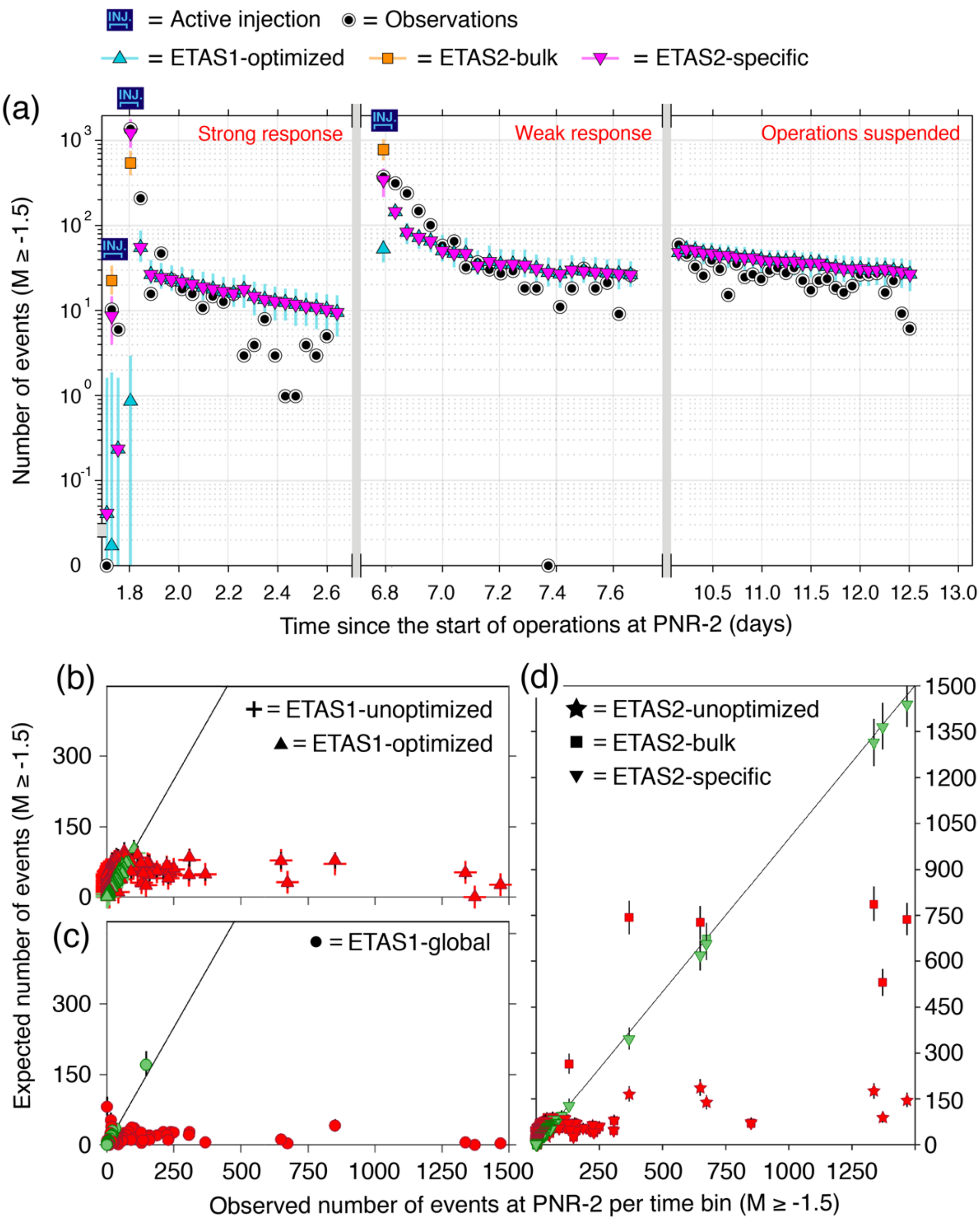

Figure 4. Observed vs. forecasted number of M ≥ -1.5 events at PNR-2. (a) Illustration of incremental 1-hour forecast timeseries vs. observations at PNR-2 injection sleeves characterized by weak and strong seismicity response as well as during the pause of operations. ETAS2-*bulk* model predictions are shown only during injection periods indicated by the "Inj." label (otherwise its forecasts are identical to ETAS1-*optimized* and ETAS2-*specific*). Black circles indicate the number of observed events in each forecast window. Other symbols represent the mean expected number from the simulations. Bars denote 95% ETAS model simulation ranges. For illustration purposes, during periods of suspended/paused injection data are plotted at 2-hour intervals. (b-d) Observed vs. expected number of events per forecast period over all injection stages. Each symbol indicates one forecast window, which is accepted if the 95% model range (black vertical bars) intersect the diagonal black line. Red symbols denote rejected forecasts (data outside model range); green symbols denote accepted forecasts.

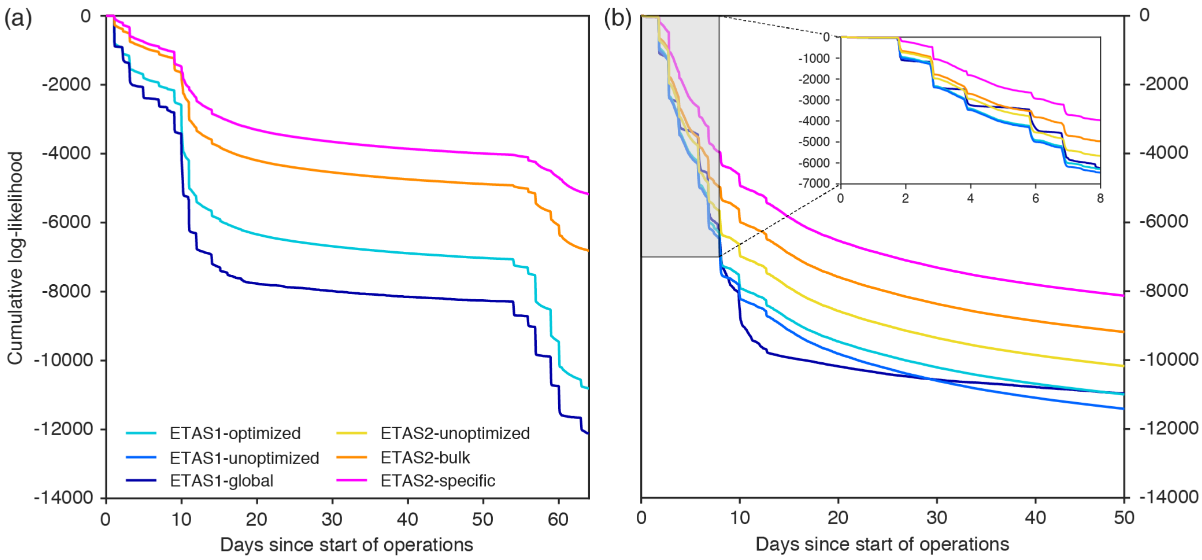

Figure 5. Cumulative log-likelihood timeseries. ETAS models tested on (a) PNR-1z and (b) PNR-2.