

<https://doi.org/10.1038/s43247-025-02881-2>

The bound growth of induced earthquakes could de-risk hydraulic fracturing



Ryan Schultz¹ , Federica Lanza¹, Ben Dyer², Dimitrios Karvounis², Rémi Fiori², Peidong Shi¹,
Vanille Ritz¹, Linus Villiger¹, Peter Meier² & Stefan Wiemer¹

The world's energy supply depends critically on hydraulic fracturing (HF) to access otherwise uneconomical resources. Unfortunately, HF also has the potential to induce larger earthquakes – with some projects being prematurely terminated because of perceived earthquake risks. To de-risk HF, we use a suite of statistical tests to discern if some physical process has restricted the growth of earthquake magnitudes. We show that all stage stimulations at both UK PNR-1z and Helsinki St1 indicate bound fracture growth, implying a more controllable operation. Contrastingly, stimulations at Utah FORGE and UK PNR-2 sequentially transitioned into unbound fault reactivation. The problematic stages (that ultimately led to the termination of PNR-2) are clearly distinguishable. We postulate that our research can discriminate fracture stimulation from fault reactivation, contributing to the de-risking of HF operations worldwide. Our statistical tests provide a framework for model falsification, which can guide physical insights into the bounding processes.

The insatiable demand for energy continues to grow each year. In recent decades, hydraulic fracturing (HF) has been revolutionary for unlocking trapped hydrocarbon energy sources: fluids injected under high-pressure cause subsurface rocks to fail, stimulating fractures that act as permeable fluid pathways needed to extract resources¹. This stimulation process can either propagate failures in tension or in shear; these failures can happen on either pre-existing fractures or newly created fractures. However, the urgency for more energy is also complicated by climate goals and decarbonisation needs. Enhanced Geothermal Systems (EGS)² could play a part in the energy transition, since they have the potential to fill this niche by producing abundant, green, sustainable, baseload power³. EGS taps the limitless geothermal potential of the deep earth through stimulation techniques (e.g., HF)⁴: stimulated fractures can also act as fluid pathways needed for heat exchange. EGS could be on the cusp of widespread adoption, considering recent successes⁵. However, the deployment of HF techniques is not without risks.

The HF process also has the potential to induce earthquakes^{6,7}. If stimulated fractures intersect a pre-existing fault, then the increase of pore pressure can cause fault slip reactivation⁸. These earthquakes can be consequential via economic/human losses, premature project termination, or resource development moratoriums^{9,10}. For example, magnitude ~5 earthquakes have caused hundreds of millions of dollars in economic losses from both petroleum and EGS operations using HF^{10,11}. On the other hand, there are also numerous HF projects that did not encounter detrimental earthquake impacts^{12–15}. Thus, there is a clear need to effectively manage these earthquakes and mitigate their consequences.

Guidelines to managing induced earthquake risks exist^{16,17}; these provide templates for pre-screening assessments, risk evaluation, monitoring, and defining operational endpoints. The core risk management framework is the traffic light protocol, which details how an operation should proceed, depending on the largest magnitude observed¹⁸. Risk-based approaches can identify magnitude thresholds for when an operation must stop, to remain below acceptable risk tolerances^{19–21}. Despite these efforts, state-of-the-art procedures can still fail to robustly discern and then react to risky/consequential projects^{16,17} – as evidenced by high-profile cases that caused nuisance, damage, human losses, premature project termination, or development moratoriums^{9–11}.

Fundamental to this risk management process is accurately anticipating the growth of the largest earthquakes²². Unfortunately, this is a poorly understood topic in seismology. What is generally accepted, is that the statistics of earthquake magnitudes (M) typically follows the Gutenberg-Richter magnitude frequency distribution (GR-MFD)²³. A natural consequence of the GR-MFD is that the largest magnitude events (M_{LRG}) scale with the size of the catalogue (N)²⁴. It is debatable whether any other physical process bounds the growth of magnitudes: i.e., an upper bound magnitude (M_{MAX}) that restricts the (statistically expected) growth of magnitudes. Some M_{MAX} models have been hypothesised. In the simplest model, an M_{MAX} can arise from the finite size of the fault extent, since magnitude is tied to shear slip across a fault area²⁵. More contentiously, M_{MAX} could also be some injection-dependent process that changes with time; for example, either evolving with seismic moment release (McGarr-like models)^{26,27} or scaling with self-arrested rupture dynamics (Galis model)²⁸. Investigating

¹Swiss Seismological Service, ETH Zürich, Zürich, Switzerland. ²Geo-Energie Suisse, Zürich, Switzerland. ✉ e-mail: Ryan.Schultz@sed.ethz.ch

the physics of these M_{MAX} models would benefit from approaches that could verify/falsify their presence in a catalogue.

However, there is a lack of empirical approaches to identify/falsify these M_{MAX} models – the identification of M_{MAX} solely from a catalogue is fraught with difficulty. Traditional approaches to this M_{MAX} problem have found it to be solvable only in ideal conditions^{29,30}, often requiring a significant degree of truncation for robust identification. On the other hand, verifying the existence of M_{MAX} within an induced earthquake sequence is significant, because then the risk of induced seismicity is inherently more controllable²². Given this gap between need/difficulty, expert opinion panels on M_{MAX} ³¹ often substitute for quantitative science out of necessity.

In this study, we consider a reformulation of this M_{MAX} problem. Here, we implement a novel suite of statistical tests to discern if a sequence of earthquake magnitudes was bound by an M_{MAX} ³². These tests compare event magnitudes (M) against ‘jumps’ in the sequence of largest earthquakes (ΔM_{LRG}), to infer the presence of a maximum upper bound (M_{MAX}). Simply put, the statistical distributions of M and ΔM_{LRG} are identical when a sequence is unbound, but distinguishable in the presence of an M_{MAX} (Fig. 1)—this is because the presence of an M_{MAX} restricts the allowable sizes of ΔM_{LRG} . Our suite of three statistical tests exploits this fact. The first test uses the Kolmogorov-Smirnov test (KS-test) to compare M and ΔM_{LRG} observations, to discern if they are drawn from the same distribution; this tests for the existence of an M_{MAX} . The second test uses a maximum likelihood estimation (MLE-test) to quantify the value of M_{MAX} . The third test computes ensemble weights (EW-test) through Akaike and Bayesian Information Criteria; this selects the best M_{MAX} model amongst several hypothesised choices, given the catalogue data. We refer to this suite of three novel tests as CAP-tests. These CAP-tests are generally more sensitive to M_{MAX} bounds than traditional approaches³². Full details on terminology, methods, guidelines to the interpretation of results, and synthetic examples are all available in the Supplementary Information (Text S1–S4).

We demonstrate how this workflow (Fig. 2 & S1) can be used to identify concerning cases of induced seismicity. This is accomplished by contrasting cases of induced seismicity related to HF (Fig. S7). In the first case, we consider Preston New Road 1z (PNR-1z)³³: a shale gas HF operation in the UK that injected $\sim 4200 \text{ m}^3$ over 17 (of 41 planned) stages and produced earthquakes up to M_L 1.6 in 2018. In the second case, we consider Helsinki St1¹⁵: an EGS operation in Finland that injected $\sim 18,200 \text{ m}^3$ over five stages and produced earthquakes up to M_W 1.9 in 2018. Our tests find that all stage stimulations in these two cases follow a bound process, dominantly suggesting fracture growth via stage stimulations. In the contrasting cases, we consider data from the Frontier Observatory for Research in Geothermal Energy (FORGE)³⁴ and Preston New Road 2 (PNR-2)³⁵. FORGE is an EGS operation in Utah that injected $\sim 18,500 \text{ m}^3$ over ten stages and produced earthquakes up to M_W 1.8 in 2022 and 2024. PNR-2 is a shale gas HF operation in the UK that injected $\sim 2500 \text{ m}^3$ over seven stages and produced earthquakes up to M_L 2.9 in 2019. Our tests find that the first handful of stages follow a bound process (i.e., restricted by a volume-based M_{MAX}). However, the subsequent (and problematic) stages were clearly unbound, dominantly suggesting fault reactivation via stage stimulations. The PNR-2 case was prematurely terminated because of induced seismicity concerns.

Bound growth

Bound growth at PNR-1z

HF was used in the UK at three wells, to extract the hydrocarbons locked within the Mississippian aged Bowland Shale. The operation at PNR-1z, near Blackpool, was the second HF operation in the UK³³. The first set of stimulations (October–November 2018) produced moderate seismic response; these events were concerning due to their magnitude above the red-light threshold, which caused an operational pause. The second set of stimulations (December 2018) resumed after this pause. The largest event (M_L 1.6) occurred on 11 December 2018³³. All events appear to be

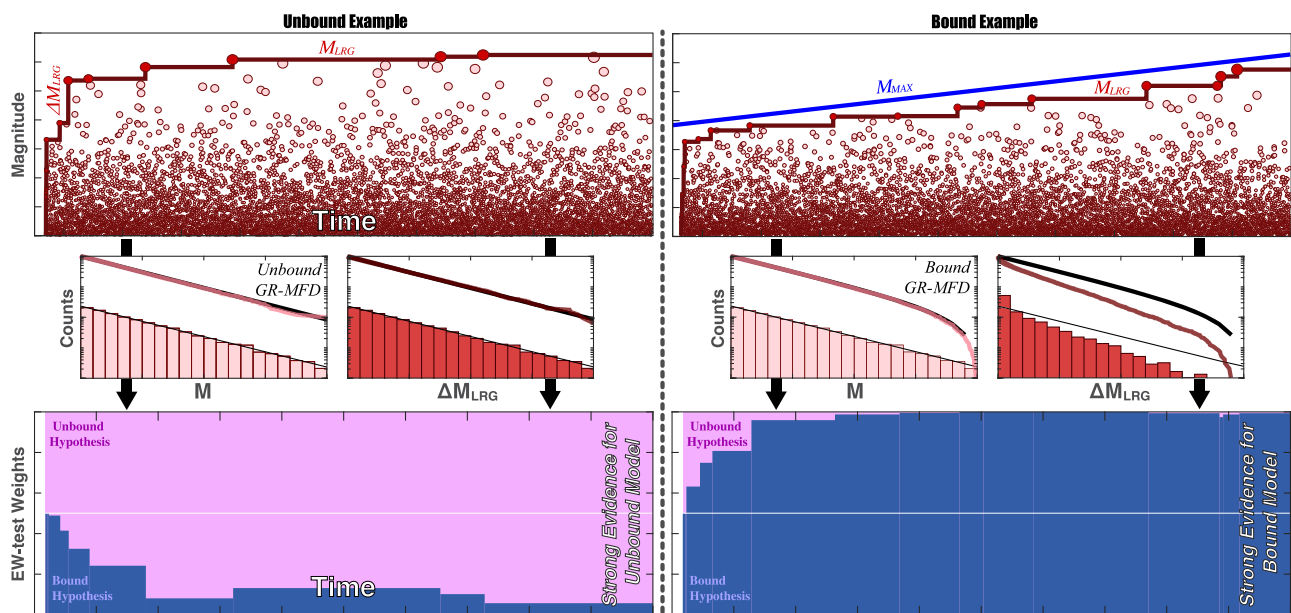


Fig. 1 | Schematic demonstration of observational and methodological concepts. The upper panels show hypothetical catalogues of earthquakes. Earthquake magnitudes (M , pink circles), the sequence of largest events (M_{LRG} , red line), and jumps in the largest event magnitudes (ΔM_{LRG}) can be directly observed. Potentially, some unobservable physical process could be bounding these catalogues (M_{MAX} , blue line). Left and right panels detail contrasting concepts for an unbound and bound catalogue, respectively. If there is an M_{MAX} , then it can be statistically inferred from the observables. Middle panels show the differences between distributions of M and M_{LRG} , when bounded by M_{MAX} (or not). The distribution of M follows a GR-MFD: analytical cumulative (thick black line) and non-cumulative (thin black line) agree

with numerical cumulative (pink line) and non-cumulative (pink bars) distributions. The distribution of ΔM_{LRG} will differ from the GR-MFD when bounded by M_{MAX} . Bottom panels show the results of EW-tests using these concepts. Weights of two hypothesised models, an unbound M_{MAX} (pink area) and a bound M_{MAX} (blue area), change as new values of M_{LRG} are observed. EW-tests can quickly infer the presence of the true M_{MAX} model, from the equivocal a priori assumption (white horizontal line). Evidence for a model ‘winning’ over another becomes substantial/positive, strong, and decisive when the ratio-of-weights (i.e., odds ratio) is 3, 10, and 100, respectively.

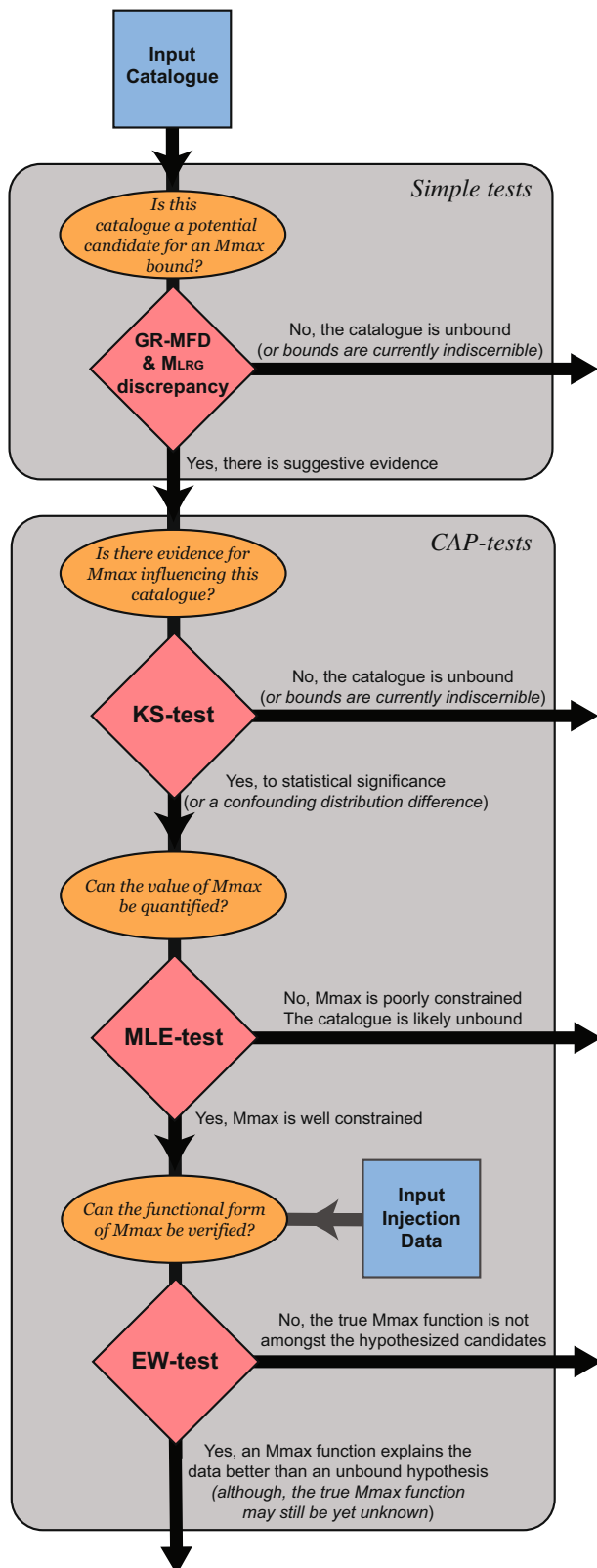


Fig. 2 | Logical sequence of statistical tests to discern M_{MAX} bounds. Simple-tests are used to pre-screen for potentially bound catalogues by considering the degree-of-truncation. Afterwards, CAP-tests can more rigorously identify bound cases.

hydraulically linked into a single cluster (Fig. S8; Text S5.1). The success of operations at PNR-1z enabled the subsequent operation at PNR-2.

To begin assessing if some process might be restricting magnitude growth, we fit the GR-MFD²³ and perform the simple-tests (Text S5.1). The

earthquakes here appear to be deficient in large magnitude events (Fig. S9). Correspondingly, the discrepancy between the largest observed event and the largest expected event magnitudes is $-0.5 M_{\text{W}}$, which is a rare occurrence ($\sim 2\%$) purely by chance. These initial assessments are suggestive of some M_{MAX} upper bound restricting the PNR-1z catalogue growth.

Next, we use the CAP-tests to detect and assess the potential for M_{MAX} more rigorously (Fig. 3 & S10). Average KS-test confidences are 99.82% (Text S5.1), where 95% is a common threshold to declare statistical significance. Next, we seek test hypotheses for the functional form of M_{MAX} : our EW-tests focus on a McGarr-like model (i.e., log-proportional to injected volume; $M_{\text{MAX}} \propto \log_{10}(V^1)$ ^{26,27}, the Galis-like model (i.e., $M_{\text{MAX}} \propto \log_{10}(V^{3/2})$ ²⁸, and the unbound null hypothesis. Similarly, EW-tests also show strong evidence for an M_{MAX} bound process. By ~ 100 event observations, respective model weights are already {0.39, 95.69, 3.92%} which means the Galis-like M_{MAX} model is ~ 24 times more likely than unbound (\log_{10} odds ratio of +1.4; Fig. 3a). By the end of stimulations, the Galis-like model is >100 times more likely than the unbound model (Fig. 3a)—which surpasses the definitive evidence threshold.

Bound growth at Helsinki St1. HF was used in a deep geothermal well to create an EGS in the Precambrian Svecofennian basement rocks near Helsinki, Finland. The St1 operation injected $\sim 18,200 \text{ m}^3$ of fluid within five stages from June–July 2018¹⁵. These stimulations produced a moderate seismic response; the largest event (M_{W} 1.9) occurred on 8 July 2018. All events appear to be hydraulically linked into a single cluster (Fig. S11; Text S5.2), which is consistent with prior interpretations¹⁵. The success of this operation allowed for later restimulation of the reservoir.

To assess if some process might be restricting magnitude growth, we first employ the simple tests based on GR-MFD analyses (Text S5.2). The earthquakes here appear to be deficient in large magnitude events (Fig. S12). Correspondingly, the magnitude discrepancy between the largest observed event and the largest expected event is $-0.7 M_{\text{W}}$, which is a rare occurrence ($< 1\%$) purely by chance. These initial assessments are suggestive of some M_{MAX} upper bound restricting the St1 catalogue growth.

Next, we use the CAP-tests to detect and assess the potential for M_{MAX} more rigorously (Fig. 4 & S13). Average KS-test confidences are 98.17% (Text S5.2). Similarly, EW-tests also show strong evidence for an M_{MAX} bound process (Fig. 4): McGarr-like and Galis-like models are >100 (\log_{10} odds ratio of +2.2) and ~ 57 (\log_{10} odds ratio of +1.8) times more likely than unbound models, respectively. This dataset is somewhat ambiguous between either McGarr-like or Galis-like models best explaining the data.

We note that we have examined supporting cases, including the EGS stimulations at Soultz-sous-Forêts in France³⁶. There, we also find some evidence for a bound process restricting the growth of earthquake magnitudes during their stimulations (Text S5.5; Figs. S27–S36).

The transition to unbound growth

The transition to unbound growth at FORGE

Utah FORGE is a geothermal experiment ($\sim 16 \text{ km}$ NE of Milford) aimed at commercialising EGS³⁴. Two deviated wells form a doublet heat exchange system (16 A injector & 16B producer). The first three stages of well 16 A were stimulated in 2022; in 2024, stage 3 was restimulated, followed by the remaining stimulations (Figs. S14–S15). Originally, nine stages were planned for well 16 A at FORGE. However, frack barriers were encountered at stage 6, preventing fracture stimulation; in response, a tenth stage was added. Correspondingly, stage 6 separates the adjacent stages into two hydraulically separate compartments. Overall, stages are compartmentalised into hydraulically independent clusters as follows: cluster 1 (stages 1–2), cluster 2 (stages 3–6), and cluster 3 (stages 7–10). Additional justifications for this clustering choice are provided in the Supplementary Information (Text S5.3) and in prior studies³⁵.

To assess the potential for bound/unbound earthquake growth, we first employ the simple GR-MFD analyses (Text S5.3). The earthquakes here appear systematically skewed in larger magnitude events: both in terms of the full catalogue (Fig. S16) and on a cluster-by-cluster basis (Fig. S17).

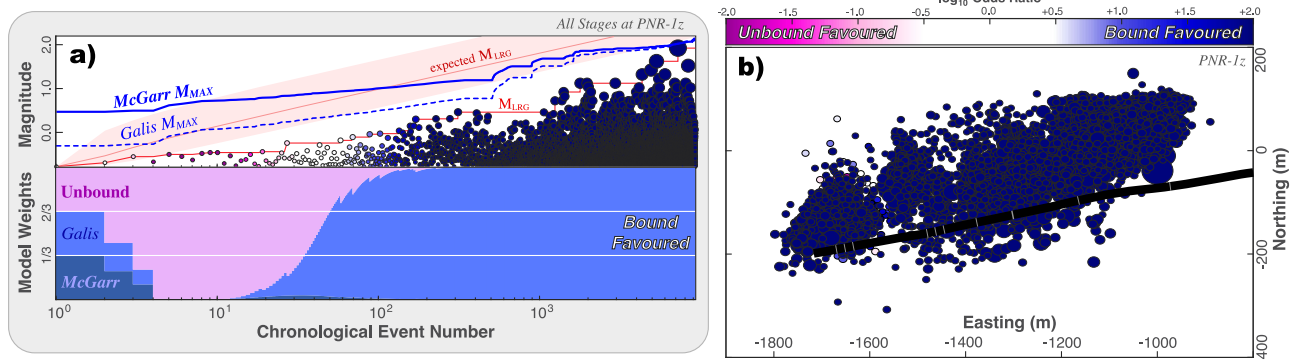


Fig. 3 | EW-tests applied to PNR-1z. **a** EW-tests are shown for all the stages at PNR-1z. The input data includes earthquake magnitudes (circles), the sequence of observed M_{LRG} (red line), alongside the expected M_{LRG} at the 10/50/90 percentiles (red area); two possible injection-based M_{MAX} models are considered (blue lines). Correspondingly, the ensemble weights from the EW-test are shown; bound models

are highly favourable. **b** Map view showing the locations of earthquakes (circles) alongside the well trajectory (black line) and stage locations (grey lines). Throughout the figure, all earthquakes are colour-coded accordingly to their likelihood of being bound or unbound from EW-tests.

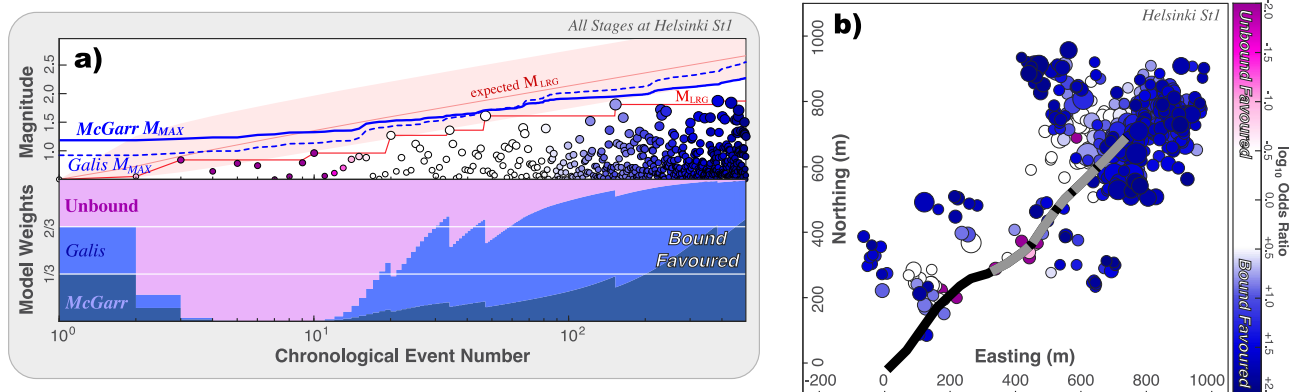


Fig. 4 | EW-tests applied to Helsinki St1. **a** EW-tests are shown for all the stages at Helsinki St1. The input data includes earthquake magnitudes (circles), the sequence of observed M_{LRG} (red line), alongside the expected M_{LRG} at the 10/50/90 percentiles (red area); two possible injection-based M_{MAX} models are considered (blue lines). Correspondingly, the ensemble weights from the EW-test are shown; bound models

are highly favourable. **b** Map view showing the locations of earthquakes (circles) alongside the well trajectory (black line) and stage locations (grey lines). Throughout the figure, all earthquakes are colour-coded accordingly to their likelihood of being bound or unbound from EW-tests.

Furthermore, we also consider the discrepancy between the largest observed event and largest expected event magnitudes. Correspondingly, this difference ranges between -0.4 and $-1.3 M_W$ for the first two clusters, but is $+0.1 M_W$ for the third cluster. These are rare occurrences ($\sim 1\%$) for the first two clusters. Together, these initial assessments are suggestive of some M_{MAX} upper bound restricting the FORGE catalogue growth for clusters 1–2.

To more rigorously examine M_{MAX} upper bounds, we use the suite of CAP-tests. Both KS-test and MLE-test results provide strong evidence for the existence of M_{MAX} upper bounds. Specifically, KS-test confidences range between 99 and 100% for clusters 1 and 2. Thus, there is strong evidence for some physical M_{MAX} process. On the other hand, cluster 3 does not meet the 95% confidence threshold. For the EW-tests on stage 3 in 2022, within ~ 100 event observations, respective model weights are already $\{70.44, 17.69, 11.87\}$ which means that the McGarr-like M_{MAX} model is ~ 6 times more likely (\log_{10} odds ratio of $+0.8$; Fig. 5a). Furthermore, clusters 1 and 2 show strong evidence for a bound process after EW-tests (Fig. 5). Specifically, the McGarr-like model appears to best explain these bound clusters (Figs. S18–S21). On the other hand, the EW-tests for cluster 3 show strong evidence for an unbound process: the unbound model is ~ 50 times more likely than the best-fitting bound

model. These results hold, even after various perturbation/sensitivity tests (Text S6–S7).

The transition to unbound growth at PNR-2

The operation at PNR-2, near Blackpool, was the most seismically problematic³⁷ and last of the three HF operations in the UK. The first 1–3 stages only produced moderate seismic response in the western-most clusters, while more concerning seismicity was encountered at the eastern-most clusters, which predominantly occurred during stages 5–7 (Fig. S22). The eastern-most cluster culminated in the M_L 2.9 event on 26 August 2019, which occurred >72 hours after the shut-in of stage 7³⁷. This event caused the premature termination of PNR-2, after only the first seven stages (of 47 planned) were completed. Ultimately, these earthquakes spurred a resource development moratorium in the UK³⁸.

To assess the potential for bound/unbound earthquake growth, we repeat the simple GR-MFD analyses. The western-most clusters appear to be deficient in large magnitude events, while the eastern-most cluster is abundant (Figs. S23–S24). Correspondingly, the discrepancy between the largest observed event and largest expected event magnitudes is $-1.2 M_W$ and $+0.8 M_W$, respectively. This is a rare occurrence ($< 1\%$) for the western-most cluster, but not the eastern-most. Together, these initial assessments

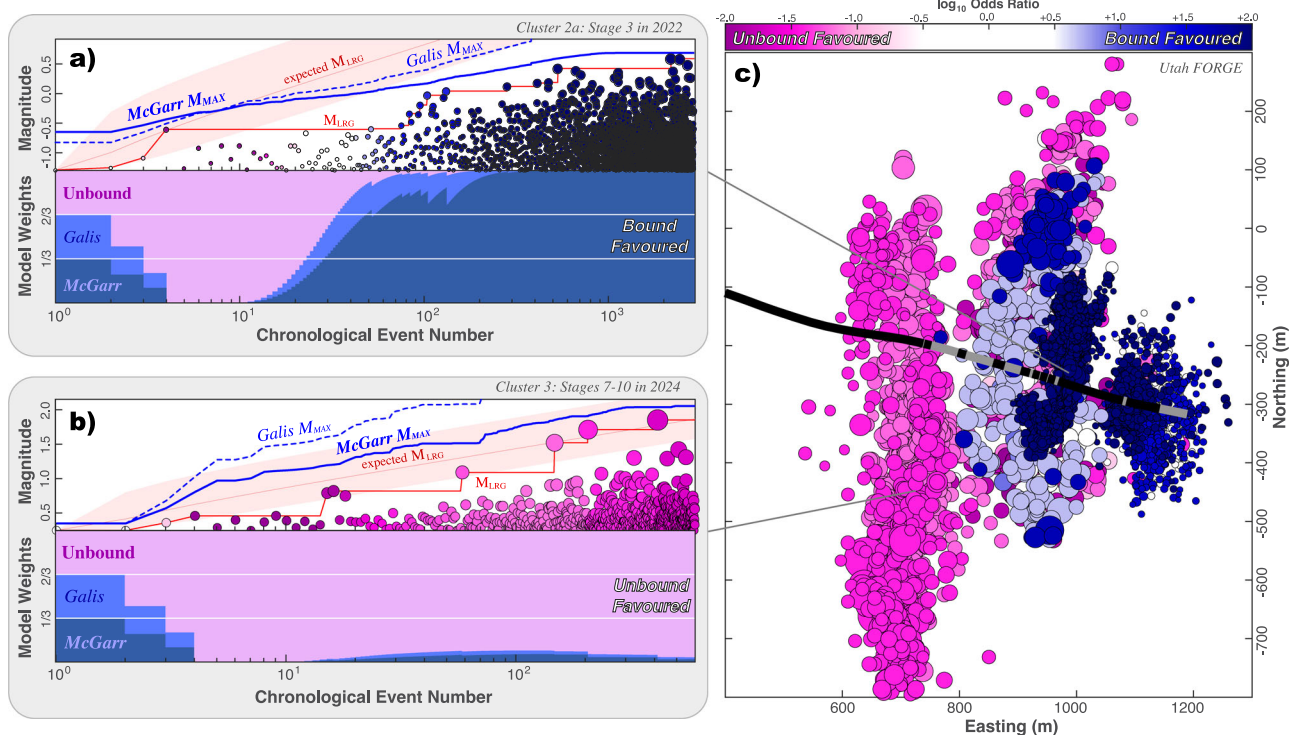


Fig. 5 | EW-tests applied to FORGE. EW-tests are shown for **a** the 2022 stage 3 and **b** for the 2024 cluster 3 (stages 7–10). The input data includes earthquake magnitudes (circles), the sequence of observed M_{LRG} (red line), alongside the expected M_{LRG} at the 10/50/90 percentiles (red area); two possible injection-based M_{MAX} models are considered (blue lines). Correspondingly, the ensemble weights from the

EW-test are shown; bound models are highly favourable in most cases (Figs. S18–S21). **c** Map view showing the locations of earthquakes (circles) alongside the 16 A well trajectory (black line) and stage locations (grey lines). Throughout the figure, all earthquakes are colour-coded accordingly to their likelihood of being bound or unbound from EW-tests.

are suggestive of some M_{MAX} upper bound restricting the PNR-2 catalogue growth at the western-most cluster, but not the eastern-most.

Again, we use the CAP-tests to detect and assess the potential for M_{MAX} more rigorously (Fig. 6). KS-test and MLE-test results show strong evidence for M_{MAX} in the western-most case (i.e., > 99.99% confidence), but not the eastern-most case (Text S5.4). Similarly, EW-tests also show strong evidence for a bound process, but only for the western-most cluster (Fig. 6a). The western-most cluster is somewhat ambiguous between either McGarr-like or Galis-like models best explaining the data. On the other hand, the eastern-most cluster definitively indicates an unbound process, being ~100 times more likely than the best-fitting bound model.

Implications

Practical risk management. Induced earthquakes pose considerable risks. However, the chances of encountering earthquake consequences for a bound/unbound stage stimulation are significantly different (Text S8; Fig. S41). Thus, the ability of CAP-tests to discern between these two extremes constitutes an important reactive risk mitigation measure: bound stages could proceed as planned, while unbound stages could be treated with caution (or even prematurely stopped and then skipped). This would be especially important for cases where earthquake magnitude growth is approaching the red-light^{19–21}—the last-possible stopping-point before a regulatory intervention is required. In this sense, CAP-tests provide information on how/if ‘jumps’ in earthquake magnitudes (or trailing seismicity) are restricted, which is consistently identified as an important factor for managing induced seismicity^{22,39,40}.

To elaborate, we separate bound/unbound clusters using just the EW-test results (Fig. S42). We use an odds ratio threshold of 3-to-1 (i.e., \log_{10} of +0.5) to discern a bound case to statistical significance; all subsequent ΔM_{LRG} events are labelled bound. Clusters below this threshold are considered unbound. We then examine the distribution of ΔM_{LRG} from these two catalogue subsets (Fig. 7); we find that they are different to statistical

significance (KS-test confidence 99.8%). The unbound cases are well-fit to the GR-MFD, with a typical b -value of 1.73 ± 0.18 (R^2 0.989) and ΔM_{LRG} magnitude jumps up to +1.2. On the other hand, the bound cases are not well-fit to the GR-MFD, with an ‘unphysical’ b -value of 3.02 ± 0.30 (R^2 0.967) and ΔM_{LRG} abruptly truncated near +0.3.

First, this finding is noteworthy, since it provides positive feedback that the CAP-tests are working correctly. Theoretically, unbound ΔM_{LRG} values are expected to follow the same distribution as their underlying magnitudes; on the other hand, bound ΔM_{LRG} values deviate from their GR-MFD³²—because M_{MAX} restricts bound ΔM_{LRG} values, creating apparently steeper b -values (Fig. 1). This expected difference in ΔM_{LRG} distributions is consistent with the observations from our EW-test separation (Fig. 7). More generally, this biasing phenomenon could potentially explain the anomalously high b -values typically associated with earthquakes encountered during HF stimulation⁴¹, since a temporally growing M_{MAX} produces a long and pronounced roll-off at the high-magnitude tail of the GR-MFD.

Furthermore, this finding is significant, since these bound cases are appreciably less hazardous. For example, we can consider a hypothetical situation where the operator approaches a red-light. If M_{LRG} was currently 0.30 M away from the red-light, the next largest event would trigger the red-light in ~30% of unbound cases but only ~4% of the bound cases (Fig. 7). This is especially important considering that earthquake consequences scale exponentially with magnitude⁴². Overall, this hypothetical scenario demonstrates how CAP-tests can be used to quantify magnitude jumps, providing insights into red-light designs and operational decision-making. Correspondingly, we note that operations that encountered significant consequences or were prematurely terminated (e.g., Basel, Pohang, PNR-2) tended to encounter unbound earthquakes³².

This approach could find near-future use in *post hoc* analyses between stage stimulations – especially in tandem with completion engineering information. Potentially even providing complementary information for identifying hydraulic connectivity between stages. While we have used

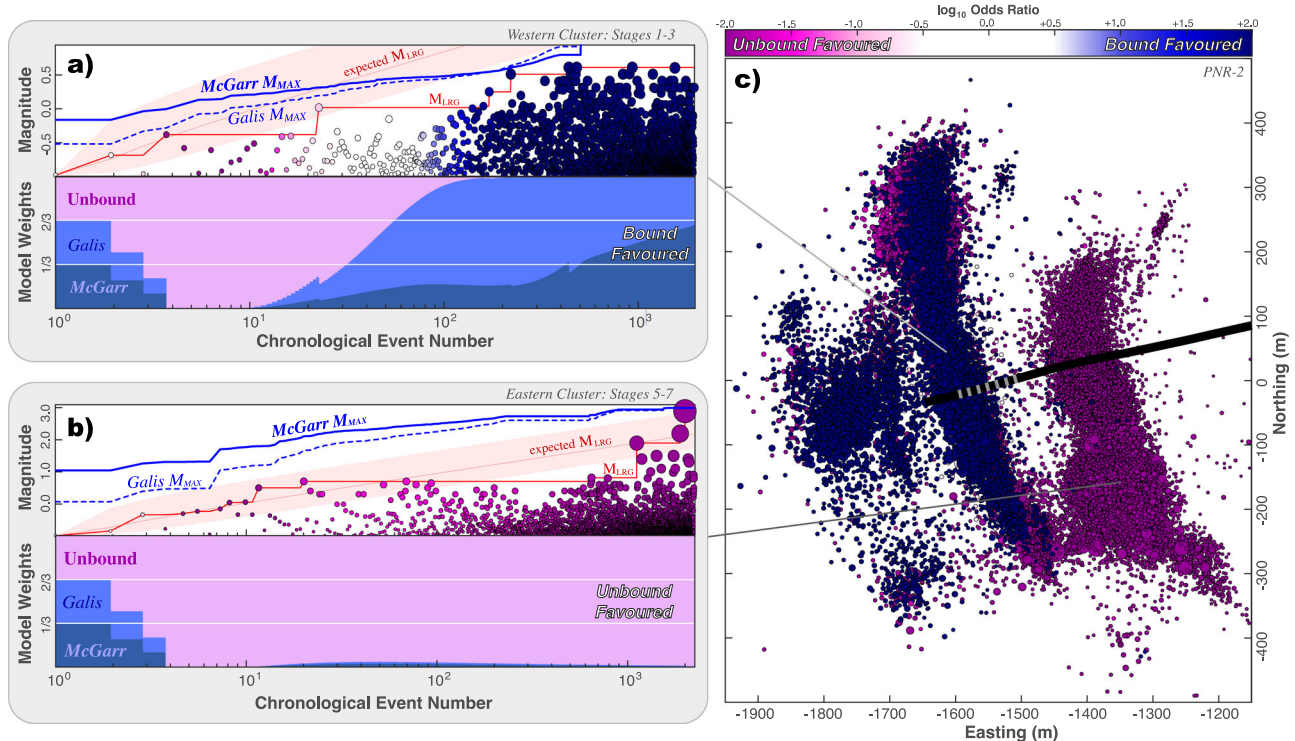


Fig. 6 | EW-tests applied to PNR-2. EW-tests are shown for **a** the western-most and **b** eastern-most clusters. The input data includes earthquake magnitudes (circles), the sequence of observed M_{LRG} (red line), alongside the expected M_{LRG} at the 10/50/90 percentiles (red area); two possible injection-based M_{MAX} models are considered (blue lines). Correspondingly, the ensemble weights from the EW-test are shown; the seismic response is spatially partitioned into eastern/western clusters, where stages

sequentially transition from bound into unbound (Figs. S25–S26). **c** Map view showing the locations of earthquakes (circles) alongside the PNR-2 well trajectory (black line) and stage locations (grey lines). Throughout the figure, earthquakes are colour coordinated according to their likelihood of being bound or unbound from EW-tests.

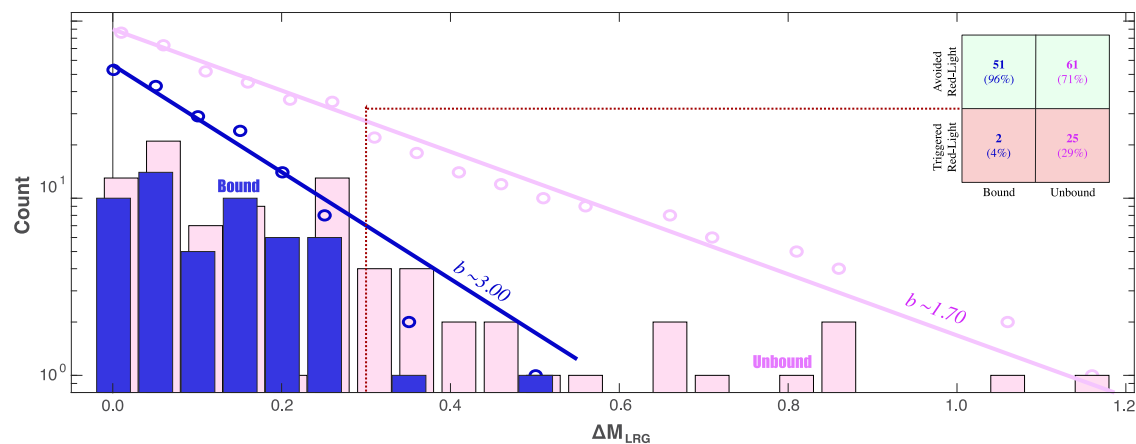


Fig. 7 | Comparing ΔM_{LRG} distributions. Clusters were separated into bound (blue) or unbound (pink) groups using EW-tests. The ΔM_{LRG} data are plotted as both the cumulative (circles) and non-cumulative (bars) distributions, alongside the

GR-MFD fit to the data (solid line). A hypothetical traffic light protocol (dashed line) informs a confusion matrix (inset), that delineates the amount/proportion of bound/unbound events that might trigger the red-light.

CAP-tests to demonstrate this approach on HF and EGS cases, we anticipate this approach will be useful for risk management of other types of induced seismicity (e.g., disposal, mining, reservoir impoundment) as well as tectonic earthquakes. Furthermore, McGarr/Galis predictions of small magnitude M_{MAX} at low volumes suggest the potential to test these models at smaller scales, or use deep catalogues to improve warning times. With further development, this type of approach could be used to quantify logic-tree weights of M_{MAX} within seismic hazard analysis³¹ or to inform seismicity forecasts with real-time data. Although, we caution that the potential for a bound-to-unbound transition (Fig. 8) needs to be considered. We

expect CAP-tests to be an important part of future tools that de-risk induced seismicity.

Stimulation optimisation. Distinguishing between bound fracture growth and unbound fault reactivation is also significant from the perspective of reservoir optimisation. Stage fluids diverted into faults hinder stimulation efficiency and thus limit reservoir productivity⁴³. The use of CAP-tests could be paired with complimentary diagnostic tests attempting to discern the presence of faults⁴⁴. Future operators could use these diagnostic tests to proactively discern stages that are linked to faults,

Fig. 8 | Interactions between stimulated fractures and fault reactivation during hydraulic fracturing. A series of HF stages of increasing complexity (text labels) are considered alongside a hypothetical well (black & grey rectangle). In the simple case, the host rock is initially split in tension via stimulated fractures (grey polygons) and may also subsequently reactivate in shear. In the complex case, small pre-existing faults (black polygons) can also be reactivated in shear slip. The clustered case hydraulically connects a series of stages/perforations into a single fracture/fault network. In the last case, stimulated fractures intersect a large fault system that is critically-stressed for slip. By watching the growth of earthquake magnitudes, CAP-tests can discern between bound cases (blue circles) and unbound cases (pink circles).

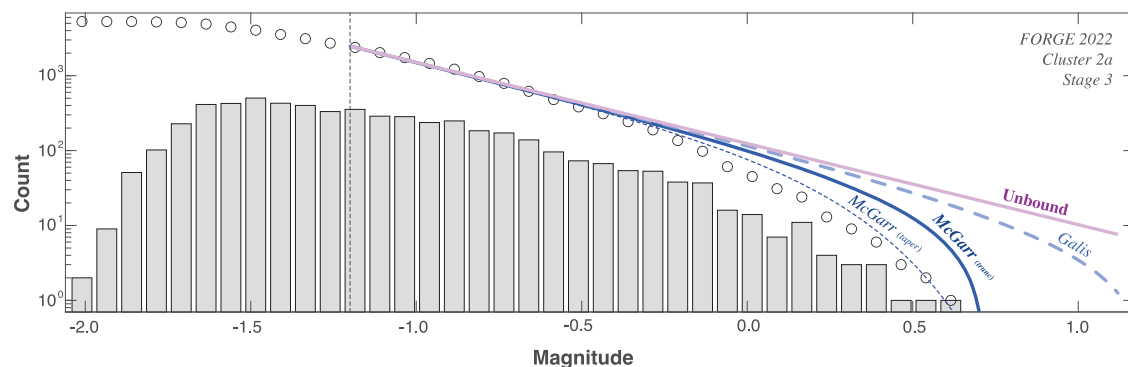
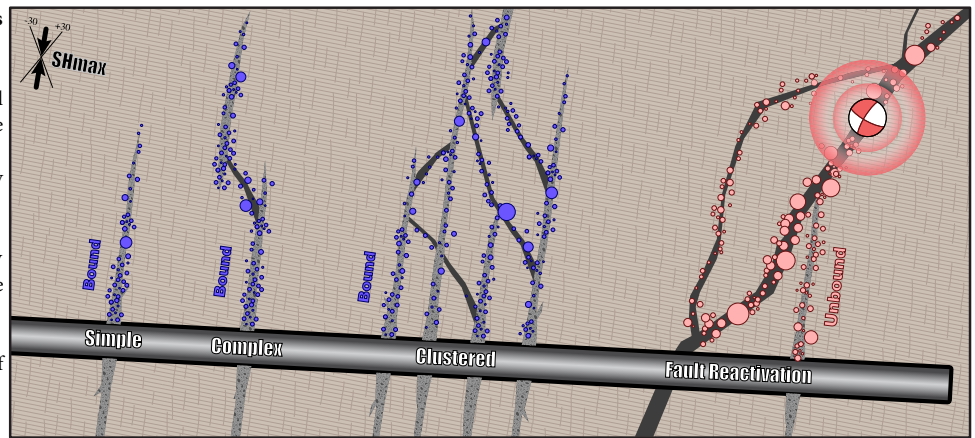


Fig. 9 | Comparison of expected fits to the GR-MFD. Data are plotted as both the cumulative (circles) and non-cumulative (bars) distribution, alongside the magnitude-of-completeness M_c (dashed lines). Expected fits to the data based on the unbound (pink line), Galis-like (pale blue dashed line), McGarr-like with a truncated

GR-MFD (dark blue solid line), and McGarr-like with a tapered GR-MFD (dark blue dashed line). Data is from stage 3 stimulation at FORGE in 2022 (cluster 2a). The magnitude-of-completeness M_c is also shown (dashed lines).

thus avoiding stimulation into unproductive and seismically risky stages. On the flip side, operators that recognise a bound case could potentially continue injecting closer to the red-light – allowing for greater fracture stimulation, and greater reservoir productivity, which enables further energy development. For example, the next largest event for ~50% of bound cases would trigger a red-light if M_{LRG} was 0.10 M away from the red-light (Fig. 7). Although, this would require further research and validation testing into the nuances and efficacy of CAP-tests (e.g., Text S6-S7). A hypothetical scenario could be a bound-to-unbound transition; for example, if bound fracture growth suddenly encounters a large fault (Fig. 8). Generally, having a strong understanding of the fault architecture and hydraulic connectivity appears to be important for discerning bound cases.

Physical mechanisms. The operation at PNR-1z, St1, and initial stages of FORGE/PNR-2 appear to have responded in similar manner to stage stimulation; all have encountered some physical process bounding magnitude growth. There is strong evidence for this from the GR-MFD and CAP-tests, that cannot be reasonably explained by statistical variance (Sections 2–3; Figs. 3–7). This means that there is a measurable impact of finite fault size restricting magnitudes in these cases, albeit with a time dependence from the injection.

Certainly, the spatial extent of stimulation earthquakes increases with time during HF operations, including FORGE^{45,46}. This finite extent of the stimulated fractures provides a potential candidate mechanism for restricted magnitude growth. As well, there is growing evidence suggesting that

injection-based processes have restricted at least some stage stimulations at HF and EGS operations^{16,47–49} and during laboratory experiments⁵⁰. Correspondingly, our results indicate that either Galis-like or McGarr-like models^{26–28} can explain the data, depending on the case. This result also corresponds with the scaling of total seismic moment and total injected volume observed during HF operations in Canada⁴⁷. We note that while Galis-like and McGarr-like are the ‘winning’ models, there could still be other unknown/untested models that explain the data even better. Future high-resolution studies (aided by CAP-tests) will be able to further elucidate the physics of the bounded fracturing process.

On the other hand, cluster 3 at FORGE and the eastern-most cluster at PNR-2 have responded in a fundamentally different way; there appears to be no (currently discernible) bounds restricting the growth of those earthquake magnitudes. Correspondingly, other studies of PNR-2 have also recognised sequential changes in stage stimulations towards the eastern-most cluster^{21,51}. We argue that this likely represents fault reactivation: the fluid/pressure diffusion process triggers slip initiation on a large fault, releasing the tectonic stresses stored over geological timeframes (Fig. 8). The earthquake magnitudes were able to grow unrestricted by any physical limitations. This interpretation is consistent with numerous other large magnitude cases that have also exhibited unbound growth, which appears to be the norm³³. Further work would be needed to better understand the processes responsible (e.g., fracture complexity, hydraulic connectivity, stress orientations) for creating this bound-to-unbound transition⁵².

In this sense, we argue that CAP-tests represent an approach able to empirically discern between triggered or driven induced earthquakes^{22,53,54}.

Earthquakes induced during bound fracture growth are measurably discernible from those triggered during fault reactivation. This approach will be useful in testing against other cases that appear to be deficient in high magnitude events^{55,56}.

Future directions for CAP-tests. We note that our CAP-tests are limited, in the sense that they can only infer the presence of M_{MAX} and then select the best model among the ensemble of M_{MAX} proposals. If the true model is missing from this ensemble, then it is impossible to select. Here, we discuss the potential for this oversight bias in our study, as well as other errors.

To examine this, we focus on the stage 3 stimulation at FORGE (Fig. 4a; cluster 2a). We use the previously fitted McGarr-like and Galis-like M_{MAX} models to build the expected GR-MFD at the end of catalogue recording. To do so, we use a mixture model that performs a summation of truncated GR-MFDs, with the M_{MAX} temporally varying over the course of injection. This allows us to visualise the expected roll-off that the McGarr-like and Galis-like M_{MAX} ought to have imposed on the data (Fig. 9). The gap between data-and-expectation at high magnitudes provides a qualitative sense of model misfit – further indicating as to why the EW-tests preferred McGarr-like to Galis-like to unbound, respectively.

Despite these efforts, there still remains a misfit gap. On one hand, this gap could be the result of an oversight-bias, which would suggest that a more conservative M_{MAX} -volume relationship is still needed. On the other hand, additional physical complexities like limited stimulated rock volume, a distribution of fault/fracture areas, or seismic moment conservation could have contributed to this gap. Our examination here considered truncated GR-MFDs, if we instead consider a mixture of tapered distributions⁵⁷, then this gap appreciably closes. We note that tapered GR-MFDs have been introduced as a potential explanation for the aforementioned physical considerations, and are generally viewed as advantageous over truncated distributions⁵⁷. Regardless, CAP-test can accurately infer the presence of this roll-off. Further interpretation between these physical mechanisms will require additional studies and information to make more sophisticated M_{MAX} proposals.

Last, there is the potential for errors lurking within the catalogue. These can be related to differing measurement types, differing catalogue building, or magnitude uncertainties, all of which can appreciably impact interpretation⁵⁸. In light of these errors, we have tried to be careful in our analysis, to consider magnitude dithering, bootstrapping and parameter perturbations (Text S6–S7). We note that the relative importance of ΔM_{LRG} for discerning M_{MAX} provides future directions for a quality-control prioritisation scheme. In cases where constraining M_{MAX} via CAP-tests are important, quality-control efforts should primarily focus on accurately capturing the M_{LRG} envelope (Fig. 1 & S1).

Conclusions

We find that our CAP tests can confidently discern cases of bound earthquake magnitude growth during HF stage stimulations at PNR-2 and FORGE (including PNR-1z, ST1, and Soultz-sous-Forêts). This includes an identification of a transition into (problematic) unbound growth, which prematurely terminated operations at PNR-2. We discuss the potential physical mechanisms controlling this binding process. We also discuss the implications for operational risks regarding reservoir optimisation and induced seismicity, including our thoughts on possible mitigation measures guided by CAP-tests. These tests could be used to de-risk future operations, aiding energy development and the transition to a green energy economy.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data (and codes) used to derive our results are available online at GitHub (<https://github.com/RyanJamesSchultz/CAPfr>). The catalogue and hydraulic

source datasets for FORGE are also available online (<https://gdr.openei.org/>). As well, the source injection data for both PNR-2 and PNR1-z are available (<https://www.nstauthority.co.uk/regulatory-information/exploration-and-production/onshore/onshore-reports-and-data/preston-new-road-well-pnr2-data-studies/>; <https://www.nstauthority.co.uk/regulatory-information/exploration-and-production/onshore/onshore-reports-and-data/preston-new-road-well-pnr-1z-data-studies/>) PNR source catalogue information is also available online (<https://webapps.bgs.ac.uk/services/ngdc/accessions/index.html#item173104>). Injection and catalogue source data for Soultz-sous-Forêts (<https://episodesplatform.eu/>). Catalogue data for Helsinki St1 (<https://doi.org/10.5880/GFZ.4.2.2019.001>).

Code availability

The codes (and data) used to derive our results are available online at GitHub (<https://github.com/RyanJamesSchultz/CAPfr>).

Received: 13 March 2025; Accepted: 6 October 2025;

Published online: 22 November 2025

References

- King, G. E. Thirty years of gas shale fracturing: what have we learned? In *SPE Annual Technical Conference and Exhibition?* (pp. SPE-133456), <https://doi.org/10.2118/133456-MS> (2010).
- Olasolo, P., Juárez, M. C., Morales, M. P. & Liarte, I. A. Enhanced geothermal systems (EGS): a review. *Renew. Sustain. Energy Rev.* **56**, 133–144 (2016).
- Ricks, W., Voller, K., Galban, G., Norbeck, J. H., & Jenkins, J. D. The role of flexible geothermal power in decarbonized electricity systems. *Nat. Energy*, 1–13, <https://doi.org/10.1038/s41560-023-01437-y> (2024).
- Li, S., Wang, S. & Tang, H. Stimulation mechanism and design of enhanced geothermal systems: a comprehensive review. *Renew. Sustain. Energy Rev.* **155**, 111914 (2022).
- Horne, R. et al. Enhanced geothermal systems for clean firm energy generation. *Nat. Rev. Clean. Technol.* **1**, 13 (2025).
- Majer, E. L. et al. Induced seismicity associated with enhanced geothermal systems. *Geothermics* **36**, 185–222 (2007).
- Schultz, R. et al. Hydraulic fracturing-induced seismicity. *Rev. Geophys.* **58**, e2019RG000695 (2020).
- Moein, M. J. et al. The physical mechanisms of induced earthquakes. *Nat. Rev. Earth Environ.* **4**, 847–863 (2023).
- Häring, M. O., Schanz, U., Ladner, F. & Dyer, B. C. Characterisation of the Basel 1 enhanced geothermal system. *Geothermics* **37**, 469–495 (2008).
- Grigoli, F. et al. The November 2017 Mw 5.5 Pohang earthquake: a possible case of induced seismicity in South Korea. *Science* **360**, 1003–1006 (2018).
- Lei, X., Wang, Z. & Su, J. The December 2018 M_L 5.7 and January 2019 M_L 5.3 earthquakes in South Sichuan basin induced by shale gas hydraulic fracturing. *Seismol. Res. Lett.* **90**, 1099–1110 (2019).
- Evans, K. F., Zappone, A., Kraft, T., Deichmann, N. & Moia, F. A survey of the induced seismic responses to fluid injection in geothermal and CO₂ reservoirs in Europe. *Geothermics* **41**, 30–54 (2012).
- Atkinson, G. M. et al. Hydraulic fracturing and seismicity in the Western Canada Sedimentary Basin. *Seismological Res. Lett.* **87**, 631–647 (2016).
- 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**, 304–308 (2018).
- Kwiatek, G. et al. Controlling fluid-induced seismicity during a 6.1-km-deep geothermal stimulation in Finland. *Sci. Adv.* **5**, eaav7224 (2019).
- Majer, E., Nelson, J., Robertson-Tait, A., Savy, J. & Wong, I. Protocol for addressing induced seismicity associated with enhanced geothermal systems. *US Department of Energy*. <http://wellbore.lbl.gov/downloads/EGS-IS-Protocol-Final-Draft-20120124.PDF> (2012).

17. Zhou, W. et al. Managing induced seismicity risks from enhanced geothermal systems: A good practice guideline. *Rev. Geophys.* <https://doi.org/10.1029/2024RG000849> (2024).
18. Bommer, J. J. et al. Control of hazard due to seismicity induced by a hot fractured rock geothermal project. *Eng. Geol.* **83**, 287–306 (2006).
19. Schultz, R., Beroza, G. C. & Ellsworth, W. L. A risk-based approach for managing hydraulic fracturing-induced seismicity. *Science* **372**, 504–507 (2021a).
20. Schultz, R., Beroza, G. C. & Ellsworth, W. L. A strategy for choosing red-light thresholds to manage hydraulic fracturing induced seismicity in North America. *J. Geophys. Res.: Solid Earth* **126**, e2021JB022340 (2021b).
21. Schultz, R., Baptie, B., Edwards, B. & Wiemer, S. Red-light thresholds for induced seismicity in the UK. *Seismica*, 2, <https://doi.org/10.26443/seismica.v2i2.1086> (2023b).
22. Bommer, J. & Verdon, J. P. The maximum magnitude of natural and induced earthquakes. *Geomech. Geophys. Geo-Energy Geo-Resour.* **10**, 172 (2024).
23. Gutenberg, B. & Richter, C. F. Frequency of earthquakes in California. *Bull. Seismol. Soc. Am.* **34**, 185–188 (1944).
24. 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**, 4575–4590 (2016).
25. Kanamori, H. & Anderson, D. L. Theoretical basis of some empirical relations in seismology. *Bull. Seismol. Soc. Am.* **65**, 1073–1095 (1975).
26. McGarr, A. Maximum magnitude earthquakes induced by fluid injection. *J. Geophys. Res.: Solid Earth* **119**, 1008–1019 (2014).
27. Hallo, M., Oprsal, I., Eisner, L. & Ali, M. Y. Prediction of magnitude of the largest potentially induced seismic event. *J. Seismol.* **18**, 421–431 (2014).
28. Galis, M., Ampuero, J. P., Mai, P. M. & Cappa, F. Induced seismicity provides insight into why earthquake ruptures stop. *Sci. Adv.* **3**, eaap7528 (2017).
29. Holschneider, M., Zöller, G. & Hainzl, S. Estimation of the maximum possible magnitude in the framework of a doubly truncated Gutenberg–Richter model. *Bull. Seismol. Soc. Am.* **101**, 1649–1659 (2011).
30. Zöller, G. & Holschneider, M. The earthquake history in a fault zone tells us almost nothing about Mmax. *Seismol. Res. Lett.* **87**, 132–137 (2016).
31. DeDontney, N. et al. Maximum magnitude of induced earthquakes in the Groningen gas field. *ExxonMobil URC External Report*, <https://nam-onderzoekrapporten.data-app.nl/reports/download/groningen/en/36dea690-c7eb-4164-9d94-d9725381c426> (2016).
32. Schultz, R. Inferring maximum magnitudes from the ordered sequence of large earthquakes. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **38**, 20230185 (2024).
33. Clarke, H., Verdon, J. P., Kettlety, T., Baird, A. F. & Kendall, J. M. Real-time imaging, forecasting, and management of human-induced seismicity at Preston New Road, Lancashire, England. *Seismol. Res. Lett.* **90**, 1902–1915 (2019).
34. Moore, J. et al. The Utah Frontier Observatory for Research in Geothermal Energy (FORGE): an international laboratory for enhanced geothermal system technology development. In *Proc. 44th Workshop on Geothermal Reservoir Engineering* (pp. 11–13). (Stanford University, 2019).
35. Niemz, P. et al. Mapping fracture zones with nodal geophone patches: insights from induced microseismicity during the 2024 stimulations at Utah FORGE. *Seismol. Res. Lett.* <https://doi.org/10.1785/0220240300> (2025).
36. Dorbath, L., Cuenot, N., Genter, A. & Frogneux, M. Seismic response of the fractured and faulted granite of Soultz-sous-Forêts (France) to 5km deep massive water injections. *Geophys. J. Int.* **177**, 653–675 (2009).
37. Kettlety, T., Verdon, J. P., Butcher, A., Hampson, M. & Craddock, L. High-resolution imaging of the M_L 2.9 August 2019 earthquake in Lancashire, United Kingdom, induced by hydraulic fracturing during Preston New Road PNR-2 operations. *Seismol. Res. Lett.* **92**, 151–169 (2021).
38. BEIS, UK Department of Business, Energy, and Industrial Strategy. Government ends support for fracking: *Press Release*. URL: <https://www.gov.uk/government/news/government-ends-support-for-fracking> (2019).
39. 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.* **25**, 301–326 (2021).
40. Schultz, R., Ellsworth, W. L. & Beroza, G. C. Statistical bounds on how induced seismicity stops. *Sci. Rep.* **12**, 1184 (2022).
41. Maxwell, S. C. et al. Fault activation during hydraulic fracturing. In *SEG International Exposition and Annual Meeting*, (pp. SEG-2009). SEG. (2009).
42. Langenbruch, C., Ellsworth, W. L., Woo, J. U. & Wald, D. J. Value at induced risk: injection-induced seismic risk from low-probability, high-impact events. *Geophys. Res. Lett.* **47**, e2019GL085878 (2020).
43. Kurison, C. & Woldeamanuel, M. B. Natural fracture density controls productivity in shale reservoirs. *Geological Society, London, Special Publications*, 546, SP546-2023, (2024).
44. Zeinabady, D. et al. Fault identification for the purposes of evaluating the risk of induced seismicity: a novel application of the Flowback DFIT. *SPE Reserv. Eval. Eng.* **26**, 1059–1072 (2023).
45. Niemz, P., McLennan, J., Pankow, K. L., Rutledge, J. & England, K. Circulation experiments at Utah FORGE: near-surface seismic monitoring reveals fracture growth after shut-in. *Geothermics* **119**, 102947 (2024).
46. Lanza, F. et al. The 2022 hydraulic stimulation at Utah FORGE: investigating fracturing mechanisms and testing forecasting approaches, submitted to *Seismica*. (2026).
47. 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**, 492–505 (2017).
48. Bentz, S., Kwiitek, G., Martínez-Garzón, P., Bohnhoff, M. & Dresen, G. Seismic moment evolution during hydraulic stimulations. *Geophys. Res. Lett.* **47**, e2019GL086185 (2020).
49. Drif, K., Lengliné, O., Kinscher, J. & Schmitbuhl, J. Induced seismicity controlled by injected hydraulic energy: The case study of the EGS Soultz-sous-Forêts site. *J. Geophys. Res.: Solid Earth* **129**, e2023JB028190 (2024).
50. Yu, J. et al. Role of critical stress in quantifying the magnitude of fluid-injection triggered earthquakes. *Nat. Commun.* **15**, 7893 (2024).
51. Minetto, R., Helmstetter, A., Edwards, B. & Guéguen, P. How injection history can affect hydraulic fracturing-induced seismicity: insights from downhole monitoring at Preston New Road, United Kingdom. *Bull. Seismol. Soc. Am.* **114**, 924–941 (2024).
52. Gischig, V. S. Rupture propagation behavior and the largest possible earthquake induced by fluid injection into deep reservoirs. *Geophys. Res. Lett.* **42**, 7420–7428 (2015).
53. McGarr, A., Simpson, D., Seeber, L. & Lee, W. Case histories of induced and triggered seismicity. *Int. Geophys. Ser.* **81**, 647–664 (2002).
54. Dahm, T., Cesca, S., Hainzl, S., Braun, T. & Krüger, F. Discrimination between induced, triggered, and natural earthquakes close to hydrocarbon reservoirs: a probabilistic approach based on the modeling of depletion-induced stress changes and seismological source parameters. *J. Geophys. Res. Solid Earth* **120**, 2491–2509 (2015).

55. Huang, Y. & Beroza, G. C. Temporal variation in the magnitude–frequency distribution during the Guy–Greenbrier earthquake sequence. *Geophys. Res. Lett.* **42**, 6639–6646 (2015).
56. Chen, X. et al. Temporal correlation between seismic moment and injection volume for an induced earthquake sequence in central Oklahoma. *J. Geophys. Res. Solid Earth* **123**, 3047–3064 (2018).
57. Kagan, Y. Y. Seismic moment distribution revisited: I. Statistical results. *Geophys. J. Int.* **148**, 520–541 (2002).
58. Herrmann, M. & Marzocchi, W. Inconsistencies and lurking pitfalls in the magnitude–frequency distribution of high-resolution earthquake catalogs. *Seismol. Soc. Am.* **92**, 909–922 (2021).

Acknowledgements

We would like to thank Kevin England and Jim Rutledge for engaging conversations about the FORGE stimulation programme that directed the insights and interpretations of this paper. We would also like to thank James Verdon & Tom Kettlety for discussions around the interpretation of PNR-1z and PNR-2 operations. We thank Max Werner for discussions on statistical aspects of the paper. Last, we thank Barnaby Fryer and the anonymous reviewers, whose comments helped improve this manuscript. This work was supported by the Seismogenic Fault Injection Test (SFIT) and De-Risking Enhanced Geothermal Energy Projects (DEEP) grants. SFIT is funded by the Swiss National Science Foundation, under project number TMPFP2_224393. DEEP is subsidised through the Cofund GEOTHERMICA by the Swiss Federal Office of Energy (SFOE), which is supported by the European Union's HORIZON 2020 programme for research, technological development, and demonstration under Grant Agreement #731117.

Author contributions

R.S. created the CAP-tests, gathered the data, analysed the cases, and wrote the manuscript. B.D., D.K., R.F. and P.M. were involved in the processing and curation of the FORGE catalogue. F.L., P.S., V.R. and L.V. were involved in the collection of FORGE data. F.L. and S.W. provided project management and administration. All authors were involved in manuscript editing and review.

Funding

Open access funding provided by Swiss Federal Institute of Technology Zurich.

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/s43247-025-02881-2>.

Correspondence and requests for materials should be addressed to Ryan Schultz.

Peer review information *Communications Earth and Communications* thanks Barnaby Fryer, Dominique Bruel and Matteo Taroni for their contribution to the peer review of this work. Primary Handling Editors: Gareth Roberts, Joe Aslin and Alireza Bahadori. A peer review file is available.

Reprints and permissions information is available at <http://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) 2025