

1 **Title**

2 Interpreting the cause of bound earthquakes at underground injection experiments

3

4 **Authors**

5 Ryan Schultz^{1*}, Linus Villiger¹, Valentin Gischig¹, Stefan Wiemer¹

6

7 ¹ Swiss Seismological Service, ETH Zürich, Zürich, Switzerland

8 * corresponding author: Ryan.Schultz@sed.ethz.ch

9

10 **Abstract**

11 Constraining the maximum possible magnitude (M_{MAX}) of an induced earthquake
12 sequence is a challenging process with important implications for managing risks. CAP-tests
13 are a suite of statistical tests that can infer, quantify, and select best-fitting M_{MAX} models via
14 an earthquake catalogue's magnitudes. We use CAP-tests to discern between bound/unbound
15 earthquake sequences at underground laboratories, where high-resolution and near-field
16 geophysical observations are abundant. There, we find clear evidence for bound sequences,
17 where magnitude growth was restricted during stimulation. Furthermore, bound sequences
18 tend to be associated with stimulations that occurred within intact rock. On the other hand,
19 unbound sequences tended to be associated with stimulations where hydraulic fractures
20 interacted with relatively large pre-existing faults/fractures. We further examine bound
21 sequences by fitting magnitude growth to a generalized family of M_{MAX} functions. This
22 process appears to be able to aggregate bound sequences into categories consistent with
23 theoretical considerations (*e.g.*, tectonic, tensile-crack, or shear-crack). These results provide
24 a basis for validating and interpreting bound sequences in controlled experiments, which is
25 important for extrapolating to larger-scale observations. Overall, CAP-tests appear to be a
26 promising avenue for constraining M_{MAX} from earthquake catalogue data.

27

28 **Short Summary**

29 We use statistical tests to infer M_{MAX} from an earthquake catalogue and focus on data
30 from three underground laboratories with controlled injection experiments. There, we find
31 clear evidence for M_{MAX} bounds and corroborate interpretations of fracture growth against
32 other geophysical studies. Unbound sequences occur when stimulation is directed towards pre-
33 existing faults. The validation of our methods against well-studied cases is encouraging and
34 will help validate future interpretations.

35

36 **Key Points**

- 37
- Bound earthquake sequences (*i.e.*, physically restricted maximum magnitudes) were
- 38 unambiguously identified in underground laboratories.
- Bound sequences tend to be associated with stimulation stages, while unbound
- 39 sequences tend to inject into pre-existing faults/fractures.
- Maximum magnitudes aggregate into four categories that are consistent with
- 40 theoretical expectations.
- 41
- 42
- 43
- 44

45 Main text

46 1. Introduction

47 Any process that has the potential to alter stresses in the subsurface also has the
48 potential to trigger earthquakes (Moein et al., 2023). To date, many types of anthropogenic
49 operations have induced earthquakes including wastewater disposal, hydraulic fracturing,
50 enhanced geothermal systems, geological carbon sequestration, reservoir impoundment, and
51 mining (Majer et al., 2007; Foulger et al., 2018; Schultz et al., 2020). In some cases, these
52 events have grown large enough to be felt, damaging, or even harmful (Atkinson et al., 2016;
53 Grigoli et al., 2018). In reaction to some of these cases, social concerns have resulted in
54 subsurface development moratoriums (Kettlety et al., 2021; Muntendam-Bos et al., 2022).

55 The need to manage the risks of induced earthquakes has been recognized (Bommer,
56 2022; Zhou et al., 2024). The *de facto* approach to risk management centralizes around the
57 traffic light protocol (Bommer et al., 2006), which delineates when an operation must stop (*i.e.*,
58 the red-light). Typically, red-lights are designed as magnitude thresholds (Schultz et al., 2021a;
59 2021b; 2023a). Despite the widespread adoption of traffic light protocols, relatively little is
60 understood about the growth of induced earthquake magnitudes. Some recent papers have
61 begun examining this topic; for example, by attempting to forecast the next largest event
62 (Mendecki, 2016; Cao et al., 2020; 2024; Schultz et al., 2023b; Verdon et al., 2023; 2024; Yin
63 et al., 2024). However, these approaches are unable to discern the adequacy of input models.
64 These topics are important for understanding trailing earthquakes (Schultz et al., 2022) and
65 magnitude jumps (Verdon & Bommer, 2021), which are among the most important factors for
66 designing risk-informed red-lights (Schultz et al., 2021a).

67 A common theme underlying these efforts is constraining the maximum possible
68 magnitude (M_{MAX}), given its prominence for induced seismicity hazard quantification
69 (Bommer & Verdon, 2024). While there are numerous theoretical models that could possibly
70 limit M_{MAX} (McGarr, 2014; Hallo et al., 2014; Galis et al., 2017; Elsworth, et al., 2025; Im &
71 Avouac, 2025; Sáez et al., 2025), there are relatively few methods available to reliably validate
72 them with empirical data (Holschneider et al., 2014; Pisarenko & Rodkin, 2022; Kijko, 2025).
73 In fact, empirically constraining M_{MAX} from a catalogue is known to be a challenging task that
74 is usually only possible in special circumstances (Holschneider et al., 2011; Zöller et al., 2016).
75 Given the conflict between needing M_{MAX} constraints against the lack of robust approaches,
76 expert solicitation is often used as a substitute (DeDontney et al., 2016). Thus, there is a need
77 to replace subjective/opinion-based methods with more quantitative measures for M_{MAX} .

78 Based on this need, CAP-tests were designed as a suite of statistical tests to infer and
79 quantify the presence of M_{MAX} within a catalogue of data. Specifically, this approach compares
80 the distribution of earthquake magnitudes (M) against the distribution of jumps in the largest
81 magnitude events (ΔM_{LRG}): when they are the same distribution, the catalogue is unbound and
82 there is no M_{MAX} (and *vice versa*) (Schultz, 2024). Synthetic testing suggests that this problem
83 reformulation can be more sensitive to M_{MAX} than traditional approaches (Schultz, 2024;
84 2026). CAP-tests are broken into three component parts: the KS-test (Kolmogorov-Smirnov
85 test) as a hypothesis test to infer the presence of M_{MAX} , the MLE-test (Maximum Likelihood
86 Estimation test) to quantify the value of M_{MAX} , and the EW-test (Ensemble Weighting test) to
87 select the theoretical M_{MAX} model that explains the catalogue data best. We provide the full
88 methodological details of CAP-tests in Section 2.

89 The application of CAP-tests to relatively larger magnitude induced seismicity cases
90 found that these cases behaved in an unbound manner (Schultz, 2024) – suggesting that the use
91 of theoretical M_{MAX} models (McGarr, 2014; Hallo et al., 2014; Galis et al., 2017; Elsworth, et
92 al., 2025; Im & Avouac, 2025; Sáez et al., 2025) are not appropriate during these fault
93 reactivations. On the other hand, smaller-scale cases of induced seismicity (e.g., Utah FORGE,
94 Preston New Road, Helsinki St1) from hydraulic stimulation indicated strong evidence for
95 bound magnitude growth via an M_{MAX} (Schultz et al., 2025). Furthermore, some stimulation
96 stages transitioned from a bound process into an unbound one; this was interpreted as these
97 particular stages reactivating pre-existing faults instead of stimulating new fractures. That said,
98 subsurface unknowns/uncertainties hamper an unambiguous interpretation for these field-scale
99 cases.

100 To better address subsurface unknowns/uncertainties, several underground laboratories
101 (UGLs) have performed hydraulic fracturing experiments with varying scales/scopes. These
102 intermediate-scale experiments aim to fill a gap between laboratory studies and field-scale
103 observations: accessing near-field geophysical observations (that are infeasible at the field-
104 scale) in a semi-realistic *in situ* setting. Said another way, UGL data was collected with the
105 intention of delineating and interpreting fracture stimulation. Thus, the complementary data
106 from UGLs is ideal for better interpreting bound sequences (found via CAP-tests). This
107 comparison has the potential to elucidate an underlying mechanism for bound induced
108 seismicity sequences. This is significant, since it starts to link statistical inferences of M_{MAX}
109 to physical processes.

110 In this study, we collect data from three UGLs: the Äspö Hard Rock Laboratory (Zang
111 et al., 2019; 2021), the Sanford Underground Research Facility (Dobson et a., 2018; Kneafsey

112 et al., 2018; Morris et al., 2018), and the Grimsel Test Site (Gischig et al., 2016; Amann et al.,
113 2018). We perform CAP-tests on these UGL datasets and observe both bound and unbound
114 stimulation stages. We cross-reference these observations against the wealth of geophysical
115 observations and interpretations from prior multi-disciplinary studies. Bound cases tend to
116 occur when stimulation is restricted to the growth of new fractures, while unbound cases tend
117 to reactivate (relatively large) pre-existing faults/fractures. Furthermore, we empirically search
118 for the functional form of M_{MAX} models; we delineate four categories that are predominantly
119 consistent with theoretical expectations. Finally, we discuss how these results create a
120 consistent picture with field-scale results.

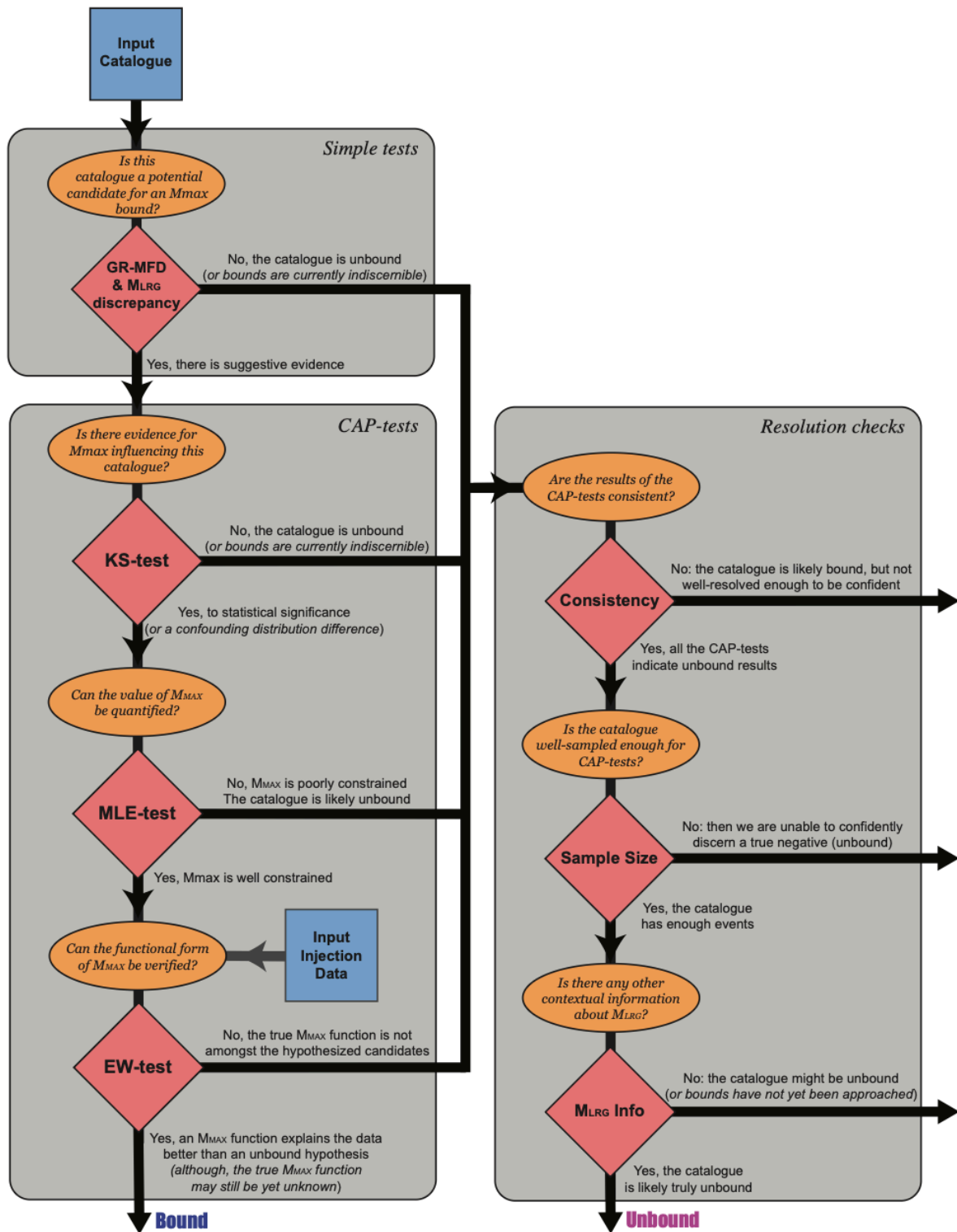
121

122 **2. Methods**

123 ***2.1 Methodological Overview***

124 The methods of this paper concern the robust inference and estimation of M_{MAX} from a
125 catalogue of earthquakes. To do so, we have composed a suite of statistical tests. First, simple-
126 tests can be used to pre-screen cases and provide qualitative or semi-quantitative inferences of
127 M_{MAX} bounds. Candidate catalogues can then be examined more deeply using CAP-tests,
128 which use a series of statistical tests to robustly discern the presence of M_{MAX} and quantify the
129 best model hypothesis. Last, we outline criteria for assessing the absence of M_{MAX} . These
130 tests are used to logically guide interpretations amongst the three possibilities: that M_{MAX}
131 bounds a catalogue, that M_{MAX} is not bounding a catalogue, or that the catalogue is not well-
132 resolved enough to make an interpretation either way. An overview of this workflow's logical
133 sequencing has been provided (Figure 1).

134



135
136
137
138

Figure 1. Workflow to discern M_{MAX} bounds. Simple-tests are used to pre-screen for potentially bound catalogues. Afterwards, CAP-tests can more rigorously identify bound cases.

139 2.2 Simple-tests

140 Here, we briefly define some simple-tests to provide an initial assessment of M_{MAX} .
141 This is accomplished through an examination of the Gutenberg-Richter magnitude frequency

142 distribution (GR-MFD) (Ishimoto & Iida, 1939; Gutenberg & Richter, 1944) and discrepancies
143 between observed largest event minus the expected largest events. This sub-section constitutes
144 the first step in our overall workflow (Figure 1).

145

146 2.2.1 Fitting the Gutenberg-Richter Magnitude Frequency Distribution

147 The GR-MFD is a relationship that describes the amount and frequency of earthquake
148 magnitudes: $N = 10^a 10^{-bM}$. The a -value describes the scaling to the total amount of
149 earthquakes N , while the b -value is the proportionality of big-to-small events. Given a
150 catalogue of earthquake magnitudes M , this relationship can be fit through Maximum
151 Likelihood Estimators (Marzocchi & Sandri, 2003). For real datasets, a lower bound truncation
152 called the magnitude-of-completeness (M_c) is introduced to account for detection
153 incompleteness. Many methods exist to evaluate M_c robustly, to account for the
154 incompleteness in the detection of low magnitude events, thusly to avoid biases during the GR-
155 MFD fitting process. In this study, M_c is selected by searching for the value that maximizes
156 the goodness-of-fit metrics like R^2 (Schultz et al., 2018) or minimizes the negative log-
157 likelihood of the GR-MFD. In cases with a goodness-of-fit plateaus/valleys, we select the
158 value of M_c closest to plateau start. Sometimes we are conservative in our M_c choice by
159 selecting a value slightly larger than optimal. Note that this conservatism will have a
160 detrimental effect on finding bound cases. The M_c value selected by this process is typically
161 near (but skewed right-ward of) the peak bin of the non-cumulative GR-MFD.

162 As part of the simple indicators of M_{MAX} , a visual comparison of the GR-MFD fit
163 against the observed data is used to qualitatively examine for the presence of M_{MAX} .
164 Catalogues that are relatively deficient in large magnitudes (compared to their fits) are possible
165 candidates for an M_{MAX} . We note that this deficiency in large magnitude events is a key metric
166 for identifying if/when M_{MAX} can be constrained (Holschneider et al., 2011; Schultz, 2024;
167 2026).

168

169 2.2.2 Appraising the deficiency of large events

170 To provide additional semi-quantitative assessments of an M_{MAX} , we examine the
171 empirical degree-of-truncation (δM_{LRG}), which is the discrepancy between the largest observed
172 event minus the largest expected event. The expected M_{LRG} can be estimated: if a GR-MFD is
173 assumed, then order statistics suggests a modal value of $M_{LRG} = M_c + \log_{10}(N)/b$ (van der
174 Elst et al., 2016). Where N is the total number of events larger than the magnitude-of-

175 completeness M_c and b is the previously described b -value. We can use the prior GR-MFD
176 fits to determine the expected value of M_{LRG} and then compare that against the observed value
177 (*i.e.*, δM_{LRG}). We also use the inverted cumulative distribution function (van der Elst et al.,
178 2016) to determine the percentile of the δM_{LRG} discrepancy – or the likelihood of this degree-
179 of-truncation occurring, assuming an unbound catalogue.

180 As part of the simple indicators of M_{MAX} , catalogues that exhibit negative δM_{LRG}
181 suggest the presence of an M_{MAX} . We note that the δM_{LRG} discrepancy is a proxy metric for
182 the resolvability of CAP-tests. Usually values of $-0.5 M$ (or less) for $M_{\text{LRG}}-M_{\text{MAX}}$ are required
183 to confidently assert the presence of an M_{MAX} (Schultz, 2024; 2026). Note that this $M_{\text{LRG}}-$
184 M_{MAX} difference is the expected value where quantitative inferences of M_{MAX} can start being
185 made with 95% confidence, from theoretical considerations (Equation 15; Holschneider et al.,
186 2011).

187

188 **2.3 CAP-tests**

189 The CAP-tests are a suite of statistical methods aimed at discerning the influence of
190 M_{MAX} on a catalogue; each test is rooted in fundamentally different statistical frameworks, to
191 ensure the robust validation of results. In this sense, when all the CAP/simple-tests suggest a
192 similar bound/unbound result, we can be (more) confident that we have reached the right
193 interpretation – even if there might be data/method issues.

194 In this sub-section, we explicitly define each of the CAP-tests. We also refer readers
195 to a past study that defines the CAP-tests in detail, provides comprehensive sensitivity tests,
196 and highlight real-data applications (Schultz, 2024; Schultz et al., 2025; Verdon & Schultz,
197 2026). These tests build upon the simple pre-screening assessments (Section 2.2), logically
198 answering a sequence of questions to discern the influence of M_{MAX} more rigorously (Figure
199 1).

200

201 2.3.1 The Kolmogorov-Smirnov test (KS-test)

202 The first test is rooted in the statistical framework of hypothesis testing. Here, we take
203 advantage of the fact that the distribution of magnitudes (M) and the distribution of jumps in
204 the sequence of large events (ΔM_{LRG}) is the same if unbound, but differ when there is an M_{MAX}
205 upper bound (Schultz, 2024; 2026). This fact ideally lends itself to hypothesis testing via the
206 KS-test (Berger & Zhou, 2014). Given a catalogue, both M and ΔM_{LRG} can be observed. We
207 can then compare these two observations against each other to test if they are drawn from the
208 same distribution (or not), via the KS-test. This approach is advantageous in that it is non-

209 parametric – *i.e.*, it is completely data-driven and imposes no assumptions about the kind of
 210 distributions M or ΔM_{LRG} were drawn from. Because of this, we do not need to fit the data to
 211 a GR-MFD or have any knowledge/estimates of the b -value to perform our KS-test.
 212 Confidence in the KS-test is reported as compliments of standard p -values, where 95% is a
 213 common threshold used to declare statistical significance.

214 Since this test is only interested in discerning the existence of an M_{MAX} , additional
 215 catalogue realizations can be drawn through reshuffling the order of events. In this sense,
 216 bootstrapping can be employed to generate numerous catalogue realizations in which the KS-
 217 test is repeated. This provides more robust p -value estimates. Testing on both synthetic and
 218 real datasets suggests that the KS-test is significantly more sensitive to discerning M_{MAX} than
 219 approaches that attempt to appraise GR-MFD fits (Schultz, 2024; 2026). As well, synthetic
 220 testing on unbound cases shows that this KS-test produces false-positives at the rate expected
 221 for a p -value distribution (Schultz, 2024). To be able to discern the influence of M_{MAX} , M_{LRG} -
 222 M_{MAX} discrepancies of $-0.5 M$ or better are usually required, consistent with theoretical
 223 expectation (Holschneider et al., 2011).

224 While this formulation of the KS-test is powerful, we also provide a word of caution
 225 towards a potential interpretation pitfall: this method is testing for differences between the
 226 distributions M and ΔM_{LRG} . The presence of an M_{MAX} is one possible reason for this
 227 difference, but others may also confound a clear interpretation (*e.g.*, temporal changes in b -
 228 value, kinked distributions, tapered distributions). Thus, the KS-test should be suitably pre-
 229 processed or complemented with other tests to increase the certainty of an M_{MAX} interpretation.

230
 231

232 2.3.2 The Maximum Likelihood Estimator (MLE-test)

233 The next test is rooted in the statistical framework of Maximum Likelihood Estimation.
 234 If there is some suggestive evidence for the existence of an M_{MAX} , the next natural step is to
 235 quantify this value. The log-likelihood function is defined as follows:

$$236 \quad \ln(\mathcal{L}(M; \boldsymbol{\theta})) = \sum_i \ln(f_M(M; b, M_C, M_{\text{MAX}})) - \sum_j \ln(f_M(\Delta M_{\text{LRG}}; b, 0, M_{\text{MAX}} - M_{\text{LRG}}))$$

237 Where the probability density function of the GR-MFD is given by $f_M(M)$, with a set of model
 238 parameters $\boldsymbol{\theta}$ (Schultz, 2024). This function essentially entails two parts: the log-likelihood
 239 for the catalogue magnitudes M (right-hand side of equation, first term) and the log-likelihood
 240 of the jumps in largest events ΔM_{LRG} (right-hand side of equation, second term). The optimal
 241 set of model parameters $\boldsymbol{\theta}$ are then solved for via numerical methods to maximize the log-

242 likelihood, given the observed catalogue data M and ΔM_{LRG} . We perform this optimization in
243 two-steps: the first using the standard log-likelihood constrain the b -value (Marzocchi &
244 Sandri, 2003) (with the optimal M_c estimates) and then using the composite log-likelihood to
245 constrain M_{MAX} . In this study, we consider the simplest M_{MAX} variant for $f_M(M)$, which is a
246 truncated GR-MFD (Schultz, 2024).

247 If this test is only interested in discerning a stationary value of M_{MAX} , additional
248 catalogue realizations can be drawn through reshuffling the order of events. Similar to the KS-
249 test, bootstrapping can be employed to generate numerous catalogue realizations in which the
250 MLE-test is repeated. This provides more robust M_{MAX} estimates. Testing on both synthetic
251 and real datasets suggests that the MLE-test is sensitive to quantifying M_{MAX} within a
252 hundredth of a magnitude unit when $M_{\text{LRG}}-M_{\text{MAX}}$ discrepancies are $-0.5 M$ or better. In cases
253 where the MLE-test is applied to unbound catalogues, bootstrapped estimates of M_{MAX} will be
254 much larger than M_{LRG} and standard deviations can be on the order of 1 magnitude unit or
255 greater.

256

257 2.3.3 The Ensemble Weighting test (EW-test)

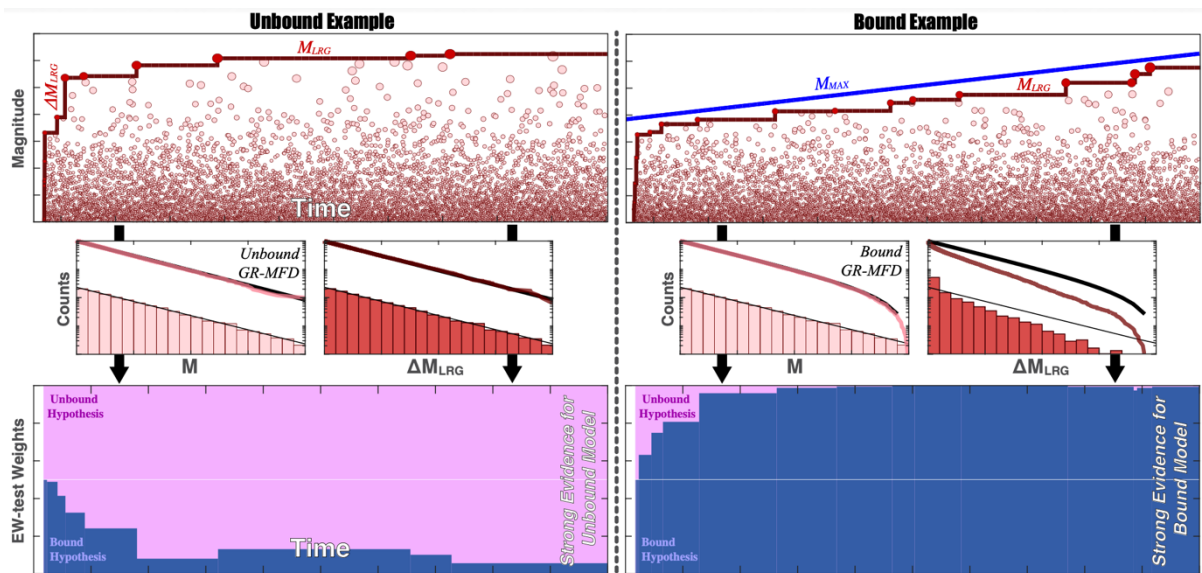
258 The third and final CAP-test is rooted in the statistical framework of likelihood
259 inference. Together, the two prior tests provide suitable evidence for the existence of M_{MAX} .
260 However, M_{MAX} may be a function of time or injected volume, becoming some non-stationary
261 value throughout the catalogue duration. Certainly, M_{MAX} processes relevant for induced
262 seismicity and hydraulic fracturing have been proposed in the past (McGarr, 2014; Hallo et al.,
263 2014; Galis et al., 2017; Elsworth, et al., 2025; Im & Avouac, 2025; Sáez et al., 2025). In this
264 sense, having an approach that can distinguish the best proposed M_{MAX} model (given the data)
265 would be insightful.

266 The EW-test starts by considering an ensemble of proposed M_{MAX} models to explain
267 the catalogue data (Schultz, 2024). Using the previously defined log-likelihood function, both
268 Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) can be defined
269 for each M_{MAX} model (Schwarz, 1978; Akaike, 1998); note that we have used the small sample
270 size corrections for AIC/BIC (Sugiura, 1978; McQuarrie, 1999). Next, we compute the
271 differences in AIC/BIC scores, by subtracting the score of the best M_{MAX} model. These score
272 differences can be translated into relative model weights by an exponential function
273 (Wagenmakers & Farrell, 2004). We combine AIC/BIC weights into a single weight by taking
274 an average between the two. In this study we will consider four standard M_{MAX} models:
275 McGarr-like (*i.e.*, log-proportional to injected volume; $M_{\text{MAX}} \propto \log_{10}(V^1)$) (McGarr, 2014;

276 Hallo et al., 2014; Elsworth, et al., 2025), Galis-like (*i.e.*, $M_{\text{MAX}} \propto \log_{10}(V^{3/2})$) (Galis et al.,
 277 2017), a constant tectonic upper bound (Kanamori & Anderson, 1975), and the unbound null
 278 hypothesis. The unbound null hypothesis consists of three model parameters (M_c , b -value,
 279 GR-MFD variance), while all bound models have one additional model parameter (*i.e.*, $K=4$)
 280 to account for the slope of the volume-based relationship.

281 The interpretation of the M_{MAX} model weights is straightforward: larger model weights
 282 indicate a better explanation of the data. The model with the largest weight is the best
 283 explanation of the data (within the ensemble). To quantify the statistical significance of weight
 284 differences between two models, the relative odds ratio can be computed as the ratio of the two
 285 model weights (larger/smaller). Ratios of 1+ are insignificant, 3+ are substantial/positive, 10+
 286 are strong, and 100+ is decisive (Kass & Raftery, 1995). We note that the best model in an
 287 ensemble does not necessarily imply the veracity of the model; there could be another
 288 (unknown) model that explains the data better than all of those yet considered.

289 In synthetic testing, the EW-test can accurately and confidently discern the true M_{MAX}
 290 model (Schultz, 2024). Usually, only a handful of ΔM_{LRG} observations are required to
 291 confidently identify the true model (*i.e.*, with odds ratios of 3-10 or better). However, like the
 292 prior CAP-tests, M_{MAX} must be influencing the catalogue for meaningful inferences to be made.
 293 Said another way, if the M_{MAX} is much larger than M_{LRG} , then the EW-test will not be able to
 294 distinguish between bound/unbound hypotheses. A schematic diagram explaining the EW-test
 295 has been provided (Figure 2).



296
 297 **Figure 2. Schematic demonstration of the EW-test.** The upper panels show hypothetical catalogues
 298 of earthquakes. Earthquake magnitudes (M , pink circles), the sequence of largest events (M_{LRG} , red
 299 line), and jumps in the largest event magnitudes (ΔM_{LRG}) can be directly observed. Potentially, some
 300 unobservable physical process could be bounding these catalogues (M_{MAX} , blue line). Left and right

301 panels detail contrasting concepts for a bound and unbound catalogue, respectively. If there is an
302 M_{MAX} , then it can be statistically inferred from the observables. Middle panels show the differences
303 between distributions of M and M_{LRG} , when bounded by M_{MAX} (or not). The distribution of M follows
304 a GR-MFD: analytical cumulative (thick black line) and non-cumulative (thin black line) agree with
305 numerical cumulative (pink line) and non-cumulative (pink bars) distributions. The distribution of
306 ΔM_{LRG} will differ from the GR-MFD when bounded by M_{MAX} . Bottom panels show the results of EW-
307 tests using these concepts. Weights of two hypothesized models, an unbound M_{MAX} (pink area) and a
308 bound M_{MAX} (blue area), change as new values of M_{LRG} are observed. EW-tests can quickly infer the
309 presence of the true M_{MAX} model, from the equivocal *a priori* assumption (white horizontal line).
310

311 **2.4 Truly unbound or just lacking data?**

312 In an ideal experiment, injection would continue indefinitely, providing excellent
313 resolving power for M_{MAX} by sampling infinitely many large events near M_{MAX} . Obviously,
314 real-data cases are sample/time-limited, however. CAP-tests have the potential to discern
315 influence from M_{MAX} in a (sample-limited but still) well-resolved catalogue. However, an
316 issue arises when attempting to perform these tests on poorly-resolved catalogues, since CAP-
317 tests cannot distinguish an M_{MAX} in this case. For example, if CAP-tests fail to indicate a
318 bound catalogue, then this could be because the catalogue is either truly unbound or simply is
319 not well-resolved enough (*i.e.*, it has a low degree-of-truncation). Distinguishing between
320 these two possibilities can be nebulous. We outline some guiding metrics to assist in making
321 this assessment, which is our last workflow step in discerning the (apparent) absence of M_{MAX}
322 more rigorously (Figure 1).

323 The first consideration is the consistency among simple/CAP-test results. Cases that
324 are truly bound tend to have all tests unanimously indicate a bound process, and *vice versa*.
325 For reference, sensitivity testing on bound real-data that was sequentially decimated indicated
326 a loss of resolving power, consistent with the degree of decimation (Schultz et al., 2025).
327 Specifically, the EW-test tends to lose resolving power first, followed by the KS-test, with the
328 MLE-test generally being the most sensitive. In fact, the use of multiple sub-tests rooted in
329 disparate statistical methods was an intentional design choice to cover strength/deficiencies of
330 each individual approach (Schultz, 2024). Note that this consistency indicator is only relevant
331 for semi-well-resolved cases that are truly bound; very-poorly-resolved cases would still appear
332 as unbound.

333 The second metric considered is the size of the catalogue N , within 1-2 magnitude units
334 of the current M_{LRG} . Synthetic tests indicate that catalogues between 10^1 - 10^2 (above the Mc)
335 are usually required to confidently assert M_{MAX} (Schultz, 2024). Similarly, real-data cases also
336 noted that EW-tests can confidently assert an M_{MAX} model after observing 30-200 events –

337 although most cases required ~50-100 events (Schultz et al., 2025). Thus, we can use catalogue
338 size as another rough indicator for how well-resolved M_{MAX} is.

339 The third metric we will consider is the currently observed M_{LRG} . Specifically, we will
340 consider M_{LRG} in a contextual comparison against relevant cases that are already known to be
341 bound. For example, if two independent neighbouring stages (with similarly sized catalogues)
342 give diverging bound/unbound responses, then the observed M_{LRG} can be used to make
343 inferences. Scenarios where the unbound stage has an observed M_{LRG} greater than the bound
344 stage is more likely to be truly unbound. On the other hand, scenarios where the unbound stage
345 has an observed M_{LRG} significantly less than the bound stage is potentially just a poorly-
346 resolved case.

347

348 **3. Data, Results, & Interpretations**

349 We apply this workflow to data collected at a few UGLs that were aimed at better
350 understanding the hydraulic fracturing process (Figure 3). UGLs include the Äspö Hard Rock
351 Laboratory (Äspö HRL), Sanford Underground Research Facility (SURF), and the Grimsel
352 Test Site (GTS). Each of these UGLs had differing scopes, scales, and aims. Here, we will
353 cover each UGL in a parallel style: we introduce each UGL, examine CAP-test results at the
354 UGL, and then briefly interprets those results. We will start with the simpler UGLs and then
355 transition into more complex cases, to pedagogically build upon the complexity of
356 interpretations.

357



358

359 **Figure 3. Global locations of datasets.** Locations of the test sites considered: the Äspö Hard Rock
 360 Laboratory (Sweden), Sanford Underground Research Facility (South Dakota, USA), and the Grimsel
 361 Test Site (Switzerland). Supporting locations from past studies (Schultz, 2024; Schultz et al., 2025),
 362 are also shown.
 363

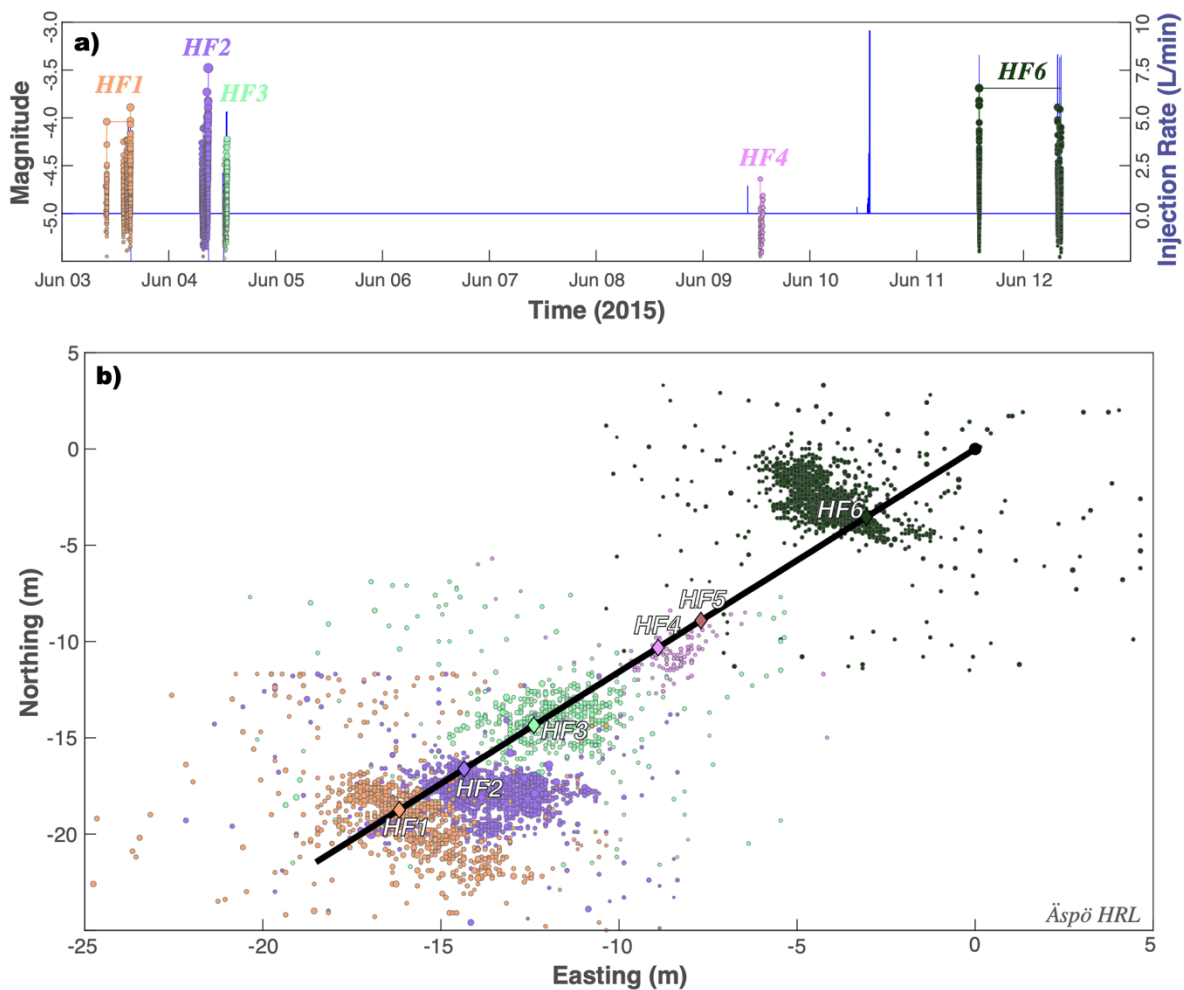
364 **3.1 The Äspö Hard Rock Laboratory**

365 3.1.1 Overview of the Äspö HRL

366 The Äspö HRL is located on the Baltic east coast of Sweden, near the Simpevarp
 367 Peninsula (Figure 3). This UGL was constructed as a testbed to study the potential for nuclear
 368 waste storage by the Swedish Nuclear Fuel and Waste Management Company starting in 1986.
 369 By 1995, construction efforts reached the 450 m target depth into the granitoids of the Trans-
 370 Scandinavian Igneous Belt (Stanfors et al., 1999).

371 The intention of recent (June 2015) injection experiments at the Äspö HRL was to test
 372 the feasibility of cyclic stimulation as an alternative means to enhance subsurface permeability,
 373 while also reducing the severity of induced seismicity (Zang et al., 2019; 2021). Six injection
 374 stages (HF1-HF6) were conducted in a single borehole (28 m long at 410 m depth), with an
 375 average stage spacing of ~3 m, and encountering either Ävrö granodiorite or fine-grained
 376 diorite-gabbro. Stage locations were chosen to avoid natural fractures (Zimmerman et al.,
 377 2019). Each stage used between 4.1-27.2 L of injected fluid, spread between 4-6 injection sub-
 378 cycles (Zang et al., 2017; 2019). Some stages (HF1, HF2, HF4, & HF6) used traditional
 379 hydraulic fracturing techniques, while other stages (HF3 & HF5) employed cyclic stimulation
 380 (Zang et al., 2019). Resulting microseismicity was predominantly recorded in stages HF1-HF2

381 (López-Comino et al., 2017; Niemz et al., 2020; 2021), with events reaching up to M_w -3.5
 382 (Kwiatek et al., 2018) and slipping with reverse or strike-slip motions (López-Comino et al.,
 383 2021). Cyclic stimulations tended to produce less seismicity (Zang et al., 2019) and more
 384 complex fracture networks (Stephansson et al., 2019; Zhuang et al., 2019), albeit with less
 385 permeability enhancement (Zimmerman et al., 2019). Data for the Äspö HRL is publicly
 386 available (Zang et al., 2024) and a spatiotemporal summary of stimulation events is plotted
 387 (Figure 4).
 388



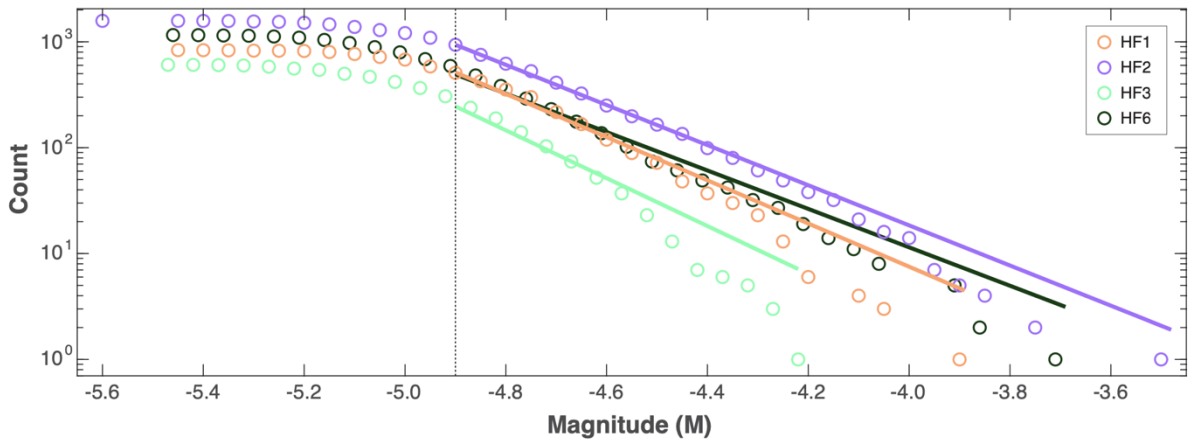
389
 390 **Figure 4. Hydraulic stimulation and earthquake response at the Äspö HRL.** a) Timings and
 391 injection rates of six stages (blue line) are compared against induced earthquake magnitudes (circles).
 392 b) Locations of six stages (diamonds; HF1-HF6), along the well bore (black line), and the resultant
 393 earthquakes (circles). Injection stages and corresponding earthquakes are colour coordinated in all
 394 panels.
 395

396 3.1.2 CAP-tests results at the Äspö HRL

397 To begin assessing if some process might be restricting magnitude growth at the Äspö
 398 HRL, we fit the GR-MFD to the catalogue data from each injection stage (*i.e.*, HF1-HF6). To

399 account for magnitude errors, we employ a 50-trial bootstrap process in which the catalogue
 400 magnitudes are dithered by ± 0.1 (this includes a dithered M_c). Only four of the stages (HF1-
 401 HF3, & HF6) have enough data to examine. Overall, the earthquakes here appear to be
 402 deficient in large-magnitude events (Figure 5). Correspondingly, the δM_{LRG} discrepancies
 403 range between -0.2 to -0.4 for all stages. Assuming an unbound catalogue, these degrees-of-
 404 truncation (δM_{LRG}) would be $<1^{st}$, 3^{rd} , $<1^{st}$, and 6^{th} percentile events, respectively. These initial
 405 assessments are suggestive of some M_{MAX} upper bound restricting catalogue growth at the
 406 Äspö HRL.

407

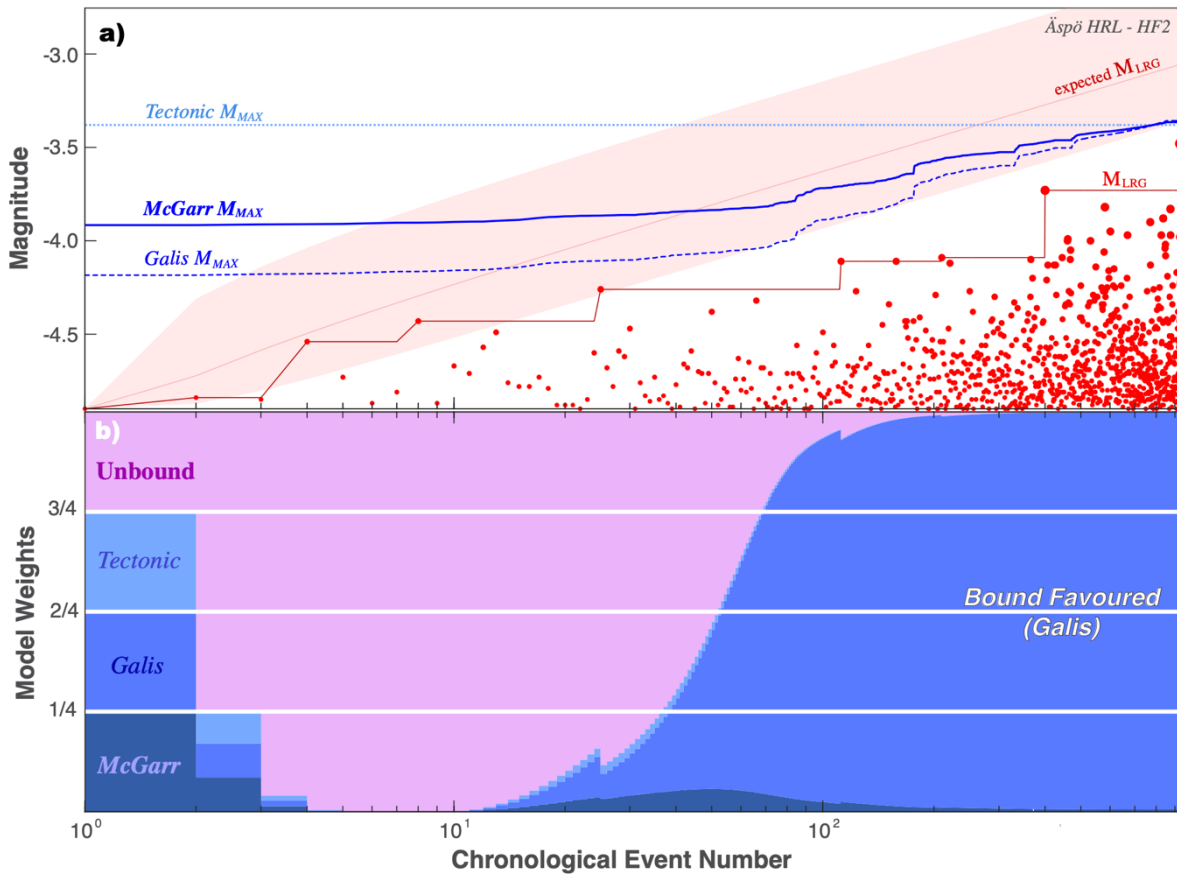


408

409 **Figure 5. Simple magnitude statistics for the Äspö HRL.** GR-MFD of cumulative events (circles),
 410 alongside best fit to the data (solid lines) and the magnitude-of-completeness (dashed line). Data is
 411 colour coordinated according to injection stages (*i.e.*, HF1-HF6).
 412

413

414 Next, we use the CAP-tests to detect and assess the potential for M_{MAX} more rigorously.
 415 The KS-test is performed 50 times, in which the catalogue magnitudes are dithered (including
 416 a dithered truncation magnitude). The KS-test also performs 100 reshuffles within each trial.
 417 Average KS-test confidences are $>99.99\%$, $>99.99\%$, 99.87% , and 68.85% for stages HF1-
 418 HF3 and HF6, respectively. We perform 50 MLE-tests using dithered catalogues and 100
 419 reshuffles within each trial. Similarly, MLE-tests suggest the standard error in fitted M_{MAX}
 420 values is 0.02, <0.01 , 0.02, and 0.18 for stages HF1-HF3 and HF6, respectively. Next, we seek
 421 test hypotheses for the functional form of M_{MAX} : our EW-tests focus on McGarr-like (McGarr,
 422 2014; Hallo et al., 2014; Elsworth, et al., 2025), the Galis-like model (Galis et al., 2017), a
 423 constant tectonic upper bound (Kanamori & Anderson, 1975), and the unbound null hypothesis.
 424 Similarly, EW-tests also show evidence for an M_{MAX} bound process, most strongly evidenced
 425 with stage HF2 (Figures 6 & S1-S3). By the end of HF2 stimulation, the Galis-like model is
 >100 times more likely than the unbound model (Figure 6).



427

428 **Figure 6. Using the EW-test to discern between M_{MAX} models for HF2 at the Äspö HRL.** a) The
 429 chronological sequence of earthquake magnitudes (red circles), the observed M_{LRG} (red line) and the
 430 expected M_{LRG} at the 10/50/90 percentiles (red area) are compared against M_{MAX} hypotheses (blue
 431 lines). b) The corresponding ensemble weights update as new data is encountered.

432

433 Last, we organize our results for all injection stages at the Äspö HRL – for convenience
 434 to the reader. These results are summarized below (Table 1).

435

Case		Simple-tests		CAP-tests				Resolution		
UGL	Stage	b -value	δM_{LRG}	KS-test	MLE-test	EW-test	M_{MAX} model	$N \geq M_C$	M_{LRG}	M_c
Äspö HRL	HF2	1.71±0.08	-0.32	>99.99%	<0.01	>100	Galis	898	-3.48	-4.9
Äspö HRL	HF1	1.91±0.11	-0.41	>99.99%	0.02	~11	Tectonic	509	-3.89	-4.9
Äspö HRL	HF6	1.90±0.11	-0.24	68.85%	0.18	~0.3	Indeterminate	540	-3.69	-4.9
Äspö HRL	HF3	2.37±0.17	-0.36	99.87%	0.02	~1.3	Tectonic	265	-4.22	-4.9

436 **Table 1. Summary of results at the Äspö HRL.** All the prior results of our simple-tests and CAP-
 437 tests are compiled here for convenience. Additionally, we have coordinated individual entries
 438 according to their interpretation: blue for bound, pink for unbound, and uncoloured for indeterminate.
 439

440

3.1.3 Interpretations for the Äspö HRL

441 All the stages at the Äspö HRL appear to have some evidence of a bound process
442 restricting the growth of earthquake magnitudes. The most convincing of the cases is also the
443 most well-resolved one: for HF2, all the simple/CAP-tests unanimously indicate a bound
444 process with strong statistical significance. As the cases become less well-resolved, the
445 statistical confidence also diminishes. For example, the HF1 and HF3 stages have most CAP-
446 tests indicating a bound process, only the EW-test of HF3 narrowly falls short of statistical
447 significance. The HF6 stage produces the most mixed results: simple-tests, MLE-test, and EW-
448 test are indeterminate, while the KS-test is unbound. It is worthwhile to mention that HF6
449 differed in stimulation approach in two ways: the first sub-stage of HF6 was the largest injected
450 volume (with reduced volume for later sub-stages) and the last three sub-stages were performed
451 on a different day.

452 Overall, we interpret stage HF2 as certainly being bound. Stages HF1 and HF3 as likely
453 to be bound, but likely needing more well-resolved catalogues to better discern M_{MAX} . We
454 hesitate to make a clear statement for HF6, which appears indeterminate due to data limitations.

455 The interpretation that all the Äspö HRL stages were (likely) bound corresponds with
456 the geophysical interpretations from relevant studies. The intention of the Äspö HRL was to
457 test various stimulation programs against the complexity/growth of hydraulic fractures (Zang
458 et al., 2019; 2021). Stage intervals were chosen to avoid natural fractures, and impression
459 packers noted the generation of new hydraulic fractures (Zimmerman et al., 2019).
460 Furthermore, the progressive growth of the HF2 hydraulic fracture plane was inferred jointly
461 from the microseismic and deformation constraints (Niemz et al., 2020; 2021). In this sense,
462 a progressively growing fracture aligns well with the bound interpretation: the finite extent of
463 a fracture limits M_{MAX} via geometric considerations (Kanamori & Anderson, 1975). As
464 stimulation continues, the fracture continues to grow; thus, the value of M_{MAX} would increase
465 alongside the injected volume. Correspondingly, each of the Äspö stages indicated bound
466 growth (albeit with varying degrees of confidence).

467 These interpretations at the Äspö HRL constitute the simplest interpretation. We
468 describe a scenario as to how hydraulic fracturing would be linked to a bound M_{MAX}
469 interpretation. Regardless of the fracture network's complexity, the finite spatial extent of
470 stimulated fractures ultimately restricts magnitude growth.

471

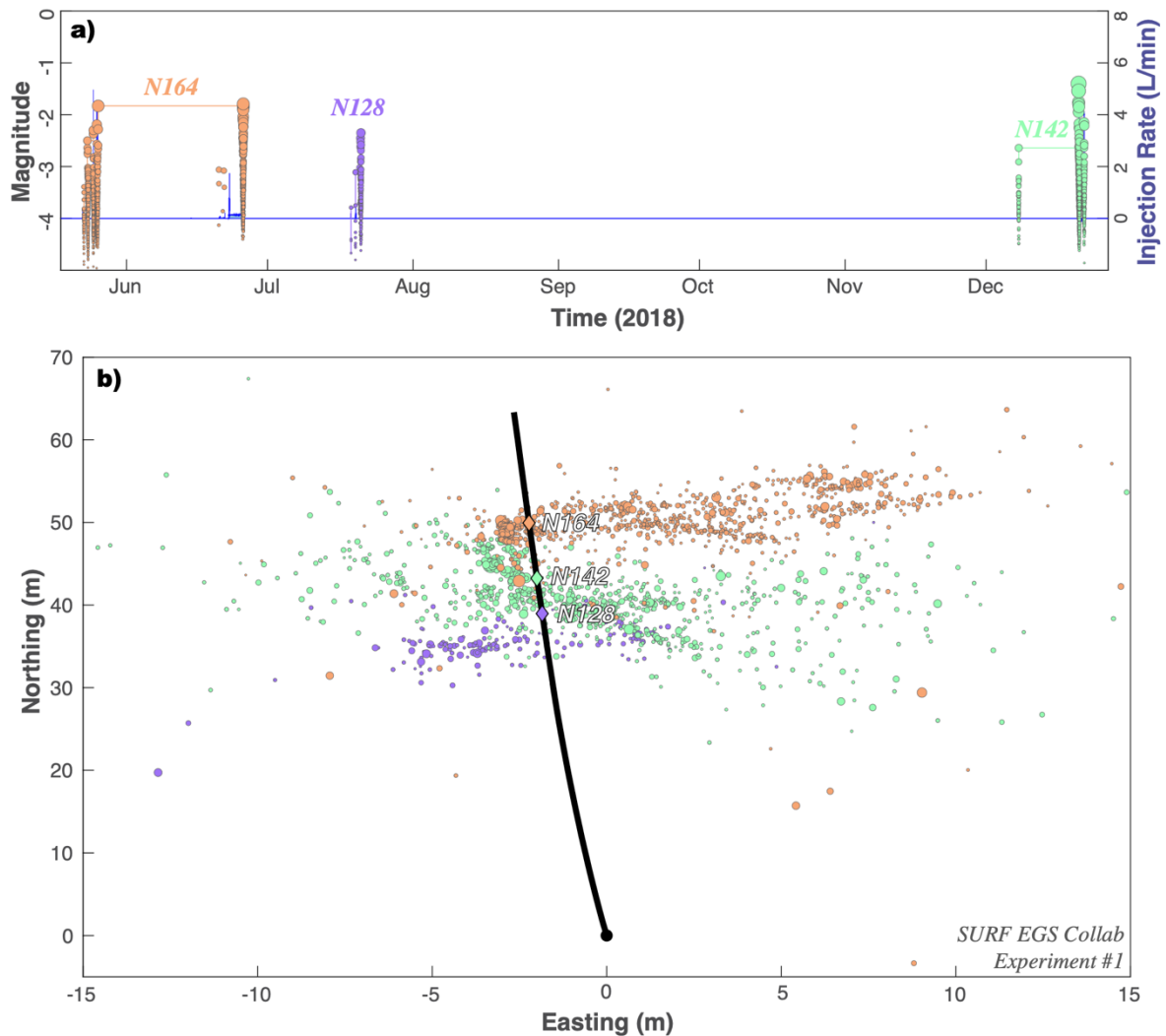
472 ***3.2 The Sanford Underground Research Facility***

473 ***3.2.1 Overview of the EGS Collab Experiment #1***

474 The SURF is located near Lead, South Dakota, and has repurposed the Homestake gold
475 mine (Figure 3). The SURF is a research facility operated by the South Dakota Science and
476 Technology Authority to study rare-process physics (Heise, 2015). The EGS Collab takes
477 advantage of the SURF, by using this facility to host hydraulic stimulation experiments at field-
478 scale depths of ~1.5 km (Dobson et al., 2018; Kneafsey et al., 2018; Morris et al., 2018).

479 The EGS Collab Experiment #1 intended to connect injection and production boreholes
480 via stimulated fractures in a controlled environment (Kneafsey et al., 2020; Morris et al., 2018).
481 Injection and production boreholes were drilled subparallel to the estimated minimum principal
482 stress direction (Oldenburg et al., 2017) in the phyllites of the Precambrian-aged Poorman
483 Formation (Kneafsey et al., 2020). These metamorphic rocks are strongly foliated and highly
484 anisotropic (Frash et al., 2019; Vigilante et al., 2017). The rock mass is generally thought to
485 be low permeability ($\sim 10^{-18}$ m²), with a thermally altered stress field (Singh et al., 2019). Some
486 natural fractures were noted in borehole cores (Fu et al., 2021), with at least one noteworthy
487 fracture that is naturally/actively flowing (Wu et al., 2021a). Starting in May 2018, stimulation
488 stages took place at three wellbore intervals (N164, N142, N128), covering both hydraulic
489 stimulation and hydraulic characterization programs; stimulation programs injected on the
490 order of 10s L of water per interval (Morris et al., 2018; White et al., 2019). Resulting
491 microseismicity was recorded for each interval, with ~2,000 located events (Schoenball et al.,
492 2019; 2020; Chai et al., 2020; Qin et al., 2024). Studies have covered topics including
493 modeling heat transport (Wu et al., 2021b), strain/deformation (Guglielmi et al., 2021), and
494 hydraulic fracture propagation (Li & Zhang, 2023). Data for the EGS Collab Experiment #1
495 is publicly available (<https://gdr.openei.org/>) and a spatiotemporal summary of stimulation
496 events is plotted (Figure 7).

497



498

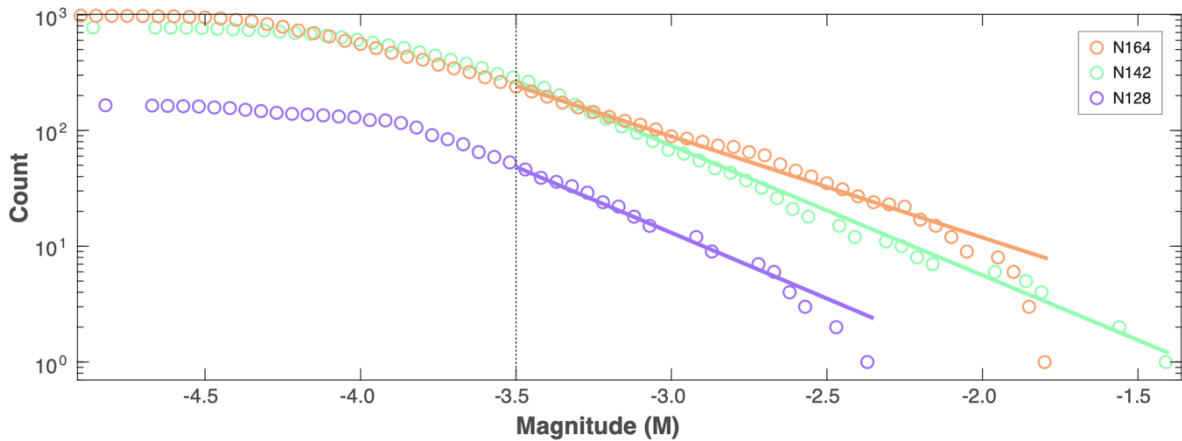
499 **Figure 7. Hydraulic stimulation and earthquake response for SURF Experiment #1.** a) Timings
 500 and injection rates of three stages (blue line) are compared against induced earthquake magnitudes
 501 (circles). b) Locations of three stages (diamonds; N164, N142, N128), along the well bore (black line),
 502 and the resultant earthquakes (circles). Injection stages and corresponding earthquakes are colour
 503 coordinated in all panels.

504

505 3.2.2 CAP-tests results from the EGS Collab Experiment #1

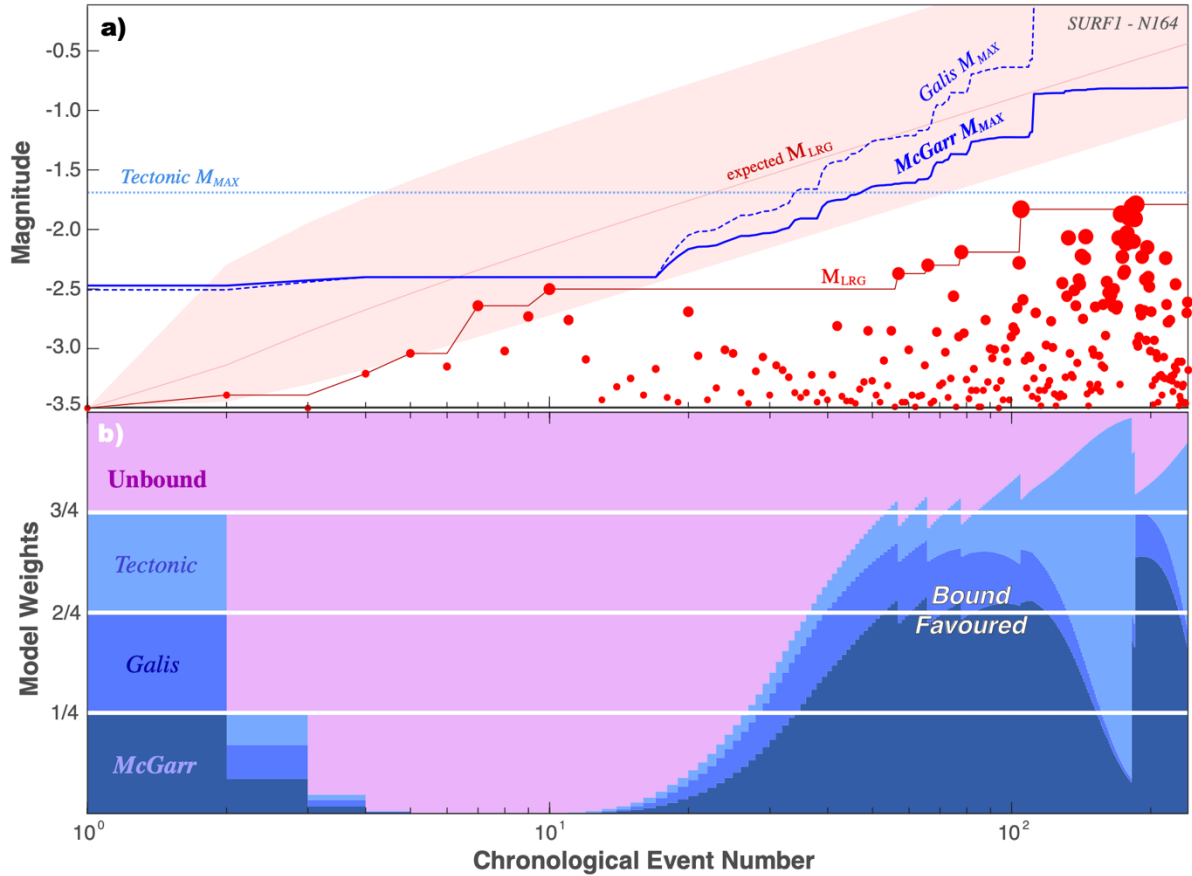
506 To begin assessing if some process might be restricting magnitude growth at SURF
 507 Experiment #1, we fit the GR-MFD to the catalogue data from each injection stage (*i.e.*, N164,
 508 N142, N128). To account for magnitude errors, we employ a 50-trial bootstrap process in
 509 which the catalogue magnitudes are dithered by ± 0.1 (this includes a dithered M_c). Only one
 510 of the stages (N164) appear to be appreciably deficient in large magnitude events (Figure 8).
 511 Correspondingly, the δM_{LRG} discrepancy is significant for N164 (-1.15 M), but fairly small for
 512 N142 (-0.05 M) and N128 (-0.28 M). Assuming an unbound catalogue, the δM_{LRG} discrepancy
 513 for N164 would be a <1st percentile event; stages N142 (10th percentile) and N128 (33rd

514 percentile) are more commonly expected occurrences. These initial assessments are suggestive
 515 of some M_{MAX} upper bound restricting catalogue growth for one stage at SURF Experiment
 516 #1.
 517



518
 519 **Figure 8. Simple magnitude statistics for SURF Experiment #1.** GR-MFD of cumulative events
 520 (circles), alongside best fit to the data (solid lines) and the magnitude-of-completeness (dashed line).
 521 Data is colour coordinated according to injection stage (*i.e.*, N164, N142, N128).
 522

523 Next, we use the CAP-tests to detect and assess the potential for M_{MAX} more rigorously.
 524 The KS-test is performed 50 times, in which the catalogue magnitudes are dithered (including
 525 a dithered truncation magnitude). The KS-test also performs 100 reshuffles within each trial.
 526 Again, KS-test results are split, with average confidences of 99.96% for stage N164, but
 527 34.23% and 74.39% for stages N142 and N128, respectively. We perform 50 MLE-tests using
 528 dithered catalogues and 100 reshuffles within each trial. MLE-tests are also split: standard
 529 error in fitted M_{MAX} values are 0.10, 2.18, and 1.31 for stages N164, N142, and N128,
 530 respectively. Similarly, EW-tests also shows suggestive evidence for an M_{MAX} bound process
 531 with N164 (Figure 9), but certainly not for stages N142 and N128 (Figures S4-S5).
 532



533

534 **Figure 9. Using the EW-test to discern between M_{MAX} models for N164 for SURF Experiment #1.**

535 a) The chronological sequence of earthquake magnitudes (red circles), the observed M_{LRG} (red line),
 536 and the expected M_{LRG} at the 10/50/90 percentiles (red area) are compared against M_{MAX} hypotheses
 537 (blue lines). b) The corresponding ensemble weights update as new data is encountered.

538

539 Last, we organize our results for all injection stages at SURF Experiment #1 – for
 540 convenience to the reader. These results are summarized below (Table 2).

541

Case		Simple-tests		CAP-tests				Resolution		
UGL	Stage	b -value	δM_{LRG}	KS-test	MLE-test	EW-test	M_{MAX} model	$N \geq M_C$	M_{LRG}	Mc
SURF #1	N164	0.83±0.02	-1.15	99.96%	0.10	~6.0	McGarr/Tectonic	239	-1.79	-3.5
SURF #1	N142	1.13±0.04	-0.06	34.23%	2.18	~0.11	Unbound	282	-1.40	-3.5
SURF #1	N128	1.18±0.05	-0.28	74.39%	1.31	~0.04	Unbound	50	-2.35	-3.5

542 **Table 2. Summary of results for the SURF EGS Collab Experiment #1.** All the prior results of our
 543 simple-tests and CAP-tests are compiled here for convenience. Additionally, we have coordinated
 544 individual entries according to their interpretation: blue for bound, pink for unbound, and uncoloured
 545 for indeterminate.

546

547 3.2.3 Interpretations for the EGS Collab Experiment #1

548 CAP-test results at SURF Experiment #1 have similarities and differences from those
 549 at the Äspö HRL. For example, stage N164 is similar to the stages at the Äspö HRL: both

550 simple-tests and CAP-tests unanimously agree on an M_{MAX} bound – with varying degrees of
551 confidence. On the other hand, stages N142 and N128 differ in that they produce strong and
552 unambiguous unbound inferences. Because of this, we interpret stages N142 and N128 to be
553 truly unbound.

554 The interest in considering UGL cases is the wealth of complementary geophysical
555 information to cross-examine against the results of CAP-tests. The stimulation at stage N164
556 is predominantly understood to be the creation of a new fracture network. This interpretation
557 comes from multiple lines of evidence: the orientation of microseismic fault planes with respect
558 to the ambient stress field (Schoenball et al., 2020), deformation constraints on fracture motion
559 (Guglielmi et al., 2021), significant recovery of injected fluid (White et al., 2019), and direct
560 evidence of new fluid jets intersecting the producing well (Fu et al., 2021). The later N164
561 injections (*i.e.*, June 2018) were performed from the producing well side, into the newly created
562 fracture network. That said, some complexity in the N164 stimulation suggested limited
563 interaction with natural fractures, via either arrests or redirected continuation of fracture growth
564 (Schoenball et al., 2020; Fu et al., 2021). Overall, these inferences are consistent with finite
565 hydraulic fractures bounding the growth of earthquake magnitudes – which are the same
566 interpretations made for the Äspö HRL.

567 On the other hand, stages N142 and N128 had significant interactions with natural pre-
568 existing fractures, and likely reactivated them in shear rather than creating a new fracture
569 network. These inferences/interpretations come from the orientation of microseismic fault
570 planes with respect to the ambient stress field, corroboration with well image logs, and
571 observations of shear deformation (Schoenball et al., 2020). The connection to fractures/faults
572 allows for seemingly unbound magnitude growth, as the spatial extent of these pre-existing
573 structures have the potential to host larger events.

574 These interpretations at SURF Experiment #1 constitute the next level of complexity in
575 interpretation. We describe a scenario where hydraulic stimulation could interact with natural
576 fractures/faults to produce an unbound process. Said another way, the interaction with natural
577 fractures/faults (of sufficient size) can facilitate unbound magnitude growth.

578

579 ***3.3 The Grimsel Test Site***

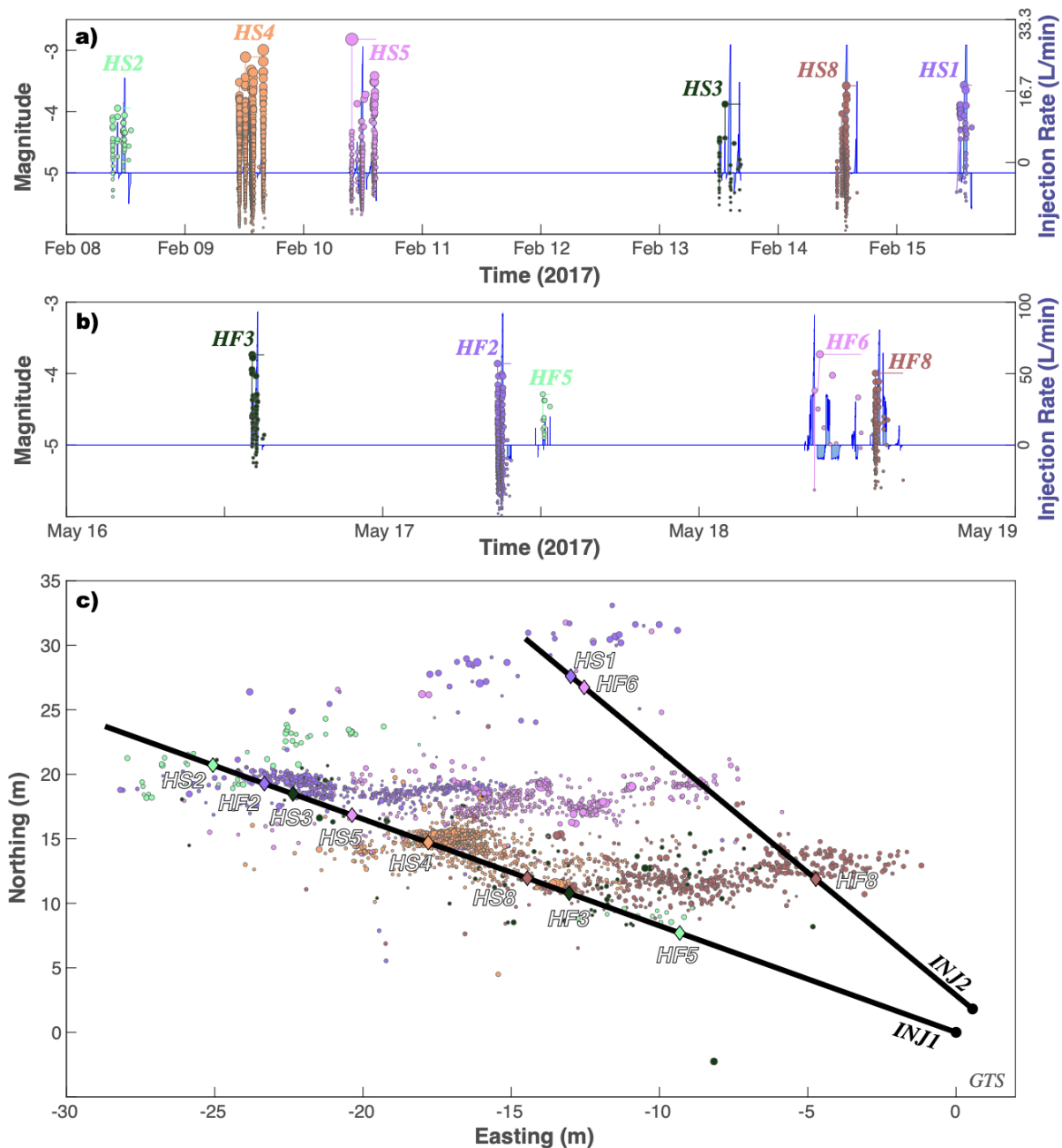
580 *3.3.1 Overview of the GTS*

581 The GTS is located near the Grimsel pass in the central Swiss Alps (Figure 3), ~450 m
582 below the Juchlistock in the Varsican-aged Aar Massif granites (Schneeberger et al., 2019).
583 The GTS was established in 1984 as an underground research facility to study the safe disposal

584 of nuclear waste. It is owned and operated by the National Cooperation for Radioactive Waste
585 Disposal (Nagra; <https://www.grimse.com/>).

586 The intention of recent (February-May 2017) injection experiments at the GTS was to
587 demonstrate the stimulation of fractures at the decameter scale and to better understand how to
588 manage induced seismicity (Gischig et al., 2016; Amann et al., 2018). Twelve injection stages
589 spanning ~1 m intervals were situated in two boreholes (~45 m long) (Gischig et al., 2020): six
590 related to stimulation via hydraulic fracturing (HF1-HF8) (Dutler et al., 2019) and six more via
591 hydroshearing (HS1-HS8) (Krietsch et al., 2020b). Each stage used (on the order of) ~1000 L
592 of injected fluid, spread between 4 injection sub-cycles. The role of natural faults and fractures
593 are a prominent focus of the GTS injection experiments, which injected into either brittle-
594 ductile shear zones, ductile shear zones, or intact rock (Doetsch et al., 2018a). The degree of
595 seismic response for individual stages was strongly heterogeneous in space, with stages HS4,
596 HS5, and HF2 being both the most seismically active and hosting the large events (M -3.0, M
597 -2.8 M, & M -3.9) (Villiger et al., 2020). Studies at the GTS were diverse, covering topics like
598 geological characterization (Krietsch et al., 2018a), stress inversion (Krietsch et al., 2019),
599 tomographic velocity changes (Doetsch et al., 2018b; Schopper et al., 2020), permeability
600 changes (Jalali et al., 2018; Brixel et al., 2020a; 2020b), and inferring fracture propagation
601 from hydromechanical response (Dutler et al., 2019; Krietsch et al., 2020a; 2020b). Data for
602 the GTS is publicly available (<https://doi.org/10.3929/ethz-b-000276170>) and a spatiotemporal
603 summary is plotted (Figure 10).

604



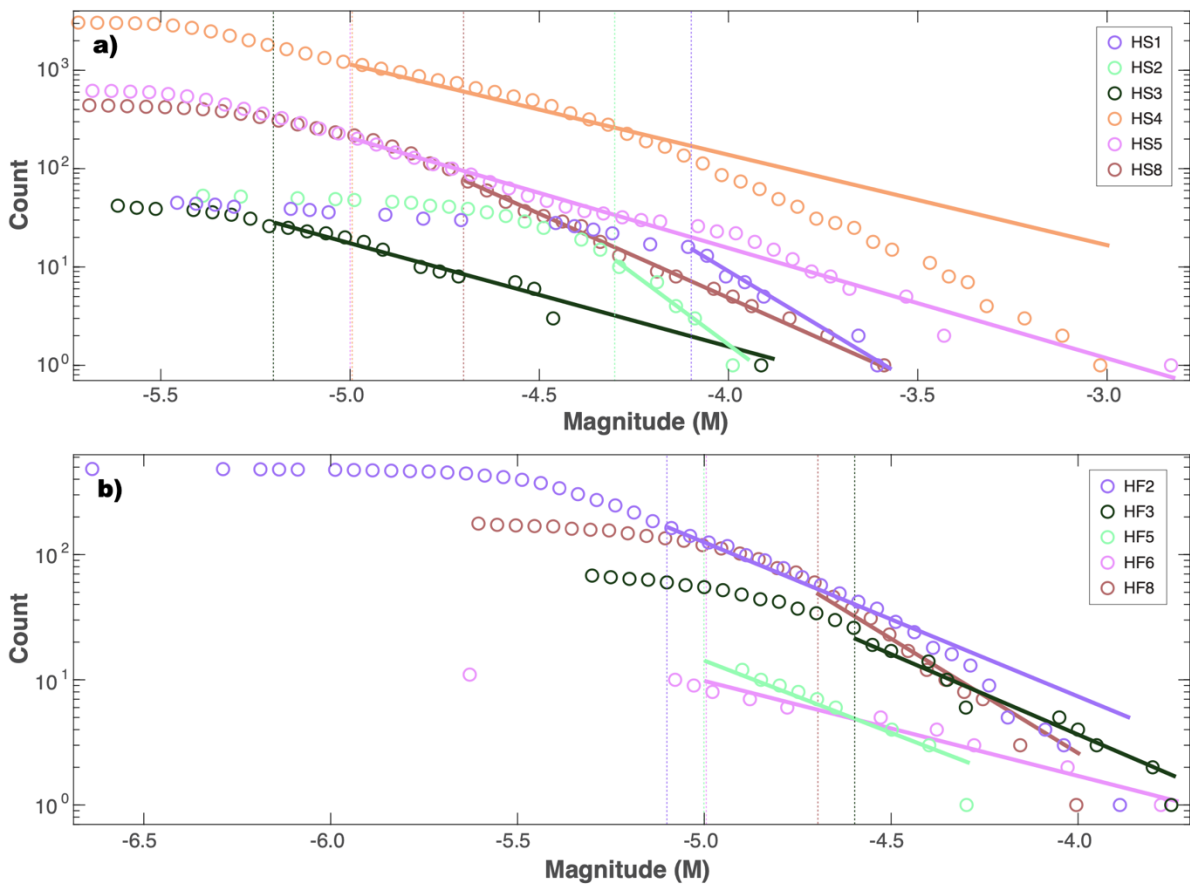
605

606 **Figure 10. Hydraulic stimulation and earthquake response at the GTS.** a & b) Timings and
 607 injection rates of twelve stages (blue line) are compared against induced earthquake magnitudes
 608 (circles). c) Locations of twelve stages (diamonds; HFX & HSX), along the well bore (black line), and
 609 the resultant earthquakes (circles). Injection stages and corresponding earthquakes are colour
 610 coordinated in all panels.
 611

612 3.3.2 CAP-tests results at the GTS

613 To begin assessing if some process might be restricting magnitude growth at the GTS,
 614 we fit the GR-MFD to the catalogue data from each injection stage (*i.e.*, HSX & HFX). To
 615 account for magnitude errors, we employ a 50-trial bootstrap process in which the catalogue
 616 magnitudes are dithered by ± 0.1 (this includes a dithered M_c). Note that many of the stages

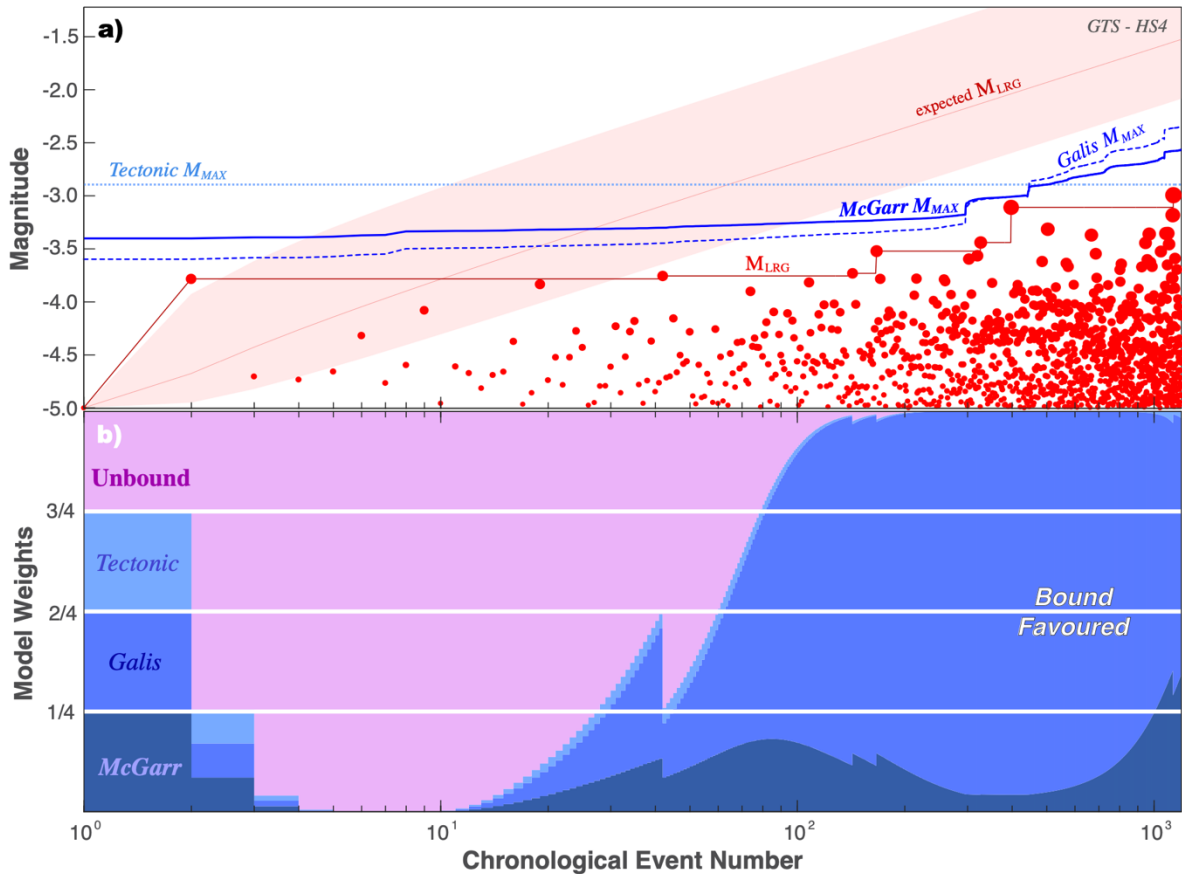
617 recorded here have too few events for a meaningful analysis; thus, we predominantly focus our
 618 discussions to a subset of stages. Only two of the stages (HS4 & HF2) appear to be appreciably
 619 deficient in large magnitude events (Figure 11). Correspondingly, the δM_{LRG} discrepancy is
 620 large for HS4 and HF2 (-1.31 M & -0.61 M), but fairly small for the other viable stages
 621 (between -0.19 and +0.08 M). Assuming an unbound catalogue, the δM_{LRG} discrepancy for
 622 HS4 and HF2 would both be <1st percentile events; all other stages are commonly expected
 623 occurrences. These initial assessments are suggestive of some M_{MAX} upper bound restricting
 624 catalogue growth for just two stages at the GTS.
 625



626
 627 **Figure 11. Simple magnitude statistics at the GTS.** GR-MFD of cumulative events (circles),
 628 alongside best fit to the data (solid lines) and the magnitude-of-completeness (dashed lines). Data is
 629 colour coordinated according to injection stage (*i.e.*, HSX & HFX) and separated into hydroshearing
 630 (a; HSX) or hydrofracturing experiments (b; HFX).
 631

632 Next, we use the CAP-tests to detect and assess the potential for M_{MAX} more rigorously.
 633 The KS-test is performed 50 times, in which the catalogue magnitudes are dithered (including
 634 a dithered truncation magnitude). The KS-test also performs 100 reshuffles within each trial.
 635 KS-test results are split, with strong confidences of >99.99% for stages HS4 & HF2, but

636 unconvincing values for all other stages. We perform 50 MLE-tests using dithered catalogues
 637 and 100 reshuffles within each trial. MLE-tests are also split: standard error in fitted M_{MAX}
 638 values are <0.01 and 0.02 for stages HS4 and HF2, but values are large (0.34 - 1.92) for all other
 639 stages. The EW-tests shows decisive evidence for an M_{MAX} bound process with HS4 (Figure
 640 12), and strong evidence for HF2 (Figure S6). All other stages have substantial-to-strong
 641 evidence for an unbound process via EW-tests (Figures S7-S10).
 642



643
 644 **Figure 12. Using the EW-test to discern between M_{MAX} models for HS4 at the GTS.** a) The
 645 chronological sequence of earthquake magnitudes (red circles), the observed M_{LRG} (red line), and the
 646 expected M_{LRG} at the 10/50/90 percentiles (red area) are compared against M_{MAX} hypotheses (blue
 647 lines). b) The corresponding ensemble weights update as new data is encountered.
 648

649 Last, we organize our results for all injection stages at the GTS – for convenience to
 650 the reader. These results are summarized below (Table 3).

651

Case		Simple-tests		CAP-tests				Resolution		
UGL	Stage	b -value	δM_{LRG}	KS-test	MLE-test	EW-test	M_{MAX} model	$N \geq M_C$	M_{LRG}	M_C
GTS	HS4	0.93 ± 0.02	-1.31	>99.99%	<0.01	>100	Galis	1194	-3.00	-5.0
GTS	HS5	1.11 ± 0.05	+0.08	32.45%	0.76	~ 0.07	Unbound	214	-2.82	-5.0

<i>GTS</i>	<i>HF2</i>	1.20±0.05	-0.61	>99.99%	0.02	~60	Tectonic	170	-3.86	-4.9
<i>GTS</i>	<i>HS8</i>	1.68±0.11	-0.02	71.14%	1.34	~0.15	Unbound	81	-3.58	-4.7
<i>GTS</i>	<i>HF8</i>	1.97±0.14	-0.19	75.82%	0.34	~0.27	Unbound	58	-4.00	-4.7
<i>GTS</i>	<i>HF3</i>	1.55±0.10	-0.06	29.02%	1.92	~0.05	Unbound	26	-3.74	-4.6
<i>GTS</i>	<i>HS3</i>	1.05±0.05	-0.04	62.04%	0.98	~0.04	Unbound	26	-3.88	-5.2
<i>GTS</i>	<i>HS1</i>	2.40±0.22	+0.03	46.97%	1.26	~0.01	Unbound	15	-3.57	-4.1
<i>GTS</i>	<i>HS2</i>	2.76±0.38	-0.03	40.94%	0.81	~0.01	Unbound	11	-3.94	-4.3
<i>GTS</i>	<i>HF5</i>	1.16±0.05	-0.22	94.34%	2.23	~0.04	Unbound	12	-4.29	-5.0
<i>GTS</i>	<i>HF6</i>	0.73±0.06	-0.01	26.17%	3.61	~0.01	Unbound	9	-3.73	-5.0

652 **Table 3. Summary of results at the GTS.** All the prior results of our simple-tests and CAP-tests are
653 compiled here for convenience. Additionally, we have coordinated individual entries according to their
654 interpretation: blue for bound, pink for unbound, and uncoloured for indeterminate.
655

656 3.3.3 Interpretations for the GTS

657 In summary, most of the GTS stages appear to exhibit unbound growth of earthquake
658 magnitudes. Although, this interpretation varies in confidence, depending on the specific stage
659 in question, since many of the stages had relatively few events recorded. The most confidently
660 unbound case is HS5, which fails all of the simple-tests and CAP-tests. On the other hand,
661 there are two exceptions to this general trend: HS4 shows clear and definitive evidence for a
662 bound process, while HF2 shows strong evidence for a bound process. That said, stages with
663 fewer than 50 events are difficult to arrive at a clear interpretation between truly unbound or
664 simply lacking data.

665 The aims and scope of the GTS were to investigate the response of injection into
666 faults/fractures at the intermediate scale (Gischig et al., 2016; Amann et al., 2018). For
667 example, all HSX stages intentionally injected into previously known fractures of shear damage
668 zones (Krietsch et al., 2018); thus, micro-seismically delineated fault planes were generally
669 consistent with fracture/fault orientations (Villiger et al., 2020), near-field hydromechanical
670 effects were consistent with pre-existing fracture dislocation (Krietsch et al., 2020b), and many
671 stages were hydraulically connected to other boreholes via these stimulated fractures (Brixel et
672 al., 2020b). In particular to HS5, seismicity was well fit to a single fault plane (~16 m diameter)
673 oriented subparallel to the targeted fracture (Villiger et al., 2020), significant pressure increases
674 (~70-75% of injection pressure) were observed in boreholes 7-8 m away (Krietsch et al.,
675 2020a), likely due to transient permeability increases driven by fracture aperture changes
676 (Krietsch et al., 2020b). Furthermore, the diameter of the HS5 fault plane (~16 m) is
677 appreciably larger than the largest event diameter (~0.65 m), assuming a circular crack. HFX
678 stages started hydraulically fracturing intact rock and subsequently propagated/stimulated new
679 fractures (Dutler et al., 2019). However, the newly propagating fractures were inferred to have

680 significant interactions with pre-existing fractures, which served as pressure sinks that arrested
681 further growth (Dutler et al., 2019; Villiger et al., 2020). Note that HF6 is unique, since it
682 injected directly into a pre-existing fracture (by mistake) and thus can be considered a
683 hydroshearing experiment. Furthermore, the propagation of microseismic events along pipe-
684 like geometries was thought to be formed via the intersection with natural fractures (Dutler et
685 al., 2019). Given the prior interpretations of CAP-test results at SURF Experiment #1, we
686 would anticipate that all GTS stimulations should be unbound – because of their prominent
687 connection to pre-existing fractures. Bound cases would require special exceptions to this
688 generalization.

689 On the other hand, the HS4 stage is a clear exception to this generalization. The HS4
690 seismicity is not well-fit by a single plane. Instead, it is best fit by four intersecting planes (C1-
691 C4), where C1-C3 are oriented subparallel with pre-existing fractures and C4 is likely a new
692 tensile failure (Villiger et al., 2020; 2021). Clustering in focal mechanism slip style also
693 corresponds to spatial clusters (Villiger et al., 2021). Each cluster has seismicity spatially
694 restricted along discrete linear streaks (Villiger et al., 2021); these streaks grow/propagate
695 alongside injection, although their spatial extent (~1-2 m) is much smaller than other HSX fault
696 planes. In tectonic contexts, streaks are often interpreted as rheological boundaries between
697 seismic and creeping/locked fault segments (Rubin et al., 1999; Waldhauser et al., 2004). In
698 other hydraulic fracturing contexts, streaks have also been observed and analogously
699 interpreted (Rutledge et al., 2004; Evans et al., 2005). Thus, the streaks and clustering at HS4
700 have been interpreted as fractures channelizing fluid-flow towards highly seismogenic
701 asperities that slip perpendicular to the fluid migration direction (Villiger et al., 2021). In this
702 sense, we argue that HS4 seismicity is spatially bound to these asperities/streaks, giving rise to
703 the bound growth of magnitudes. Correspondingly, the GR-MFD for HS4 (Figure 11a) starts
704 to roll-off around $M -4.3$ and $M -3.0$ was the final M_{LRG} value; note that $M -3.0$ roughly
705 corresponds to a circular crack diameter of ~0.5-1.0 m, which is comparable to the spatial
706 extent of the streaks (~1-2 m).

707 Similarly, HF2 is also an exception to this generalization. The HF2 seismicity was best
708 fit by two intersecting planes, the first which resembles a newly created hydraulic fracture and
709 then a deflection to a more E-W orientation (Dutler et al., 2019; Villiger et al., 2020). It has
710 been suggested that this stage is exceptional in that it potentially has limited interaction with
711 pre-existing structures, being able to propagate fracture growth before leak-off into potential
712 fracture connections (Dutler et al., 2019; Villiger et al., 2020). In fact, the growing hydraulic
713 fracture intersected monitoring boreholes, reaching further than suggested by the microseismic

714 events (Dutler et al., 2019). In this sense, we would argue that HF2 is most similar to the N164
715 stage at SURF Experiment #1. Correspondingly, both HF2 and N164 express a bound process
716 via CAP-tests.

717 These interpretations at the GTS constitute the greatest level of interpretive complexity
718 we will consider in this study. We have reiterated scenarios where hydraulic stimulation could
719 interact with natural fractures/faults to produce an unbound process. Regardless of the
720 stimulation program, the interaction with natural fractures/faults (of sufficient size) can
721 facilitate unbound magnitude growth. Exceptions come in the form of hydraulic fracturing
722 with limited interactions to natural fractures (HF2) and shear reactivation with seismicity
723 limited to streaks/asperities (HS4).

724

725 **4. Discussion**

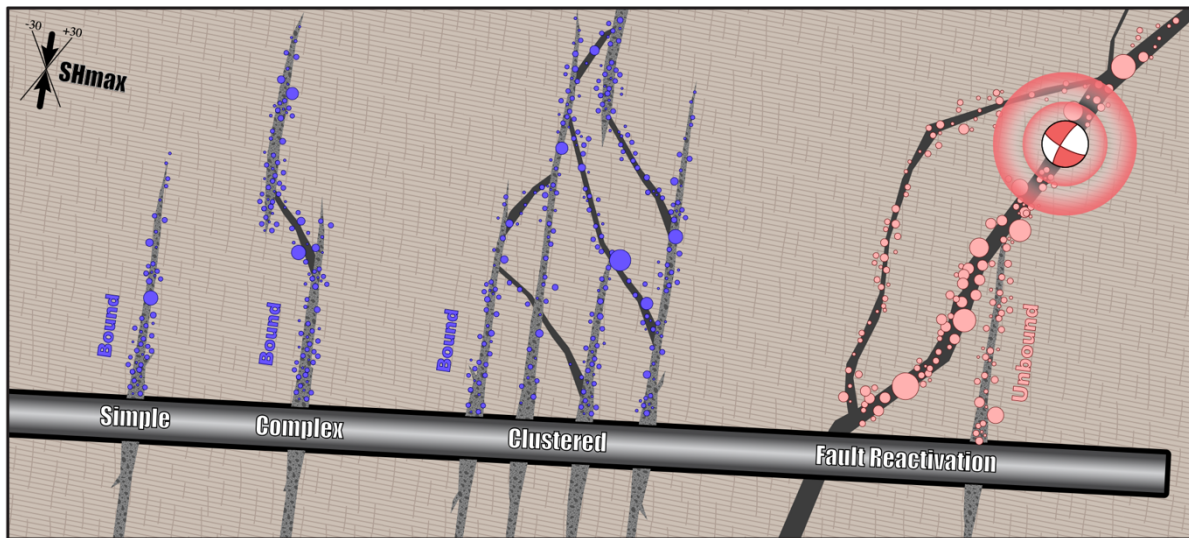
726 **4.1 Consolidating M_{MAX} interpretations**

727 In this sub-section, we synthesize the results and interpretations of this study in
728 comparison with relevant literature. We begin to jointly interpret the physical rationale behind
729 bound/unbound earthquake sequences. We start by reiterating the interpretations of CAP-tests
730 results at the UGLs and then discuss the correspondence with CAP-tests at the field-scale.

731 In the simplest case of the Äspö HRL (Section 3.1), stages were stimulated among intact
732 rock, creating new hydraulic pathways via stimulated fractures. The finite geometric extent of
733 these stimulated fractures limits the size of the largest events that can occur, at any given instant
734 (Figure 13). Although, the spatial extent of this fracture-bound M_{MAX} will change with
735 time/injection. Correspondingly, all Äspö HRL cases indicated a bound M_{MAX} process via
736 CAP-tests. While newly stimulated fractures were the intention of the Äspö HRL, bound cases
737 were also observed for analogous stages at SURF Experiment #1 (N164; Section 3.2) and the
738 GTS (HF2; Section 3.3), despite their different experiment objectives.

739 On the other hand, stages encountering more complicated interactions with pre-existing
740 fractures/faults tended to exhibit an unbound catalogue. For example, the simplest case of
741 fracture interaction was noted at SURF Experiment #1 (Section 3.2): the two stages (N142 &
742 N128) that intersected a natural fracture were also unbound via CAP-tests. Furthermore,
743 injection into faults/fractures was an aim of the GTS (Section 3.3), and these stages
744 predominantly exhibited an unbound process. We interpret this correspondence via
745 geometrical considerations again (Figure 13): an immediate hydraulic connection to a relatively
746 large-scale fracture system provides an opportunity for (apparently) unrestricted magnitude

747 growth. While we have not observed it at these three UGLs (to our knowledge), interactions
748 with smaller-scale pre-existing fractures would be expected to produce a bound sequence.
749



750
751 **Figure 13. Interactions between stimulated fractures and reactivated faults during hydraulic**
752 **fracturing.** A series of stages of increasing complexity (text labels) are considered alongside a
753 hypothetical well (black & grey rectangle). In the simple case, the host rock is split in tension via
754 stimulated fractures (grey polygons). In the complex case, small pre-existing faults (black polygons)
755 can also be reactivated in shear slip. The clustered case hydraulically connects a series of
756 stages/perforations into a single fracture/fault network. In the last case, stimulated fractures intersect a
757 large fault system that is critically-stressed for slip. By watching the growth of earthquake magnitudes,
758 CAP-tests can discern between bound cases (blue circles) and unbound cases (pink circles).
759

760 It is worth emphasizing that stage HS4 at the GTS is exceptional (Section 3.3), in that
761 a bound M_{MAX} process was observed for a case with strong natural fracture interactions. Thus,
762 this stage is an outlier to the general interpretive theme regarding natural faults/fractures. That
763 said, there is also good evidence for an alternative bounding mechanism at HS4. Earthquakes
764 were restricted to smaller streaks along larger fault plane trends (Villiger et al., 2021); these
765 streaks grew in spatial extent with increased injection volume. Streaks are interpreted as fault
766 heterogeneity, where seismic asperities are surrounded by a broader creeping/locked fault
767 segment (Rubin et al., 1999; Waldhauser et al., 2004; Rutledge et al., 2004; Evans et al., 2005).
768 Furthermore, the spatial extent of these streaks is comparable to the fault plane area of the
769 largest events observed. In this sense, HS4 seismicity is likely bound to these asperities/streaks,
770 giving rise to the bound growth of magnitudes. Following this logic, it should be possible to
771 test seismic asperities for a bound M_{MAX} process (via CAP-tests) in tectonic settings too.

772 A recent study performed CAP-tests on field-scale experiments (Schultz et al., 2025),
773 including both shale gas hydraulic fracturing at Preston New Road in the UK (Clarke et al.,

2019; Kettlety et al., 2021) and the enhanced geothermal system at Utah FORGE (Moore et al., 2019; Niemz et al., 2025). CAP-tests indicated that many clustered stages followed a bound process, while the more hazardous stages were unbound. Interpretations in these field-scale cases followed a similar interpretation to UGL stages in this study (Section 4.1.1), albeit with larger uncertainties. This interpretation was partly driven by clustering sensitivity tests that showed omitting/ignoring earlier bound stages tended to diminish statistical confidence (Schultz et al., 2025) – providing an inference to the importance of accurate clustering and hydraulic connectivity for bound sequences. Certainly, natural fractures have demonstrably influenced the propagation of stimulated fractures (Jin et al., 2024). In this sense, the verification of this interpretation at UGLs provides a potential conceptual linkage to the field-scale, given their observational similarities. Prior results for bound processes are consistent with stimulating a fracture system with a restricted extent, while unbound cases likely have reactivated larger fault systems (Schultz, 2024).

787

788 ***4.3 Empirically constraining the functional form of M_{MAX} vs V***

789 Up to this point, our study has focused on discerning between bound/unbound
790 sequences – and establishing the importance of natural fault/fracture systems within this
791 dichotomy. However, our discussion has neglected the underlying physical mechanisms
792 responsible for creating this bound M_{MAX} process. Here, we further discuss the underlying
793 physical mechanisms, for all the known bound sequences with data available (Figure 3) – both
794 in this study and past studies (Schultz et al., 2025).

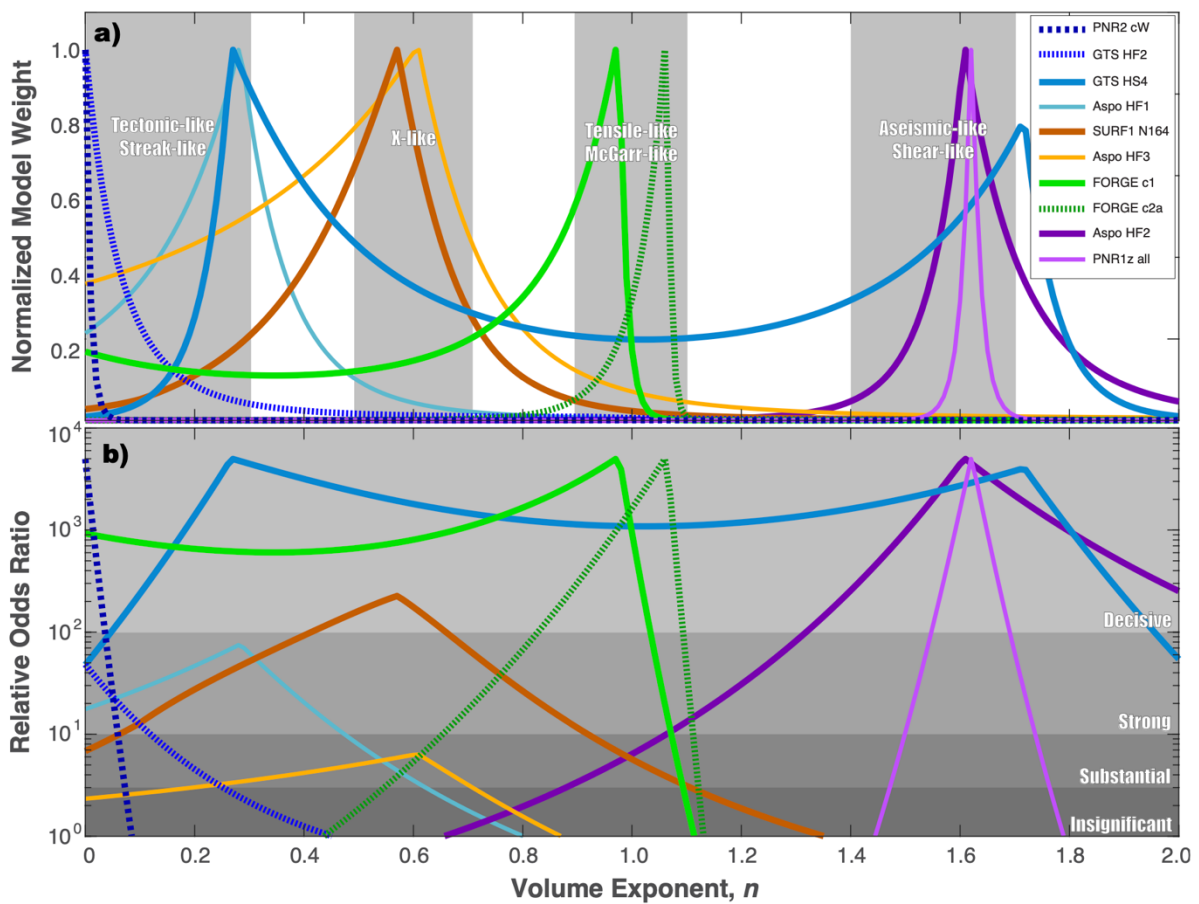
795 Numerous theoretical models of M_{MAX} and fracture propagation have been proposed in
796 the literature (Eaton & Igonin, 2018). The simplest model considers a physical limitation based
797 on the geometry of the finite fault extent (Kanamori & Anderson, 1975); For a circular fault
798 with radius R and constant stress drop $\Delta\sigma$, the seismic moment release is given via the equation
799 $M_0 = \frac{16}{7} \Delta\sigma R^3$. Seismic moment can then be translated into an M_{MAX} magnitude via the
800 moment magnitude relationship (Hanks & Kanamori, 1979). Non-stationary M_{MAX} models
801 have been proposed for induced seismicity based on the footprint of stress perturbation along
802 a fault (Shapiro et al., 2011), the evolution of seismic moment (McGarr, 2014; Hallo et al.,
803 2014; Elsworth et al., 2025), self-arrested rupture dynamics (Galis et al., 2017), or time-
804 dependent pressure diffusion (Shapiro et al., 2021). Additionally, we can consider the finite
805 extent of fracture propagation via an aseismic shear crack (Danré et al., 2024) or a tensile crack
806 (Davis et al., 2020). These two models suggest time-varying fracture radii based on the

807 equations $R(t) = \sqrt{s_d V(t)}$ and $R(t) = \sqrt[3]{k V(t)}$, respectively. Here, s_d and k are injection
808 propagation coefficients. M_{MAX} models have been investigated in numerous prior studies
809 (Kwiatek et al., 2018; Li et al., 2021; Yu et al., 2024; Langenbruch et al., 2024; Lanza et al.,
810 2026).

811 Examining the theoretical models illustrates a common mathematical theme: seismic
812 moment release (M_0) is proportional to injected volume, raised to some exponent. For example,
813 the McGarr-like model considers $M_0 \propto V^1$ (McGarr, 2014), while the Galis-like model
814 considers $M_0 \propto V^{3/2}$ (Galis et al., 2017). Similarly, a growing tensile crack is equivalent to M_0
815 $\propto V^1$ (Kanamori & Anderson, 1975; Davis et al., 2020), while a growing aseismic shear crack
816 is $M_0 \propto V^{3/2}$ (Kanamori & Anderson, 1975; Danré et al., 2024; Im & Avouac, 2025; Sáez et
817 al., 2025). Even purely tectonic/geometric limitations can be cast into this form as $M_0 \propto V^0$.
818 Given this commonality, we consider a generic family of M_{MAX} models that have a functional
819 form of $M_{\text{MAX}} \propto \log_{10}(V^n)$. This generalization is advantageous, as the EW-test can consider
820 any arbitrary M_{MAX} function. Following this logic, we use these arbitrary functions within the
821 EW-test to find the best-fitting n -exponent for each bound case. To ensure a fair comparison,
822 each competing/candidate M_{MAX} model is first fit to the case data before the EW-test selects
823 the best model.

824 This investigation indicates that the underlying physical mechanisms are varied among
825 cases (Figure 14). However, the results tend to aggregate into four distinct categories:
826 geometric/streak-like (n 0.0-0.3), X-like (n 0.5-0.7), McGarr/tensile-like (n 0.9-1.1), and
827 aseismic/shear-like (n 1.4-1.7). Bounding mechanisms can be varied, even between similar
828 stages (*e.g.*, Äspö HRL, HF1-HF3). Higher n -exponents tend to be less confident, likely
829 because aseismic/shear-like exponents emulate/approach unbound-like magnitude growth (van
830 der Elst et al., 2016; Galis et al., 2017). Other geophysical data/studies to corroborate the
831 inferences of physical mechanisms are generally limited to a few cases. For example, the
832 propagation of microseismic events at Utah FORGE stages 1-3 were best-fit by tensile-like
833 models (Clarkson et al., 2025; Lanza et al., 2026; Niemz et al., 2026), which independently
834 corresponds to our best-fit n -exponents (0.97 & 1.06) for clusters 1 (stages 1-2) and 2 (stages
835 3-6) (Figure 14), as defined in Schultz et al. (2025). Interestingly, the low n -exponent
836 associated with HS4 at the GTS is consistent with the observation of tectonic streaks/asperities
837 that exhibited limited spatial growth during injection (Villiger et al., 2021). The X-like
838 category is exceptional, since there is (to our knowledge) no previously suggested model for
839 this n -exponent. This could be potentially explained by M_{MAX} models that do not fit the V^n

840 generalization (Shapiro et al., 2011; 2021), an as-of-yet undiscovered M_{MAX} model, or
 841 observational biases for encountering a relatively large pre-existing fracture network (V^0) that
 842 also starts growing alongside injection (V^1 or $V^{3/2}$). Currently, we favour the observational
 843 bias interpretation for the X-like category, based on the results of drop-out tests that
 844 demonstrate how an intermediate sized pre-existing fracture network that also grows alongside
 845 injection can produce a mixed n -exponent with poorly-resolved data (Supplementary Section
 846 S1 & Figure S16). We consider it likely that these considerations were a factor for the SURF
 847 N164 and Äspö HF3 cases, which indicated the next closest M_{MAX} model when restricting
 848 choices to our three bound options (Tables 1 & 2). Accounting for some prior growth appears
 849 to be an important consideration for discerning the true M_{MAX} model.



850
 851 **Figure 14. Searching for volume-exponents.** The EW-test fit to case data (lines) for both a)
 852 normalized model weights and b) odds ratios (relative to the unbound null hypothesis). Strength of
 853 statistical confidence and n -exponent categories are shown with grey shaded areas (and text callouts).
 854 Note that relative odds ratios are capped at 5000 to facilitate easier comparison via plotting.
 855

856 Additional UGL study alongside refinements to CAP-tests (potentially also considering
 857 other candidate M_{MAX} models) will likely shed light on the physical processes underlying
 858 bound sequences. For example, CAP-tests could be used to systematically search for M_{MAX}

859 models in other datasets. We note that this approach could be straightforwardly adapted to
 860 forecast the ‘next largest event’ (Cao et al., 2020), using these weights as part of either a logic-
 861 tree (Bommer & Verdon, 2024) or an ensemble of models (Schultz et al., 2023c).

862

863 *4.3 Exceptions to bound/unbound interpretations*

864 Thus far, we have predominantly interpreted the difference between bound/unbound
 865 sequences via distinctions between fracture stimulation and fault reactivation (Figure 13).
 866 While this appears to capture the first-order effects, there likely exists subtleties that can deviate
 867 from this conjecture.

868

869 *4.3.1 Exceptions to the bound-fracture interpretation*

870 We have predominantly interpreted bound sequences arising from the finite (and
 871 time/volume dependent) extent of a propagating fracture restricting available slip area. While
 872 this interpretation matches both expectation and observation, we can still envision a scenario
 873 in which a finite fracture could produce an unbound sequence. For example, in the case that
 874 the finite extent of the fracture grows faster than the equivalent area of an unbound M_{LRG} . In
 875 this sense, while the sequence is technically bound, it would never be empirically inferable by
 876 magnitude information. Criteria for this condition are given below: for the V^n generalization
 877 of M_{MAX} as well as the tensile-like and shear-like fractures. There, the expected M_{LRG}
 878 formulation is a combination of the population expectation (van der Elst et al., 2016) coupled
 879 with a proportionality between injected volume and event counts (Shapiro et al., 2010). Note
 880 that Σ is the seismogenic index, a parameter indicating a fault’s potential to induce earthquakes
 881 per unit of injected fluid volume (Shapiro et al., 2010).

$$882 \quad \left(\frac{2n}{3} - \frac{1}{b}\right) \log_{10}(V(t)) + \left(\frac{2}{3}\right) \log_{10}(c) - 9.1 \gg \frac{\Sigma}{b} \quad (V^n \text{ generalization})$$

$$883 \quad \left(1 - \frac{1}{b}\right) \log_{10}(V(t)) + \left(\frac{2}{3}\right) \log_{10}\left(\frac{16}{7} \Delta\sigma s_d^{3/2}\right) - 9.1 \gg \frac{\Sigma}{b} \quad (\text{tensile fracture})$$

$$884 \quad \left(\frac{2}{3} - \frac{1}{b}\right) \log_{10}(V(t)) + \left(\frac{2}{3}\right) \log_{10}\left(\frac{16}{7} \Delta\sigma k\right) - 9.1 \gg \frac{\Sigma}{b} \quad (\text{Shear fracture})$$

885 As a reminder of terms: $V(t)$ is the injection volume time-series, n is the volume exponent, b
 886 is the b -value, $\Delta\sigma$ is the stress drop, while c , s_d and k are injection propagation coefficients
 887 for different growth.

888

889 *4.3.2 Exceptions to the unbound-fault interpretation*

890 We have predominantly interpreted unbound sequences arising from fractures
891 interacting with relatively-large pre-existing faults. While this matches our expectation and
892 observations, we can still envision scenarios where pre-existing faults could produce bound
893 sequences. For example, M_{LRG} could be limited by the pre-existing faults being too small, the
894 *in situ* stress resolved on the fault (Gischig, 2015; Norbeck & Horne, 2018), time-dependence
895 of stress perturbations (Segall & Lu, 2015), fault geometry/complexity (Lee et al., 2024), fault
896 roughness (Mauer et al., 2020; Wang et al., 2024), material/stress heterogeneity (Kroll &
897 Cochran, 2021), rheology, or the extent of asperities. In fact, HS4 at the GTS is one such
898 exception: where streaks outlined the rheological boundary between seismic asperities hosted
899 within an aseismic fault. If these factors are limiting M_{LRG} , CAP-tests should be able to discern
900 their effects.

901

902 **5. Conclusions**

903 In summary, we have rigorously tested CAP-tests against data from controlled injection
904 experiments at three underground laboratories. Bound sequences are consistent with
905 observations of new hydraulic fracturing growth, while unbound typically reactivate larger pre-
906 existing structures. Furthermore, the EW-test appears to be able aggregate generalized M_{MAX}
907 functions into categories consistent with theoretical considerations for tectonic/geometric
908 limits, tensile/McGarr-like growth, or aseismic/shear-like growth. This process is potentially
909 able to identify the underlying physical mechanism responsible for M_{MAX} , specific to an
910 individual earthquake sequence. Overall, the ability to robustly identify and constrain M_{MAX}
911 will likely be important for both induced seismicity and natural/tectonic earthquake sequences.

912

913 **Acknowledgments**

914 We would like to thank Luigi Passarelli for insightful conversations that helped the
915 interpretation of this study. We also thank Peter Niemz and an anonymous reviewer, whose
916 comments helped to improve this manuscript. We are grateful for the high-quality UGL studies
917 that facilitated our CAP-test interpretations, which would have been impossible otherwise.
918 This work is supported by the Seismogenic Fault Injection Test (SFIT), which is funded by the
919 Swiss National Science Foundation, under project number TMPFP2_224393. This work is
920 also supported by the CETPartnership, which is funded by the Clean Energy Transition
921 Partnership under the 2023 joint call for research proposals, co-funded by the European
922 Commission (GA N°101069750) and with the funding organizations detailed on
923 <https://cetpartnership.eu/funding-agencies-and-call-modules>.

924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941

Data Availability

The codes (and data) used to derive our results are available online at GitHub (**LINK**). The catalogue and hydraulic datasets are also available online, from their source: for the Äspö HRL (Zang et al., 2024), SURF EGS Collab Experiment #1 (<https://gdr.openei.org/>), and the GTS (<https://doi.org/10.3929/ethz-b-000276170>).

Author Contributions

R.S. created the CAP-tests, gathered the data, analysed the cases, and wrote the manuscript. L.V., and V.G. were involved in the curation of GTS catalogue and interpretation of the results. S.W. provided project management. All authors were involved in manuscript editing and review.

Competing Interests

The authors declare no competing interests.

942 **References**

- 943
- 944 Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle.
945 In *Selected papers of hirotugu akaike* (pp. 199-213). New York, NY: Springer New York.
- 946
- 947 Amann, F., Gischig, V., Evans, K., Doetsch, J., Jalali, R., Valley, B., ... & Giardini, D.
948 (2018). The seismo-hydromechanical behavior during deep geothermal reservoir
949 stimulations: open questions tackled in a decameter-scale in situ stimulation experiment.
950 *Solid Earth*, 9(1), 115-137, <https://doi.org/10.5194/se-9-115-2018>.
- 951
- 952 Atkinson, G. M., Eaton, D. W., Ghofrani, H., Walker, D., Cheadle, B., Schultz, R., ... & Kao,
953 H. (2016). Hydraulic Fracturing and Seismicity in the Western Canada Sedimentary Basin.
954 *Seismological Research Letters*, 87(3), 631-647, <https://doi.org/10.1785/0220150263>.
- 955
- 956 Berger, V. W., & Zhou, Y. (2014). Kolmogorov–Smirnov test: Overview. *Wiley statsref:*
957 *Statistics reference online*, <https://doi.org/10.1002/9781118445112.stat06558>.
- 958
- 959 Bommer, J. J., Oates, S., Cepeda, J. M., Lindholm, C., Bird, J., Torres, R., ... & Rivas, J.
960 (2006). Control of hazard due to seismicity induced by a hot fractured rock geothermal
961 project. *Engineering Geology*, 83(4), 287-306, <https://doi.org/10.1016/j.enggeo.2005.11.002>.
- 962
- 963 Bommer, J. J. (2022). Earthquake hazard and risk analysis for natural and induced seismicity:
964 towards objective assessments in the face of uncertainty. *Bulletin of Earthquake Engineering*,
965 20(6), 2825-3069, <https://doi.org/10.1007/s10518-022-01357-4>.
- 966
- 967 Bommer, J., & Verdon, J. P. (2024). The maximum magnitude of natural and induced
968 earthquakes. *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, 10, 172,
969 <https://doi.org/10.1007/s40948-024-00895-2>.
- 970
- 971 Brixel, B., Klepikova, M., Jalali, M. R., Lei, Q., Roques, C., Kriestch, H., & Loew, S.
972 (2020a). Tracking fluid flow in shallow crustal fault zones: 1. Insights from single-hole
973 permeability estimates. *Journal of Geophysical Research: Solid Earth*, 125(4),
974 e2019JB018200, <https://doi.org/10.1029/2019JB018200>.
- 975
- 976 Brixel, B., Klepikova, M., Lei, Q., Roques, C., Jalali, M. R., Krietsch, H., & Loew, S.
977 (2020b). Tracking fluid flow in shallow crustal fault zones: 2. Insights from cross-hole forced
978 flow experiments in damage zones. *Journal of Geophysical Research: Solid Earth*, 125(4),
979 e2019JB019108, <https://doi.org/10.1029/2019JB019108>.
- 980
- 981 Cao, N. T., Eisner, L., & Jechumtálová, Z. (2020). Next record breaking magnitude for
982 injection induced seismicity. *First Break*, 38(2), 53-57, <https://doi.org/10.3997/1365-2397.fb2020010>.
- 983
- 984
- 985 Cao, N. T., Eisner, L., Jechumtálová, Z., Verdon, J., & Waheed, U. B. (2024). Upper limit
986 magnitudes for induced seismicity in energy industries. *Geophysical Prospecting*,
987 <https://doi.org/10.1111/1365-2478.13553>.
- 988
- 989 Chai, C., Maceira, M., Santos-Villalobos, H. J., Venkatakrishnan, S. V., Schoenball, M., Zhu,
990 W., ... & EGS Collab Team. (2020). Using a deep neural network and transfer learning to

991 bridge scales for seismic phase picking. *Geophysical Research Letters*, 47(16),
992 e2020GL088651, <https://doi.org/10.1029/2020GL088651>.
993

994 Clarke, H., Verdon, J. P., Kettlety, T., Baird, A. F., & Kendall, J. M. (2019). Real-time
995 imaging, forecasting, and management of human-induced seismicity at Preston New Road,
996 Lancashire, England. *Seismological Research Letters*, 90(5), 1902-1915,
997 <https://doi.org/10.1785/0220190110>.
998

999 Clarkson, C.R., Alkhayyali W., & Zeinabady, D. (2025). Characterization and modeling of
1000 enhanced geothermal systems using methods developed for unconventional hydrocarbon
1001 reservoirs, *SPE Annual Technical Conference and Exhibition*, SPE-228230.
1002

1003 Danré, P., Garagash, D., De Barros, L., Cappa, F., & Ampuero, J. P. (2024). Control of
1004 seismicity migration in earthquake swarms by injected fluid volume and aseismic crack
1005 propagation. *Journal of Geophysical Research: Solid Earth*, 129(1), e2023JB027276,
1006 <https://doi.org/10.1029/2023JB027276>.
1007

1008 Davis, T., Rivalta, E., & Dahm, T. (2020). Critical fluid injection volumes for uncontrolled
1009 fracture ascent. *Geophysical Research Letters*, 47(14), e2020GL087774,
1010 <https://doi.org/10.1029/2020GL087774>.
1011

1012 DeDontney, N., Gans, C., Burnett, W., Burch, D., Garzon, J., Gist, G., ... & Younan, A.
1013 (2016). Maximum magnitude of induced earthquakes in the Groningen gas field. *ExxonMobil*
1014 *URC External Report*, [https://nam-onderzoeksrapporten.data-](https://nam-onderzoeksrapporten.data-app.nl/reports/download/groningen/en/36dea690-c7eb-4164-9d94-d9725381c426)
1015 [app.nl/reports/download/groningen/en/36dea690-c7eb-4164-9d94-d9725381c426](https://nam-onderzoeksrapporten.data-app.nl/reports/download/groningen/en/36dea690-c7eb-4164-9d94-d9725381c426).
1016

1017 Dobson, P., Kneafsey, T., Morris, J., Singh, A., Zoback, M., Roggenthen, W., ... & White, M.
1018 (2018). The EGS Collab hydroshear experiment at the Sanford Underground Research
1019 Facility—Siting criteria and evaluation of candidate sites. *Geothermal Resources Council*
1020 *Transactions*, 42, 708-723.
1021

1022 Doetsch, J., Gischig, V., Krietsch, H., Villiger, L., Amann, F., Dutler, N., Jalali, M., Brixel,
1023 B., Roques, C., Giertzuch, P., Kittilä, A., and Hochreutener, R. (2018a). Grimsel ISC
1024 Experimental Description, *ETH Zurich Report*, <https://doi.org/10.3929/ethz-b-000310581>.
1025

1026 Doetsch, J., Gischig, V. S., Villiger, L., Krietsch, H., Nejati, M., Amann, F., ... & Giardini, D.
1027 (2018b). Subsurface fluid pressure and rock deformation monitoring using seismic velocity
1028 observations. *Geophysical Research Letters*, 45(19), 10-389,
1029 <https://doi.org/10.1029/2018GL079009>.
1030

1031 Dorbath, L., Cuenot, N., Genter, A., & Frogneux, M. (2009). Seismic response of the
1032 fractured and faulted granite of Soultz-sous-Forêts (France) to 5 km deep massive water
1033 injections. *Geophysical Journal International*, 177(2), 653-675,
1034 <https://doi.org/10.1111/j.1365-246X.2009.04030.x>.
1035

1036 Dutler, N., Valley, B., Gischig, V., Villiger, L., Krietsch, H., Doetsch, J., ... & Amann, F.
1037 (2019). Hydraulic fracture propagation in a heterogeneous stress field in a crystalline rock
1038 mass. *Solid Earth*, 10(6), 1877-1904, <https://doi.org/10.5194/se-10-1877-2019>.
1039

1040 Eaton, D. W., & Igonin, N. (2018). What controls the maximum magnitude of injection-
1041 induced earthquakes?. *The Leading Edge*, 37(2), 135-140,
1042 <https://doi.org/10.1190/tle37020135.1>.
1043

1044 Elsworth, D., Li, Z., Yu, P., An, M., Zhang, F., Huang, R., ... & Liu, S. (2025). Constraints
1045 on triggered seismicity and its control on permeability evolution. *Journal of Rock Mechanics
1046 and Geotechnical Engineering*, 17(1), 20-30, <https://doi.org/10.1016/j.jrmge.2024.11.035>.
1047

1048 Evans, K. F., Moriya, H., Niitsuma, H., Jones, R. H., Phillips, W. S., Genter, A., ... & Baria,
1049 R. (2005). Microseismicity and permeability enhancement of hydrogeologic structures during
1050 massive fluid injections into granite at 3 km depth at the Soultz HDR site. *Geophysical
1051 Journal International*, 160(1), 388-412, <https://doi.org/10.1111/j.1365-246X.2004.02474.x>.
1052

1053 Foulger, G. R., Wilson, M. P., Gluyas, J. G., Julian, B. R., & Davies, R. J. (2018). Global
1054 review of human-induced earthquakes. *Earth-Science Reviews*, 178, 438-514,
1055 <https://doi.org/10.1016/j.earscirev.2017.07.008>.
1056

1057 Frash, L. P., Carey, J. W., & Welch, N. J. (2019, February). EGS collab experiment 1
1058 geomechanical and hydrological properties by triaxial direct shear. In *44th workshop on
1059 geothermal reservoir engineering*.
1060

1061 Fu, P., Schoenball, M., Ajo-Franklin, J. B., Chai, C., Maceira, M., Morris, J. P., ... & EGS
1062 Collab Team. (2021). Close observation of hydraulic fracturing at EGS Collab Experiment 1:
1063 Fracture trajectory, microseismic interpretations, and the role of natural fractures. *Journal of
1064 Geophysical Research: Solid Earth*, 126(7), e2020JB020840,
1065 <https://doi.org/10.1029/2020JB020840>.
1066

1067 Galis, M., Ampuero, J. P., Mai, P. M., & Cappa, F. (2017). Induced seismicity provides
1068 insight into why earthquake ruptures stop. *Science Advances*, 3(12), eaap7528,
1069 <https://doi.org/10.1126/sciadv.aap7528>.
1070

1071 Grigoli, F., Cesca, S., Rinaldi, A. P., Manconi, A., Lopez-Comino, J. A., Clinton, J. F., ... &
1072 Wiemer, S. (2018). The November 2017 Mw 5.5 Pohang earthquake: A possible case of
1073 induced seismicity in South Korea. *Science*, 360(6392), 1003-1006,
1074 <https://doi.org/10.1126/science.aat2010>.
1075

1076 Gischig, V. S. (2015). Rupture propagation behavior and the largest possible earthquake
1077 induced by fluid injection into deep reservoirs. *Geophysical Research Letters*, 42(18), 7420-
1078 7428, <https://doi.org/10.1002/2015GL065072>.
1079

1080 Gischig, V., Jalali, R., Amann, F., Krietsch, H., Klepikova, M., Esposito, S., ... & Madonna,
1081 C. (2016). Impact of the ISC Experiment at the Grimsel Test Site-Assessment of Potential
1082 Seismic Hazard and Disturbances to Nearby Experiments and KWO Infrastructure. *ETH
1083 Zurich Report*, <https://doi.org/10.3929/ethz-b-000189973>.
1084

1085 Gischig, V. S., Giardini, D., Amann, F., Hertrich, M., Krietsch, H., Loew, S., ... & Valley, B.
1086 (2020). Hydraulic stimulation and fluid circulation experiments in underground laboratories:
1087 Stepping up the scale towards engineered geothermal systems. *Geomechanics for Energy and
1088 the Environment*, 24, 100175, <https://doi.org/10.1016/j.gete.2019.100175>.
1089

- 1090 Guglielmi, Y., Cook, P., Soom, F., Schoenball, M., Dobson, P., & Kneafsey, T. (2021). In
1091 situ continuous monitoring of borehole displacements induced by stimulated hydrofracture
1092 growth. *Geophysical Research Letters*, 48(4), e2020GL090782,
1093 <https://doi.org/10.1029/2020GL090782>.
- 1094
1095 Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of*
1096 *the Seismological Society of America*, 34(4), 185-188,
1097 <https://doi.org/10.1785/BSSA0340040185>.
- 1098
1099 Hanks, T. C., & Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical*
1100 *Research: Solid Earth*, 84(B5), 2348-2350, <https://doi.org/10.1029/JB084iB05p02348>.
- 1101
1102 Hallo, M., Oprsäl, I., Eisner, L., & Ali, M. Y. (2014). Prediction of magnitude of the largest
1103 potentially induced seismic event. *Journal of Seismology*, 18, 421-431,
1104 <https://doi.org/10.1007/s10950-014-9417-4>.
- 1105
1106 Heise, J. (2015). The Sanford underground research facility at Homestake. In *Journal of*
1107 *Physics: Conference Series* (Vol. 606, No. 1, p. 012015). IOP Publishing,
1108 <https://doi.org/10.1088/1742-6596/606/1/012015>.
- 1109
1110 Holschneider, M., Zöller, G., & Hainzl, S. (2011). Estimation of the maximum possible
1111 magnitude in the framework of a doubly truncated Gutenberg–Richter model. *Bulletin of the*
1112 *Seismological Society of America*, 101(4), 1649-1659, <https://doi.org/10.1785/0120100289>.
- 1113
1114 Holschneider, M., Zöller, G., Clements, R., & Schorlemmer, D. (2014). Can we test for the
1115 maximum possible earthquake magnitude?. *Journal of Geophysical Research: Solid Earth*,
1116 119(3), 2019-2028, <https://doi.org/10.1002/2013JB010319>.
- 1117
1118 Im, K., & Avouac, J. P. (2025). Maximum Magnitude of Induced Earthquakes in Rate and
1119 State Friction Framework. *Seismological Research Letters*, 96(3), 1654-1664,
1120 <https://doi.org/10.1785/0220240382>.
- 1121
1122 Ishimoto, M., & Iida, K. (1939). Observations of earthquakes registered with the
1123 microseismograph constructed recently. *Bull. Earthq. Res. Inst. Univ. Tokyo*, 17, 443-478.
- 1124
1125 Jalali, M., Gischig, V., Doetsch, J., Näf, R., Krietsch, H., Klepikova, M., ... & Giardini, D.
1126 (2018). Transmissivity changes and microseismicity induced by small-scale hydraulic
1127 fracturing tests in crystalline rock. *Geophysical Research Letters*, 45(5), 2265-2273,
1128 <https://doi.org/10.1002/2017GL076781>.
- 1129
1130 Jin, G., Ning, Y., Gale, M., Simmons, J., & Tura, A. (2024). Impact of natural fractures on
1131 hydraulic fracture propagation in Denver-Julesburg Basin: Insights from a decade of research.
1132 *The Leading Edge*, 43(12), 806-814, <https://doi.org/10.1190/tle43120806.1>.
- 1133
1134 Kanamori, H., & Anderson, D. L. (1975). Theoretical basis of some empirical relations in
1135 seismology. *Bulletin of the Seismological Society of America*, 65(5), 1073-1095,
1136 <https://doi.org/10.1785/BSSA0650051073>.
- 1137
1138 Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical*
1139 *Association*, 90(430), 773-795, <https://doi.org/10.1080/01621459.1995.10476572>.

1140
1141 Kettlety, T., Verdon, J. P., Butcher, A., Hampson, M., & Craddock, L. (2021). High-
1142 resolution imaging of the ML 2.9 August 2019 earthquake in Lancashire, United Kingdom,
1143 induced by hydraulic fracturing during Preston New Road PNR-2 operations. *Seismological*
1144 *Research Letters*, 92(1), 151-169, <https://doi.org/10.1785/0220200187>.
1145
1146 Kijko, A. (2025). Bayesian Assessment of the Maximum Possible Earthquake Magnitude m
1147 max. *Journal of the Geological Society of India*, 101(6), 764-769,
1148 <https://doi.org/10.17491/jgsi/2025/174157>.
1149
1150 Krietsch, H., Doetsch, J., Dutler, N., Jalali, M., Gischig, V., Loew, S., & Amann, F. (2018).
1151 Comprehensive geological dataset describing a crystalline rock mass for hydraulic
1152 stimulation experiments. *Scientific data*, 5(1), 1-12, <https://doi.org/10.1038/sdata.2018.269>.
1153
1154 Krietsch, H., Gischig, V., Evans, K., Doetsch, J., Dutler, N. O., Valley, B., & Amann, F.
1155 (2019). Stress measurements for an in situ stimulation experiment in crystalline rock:
1156 integration of induced seismicity, stress relief and hydraulic methods. *Rock Mechanics and*
1157 *Rock Engineering*, 52, 517-542, <https://doi.org/10.1007/s00603-018-1597-8>.
1158
1159 Krietsch, H., Villiger, L., Doetsch, J., Gischig, V., Evans, K. F., Brixel, B., ... & Amann, F.
1160 (2020a). Changing flow paths caused by simultaneous shearing and fracturing observed
1161 during hydraulic stimulation. *Geophysical Research Letters*, 47(3), e2019GL086135,
1162 <https://doi.org/10.1029/2019GL086135>.
1163
1164 Krietsch, H., Gischig, V. S., Doetsch, J., Evans, K. F., Villiger, L., Jalali, M., ... & Amann, F.
1165 (2020b). Hydromechanical processes and their influence on the stimulation effected volume:
1166 observations from a decameter-scale hydraulic stimulation project. *Solid Earth*, 11(5), 1699-
1167 1729, <https://doi.org/10.5194/se-11-1699-2020>.
1168
1169 Kroll, K. A., & Cochran, E. S. (2021). Stress controls rupture extent and maximum
1170 magnitude of induced earthquakes. *Geophysical Research Letters*, 48(11), e2020GL092148,
1171 <https://doi.org/10.1029/2020GL092148>.
1172
1173 Kneafsey, T. J., Dobson, P., Blankenship, D., Morris, J., Knox, H., Schwering, P., ... &
1174 Valladao, C. (2018). An overview of the EGS Collab project: field validation of coupled
1175 process modeling of fracturing and fluid flow at the Sanford Underground Research Facility,
1176 Lead, SD. In *43rd Workshop on Geothermal Reservoir Engineering* (Vol. 2018).
1177
1178 Kneafsey, T. J., Blankenship, D., Dobson, P. F., Morris, J. P., White, M. D., Fu, P., ... &
1179 Valladao, C. (2020). The EGS collab project: Learnings from Experiment 1. In *Proceedings*
1180 *of the 45th Workshop on Geothermal Reservoir Engineering* (pp. 10-12). Stanford, CA:
1181 Stanford University.
1182
1183 Kwiatek, G., Martínez-Garzón, P., Plenkers, K., Leonhardt, M., Zang, A., von Specht, S., ...
1184 & Bohnhoff, M. (2018). Insights into complex subdecimeter fracturing processes occurring
1185 during a water injection experiment at depth in Äspö Hard Rock Laboratory, Sweden.
1186 *Journal of Geophysical Research: Solid Earth*, 123(8), 6616-6635,
1187 <https://doi.org/10.1029/2017JB014715>.
1188

1189 Langenbruch, C., Moein, M. J., & Shapiro, S. A. (2024). Are maximum magnitudes of
1190 induced earthquakes controlled by pressure diffusion?. *Philosophical Transactions A*,
1191 382(2276), 20230184, <https://doi.org/10.1098/rsta.2023.0184>.
1192

1193 Lanza, F., Rinaldi, A.P., Passarelli, L., Ritz, V.A., Repollés, V.C., Schultz, R., Ciardo, F.,
1194 Dyer, B., Ermert, L., Grigoratos, I., Karvounis, D., Meier, P., Mignan, A., Moore, J. Pankow,
1195 K., Scarabello, L., Schmid, N., Shi, P., Tuinstra, K., & Wiemer, S. (2026). The 2022
1196 hydraulic stimulation at Utah FORGE: investigating fracturing mechanisms and testing
1197 forecasting approaches. Submitted to *Seismica*.
1198

1199 Lee, J., Tsai, V. C., Hirth, G., Chatterjee, A., & Trugman, D. T. (2024). Fault-network
1200 geometry influences earthquake frictional behaviour. *Nature*, 1-5,
1201 <https://doi.org/10.1038/s41586-024-07518-6>.
1202

1203 Li, Z., Elsworth, D., & Wang, C. (2021). Constraining maximum event magnitude during
1204 injection-triggered seismicity. *Nature Communications*, 12(1), 1528,
1205 <https://doi.org/10.1038/s41467-020-20700-4>.
1206

1207 Li, S., & Zhang, D. (2023). Three-dimensional thermoporoelastic modeling of
1208 hydrofracturing and fluid circulation in hot dry rock. *Journal of Geophysical Research: Solid*
1209 *Earth*, 128(2), e2022JB025673, <https://doi.org/10.1029/2022JB025673>.
1210

1211 López-Comino, J. Á., Cesca, S., Heimann, S., Grigoli, F., Milkereit, C., Dahm, T., & Zang,
1212 A. (2017). Characterization of hydraulic fractures growth during the Äspö Hard Rock
1213 Laboratory experiment (Sweden). *Rock Mechanics and Rock Engineering*, 50, 2985-3001,
1214 <https://doi.org/10.1007/s00603-017-1285-0>.
1215

1216 López-Comino, J. Á., Cesca, S., Niemz, P., Dahm, T., & Zang, A. (2021). Rupture directivity
1217 in 3D inferred from acoustic emissions events in a mine-scale hydraulic fracturing
1218 experiment. *Frontiers in Earth Science*, 9, 670757,
1219 <https://doi.org/10.3389/feart.2021.670757>.
1220

1221 Majer, E. L., Baria, R., Stark, M., Oates, S., Bommer, J., Smith, B., & Asanuma, H. (2007).
1222 Induced seismicity associated with enhanced geothermal systems. *Geothermics*, 36(3), 185-
1223 222, <https://doi.org/10.1016/j.geothermics.2007.03.003>.
1224

1225 Marzocchi, W., & Sandri, L. (2003). A review and new insights on the estimation of the b-
1226 value and its uncertainty. *Annals of Geophysics*, 46(6), 1271-1282.
1227

1228 Maurer, J., Dunham, E. M., & Segall, P. (2020). Role of fluid injection on earthquake size in
1229 dynamic rupture simulations on rough faults. *Geophysical Research Letters*, 47(13),
1230 e2020GL088377, <https://doi.org/10.1029/2020GL088377>.
1231

1232 McGarr, A. (2014). Maximum magnitude earthquakes induced by fluid injection. *Journal of*
1233 *Geophysical Research: solid earth*, 119(2), 1008-1019,
1234 <https://doi.org/10.1002/2013JB010597>.
1235

1236 McQuarrie, A. D. (1999). A small-sample correction for the Schwarz SIC model selection
1237 criterion. *Statistics & Probability Letters*, 44(1), 79-86, [https://doi.org/10.1016/S0167-7152\(98\)00294-6](https://doi.org/10.1016/S0167-7152(98)00294-6).
1238

1239
1240 Mendecki, A.J., 2016. Mine Seismology Reference Book: Seismic Hazard: Institute of Mine
1241 Seismology, Tasmania, Australia. p 88. ISBN 978-0-9942943-0-2.
1242
1243 Moein, M. J., Langenbruch, C., Schultz, R., Grigoli, F., Ellsworth, W. L., Wang, R., ... &
1244 Shapiro, S. (2023). The physical mechanisms of induced earthquakes. *Nature Reviews Earth
1245 & Environment*, 4(12), 847-863, <https://doi.org/10.1038/s43017-023-00497-8>.
1246
1247 Moore, J., McLennan, J., Allis, R., Pankow, K., Simmons, S., Podgorney, R., ... & Rickard,
1248 W. (2019). The Utah Frontier Observatory for Research in Geothermal Energy (FORGE): an
1249 international laboratory for enhanced geothermal system technology development. In *44th
1250 Workshop on Geothermal Reservoir Engineering* (pp. 11-13). Stanford University.
1251
1252 Morris, J. P., Fu, P., Dobson, P., Ajo-Franklin, J., Kneafsey, T. J., Knox, H., ... & EGS
1253 Collab Team, _ . (2018). Experimental design for hydrofracturing and fluid flow at the DOE
1254 EGS collab testbed. In *ARMA US Rock Mechanics/Geomechanics Symposium* (pp. ARMA-
1255 2018). ARMA.
1256
1257 Muntendam-Bos, A. G., Hoedeman, G., Polychronopoulou, K., Draganov, D., Weemstra, C.,
1258 van der Zee, W., ... & Roest, H. (2022). An overview of induced seismicity in the
1259 Netherlands. *Netherlands Journal of Geosciences*, 101, e1,
1260 <https://doi.org/10.1017/njg.2021.14>.
1261
1262 Niemz, P., Cesca, S., Heimann, S., Grigoli, F., von Specht, S., Hammer, C., ... & Dahm, T.
1263 (2020). Full-waveform-based characterization of acoustic emission activity in a mine-scale
1264 experiment: a comparison of conventional and advanced hydraulic fracturing schemes.
1265 *Geophysical Journal International*, 222(1), 189-206, <https://doi.org/10.1093/gji/ggaa127>.
1266
1267 Niemz, P., Dahm, T., Milkereit, C., Cesca, S., Petersen, G., & Zang, A. (2021). Insights into
1268 hydraulic fracture growth gained from a joint analysis of seismometer-derived tilt signals and
1269 acoustic emissions. *Journal of Geophysical Research: Solid Earth*, 126(12), e2021JB023057,
1270 <https://doi.org/10.1029/2021JB023057>.
1271
1272 Niemz, P., Pankow, K., Isken, M.P., Whidden, K., McLennan, J., & Moore, J. (2025).
1273 Mapping fracture zones with nodal geophone patches: Insights from induced microseismicity
1274 during the 2024 stimulations at Utah FORGE. *Seismological Research Letters*,
1275 <https://doi.org/10.1785/0220240300>.
1276
1277 Niemz, P., Petersen, G., Rutledge, J., Whidden, K., & Pankow, K. (2026). Isotropic
1278 components of microseismic moment tensors at Utah FORGE reveal a diversity of fluid
1279 pathway creation processes in EGS development. *Scientific Reports*,
1280 <https://doi.org/10.1038/s41598-026-42493-0>.
1281
1282 Norbeck, J. H., & Horne, R. N. (2018). Maximum magnitude of injection-induced
1283 earthquakes: A criterion to assess the influence of pressure migration along faults.
1284 *Tectonophysics*, 733, 108-118, <https://doi.org/10.1016/j.tecto.2018.01.028>.
1285
1286 Oldenburg, C., Dobson, P., Wu, Y., Cook, P., Kneafsey, T., Nakagawa, S., ... & Heise, J.
1287 (2017). Hydraulic fracturing experiments at 1500 m depth in a deep mine: Highlights from

1288 the KISMET project, in *42nd Workshop on Geothermal Reservoir Engineering*, p. 9, Stanford
1289 University.

1290

1291 Pisarenko, V. F., & Rodkin, M. V. (2022). Approaches to solving the maximum possible
1292 earthquake magnitude (Mmax) problem. *Surveys in Geophysics*, 1-35, doi:
1293 <https://doi.org/10.1007/s10712-021-09673-1>.

1294

1295 Qin, Y., Li, J., Huang, L., Schoenball, M., Ajo-Franklin, J., Blankenship, D., ... & EGS
1296 Collab Team. (2024). Source mechanism of kHz microseismic events recorded in multiple
1297 boreholes at the first EGS Collab testbed. *Geothermics*, 120, 102994,
1298 <https://doi.org/10.1016/j.geothermics.2024.102994>.

1299

1300 Rubin, A. M., Gillard, D., & Got, J. L. (1999). Streaks of microearthquakes along creeping
1301 faults. *Nature*, 400(6745), 635-641, <https://doi.org/10.1038/23196>.

1302

1303 Rutledge, J. T., Phillips, W. S., & Mayerhofer, M. J. (2004). Faulting induced by forced fluid
1304 injection and fluid flow forced by faulting: An interpretation of hydraulic-fracture
1305 microseismicity, Carthage Cotton Valley gas field, Texas. *Bulletin of the Seismological
1306 Society of America*, 94(5), 1817-1830, <https://doi.org/10.1785/012003257>.

1307

1308 Sáez, A., Passelègue, F., & Lecampion, B. (2025). Maximum size and magnitude of
1309 injection-induced slow slip events. *Science Advances*, 11(19), eadq0662,
1310 <https://doi.org/10.1126/sciadv.adq0662>.

1311

1312 Schneeberger, R., Kober, F., Spillmann, T., Blechschmidt, I., Lanyon, G. W., & Mäder, U. K.
1313 (2019). Grimsel Test Site: Revisiting the site-specific geoscientific knowledge. *Nagra
1314 Technical Report 19-01*. <https://nagra.ch/en/downloads/technical-report-ntb-19-01-2/>.

1315

1316 Schoenball, M., Ajo-Franklin, J., Fu, P., & Templeton, D. (2019). *Microseismic monitoring
1317 of meso-scale stimulations for the DOE EGS Collab project at the Sanford Underground
1318 Research Facility* (No. LLNL-CONF-767025). Lawrence Livermore National Lab.(LLNL),
1319 Livermore, CA (United States).

1320

1321 Schoenball, M., Ajo-Franklin, J. B., Blankenship, D., Chai, C., Chakravarty, A., Dobson, P.,
1322 ... & EGS Collab Team. (2020). Creation of a mixed-mode fracture network at mesoscale
1323 through hydraulic fracturing and shear stimulation. *Journal of Geophysical Research: Solid
1324 Earth*, 125(12), e2020JB019807, <https://doi.org/10.1029/2020JB019807>.

1325

1326 Schopper, F., Doetsch, J., Villiger, L., Krietsch, H., Gischig, V. S., Jalali, M., ... & Maurer,
1327 H. (2020). On the variability of pressure propagation during hydraulic stimulation based on
1328 seismic velocity observations. *Journal of Geophysical Research: Solid Earth*, 125(2),
1329 e2019JB018801, <https://doi.org/10.1029/2019JB018801>.

1330

1331 Schultz, R., Atkinson, G., Eaton, D. W., Gu, Y. J., & Kao, H. (2018). Hydraulic fracturing
1332 volume is associated with induced earthquake productivity in the Duvernay play. *Science*,
1333 359(6373), 304-308, <https://doi.org/10.1126/science.aao0159>.

1334

1335 Schultz, R., Skoumal, R. J., Brudzinski, M. R., Eaton, D., Baptie, B., & Ellsworth, W. (2020).
1336 Hydraulic fracturing-induced seismicity. *Reviews of Geophysics*, 58(3), e2019RG000695,
1337 <https://doi.org/10.1029/2019RG000695>.

1338
1339 Schultz, R., Beroza, G. C., & Ellsworth, W. L. (2021a). A risk-based approach for managing
1340 hydraulic fracturing–induced seismicity. *Science*, 372(6541), 504-507,
1341 <https://doi.org/10.1126/science.abg5451>.
1342
1343 Schultz, R., Beroza, G. C., & Ellsworth, W. L. (2021b). A strategy for choosing red-light
1344 thresholds to manage hydraulic fracturing induced seismicity in North America. *Journal of*
1345 *Geophysical Research: Solid Earth*, 126(12), e2021JB022340,
1346 <https://doi.org/10.1029/2021JB022340>.
1347
1348 Schultz, R., Ellsworth, W. L., & Beroza, G. C. (2022). Statistical bounds on how induced
1349 seismicity stops. *Scientific Reports*, 12(1), 1184, [https://doi.org/10.1038/s41598-022-05216-](https://doi.org/10.1038/s41598-022-05216-9)
1350 [9](https://doi.org/10.1038/s41598-022-05216-9).
1351
1352 Schultz, R., Baptie, B., Edwards, B., & Wiemer, S. (2023a). Red-light thresholds for induced
1353 seismicity in the UK. *Seismica*, 2(2), <https://doi.org/10.26443/seismica.v2i2.1086>.
1354
1355 Schultz, R., Park, Y., Aguilar Suarez, A. L., Ellsworth, W. L., & Beroza, G. C. (2023b). En
1356 echelon faults reactivated by wastewater disposal near Musreau Lake, Alberta. *Geophysical*
1357 *Journal International*, 235(1), 417-429, <https://doi.org/10.1093/gji/ggad226>.
1358
1359 Schultz, R., Ellsworth, W. L., & Beroza, G. C. (2023c). An ensemble approach to
1360 characterizing trailing-induced seismicity. *Seismological Research Letters*, 94(2A), 699-707,
1361 <https://doi.org/10.1785/0220220352>.
1362
1363 Schultz, R. (2024). Inferring maximum magnitudes from the ordered sequence of large
1364 earthquakes. *Philosophical Transactions A*, 382(2276), 20230185,
1365 <https://doi.org/10.1098/rsta.2023.0185>.
1366
1367 Schultz, R. (2025). Reining-in the spring-slider with reinforcement learning. *Journal of*
1368 *Geophysical Research: Solid Earth*, 130(3), e2024JB029697,
1369 <https://doi.org/10.1029/2024JB029697>.
1370
1371 Schultz, R., Lanza, F., Dyer, B., Karvounis, D., Fiori, R., Shi, P., Ritz, V., Villiger, L., Meier,
1372 P., & Wiemer, S. (2025). The bound growth of induced earthquakes could de-risk hydraulic
1373 fracturing, *Communications Earth & Environment*, [https://doi.org/10.1038/s43247-025-](https://doi.org/10.1038/s43247-025-02881-2)
1374 [02881-2](https://doi.org/10.1038/s43247-025-02881-2).
1375
1376 Schultz, R. (2026). Improving the resolvability of M_{MAX} truncation via deeper order statistics.
1377 *Geophysical Journal International*, ggag110, <https://doi.org/10.1093/gji/ggag110>.
1378
1379 Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 461-464.
1380
1381 Segall, P., & Lu, S. (2015). Injection-induced seismicity: Poroelastic and earthquake
1382 nucleation effects. *Journal of Geophysical Research: Solid Earth*, 120(7), 5082-5103,
1383 <https://doi.org/10.1002/2015JB012060>.
1384
1385 Shapiro, S. A., Dinske, C., Langenbruch, C., & Wenzel, F. (2010). Seismogenic index and
1386 magnitude probability of earthquakes induced during reservoir fluid stimulations. *The*
1387 *Leading Edge*, 29(3), 304-309, <https://doi.org/10.1190/1.3353727>.

1388
1389 Shapiro, S. A., Krüger, O. S., Dinske, C., & Langenbruch, C. (2011). Magnitudes of induced
1390 earthquakes and geometric scales of fluid-stimulated rock volumes. *Geophysics*, 76(6),
1391 WC55-WC63, <https://doi.org/10.1190/geo2010-0349.1>.
1392
1393 Shapiro, S. A., Kim, K. H., & Ree, J. H. (2021). Magnitude and nucleation time of the 2017
1394 Pohang Earthquake point to its predictable artificial triggering. *Nature Communications*,
1395 12(1), 6397, <https://doi.org/10.1038/s41467-021-26679-w>.
1396
1397 Singh, A., Neupane, G., Dobson, P., Zoback, M., Morris, J., Fu, P., ... & Johnston, B. (2019).
1398 *Slip tendency analysis of fracture networks to determine suitability of candidate testbeds for*
1399 *the EGS Collab hydroshear experiment* (No. INL/CON-19-53585-Rev001). Idaho National
1400 Laboratory, Idaho Falls, United States.
1401
1402 Stanfors, R., Rhén, I., Tullborg, E. L., & Wikberg, P. (1999). Overview of geological and
1403 hydrogeological conditions of the Äspö hard rock laboratory site. *Applied Geochemistry*,
1404 14(7), 819-834, [https://doi.org/10.1016/S0883-2927\(99\)00022-0](https://doi.org/10.1016/S0883-2927(99)00022-0).
1405
1406 Stephansson, O., Semikova, H., Zimmermann, G., & Zang, A. (2019). Laboratory pulse test
1407 of hydraulic fracturing on granitic sample cores from Äspö HRL, Sweden. *Rock Mechanics*
1408 *and Rock Engineering*, 52, 629-633, <https://doi.org/10.1007/s00603-018-1421-5>.
1409
1410 Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the
1411 finite corrections. *Communications in Statistics-Theory and Methods*, 7(1), 13-26,
1412 <https://doi.org/10.1080/03610927808827599>.
1413
1414 van der Elst, N. J., Page, M. T., Weiser, D. A., Goebel, T. H., & Hosseini, S. M. (2016).
1415 Induced earthquake magnitudes are as large as (statistically) expected. *Journal of*
1416 *Geophysical Research: Solid Earth*, 121(6), 4575-4590,
1417 <https://doi.org/10.1002/2016JB012818>.
1418
1419 Verdon, J. P., & Bommer, J. J. (2021). Green, yellow, red, or out of the blue? An assessment
1420 of Traffic Light Schemes to mitigate the impact of hydraulic fracturing-induced seismicity.
1421 *Journal of Seismology*, 25, 301-326, <https://doi.org/10.1007/s10950-020-09966-9>.
1422
1423 Verdon, J.P., Pullen, B., & Rodríguez-Pradilla, G. (2023). Growth and stabilisation of
1424 induced seismicity rates during long-term, low pressure fluid injection. *Philosophical*
1425 *Transactions A*, 382(2276), 20230183, <https://doi.org/10.1098/rsta.2023.0183>.
1426
1427 Verdon, J. P., & Eisner, L. (2024). An Empirically Constrained Forecasting Strategy for
1428 Induced Earthquake Magnitudes Using Extreme Value Theory. *Seismological Research*
1429 *Letters*, <https://doi.org/10.1785/0220240061>.
1430
1431 Verdon, J. P., & Schultz, R. (2026). Induced earthquakes in the southern Delaware Basin,
1432 Texas, are bound by a geomechanically controlled maximum magnitude. *Geophysical*
1433 *Research Letters*, 53(3), e2025GL117419, <https://doi.org/10.1029/2025GL117419>.
1434
1435 Vigilante, P. J., Sone, H., Wang, H. F., Haimson, B., & Doe, T. W. (2017, June). Anisotropic
1436 strength of Poorman Formation rocks, kISMET project. In *ARMA US Rock*
1437 *Mechanics/Geomechanics Symposium* (pp. ARMA-2017). ARMA.

1438
1439 Villiger, L., Gischig, V. S., Doetsch, J., Krietsch, H., Dutler, N. O., Jalali, M., ... & Wiemer,
1440 S. (2020). Influence of reservoir geology on seismic response during decameter-scale
1441 hydraulic stimulations in crystalline rock. *Solid Earth*, 11(2), 627-655,
1442 <https://doi.org/10.5194/se-11-627-2020>.
1443
1444 Villiger, L., Gischig, V. S., Kwiatek, G., Krietsch, H., Doetsch, J., Jalali, M., ... & Wiemer, S.
1445 (2021). Metre-scale stress heterogeneities and stress redistribution drive complex fracture slip
1446 and fracture growth during a hydraulic stimulation experiment. *Geophysical Journal*
1447 *International*, 225(3), 1689-1703, <https://doi.org/10.1093/gji/ggab057>.
1448
1449 Wang, L., Kwiatek, G., Renard, F., Guérin-Marthe, S., Rybacki, E., Bohnhoff, M., ... &
1450 Dresen, G. (2024). Fault roughness controls injection-induced seismicity. *Proceedings of the*
1451 *National Academy of Sciences*, 121(3), e2310039121,
1452 <https://doi.org/10.1073/pnas.2310039121>.
1453
1454 Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights.
1455 *Psychonomic Bulletin & Review*, 11, 192-196, <https://doi.org/10.3758/BF03206482>.
1456
1457 White, M., Johnson, T., Kneafsey, T., Blankenship, D., Fu, P., Wu, H., ... & Zhang, Y. (2019,
1458 February). The necessity for iteration in the application of numerical simulation to EGS:
1459 Examples from the EGS Collab test bed 1. In *44th Workshop on Geothermal Reservoir*
1460 *Engineering*.
1461
1462 Waldhauser, F., Ellsworth, W. L., Schaff, D. P., & Cole, A. (2004). Streaks, multiplets, and
1463 holes: High-resolution spatio-temporal behavior of Parkfield seismicity. *Geophysical*
1464 *Research Letters*, 31(18), <https://doi.org/10.1029/2004GL020649>.
1465
1466 Wu, H., Fu, P., Morris, J. P., Mattson, E. D., Neupane, G., Smith, M. M., ... & EGS Collab
1467 Team. (2021a). Characterization of flow and transport in a fracture network at the EGS
1468 Collab field experiment through stochastic modeling of tracer recovery. *Journal of*
1469 *Hydrology*, 593, 125888, <https://doi.org/10.1016/j.jhydrol.2020.125888>.
1470
1471 Wu, H., Fu, P., Frone, Z., White, M. D., Ajo-Franklin, J. B., Morris, J. P., ... & EGS Collab
1472 Team. (2021b). Modeling heat transport processes in enhanced geothermal systems: A
1473 validation study from EGS Collab Experiment 1. *Geothermics*, 97, 102254,
1474 <https://doi.org/10.1016/j.geothermics.2021.102254>.
1475
1476 Yin, X., Jiang, C., Yin, F., Zhai, H., Zheng, Y., Wu, H., ... & Li, J. (2024). Assessment and
1477 optimization of maximum magnitude forecasting models for induced seismicity in enhanced
1478 geothermal systems: The Gonghe EGS project in Qinghai, China. *Tectonophysics*, 886,
1479 230438, <https://doi.org/10.1016/j.tecto.2024.230438>.
1480
1481 Yu, J., Eijssink, A., Marone, C., Rivière, J., Shokouhi, P., & Elsworth, D. (2024). Role of
1482 critical stress in quantifying the magnitude of fluid-injection triggered earthquakes. *Nature*
1483 *Communications*, 15(1), 7893, <https://doi.org/10.1038/s41467-024-52089-9>.
1484
1485 Zang, A., Zimmermann, G., Hofmann, H., Stephansson, O., Min, K. B., & Kim, K. Y.
1486 (2019). How to reduce fluid-injection-induced seismicity. *Rock Mechanics and Rock*
1487 *Engineering*, 52, 475-493, <https://doi.org/10.1007/s00603-018-1467-4>.

1488
1489 Zang, A., Zimmermann, G., Hofmann, H., Niemz, P., Kim, K. Y., Diaz, M., ... & Yoon, J. S.
1490 (2021). Relaxation damage control via fatigue-hydraulic fracturing in granitic rock as inferred
1491 from laboratory-, mine-, and field-scale experiments. *Scientific reports*, 11(1), 6780,
1492 <https://doi.org/10.1038/s41598-021-86094-5>.
1493
1494 Zang, A., Niemz, P., von Specht, S., Zimmermann, G., Milkereit, C., Plenkers, K., & Klee,
1495 G. (2024). Comprehensive data set of in situ hydraulic stimulation experiments for
1496 geothermal purposes at the Äspö Hard Rock Laboratory (Sweden). *Earth System Science*
1497 *Data*, 16(1), 295-310, <https://doi.org/10.5194/essd-16-295-2024>.
1498
1499 Zhou, W., Lanza, F., Grigoratos, I., Schultz, R., Cousse, J., Trutnevite, E., Muntendam-Bos,
1500 A., & Wiemer, S. (2024). Managing induced seismicity risks from enhanced geothermal
1501 systems: A good practice guideline. *Reviews of Geophysics*,
1502 <https://doi.org/10.1029/2024RG000849>.
1503
1504 Zhuang, L., Kim, K. Y., Jung, S. G., Diaz, M., Min, K. B., Zang, A., ... & Hofmann, H.
1505 (2019). Cyclic hydraulic fracturing of pocheon granite cores and its impact on breakdown
1506 pressure, acoustic emission amplitudes and injectivity. *International Journal of Rock*
1507 *Mechanics and Mining Sciences*, 122, 104065, <https://doi.org/10.1016/j.ijrmms.2019.104065>.
1508
1509 Zimmermann, G., Zang, A., Stephansson, O., Klee, G., & Semiková, H. (2019). Permeability
1510 enhancement and fracture development of hydraulic in situ experiments in the Äspö Hard
1511 Rock Laboratory, Sweden. *Rock Mechanics and Rock Engineering*, 52, 495-515,
1512 <https://doi.org/10.1007/s00603-018-1499-9>.
1513
1514 Zöller, G., & Holschneider, M. (2016). The earthquake history in a fault zone tells us almost
1515 nothing about Mmax. *Seismological Research Letters*, 87(1), 132-137,
1516 <https://doi.org/10.1785/0220150176>.
1517
1518