Deliverable

D3.4 Scalability of new OEF techniques from the field to the laboratory to Bedretto URL

Deliverable information	
Work package	WP3: Advancing operational earthquake forecasting and earthquake predictability
Lead	UNINA
Authors	Antonio P. Rinaldi, ETHZ; Paul Selvadurai, ETHZ; Zihua, Niu, ETHZ; Patrick Bianchi, ETHZ; Sofia Michail, ETHZ; Antonio Salazar, ETHZ; Claudio Madonna, ETHZ; Federico Ciardo, ETHZ; Luigi Passarelli, ETHZ
Reviewers	Warner Marzocchi, UNINA; Stefan Wiemer, ETHZ
Approval	Management Board
Status	Draft
Dissemination level	Internal
Delivery deadline	31.10.2021
Submission date	31.10.2021
Intranet path	DOCUMENTS/DELIVERABLES/Deliverable_3.4.pdf

Table of contents

Summary

Experiments conducted at the cm to decametre/hectometre scale provide a fundamental under-standing of physical processes leading to fracture creation and reactivation. The current development of monitoring techniques, including very sensitive earthquakes sensors as well as deformation monitoring will allow to significantly lower the completeness magnitude and hence bring the OEF models to a new level. In this deliverable, we present two datasets at different scale and pave the way to the development of new OEF models that account for a more advanced physical understanding of the earthquakes processes.

1.   Introduction

Laboratory and in-situ experiments provide essential information at small scale. It has been often suggested that the processes occurring at the cm to decametre scale are representative of the processes occurring in a real fault zone. This makes the lab experiments fundamental to understand the physical processes leading to ruptures. However, to which extend the OEF models can be scaled from/to this scale is uncertain, and novel approaches will need to be implemented as new information are collected at the laboratory and in-situ scale (rock lab).

Laboratory experiments provides a perfect environment to create repeatable conditions, while in-situ experiments are all unique but in controlled conditions and forcing. The development of new technologies to monitor extremely small earthquakes (at the nano and pico-scale) combined with the technological advancement on monitoring deformation will open the path to the creation of new empirical models, potentially scaling from this small scale to the natural environment.

In this deliverable, we present the data collected at both laboratory and rock-lab scale, and present the preliminary analysis aimed at fostering the use of these dataset to develop new OEF models. The laboratory dataset has been acquired at the newly create LabQuake machine at ETH Zürich, while the rock-lab scale dataset was collected during a recent injection experiment at the Bedretto Underground Laboratory for Geoenergy and Geosciences. In particular for this latest, we tried operational forecasting with simplified models and reported on the lessons learned by doing such an exercise: this will be pivotal for the development of new OEF models.

1.1   LabQuake

LabQuake is a state-of the-art triaxial rock testing system for the geomechanical testing of both soft and hard rocks with a number of unique capabilities targeted at enhancing laboratory seismology (Figure 1a). LabQuake can be used for systematic experimental investigations of the effects of pore-fluid variation in fractured media by measuring poroelastic, mechanical, hydraulic and, thermal properties, as well as acoustic emissions of intact, fractured rocks and saw cut samples. We are able to consider a wide range of stress regimes and fracture orientations that is representative of faulting conditions in the upper crust. The system can achieve confining pressures up to 170 MPa, pore pressures of up to 158 MPa, and exhibits a frame stiffness of 2500 kN/mm, making it capable to perform post-peak failure tests on hard rocks and can impose complex fracture networks through failure tests.

Figure 1c shows a schematic diagram of the sample layout and general loading of a triaxial sample for intact, fractured and (saw-cut) frictional rock testing. Advances in triaxial testing in LabQuake is apparent with deployment of distributed fibre optic (FO) sensing and broadband AE sensor (Figure 1b) but also with advanced processing tools thanks to the partnership with Elsys Instruments AG. Since 2018 they have provided support in the development of the unique and flexible data acquisition system. They have provided advanced hardware and software solutions provided for AE sensor development including a signal filter bank and the ability to perform tomographic analysis and continuous data recording simultaneously.
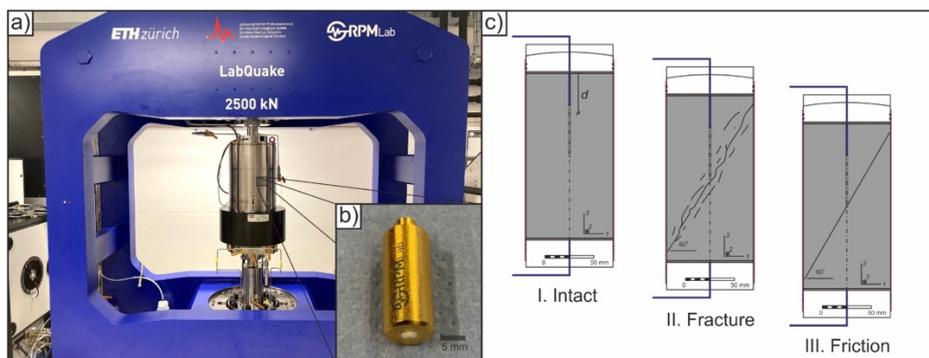
Figure 1: (a) Picture of the fully-operational LabQuake machine installed at ETH Zurich in the Rock Physic and Mechanics Laboratory. (b) Version of high-fidelity, broadband, high-temperature and high-pressure AE sensor designed exclusively for use inside the triaxial cell. (c) Configuration of the sample geometries that can be studied in the LabQuake machine.

1.2   The Bedretto Underground Laboratory for Geoenergy and Geosciences (BULGG)

The Bedretto Underground Research Laboratory (Bedretto URL), currently known as Bedretto Underground Laboratory for Geoenergy and Geosciences (BULGG), is a research infrastructure located in the Bedretto Tunnel in the Swiss Central Alps, which is a 5218 m long that connects the Furka railway tunnel with the Bedretto Valley (Figure 2). Since construction in 1982, the Bedretto tunnel remained unlined and unpaved and was mostly used for ventilating and draining the Furka tunnel. In 2018, the Bedretto tunnel has been made available by its owner (the railway operator MGB) to ETH Zürich to conduct research related to geoenergy and other topics. The tunnel runs from NW to SE approximately at an elevation of 1562 m a.s.l. at the junction with the Furka tunnel to 1480 m at the southern portal. The maximum overburden is ~1593 m at tunnel meter (TM) 3100 measured from the south portal. At the laboratory level, which occupies a 100 m long enlarged section of the tunnel at 2000 – 2100 TM, the overburden is about 1000 m. The host rock of the laboratory is a granitic body, the Rotondo granite, which has a boundary to metamorphic crystalline rock units at TM1138 and reaches beyond the junction to the Furka tunnel (Figure 2b). At end of 2020, during which some injection activities took place, there were nine boreholes drilled (Figure 1c-d). Four boreholes were fully dedicated to geophysical monitoring (MB1 to MB4) and two dedicated to fluid injection/monitoring (ST1 and ST2).
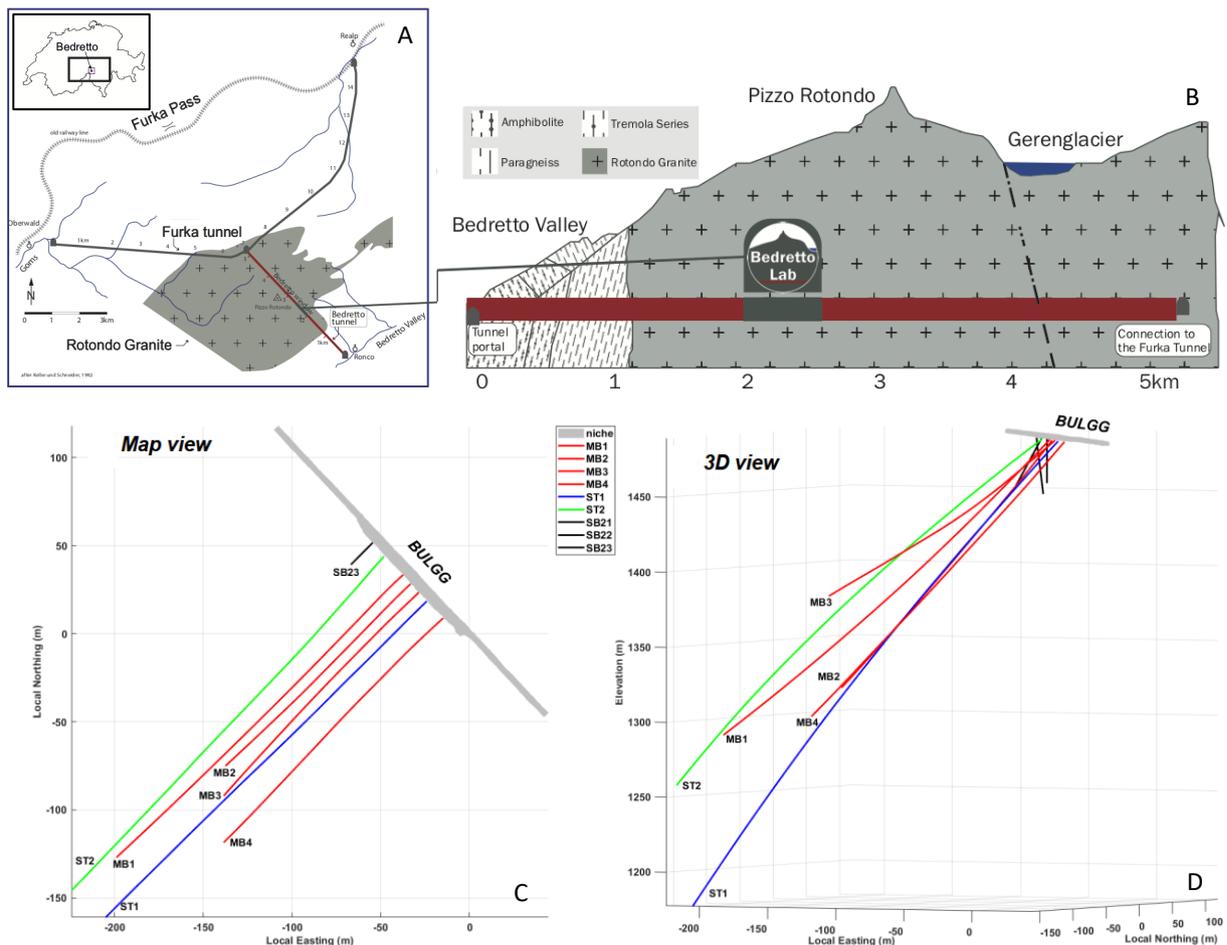
Figure 1: (A,B) Overview and geological context of the BULGG (adopted from SED) and (C, D) borehole configuration at the BULGG at the end of 2020.

2. Fracture creation at cm scale (LabQuake)

The work presented here is substantially based on the results by Niu (2021). Three confined compression tests on the intact cylindrical Rotondo granite samples have been conducted, through which the loading protocols and the parameters in the acquisition system are improved step by step. One particular experiment (LBQ2) shows seismic and aseismic processes before the failure of the intact Rotondo granite sample have been studied.

After the first two stages aimed at characterizing the rock sample, the main goal of Stage III is to study the seismic and aseismic deformation of the sample before and during the failure of the sample. The loading rate was set to 0.10 mm/min to slow down the approaching of the rock sample to failure. Ultrasonic tomography was conducted every 100 s. The peak stress of ~225 MPa was reached at around 893 s after the starting of the ramp up. The instant when the stress started to drop, the loading system was set to hold position. This enabled to observe a less violent stress drop of about 48.4 MPa for around 40 seconds (Figure 3).
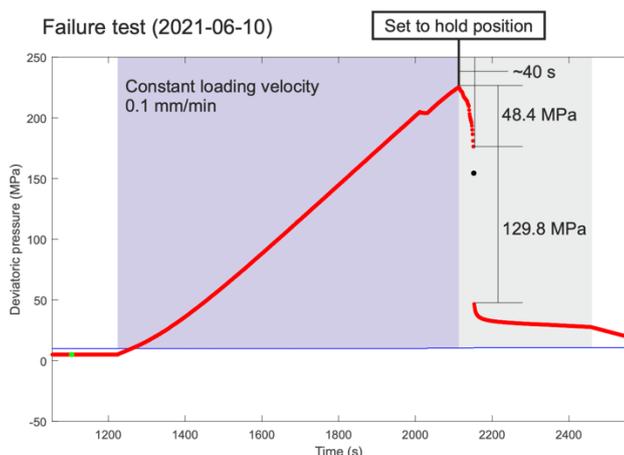
Figure 3: The loading protocols in the Stage III of LBQ2. The sample was directly loaded to failure.

During the experiment, Acoustic Emission (AE) were recorded, and the event selection is done with the short-time average/long-time average (STA/LTA) algorithm (Trnkoczy, 1998).

In the experiment LBQ2, after the event selection and classification, the use of an ad-hoc adapted AIC picker (Ada-AIC) enables getting the P-wave arrivals on all the traces, rule out the events that are not picked accurately and reduced the temporal resolution in detection to 2 ms. AE events are then located by using FaATSO: an inversion code that is capable of simultaneously locating seismic events and per- forming ultrasonic tomograph (Brantut, 2018).

Fiber optics cable are wrapped around the sample to measure volumetric strain. The fiber optics interrogators used in the present experiments were developed following the Rayleigh-based opti-cal frequency domain reflectometer (c-OFDR). Due to existence of AE sensors, the fiber optics cables cannot cover the entire surface of the rock sample. The interpolation in space is completed with a piecewise cubic, continuously differentiable and approximately curvature-minimizing poly-nomial surface (Nielson, 1983).

2.1 Fracture development and acoustic emissions

The above procedures eventually leave us with 1978 events whose locations and moment tensors can be inverted. The magnitudes of the moment tensors from LBQ2 are shown in Figure 4. From the FMD plot, the b-value is estimated as 0.88±0.02 with a magnitude completeness of Mw -7.9. This magnitude is at the same scale that grains of a few millimeters can generate (Manthei and Plenkers, 2018), which is the size of the most minerals in Rotondo granite (Rast, 2020). It is also observed that the data significantly deviates from the fitting results at the magnitude larger than around Mw -6.5.

More experiments are required to understand if this deviation at large magnitude is introduced by the physics of the sample failure or from the processing methodology. We noticed that early in

the loading, events of large magnitude occurred and resulted in swarms of earthquakes (Mogi, 1966) and a step increase in the cumulative seismic moment. However, most major events are observed in the few seconds before the failure of the entire sample. From the swarms of earthquakes in Figure 4c, three major groups, between which no event occurs, are observed. The average magnitude and the number of events in each group increase with time. The first group between 870 and 880 s contains only a few events and is located at the bottom of the events cluster that propagated upward (see Figure 4b), probably related to the initial coalesce of microcracks. The second group occurs between around 890 and 900 s along with the upward propagation of events. After 900 s and before the eventual failure of the sample, much more events appear and corresponds to the migration of events downward along the changed direction.
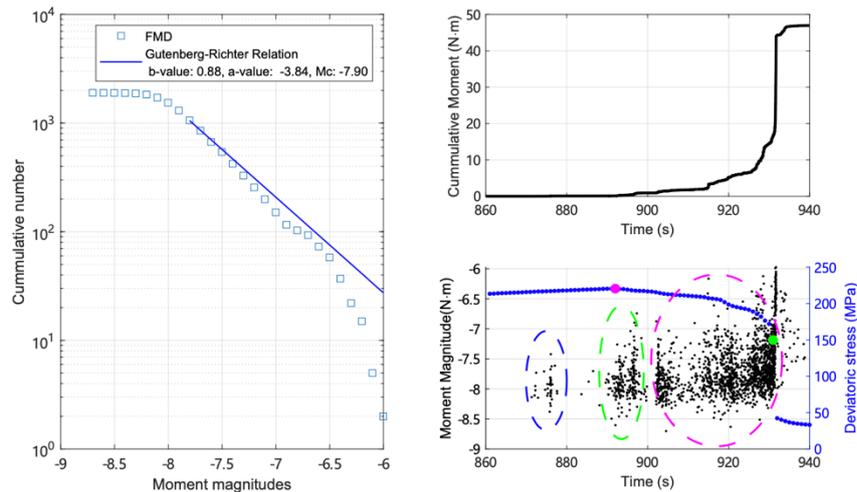
Figure 4: (a) Fitting of the frequency magnitude distribution of the events with estimation of the magnitude of completeness. (b) The cumulative moment release in time. (c) The magnitude of all events over time. The three dashed circles in (c) marks the three different clusters.

Aseismic deformation may be the predominant process during the failure of rock in the confined compression test. In LBQ2, it is found that the seismic deformation only accounts for $[0.07$ to $4] \times 10^{-2}\%$ of the total anelastic deformation and it seems to play a more important role (two orders of magnitude larger) as the sample approaches failure. Such relationship is based on the measurement of the distributed strain field and is, therefore, more representative of the entire volume than the point measurement.

2.2 Preliminary statistical analysis

Temporal distribution of the occurrence of the events is shown Figure 5. The cumulative number of events and the occurrence rate is plotted against the time to failure. The occurrence rate curve is fitted against the modified Omori's law (Lei and Ma, 2014). The exponential index is around -0.83, close to the value pointed out by Lei et al. (2003).

In Figure 6, temporal variation of fractal dimension, b-value and types of moment tensors are shown together for comparison.

Spatial distribution is characterized by the change of fractal dimension with time. As known from the locations of the events, a fault first propagated upwards from around 890 s to around 900 s. During this period the fractal dimension increased indicating the delocalization of events as the fault propagated. Next, the events localized again from 900 s to around 906 s, representing the nucleation of the new fault that would propagate in a different direction later. Then, during the propagation of new faults, the fractal dimension oscillated around 1.5, probably indicating the stop and propagation of the fault before failure. Immediately before the failure (around 931 s), the fractal dimension dropped from ~1.6 to ~1.4, which may correspond to the nucleation of the rupture front at a large aperture.
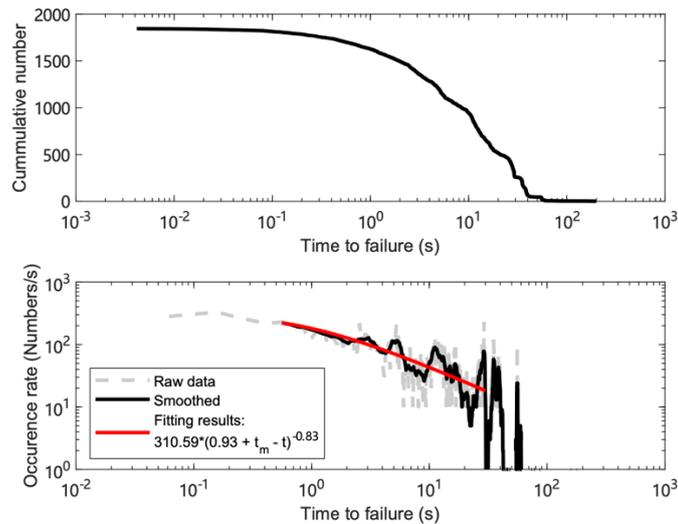
Figure 5. (a) The cumulative number of events against the time to failure. (b) The occurrence rate against the time to failure at the log-log scale. The dashed gray curve is the raw occurrence rate computed from the full seismic catalog in LBQ2. The solid black curve smooths the variation for better visualization of the trend and red curve is fitted using the raw data against the modified Omori's law.

b-value is an often-used indicator of seismicity in the sample. The lower the b-value, the higher the probability that large seismic events may occur. From Figure 5, from around 890 to around 898 s, during the propagation of the first fault, b-value drops from around 1.32 to around 0.88. Then, the position of the loading piston was held, accompanied by a small recovery of b-value. It is kept around 1.00 from 903 to 923 s until a further drop to around 0.80 right before the failure of the rock.

In Figure 6c, we looked at the types of events that also vary with the development of damage in the rock sample. During the propagation of the first fault, angles between the normal and the slip vectors increase from around 100 to 108 degrees, together with the drop of DC component. It shows that the sample is still shrinking and the events are becoming more compressive as the fault propagates. But since the change of the direction of the fault, the angles start to drop and reach 90 degrees at around 923 s. Then, the angles oscillate around 87 degrees. This indicates that events become predominantly expansive before the failure of the sample.

Previous studies tried to find correlations between the b-value and the fractal dimension. Lei et al. (2000; 2003) found that scattered background AEs results in relatively higher b-values. By assuming constant stress drop and isotropic rupture, it was estimated and also observed by Goebel et al. (2017) that the fractal dimension is nearly proportional to b-value with a slope of around 2. However, the spatio-temporal variation of the two variables was not looked at.
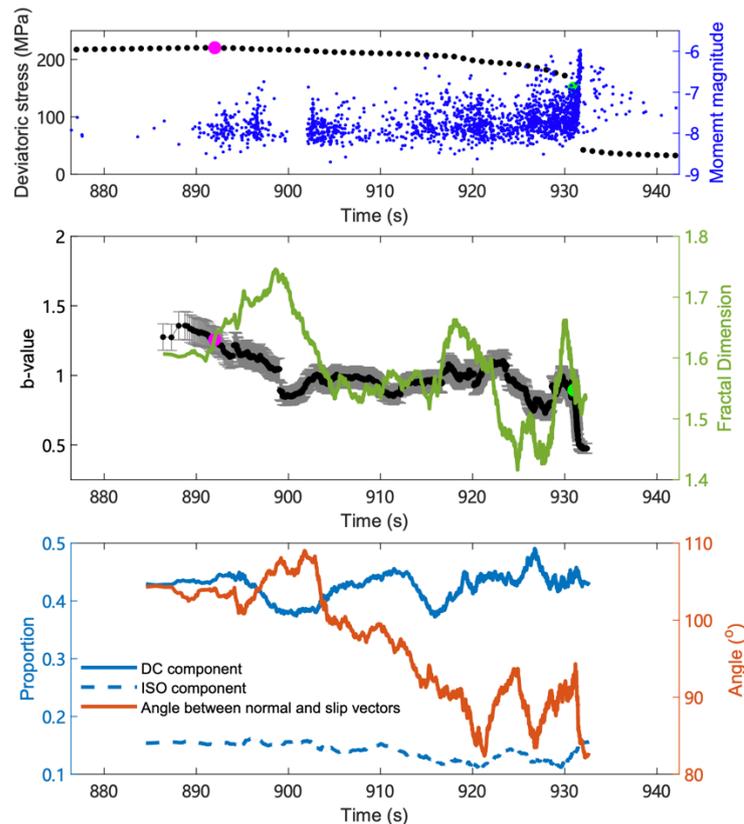
Figure 6. (a) The change of deviatoric stress (black dots) and moment magnitude of events (blue dots) in time serving as the indicator of the loading stage as statistical variables varies. The pink dots indicate the start of nucleation and propagation of the first fault, while the green dots mark the failure of the sample (the time when the sudden stress drop occurs). **(b)** Temporal variation of the b-value and the fractal dimension. **(c)** Temporal variation of different components of the moment tensors (the solid and dashed curves) and of the angle between normal and the slip vectors (the red curve).

2.3   Future activities and potential added value for OEF

Much more improvement can be made technologically, but the ultimate goal is to assist a better understanding of the physical processes in the failure of rocks. We highlight here some point that could provide a step-forward in the development of new OEF models:

- It is estimated that seismic deformation takes only a tiny proportion of anelastic defor-mation. The combination of FO and AE provides an enhanced understanding on how both aseismic and seismic processes occur. A model (empirical or hydro-mechanical) accounting for such behaviour may provide advanced forecasting capabilities, with the overall goal of predicting when the failure occurs.
- Near-real-time stress inversion is potentially a powerful tool for understanding the aseis-mic processes inside the rock if the its reliability can be justified. If the techniques are upscaled to natural seismicity, this would provide a unique tool to better forecast seismic-ity.
- The current developed methods also lay the foundations for the study of the failure of rock during the injection of fluid, which is the major task of LabQuake. The physical processes can be more complex since hydro-mechanical coupling will be involved. But with the frame-work developed and the familiarities gained so far, it is ready to overcome the forthcoming challenges.
- LabQuake has the capabilities to perform injection experiment, which would show if and how the b-value depends on human operations. While it may not be too relevant for natural OEF models, the identified behaviour may lead the path to better forecasting and planning the seismicity linked to geoernergies.

3.   Fracture stimulation at the decameter scale (Bedretto URL)

During the November/December 2020 a multi-stage stimulation was performed through boreholes ST1 and ST2. The injection occurred in packed interval, with the target depth varying inside each borehole. The injection depth varies between 268 and 344 m along borehole ST1 length, and 306 and 345 m along borehole ST2 length.

The activities started on Nov. 11th, 2020 in borehole ST2 at a target depth between 306-312 m, followed by progressively increasing the depth of the two-packer system until Nov. 29th, 2020. For most intervals, a test stimulation of about 5 m^3 was performed before 50 – 60 m^3 were actively injected into the rock (i.e. accounting for recorded bypass). Such test stimulation (referred to as TS-TLS) was needed to update some forecasting models and get a first understanding on how risky was to stimulate that interval and to test forecasting capabilities of the model. Then some minor re-stimulation occurred for the uppermost intervals, but with volumes limited to about 10 m^3. In terms of seismicity, during the TS-TLS between 64 and 128 events were located, with 41 to 75 events above the magnitude of completeness Mc in the range -3.15 and -3.01 and a maximum observed magnitude between Mw -2.77 and -1.88. During the main injection cycles, the number of located events ranged between 94 and 383, of which 57 to 231 above completeness that increased to the range Mc -3.15 to -2.91 with the maximum magnitude Mw between -1.88 and -1.71. The calculated b-value depended on the forecasting model, but the average for the whole ST2 stimulation was estimated as 2.2, calculated for a total of 861 events above completeness (Mc = -3.01). Figure 7 shows the recorded seismicity and pressure for the entire stimulation in ST2, as well some forecasting models fitting (more details below).
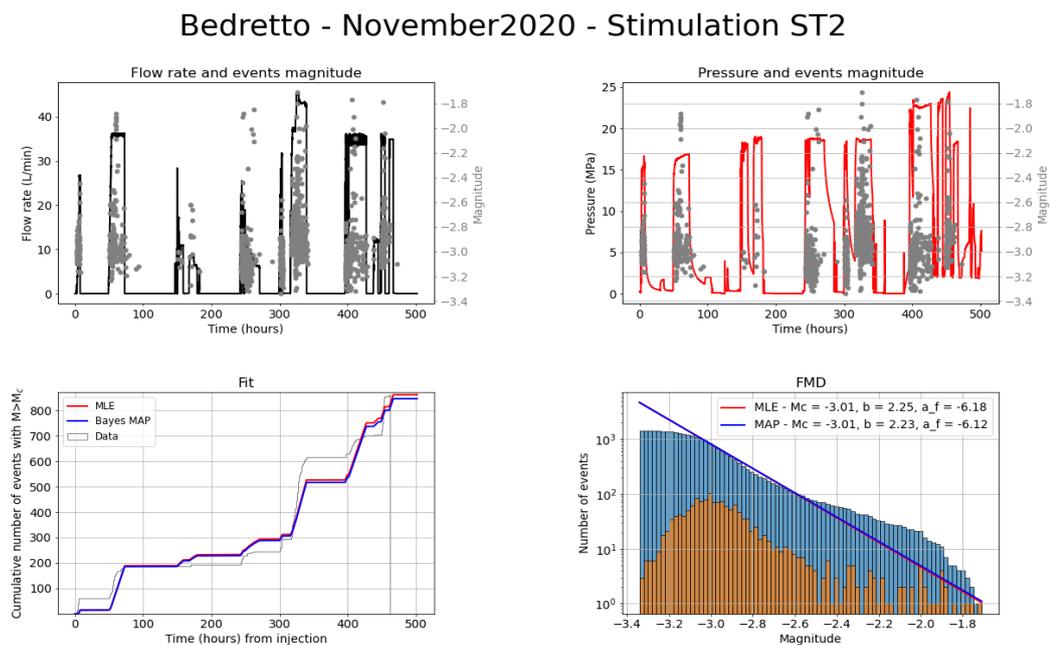
Figure 7. ST2 stimulation November/December 2020. (top left) imposed flow rate (black line) and recorded seismicity (gray dots).; (top right) recorded pressure (red line) compared to recorded seismicity (gray dots); (bottom left) models fitting to the evolution of total number of events; (bottom right) frequency-magnitude distribution and GR fitting for two forecasting models.

The stimulation of borehole ST1 started on Dec. 12th, 2020 at a target depth between 278-287 m. For this borehole, no TS-TLS injection was performed and a more powerful pump was available, that allowed injecting a larger volume of fluid (between 65 – 160 m^3). In this case the stimulation did not proceed with increasing the injection depth, as the two-packer system was moved first in between the most promising fractures identifying in the logs with the most bottom interval stimulated on Dec. 19th, 2020. The fracture systems seem much more permeable in this borehole compared to ST2, and indeed only four out of seven intervals showed a good pressurization to induced seismicity. The number of located events was much smaller compared to the ST2, ranging

between 1 and 301 events located, and when possible, with 20 to 152 events above completeness that ranged in between -2.86 and -2.7. The maximum observed magnitude Mw was in the range −2.54 and -2.24. For the whole ST1 stimulation, the average b-value was in the order of 3, calculated for 500 events above completeness Mc -2.82. Figure 8 shows the recorded seismicity and pressure for the entire stimulation in ST1, as well some forecasting models fitting (more details below).
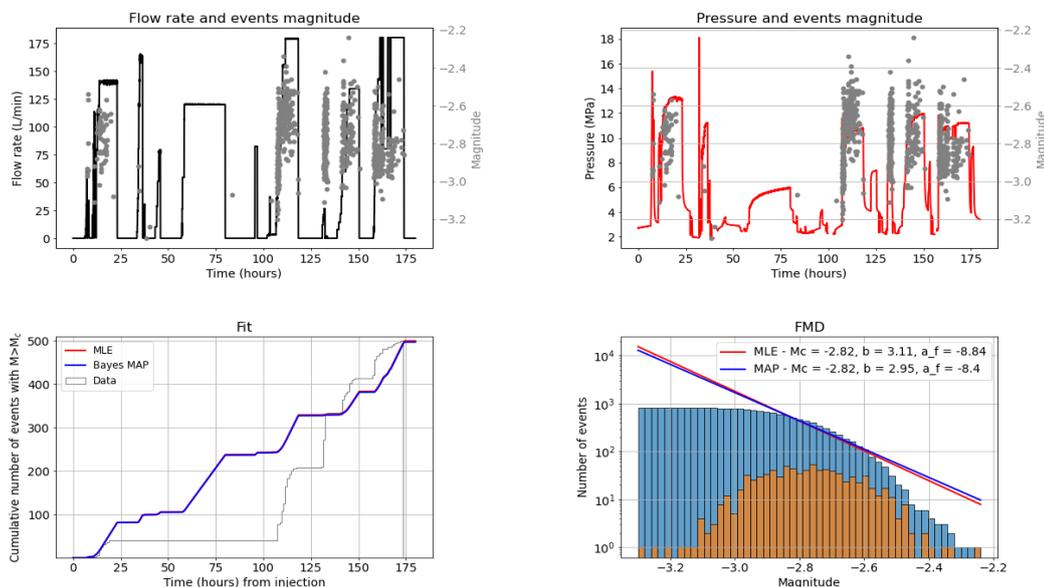
Figure 8. ST1 stimulation November/December 2020. (top left) imposed flow rate (black line) and recorded seismicity (gray dots).; (top right) recorded pressure (red line) compared to recorded seismicity (gray dots). To note as some cycles produced very low pressure with little to no seismicity; (bottom left) models fitting to the evolution of total number of events; (bottom right) frequency-magnitude distribution and GR fitting for two forecasting models.

3.1 Lessons learned while attempting real-time forecasting

We took the opportunity of the TS-TLS to test forecasting capabilities of induced seismicity models. While such models are not really usable for natural seismicity, here we report the lessons learned while attempting real-time forecasting and such a lesson is relevant for the future development of OEF models.

We tested three different models, the first two are related to the same class of "empirical" Non-Homogeneous Poisson Processes (NHPP) models and are described in details by Broccardo et al. (2017). In substance, the model has three parameters to link the injection flow rate to rate of seismicity, and it is based on the pioneering study of Shapiro et al. (2007). The two variations considered here accounts for a Maximum-Likelihood estimate of the three parameters (EM1_MLE) or for a Bayesian estimate of the parameters (EM1_BH).

The third model, referred to as HM0, accounts for a simplified hydro-mechanical formulation, in which the flow rate follows an analytical solution for a cylinder and the seismicity is computed as function of the simulated pressure front.

The main lessons learned are:

- One of the first questions we wanted to answer was: How much data do we need to forecast seismicity? By doing the TS-TLS experiment we realized that it was impossible to answer, as it really depended on the given interval being stimulated. Figure 9(top row) shows an example on how the models clearly were not able to predict the main stimulation based only on the first 5 m^3, while a little more data may actually help in having a better forecasting (bottom row). However, this could be strictly linked to the given interval, and to verify a general pattern a full pseudo-forecasting test would need to be performed. In general, it is important to have real-time assessment, have continuous data-feeding and

to look at the short term only (next hour), rather than trying to predict long-term with little information.
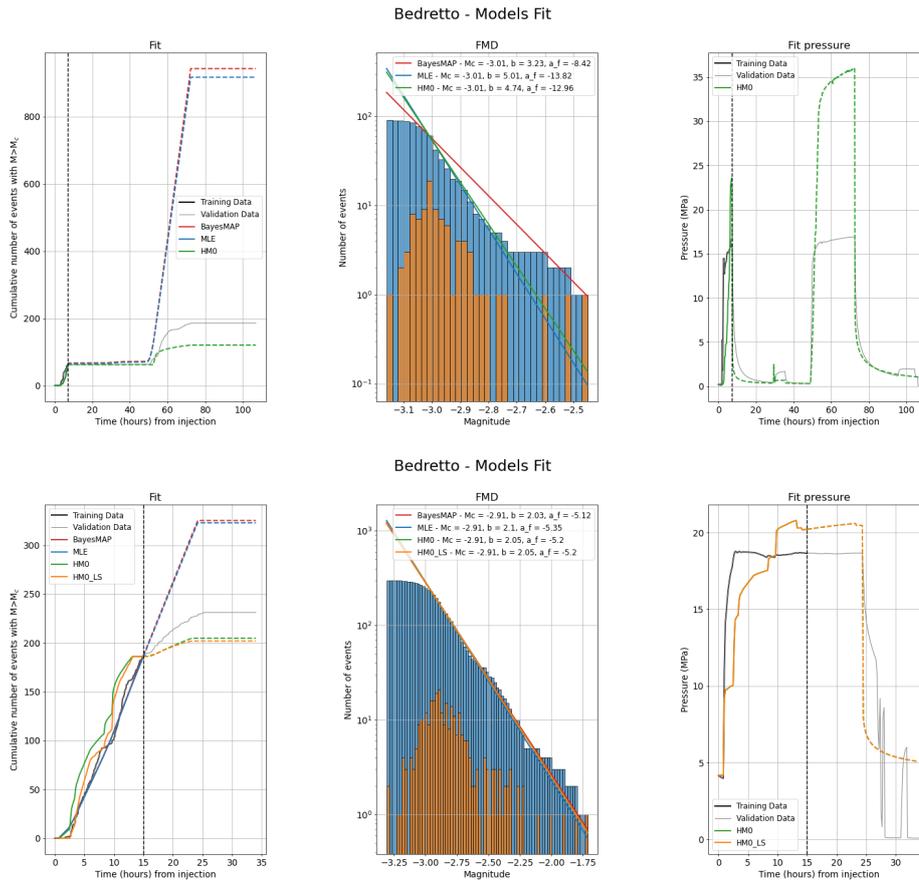
Figure 9. Testing forecasting capabilities of the three models. The top row shows an example of forecasted seismicity based on the TS-TLS only (i.e. the first 5 m³). The bottom row shows how the forecasting may improve with more data (30 m³).

- Was the seismicity detection in "real-time" accurate? Events picked by hand, so at time of forecast are not necessarily accurate. Figure 10 shows how the number of events strongly changed from the time of forecast (left) and in post-processing (right): at the time of forecast, the number of events was definitely too little to provide a reliable estimate of the model parameters. Full real-time forecasting would have updated automatically the fit/forecast.
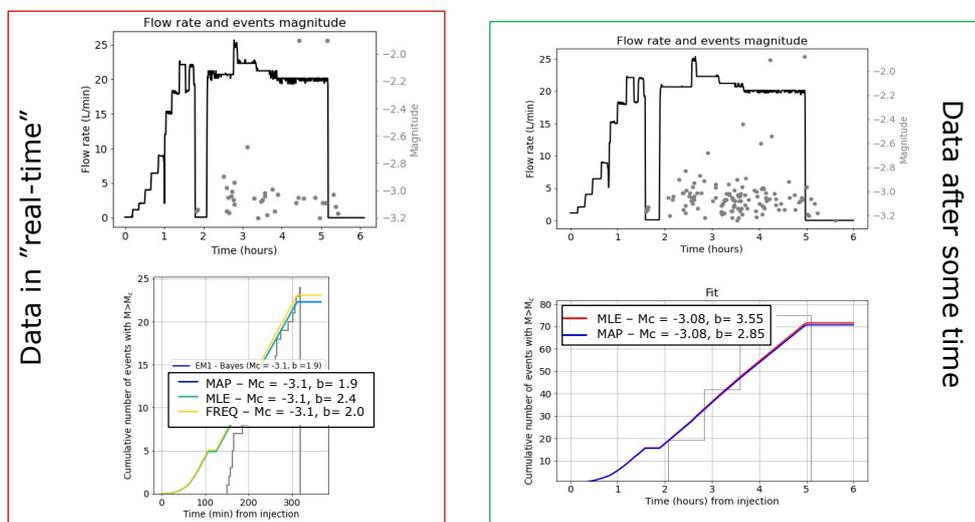
Figure 10. Comparison of real-time data and pot-processed.

- Magnitudes could be a problem for OEF models. Current models tested at Bedretto do have the ability to "communicate" when the fit is unreliable. Figure 11 shows the example of the FMD recorded for the stimulation of interval 1 in ST2, hard to understand what is the real Gutenberg-Richter fit or even if the data are reliable or indicative of a physical process. Understanding some observations have a real physical meaning in real-time is nearly impossible.
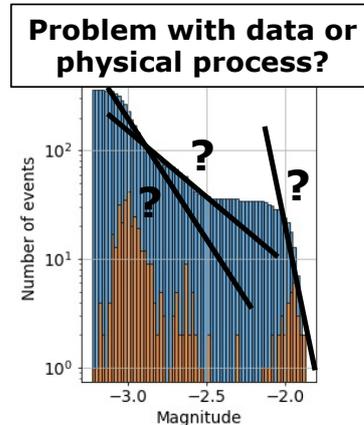
Figure 11. Frequency magnitude distribution recorded during the stimulation of Int. 1 in borehole ST2.

- An extremely bad case for the current models: high flow rate but little pressure increase and no seismicity observed. This occurred for a couple of intervals in boreholes ST1. For the current models would be impossible to state if a large event will occur. We could rely on pressure (i.e. HM0), but stimulating high transmissive interval is challenging, fluid can migrate far, and trigger seismicity at shallower depth.

3.2 From DEEP to RISE: Adapting CSEP model comparison for induced seismicity

In the framework of developing a forecasting system as described above, it is fundamental to compare model forecasts. In induced seismicity, we refer to the Adaptive Traffic Light System (ATLS) as a tool that embeds sophisticate methods for characterizing the seismicity and producing forecasts on the seismicity rate/magnitude to evaluate hazard and risk. Such an ATLS is developed in the DEEP project (http://deepgeothermal.org/home/) and does not constitute a direct output of RISE, but the development of the tools needed are of interested and could be easily adapted for forecasting natural seismicity in the context of OEF. We report here the current methodology employed to compare models.

We used the data for the interval 5 of ST2 to test a model comparison approach substantially based on CSEP approach but adapted to the conditions posed by the models described above (i.e. only temporal evolution and not spatial forecasting). In the given interval stimulation, a total of 115 earthquakes were recorded with magnitudes ranging from -3.2 to -1.8. We split the dataset at time 15 hours after the start of the injection in a training phase which contains 80 earthquakes with the rest belonging to the validation phase.

We fit the model described above and parameter uncertainties in the training phase for both EM1 and HM0 models (Fig. 12). After time 15 hours, we use the model in a predictive way to forecast the validation phase by only using hydraulic data as input (Fig. 12). For EM1 models we produced 1000 synthetic seismicity catalogs by tapping the parameters from their uncertainty distributions, while for HM0 we simulated a single realization. The results are reported in Fig. 12, where we present for EM1 models the median as best estimate of the number of events together with the variability of the simulated catalogs and parameters (shaded areas in Fig. 12 left and middle panel), while for HM0 we also show the single simulated catalog along with the input pressure profile used (Fig. 12 right panel). The median forecast catalogs of EM1 models overestimate the observed seismicity rate (thin black line in Fig. 12, left panel). However, the larger variability of forecast catalogs of EM1_BH can well reproduce the observed earthquake rate (see blue shaded area in Fig. 12 left panel). This large variability in the forecast arises from the larger uncertainties

associated with this EM1_BH (see Fig. 12 middle panel) compared with EM1_MLE. This implies that even if EM1_BH has larger uncertainties associated to the model parameter and apparently less precise, this model is more accurate in the simulation of forecast catalogs and better describe the temporal evolution of the observed seismicity of the validation phase. HM0 instead slightly underestimates the observed seismicity in the validation phase suggesting that Monte Carlo simulation of this model can be a promising path to explore.
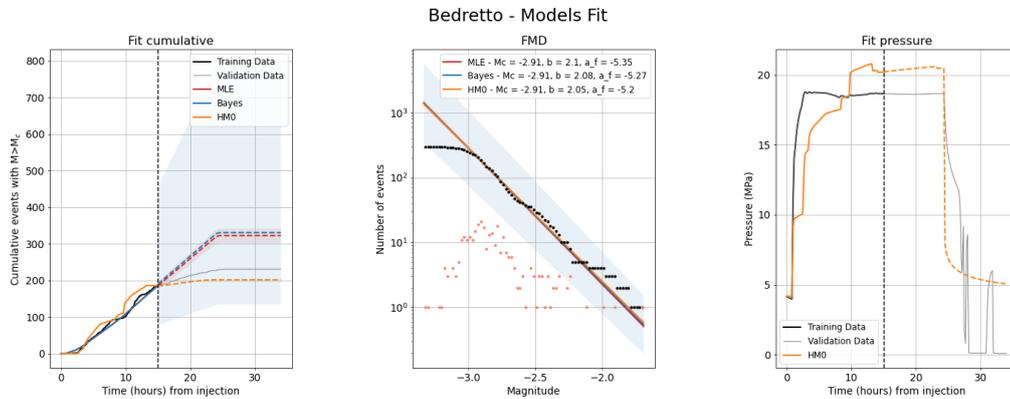
Figure 12. From left to right: panel 1) Cumulative number of events for observed data and simulated by the three seismicity models used in main text. Vertical dashed line indicate separation in training and validation data. The light blue and red shaded areas represent the 25th and 75th percentiles plotted around the median value of the N synthetics catalogues simulated from EM1_MLE and EM1_BH., respectively. Panel 2) Frequency magnitude distribution and fit, shaded area (light blue and red) indicates the uncertainties in the parameters of Eq.1. Panel 3) Pressure profile (black/gray) and fit (orange) for model HM0.

In order to proper compare the different models, we calculate the Probability Gain (PG) as logarithm of the empirical probability of reproducing the observed number of events in the given time bin. If a given model does not foresee the simulation of a number of realizations to compute an empirical distribution, we use the classical CSEP approach and take as PG the Poissonian Log-Likelihood. We have calculated the information gain as punctual and total PG using EM1_MLE as a reference model in time bins of 1800 sec. We present the results in figure 13 in terms of the cumulative evolution of the log-likelihood/log-probability values (left panel) for the three models and the PG curves form EM1_BH and HM0 against EM1_MLE (right panel). As expected, the best forecast performance is of EM1_BH which outperforms the other two models.
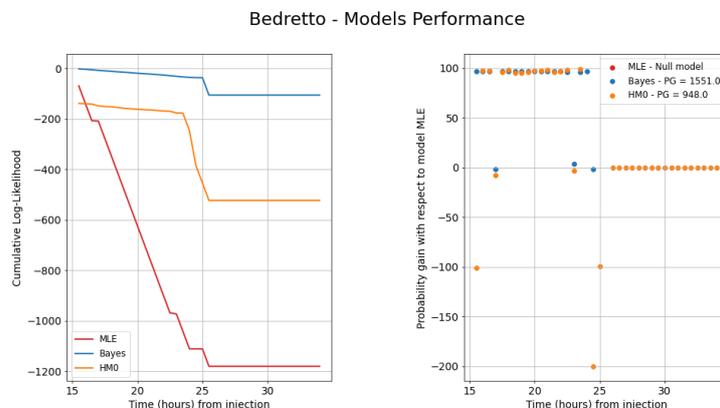
Figure 13. (left) Cumulative value of the log-probability/log-likelihood. (right) Punctual probability gain (PG); the total PG for each model is indicated in the legend. The PG is calculated against the EM1_MLE model and indicates that EM1_BH and HM0 perform almost always better in forecasting seismicity than EM1_MLE.

In summary, for the dataset from the Bedretto Underground Lab, where the number of induced earthquakes is relatively small, the best model is the one of larger parameter uncertainties, i.e. EM1_BH. Worth pointing out that both EM1_MLE and the median model EM1_BH (blue and red

dashed lines in Fig. 12 left panel) are very similar in the forecast performance. This arises from the use of the small amount of data in the likelihood function of the two models. The difference with the Bayesian approach is that can incorporate additional information in the prior distribution model that combined with the likelihood model (data model) produce a better estimation of the epistemic uncertainties when a small sample dataset is used to fit the model. This is a well-known advantage of using Bayesian hierarchical modeling in cases where only the data per se are not enough to capture the variability of the process. In essence, the larger uncertainties in the model parameters are not always indicative of a poor fit, rather they indicate that data alone cannot suffice to fully capture the epistemic uncertainties in the studied process.

3.3 Future activities and potential added value for OEF

The November/December 2020 stimulation at Bedretto was only the first attempt to enhance the hydraulic connection between wells ST1 and ST2. New activities are foreseen in the Fall 2021, when the upper part of the boreholes will be stimulated. This are the future advancements that will help producing more reliable forecasting:

- A more advanced method for characterizing the seismicity will be in place, with AE sensor similar to the ones used in LabQuake that will allow to lower the completeness theoretically to -5 and hence largely increasing the total number of recorded events. The sensor coverage will also strongly improve, as the stimulation will occur in the dense region of the MBs boreholes (Fig. 2)
- We have further refined the system that manages the model execution (referred to as RT-Ramsis). This will allow for full real-time forecasting of the number of events and we will better test the forecasting capabilities of our models. This will be a unique opportunity to have a forecasting model actually performing in real time at this "scientific" scale.
- While the models will not be adapted at the current stage, in the future we foresee the use of standard OEF models also for induced seismicity. In particular it will be important to implement models that features spatial forecasting of seismicity (e.g. ETAS) to be compared with more advanced hydro-mechanical and physics-based approach.

4. Conclusion

In this deliverable, we have presented the preliminary results from experiments performed at laboratory and rock-lab scale. Both environments allow for controlled experiments at different scales, ranging from cm of a small samples, to decametre and hectometre scale in the underground laboratories.

The newly developed LabQuake machine at ETH Zurich features enhanced real-time characterization of acoustic emission and deformation, and it constituted the perfect, controlled playground to test repeatability of physical processes to be included in the future OEF models. In this deliverable we show an example experiment, featuring already some preliminary statistical analysis and showing how the b-value changes when approaching failure. If the models produced within RISE are scalable, LabQuake will provide dataset to test such a scalability.

On the other side, the dataset produced at the Bedretto Underground Laboratory are unique, despite the control played by modulated injection activities. In this deliverable we have summarized the lessons learned in the first attempt of performing live forecasting (albeit not exactly real-time). Such an exercise was accompanied by developing of models and model comparison, in a similar approach to what is foreseen in CSEP. The lessons learned will be pivotal for the future activities and the dataset produced at Bedretto in the Fall 2021 will be the first of its kind at this scale, and it will constitute the ultimate testbench for the scalability of OEF models.

References

- Brantut, N., 2018. Time-resolved tomography using acoustic emissions in the laboratory, and application to sandstone compaction. Geophysical Journal International 213, 2177–2192.
- Broccardo et al., 2017. Hierarchical Bayesian Modeling of Fluid-Induced Seismicity. *Geophys. Res. Lett.* 44, 11,357–11,367.
- Goebel et al., 2017. Detecting significant stress drop variations in large micro-earthquake datasets: A comparison between a convergent step-over in the san andreas fault and the ventura thrust fault system, southern california. Pure and Applied Geophysics 174, 2311–2330.
- Lei, X., Ma, S., 2014. Laboratory acoustic emission study for earthquake generation process. Earthquake Science 27, 627–646.
- Lei et al., 2000. Quasi-static fault growth and cracking in homogeneous brittle rock under triaxial compression using acoustic emission monitoring. Journal of Geophysical Research: Solid Earth 105, 6127–6139.
- Lei et al., 2003. The hierarchical rupture process of a fault: an experimental study. Physics of the Earth and Planetary Interiors 137, 213–228.
- Manthei, G., Plenkers, K., 2018. Review on in situ acoustic emission monitoring in the context of structural health monitoring in mines. Applied Sciences 8, 1595.
- Mogi, K., 1966. Some presice measurements of fracture strength of rocks under uniform compressive stress. Felsmech. Ingenieurgeol. 4, 41.
- Nielson, G.M., 1983. A method for interpolating scattered data based upon a minimum norm network. Mathematics of computation 40, 253–271.
- Niu, Z., 2021. Experimental study on the seismic and aseismic deformation during the failure of granitic rock. ETH Zurich.
- Rast, M., 2020. Geology, geochronology and rock magnetics along bedretto tunnel (gotthard massif, central alps) and numerical modelling of quartz-biotite aggregates. ETH Zurich. https://doi.org/10.3929/ETHZ-B-000454117
- Shapiro et al., 2007. Probability of a given-magnitude earthquake induced by a fluid injection. *Geophys. Res. Lett.* 34, L22314.
- Trnkoczy, A., 1998. Understanding and setting sta/lta trigger algorithm parameters for the k2. Appl Note 41, 16–20.