Investigating the dynamic response of rock mass to reservoir drainage at Grimsel Test Site, Switzerland, as an analogue for Glacial Retreat

Marianna Kinali

Department of Civil and Environmental Engineering

University of Strathclyde

This dissertation is submitted for the degree of

Doctor of Philosophy

University of Strathclyde April 2020

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. This dissertation contains less than 65,000 words including appendices, bibliography, footnotes, tables and equations and has less than 150 figures.

Marianna Kinali

April 2020

Acknowledgements

This work is supported by Radioactive Waste Management Limited (RWM) and is part of the collaborative LASMO (LArge Scale MOnitoring) program at Grimsel Test Site. I would also like to acknowledge Sally Thompson from RWM and Florian Kober and Thomas Spillmann from Nagra (National Cooperative for the Disposal of Radioactive Waste) for facilitating my experiments and assisting in a smoother collaboration among all involved parties.

I would like to thank my supervisors Rebecca Lunn and Stella Pytharouli for choosing me as their PhD student and giving me the opportunity to be involved in this wonderful experience. I would like to express my sincere gratitude to both of them for understanding and supporting me through the difficult times as well. Their guidance has helped me in both research and write up of this thesis.

Finally, I would like to dedicate this thesis to my loving parents and my brother who have been supporting me throughout the years. A special thank you to my brother for training me, motivating me and being my rock through the rough and desperate times of this adventure.

Abstract

An effective solution for the geologic disposal of nuclear waste, with no environmental risk (i.e. avoidance of harmful release of radioactive material), is a fundamental issue for the environment protection, and for the future continued reliance on nuclear power. Although geological disposal is considered as the best option, there are still elements of risk to be addressed, such as glacial retreat, which could impact the safety performance of a geological disposal facility.

In this project two consecutive annual cycles of a reservoir in the Swiss Alps are used as a small scale analogue of the glacial retreat cycles, in order to investigate the response of granitic rock (as a host rock to a geologic disposal facility) to significant load changes. Assuming that the reservoir's stress changes cause the fractured and weakened rock slopes to slip, I chose to use microseismic monitoring as a tool to monitor the reservoir induced seismicity. A seismic network was deployed in the tunnels adjacent to the reservoir and recorded continuously ground movement over a 3.5-year period (Nov 2014 – Aug 2018).

In order to be able to detect microseismic slips in the acquired real field dataset I explore various algorithms from the literature and develop my own methodology. The two main problems my research focuses on are the length of the dataset (big data issues) and the signal to noise ratio of the events I want to detect (small magnitude events in a varying noisy background). My results show, albeit not all of the seismic signals were possible to locate or characterise, that the reservoir unloading increases the frequency of occurrence of microseismic events for a short time period in the region surrounding the reservoir. It is possible therefore that the construction of a geologic disposal facility will have a similar effect. However, the magnitudes of the induced events are very small and hence unlikely to have a significant effect as part of a safety case for a geologic disposal facility.

The contributions of this thesis can be summarised to: (i) using a reservoir as a small-scale test site analogue for exploring the seismic hazard in radioactive deep geologic disposal facilities due to glacial retreat; (ii) sensor deployment design and sensor data cleaning with noise characterisation for microseismic monitoring over several years; (iii) proposal of a new algorithm (NpD) for detecting potential seismic signals under not well-constrained conditions and without requirement of a priori knowledge about the expected signal frequencies and amplitudes; (iv) the NpD detection algorithm and acquired 3.5 years dataset are made freely available; (v) detailed discussion of onset time picking and hypocentre localisation methodologies, where again novelty lies in using, comparing suitability and adjusting a number of well-known approaches for the purposes of my project; (vi) compilation of a seismic catalogue related to the dynamic response of the rock mass to reservoir drainage.

Abbreviations

Above sea level (a.s.l.)

AIC-wav (onset time determination algorithm proposed in the present study)

Akaike Information Criterion (AIC)

Analogue to Digital (A/D)

Artificial neural network (ANN)

Coordinated Universal Time (UTC)

Edge-preserving smoothing (EPS)

Electromagnetic (EM)

European Geosciences Union General (EGU)

Geologic disposal facility (GDF)

Global Positioning System (GPS)

Grimsel Test Site (GTS)

Hertz (Hz)

International Seismological Centre (ISC)

Joint Energy Ratio (JER) method

Kraftwerke Oberhasli AG (KWO)

LASMO (LArge Scale MOnitoring)

Long Term Diffusion (LTD)

Long-term Cement Studies (LCS)

Modified Coppens's method (MCM)

Modified Energy Ratio (MER)

Nagra (National Cooperative for the Disposal of Radioactive Waste)

Non-parametric detection algorithm (NpD)

Power Spectral Density (PSD)

Quartile (Q)

Radioactive Waste Management (RWM)

Ratio of the peak eigenvalues (PER)

Reservoir Induced Seismicity (RIS)

Root-Mean-Square (RMS)

Samples per second (sps)

Short Time Average over Long Time Average (STA/LTA)

Signal to Noise Ratio (SNR)

Swiss Seismological Service (SED)

Three-dimensional (3D)

Time intervals (TI)

TOPological INVersion (TOPINV)

Underground Research Laboratory (URL)

Virtual Field Optimization Method (VFOM)

Contents

List of Figures

List of Tables

Chapter 1 Preface

1.1 Introduction

Radioactive waste is produced as a by-product from a range of activities, i.e. power generation, medical, industrial, research. The UK alone generates about 200 million tonnes of nuclear waste from households and other industries per year (Ripper, 2017). Most of the radioactive waste is produced by the decommissioning of nuclear power reactors.

Although there are a number of different options investigated for long-term disposal of radioactive waste, such as sea disposal, sub seabed disposal, deep well injection and even disposal in outer space, there are two commonly accepted disposal options. Near-surface geological disposal (at ground level or at depths of tens of metres below ground level), which is implemented in countries as UK and USA, Finland and Sweden for Low Level Waste and Scotland for Intermediate Level Waste; and deep geological disposal (at depths between 200m and 1km for mined repositories or between 2km and 5km for boreholes), which is official policy in several countries, while in the UK and Canada the site selection process has commenced (International Atomic Energy Agency, 2018).

Resulting in an effective solution for the geologic disposal of nuclear waste, with no environmental risk (i.e. avoidance of harmful release of radioactive material), is a fundamental issue for the environment protection, and for the future continued reliance on nuclear power. Although geological disposal is considered the best option, there are still elements of risk that need to be addressed and that may be of natural origin (e.g. glacial retreat, variation of earth orbital parameters, tectonics and meteorite impact) or of human origin (non-detected features, sealing defects, human intrusion and war effects) (Tshibangu K and Descamps, 2011). Out of the aforementioned elements of risk which are least understood is the impact of glacial retreat (Talbot, 1999), or load changes above a geologic disposal facility (GDF), and research is needed to provide further knowledge.

1.2 Glacier retreat and hydromechanical changes

The greatest natural threats to the integrity of the geological barriers of deep geological disposal facilities are due to rapid retreats of ice sheets (Talbot, 1999). There is a number of studies related to deep geological disposal facilities' performance assessments, which have identified glaciation/deglaciation capable of causing major impact on the long-term performance of the repository system (e.g. Boulton et al., 1999; Hedin and Kautsky, 2000). In particular, Boulton et al. (1999) developed an ice sheet model to simulate the ice sheet behaviour in time and space and concluded in the following: that coupling occurs between ice sheet dynamics and permeability; hydrofracturing and shear fracture may be widespread beneath the ice sheet; and that even intact granitic rocks can be fractured beneath the divide zone. Hedin and Kautsky (2000) conducted a post-closure safety assessment for a deep repository of spent nuclear fuel in granitic bedrock and concluded that in general the prospects are very good, i.e. the isolating capacity of the repository is not threatened in terms of thermal, hydrological, mechanical and chemical evolutions.

The actual phenomenon is explained as follows: as ice overrides the terrain, it acts as a direct glacial vertical loading to the ground surface, compressing the underlying geological media. Thinning and retreat of an ice sheet removes the load, allowing media to dilate. But the mechanical effects of glacial retreat are not solely attributed to this direct compression / dilatation. Flexural loading (i.e. stresses resulting from flexure or bending of the lithosphere under ice weight) appears capable of causing comparable or greater stress changes (Neuzil, 2012). Flexure lowers the ice sheet and the land surface elevation which in effect reduces the effective hydraulic head imposed by the ice sheet (Bense and Person, 2008). In addition, the surface rebound that follows the ice retreat results in changing the topographic gradient which drives the groundwater flow. The glaciation cycle – ground dynamics is demonstrated in Figure 1.

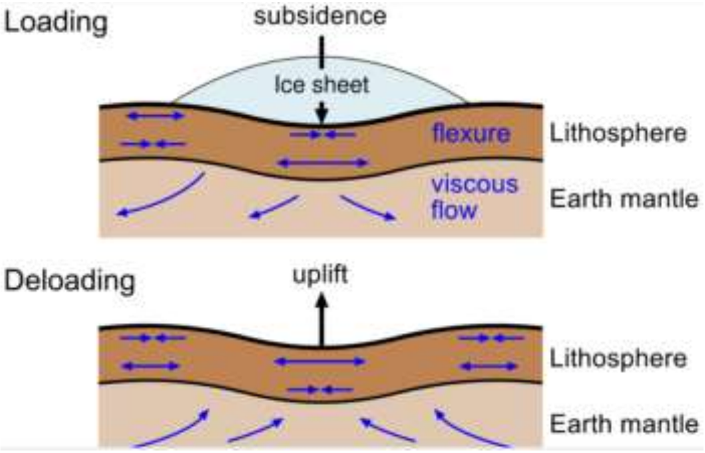


Figure 1: Glacial cycle; sketch reproduced from https://www.gfz-potsdam.de/en/section/earth-system-modelling/topics/solid-earth-dynamics/glazial-isostatic-adjustment-gia/

Taking a closer look at the medium, when this is porous and is loaded or unloaded small changes occur in its pore water volume, which result in pore water pressure changes. During glaciation the evolution of the subsurface hydraulic head follows the growth (mostly driven by precipitation rates) or decay (driven by rising temperatures) of the ice sheet (Chan et al., 2005). During ice-sheet growth rapid rise in hydraulic head, high transient hydraulic gradients and groundwater velocities 2–3 orders of magnitude higher than under non-glacial conditions are observed (Chan et al., 2005). Upon ice-sheet decay, hydraulic heads decline as a result of the unloading and the ground surface elevation increases as a result of isostatic readjustment (Bense and Person, 2008). Residual elevated heads may be preserved for 10,000s in the cases of low-diffusivity rocks (Chan et al., 2005).

How does the glacial phenomenon however cause failure? As per Grämiger's et al. (2017) simulations, glacial cycles, as purely mechanical loading and unloading phenomena, produce moderately restricted new damage. However, ice load fluctuations may increase the criticality of fractures, which may in effect increase the efficacy of fatigue; i.e. an already weakened rock slope is more susceptible to damage from glacier loading and unloading and may fail completely. Grämiger et al. (2017) found also that damage kinematics is controlled by discontinuity geometry (size, spacing and connectivity) and the relative position of the glacier. It has been considered by scientists that because the amount of loading during ice sheet growth and decay is effectively equal, its effects would balance out (e.g. Corbet and Bethke (1992)), however this hypothesis has been proven wrong by Tarasov and Peltier (1997). Their research showed that the ice sheet decay is usually more rapid than the growth of an ice sheet. Hence, even if the resulting mass equilibrium is achieved, the imbalance of rates of unloading and reloading may still affect the dynamics of the underlying structures.

1.3 Research question and objectives

Glacial stress changes are expected to occur, due to glacial cycles, whilst waste in the GDF remains radioactive, i.e. due to the long-lived nature of radionuclides (e.g. transuranic wastes remain hazardous for 1000 years https://www.nrc.gov). Therefore, it is crucial to consider the effect these stresses would have, being exerted on the repository and the surrounding rocks. This study focuses on exploring the response of a radioactive disposal facility to glacial retreat, using as a small-scale analogue the draining and refilling of a reservoir. The host rock studied here has granitic characteristics.

The reservoir is situated in the central part of the Swiss Alps, an artificial lake called Raeterichsboden. The lake was drained and let to refill with rainwater and snowmelt twice in the period of two years. There was an opportunity to monitor the lake level changes of this particular reservoir and their effect on the surrounding granitic rocks because of the convenient positioning of an adjacent underground research laboratory and its access tunnels which allowed for the deployment of monitoring equipment.

Based on literature (e.g. Talwani (1997) and Chan et al. (2005)), it is assumed that the load changes of the lake during its unloading will induce seismicity and change flow pathways because of the stress redistribution with no time delay. It is also assumed that during the lake's reloading, the groundwater head rises with a time delay governed by flow rates causing seismicity.

In order to be in position to detect the instances this reservoir induced seismicity occurs, continuous monitoring of the reservoir's surrounding area is needed. The reservoir water level changes (as an analogue to glacial retreat changes) are expected to increase the criticality of fractures, i.e. already weakened granitic fabric will potentially fail, resulting in small cracks opening or closing, or critical slopes slipping. This material failure will be transmitted as seismic signals of small magnitudes through the medium in the surrounding rocks. One potential way to capture this small magnitude induced seismicity through continuous monitoring of an extended area is through microseismic monitoring and analysis.

Microseismic monitoring refers to the recording and detection of small in magnitude (less than ML 3) earthquakes. It was mainly developed in the framework of the Test Ban Treaty (late 1950s) for the monitoring of the relaxation of the rock mass after nuclear weapon testing (Kinali et al., 2018; Lee and Stewart, 1981). In such a demanding environment, microseismic monitoring proved to be a powerful tool, tuned to detect weak seismic signals in low signal-to-noise ratios (SNR; Kinali et al., 2018). Induced (RIS) or Triggered Seismicity (RTS) mainly consists of sequences of microearthquakes with magnitudes ML 3 or less. Unless there are specific concerns of the occurrence of RIS/RTS, the phenomenon is usually monitored by existing national seismic networks with completeness magnitudes usually down to M = 2 or 1. Microseismic monitoring based on temporary installations has the potential to provide missed information on the occurrence of shocks with magnitudes ML=0 or even less than that (e.g. Pytharouli et al. (2011)). Hence, its applications have expanded into a wide range of projects related to RIS/RTS including the monitoring of rockslides and landslides (Helmstetter and Garambois, 2010; Torgoev et al., 2013; Yfantis et al., 2014), the monitoring of fracking processes (Maxwell, 2011), reservoir monitoring for geological CO2 (Zhou et al., 2010) and radioactive waste disposal (Young and Martin, 1993). A microseismic monitoring configuration mainly consists of short-period seismic arrays, with the components (seismometers) placed in a grid or triangular geometry, depending on their number.

The objectives of my research are as follows:

- Find a suitable field analogue to represent glacier retreat.
- Deploy microseismic instrumentation and be in charge of any on site works and troubleshooting.
- Collect microseismic data to describe the hydromechanical behaviour of the rock mass at the field site.
- Develop a methodology for the analysis of these data.

- Investigate the inflicted geomechanical responses (if any) to the lake load changes.

1.4 Novel research contributions

The key contributions of the thesis are listed below:

- A small-scale test site of a reservoir is chosen as an analogue for exploring the seismic hazard in radioactive deep geologic disposal facilities due to glacial retreats; this is described in Chapter 2.
- Although sensor deployment design and sensor data cleaning with noise characterisation for micro-seismicity monitoring is not an unmapped area, in this PhD this is explored in detail and for an extended period of several years; this is described in Chapter 3.
- A new algorithm (NpD) for detecting potential seismic signals under not well-constrained conditions and without requirement of a priori knowledge about the expected signal frequencies and amplitudes is proposed in Chapter 4.
- The NpD algorithm and the data set are open source.
- Although there are various onset time picking algorithms suggested by researchers (e.g. Irving‘s method, Akazawa’s method, JER-AIC, e.t.c. (Akram and Eaton, 2016)) the novelty of this research lies by using, comparing suitability and adjustment of a number of well-known approaches from the literature in conjunction for real field data and for microseismic events detection purposes (Chapter 4).
- A seismic catalogue related to the dynamic response of the rock mass to reservoir drainage is presented in Chapter 5.

1.5 Academic outputs

The outputs of this research are listed as follows:

- NpD algorithm was published as: M. Kinali, S. Pytharouli, R. J. Lunn, Z. K. Shipton, M. Stillings, R. Lord, S. Thompson; Detection of Weak Seismic Signals in Noisy Environments from Unfiltered, Continuous Passive Seismic Recordings. *Bulletin of the Seismological Society of America*; 108 (5B): 2993–3004. doi: https://doi.org/10.1785/0120170358. The algorithm is open source and is described in Chapter 4.
- Part of the project was presented in the Geology of Geomechanics conference (Oct. 2015) in London (https://www.geolsoc.org.uk/expired/Geology-of-Geomechanics-15).
- The NpD algorithm was presented (PICO presentation) in the European Geosciences Union General (EGU) Assembly 2018 in Austria, which was an international conference hosting 15,075 scientists from 106 countries (https://www.egu2018.eu/).
- Quarterly and annual reports detailing the progress of the research for Nagra (National Cooperative for the Disposal of Radioactive Waste; currently unpublished).
- A seismic catalogue related to the dynamic response of the rock mass to reservoir drainage is presented in Chapter 5 and Appendix B.
- The 3.5 years long dataset is freely available upon request. The dataset can be used for signal analysis purposes and / or for exploring the seismicity in the Grimsel area.

1.6 Thesis structure

The thesis is following the structure that follows: **Chapter 1** is a general introduction to the subject and also defines the research questions, novelty contributions, and research outputs. **Chapter 2** discusses the use of a reservoir as a field analogue for exploring the response of deep geologic radioactive disposal facilities to glacial retreat. In the chapter, further information of the geology of the investigated area is also provided. **Chapter 3** offers a description of the site activities (e.g. sensors deployment, maintenance and operation of the seismic network, preliminary noise investigations) as well as the seismic instrumentation used in the present study. **Chapter 4** includes the literature review explored with reference to detection and onset time determination of microseismicity, methodology used and approaches that were tested and were partially/not effective. **Chapter 5** includes the literature review explored with reference to deriving seismic source locations and source parameters and application in the present study. **Chapter 6** is the discussion on the work performed, as well as the novelty contributions, limitations and future improvements from the aspect of methodology proposed and results. **Chapter 7** summarises the conclusions of the research. **Appendix A** includes various algorithms written for this research and **Appendix B** includes the seismic catalogue with seismic and reservoir induced events.

Chapter 2 A field analogue for glacial retreat

2.1 Introduction

For the safety case consideration of a deep radioactive waste repositories research and monitoring are on-going (IAEA, 2001). The LASMO project (LArge Scale MOnitoring; January 2003 to 2018) is a monitoring program, established by Nagra, which investigates potential monitoring techniques for the safety assessment of such repositories. The scientific work was undertaken in the Underground Research Laboratory (URL) called Grimsel Test Site (GTS; “LASMO” (n.d.)). The GTS was established in 1984 and has hosted a wide range of repository-relevant research ever since. LASMO project benefits therefore from 30 years of GTS research.

The location of the GTS is in a catchment, in the Swiss Alps (Figure 2), where a number of hydroelectric power plants, owned and operated by Kraftwerke Oberhasli AG (KWO; grimselstrom.ch), are embedded. Out of eight in total storage lakes, two of them are located in the direct vicinity of the GTS: Lake Grimsel in the south and Lake Raeterichsboden in the east with peak water levels of 1,909 and 1,767 m asl respectively (Figure 2).

The GTS tunnels are located at depths from 1,727 m above sea level (a.s.l.) to 1,732 m a.s.l., while the GTS sits approximately 450m beneath Juchlistock mountaintop (see also Figure 2), 37m below the top of Raeterichsboden lake and 200 – 600m away laterally on the western side of the reservoir. Figure 2 shows the location of the GTS site and one of the main tunnels.

Figure 2: Panoramic picture of the general area where the GTS sits. Insets show the location of the GTS in Switzerland (in an orange dot) and a photograph from inside the main tunnel. Figure adapted from http://www.grimsel.com/gts-phase-vi/lasmo/lasmo-project-perimeter-gts-layout

The GTS tunnelling system spans more than 1 km in length and the average diameter of the tunnels is 3.5m. The central GTS facilities can be reached via an access tunnel from a northern entry (Gerstenegg tunnel) which reaches up till one of KWO's electrical power plants (Grimsel II; Figure 3). In Figure 3 the branching tunnel system of the GTS, the central facilities and the access tunnel are clearly visible. Projects undertaken by various research groups are also noted down on the map in the respective areas that each of them takes place. Finally, some information in regards to the geological setting of the GTS area and mapped shear zones are also provided.

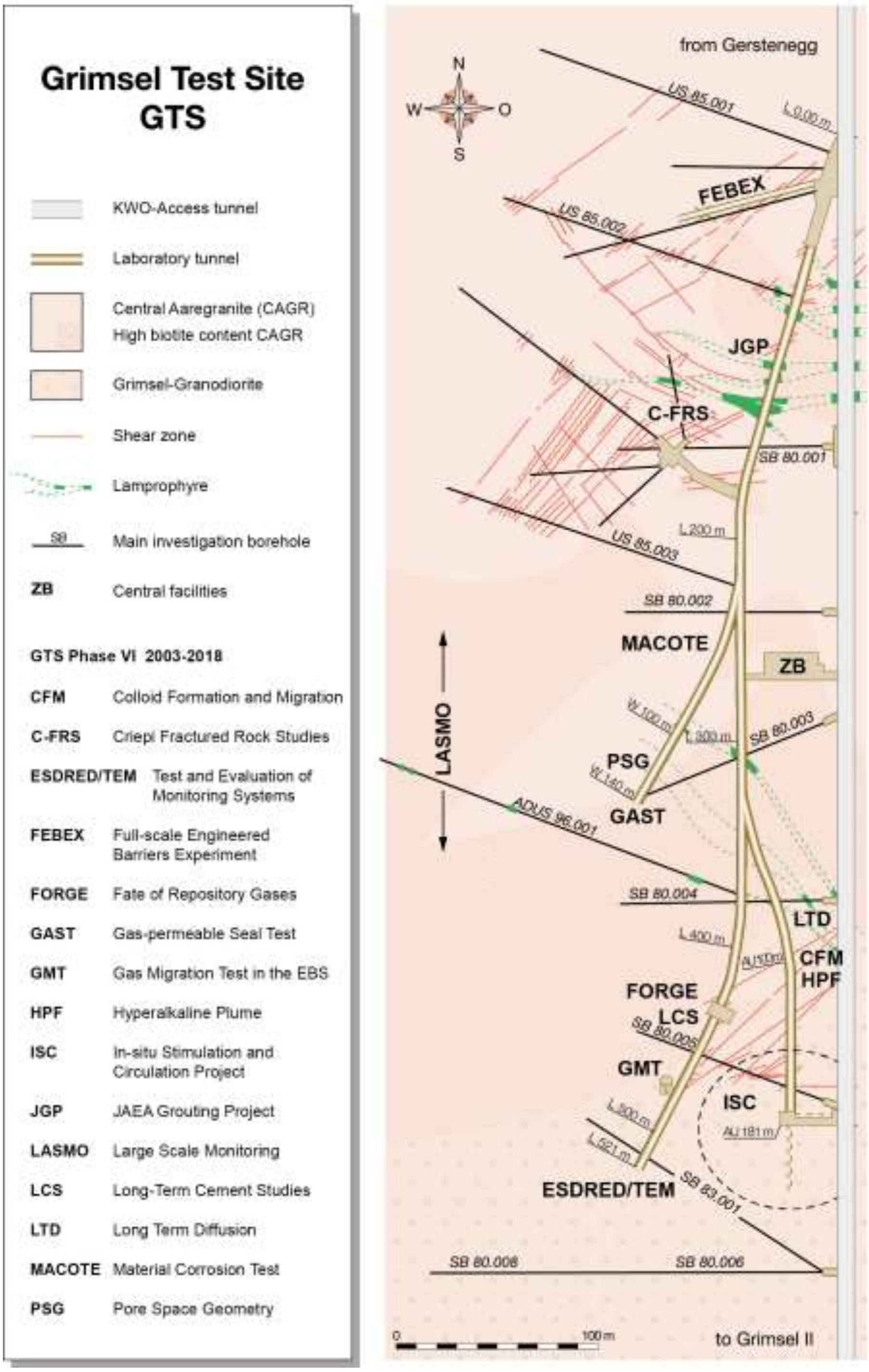


Figure 3: Map illustrating the tunnels within the GTS (AU, BK, EM, FRI, GS, HPA, MI, MOD, NM, UR, US, VE, WT) with the corresponding experiments that take place within these (Majer et al., 1990b). Figure reproduced from LASMO (2016).

2.2 A field analogue for Glacier retreat

A natural analogue is an important means of validation of long term geologic disposal of radioactive waste because it allows for the examination of processes that occur over geological timescales rather than the short timescales of the usual desk analysis (Chapman et al., 1984). For this PhD, the planned construction works of KWO, which led to the unloading of Raeterichsboden lake twice in a period of two years, as well as the conveniently located underground GTS tunnels rendered the lake as an ideal natural analogue for glacier retreat.

Figure 4 summarises the interaction concept among geomechanics, hydrogeology and geochemistry caused by the unloading and reloading of Raeterichsboden lake. The LASMO project combines geomechanical and hydrogeochemical findings from different research teams. The structural geology and hydrogeochemistry monitoring of this site was conducted by Raphael Schneeberger as part of his Masters and PhD in the University of Bern. Stress measurements and internal rock anisotropy research was conducted by Staš, L from the Institute of Geonics (UGN) of the Czech Academy of Sciences while measurements of 3D displacements by Josef Stemberk. Mark Stillings as part of his PhD in Strathclyde University explored the groundwater composition and the water-conducting features in the GTS. The aforementioned research, along with my PhD which explores the reservoir induced seismicity, will eventually all be bound together in a geosphere model. In the following paragraphs I will expand a bit more on the aforementioned interactions that influence RIS, i.e. the water level changes and the geology of the surrounding area.

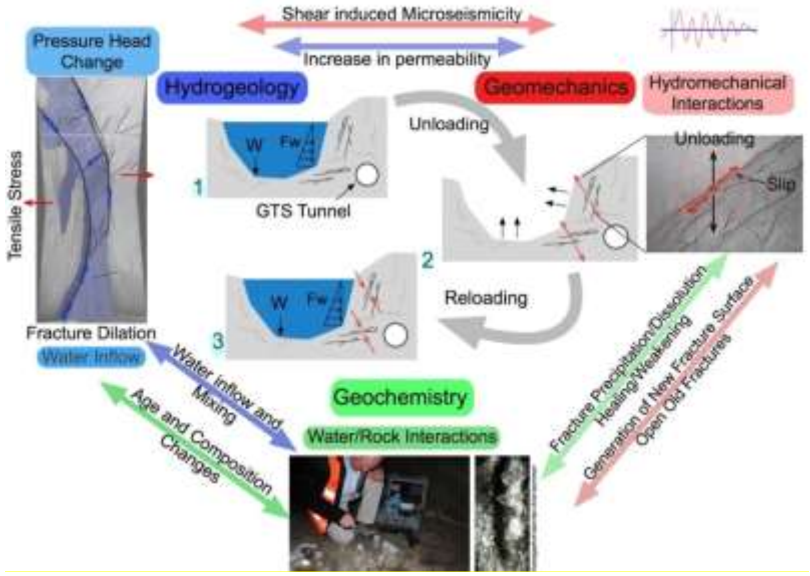


Figure 4: Interaction concept of LASMO project

2.3 Reservoir Induced Seismicity

Seismic triggering has a causal relationship with water level changes in large reservoirs (Simpson et al., 1988). Reservoir induced seismicity may follow the impoundment, large water level variations, or filling at a higher water level than the level achieved until then (Talwani, 1997). In general, induced seismicity is initiated because of a complex interaction mechanism between the pre-existing stress regime, the increase of the elastic stress, and the cumulative effects of increased pore pressure from elastic and diffusion mechanisms (Roeloffs, 1988).

RIS can be attributed to two processes of stress adjustment: the direct effect of loading, which increases the elastic shear stress and the pore pressure; and the effect of diffusion of pore pressure from the reservoir into the hypocentral zone (Gupta, 2002). In the first case the response is rapid, consists of low magnitude, shallow seismic events occurring below or in the immediate vicinity of the reservoir area, while in the latter the response is delayed and may extend well beyond the reservoir confines ($\geq$ 10km), not show an immediate correlation with major changes in reservoir level and be associated with large magnitude earthquakes, at deeper depths (Simpson et al., 1988).

Both water loading and pore pressure diffusion produce stress changes in the underlying faults (Lei et al., 2013, 2008). When the stress of a fault accumulates to the peak strength of a new rupture, a small stress disturbance can accelerate ruptures to trigger earthquakes (Alt and Zoback, 2017). As per Costain et al. (1987), in most cases RIS is attributed to a small stress increase triggering the release of large pre-stress, i.e. a pore pressure stress increase triggering failure along a pre-existing fault already tectonically stressed close to failure, rather than because of a great stress increase, since the deepest reservoirs add stress loads of up to 20 bars.

RIS has been reported by various researchers to be affected by the stress state of faults, hydraulic properties of fault rocks and connectivity of fluid pathways (Lei et al., 2008; Roeloffs, 1988). Roeloffs (1988) and Zoback and Harjes (1997) claim that the induced stresses are very small in cases of RIS and only affect the ambient stress field. Simpson (1976) suggest that RIS is more often encountered in strike-slip or normal faulting. Talwani (1997) report that the location of induced seismicity is governed by the nature of faulting below and near the reservoir. Finally, RIS has been observed to continue for years in some reservoirs, depending on the frequency and amplitude of lake-level changes, reservoir dimensions and hydromechanical properties of the substratum (Roeloffs, 1988). Gupta et al. (1972a), (1972b), and Simpson and Negmatullaev (1981) claim that the rate of increase of the reservoir water level, duration of loading, maximum levels achieved and duration of retention of high levels are factors influencing the frequency and magnitude of triggered earthquakes.

According to previous research by Gupta (2002), for the majority of the cases that reservoir induced seismicity has been reported, the magnitudes of the events are less than 4 M_L; i.e. in 95 sites globally (as per 2002 data) there were 4 cases of magnitudes more than 6 M_L, 10 cases of magnitudes between 5 and 5.9 M_L; 28 cases of magnitudes between 4 and 4.9 M_L, and 53 cases of magnitudes less than 4 M_L. Most of the RIS reported in the vicinity of Hsinfengkiang reservoir in eastern Chine in 1962 (Gupta, 1992) occurred near intersections of major faults and in rock masses with interbedded weak layers and their focal depths were shallow, i.e. 1 to 11km. The most seismically active areas and the strongest shocks location were reported to be in the vicinity of the dam where the water depth was deepest (Gupta, 1992).

2.4 Lake water level data

The water level data of Grimsel and Raeterichsboden lakes in the Grimsel region were provided by Nagra. Water level measurements are taken every 15 minutes. The graphs in Figure 5 show the water levels of the lakes expressed in metres above sea level over several years. The annually repeated pattern caused by snow melt and subsequent rainfall causing in turn drop and rise of the lakes' water levels is obvious in the figure. Keusen et al. (1989) observed periodical variations of water inflow in the GTS within the sphere of influence of Grimsel lake due to the effects of the snow-melt. The high lake levels in the second half of each year are controllably lowered by KWO to account for the snow melt in the year to follow.

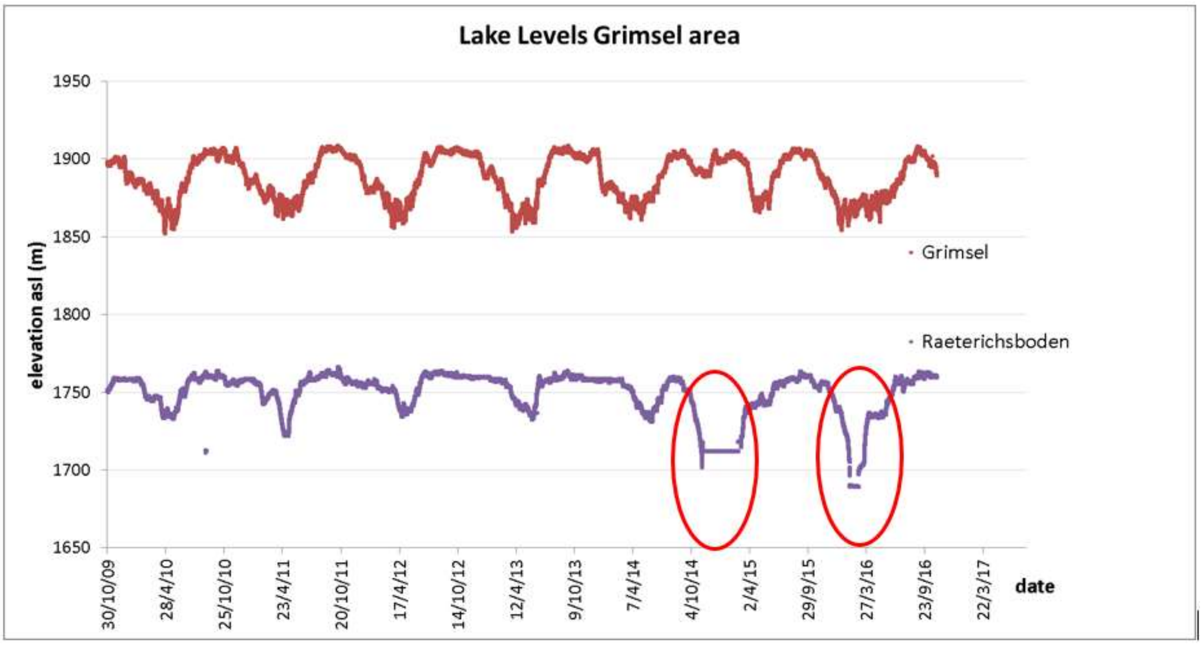


Figure 5: Annually repeated pattern caused by snow melt and rainfall effect in the Grimsel area lakes height data over several consecutive years (2009 – 2017). Lake water level data are given in elevation (in meters) above sea level. With red ovals are noted the two periods that the Raeterichboden lake was drained due to maintenance works of KWO.

With red ovals are noted the two periods that the Raeterichsboden lake was drained due to maintenance works of KWO and is also shown in Figure 6 in greater detail. Data exhibited include the period from 01.09.2014 00:00 to 01.09.2017 00:00. The maximum and minimum storage levels of Raeterichsboden lake within that time are 1763m and 1689m a.s.l.

respectively. As can be observed from the graph, the lake was controllably drained, gradually and not during the course of a single / couple of days and the level of the lake reached only during the day of the final drainage its minimum level (i.e. in the case of the first drainage period: 1703m a.s.l.). During the rest of KWO's maintenance works the taps remained open and the level of the lake remained at a constant elevation of 1712m and 1689m a.s.l. respectively for the first and second cycles. Some of the fluctuations observed in the graph are caused due to collection of water samples from productive boreholes within the GTS, together with flow rate measurements for other active GTS projects. Other fluctuations can be attributed to weather conditions (rain, snowfall, etc.). Under the graph there are two photographs of the lake, taken on the 07.11.2014 (right before the first drainage) and 05.05.2015 (during the refilling period).

The effect that the annual and smaller frequency variations of lake levels have on hydraulic head measurements around the GTS has been documented in the course of various experiments, for instance in the Long Term Diffusion experiments at the GTS (LTD; Havlova et al. (2013)), the Long-term Cement Studies (LCS; Sarout et al. (2012)); see also Figure 3 for locations of experiments. In the aforementioned experiments, direct correlation is difficult due to the additional head variation caused by earth tide effects. Overall, the similarity of hydraulic head responses to the lake levels changes observed in different experiments conducted in different parts of the GTS suggests a coupled mechanical response to the loading of the Juchlistock (LASMO Team, 2016; see also Figure 1).

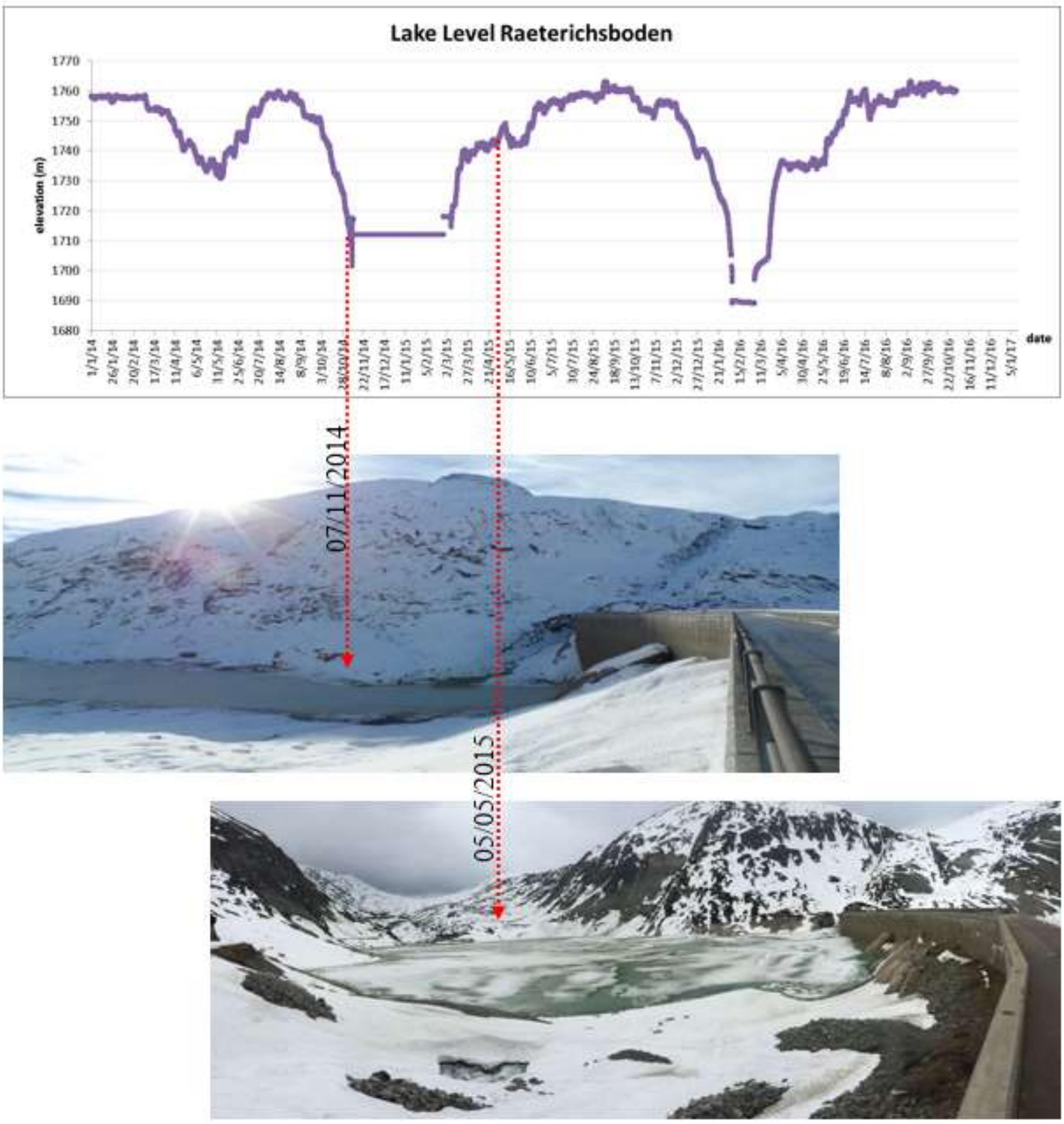


Figure 6: The two drainage and natural refill cycles of Raeterichboden lake due to KWO's maintenance works along with photographs taken during the first cycle.

2.5 Geological and Tectonic background

The Grimsel Test Site is located within the granitic formations of the Aar Massif (Konietzky, 1995). Majer et al.(1990) studied the major structures at the surface above the Grimsel laboratory. According to his paper, the geology comprises of magma - formed granite, granodiorite and dyke rocks (lamprophyres). Along the tunnels deformation structures, such as ductile shear zones, brittle faults, fracture systems and water pathways are visually evident. Both granite and granodiorite are foliated and the foliation strikes northeast, and dips steeply to the southeast. The foliation is defined by a grain-scale fabric combination of biotite and mylonite. Another steeply-plunging linear fabric element within the foliation is defined by elongated feldspar grains (Majer et al., 1990).

The hydrological important rock structures are the: S-zones, K-zones, and lamprophyres (Nagra (1985)). The S-zones are fracture-bearing shear zones, which generally dip parallel to the foliation (Keusen et al., 1989). The K-zones are fracture zones that generally strike west or northwest, cutting the host rock fabric. The lamprophyres are dikes, dipping steeply and generally striking west or northwest. The lamprophyres are highly discontinuous and are widely distributed. The leakage of water from the shear zone indicates that it is hydrologically active. There are also exfoliation joints present, suggesting subcritical fracturing conditions but also greatly varied orientations of near-surface stresses, possibly attributed to thermal or cryogenic stresses, or even a combination of processes (Ziegler et al., 2013; 2014a and b).

More recently, Schneeberger R. developed a surface fault map of the area surrounding the GTS through combination of fieldwork data and published maps (Baumberger, 2015; Keusen et al., 1989; Vouillomaz, 2009; Wehrens, 2015; Wicki, 2011). The map can be seen in Figure 7 (Schneeberger et al., 2016). The faults are categorised in different groups according to their orientation. Group A are mainly steep SE-plunging stretching-lineations resulting from ductile shearing with an average orientation (dip azimuth / dip) is 149º/74º. Group B are mainly steep S-dipping faults with mean orientation of 178º/72º. Group C are SW-dipping faults with an average orientation of 196º/72º. Group C faults coeval with group B and are subparallel to meta-basic dykes and often co-occur spatially with the latter. Along individual faults the standard deviation among multiple orientation measurements of the mean dip azimuth and dip is below 15° and 10° respectively. Generally, a greater number of faults are observed in the southern area (Figure 7). In the Figure the fault exposure lines that cross the GTS are labelled (also discussed in Figure 8b). The fault lines over uncertain areas are mapped as dashed lines.

Detailed underground mapping resulted in a lithological (Figure 8a) and structural map of the GTS (Figure 8b). Meta-basic dykes occur as three distinct groups. The two northern groups have a NW-SE strike, while the southern one strikes E-W. Numerous dykes are filled with an Alpine foliation or by localized ductile and brittle deformation, i.e. shear zones and fault gouges. As seen in Figure 8b faults occur along three NE-SW (Group A), two E-W (Group B) and two NW-SE (Group C) trending groups, leading to a heterogeneous strain distribution. Group A faults have an average 16m spacing and can be further subdivided into moderately to steeply dipping faults (between 45 and 75°) and subvertically dipping ones (>80°). Faults in the Central Aar Granite area (northern part) seem to preferentially localize along pre-existing anisotropies, i.e. brittle fractures or meta-basic dykes, hence form centimetre-sized discrete faults (inset in Figure 7 and 8). On the contrary, faults in the Grimsel Granodiorite area (southern part) form metre-sized strain gradients (inset in Figure 7 and 8).

Overall, the faults can further be discriminated to: 8 major group A faults (dip<80º) and 23 relay structures, 6 major group B faults and 7 relay faults, and 6 major group C faults and 32 relay structures (Schneeberger, 2017). Were all faults to be combined that would yield an average spacing of 25.4m. The lengths of the mapped lineaments range from 5m to 1941m. Faults in the GTS area converge with depth.

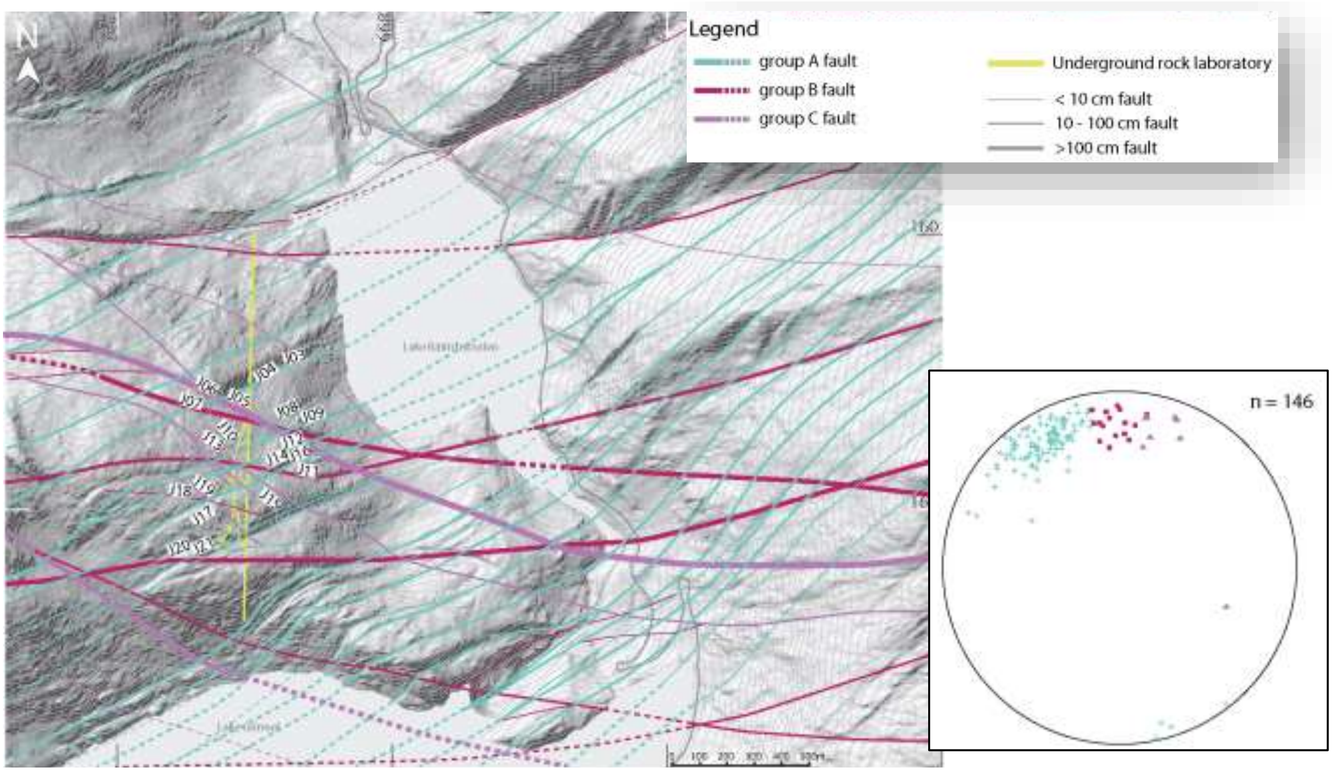


Figure 7: Figure reproduced from Schneeberger et al. (2017). Surface fault map with faults grouped by strike orientation (group A, B, C). Fault exposure lines are dashed over uncertain areas and are labelled in cases where a connection to GTS exists. Lower hemisphere equal area projection with planes poles grouped according to strike.

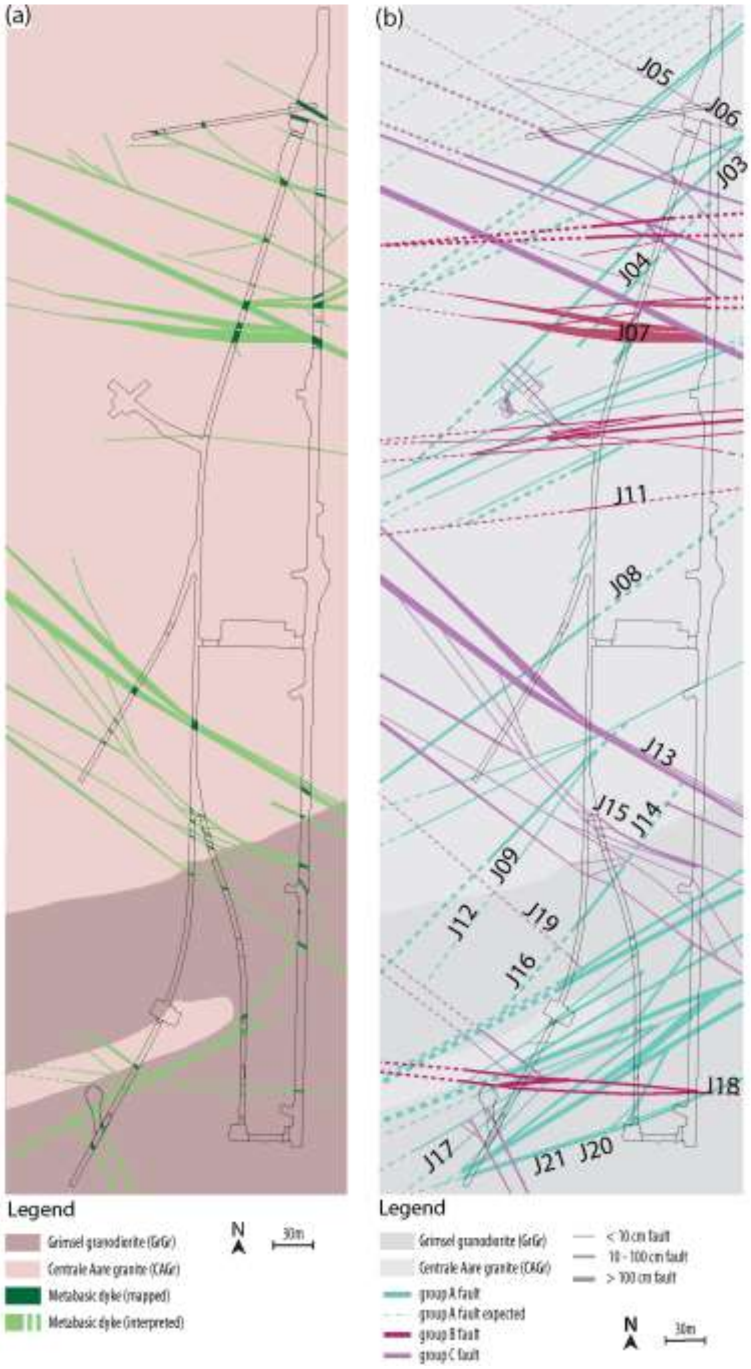


Figure 8: (a) Figure reproduced from Schneeberger et al. (2017). Petrographic underground map. (b) Structural mapping (1:1000) of the underground rock laboratory (GTS) with faults grouped according to their strike. Indicated labels correspond to surface fault labelling and represent ‘maximum a posteriori’ interpolation.

During one of my field trips to the GTS (16th till the 23rd of March 2015) I also proceeded in mapping the great features of the Gerstenegg tunnel along with another Strathclyde PhD student. At the time no geological map of Gerstenegg was available. The mapping of the features was thought at the time that it would enable me to link detected events with potential re-activation of the weak zones in the Gerstenegg tunnel and/or mapped geological features in the GTS. The map (Figure 9) was then to be digitised but this work was superseeded by Schneeberger et al. (2017).

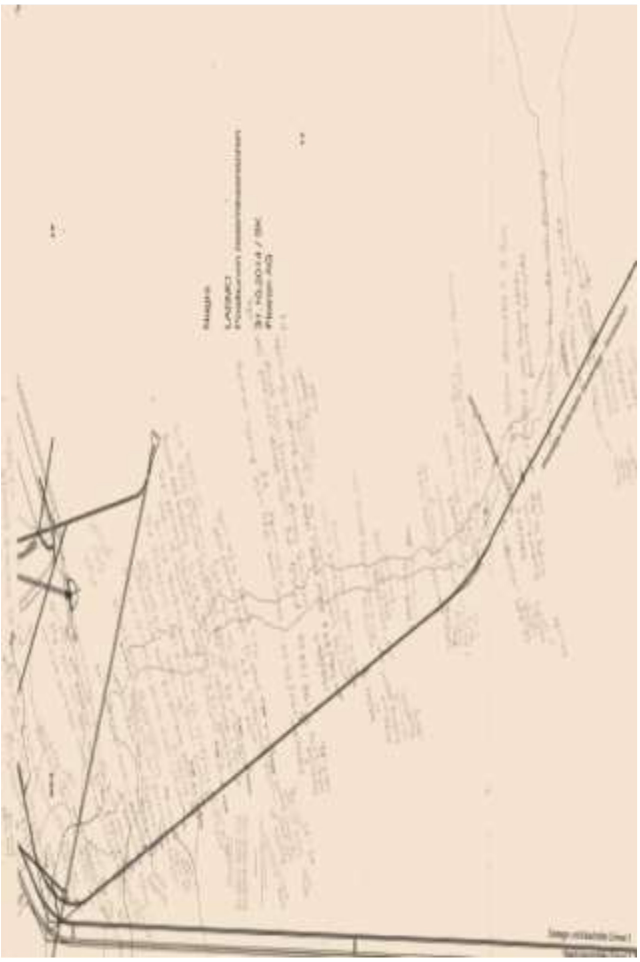


Figure 9: Geological mapping of Gerstenegg tunnel

2.6 Summary

In this chapter, I introduced the chosen small-scale field analogue to glacial retreat: the drainage and refilling of Raeterichsboden lake in the Swiss Alps (subsection 2.2) which is one of the novelty contributions of this PhD. The reason why RIS has been chosen as an analogue to glacial retreat induced seismicity was explained as follows: both RIS and glacial retreat induced seismicity have been documented to follow water level variations, in a similar way, because both phenomena cause the same two processes of short-term stress adjustment: the direct effect of loading and pore water pressure changes (subsection 2.3). In particular in this project, the planned construction works of KWO, which led to the unloading of Raeterichsboden lake twice in a period of two years, as well as the conveniently located underground GTS tunnels rendered Raeterichsboden lake as an ideal reservoir to monitor for RIS.

In this chapter it has also been explained, based on literature, that RIS is governed by the nature of faulting below and near the reservoir; hence the geology of the surrounding area of Raeterichsboden lake was investigated and is presented in subsection 2.6. As an overview, the area is governed by granitic rocks with 3 main groups of fractures, A, B, and C, dipping southeast, south and southwest respectively. The majority of the faults are in the south of the Raeterichboden lake and GTS. Not all fractures that are contained within these groups are of the same size, i.e. group A consists of 8 major faults (dip<80°) and 23 relay structures, while group B from 6 major and 7 relay, and finally group C from 6 major and 32 relay.

The granitic rocks' response to Raeterichsboden lake's water level changes is going to be monitored using microseismic monitoring, as already suggested in subsection 1.3; this is further discussed in the following chapter.

Chapter 3 Microseismic monitoring

3.1 Introduction

Microseismic monitoring was selected as the optimum approach to monitor the effect of the Raeterichsboden Lake levels variations to the surrounding rocks for a number of reasons:

(a) it allows for the recording of very weak events of energy dissipation;
(b) it is a continuous monitoring method;
(c) data can be transmitted in real-time from the monitoring site to the University premises, i.e. staff on-site are required only for the seismic sensors installation and regular maintenance;
(d) data can be analysed automatically by software routines;
(e) seismometers can be easily and rapidly deployed at a close proximity to the reservoir to record events;
(f) seismometers are easily retrievable; and finally
(g) the necessary equipment was already purchased for a previous PhD student by Strathclyde University and was readily available for this application.

Establishing a new microseismic network can be quite challenging and its intended goal largely defines the optimal technical design for it, i.e. detection of tectonic earthquakes or microseismic events, for a short time period or an extended amount of time. Not every design serves equally well for all applications and potential factors to consider are: the type of sensor (seismometer and/or accelerometer) and number of components per station; seismometer characteristics (sensitivity, dynamic range, frequency range of operation); number of sensors; network configuration, i.e. number of sensors and layout; communication infrastructure availability, i.e. processes for data transmission, storage, GPS (Global Positioning System) and power availability; and site selection. The network design for the present project will be discussed in the following subchapters.

3.2 Seismic sensors

Traditionally, accelerometers have been considered for strong motion only and broadband seismometers for weak motion. However, the latest generations of accelerometers are nearly as sensitive as standard short-period seismometers down to frequencies of 1 Hz and have a large dynamic range (up to more than 110 dB). At the same time, broadband sensors have a larger dynamic range. In terms of signal processing, there is no difference in using a seismometer or an accelerometer and correcting digital data for differences in the instrument response can make the signals look identical.

Surface arrays and downhole sensors have different advantages / disadvantages (e.g. see Eisner et al. (2010)) and it is common practice to combine their use in field to counterpart their shortcomings, i.e. borehole sensors / arrays suffer from lateral positioning accuracy while array sensors from low Signal to Noise Ratio. In this project, the deployed microseismic instrumentation consisted of 2 short-period surface arrays (Lennartz electronic GmbH) and a downhole 3D seismometer (IESE, Auckland University, New Zealand). The sensors were deployed and monitored seismic activity for a period of 3.5 years.

Each of the short-period surface arrays consists of one three-directional component and three single component seismometers (Figure 10(i)), with a flat response between 1 and 80 Hz and an eigenfrequency of 1 Hz. The 3D seismometer is of a diameter of 97mm and a height of 68mm, while the vertical component is of a diameter of 85mm and a height of 55mm. The downhole, three-directional component sensor (Figure 10(ii)) is a sonde designed for permanent installations in high-pressure environments and has a flat response between 6 to 100 Hz and an eigenfrequency of 4.5 Hz. The sensor has an exterior casing of stainless steel, weighs 30kg, a height of 1.07m and a diameter of 88mm. The maximum depth it can be deployed is 245m.

The sensors are designed to have a flat frequency response over a given bandwidth, which can be read as a prescribed frequency range. Events with bandwidths outside of the flat frequency response will be distorted by the instrumental response and yield unreliable magnitude estimates if they are detected at all. The bandwidth of the instrument needs to capture a robust estimate of the spectral plateau (the high-frequency decay) and not just the corner frequency (Baig and Urbancic, 2010). This last requirement practically means being able to sample at a rate that is about a factor of 8 greater than the corner frequency. For microseismic events, with magnitudes down to -2, corner frequencies can be 500 Hz and therefore sample rates of 4000 Hz are needed. In this project, there was the additional restriction of data storage and data handling, i.e. since the anticipated duration of monitoring would reach 3.5 years, it was thought practical to choose a small sampling rate of 250Hz. That would yield a total of 13,500,000 data points per hour and 1.1826E11 per year. Hence, with an 80Hz flat response and 250Hz sampling rate, the detectable microseismic events would be expected to be of a 31Hz corner frequency and up till -1.49 magnitudes. Overall, the seismic sensors fit the purposes of the present research since their sensitivity targets weak motion and since I am aiming to capture the immediate responses of the granitic rocks to the reservoir changes (i.e. low magnitude, shallow seismic events occurring below or in the immediate vicinity of the reservoir area, as discussed in subchapter 2.2).

In regards to their actual physical connections, in the case of the arrays, all three single component seismometers need to first connect to a data distributor and then, via a single geophone cable, to the data logger. The 3D surface and downhole sensor connect straight to the data logger. The data logger can operate either with a battery or connect directly to the mains. The data logger is also connected to a GPS antenna (continuously or whenever needed) for its internal clock mechanisms to remain synchronised. For all connections Lennartz-supplied cables, mating female connectors, open wires or prolongation cables were used.

The data loggers used for the surface arrays were the units AC16 and BD57 and for the downhole sensor the B961. Both surface arrays data loggers have in-built the same Analogue to Digital (A/D) converters, their sensitivity is 400 V/m/sec but AC16 unit has bit weight equal to 4.9E-8 V/counts while BD57's is 8.6E-8 V/counts. B961 unit has a smaller sensitivity equal to 70 V/m/sec and its bit weight is 2.761 μV/count (see also Table 1). These technical characteristics are needed for the conversion of the raw data (in counts) to ground velocity (m/s) (see **Appendix A**).

Table 1: Technical characteristics of data loggers

	North array	South array	Borehole sensor
A/D converter	RT505	RT505	RT649
Sensitivity	400 V/m/sec	400 V/m/sec	70 V/m/sec
Bit weight	8.616e-08 V/counts	4.945e-08 V/counts	2.761e-6 V/counts

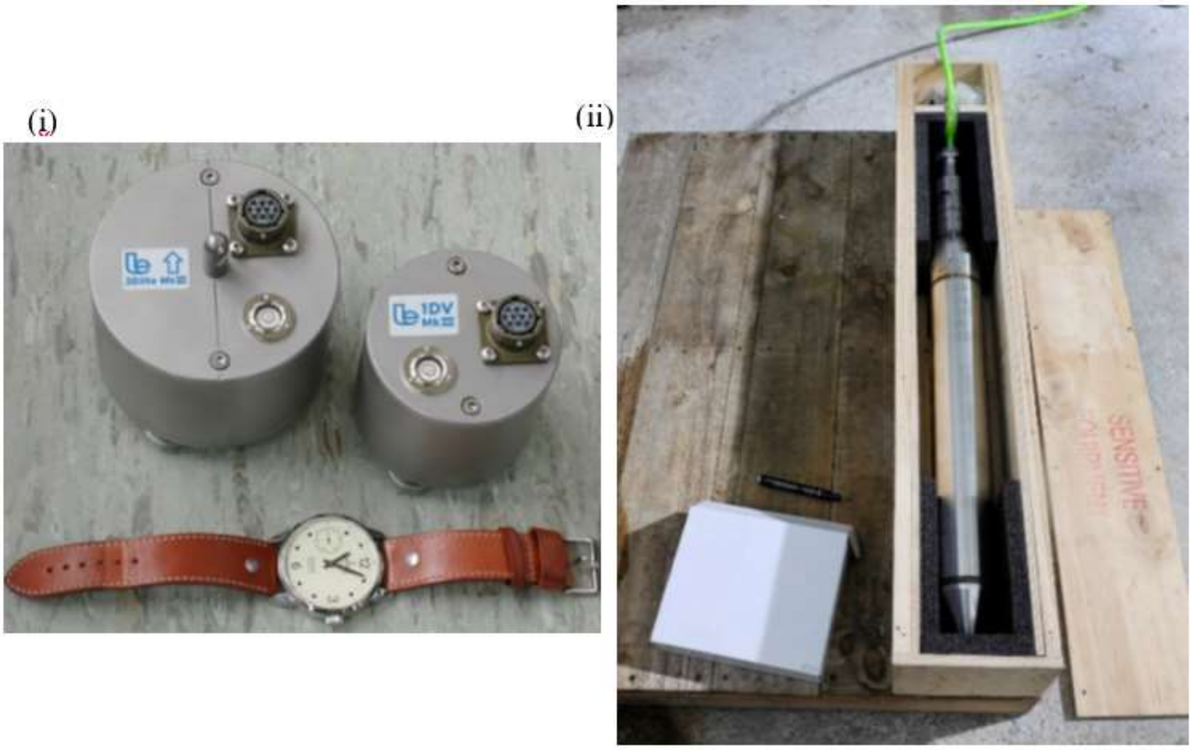


Figure 10: (i) Geophones: a 3D (left) and a 1D vertical (right) short-period seismometer. Figure reproduced from Lennartz-electronic.de (2015) (ii) Downhole seismometer HS-1-LT mini

3.2.1 Seismic network deployment

In theory, the minimum amount of records / readings for the source location determination is four, i.e. three p-arrival times and one S-arrival time; hence the number of sensors (15 channels in total) was considered adequate for microseismic monitoring (Joswig et al., 1992).

Apart from the number of the sensors, the geometry of the network determines the accuracy of location in different directions. The two arrays and the 3-D borehole sensor were to be deployed in three distinct areas forming a triangle. The process of site selection began with choosing potential locations for sensors deployment, as close as possible to Raeterichsboden Lake and within the tunnels under the jurisdiction of Nagra. The suggested locations had further limitations of tunnels accessibility, for instance equipment were not allowed to be deployed in the access tunnels. In addition to the above, the estimated events locations within the triangular area defined by the sensors would be more accurate than those outside the seismic network; hence the sensors were spread as far away from each other as the tunnels allowed.

Physical aspects such as the geology of the area and sources of cultural noise were also taken into consideration for the choice of the seismic network location. As per the first consideration point, the lithology and the existence of faults in the host material are liable to affect the seismic wave velocity anisotropy and the splitting of the surface waves. The location of the seismic network was chosen so that it lies within tunnels of granite (while further away from the sensors, but within the planned monitoring area, there are areas of granodiorite). Exposed faults were taken into consideration for the deployment of the seismic arrays so that the chosen locations were free of fracture zones. As per the second consideration point, anthropogenic noise, i.e. excessive bursts or spike-type, man-made seismic noise, results in high trigger thresholds and therefore poor event detectability. It was deemed necessary therefore to locate deploy the sensors in the least active tunnels.

Taking all the above into consideration I deployed the instrumentation between the 29th and 30th of October 2014 (see plan view in Figure 11). In particular, the two surface arrays were deployed within the Gerstenegg tunnel and the downhole sensor in the vertical borehole "Bohrung 3" (GS 84.041A) with grid ref (46.58°, 8.32°) in the GTS tunnel. The triangular coverage area was approximately 340m^2 (marked with orange dashed lines in Figure 11). This network geometry does not offer good hypocentral coverage by the arrays sensors given that they were deployed in approximately the same depth; a limitation that the borehole sensor would counterbalance. The Gerstenegg tunnel was used neither by KWO nor by Nagra stuff, hence minimising obstruction and anthropogenic noise levels. However, within the GTS tunnel, where the borehole sensor was deployed, the anthropogenic noise levels were expected to be high given the numerous projects that were simultaneously in operation. This was a situation that could not be avoided and it would be part of the project to find ways to mitigate noise in seismic detection.

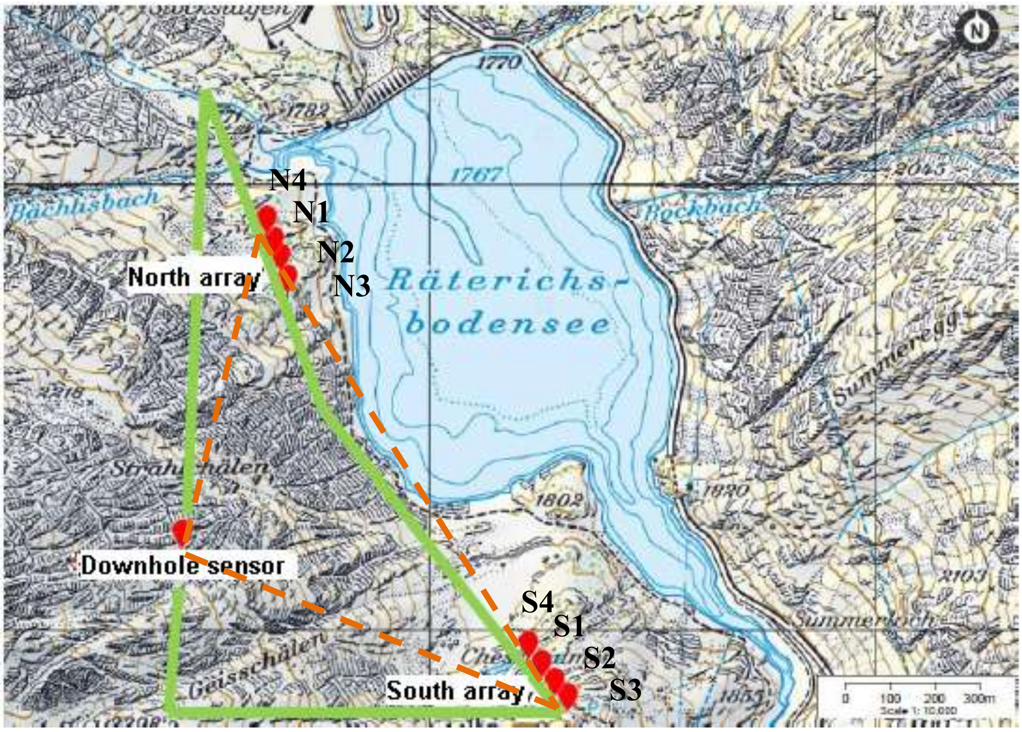


Figure 11: Location of seismic network on map (public.geo.admin.ch & swisstopo, n.d.), where the tunnels are noted with green and the sensors positions with red (South array: S1:4 and North array N1:4, where 4 corresponds to the 3D sensors). The triangular area that the seismic geometry offers best coverage is denoted with orange dashed lines.

3.2.2 Surface arrays

The surface sensors were placed upright on top of circular concrete pillars constructed on the tunnel floor of diameter 30 cm and height 30 cm (Figure 12 (i)). Both arrays had a 135metres aperture while the radial distances between successive sensors were at 45m. The two arrays were approximately at 1km distance from each other.

The sensors were connected to geophone cables which were mounted on the tunnel walls. Then the best possible levelling was achieved by adjusting the three foot screws of the sensors so that the bull's eye bubble located on top of each sensor was centred. With the sensors in place, oriented, and levelled, and with the cables connected on the sensors side, the cables to the data logger side were attached.

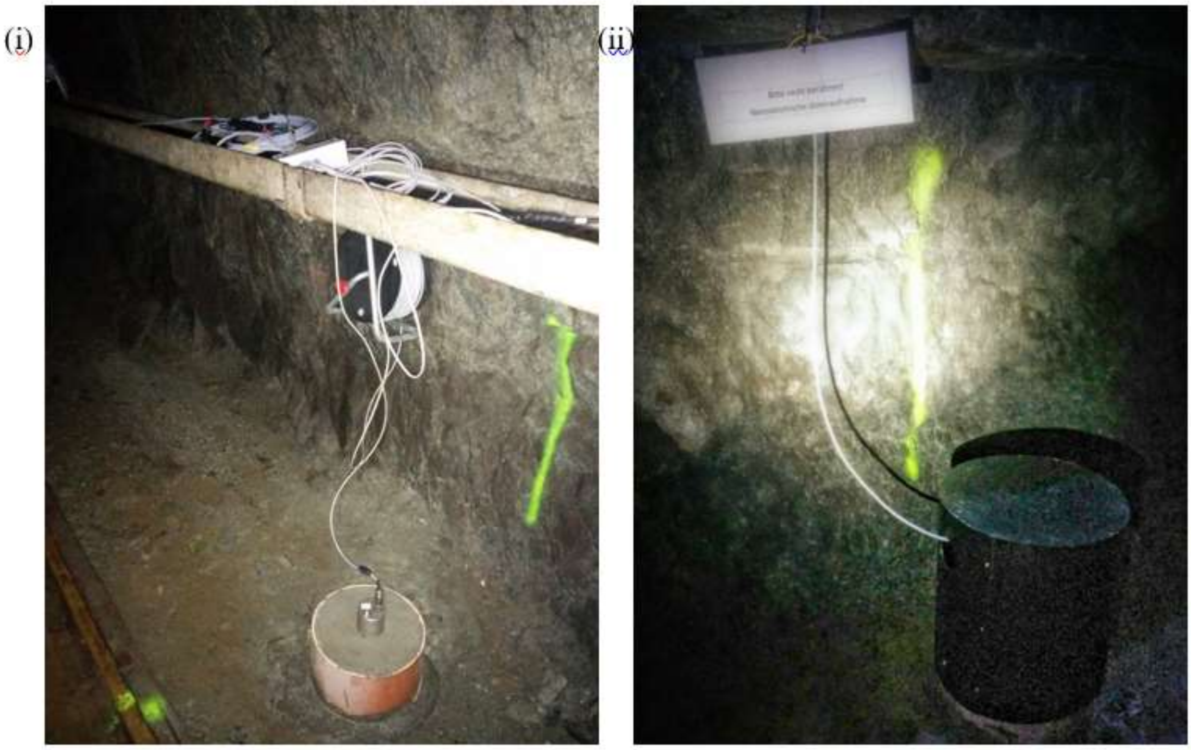


Figure 12: (i) Surface arrays’ seismometer and (ii) black casings for long-term protection

The LE-xD sensors are remarkably insensitive to (gentle) temperature and air pressure variations. Consequently it was not considered necessary to construct complicated styrofoam / copper foil / glass jar contraptions to seal the sensors. They are also supposed to be completely protected against dust and low pressure jets of water. However, and since water leakage (droplets) from the tunnel walls and ceiling was observed, plastic casings and insulating tape for all cable connections were considered essential for the long-term proper functioning of the seismometers.

The data logger (black device in Figure 13) was connected to an external power supply. Considering that the nominal input voltage was 12 V DC and typical current drains at 12 V DC were: LE-1D: 3 mA, LE-3D: 8 mA, 6 channel data logger: 0.17A, power consumption calculations led to a choice of a rechargeable Pb-acid battery of 12V and 38Ah.

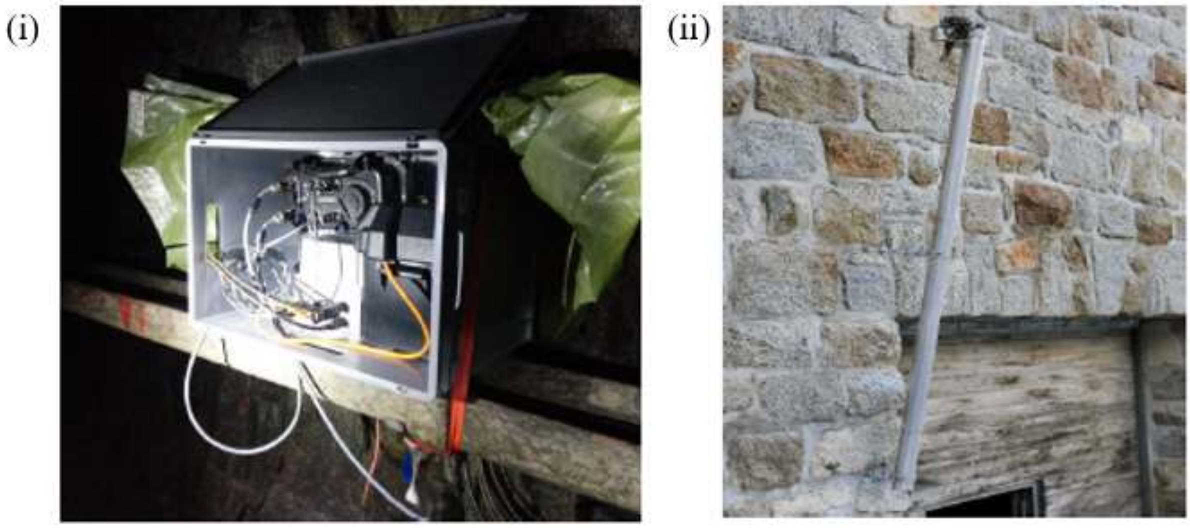


Figure 13: (i) Datalogger, fiber optics central box and battery were placed in a box to be protected from humidity and (ii) GPS antenna

Each of the data loggers was in turn connected to a GPS antenna (Figure 13(ii)), mounted on a pillar outside the north exit of the Gerstenegg tunnel to achieve accurate timings and synchronisation. The GPS signal was transferred in the tunnels via fiber optics (Figure 13).

The measured powered consumption with the GPS use was 107Wh at 13°C. Hence for the proper function of the surface array system the batteries needed to be changed twice per week. On the 5th of December 2014 a special control circuit was attached at the sensors network and the GPS fiber optics converter. The circuit allowed a twice per day synchronisation of the GPS extending the use of the batteries from 3.5 days to 10 days. On the 23rd of December new electricity lines were installed in the Gerstenegg tunnel to stop relying on batteries because the long term viability of this plan was deemed non sustainable.

The exact coordinates of the sensors, as derived by Nagra's surveying team, are given in Table 2. The coordinates are given in CH1903 (Easting and Northing in metres, as measured by the Swiss coordinate system) and WGS84 (latitude and longitude in degrees), and elevation, measured on top of the pillars, is given in metres above sea level. The Swiss coordinate system is a geographic coordinate system used in Switzerland by the Swiss Federal Office of Topography. The geodetic datum CH1903 uses as fundamental point the old observatory of Bern (latitude: 46°57′3.9″N, longitude: 7°26′19.1″E (WGS84)) and corresponds to the coordinates 600,000 Easting and 200,000 Northing. WGS84 is the standard U.S. Department of Defence definition of a global reference system for geospatial information and is the reference system for the GPS.

The three directional component sensor and the data distributor were north-oriented with the use of known survey points and bearings. The survey z-coordinate of the borehole sensor is also presented in the table and is the measurement in the top of the borehole (ground level of the GTS tunnel at that position).

Table 2: Coordinates of sensors (see Figure 11 for their respective locations)

Sensors	**CH1903**			**WGS84**	
	Easting (m)	Northing (m)	Elevation from sea level (m)	Latitude (decimal degrees)	Longitude (decimal degrees)
S4	668222	158947	1743.36	N46.5783	E8.3286
S1	668251	158913	1744.29	N46.5780	E8.3290
S2	668280	158878	1745.20	N46.5777	E8.3294
S3	668309	158844	1746.07	N46.5774	E8.3298
N4	667635	159910	1721.57	N46.5870	E8.3211
N1	667650	159868	1722.32	N46.5867	E8.3213
N2	667666	159825	1723.22	N46.5863	E8.3215
N3	667681	159783	1723.95	N46.5859	E8.3217
Borehole	667445	159197	1729.72	N46.5806	E8.3186

3.2.3 Borehole sensor

Prior to the deployment of the downhole sensor a sand layer was formed at the bottom of the borehole to facilitate good coupling between the borehole sensor and the walls of the borehole. The sensor sat at the bottom of the borehole in a practically vertical position (considering its dimensions and that the inner diameter of the hole is 146 mm) and part of it was surrounded by sand. Borehole depth measurements took place prior to the addition of the sand layer and after (171.40 m and 166.60 m respectively). Lowering a dummy (Figure 14(i)) down the borehole before attempting to lower the sensor convinced us that of the unobstructed passage and retrieval of the sensor.

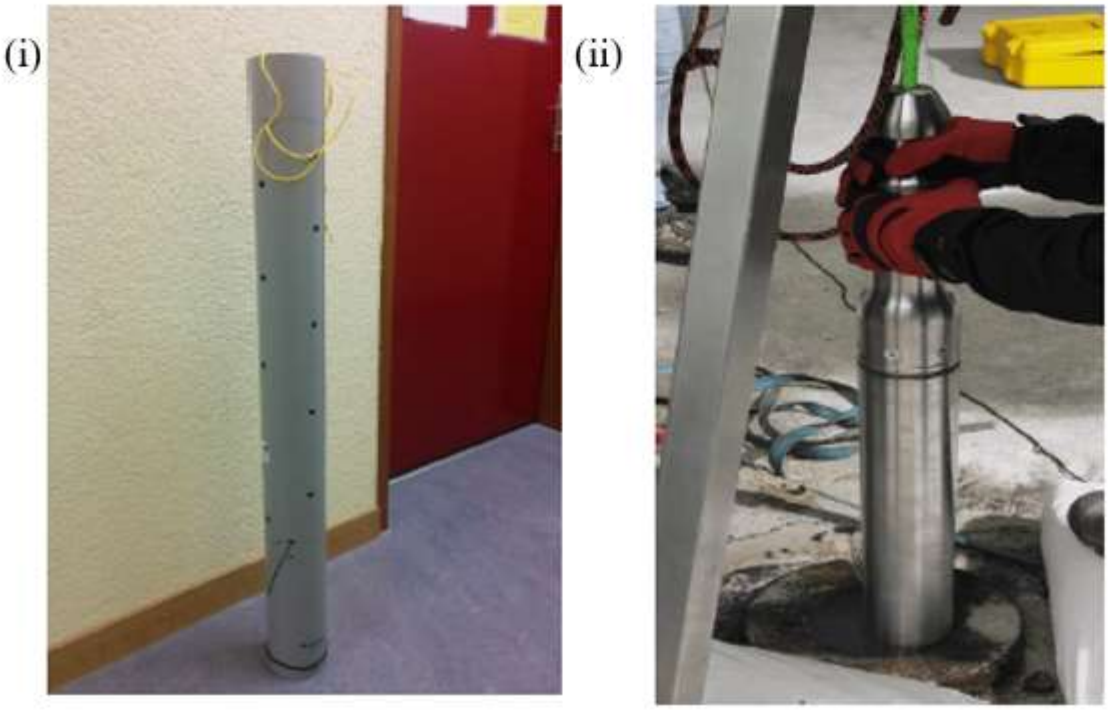


Figure 14: (i) Dummy used to ensure free passage throughout the borehole and (ii) Deployment of downhole sensor

After ensuring the operation of the downhole sensor with the use of a multimeter, the sensor was lowered carefully in the borehole (Figure 14(ii)) and was kept under tension with the use of ropes for three days. Based on the depth of the borehole, its diameter and the density of the sand used, this was the estimated time period required for the sand to settle. Till then, the ropes played the role of a stabiliser to ensure the borehole sensor remained vertical in the centre of the borehole. After that time period an extra cable was fed into the borehole and the sensor assumed its final position.

The downhole sensor was connected to a three channel data logger (B961) which was in turn connected to external power. The borehole lied within the main GTS tunnels which had electric lines already in place hence in this case the data loggers were directly charged by a battery which in turn was continuously connected to a charger and the charger to the mains.

GPS was not possible to be used in the GTS tunnel, therefore timings were kept using the in-built clocks of the data logger. A synchronisation with the GPS once per two weeks was deemed to have a non-linear drift from real time which varied from µsec to msecs, a drift that was considered feasible to be taken into account and corrected on the data of interest. Power black-outs and bad weather conditions that did not allow the synchronisation to happen as planned increased the clock drift at times. See also subchapter 3.5.2 for more details.

Additional calibration checks and a procedure to later allow for the orientation of the horizontal components of the borehole sensor took place: hammer hits were recorded by both the arrays 3D seismometer and the borehole using the same data logger.

3.3 Seismic Data

All desired acquisition parameters were assigned and communicated to the seismometers with the use of a PFC_130 (Palm Field Controller). The parameters used for the continuous acquisition recordings were:

(i) gain was set as high = 32 in the case of the surface arrays sensors and unity for the borehole sensor. Higher gain increases the amplitude of signals, hence increases the SNR making the signals more easily detectable. The borehole sensor, given that it is deployed deep within a borehole does not suffer from environmental and cultural noise hence there was no need for a high gain.

(ii) sampling rate = 250 samples per second (sps).

The data acquisition was initiated on the 4th of November 2014, 4 days prior to when the water level in Lake Raeterichsboden was at its lowest for the 2014 drainage period (see Figure 5). The seismometers were set to record data continuously. Raw data were stored digitally as counts in hourly files, with counts being a measure of electric current (voltage). The sampling rate was 250 Hz (i.e. sampling time step = 0.004s) which resulted in 900,000 data points per channel per hour, 15 channels in total.

Given that the data collection period was approximately 3.5 years, it was early understood that the visual check would only constitute a way of understanding the signals recorded in the area and that a more automated way of processing would be needed for the actual data analysis and features identification. A suitable platform for signal analysis was believed to be MATLAB, mainly because of the vast database of forums and relevant online information availability.

3.4 Pre-processing

The data acquired were in need for pre-processing, i.e. (i) the transformation of raw data (measured in counts) to ASCII format so that they can be read and further processed using MATLAB, (ii) conversion of the counts in velocities and (iii) the formation of data series (timings assigned to all data points and formation of data series matrices). All relevant algorithms for this stage can be found in **Appendix A**. Additional algorithms were created along the progress of the project to make the data series more easily manipulated by the relevant detection or onset time picking algorithms. For instance, it was observed that the raw data collected from the South array had a time lag in the start times of the hourly records. Also, the data record start time was floating, with a non-steady change period. To address this issue the ‘fixStarts1’ script was created.

Another script that was deemed necessary was the conversion of the timeseries from cell arrays that were used for the initial codes to structure arrays that were necessary for the use of Archie-West supercomputer. Archie-West was used in order to accelerate data analysis by taking advantage of multicore processors through parallel computing. NpD algorithm is an example code that takes full advantage of the processing power of Archie-West; i.e. parallel for-loops (parfor command) and special array types (structure arrays). All aforementioned algorithms can be found in **Appendix A**.

3.5 Maintenance and operation

There were several field visits in the GTS. The first one (September 2014) was a reconnaissance visit; the second (22nd October - 2nd December 2014) was mainly for the deployment and monitoring of the network to ensure smooth operation and efficient and undisturbed data acquisition. Further works, such as the protection of the array sensors, the survey of the sensors' position coordinates, checks to define the desired frequency of the borehole sensor's GPS synchronisation and the frequency of the surface arrays' batteries and data disks change, were performed. During the third and fourth visit to the GTS (16th till the 23rd of March 2015 and 8th till the 16th of June 2015) the main activities that took place are listed as follows:

- Correlation of specific patterns, which were repeatedly identified in the recordings during the preliminary analysis, with possible man-caused activities; time was spent at the sensor locations recording various activities/ambient noise and noting the corresponding timings (see also paragraph 3.6). The purpose of the noise sources investigation was to understand the acquired dataset and determine the needs (pre-processing, signal analysis techniques / tools) for my project.
- Inspection of the hardware connections and the conditions under which the sensors were operating. Main reasons for concern were humidity in the tunnels, rockfalls or displacements / disorientations of the sensors from potential KWO or Nagra works during my absence. During these inspections the cables of the surface arrays were replaced (initial positioning as shown in Figure 12) so that they were hanging from strings at a safe distance from the rest of the electricity cables to minimise effects of potential electromagnetic interference.
- Repetition of all the necessary operations that, in my absence, the Nagra staff undertook; these were repeated to ensure the correct operational process (GPS

synchronisation of the downhole sensor, disks change and download; see also paragraphs 3.5.1 and 3.5.2.1).

3.5.1 Data downloading, transfer and back-up

In the beginning of the project the recorded data were stored in two 8Gb disks (placed within each data logger) which needed to be replaced after six weeks of data recordings, and then downloaded and transferred via ftp.grimsel.com (Nagra's cloud server). For the long-term operation this procedure was substituted by the data transfer with the use of a cloud server (weconnect.westermo.com). In particular, this solution required the installation of Ethernet cables (for the upload of the data to the cloud service) and specialised software, i.e. RTPD and RTCC by REFTEK (for the transmittance / download at a university computer). This solution allowed for central data gathering and almost real-time back-up on the university's facilities and hence minimized the risk of data loss. Since August 2015 the real-time communication with all the data loggers was effective. The simplified schematic display of the final network map for the connection between all data loggers in the GTS and the Strathclyde computer can be seen in Figure 15.

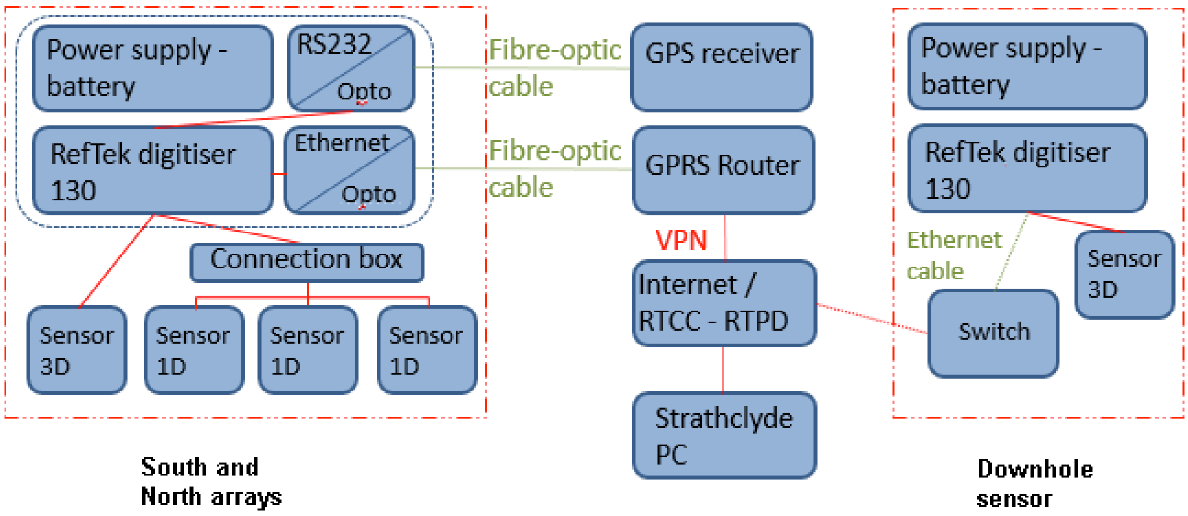


Figure 15: Final network map of the communication network

3.5.2 Downhole sensor

3.5.2.1 GPS synchronisation malfunction

As already mentioned GPS was not possible to be used in the tunnel. As the installation of new lines was too expensive to be considered as an option, timings were kept using the in-built clocks of the data logger, re-synchronised every two weeks. The procedure followed was to disconnect the data logger from the sensor, connect it to the GPS antenna and take it outside from the tunnels until it connected with the satellites signal and reinstall it to the sonde again. The times would be noted down and the internal clock health files would inform of the time drift.

Past experiments of Pytharouli, S. suggested a non-linear drift from real time which would vary from *μsec* to *msecs*, a drift that was considered feasible to be taken into account and corrected on the data of interest. Indeed in the beginning that was the case but as power black-outs and bad weather conditions did not allow the synchronisation to happen as planned the clock drift at times increased to seconds.

3.5.2.2 General malfunction

During the visual inspection of the first 3 months of data recordings, the recordings of the E-W horizontal channel of the downhole sensor showed consistently high frequency and low amplitude noise. Some tectonic earthquakes were still noticeable as can be seen in Figure 16 but lower magnitude ones could be masked.

Initially, the downhole cable was coiled around a drum, which stood on a tripod. After discussions with colleagues I thought that the high noise levels were a result of the electromagnetic (EM) interference of the power installation and the coiling of the cable, hence I readjusted the cable's positioning, however without any positive effect on the recordings. Tests using a multimeter were conducted during a GTS visit and after discussions with the manufacturing company and guided testing of the sonde's components, we concluded that the internal components of the 3rd channel were damaged. Hence, the data acquired by the BH sensor were not included in further analysis.

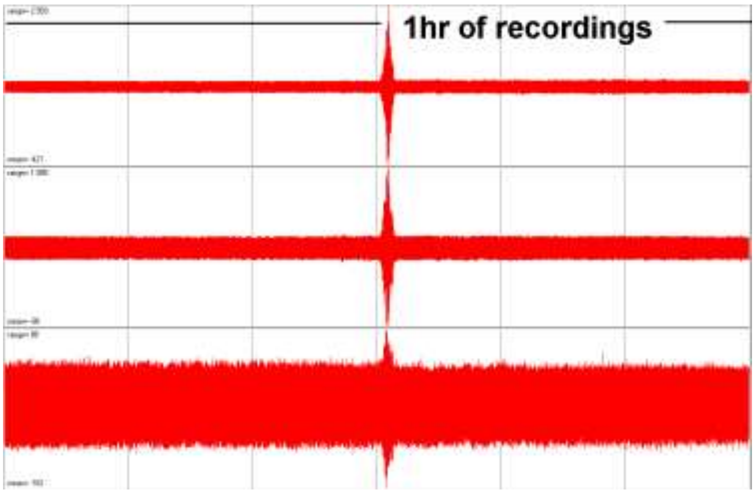


Figure 16: Illustration of the amount of noise included in the borehole channels. Even with extensive filtering the borehole sensor did not seem to have been triggered by any of the events that the seismic arrays were detecting.

3.5.3 Gaps in the Data recordings

Gaps in the data were unavoidable during the project's duration and were mainly caused because of power shortages in the tunnels. The greatest in duration data gap, of approximately one month, occurred while the connection to the cloud server (Westermo) was in progress (June - July 2015). During cable routing routes on February 8th 2016 in Gerstenegg tunnels the sensors were found to have been knocked off their positions. Ten days of data were discarded as a result of this accident and even after the sensors were put back in place and oriented (the first 2 days), some recordings were still showing perturbed signals. Testing of post-accident data showed there was no permanent damage to the sensors.

3.6 Preliminary investigation of data noise

A preliminary analysis took place in order to achieve a first understanding of the data obtained and check the effectiveness of the microseismic monitoring network. Signal processing MATLAB algorithms, for the spectral analysis of the data were developed and the data (04.11.2014 – 11.12.2014) were scanned to identify any microseismic events related to the drainage of Raeterichsboden Lake. Most of the initial attempts led to identification of tectonic earthquakes, as that presented in Figure 17, located further away from the Grimsel area at distances > 20 km.

These earthquakes were verified using the published earthquake lists of the Swiss Seismic Network. The identification of small tectonic earthquakes demonstrated that the sensors were able to successfully record small seismic events even at long distances (1.6 M at 127 km).

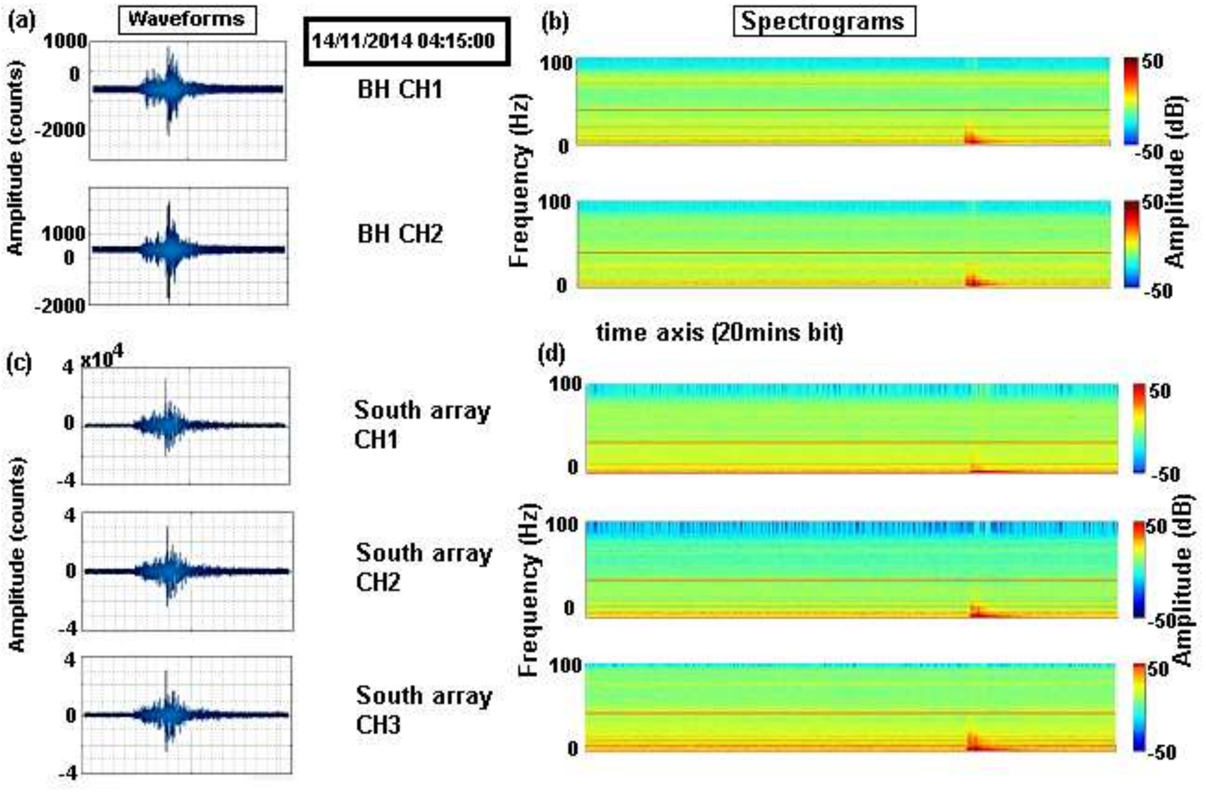


Figure 17: A 3.1 M earthquake as recorded by Strathclyde's microseismic network. The epicentre of the earthquake was 90.48 km away from GTS (Swiss Seismic Network). Waveforms as recorded by the borehole sensor (a) and the South array (c). The representation of the vertical components (CH1-3) in the time – frequency domain (spectrograms) show clearly the occurrence of the event (b, d for the borehole and South array respectively). The colour scale visible in the spectrograms indicates the amplitude of a particular frequency at a particular time with red being the highest magnitude and blue the lowest.

The first two weeks of data were visually checked in great detail to identify noise patterns. From this visual check, it is very clear that there are variations in the background noise, even within a single hour of data. Figure 18 shows spectral plots for four 20 minute intervals, all taken within a single day for a single sensor (at the South array). Apart from the mains power interference which is visible in all plots, Figure 18a shows high amplitude, fluctuating noise levels up to 10Hz, and visible harmonics of the mains power (i.e. lines at 17.5&25Hz); Figure 18b has a lower amplitude, but still intense background noise up till 8Hz; Figure 18c has noise between 4 and 8Hz, with distinct harmonics from the mains (17.5&25Hz); and Figure 18d shows very little noise, almost all of which is at frequencies less than 5Hz.

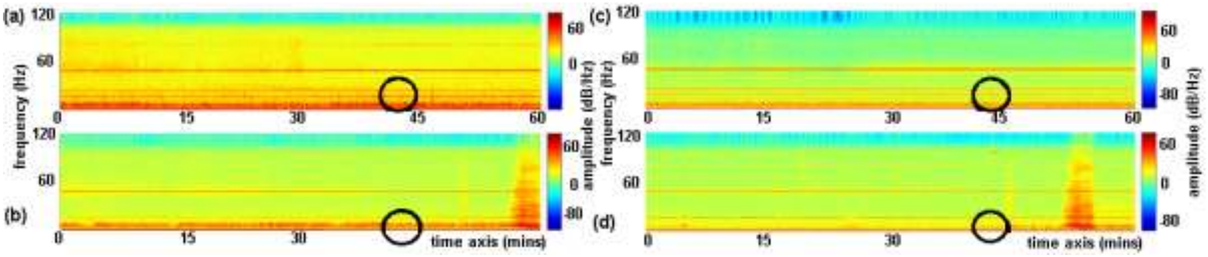


Figure 18: Recognition of different levels of background noise.Here you can see 20mins sections for 4 different hours within a day at the sensors at location 1. Figure a shows high, fluctuating noise levels up to 10Hz, visible harmonics of mains power (17.5&25Hz), figure b has less amplitude but still intense background noise up till 8Hz, figure c has noise from 4 to 8Hz, distinct harmonics of mains(17.5&25Hz), while figure d is the least noisy from the files shown, with background noise at only 5Hz.

Throughout the visual check of the data, similar reoccurring patterns were observed at many recorded waveforms. See some typical encountered patterns in Figure 19.

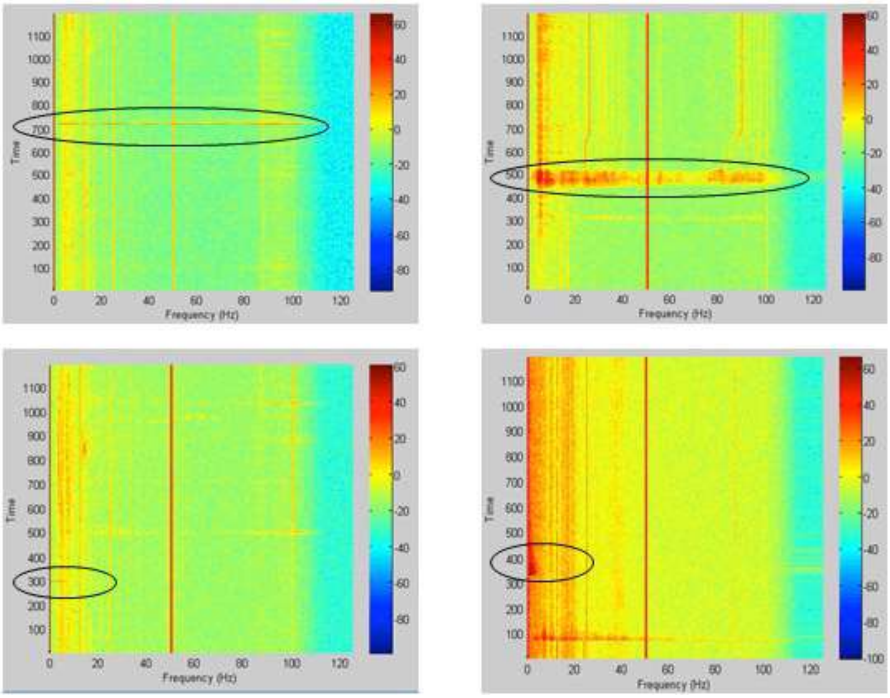


Figure 19: Patterns encountered in the recordings as seen in the spectrograms.

It became obvious that in order to effectively distinguish valuable data, containing microseismic information and geological information of the medium, from random noise and the response of the sensor/recorder, the potential noise sources had to be identified and are further analysed as follows.

3.6.1 Noise related to hydroelectric activity

KWO is one of the leading hydroelectric companies in Switzerland with 9 power plants containing 26 turbines in total, 11 of which are in the area of the GTS and potentially affect the recordings. The hydroelectric stations in the vicinity of the GTS are presented in Table 3. The locations of the hydropower stations, alongside those of the seismic arrays, are shown in Figure 20(a) and an example of machines in Grimsel 2 power plant is shown in Figure 20(b). The turbines are set on and off on a water-demand basis and do not operate on steady hours nor fixed durations.

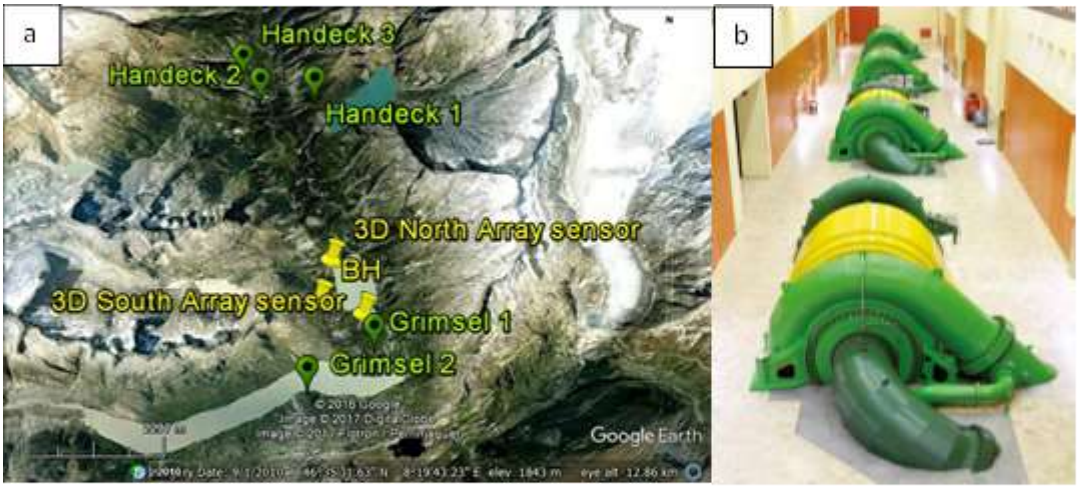


Figure 20: (a) Locations of the hydroelectric stations described in Table 1, the microseismic arrays and the borehole sensor. (b) Machine room Grimsel 2 power plant. (http://www.grimselstrom.ch/electrical-energy/power-plants-and-dams/power-plants/).

Table 3: Description of the hydroelectric stations for which data were provided by KWO (http://www.grimselstrom.ch)

Handeck 1	Water for the power plant comes from the Gelmer Lake situated 550 metres above the plant. The power plant has 4 Pelton turbines. Pelton turbines have natural frequencies of about 580Hz (Shelke, 2017).
Handeck 2(HA2)	An underground plant that is used to produce peak energy and to regulate the grid. It also contains 4 Pelton turbines.
Handeck 3	Complex power plant that fulfills a number of different functions. Water can be moved from Lake Raeterichsboden to the Handeck reservoir or sent over to the Gadmen valley. Water can also come from the Gadmen valley and then be pumped up to Lake Raeterichsboden. The runner wheels can be used as both pump and turbine.
Grimsel 1	A double Pelton turbine (Oberaar machine group) uses the gradient between the Oberaar and Raeterichsboden Lakes (around 530 m) and a 32 MW turbine (Grimsel) uses the gradient between the Grimsel and Raeterichsboden Lakes (around 140 m).
Grimsel 2(GR2)	Contains 4 machine groups, each with a pump and turbine wheel on the same shaft, and uses the gradient between the Oberaar and Grimsel Lakes. Water is pumped, using surplus grid electricity, from the Grimsel 2 pump storage plant to a lake at higher elevations where it is stored to produce electricity at a later date.

Due to the proximity of some of the hydroelectric plants, I investigated whether noise caused by the pumps activity in the area, had a straightforward effect on the data, that could be removed. A short period of data for the hydroelectric stations described in Table 3 was provided by KWO for comparison with the recordings.

Figure 21(a-f) shows the seismic data for four full days between the 8th and 14th of November as recorded by the vertical component of the 3D sensors of the North and South arrays. Also plotted on the graphs are the data provided by KWO. GR2 corresponds to Grimsel 2 pumps 1, 2 and 3 (Ma1, 2 and 3 respectively), while HA2 corresponds to the pumps of Handeck 2. Negative values correspond to pumping, positive values represent energy production. All values higher than -500 and less than 0 imply operation of the transformer only (i.e. no generator). GR2 has pumps/turbines that rotate (i.e. there are positive and negative values) which will produce significant ground vibration.

Visual analysis of Figure 21 shows that there is no clear one-to-one relationship between the pump data and the seismic recordings. However, some conclusions can be drawn. There seems to be only one pump that has a consistent effect on noise recordings – machine 1 at Grimsel 2.

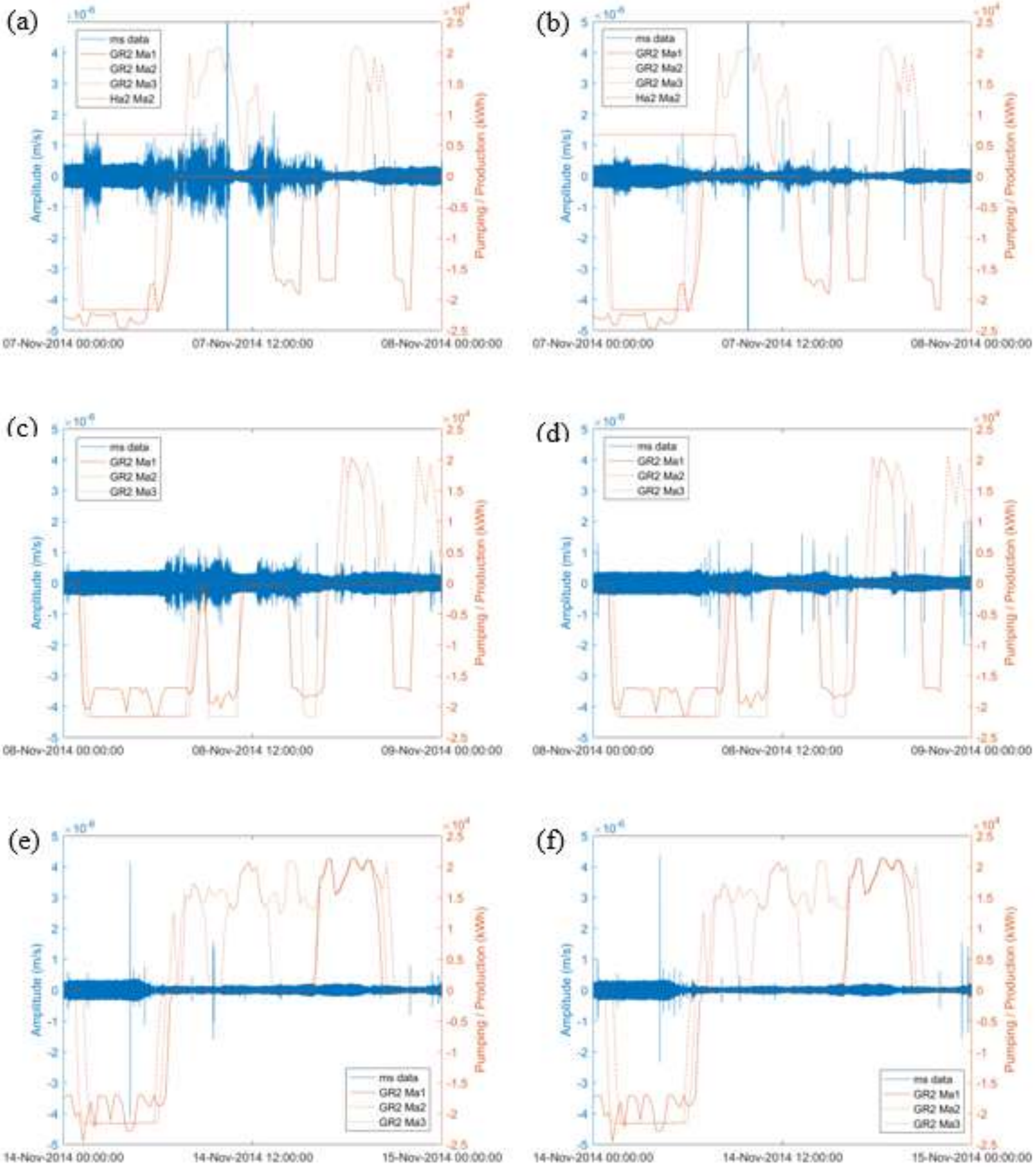


Figure 21: Full days of seismic data: 7th November of 2014 (a, b), 8th November of 2014 (c, d), 14th November of 2014 (e, f) as recorded by the vertical component of the 3D sensor of the northern (a, c, e) and southern (b, d, f) arrays and corresponding pumping data from KWO.

During most periods of pumping and production from this turbine the seismic data become far noisier. This is denoted by the change in the amplitude of the seismic recording visible in all plots of Figure 21. For example, in Figure 21(e) the amplitude of the seismic recording is almost double during hours 0:00 and around 06:00 (when GR2 Ma1 is active) compared to the amplitude during the hours between 06:00 and 15:00 (when GR2 Ma1 is not active). There may also be an effect of switching off and on this pump, evident by a gradual decrease or increase in the amplitude of the seismic recording respectively (as seen in Figure A4(d), for example). This visual analysis of the data in Figure 21 shows that it is unlikely that KWO’s operational data could be used to support the removal of high amplitude seismic noise. Even in the case of Grimsel 2, machine 1, a comparison of different days shows that the effect is not straightforward.

3.6.2 Noise related to construction activities

To increase water flow capacity, between the power plant in Handeck and Raeterichsboden lake, a new water bearing tunnel was excavated 50 m south of the existing tunnel (box a in Figure 22). This was the main reason for the 1st drainage of Raeterichsboden lake, in the winter of 2014-2015; the construction of the inlet and connection with the lake. Significant vibrations were observed at a monitoring station (location shown on map with a red asterisk) between 28.01.2015 and 21.02.2015 with maximum velocities of nearly 63 mm/s (LASMO Team, 2016). A small niche was excavated approximately midway in Gerstenegg tunnel (box b in Figure 22) between the 10th of February and the 5th of March 2014. A drill-and-blast excavation of an access tunnel was initiated south of Grimsel Pass in the summer of 2015, followed by the excavation of a cavern in winter 2015/16 (LASMO Team, 2016). The blast induced ground vibrations varies greatly depending on the parameters of rock mass, explosive characteristics and blast design (Khandelwal and Singh, 2006). TBM vibrations during excavation, for instance, have been reported to have dominant frequencies in the 300-500 Hz range (Mooney et al., n.d.).

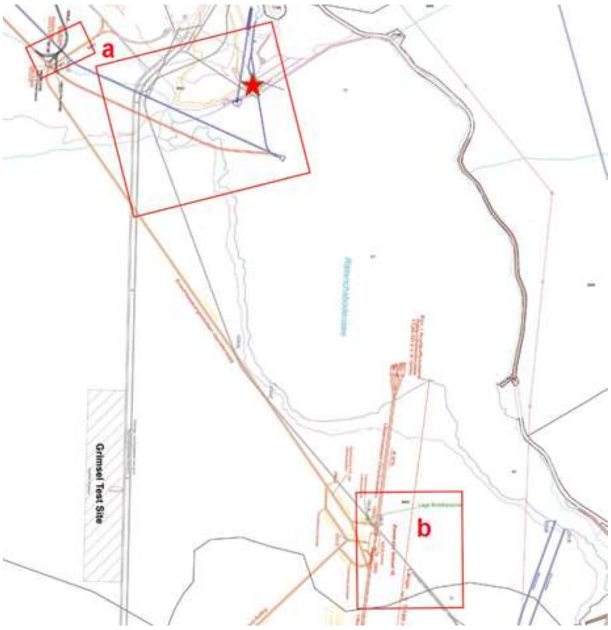


Figure 22: Gerstenegg area with location of man-made far-field perturbations. Figure adapted from LASMO Team (2016).

3.6.3 Noise related to water induced seismicity

Another source of seismicity that could be regarded as noise for the purposes of my experiments is the water induced seismicity and in particular seismicity induced by forced fluid injection at depth, i.e. hydrofracking (Phillips, 2000; Zoback and Harjes, 1997). The In-situ Stimulation and Circulation (ISC) test was a geothermal research project directed by the ETH Zurich aiming to improve the understanding of geomechanical processes underpinning permeability creation during hydraulic stimulation and related induced seismicity. The project started in the second half of 2015 site and consisted mainly of intensive field work, such as drilling and borehole investigations in the GTS tunnels (Figure 3) while drilling continued in the beginning of 2016. The stimulation and fluid circulation phase in pre-existing and stimulated fractures followed in 2017. The pulse used in hydrofracking can vary from 100 to several 1000s Hz.

Naturally induced pore pressure changes which lead to seismicity have also been reported by researchers: by seasonal groundwater recharge (Roth et al., 1992; Saar and Manga, 2003), seasonal snow melt or variations in precipitation (Jiménez and García-Fernández, 2000; Kraft et al., 2006; Lorraine et al., 1997). All the above types of triggered earthquakes are usually of small magnitudes (M<2; Husen et al., 2007).

3.6.4 Noise related to large seismic events

The Swiss Seismological Service (SED) at ETH Zurich monitors earthquake activity since 1975 and currently has one of the densest seismic networks worldwide. The GRIMS station is the SED's closest station to the GTS (Grimsel Pass) and it includes a broad-band station and an accelerometer. The SED data are freely available on seismo.ethz.ch website. Currently, the lowest magnitude of a seismic event that can be reliably detected by the SED Network and pinpointed across Switzerland is 2.0.

Figure 23a, reproduced from the SED website presents historical data for the recorded earthquakes in Switzerland between 1975 and 2014. The earthquakes are represented by circles of different sizes depending on the local magnitudes of the events. Figure 23b is the cross-section AA' (noted in Figure 23a) showing the distribution of the events in depth. From the orange circles, which flag the GTS area, we can observe that, in particular during the 1975-2014 period, earthquakes of magnitude up till $3M_L$ were detected. The vertical cross section documents that during the same period the earthquakes in the GTS area reach up till 12km depth.

Water induced seismicity can also be triggered by the passage of low-frequency, large-amplitude surface waves of distant earthquakes (Husen et al., 2004; Prejean et al., 2004) hence it is valuable to be aware of the tectonic occurrences in the area of experiments.

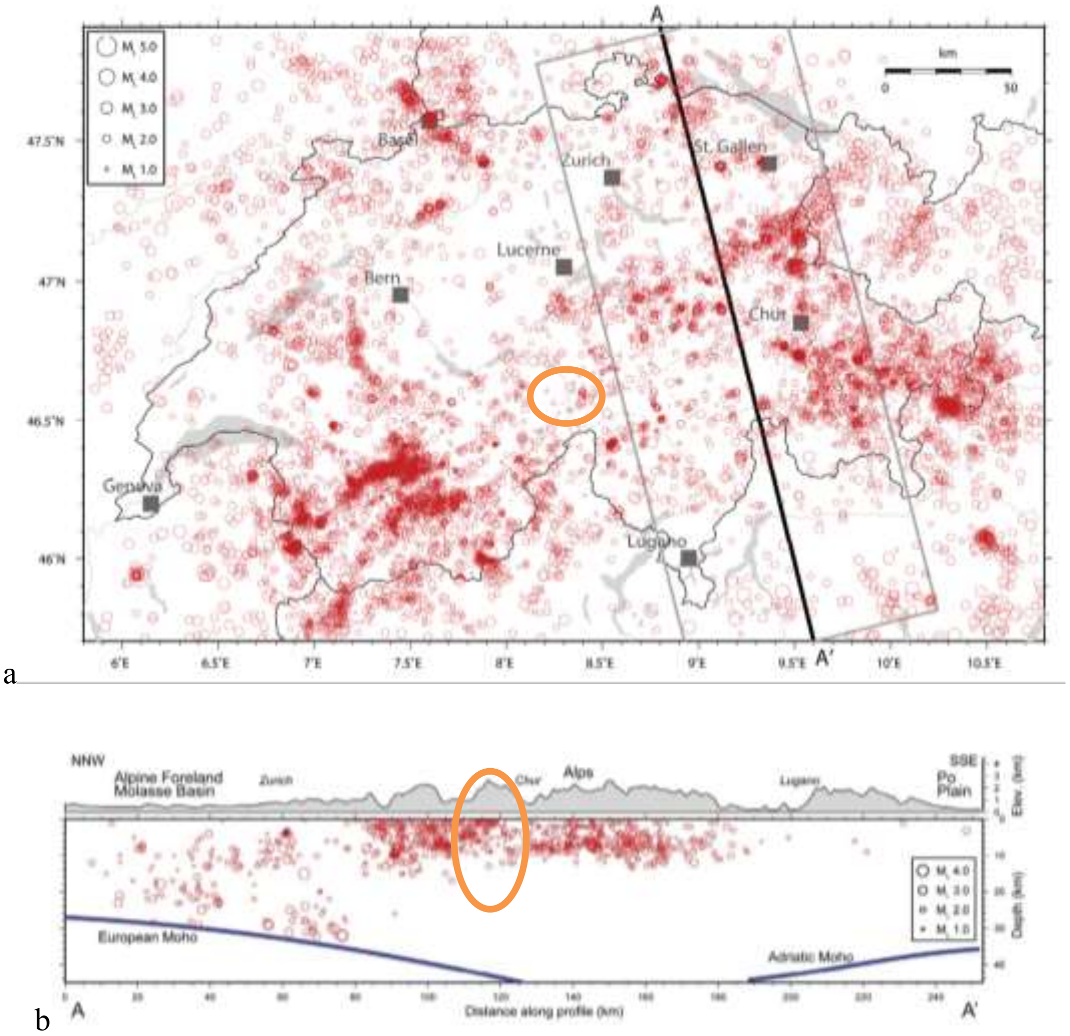


Figure 23: Figures reproduced from www.seismo.ethz.ch. (a) Earthquakes with a magnitude of 1 or more in Switzerland between 1975 and 2014. The size of the circles indicates the local magnitude (M_L) of the earthquakes. The thick black line shows the location of the deep cross section (see b; only quakes within the gray rectangle were used for the profile). (b) Vertical cross section through Switzerland documenting the depth distributions of earthquakes. In orange circes is surrounded the area around the GTS.

3.6.5 Noise related to landslides

The reported frequency content of landslide induced signals varies with study. More specific, in the cases of rockslides and during failure the frequency content, as per Wust-Bloch (2010), varies between 10Hz and 75Hz with most of the seismic energy concentrated above 20Hz. Senfaute et al. (2009) reported that signals with frequencies between 100Hz and 1000Hz are observed before crack formation, while the actual rock fall events produced signals with frequencies below 100Hz. Got et al. (2010) recorded signals with frequency content below

120Hz with dominant frequencies around 25Hz, and between 40Hz and 50Hz, and Levy et al. (2011) reported frequencies between 10Hz and 80Hz for rock thrown manually from the cliff. Rockslides have frequency content below 100Hz with few reported events above 50Hz (Helmstetter and Garambois, 2010). Debris flows have frequency content between 10Hz and 150Hz with larger events producing lower frequencies (Huang et al., 2007). There are also some studies that reported aseismic landslide deformations (e.g. Spillmann et al. (2007)).

Small seismic arrays are able to record landslides kilometres away from their deployment area they were meant to monitor. Walter et al. (2012), for instance, reported that seismic arrays consisting of short-period (1Hz) seismometers were able to record a 15,000m^3 rock fall 5,000m away from the monitoring area, while Moran et al. (2008) reported that a large rockfall (50,000m^3 to 100,000 m^3) was recorded by two seismic arrays deployed 400m and 13,400m away from the location of the event.

The waveform of a landslide recording in Washington as recorded by two sites is shown in Figure 24. The rockfall, which corresponded to a shallow earthquake of 3.1 magnitude, and the associated recordings from two different sites: the SEP with a short-period 3D seismometer, situated at 0.4km distance from the source and the CDWR with a broadband seismometer, situated at 13.4km, consisted of complex, large amplitude signals of several minutes duration as can be seen below.

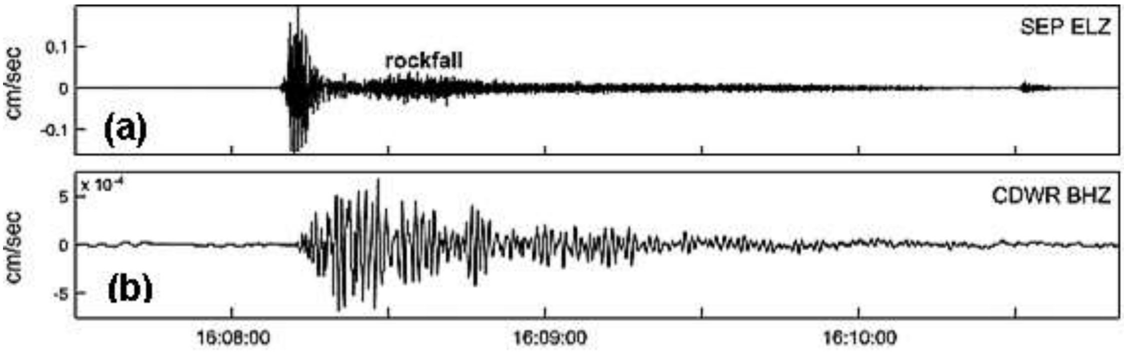


Figure 24: (a) Low-gain vertical-component record from station SEP. (b) Broadband vertical-component record from station CDWR.

3.7 Network design analysis

Not every seismic network design serves equally well for all applications and the intended goal of the network should inform the design. In this PhD there were a number of restrictions which influenced greatly the developed methodology.

One of them was the budget restriction which resulted in using specific instrumentation; the borehole and seismic arrays were already purchased for landslide monitoring experiments. The ideal instrumentation fit for detection of microseismicity would be able to capture a frequency range between 1 and 10,000Hz. This is because the bandwidth of the instrument needs to capture a robust estimate of the spectral plateau that includes the corner frequency and the high-frequency decay. To be able to cover such a wide range of frequencies a combination of sensors such as the one attempted for the present study, i.e. typical downhole 15 Hz geophone with flat responses in the expected frequency ranges for events between M_w -3 and M_w -1 (the IAGE sonde was 4.5Hz) as well as 4.5Hz near-surface arrays sensors to detect higher magnitude (>M_w-1), lower frequency events, are usually preferred.

Microseismic events of -2M_L may have 500Hz corner frequencies and therefore the corresponding sample rates should be 4kHz (8 times greater; rule of thumb). In this case, and because:

- the experiment was meant to last for years,
- one hour of raw data with sampling rate of 250Hz results in 34.5MB,
- the storing facilities were some terabytes in a local University PC,

I decided to use the 250Hz sample rate out of the 2 settings that I had as options (the other being 1000Hz). Although the majority of microseismic studies suggest using 2-4kHz for short-term monitoring, when these relate to long-term continuous monitoring STA/LTA triggering is preferred or the sampling drops to lower rates, e.g. 500Hz (Goertz-Allmann et al., 2014). The latter seemed like a practical solution albeit I would miss locating microseismic events of small magnitudes; smallest detected event was -1.2M_L. This was expected (see also Chapter 3.2) from the literature review I conducted which predicted that sensors with 80Hz flat response maximum detected frequencies of 31Hz which translates to magnitudes of up till -1.49M_L.

Another restricting factor to the accuracy of estimated locations is the geometry of the network used. The network configuration used in this study was the best solution given the restrictions of tunnels availability and accessibility. The malfunction of the borehole sensor hindered the accurate locations of events in the E-W direction since without it there was no longer enough lateral spread of sensors. Another limitation of the network geometry because of the borehole sensor malfunction is the not so good hypocentral coverage by the arrays sensors given that both arrays were deployed at approximately the same depth. In order to investigate the number of sensors that would be able to capture an event should it occur in different locations at different angles around the sensors, InSite seismic processing software was used.

InSite's analysis tool, that determines shadow space, estimates the number of sensors visible for each point of a user-defined search grid, using direct line of sight through solid material (Itasca Consulting Group, 2015). In other words, the parameter that affects event detectability depending on the source orientation in this case is the tunnels. The output model of this investigation that examines shadow space in a 4km 3D search grid is presented in Figure 25. Both Gerstenegg and main access tunnels were used in this model (shown in dashed lines in the figure), while the sensor arrays are placed at the eastern wall side of the Gerstenegg tunnel. It is evident from the model that Gerstenegg tunnel is acting as a barrier and shadows the area on the west and between the two arrays, reducing the maximum number of receivers that can record an event should it occur there to 4 (light green colour). There is only a thin corridor within which just 3 sensors can record an event (cyan colour). There is also a shadow caused from the main access tunnel, rendering all events occurring to the west of the tunnel (GTS area) more difficult to be recorded by the arrays sensors, i.e. maximum number of receivers that can record these falls to 4 for a distance of at least 2km from the arrays. As is derived from the model, there is no shadowing caused on the east side of Gerstenegg tunnel and all 8 sensors can record the events. This can be explained from the orientation of the sensors (all are oriented towards the North) and the relevant position of the tunnel (gap) which is to their west.

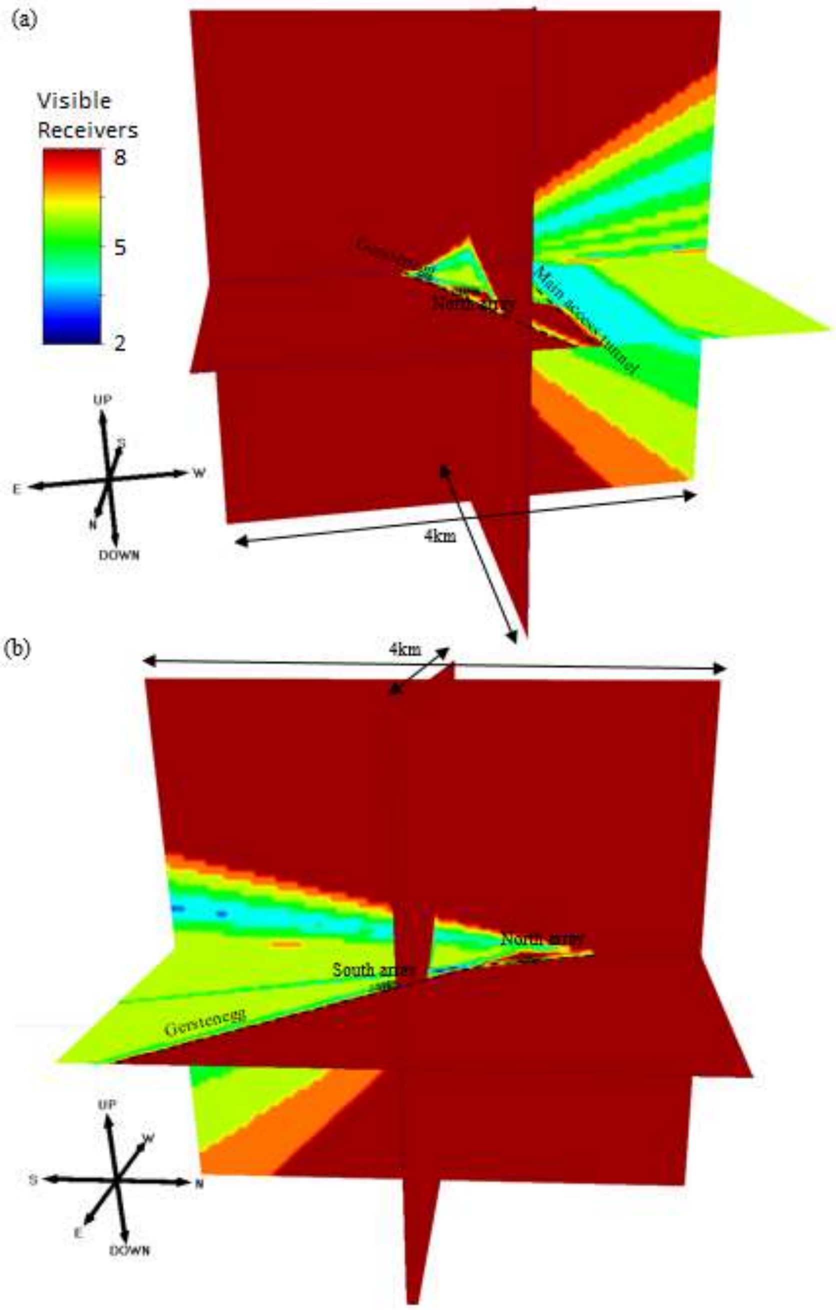


Figure 25: Shadow space for network geometry: Gerstenegg and main access tunnels (in dotted lines) act as a barrier shadowing the aligned receivers; a and b show the EW and NS views respectively of the 3D model. The color scale demonstrates how many receivers are affected by the shadow caused by Gerstenegg and main access tunnels.

The shadow space modelling was further expanded in 3 additional different cases of network geometry (Figure 26), i.e. I assumed the receivers were situated on the west side of the Gerstenegg tunnel (b; left and right hand side show the same model from a different angle: N-

S and E-W) and on the east and west side of the main access tunnel (c and d respectively), and compared the caused shadow to that of my experiments' network geometry (a; east side of Gerstenegg tunnel). The search grid is set once more to a 4km 3D grid space. The model outputs demonstrate that the network geometry (a) causes shadow to less areas than (b). That can be explained as in case (a) both Gerstenegg and the main access tunnel limit the receivers visibility to the west side of the array while in case (b) Gerstenegg causes some shadow to the east of the receivers and at the same time the main access tunnel limit the receivers visibility to the west. In (b) it seems that the only area that is less affected than in (a) is the area to the southwest of the south array.

Similarly, when the receivers are situated on the east side of the main access tunnel (c), their detectability is affected slightly less than if they were on the west side of the same tunnel (d) as in both cases the shadow caused to the west of the arrays is of similar impact while in the latter case there is some additional shadowing caused to the northeast and the southeast of the receivers (influence of the main access). It seems that when the receivers are situated in the main access tunnel the area between Gerstenegg and the main access is no longer shadowed as when they are placed within Gerstenegg. The area to the west of the main access tunnel is always shadowed in all cases tested with minor differences as per the intensity of the effect. Therefore, it seems that the location chosen for the arrays to be placed suited the needs of my PhD as it provides good coverage to the northeast of Gerstenegg (where Raeterichsboden lake lies). My work wold have potentially benefited if further arrays were to be deployed on the main access tunnel (to cover the area between Gerstenegg and main access tunnel). This modelling is using straight lines with the dimensions of the tunnels but remains an approximation of reality, since the tunnels network is much more complicated and caused shadows are believed to extend further and impact on the recording capabilities of the sensors.

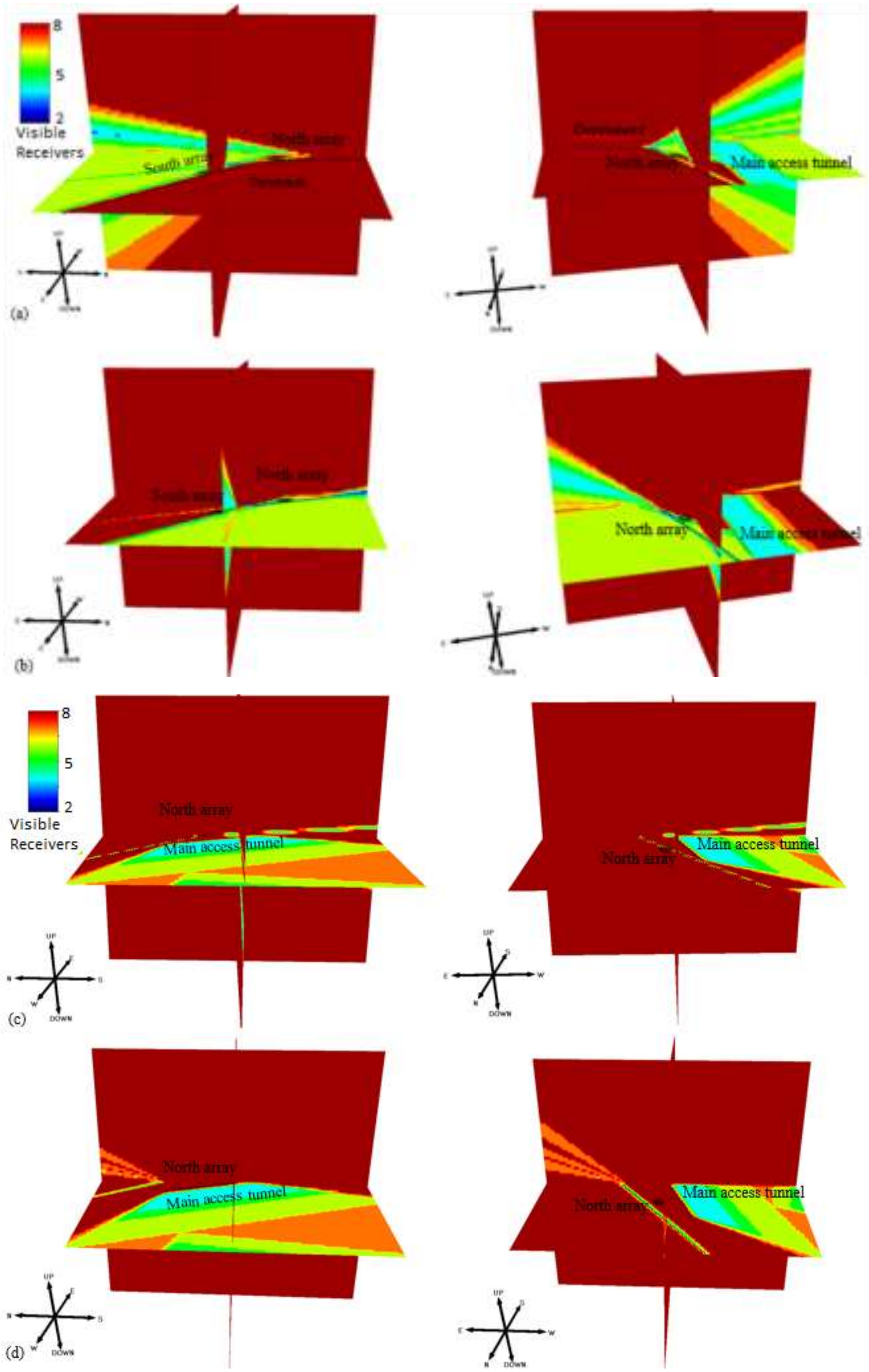


Figure 26: Comparison of shadows for 4 different cases of network geometry. For this amalysis sensors were assumed to be placed on the east and west side of the Gerstenegg tunnel (a and b respectively) and on the east and west side of the main access tunnel (c and d respectively). The color scale demonstrates how many receivers are affected by the shadow caused by Gerstenegg and main access tunnels.

Further array analysis, using the InSite software, took place to investigate the magnitudes threshold that the chosen network configuration would be able to capture. InSite's analysis tool, that determines magnitude sensitivity, calculates the minimum theoretical moment magnitude for an event to be successfully detected at each point in a user-defined monitoring space (cube of 4km edge). The analysis uses seismological theory, generalised scaling relations and waveform processing principles (e.g. Gibowicz and Kijko, 1994). It is assumed that at least four arrival time picks are required to calculate a location and that all sensors have an equal noise threshold. The input parameters are the attenuation factor Q = 650, the average crustal density Ro = 2800kg/m^3, v_P = 5.4km/s, and v_s=2.9km/s. The default InSite value of unity was used for the SNR - lower SNRs would yield lower magnitude sensitivity values.

The magnitude threshold distribution over the monitoring space is presented in Figure 27 as projections of the calculated iso-surfaces, while the inset shows the isosurface of M_L-4 calculated by InSite; all points on the iso-surface are of the same magnitude. The model exhibits that the magnitude threshold of up till -3M_w is 0.4km (green shaded), i.e. such small events can be detected only in the close proximity of each array. As per previous shadow modelling those are the same areas that are affected by the tunnels shadows rendering the detection of such magnitudes even more difficult. Events up till -2.30M_w can be detected to a distance of up till 0.8km around each array and after 1km detectability falls to magnitudes of -1.58 and higher. Overall, this magnitude coverage satisfies the needs of my PhD.

Another topic for discussion is the general deployment area: the microseismicity because of the lake heights variations and associated stresses was assumed to occur beneath the reservoir and in its exact vicinity. By deploying the sensors on its west side within the tunnels, I, by default, undermined the accuracy of locations occurring in the east side of the reservoir. A more precise location of events would be feasible provided the deployment of sensors was surrounding the lake and the GTS tunnels; i.e. a denser seismic network that included more than 8 stations in more than two locations and in various depths/elevations. However, the fact that the site is in the Swiss Alps and neither myself nor Nagra's staff were in position to guard or maintain sensors situated outside of Nagra's jurisdiction (GTS or Gerstenegg tunnels) on a permanent basis for such a long time and with less than optimal weather conditions. In addition to the above deploying sensors in a location that far away from power and internet availability rendered such a scenario non feasible.

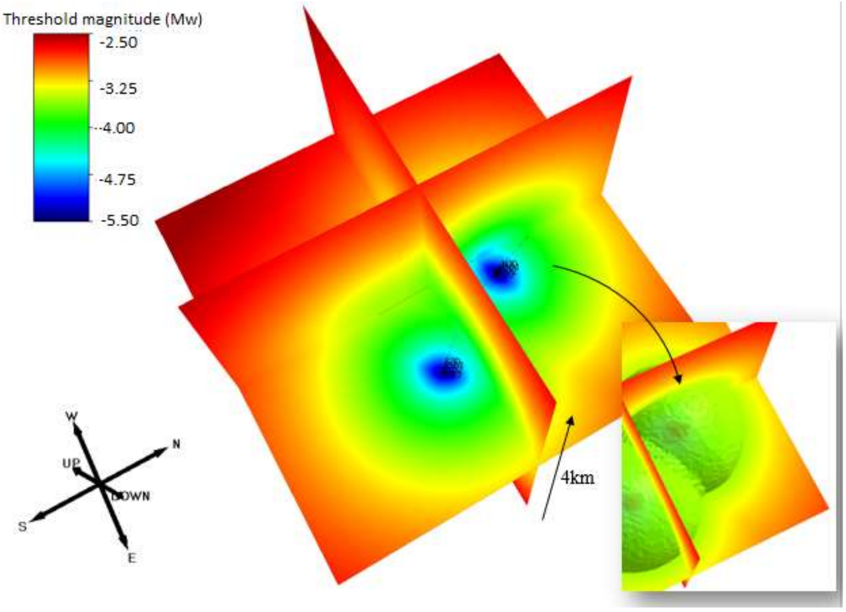


Figure 27: Model of the performance of the seismic network configuration in regards to its magnitude threshold (color scale) as a function of distance. Inset shows the magnitude isosurface of M_L-4.

3.8 Summary

In this Chapter I have detailed the works undertaken with reference to the seismic network deployment, maintenance and operation, pre-processing of acquired data, as well as some preliminary investigation of seismic noise which served as a basis of understanding some of the problems I had to tackle within my PhD. Although sensor deployment design and sensor data cleaning with noise characterisation for micro-seismicity monitoring are not unmapped areas, in this PhD they are explored in detail and for an extended period of several years and are two of the novelty contributions of this research.

In particular, it was discussed that the seismic network design depends on the scope of the application but, since it influences greatly the localisation of events, the researcher should invest time in choosing a seismic configuration that suits their project's needs. There are always going to be limitations in the seismic network design, for instance, in this particular application, the available instrumentation consisted of two surface arrays (of three 1-D sensors and one 3-D sensor each) and a borehole 3-D sensor, while the availability of locations for sensors deployment was limited by external constraints such as accessibility and jurisdiction. Because of the borehole sensor's malfunction the final number of sensors that were used in this project were eight while the total number of channels were twelve.

The pre-processing of the acquired data was also presented, i.e. the transformation of raw data to a format that could be read and processed using MATLAB, the conversion of the data from counts in velocities, the formation of data series, the corrections of time lags in the start times of the hourly records, the conversion of the timeseries from cell to structure arrays.

A preliminary visual observation of the acquired data was undertaken to allow for an understanding of the noise sources that influenced the recordings and check the effectiveness of the microseismic monitoring network. Various reoccurring noise signals of different frequencies and amplitudes were evident in the recordings. The sources of noise that were frequent in the area were related to hydroelectric activity, to construction activities, hydrofracking experiments, large seismic events and landslides. The identification of small tectonic earthquakes, which were also found in the Swiss Seismic Network database, confirmed that the sensors were able to successfully record small seismic events even at long distances.

The network design analysis demonstrated that the Gerstenegg tunnel is shadowing the area on its west and between the two arrays, while the main access tunnel on its west (GTS area). Both tunnels act as a barrier, rendering all events occurring in the areas to the west of and between the deployed arrays more difficult to be recorded. No shadowing was observed on the east side of Gerstenegg tunnel. The tunnels shadowing is expected to affect the recording and localisation of events. Further analysis indicated that sensors arrays, provided more were available, could be deployed in the main access tunnel to allow for better recordability in the area between Gerstenegg and the main access tunnel. The magnitude threshold analysis demonstrated that magnitudes of up till -3M_w can be detected only in the close proximity of each array (0.4km) which coincides with the area that suffers from the Gerstenegg tunnel's shadowing. At longer distances the detectability falls to magnitudes of -2.30M_w (0.8km) and -1.58 (after 1km).

In the following chapter, I am discussing about the research undertaken and the methodology followed to detect all recorded events and to determine the onset time of the signals' phases.

Chapter 4 Analysis of recorded microseismic data: Event detection

4.1 Introduction

After the collection of a vast dataset (each hour of data comprises of 10.8*10^6 data points for all the channels of both surface arrays) in varying background noise conditions, the need for automated algorithms that would successfully process such a dataset became evident.

As per Kinali et al. (2018), the high sensitivity of a microseismic monitoring system is also its main caveat. Seismometers record every vibration of the ground that is caused by any type of sources, at distances that can extend to tens of kilometres depending on the site conditions and the energy emitted by the seismic source. In addition, instrumental self-noise is present at all times. As a result, it can be extremely difficult to distinguish between the microseismicity that is of interest to a project and everything else. Such circumstances may be less problematic for projects such as hydro-fracturing, where the likely location and time of occurrence of microseismicity is known a priori. But for the vast majority of applications, this is not the case and peaks in ambient noise can be mistakenly regarded as microseismic events. A false increase in the recorded frequency of microseismic events will bias project results. Furthermore, manual verification of each event will result in significant data processing time, yet neglecting verification can lead to other adverse economic impacts; for example, unnecessary road closures due to the false triggering of an early warning system for landslides. By contrast, relaxing event detection criteria to avoid false alarms can result in excess risk, with microseismic events remaining undetected. Monitoring for longer than a couple of days and with a sampling rate between 200 – 250 Hz (a range adequate for the needs of most projects requiring microseismic monitoring) leads to vast datasets that are not cost effective for visual inspection and require a computational detection approach.

The following sections of the present chapter are structured as follows: Section 4.2 is a brief literature review of available detection algorithms while Section 4.3 further describes the existing onset time picking algorithms (i.e. detection algorithms that can be used for accurate locations). The methodology used for detection and onset time determination of potential seismic events of interest in the present study are then presented in Sections 4.5 and 4.6. Sections 4.2 and 4.4 include text published as part of the Kinali et al. (2018) paper. The paper presents the NpD algorithm using 3 hours of data and not the full data set of passive seismic data acquired between 1st November 2014 and 12th August 2017 which is presented in subsection 4.4.5.

4.2 Event detection literature review

A number of automatic detection approaches have been developed that work in the time or frequency domain or both e.g., Freiberger, 1963; Gibbons and Ringdal, 2006; Goforth and Herrin, 1981; Joswig, 1990; Küperkoch et al., 2010; Vaezi and Van Der Baan, 2014, and are further discussed as follows.

4.2.1 Automatic detection in the time domain

The most widely used event detection algorithm at present is the STA/LTA (Bormann, 2012) which operates in the time-domain. The ratio of two moving averages STA/LTA is computed continuously at each time t for recorded data x_t:

$$STA_t = \frac{1}{N_S}\sum_{n=t}^{t+N_S} y_n \qquad \textit{(eq. 1)}$$

$$LTA_t = \frac{1}{N_L}\sum_{n=t}^{t+N_L} y_n \qquad \textit{(eq. 2)}$$

where STA is the N_S-point Short-Term Average, LTA is the N_L-point Long-Term Average and the parameter y_t denotes a characteristic function (CF) $y_t = g(x_t)$. The characteristic function

CF is chosen so that it enhances any signal changes in the time-series; common CF choices include energy (x_t^2) (McEvilly and Majer, 1982), absolute value ($|x_t|$) (Swindell and Snell, 1977) and the envelope function ($\sqrt{x_t^2 + \bar{x}_t^2}$, where $\bar{x}$ is the Hilbert transform) (Earle and Shearer, 1994), or even higher-order statistics where skewness and kurtosis are calculated in the sliding windows (Saragiotis et al., 2002; Küperkoch et al., 2010). The raw data are demeaned and then the ratio STA/LTA is compared to a user-selected threshold: when the ratio exceeds the user-selected threshold, an event is detected. The end time of the event is defined by the time when the ratio falls below a detrigger threshold (also chosen by the user). N_s should be chosen approximately equal to the dominant period of the events the algorithm aims to trigger. LTA is a measure of background noise variations, so N_L should be set to some value longer than the period of the lowest frequency seismic signal of interest. The STA, LTA windows are usually chosen as non-overlapping (Trnkoczy, 2002).

A different approach is suggested by Stewart (1977). This method uses a high-pass non-linear filtering process, to determine whether a seismometer is operating within acceptable limits of noise before its data are accepted to be used. If accepted, the algorithm sets some requirements for detection and tentative confirmation in the time domain, i.e. setting different lower bounds for the triggering threshold, the SNR; the number of times the waveform exceeds the triggering threshold; the consecutive time the waveform stays within the threshold; and the maximum amplitude of the waveform once the signal is detected.

4.2.2 Automatic detection in the frequency domain

Most algorithms in the frequency domain use Fourier transforms. One of the first mathematically based signal detectors is the one proposed by Freiberger (1963) who developed the theory of maximum likelihood by applying an approximate comparison of spectral densities, based on the Toeplitz approximation forms, for the detection of Gaussian signals in Gaussian noise.

Goforth and Herrin (1981), in order to overcome the challenge of a varying non-normal background noise, developed an automatic seismic signal detector based on the Walsh transform, which is a series of rectangular waveforms with amplitudes of +1 or -1, instead of the sines and cosines of Fourier. Once the data are filtered in the time domain, segmented in overlapping windows and transformed, the Walsh coefficients are assigned a weight such that the noise spectrum is whitened and the expected signal is isolated. The values of the weights need to be chosen by the analyst, after manual inspection of the appropriate noise segments. At each time window, the current sum of the absolute values of the weighted Walsh coefficients is compared to a threshold,

$$Threshold = V_{50} + K(V_{75} - V_{50}), \qquad \textit{(eq. 3)}$$

where V_{50} is the median of the distribution of previous 512 values, V_{75} is the 75th percentile of the distribution of previous values, and K is the arbitrary constant set by operator. If the current value exceeds the threshold, it results in a signal detection; if not, the current sum is ranked among the previous number of predefined values and the oldest sum is discarded.

Michael et al. (1982) modified the Goforth and Herrin approach to develop a real-time event detection and recording system for the MIT Seismic Network. Their algorithm uses the power spectrum to remove the effects of phase shifts and instead of the Walsh coefficients (energy spectrum) they use power Walsh coefficients (i.e. the Walsh coefficients are squared and each pair is summed). They also add a minimum duration that the coefficients need to be above threshold; an event termination criterion; and accept events only if they are correlated by at least three stations.

Vaezi and Van de Baan (2014, 2015) developed an algorithm for the detection of induced microseismicity during hydrofracturing. They compared the moving average Power Spectral Densities (PSDs) of small segments of their data record to the averaged background noise PSD of quiet segments of their data record, resulting in the picking of all signals that stand out in a statistical sense from background noise. The PSD of a signal refers to the spectral energy distribution per unit time and is simply the representation of the signal in the frequency domain (Press et al., 2007), measured in squared magnitude units of the time series data per unit frequency. The outcome of this comparison, i.e. the normalized misfit $u_t(f)$, is calculated by the following equation (eq.4) and for a clearer depiction of the events, only the positive values are kept:

$$u_t(f) = \begin{cases} \frac{PSD_n^t(f)-\overline{PSD}(f)}{std(f)}, if\ u_t(f) > 1 \\ 0 \qquad\quad , otherwise \end{cases} \qquad \textit{(eq. 4)}$$

where $std(f)$ is the standard deviation at frequency f computed from the PSDs of the noise segment $PSD_m'(f)$, $PSD_n^t(f)$are the PSDs of small segments of the original data x(t) estimated (eq.5), using rolling (overlapping) windows of predetermined length L, and $\overline{PSD}(f)$ is the total average PSD of the quiet sections of the data x'(t) (eq.6). To isolate only the quiet sections they discarded all the absolute amplitudes greater than a multiple of the original record's root-mean-square (RMS) amplitude.

The individual moving average PSDs are estimated using Welch's modified periodogram method as follows:

$$PSD_n^t(f) = \begin{cases} \frac{a\left|\sum_{l=1}^{L} x_n(t_l)\omega(t_l)e^{-j2\pi fl}\right|^2}{f_s LU} & if\ f = 0, f_{Nyq} \\ \frac{2a\left|\sum_{l=1}^{L} x_n(t_l)\omega(t_l)e^{-j2\pi fl}\right|^2}{f_s LU} & if\ 0 < f < f_{Nyq} \end{cases}, n = 1, 2, \dots, N \qquad \textit{(eq. 5)}$$

where a is a scale factor that accounts for variance reduction which depends on the type of the taper w, f_{Nyq} is the Nyquist frequency in Hz, fs is the sampling frequency in Hz, j =$\sqrt{-1}$ and U is the window normalization constant that ensures the modified periodograms are asymptotically unbiased and is given by: $U = \frac{1}{L}\sum_{i=1}^{L} \omega(t_i)^2$.

The average PSD estimate is calculated by averaging the PSD estimates of the quiet data record:

$$\overline{\mathrm{PSD}}(\mathrm{f}) = \frac{1}{\mathrm{M}}\sum_{\mathrm{i}=1}^{\mathrm{M}} \mathrm{PSD'}_{\mathrm{m}}(\mathrm{f}) \qquad \textit{(eq. 6)}$$

where $PSD_m'(f)$ denotes the PSD estimate of the mth noise segment as a function of frequency f and is given by eq.5 where instead of the original data x(t) I are now using the quiet data record x'(t).

The triggering criterion can either be the summation of the positive misfits $(u_t(f))$ over the total number of frequencies and normalized by division with the standard deviation, or the summation of the squared positive misfits over the total number of frequencies normalized by division with the standard deviation. When the triggering criterion exceeds a user-selected threshold an event is declared.

Shensa (1977) had developed a methodology to adapt to a dynamic noise environment with a variety of (weak) signals with widely different spectra. He computed the PSDs of small segments of the data and depending on the relation between noise and signal he developed 3 algorithms: (a) the average power detector, for signals that exceed noise uniformly over a relatively broad frequency index range when both noise and signal are stable; (b) the maximum deflection detector, for signals that exceed noise over at least one narrow frequency band; and (c) the average deflection detector, for signals that exceed background noise uniformly over a relatively wide frequency index range when both signal and noise are unstable. The relevant detectors are formed accordingly:

$$Det_a = \frac{\frac{1}{N}\sum_{k=n_1}^{n_2} PSD_i(k)-\mu}{\sigma}, N = n_2 - n_1, \qquad \textit{(eq. 7)}$$

$$Det_b = \max\left[\frac{P_i(k)-\mu(k)}{\sigma(k)}\ (k=0), \frac{P_i(k)-\mu(k)}{\sigma(k)}(k=1), \ldots, \frac{P_i(k)-\mu(k)}{\sigma(k)}\ (k=N)\right] \qquad \textit{(eq. 8)}$$

$$Det_c = \frac{1}{N}\sum_{k=n_1}^{n_2} \frac{P_i(k)-\mu(k)}{\sigma(k)}, N = n_2 - n_1, \qquad \textit{(eq. 9)}$$

where index range $n_1 \leq k \leq n_2$, μ and σ the mean and standard deviation, respectively. The parameters μ and σ must be estimated from noise-only data sections (i.e. no signal present).

4.2.3 Automatic detection in the time-frequency domain

Algorithms that work in the time-frequency domain are also common. Joswig (1990) proposed a pattern recognition technique using characteristic event features in spectrograms. His algorithm defines a knowledge base of images of the typical earthquakes and noise bursts in the time-frequency domain, using Fourier transforms, each of which is defined by a matrix and a scaling factor (to account for magnitude differences). The sonogram-detector matches patterns for the events that are above a user-defined set of thresholds and provides one message per detected event stating the detection time, the maximum pattern fit and maximum amplitude of the detected event.

Another pattern recognition technique was proposed by Bodenstain and Praetorius (1977) aimed at the automatic detection of electroencephalogram signals (0.5 – 30Hz signals). According to their research, the data record can be segmented into elementary patterns (e.g. seismic signals and transients) using linear predictive filtering, leading to the extraction of features (power spectra and the signal's time structure) which in turn can be combined (clustering procedures, classification) so that they represent the seismic signal as a whole.

During the last years, wavelet transforms have increasingly been preferred over Fourier transforms. The main reason being the simultaneous time- and frequency-domain localization of the wavelets, in contrast to the only frequency-domain localization of the standard Fourier transform, or the frequency-time resolution trade-off of the Short-time Fourier transform which depends on the width of the window function used (Ching et al., 2004; Sifuzzaman et al., 2009). Anant and Dowla (1997) use polarization and amplitude information contained in the wavelet transform coefficients of the signals to construct "locator" functions that identify the P and S arrivals. High-pass and low-pass filters are used (wavelet and scaling filters respectively) which must belong to a perfect reconstruction filter bank.

4.2.4 Discussion on detection algorithms

All detection algorithms have advantages and shortcomings with no algorithm being clearly optimal under all source, receiver, path and noise conditions (Withers et al., 1998). The most widely used event detection algorithm at present is the STA/LTA (Bormann, 2012) which operates in the time-domain. STA/LTA is an excellent onset time detector for adequately high SNR events; a condition that may not be true in the case of weak microseismic events. Also, the method can lead to false triggers unless the data used are optimally filtered to minimize the effect of noise; this is difficult to achieve in a varying noise background. In fact, in all algorithms where bandpass filtering is part of the detection process (STA/LTA or Goforth's and Herrin's algorithm), some kind of a priori knowledge on the expected signals is assumed. The choice of the filter to be used is important, as inappropriate filtering can result in the removal of useful information from the data.

The method by Vaezi and Van de Baan (2014) was found to outperform the STA/LTA technique by detecting a higher number of weak events while keeping the number of false alarms at a reasonable level (Vaezi and Van Der Baan, 2015). It requires, however, some pre-processing where all noise bursts or transients that may exist in the data are removed. It also assumes stationary noise that follows a normal distribution and, therefore, employs the mean and standard deviation as statistical tools. Although this might be a good approximation for recordings with high SNR, it is not the case for seismic data with low SNR. In such cases, the average PSD is not representative of the central tendency of noise and as such any detection criteria based on deviation from the mean could lead to a large number of 'false' detections. This is particularly important where long, continuous recordings are available as it can significantly increase the processing time and bias the results.

4.3 Onset time picking literature review

There is a number of onset time picking methods in the literature and they can be classified in categories as single-level methods (e.g., window-based: energy-ratio methods and non-window based: Akaike information criterion), multilevel or array based methods (e.g., crosscorrelation approaches), and hybrid methods that combine a number of single-level methods (e.g., Akazawa's method). Some detection methods are also used for onset time picking. Some of these onset time picking methods are presented as follows.

Representative single – level window based methods are STA/LTA (discussed in the previous subsection) and modified energy ratio (MER) method introduced by Han et al. (2009). MER method is defined as the ratio of energies of the window that follows a test point over that of the preceding window. Both windows are of equal length. Both methods rely on ratio calculations, differing only on the positions and lengths of windows and ratios used for the first arrival determination. For the STA/LTA method, the first arrival times are picked at the rising slope maximum of the STA/LTA ratio while for the MER method, at the MER peak.

Modified Coppens's method (MCM), as introduced by Sabbione and Velis (2010), calculates the energy ratio of the seismic trace within two nested windows: a fixed, leading and selected a priori window and an increasing with time window. MCM assumes the window preceding the event includes only noise while the one that follows includes noise and signal. Sabbione and Velis use edge-preserving smoothing (EPS) as part of their methodology and determine the first-break onset of the signal as the maximum derivative of the filtered energy ratio.

A widely used non-window based methodology for onset time determination is presented by Maeda (1985) who uses the Akaike Information Criterion (AIC; Akaike (1974). AIC is a technique that estimates the likelihood of a model to predict/estimate the future values. A good model is the one that has minimum AIC among all the other models. Maeda computes AIC directly from the time series: $AIC\ (k) = k \log var\{x(1,k)\} + (N-k-1) \log var\{x(k+1,N)\}$, where k ranges through all samples of the input microseismic waveform and $var\{x\}$is the variance function and considers as onset time arrival the position that AIC is minimised.

Higher order statistics (kurtosis) can be another effective tool in the signal identification process with the assumption that the noise in the input data is close to a Gaussian distribution and the signal is non-Gaussian. Saragiotis et al. (PAI-K, 2002, 2004) propose a single level non-window based algorithm where the characteristic curve is formed from the kurtosis values on a sliding window for the entire input waveform length and the arrival time is picked on the maximum slope of the characteristic kurtosis curve. The arrival times are picked on the maximum slope of the local maxima in the corresponding P- and S-wave intervals.

Array-based methods such as the crosscorrelation method presented by Irving et al. (2007) require a known, manually picked arrival time at a reference (or pilot) waveform. All receiver levels in the array are then crosscorrelated with the pilot waveform and aligned using the lag value. The stacked aligned waveforms form a new reference waveform which is then again crosscorrelated with all waveforms and the process is repeated until the alignment matches user's specifications.

Another crosscorrelation workflow is presented by De Meersman et al. (2009). They use a location software to coarsely identify the P- and S- phases of microseismic events, group them in multiplet groups, and then apply cross-correlation in each multiplet to refine the picking of the phases. Multiplets are signals originating from the same fault and by the same mechanism (Geller and Mueller, 1980). Therefore, they have similar waveforms and, once identified, can be used to more accurately obtain the location. Within each multiplet, waveforms are aligned, using the known picks and are rescaled to equalize to the pre-event noise level. A really small pilot waveform window is then used, centred on the picked arrival and correlated with all waveforms to update the time-shift. The maxima of the crosscorrelation functions indicate the corrections needed to optimally realign the waveforms in relation to the pilot waveform. This process is repeated until the time lag converges to a value that is less than a user-defined threshold value.

Numerous hybrid approaches exist that merge single-level algorithms in different combinations. Anant and Dowla (1997) combine a wavelet transform method with polarization attributes to pick onset time arrivals (AD's method). In particular they compute the wavelet decomposition of single station, three component seismograms and, for each component and scale, the eigenvalues from the covariance matrix generated on a sliding window. Using the largest and intermediate eigenvalues Anant and Dowla estimate a rectilinearity function (a metric that measures the degree of linear polarization of a wave) for each scale. Then, a composite rectilinearity function is developed, which is an interscale comparison of rectilinearity, and the position that it reaches its maximum is denoted as the P arrival time.

A hybrid approach that combines the Akaike Information Criterion with autoregressive series is the method presented by Leonard (2000). His AIC algorithm is based on the concept that microseismic signals are nonstationary and can be approximated by dividing the waveform into locally stationary segments, where each segment is treated as an autoregressive process. The order of the AR process required to resolve the detail of the spectra is higher for a complex signal than for the preceding noise (Leonard and Kennett, 1999). He initially bandpass filters the data and then calculates a foreword AR process before the beginning of the signal (just noise) and applies an error prediction filter to the whole data window which removes this noise. The inverse process then follows: an AR model of the signal plus noise, based on a data section starting after the initial pick estimate, is calculated and an error prediction filter is applied to the data in the opposite direction. This removes the signal from the data. The joint AIC is calculated from the two AR-filtered time series and its minima is considered the onset time. The AIC function provides a measure of the fit of the autoregressive model. The optimal separation of the two stationary time series (noise and signal plus noise) is indicated by the time index associated with the minimum value of AIC (Ahmed and Sharma, 2007; Tronicke, 2007).

Zhang et al. (2003) (ZTR's method) aims to detect singularity with multiscale wavelet analysis and application of AIC. The methodology is based on the idea that while significant features in the signal are retained over several scales (of the wavelet decomposition), random noise or other incoherent features disappear. Zhang's method applies a soft thresholding scheme, instead of filtering, to reduce noise in the seismogram, and then decomposes the wavelet and computes AIC on each scale. A preliminary arrival time is declared on scale two of the decomposition only if the AIC values for all scales are close to each other and inside a user-specified interval. AIC is then reapplied on the data components in a window surrounding the preliminary pick and the arrival time is denoted as the time AIC reaches its minimum.

Another window-based hybrid method is presented by Akram et al. (2013) and combines two energy ratios to enhance signal coherency and improve confidence in arrival time picking of low S/N microseismic data (Joint Energy Ratio (JER) method). At the i-th time sample the energy ratio is defined as $ER_{\{i\}} = \frac{\sum_{j=i}^{i+w} x_j^2}{\sum_{j=i-w}^{i} x_j^2}$ and the ratio of the peak eigenvalues (PER) as $PER_{\{i\}} = \frac{\lambda_{1A}}{\lambda_{1B}}$, where λ_{1A} and λ_{1B} are the peak eigenvalues in windows after and before the i-th sample, respectively. This is followed by the computation of STA/LTA and JER at the ith time sample: $JER_{\{i\}} = PER_{norm\{i\}} SL_{norm\{i\}}$, where $PER_{norm}$and SL_{norm} are PER and STA/LTA normalized by their respective maximum value. The onset time arrival corresponds to the JER curve local maxima. The JER algorithm combined with AIC, for the latter to work more efficiently, is presented as a methodology by Akram (2014; JER-AIC).

Tan et al. (2014) present a joint STA/LTA-polarization-AIC method (TYFH's method). They compute the STA/LTA ratio for the input data and the degree of polarization in a sliding time window. Another function is computed from the degree of polarisation according to Moriya (2008; polarisation function), whose local maxima indicates the arrival time. Both STA/LTA and polarisation are normalised and their product is denoted as a new picking function. Likelihood functions of the autoregressive series of the waveform are computed and normalised and finally multiplied with the picking function. The maximum of the resulting response curve is considered the signal's arrival time.

Akazawa's method (2004) combines STA/LTA and AIC to pick the onset time arrival. He computes a cumulative envelope function of the input waveform and then STA/LTA on the function. Two rounds of AIC application then follow: (i) firstly, AIC is applied on the envelope function in the interval between the first index and the one associated with the maximum envelope value (*Emax*) and then (ii) AIC is repeated on the shorter interval between the index associated with the minimum value of the previous AIC application and the index Emax. The time index associated with the minimum value of the 2nd application of AIC is considered as the arrival time.

Hybrid methods combining high order statistics and other single level methods are documented by many researchers. For instance, Galiana-Merino et al. (2008) pick P-wave arrivals using a kurtosis-based criterion in the stationary wavelet domain. Li et al. (2014) propose a short-term kurtosis and long-term kurtosis ratio (S/L-Kurt) based approach. The arrival times are picked on the maximum slope of the local maxima in the corresponding P- and S-wave intervals.

More sophisticated methods of onset time picking are the neural network approaches. Dai and MacBeth (1995) propose an artificial neural network (ANN) according to which they use the vector modulus of three-component seismic recordings as the network input and employ a discriminant function (the trained ANN output) to define the onset time arrival. They use nine P waves and noise segments to train the ANN. The ANN can be improved by adding or adjusting the training data. Wang and Teng (1995) propose two variations of ANN for real-time earthquake detection: type A (AND-A) which uses the recursive STA/LTA time series as input data, and type B (AND-B) which uses moving window spectrograms. Zhao and Takano (1999) present an ANN approach for broadband seismic detection which combines the results from three backpropagation neural networks: a long, a mid- and a short term.

4.3.1 Discussion on onset time picking algorithms

There is a plethora of studies that have attempted to identify the best onset time picking algorithm. Mitchell et al. (1998) compare selected trigger algorithms for automated global seismic phase and event detection and conclude that no algorithm is clearly optimal under all source, receiver, path and noise conditions. Sharma et al. (2010) also reach a similar conclusion, i.e. out of the presently known and used real time trigger algorithms, from the very simple amplitude threshold type to the sophisticated pattern recognition, adaptive methods and neural network based approaches, none of these perform optimally on all data sets, but that there are some, which provide more accurate and precise picks.

The reason why none of the onset time picking algorithms seem to perform optimally in all cases lies on two facts: that (i) all algorithms present dependencies on specific parameters depending on the general category they belong to and (ii) the signals that need to be picked may vary greatly in waveform shape, amplitudes, coherency due to the existence of noise and duration.

In respect to the individual algorithms parameters, the accuracy of the window based methods for instance, i.e. STA/LTA, depends on the length of the time window used and the predefined user-set threshold (Trnkoczy, 1999). With MER and MCM, which are again window-based, energy ratio methods, the results are still affected by the length of the time window but also by the shape of the waveform. In all three algorithms the window size should be greater than a few periods of the microseismic signal to avoid false picks that arise from noise fluctuations and to pick the signal changes properly. A window size of approximately (2-3)τ_{dom}, where τ_{dom} is the apparent dominant period of the signal, is recommended for STA/LTA and MER while for MCM (1-3)τ_{dom} (Akram and Eaton, 2016). LTA window should be between 5 and 10 times the size of the STA window for optimal performance. MCM requires additionally an EPS filter and the window that is used is not a fixed length one as in the previous cases. JER algorithm depends on window size parameters (since part of its methodology is based on STA/LTA) and on the window size for peak eigenvalue ratio (which is recommended to be approximately (2-3)τ_{dom}. As expected, single level non-window based algorithms such as the kurtosis-based algorithms (PAI-K and S/L Kurt) have improved accuracy as compared to STA/LTA.

Crosscorrelation methods, such as IKK's method, are semi-automated because they require the user's selection of the pilot traces and of the window size and can be cumbersome for large data sets. The choice of the optimal pilot waveform is critical for the success of a crosscorrelation method (Bagaini, 2005) and it can be (i) the waveform with a high SNR, (ii) a synthetic waveform produced from the stacking of all receivers' waveforms, or (iii) it can iteratively be updated each time and substituted by the waveform with the highest SNR. The correlation window size in cases of large deviations in the estimations of the initial picks should be $\leq 10\tau_{dom}$ while in cases of small deviations it should be chosen as$\leq 3\tau_{dom}$.

AIC along with an autoregressive model requires the choosing of the model order, which is estimated by trial and error. The algorithm also requires picking on rotated data or initial estimates of the arrival windows to perform optimally (Zhang et al., 2003). Because the AIC picks are based on the global minimum of the checked data window, this should be a small time interval containing the arrival of one signal only; in cases of multiple arrivals the algorithm underperforms (Zhang et al, 2003).

Algorithms that are based on wavelets require selection of a single wavelet function and a number of decomposition levels and are computationally expensive because of the latter. ZTR is based on the interactive selection of the wavelets function and decomposition levels needed and the window size (Zhang et al, 2003).

Real time microseismic analysis requires algorithms that combine simplicity, computational speed and enough precision to reach meaningful locations of events. That is the reason why the most popular algorithms are STA/LTA and MER. In post processing scenarios and in the times we live in that processors have made data processing a much easier task, precision and a small number of parameters are a priority. In particular in the present study that aims to analyse real field microseismic data I chose to rely on the findings from the comparative study by Akram and Eaton (2016). According to their research, statistical, hybrid, and multilevel crosscorrelation methods are more efficient in terms of accuracy and precision, however, the performance of the algorithms is not found to be consistent between the field data and the synthetic data records comparison. For instance, S/L-Kurt outperforms the PAI-K algorithm on field data, but not on pseudo-synthetic data; this is attributed by the authors to the weak P-waves and complex waveforms of the field data.

In Figure 28 the percentages of pick errors within a [-2, 2] ms error range are illustrated for 14 different algorithms and for field microseismic three component data. The pick errors are obtained by subtracting the manual picks from the automatic picks. According to the pick error pie charts the best performing algorithms are Irving's method, Akazawa's method, JER-AIC, ZTR's method and S/L Kurt. Lesson learnt from this research is to enhance the S/N of the data using noise reduction techniques, such as wavelet-based denoising approaches. Akram and Eaton also recommend that an interactive quality-control process should follow any automatic picking workflow to ensure the quality of arrival-time picks.

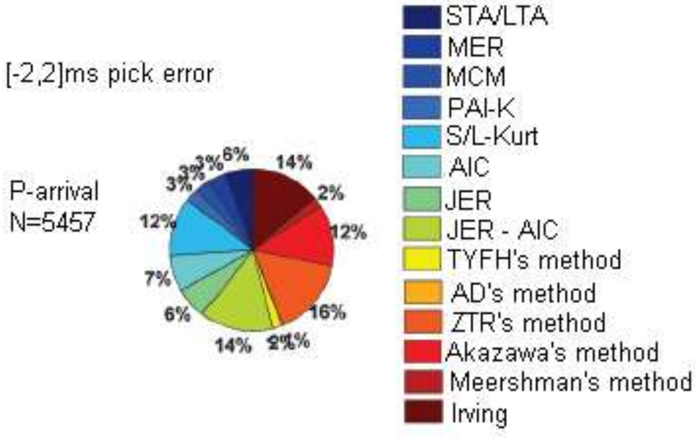


Figure 28: Figure adapted from Akram and Eaton (2016) paper, demonstrating the performance of various algorithms in picking P-arrivals on field data. According to the pick error pie charts in the [−2 ms, 2 ms] error interval Irving's method (14%), Akazawa's method (12%), JER-AIC (14%), ZTR's method (16%), and S/L-Kurt (12%) perform better than the other methods that were tested.

4.4 Event detection in the present study

For the event detection in the present study a new algorithm is proposed. This is partly based on the observation that microseismic events have been found to represent stronger spectral content over a frequency band that depends on the nature of the event, than that of background noise (Vaezi and Van Der Baan, 2014). According to this, a microseismic event can be regarded as an outlier, i.e. a data value or values that are outwith an expected range which represents the noise. The challenge is to define the upper bound of this range when no a priori knowledge of the expected signal (in terms of amplitude and frequency content) is available.

In statistical analyses, for populations that are normally distributed, the detection of outliers is usually based on the 3σ criterion, where σ is the standard deviation of the data (Barnett and Lewis, 1994). Any values that are outwith the $\pm 3\sigma$ range are considered outliers. This range includes 99.7% of the data. For populations that are not normally distributed though, this criterion could lead to erroneous results as the mean is not necessarily the best quantity to describe the central tendency of the data. Even if the PSD values are indeed normally distributed for one hour of data, it does not guarantee that this will be the case for the full duration of the data set. A robust method for the characterization of the background noise and the determination of an upper bound for the noise PSD value is required.

4.4.1 Spectral characterisation of background noise

Background seismic noise can result from numerous sources: natural perturbations, e.g. tides, tectonics, seasonal changes, etc., and man-made perturbations. Perturbations can have a periodic or transient nature; their durations may differ from instantaneous bursts to elevated noise that lasts for hours, days or even months; in the case of seismic arrays, noise amplitudes may vary between seismometers at different locations. Investigations of the seismic noise in hydrofracking sites have shown that noise has nonstationary properties, correlated in both time and space (e.g. Chambers et al. (2010)). Despite this, most detection algorithms assume normality for the noise distribution (e.g., Vaezi and Van der Baan (2014) and (2015)).

The following methodology allows for the determination of a characteristic level of background noise in the frequency domain through examination of the statistical distribution of its PSD spectrum. First, I determine whether the background noise followed a normal distribution in order to choose appropriate statistics and then check if there are significant temporal or spatial variations in background noise. Knowing the distribution allows for the determination of the appropriate statistics, i.e. parametric or non-parametric, to be used in further analysis.

I therefore compute the individual PSDs for N_w non-overlapping (to ensure that the data between segments are statistically independent) segments of duration t_l for the frequency range 0 - Nyquist frequency, f_{Nyq} using the Welch's modified method (Welch (1967)). The PSD is calculated at discrete frequencies within this range. The total duration of the data set is then

$N_w \times t_l$. In general, the duration of an individual segment should include at least two full cycles of the expected signal. I suggest a duration of 0.5 to 2 seconds is adequate for microseismicity due to shear failure (Obert and Duvall, 1967). For research on other types of microseismic events, such as those induced during a landslide, segments with longer durations are recommended. Upon completion of the PSD calculations for each individual segment, there are N_w PSD values for each discrete frequency in the range 0 - f_{Nyq}. To determine normality in the PSD values for a specific frequency, graphical methods, i.e. histograms, probability plots and boxplots, can be used. An alternative to graphical methods are normality tests such as Shapiro-Wilk test S-W (Razali and Wah, 2011) and Kolmogorov-Smirnov K-S test (Massey, 1951).

If the normality check results in normally distributed PSD values for each frequency of the PSD spectrum, then a mean PSD value and a standard deviation (σ) for each specific frequency can be calculated. The *Noise PSD* (i.e. the characteristic upper bound) value for each individual frequency can then be specified by applying the $\pm 3\sigma$ criterion or any other suitable combination between the mean and the standard deviation as an upper threshold, e.g. mean $\pm \sigma$.

If the normality testing reveals a non-normal distribution, an upper bound for the background noise can be determined using non-parametric statistics, i.e. percentiles. I recommend that a high percentile, between 75 and 90, is chosen, based on the 3σ criterion. The *Noise PSD* is then defined by the chosen percentile PSD value at each discrete frequency f.

As discussed, to determine appropriate statistics for the analysis I first assess the assumption of normality for the distributions of all PSD values for the frequencies within the interval 0 – 125 Hz. I compute the PSDs, for all non-overlapping 2 second time windows within quiet hours at the frequency range (0-125 Hz). Hours outside of the GTS working hours, during which no tectonic events were reported by the Swiss Seismological Service catalogue (public.geo.admin.ch and swisstopo, n.d.), are randomly chosen to be used for this analysis. To determine if random samples of independent PSD observations are normally distributed, different graphical methods (histograms, probability plots and boxplots) and the Shapiro-Wilk test S-W (Razali (2011)) and Kolmogorov-Smirnov K-S test (Massey (1951)), were applied. Here I present indicatively, random hours within 04/11/2014 and 16/05/2015. Both S-W and

K-S tests reject the null hypothesis of normality in all cases checked ($p<0.05$). In fact, the noise PSD histograms are negatively skewed with positive kurtosis; examples for the frequencies of 30 and 85 Hz for the vertical component of the 3-component sensor in the North and of the South Array, located approximately 1km apart, are presented in Figure 29a and 2b respectively. The histograms are clearly not derived from normally distributed data, hence non-parametric statistics for noise characterization are appropriate. Also, the histograms for each sensor are different, hence background noise at each sensor is not the same. Figure 29a and 2b also show the two Noise PSDs derived for the same hour, using a characteristic upper bound of the 75th percentile. The value of the 75th percentile for each frequency and how this is related to the noise PSD is clearly annotated on the figure.

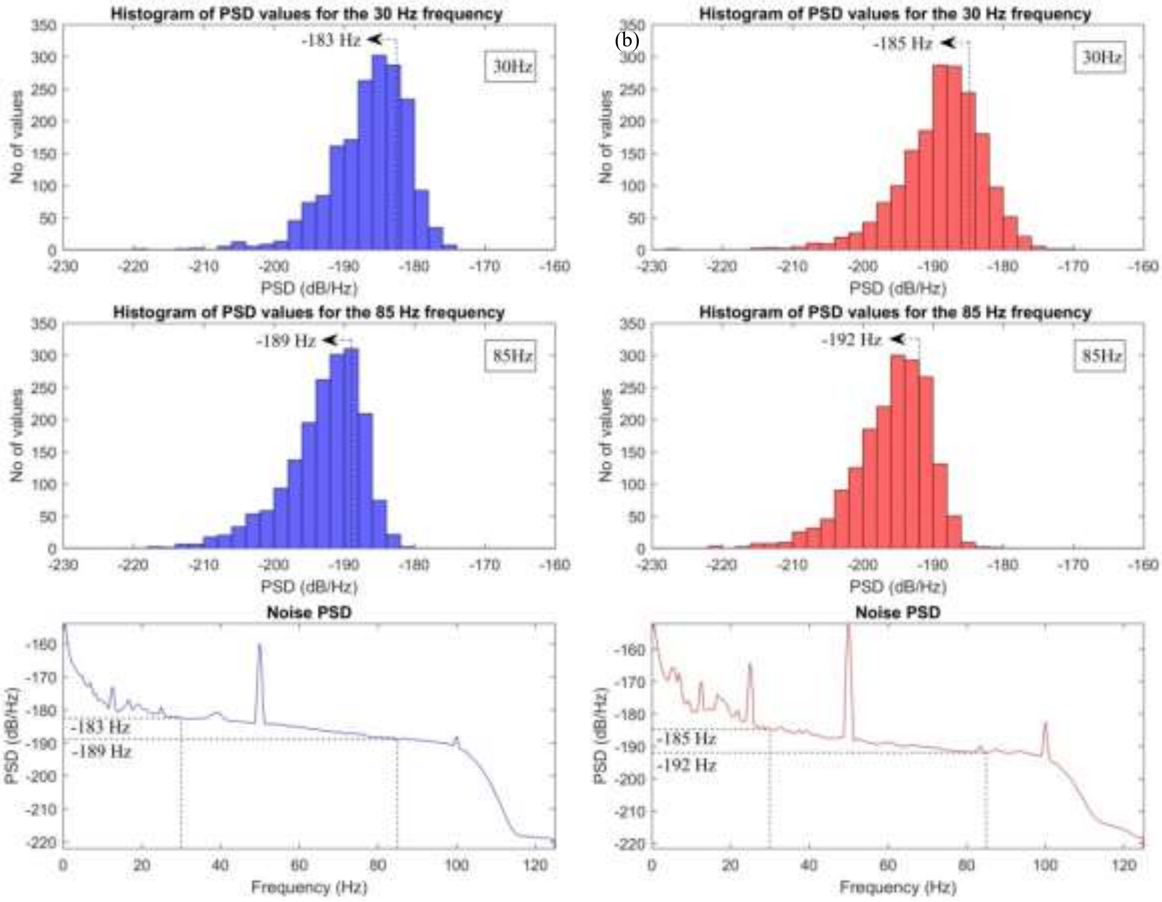


Figure 29: Calculation of the *Noise PSD* for one hour of data recorded by the vertical component of the 3-component seismometer of (a) the North and (b) the South array. The histograms of the PSD values at frequencies 30 Hz and 85 Hz and the value of a characteristic upper bound (here the 75th percentile) are shown as an example. These values are then used as the *Noise PSD* values at 30 Hz and 85 Hz frequencies, respectively. The values of the characteristic upper bound for all frequencies constitute the *Noise PSD* (bottom plots in (a) and (b)). All histograms are for data from the same day and hour.

In Figure 30 we observe examples of the PSDs plotted against the frequency range used for the temporal and spatial comparisons of the background noise. The temporal subpart is composed of upper Noise envelopes of different hours for one of the seismometers, while the spatial subpart comprises of upper Noise envelopes of different seismometers. Just by visual observation of the Figure 30a it is evident that the noise is different not only for different days but also for different hours within the same day. As it concerns the spatial variation, Figure 30b shows the PSD spectrum of one hour of data obtained from the seismometers of the North and South array. It can be seen that the spectra differ even for the seismometers of the same array (distances between adjacent sensors less than 50 m).

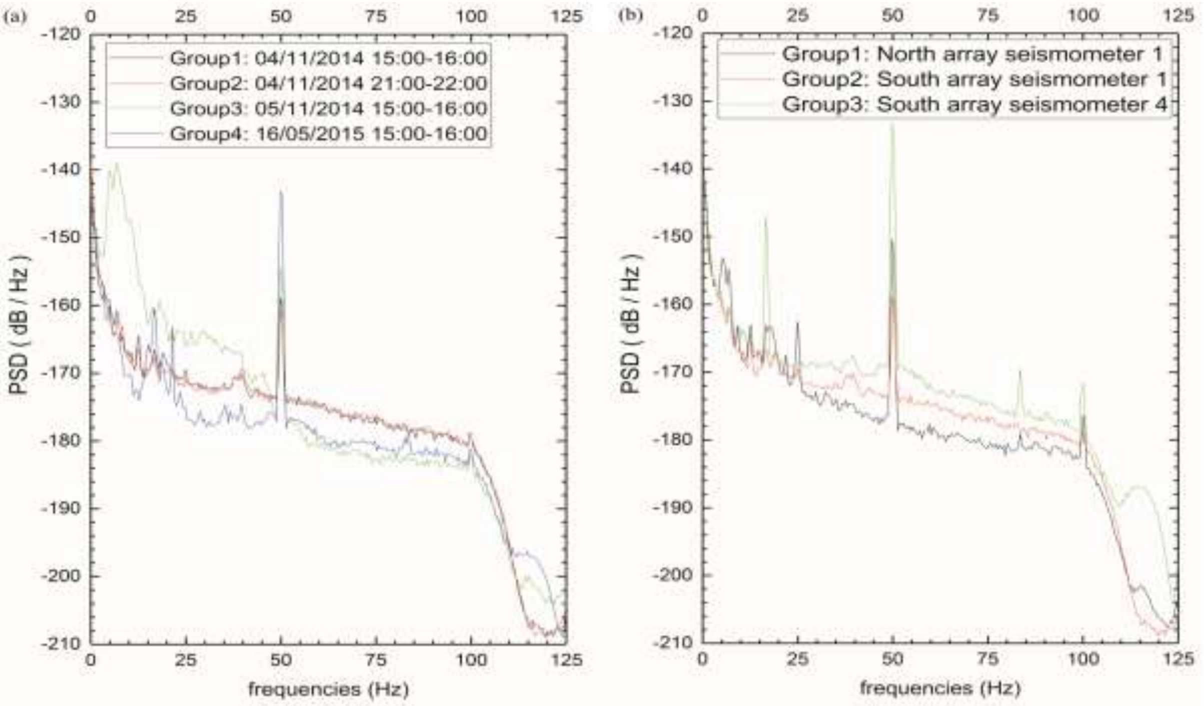


Figure 30: (a) Temporal variation of background noise and (b) spatial variation of background noise

For the temporal subpart I perform an observational study for 4 independent time intervals (TI) (TI 1:4, see inset of Figure 30a). TI1 is the Noise envelope for hour 15:00-16:00 on the 04/11/2014 (working hour), TI2 for hour 21:00-22:00 on the same day (out of working hours, diurnal variation), TI3 for hour 15:00-16:00 (same as TI1) on the 05/11/2014 (monthly variation) and TI4 for hour 15:00-16:00 (again same hour) on the 16/05/2015 (annual variation). For the spatial subpart a cross-sectional study for 3 independent TIs (TI 1:3, see inset of Figure 30b). TI1 is the Noise envelope for a vertical seismometer of the North array for hour 15:00-16:00 on the 04/11/2014, TI2 is for a vertical seismometer of the South array (temporal variation between arrays) while TI3 is the Noise envelope for the 3D vertical seismometer of the South array (temporal variation between different sensors within one array).

At both temporal and spatial analysis subparts the Kruskal-Wallis test (Chan and Walmsley, 1997) is applied. In both the temporal and spatial analysis the [medians (Q_1, Q_3)] are found to be significantly different between TIs at the level of significance 0.05 (see Table 4 for the descriptive statistics of each subpart).

Table 4: Descriptive statistics for temporal and spatial subparts of nonparametric analysis

	Median	**Q1**	**Q3**
		Temporal subpart	
TI1	-175.24	-179.32	-171.46
TI2	-175.26	-179.14	-171.62
TI3	-181.14	-183.63	-165.34
TI4	-179.89	-182.17	-175.94
		Spatial subpart	
TI1	-179.16	-181.54	-173.43
TI2	-175.24	-179.32	-171.46
TI3	-172.56	-177.08	-168.62

This analysis of the background noise demonstrate extremely large, highly unpredictable variations in background noise both between sensors and between consecutive hours/days on a single sensor. No repeatable pattern are determined.

4.4.2 The NpD event detection algorithm

The NpD event detection algorithm (Non-parametric detection) enables microseismic events to be discriminated without any prior filtering of the data.

The algorithm is an alternative detection approach for data sets with low SNR. It is based in the frequency domain by searching and detecting any changes in the PSD spectrum of the data recordings compared to the *Noise PSD*.

The algorithm is described on the basis of continuous recordings $x(t)$ of any duration, though 1-hour durations provide computational and time efficiency. The algorithm is executed in two Steps in order to minimize the computational time required. At the first step, (Step 1) a scan is performed to identify time segments that could potentially contain a microseismic event (or any other signal of interest in the more general case). Only those time segments that are picked in Step 1 are further investigated to detect potential microseismicity, or rejected altogether. The procedure is demonstrated in the schematic in Figure 31 and described in detail below:

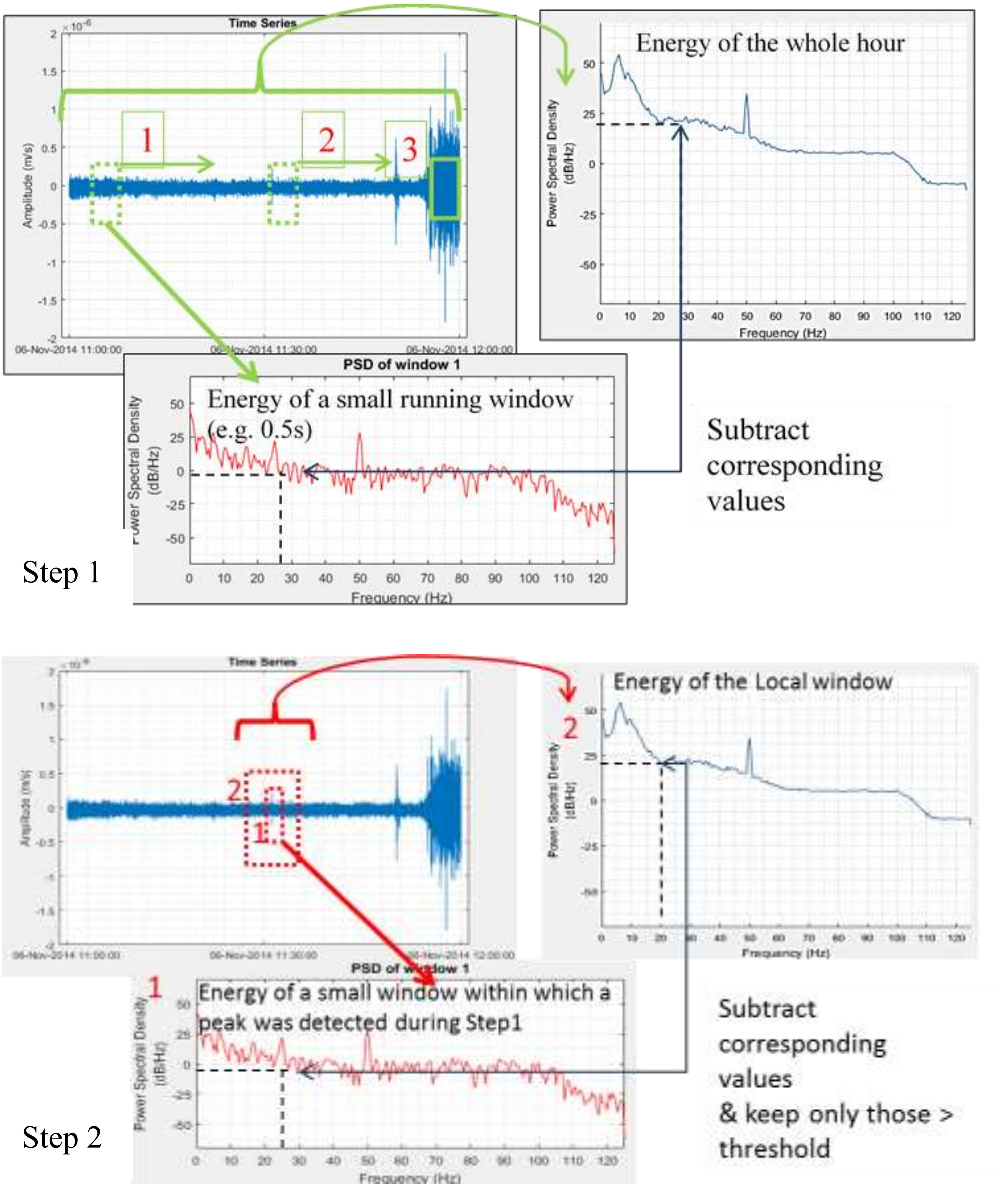


Figure 31: Schematic demonstrating NpD methodology

4.4.2.1 Step 1- Calculation of the *excess energy* over a continuous data record

Following the background noise spectral characterization methodology described in the section 4.4.1, the *Noise PSD* for each data record $x(t)$ is calculated. The individual time segment duration t_l to which the data record is divided, is chosen large enough to be able to accommodate the energy of a microseismic event or a representative energy section of a long-period long-duration event (Das and Zoback, 2011) whilst at the same time small enough to be able to pick closely-spaced events. It is not necessary for the NpD algorithm to include full cycles of the expected signal.

Next, the *Noise PSD* is subtracted from the PSD of each individual time segment forming a set of differences. Within each one of the N_w individual time segments, only the positive differences are kept and summed. This sum is termed *excess energy* which, for each individual time segment starting at time *t*, is given by:

$$PSD\_excess_n^t = \begin{cases} \sum_{f=0}^{Nyq}(PSD_n^t(f) - Noise\ PSD(f)), & if\ PSD_n^t(f) - Noise\ PSD(f) > 0 \\ 0 & , otherwise \end{cases}, \quad (eq.\ 10)$$

where $n = 1,2,..,N_w$

The total number of non-zero only, *excess energy* values, described here as N_I, is equal to or less than the number N_w of the individual time segments that the data record is split to. The results of this process can be graphically presented as a scatterplot with each point's coordinates being pairs of ($PSD\_excess_n^t$, t), with t being the start time of the n^{th} individual time segment.

Not all N_I *excess energy* values are accepted. In data records with highly variable background noise, the detection procedure described so far might result in a number of incorrect detections that do not correspond to events. In order to minimize this possibility, I introduce a threshold value and only accept those ($PSD\_excess_n^t$, t) pairs for which the *excess energy* is above this threshold.

The threshold is determined based on the statistical properties of the *excess energy* values over the duration of the data record analysed; more specifically, the first (*Q1*) and third quartiles (*Q3*) of the *excess energy* values. I then define the threshold value as:

$$Threshold = Q3 + 0.5 \times IQR, \quad (eq.\ 11)$$

where $IQR = Q3 - Q1$.

For the detection of outliers using the quartile values, a commonly used threshold is given by $Q3 + 1.5 \times IQR$, with the 1.5 factor justified by the standard normal distribution and leading to a probability of 99.3% for correctly detecting no outliers (Sun and Genton (2012)). I adopt the value 0.5 as a more conservative threshold.

Only N_2 (out of the total N_1) *excess energy* values are eventually above the *Threshold* and these are processed in the next Step of the analysis (Step 2). This reduces the calculation time significantly.

4.4.2.2 Step 2 - Calculation of the *excess energy* over a local time window

Step 2 is exactly the same as Step 1, but now the *Noise PSD* refers to a local time window rather than the duration of the full data record $x(t)$. This *local time window*, has a predetermined length and is centred around the starting time t of each of the N_2 individual time segments that fulfilled the criteria of Step 1. The total number of *local time windows* used in Step 2 is N_2 and as a result the methodology of Step 1 is repeated N_2 times in Step 2: A *Noise PSD* and then the *excess energy* and *Threshold* are calculated for each one of the N_2 local time windows as described previously.

The times corresponding to the *excess energy* values that are higher than the *Threshold* for each of the local time windows in Step 2 constitute the approximate times where a potential event occurred.

4.4.3 Detected events: Microseismicity or local noise?

A detected potential event from Step 2 could still represent local noise, e.g. steps, drilling noise or even an instrumental glitch. This possibility can be minimized by combining the NpD results from multiple seismometers, for example, from a whole array (voting scheme; Trnkoczy (1999)). A real microseismic event, irrespectively of how small it is, should be recorded by neighbouring seismometers. This is not the case for a local noise burst that is usually recorded by the seismometer closest to it, nor for a mechanical glitch.

The number of seismometers that are required to have recorded the same event depends on the application and the distance between them. A time delay between seismometers for the same event should also be considered.

To avoid having multiple true positives (i.e. correctly identified events) corresponding to different phases of the same event (i.e. different peaks in the same microseismic waveform), I decided to 'clean-up' consecutive events that are detected in consecutive PSD time segments. Consequently, only the first arrival from the consecutives is considered a trigger. This decision was verified during a sensitivity analysis for several hours of data, to ensure that it does not result in missed true positives.

4.4.4 Demonstration of efficiency of the NpD algorithm

Three hours of microseismic data recordings from the North array over two consecutive days are chosen to test the sensitivity of the algorithm to two input parameters: the *percentile* used for the calculations of the *Noise PSD* (Step 1) and the length of the *local time window* (in Step 2): Hour 1: 15/03/2016 18:00 - 19:00 (UTC); Hour 2: 15/03/2016 19:00 - 20:00 (UTC); and Hour 3: 16/03/2016 05:00 - 06:00 (UTC).

Hours 1 and 2 are chosen because after visual inspection were found to contain a number of potential microseismic events. Hour 3 is chosen as a 'quiet hour' with no events visually confirmed. I located a random selection of the visually observed events to confirm that they are indeed events occurring in the surrounding area (within 8 km from the arrays). Three of them were subsequently found in the Swiss Seismological Service catalogue (http://www.seismo.ethz.ch/en/earthquakes/switzerland/all-earthquakes/), having magnitudes down to M_L -0.6.

The visual inspection took place prior to applying the NpD algorithm. For the visual inspection, a bandstop, bidirectional two-pole Butterworth filter was applied to all Hours to remove the AC effect (the arrays were connected to the mains for power supply), as well as a high-pass 2 Hz filter to suppress ambient noise. This was only done for the purpose of visually picking potential events. For the NpD algorithm I used raw data.

Figure 32 and Figure 33 show plots of the filtered waveforms of Hours 1 and 2. The vertical lines above the waveforms indicate the visually observed events that are expected to be detected by the algorithm. I then apply the NpD algorithm for various combinations of percentiles within the range 75 – 95 (for the calculation of the *Noise PSD*) and *local time window* lengths. Table 5 and Table 6 show the best outputs from the sensitivity analyses for these two hours, for each of the arrays individually. The number of the visually observed events is represented by the *Actual no of events* parameter. The number of events that the algorithm detects is represented by the *Detected events* parameter. Those events amongst the detected events that are also within the actual no of events, i.e. visually observed, are the *True positives*. The ratios $R1 = \frac{true\ positives}{detected}\ events \cdot 100\%$ and $R2 = \frac{true\ positives}{actual\ no\ of}\ events \cdot 100\%$ were formed to investigate the efficiency of the various combinations of parameters, where *R1* is the equivalent of precision in computer science, and *R2* is the equivalent of recall/sensitivity. Ratios *R1* and *R2* were introduced to quantify the tendency of the algorithm to trigger false positives, e.g. noise mistakenly picked as an event, and their detection efficiency, respectively. *R1* and *R2* take values between 0 and 100%. A high value for *R1* would indicate a small amount of false positives, while for *R2*, a high value indicates high detection capability. Using these ratios the most efficient combination of parameters was chosen to be the one for which both *R2* and *R1* are at their highest values. From *R1* and *R2* we can derive an f1 score which is the harmonic mean of precision and sensitivity, i.e., $f1=2.(R1.R2)/(R1+R2)$.

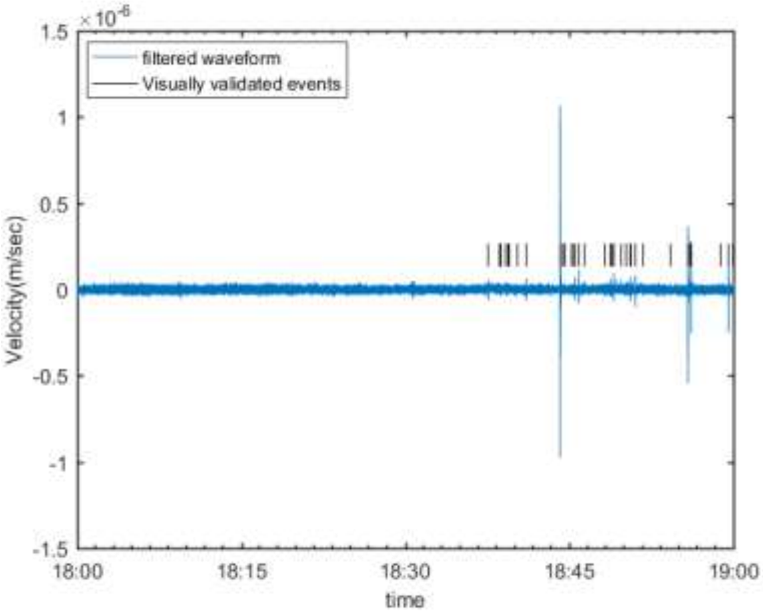


Figure 32: Hour 1: Filtered waveform and visually identified events are shown with vertical lines.

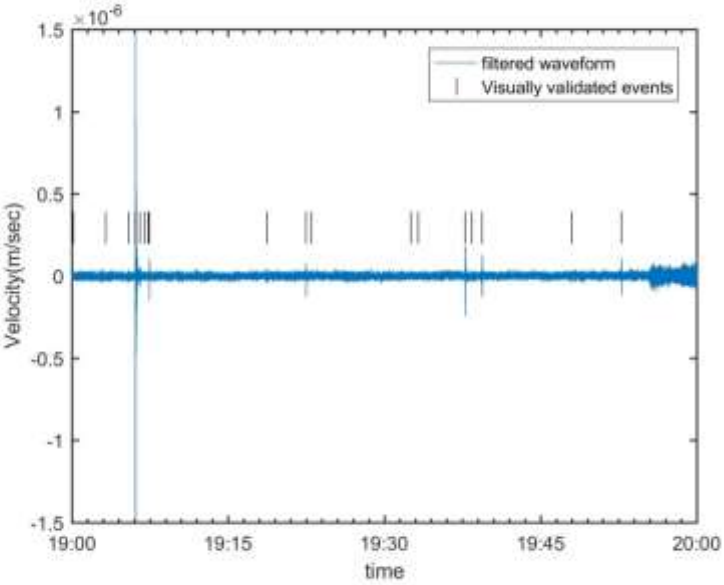


Figure 33: Hour 2: Filtered waveform and visually identified events are shown with vertical lines.

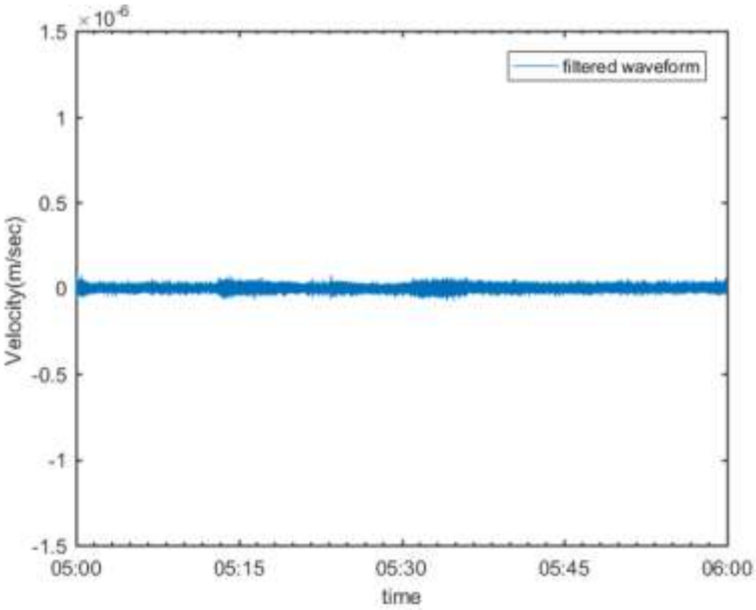


Figure 34: Hour 3: Filtered waveform. Hour with no visually identified events.

As shown in Table 5 and 6, all {*percentile*, *local time window*} combinations yield quite high *R2* ratios (>84%), depending on the location and hour. The differentiating factor is the *R1* ratio. Upon checking other combinations of parameters from the two Tables we also see that the *R1*, *R2* ratios do not vary drastically within a particular hour and array. This means that the assumption that we can treat seismic events as outliers and the choice of a dynamic threshold which adapts well to the statistical properties of each examined segment work well. In the case Hour 3 (Figure 34), the hour for which no visually observed events existed, the low number of events that the algorithm detected was acceptable (Table 7).

Table 5: Hour 1: 15 March 2016, 18:00 – 19:00. Comparison of results for different values of the parameters of *Noise PSD percentile* and *local time window* length.

	Noise PSD percentile:	**70**			**75**			**80**			**85**		
	Local window:	**150**	**300**	**450**	**150**	**300**	**450**	**150**	**300**	**450**	**150**	**300**	**450**
North Array	Detected events:	34	34	34	37	37	37	38	37	36	37	37	36
	True positives:	31	30	29	32	32	32	32	32	32	32	32	32
Actual No of events: 34	*R1*	91%	88%	85%	86%	86%	86%	84%	86%	89%	86%	86%	89%
	R2	91%	88%	85%	94%	94%	94%	94%	94%	94%	94%	94%	94%
	f1	91%	88%	85%	90%	90%	90%	89%	90%	91%	90%	90%	91%
South Array	Detected events:	28	29	28	28	29	28	28	29	28	27	29	29
	True positives:	25	25	25	24	25	25	25	25	25	25	25	25
Actual No of events: 27	*R1*	89%	86%	89%	86%	86%	89%	89%	86%	89%	93%	86%	86%
	R2	93%	93%	93%	89%	93%	93%	93%	93%	93%	93%	93%	93%
	f1	91%	89%	91%	87%	89%	91%	91%	89%	91%	93%	89%	89%

Table 6: Hour 2: 15 March 2016, 19:00 – 20:00. Comparison of results for different values of the parameters of *Noise PSD percentile* and *local time window* length.

	Noise PSD percentile:	**70**			**75**			**80**			**85**		
	Local window:	**150**	**300**	**450**	**150**	**300**	**450**	**150**	**300**	**450**	**150**	**300**	**450**
North Array	Detected events:	30	30	31	30	31	32	29	32	31	28	35	32
	True positives:	18	18	18	18	18	18	18	18	18	18	18	18
Actual No of events: 18	R1	60%	60%	58%	60%	58%	56%	62%	56%	58%	64%	51%	56%
	R2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	f1	75%	75%	73%	75%	73%	72%	77%	72%	73%	78%	68%	72%
South Array	Detected events:	20	24	25	20	24	29	20	28	31	23	32	34
	True positives:	16	16	16	16	16	16	16	16	16	16	16	16
Actual No of events: 19	R1	80%	67%	64%	80%	67%	55%	80%	57%	52%	70%	50%	47%
	R2	84%	84%	84%	84%	84%	84%	84%	84%	84%	84%	84%	84%
	f1	82%	75%	73%	82%	75%	66%	82%	68%	64%	76%	63%	60%

Table 7: Hour 3: 16 March 2016, 05:00 – 06:00. Comparison of results for different values of the parameters of *Noise PSD percentile* and *local time window* length.

	Noise PSD percentile:	**70**			**75**			**80**			**85**		
	Local window:	**150**	**300**	**450**	**150**	**300**	**450**	**150**	**300**	**450**	**150**	**300**	**450**
North Array Actual No of events:0	Detected events:	3	2	2	3	3	3	5	3	3	6	5	3
	True positives:	0											
South Array Actual No of events:0	Detected events:	8	7	9	8	7	12	8	9	15	9	11	15
	True positives:	0											

For this project, the combination of parameters that best suited my data for identifying as many seismic events with the least possible false positives was the 75th percentile for the calculation of the *Noise PSD* (Step 1) and a 300 second duration for the *local time window* (Step 2 of the NpD algorithm) as this combination demonstrated consistently high *f1* values.

4.4.5 Detected events

The full data set of passive seismic data acquired between 1st November 2014 and 12th August 2017 was scanned using the NpD algorithm resulting in more than 241,000 detections. Figure 35 shows the number of detections per day over the whole monitoring time period, along with the changes of the water level in the Raeterichsboden lake over the same time period. The black, green and red markers show detections at the South, North and borehole arrays respectively. Power failures or mechanic malfunctions of any of the sensors in each of the arrays resulted in broken data records. Some of the data, even for several consecutive days, were corrupt and could not be processed (discussed in 3.5.3). Since the NpD algorithm uses all sensors of each array to identify events, lack of or corrupted data even for just one sensor results in the data gaps visible in Figure 35 (Jun.-Jul. 2015, Jan.-Feb. 2016 and Mar.-May 2017). That could be future work, i.e. to modify the NpD algorithm so that it detects events with variable numbers of events and then assigns weights to the results depending on the number of sensors that successfully detected the potential event.

Several features can be observed from Figure 35 and are discussed in more detail in the following paragraphs:

- the numbers of events detected by each of the individual arrays are broadly consistent,

- an anomalously high number of events are detected by the North and South arrays in early 2015, which may be as a result of reservoir drainage and refilling,
- the number of events appears to peak annually in the summer months (possibly related to the increase in glacial meltwater and consequent increase in reservoir levels, e.g. Johnson et al. (2017) and Roth et al. (1992),
- a peak in event numbers may be associated with reservoir drainage.

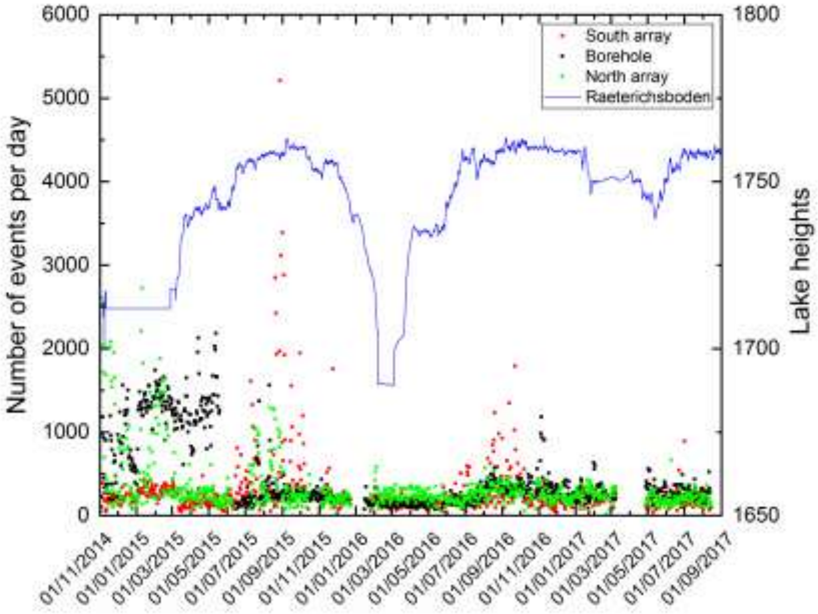


Figure 35: Detected events after the application of the NpD algorithm to the whole passive seismic data set (1st November 2014 - 12th August 2017), along with the changes of the water level in Raeterichsboden lake. Events detected from the borehole sensor, the North and the South array are presented with black, green and red dots, respectively.

Figure 36 shows the number of detected microseismic events per day over the whole monitoring time period, along with the changes of the water level in both the Raeterichsboden and Grimsel lakes over the same time period. All undertaken and presented analysis from here onwards will focus on the seismic data collected at the two surface arrays as explained in previous chapter. Hence the black and green markers show events detected at the South and North arrays respectively, while the borehole data (shown in Figure 36) were omitted from this graph. Additionally in Figure 36 are noted with red ovals, the periods that either one or both arrays have reported a significant increase in detected event frequency. Marked with red arrows in Figure 36 are noted the end days of the two drainage periods of Raeterichsboden lake. It is obvious that the two seismic arrays express increases in detections at the same times. Regardless whether these end up to actually be microseismic events (to be confirmed by the next step of the analysis), the two arrays seem to be consistently triggered.

Grimsel lake is at a greater distance from the GTS (1km to the south of the mid-point between arrays) than Raeterichsboden lake (60m to the east) but is of greater volume and hence the lake height differences (caused mainly due to hydropower management operations and temporal variations, e.g. from rain, snow) could potentially have some contributory effect to the number of recorded microseismic events. Of the two arrays, detection of events by the South array is most likely to be influenced by seasonal variations in the water level of Grimsel lake, since this array is situated approximately equidistant from both lakes. This is confirmed by analysis of the data in Figure 36 with the high frequencies of microseismic events in August-Oct of each year being predominantly recorded by the Southern array. Apart from the seasonal variations apparent in Figure 36 we also observe an increase in the number of events detected by the Northern array, possibly associated with lake drainage (11/2014 – 03/2015 and 02/2016).

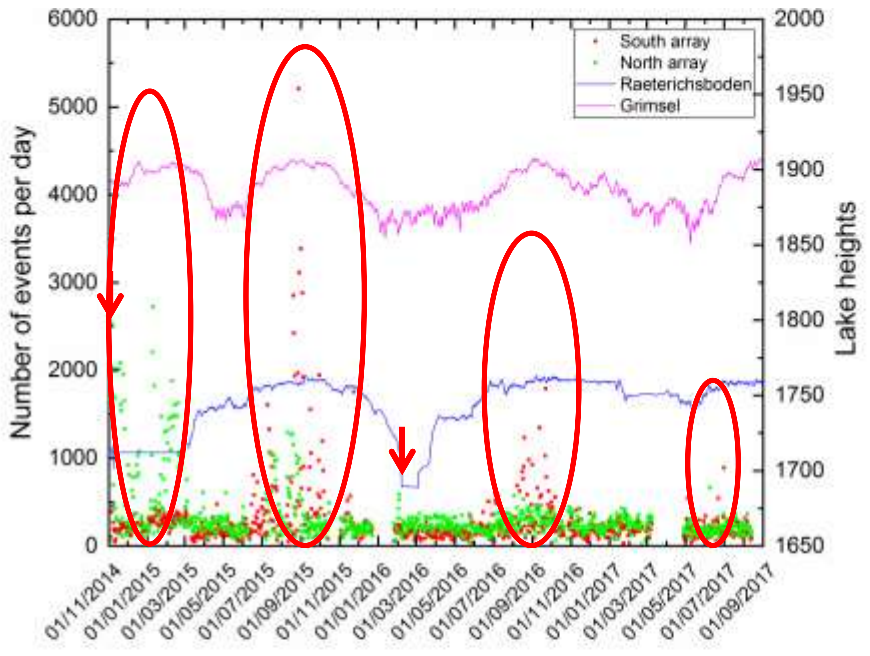


Figure 36: Detected events from the North and the South array, presented with green and red dots, along with the changes of water level of Raeterichsboden and Grimsel lakes (blue and magenta). With red ovals the periods that either of the 2 arrays has reported a significant increase in detections are noted. With red arrows the end of the two drainage periods of Raeterichboden lake are noted.

A quantitative analysis of the detected events during the time periods denoted in Figure 37 by the first three oval circles (01/2015: 03/2015, 07/2015: 10/2015 and 07/2016: 10/2016) is shown in Figure 37. For this analysis box plots were created for the aforementioned time periods and are shown in the figure in red and green for the South and North arrays respectively. Also plotted in Figure 37 is a box plot of daily event frequency for all of the remaining time periods (i.e. all the data not within the first three red ovals). In the plot, the mean value is represented with a square and the maximum and minimum values with horizontal lines. A horizontal line within the boxplot represents the median while the range of each boxplot is from the 25th till the 75th percentile. The whiskers are calculated with 1.5 coefficients and the outliers are represented with Xs. It can be visually observed that there is a clear difference between the statistics for the background seismicity (pair of box plots on the right hand side of the plot), the detected seismicity during the period of lake drainage (the first pair of box plots) and the peaks attributed to seasonal variations (middle two pairs of box plots). It is confirmed by the boxplots also that of the two arrays, the South array is more influenced by seasonal variations, which was previously attributed to its relevant location (equidistant) to lakes Grimsel and Raeterichsboden.

To investigate whether similar seasonal variations in event frequency are apparent in the regional seismicity (i.e. the larger magnitude events detected by the Swiss seismic network) I used data from the International Seismological Centre (ISC; isc.ac.uk) database. A bar chart of the number of events per month within a 30km radius from the GTS is presented in Figure 38. The lake height changes for Lake Grimsel are also shown in blue for comparison. The frequency of these larger magnitude events also exhibits some seasonal variation, but the pattern is not as clear. This is to be expected, since larger events are likely to be deeper and hence less affected by variations in the surface water and groundwater pressure heads.

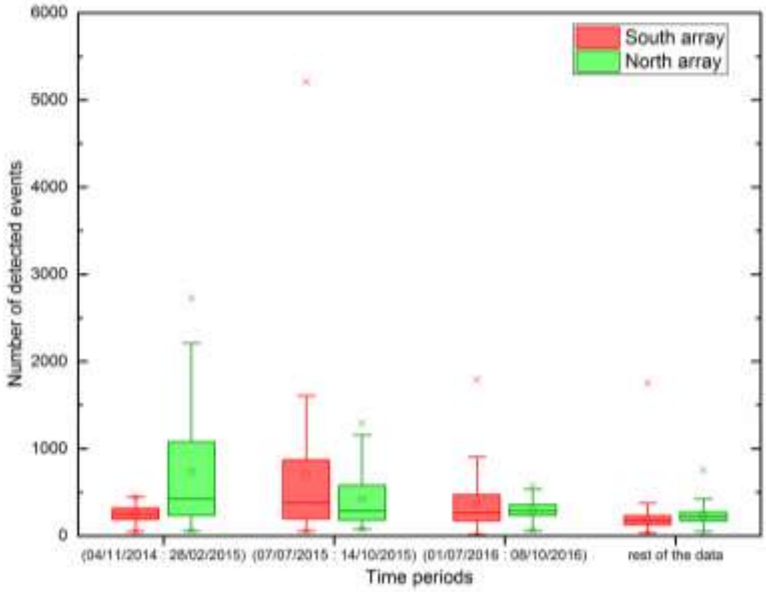


Figure 37: Boxplots for 3 different time periods (01/2015 - 03/2015, 07/2015- 10/2015 and 07/2016 - 10/2016) for the South and North array (shown in red and green respectively). The mean value is represented with a square and the maximum and minimum values with horizontal lines. A horizontal line within the boxplot represents the median while the range of each boxplot is from the 25th till the 75th percentile. The whiskers are calculated with 1.5 coefficients.

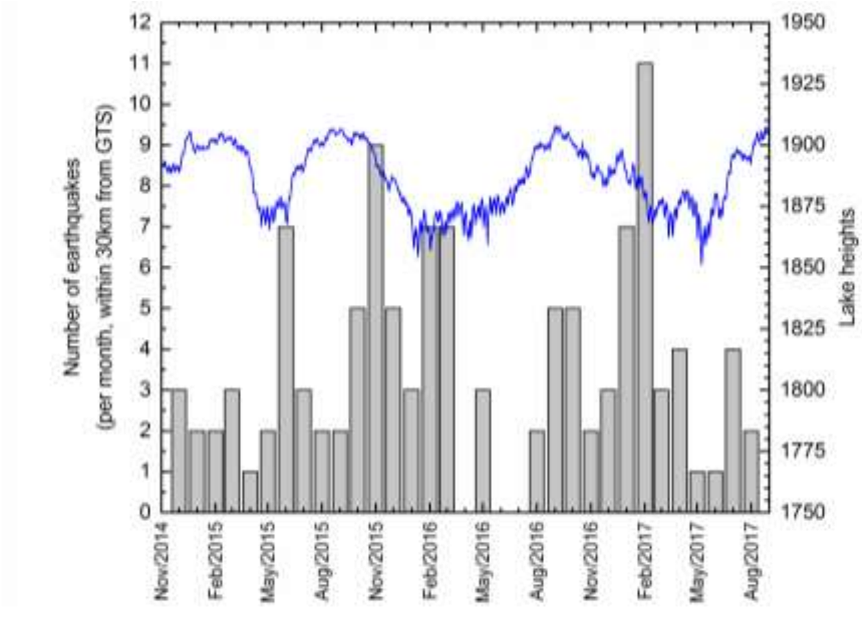


Figure 38: Reported earthquakes per month within an area of 30km radius from the GTS for the period Nov 2014 – Aug 2017. Data retrieved from the International Seismological Centre (ISC). With a blue line Raeterichsboden lake height levels are noted.

Of course to attribute the source of increase of event frequency to the lake drainage or refilling is not an exact science. As discussed in literature, a seismic response can be the result of previous unloading and reloading of the reservoir expressed with a time delay. Or it may be the case that the response to the lake load change in years 2014-2016 is expressed during the next years and is not part of the events I am to analyse. It is also possible that the unloading / reloading does not have a direct effect on the granitic rocks which would potentially need a stress change higher than historic stress changes to respond in turn by demonstrating seismic activity / discharge of energy. Let's not forget that the Grimsel area has been a highly investigated area and hence the surrounding rocks have developed some sort of tolerance to continuous stresses changes as part of different experiments and is not responding in the same way that a more virgin granitic environment would have. It may well also be that this increase of events is not at all an effect of the lake load changes; it is essential to first locate the events and then proceed with interpretations since the source locations, regardless of their accuracy (but always within acceptable limits), will shed some light and inform on seismicity trends.

4.4.6 Comparison of NpD with other approaches

In order to check the effectiveness of the NpD algorithm I compare its performance to that of the most commonly used detection algorithm, namely STA/LTA, and the algorithm suggested by Vaezi and van der Baan (2014).

For the comparison, I choose the same three hours (15/03/2016, hours 18:00-19:00 and 19:00-20:00 and 16/03/2016, hour 05:00-06:00 shown respectively at Figure 32, Figure 33 and Figure 34) from the GTS data set with varying background noise levels and with, and without, events. Table 8 shows the parameters used for each of the three detection methodologies used in the comparison. The detection thresholds in all methods are selected in such a way as to give the best balance between false positives and missed events for each algorithm. The *minimum event duration* parameter for the STA/LTA method is the minimal time length between the time of an event triggering and detriggering. The *minimum event separation* parameter specifies the minimal time length between the end of a previous event and the beginning of a new event. The STA and PSD window lengths are kept the same and equal to 0.5s to allow for a valid comparison of the algorithms. The same applies for the LTA window length and local window. The *consecutive events cleaning* parameter presumes that when the output peaks are consecutive within distances of 0.5s they correspond to the same event. All algorithms have been implemented in a multi-channel strategy in which events are detected only if they are detected by all vertical channels of each array.

Table 8: The parameters used for the STA/LTA, PSD technique and NpD methods.

STA/LTA parameters		***PSD technique parameters***		***NpD parameters***	
STA window length	0.5s	PSD window length	0.5s	Individual time segment duration	0.5s
Minimum event duration	0.005s	Window overlap	50%	*Noise PSD*	75th
Minimum event separation	0.5s	Minimum event separation	0.5s	Consecutive events cleaning	0.5s
LTA window length	5mins	_		Local window	5mins
STA/LTA detection threshold	2.5	PSD detection threshold	0.50	Dynamic detection threshold	Q3+0.5IQR

Results are summarized in Figure 39 and Table 9. Figure 39 shows the filtered (bandstop 48-52 Hz to remove the AC effect) waveforms of the three hours examined previously, both as recorded from the North (a, c, & e) and the South Array (b, d & f). The vertical lines show the detection times obtained by the STA/LTA, PSD and NpD algorithms (see inset for details). From just visual inspection, it is noticeable that the STA/LTA detects very few events and the PSD algorithm detects many more events than the NpD. In Table 9 we can see the breakdown of these detected events to true and false positives. The ratios *R1* and *R2* are once again used to quantify the fraction of the total number of detected events that were visually observed (*R1*) and the fraction of the visually observed events that were detected (*R2*).

Table 9: Summary of detections using the STA/LTA, the PSD and the NpD algorithms for hours 1, 2 and 3, for both North and South arrays.

			STA/LTA algorithm	**PSD Picker**	**NpD algorithm**
Hour 1: 15/03/2016, 18:00 - 19:00	**North Array**	Detected events:	4	123	37
	Actual No of events: 34	True positives:	4	32	32
		R1	100%	26%	86%
		R2	12%	94%	94%
		f1	21%	41%	90%
	South Array	Detected events:	3	102	29
	Actual No of events: 27	True positives:	3	24	25
		R1	100%	24%	86%
		R2	11%	89%	93%
		f1	20%	38%	89%
Hour 2: 15/03/2016, 19:00 - 20:00	**North Array**	Detected events:	12	97	31
	Actual No of events: 18	True positives:	3	18	18
		R1	25%	19%	58%
		R2	17%	100%	100%
		f1	20%	32%	73%
	South Array	Detected events:	13	140	24
	Actual No of events: 19	True positives:	1	16	16
		R1	8%	11%	67%
		R2	3%	47%	84%
		f1	4%	18%	75%
Hour 3: 16/03/2016, 05:00 - 06:00	**North Array** Actual No of events: 0	Detected events:	1	1	3
		True positives:	0	0	0
	South Array Actual No of events: 0	Detected events:	3	15	7
		True positives:	0	0	0

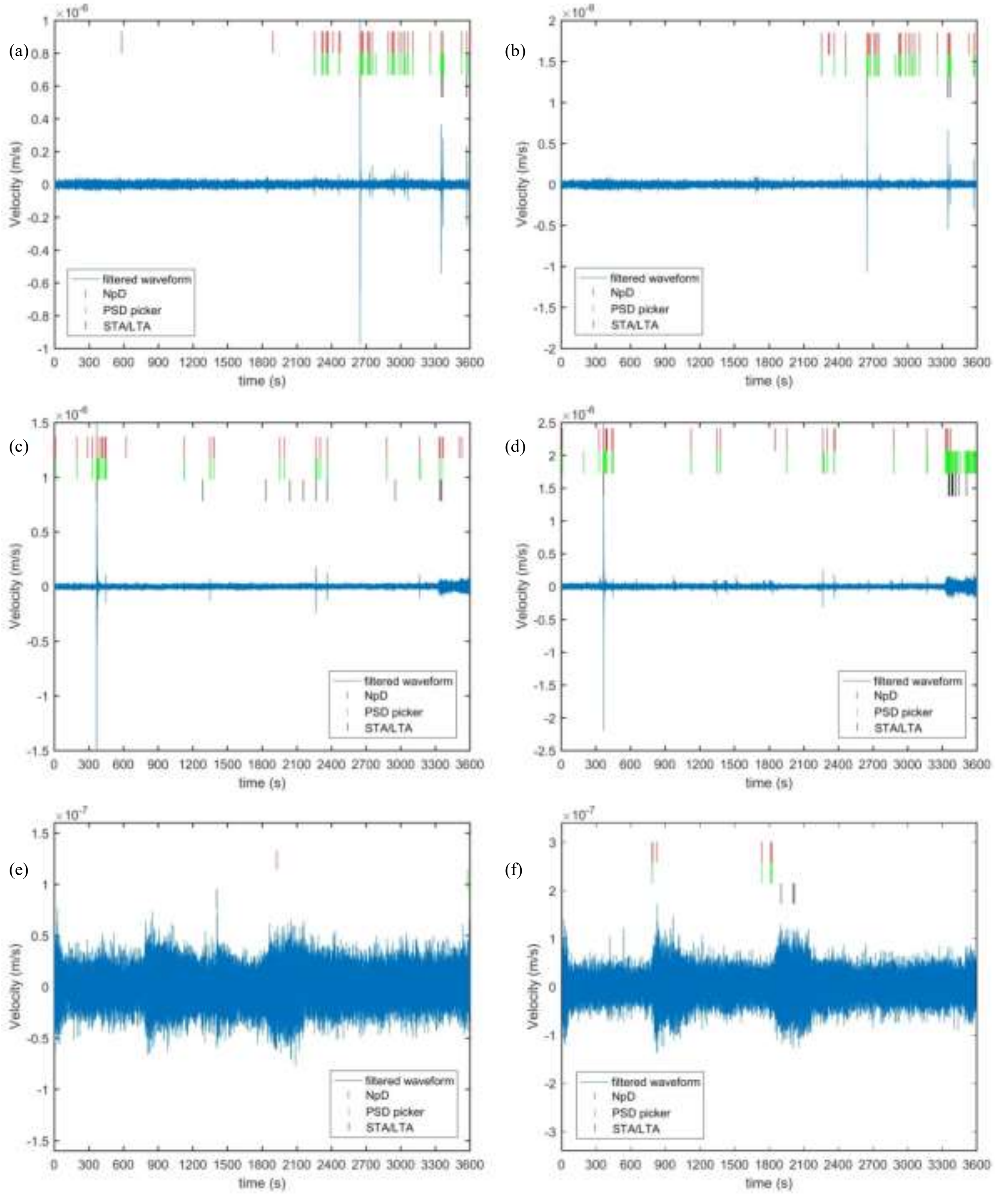


Figure 39: Velocity vs time for the filtered waveforms of (a & b) 15/03/2016, 18:00-19:00, (c & d) 15/03/2016, 19:00-20:00, and (e & f) 16/03/2016, 05:00-06:00 as recorded from the North and South array respectively. With vertical lines the events detected by the NpD algorithm, the PSD technique and the STA/LTA algorithm are noted.

As seen from Table 9, the STA/LTA algorithm is outperformed by both the PSD picker and the NpD algorithm as its ability to detect events, when using unfiltered recordings is significantly smaller (small values of the *R2* ratio). The NpD algorithm also outperforms the PSD picker. For those hours containing events, the NpD algorithm detects the same number of true events as the PSD picker. However, the value of the *R1* ratio is consistently higher for the NpD algorithm than the PSD picker, indicating that the number of false positives from the NpD algorithm is significantly smaller. The harmonic mean of precision and sensitivity *f1* is, therefore, consistently higher in the case of NpD, i.e. for instance, in the North array, 90%, 41%, 21% are the *f1* percentages for NpD, PSD Picker and STA / LTA respectively in the case of Hour 1, and 73%, 32%, 20% in the case of Hour 2.

In the last tested hour, where there are no seismic events, the STA/LTA, PSD and NpD algorithms detected 1, 1 and 3 at the North and 3, 15 and 7 false positives respectively. For this hour, the STA/LTA is the best performing algorithm, with the smallest number of false positives. However, the other two hours show that this is at the cost of missing large numbers of small events with amplitudes close to noise level (low SNR). If a seismic array is deployed for decision-making processes, such as an early-warning system for landslides, then visual validation of detected events may be required by the operator (e.g. if road closure results in a long detour). This manual quality control is a time-consuming procedure. The very low number of false positives that my NpD algorithm detects, by comparison to the STA/LTA and PSD detection algorithms, ensures that expensive operator time is minimized.

4.5 Onset time picking in the present study

In this PhD I chose to manually pick the arrival times (further discussion on this is provided in Chapter 4.6). Manual picking of the arrival times involves an extensive amount of pattern recognition and the final decision is partly subjective. The researcher visually inspects the amplitudes changes and waveform frequency changes. It is very important to pick the first breaks in a consistent and, as possible, objective way. For the visual inspection of the data records within the time periods of interest RTQT_View from REFTEK was used. The data in counts were filtered with a recursive Butterworth bandstop filter for the frequencies between 48 and 52 Hz.

Manual picking of P and S waves proved to be not just time consuming but also very challenging. The sampling rate of 250Hz proved to not be adequately high to differentiate the arrival time differences of certain signals between seismometers. In Figure 40 for instance, we can see that there is a sequence of events lasting approximately 7.5s. However, the exact P- and S-waves arrival times cannot be recognised as the boundaries of each signal are mixed within the coda. Out of this sequence it is possible to estimate only for 2 events approximate epicentral distances from the 3D seismometer of the North Array being less than 200 m (using the single location method) with magnitudes M_L -0.28 and -0.78.

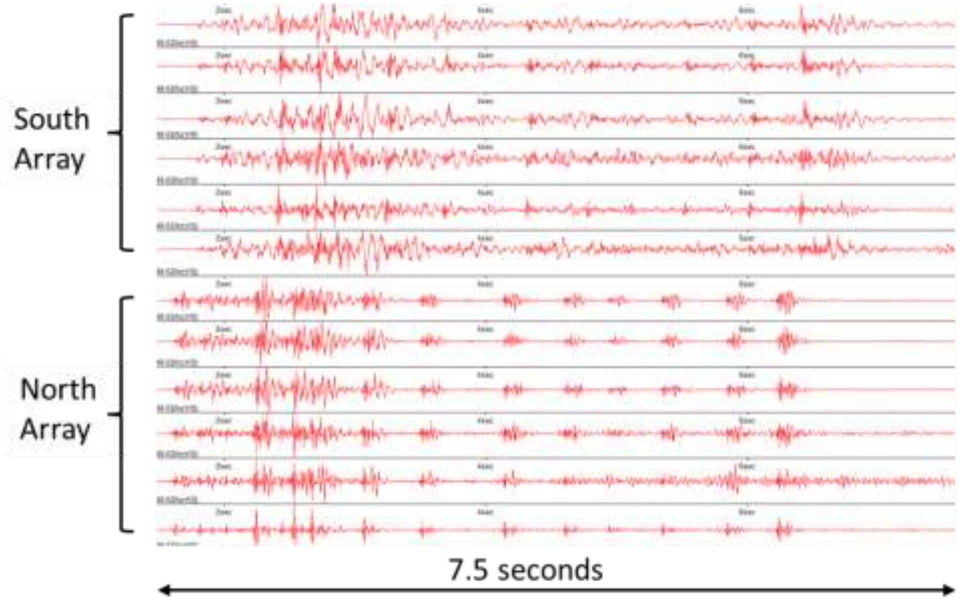


Figure 40: Detected events that could not be located accurately: 26 Nov 2014, 15:28:01

I manually re-picked individual events, identified with NpD algorithm, within periods of particular interest (seen in Figure 41). These time periods are chosen with specific criteria in mind: the chosen dates followed either after the unloading of Raeterichsboden Lake (see grey shaded areas), were two days before and after the times that a PH peak was detected in the PH measurements, or within the specific hours that a stress peak was detected in the Czech measurements (see orange arrows). These dates were expected to be the most probable for a microseismicity reaction to a stress change. They were also the ones that I could combine the scientific findings with the LASMO hydrogeological research conducted.

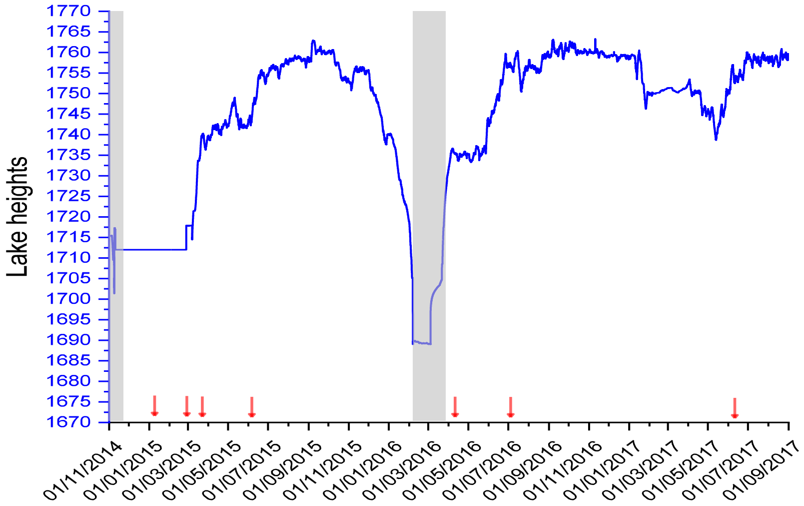


Figure 41: Figure showing the time periods with particular interest within which manual picking of events was undertaken. Manual picking also occurred on the dates shown with red arrows.

4.6 Discussion

4.6.1 NpD algorithm

In this chapter I presented a new algorithm for the detection of microseismic events at environments with low SNR. The main advantage of my approach is that it does not require any pre-filtering of the data as would be the case for detection of weak signals with most other methodologies. Pre-filtering assumes a priori knowledge of the expected microseismic signals which is seldom the case for passive monitoring applications. As a result, pre-filtering could remove information from the recordings, discarding it as noise, especially in cases of low SNR data. Avoiding pre-filtering altogether, minimizes the possibility of information loss in these low SNR recordings.

Another advantage of the NpD algorithm is that it is suitable for non-stationary background noise since the upper bound to the spectral amplitude of background noise, above which an event is detected, varies over both space and time; significant differences were observed in hourly noise characteristics between sensors 1km apart. The approach is also equally effective with non-parametric data i.e. an assumption of normality is not required.

Although considerable effort was put to avoid having multiple NpD peaks for the same signal, i.e. consecutive events in consecutive PSD time segments were "cleaned up" and only the first arrival from the consecutives was considered a trigger, it was observed, that in the cases of landslides for instance (see Figure 42) this was ineffective. The initial crack (inset) is perceived as a separate event to the main body of the landslide and NpD triggers multiple peaks for an individual event in this occasion. This could be a topic to be investigated in the future in order to minimise detections of non microseismic events; a potential way to deal with this is to explore classification algorithms (e.g. machine learning, neural networks, etc.).

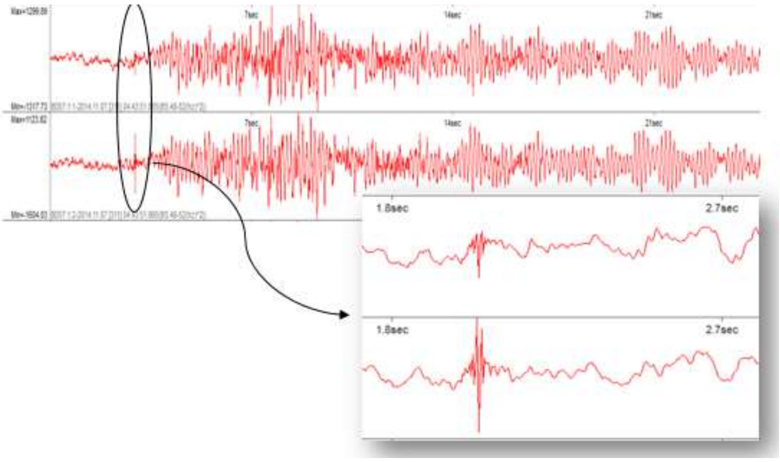


Figure 42: Landslides were generally a bug for NpD as they were picked as more than one distinct event (initial crack and main body of landslide).

The NpD algorithm is a powerful microseismicity detection tool but its output does not include accurate onset times for the detected events. Its accuracy depends on the duration of the individual time segments to which each recording is divided. For windows of duration 0.5 seconds, such as those used in this case study, it means that the onset time is within a 0.5 second frame centred around the estimated NpD time of the 'event'. For a more accurate determination of the onset time, the NpD would need to be combined with other existing automated picking algorithms, such as autoregressive techniques (Kong, 1997; Leonard and Kennett, 1999; Oye and Roth, 2003). Another idea for further improvement of the NpD code is to explore whether minor tweaks can turn it from a detection code to an onset time determination code. Potential ideas to explore would be to check the efficiency of wavelet decomposition instead of Fourier transform as wavelet decomposition does not compromise time for frequency accuracy. Also, another recommendation would be to record timeseries with a high sampling rate which would allow for smaller windows hence greater accuracy in the detections.

Such a deviation between the real arrival and the P-wave picking of the NpD algorithm would produce highly inaccurate locations of events. To be able to successfully identify the onset time of the thousands of events detected by the NpD algorithm is not an easy task given the complexity of the coda and the weak S and sometimes P phases. The quality of the P phase picking is related to the near-surface structure, source type, and S/N ratio (Sabbione and Velis, 2010). Hence, the initial onset time picking attempts (prior to proceeding with the manual picking of the onset times) focused on automated approaches.

4.6.2 Partially effective automated onset time determination approaches

An initial attempt to use an already available commercial software, InSite, from IMaGE (http://www.itasca-image.com/software/insite) proved unsuccessful. InSite, as other available off-the-shelf software, cannot be applied to continuous recordings; it works with seismic recordings acquired using 'triggered' as opposed to 'continuous' mode. As a consequence, the seismic data form a set of signal waveforms, each saved in a separate file, or in the same file but with a header before it. That would demand an enormous amount of disk storage space for the thousands of detected events. Therefore, this option was rejected.

Instead, an algorithm was developed in R studio programming language; a hybrid approach, with a simple formulation and parameter settings, that combines single and multilevel based existing methodologies for the accurate picking of the P and S arrival times. The algorithm is very similar to the workflow presented by Zhang et al. (2003) but was tweaked to fit my real field microseismic data needs (see chapter 4.6.2.2). The workflow used, called *AIC-wav*, is presented as follows.

The post-acquisition processing algorithm *AIC-wav* takes small windows (approximately 3 seconds duration) around the times of potential events derived by the NpD algorithm. To these windows of data (velocity values), a discrete wavelet transform (Daubechies wavelet transform filter of length 2) with 2 levels of decomposition is applied (Percival and Walden 2000).

Following this, the Akaike information criterion (AIC) of the wavelet coefficients for each of the vertical channels and for all scales is computed. To circumvent the border effect of the window, I define a border-effect region at both ends of the time window within which region AIC is not computed. For each detected event, the P time onset should coincide with the time of occurrence of the minimum AIC value. If this P onset time lies within a range of 50 data points (i.e. 0.2 seconds) in all vertical channels of each array, the P time onset value is accepted. If not, then the event detected by the NpD is rejected as not having a discernible P-wave.

The same procedure is repeated for the S-wave, using a smaller window that starts from the P wave onset time up till the end of the original window, avoiding the border effect region. Once again, the time of occurrence of the minimum AIC value is identified as being the S wave onset time in each of the two horizontal channels of the 3D sensor (as these detect S-waves more strongly). If the S wave onset time lies within 50 data points (i.e. 0.2 seconds) range in both horizontal channels recordings, then the event is considered potentially locatable. As a last step for each array, the S onset times are checked for consistency against the waveforms recorded at the three one-component vertical sensors of the same array.

4.6.2.1 *AIC-wav* results

The *AIC-wav* worked exceptionally well in cases that the signal was visually distinct to noise. By way of example, the automatic P and S picks determined for a single NpD-detected event recorded at the North array is shown in Figure 43. The waveforms of the detected event, as recorded by all 6 channels of the North array, are shown in black for a time window that spans approximately 400 sample points (1.6 seconds). The top three plots correspond to the three vertical component sensors, while the last three correspond to the channels of the three-component sensor (top: vertical, second last: horizontal North-South, last: horizontal East-West). The arrival times for the P and S waves are shown with red and blue vertical lines, respectively. Note that while the y-axis represents velocity (m/sec), the x-axis is number of samples rather than time. As expected, the P and S waves reach each individual sensor at slightly different times consistent with the direction of travel from the earthquake source.

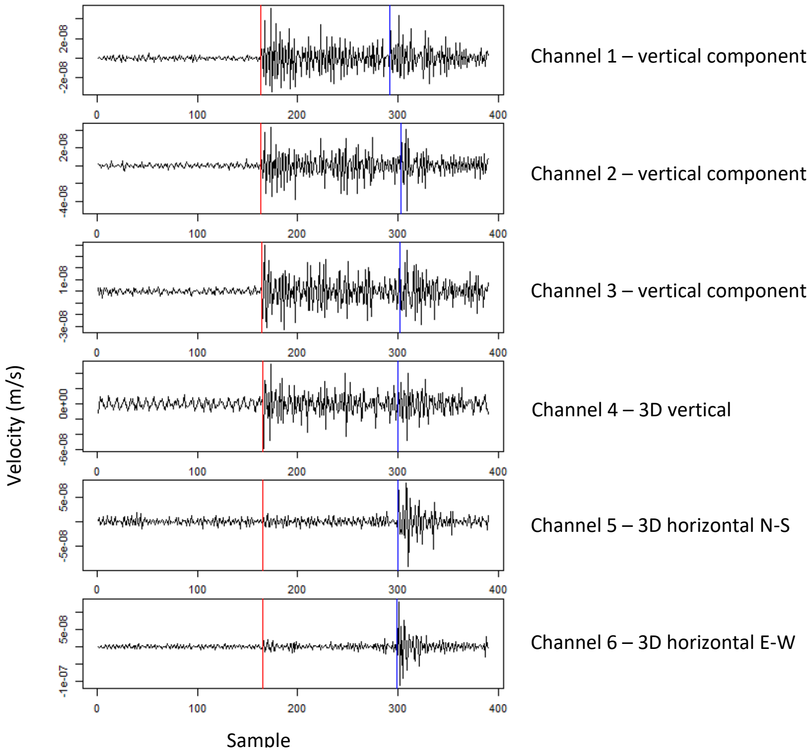


Figure 43: Example of automatically picked P and S wave arrival times at the recordings of the North array for a detected event. Red lines: P wave arrival time. Blue lines: S wave arrival times.

The full record (1st November 2014 - 12th August 2017) was scanned to determine P and S wave arrival times at all sensors. Out of the 281,351 and 239,280 potential events detected by the NpD for the North and South array respectively, a total of 29,214 events (North array) and 22,820 events (South array) were the output of *AIC-wav* workflow. Figure 44 demonstrates the output of the *AIC-wav* algorithm.

It is important to highlight that in the case that P- and S- onset times were not able to be consistently picked for all vertical sensors and one horizontal of each array the *AIC-wav* algorithm was programmed as such that would not return any results.

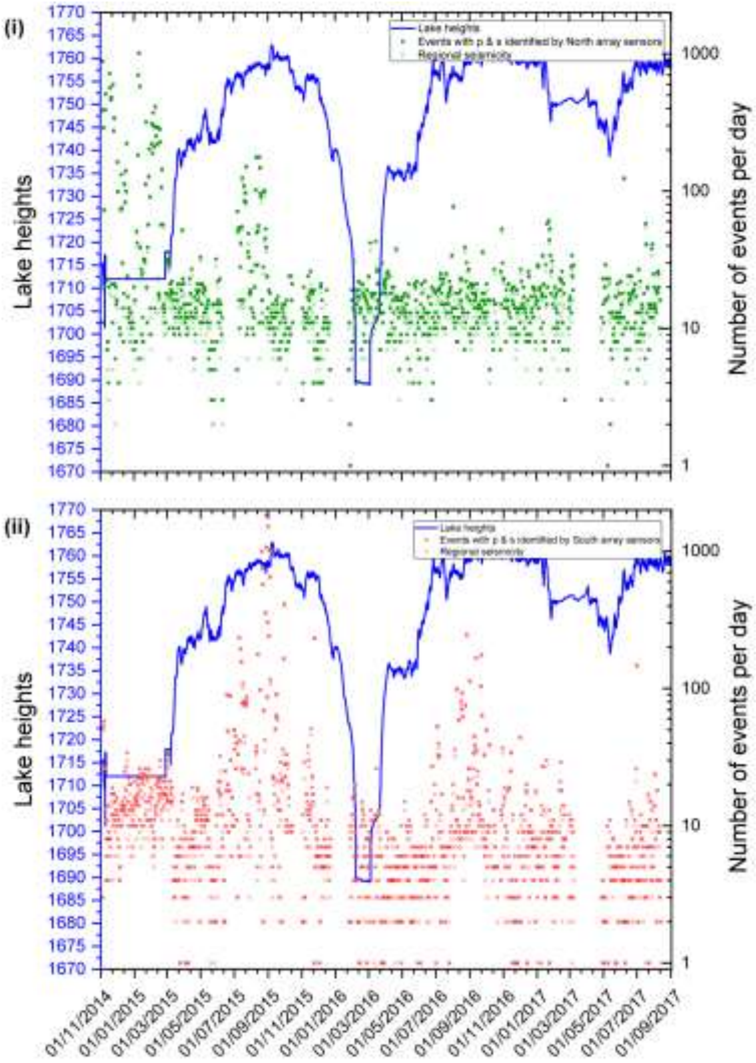


Figure 44: Onset time picked events of (i) North and (ii) South array (to be sent for location) excluding events located within the Gerstenegg tunnels (150m distance form sensors crossing line), along with the changes of the water level in Raeterichsboden lake. Events detected from *AIC-wav* and the NpD algorithm are presented with dark green/red dots and green/red x-s, respectively. Seasonal variation is a sign of localised seismicity.

4.6.2.2 *AIC-wav* workflow variations

Using an hour of data with manually pre-picked events, different variations of the workflow presented above were explored and dismissed due to ineffectiveness. These included:

- checking between the pick times at scale 1 and scale 2 are within 24 samples and those between scales 2 and 3 within 48 samples, and then declare the pick at level 2 as a P-wave arrival (Zhang, 2003). Having checked the three details obtained, it was found that in the third stage detail seismic noise was more coherent (Hafez et al., 2010) and the algorithm recognized P-wave arrivals only in high S/N ratios.

- using correlation among the pilot waveform of one of the horizontal channels and waveforms of the vertical sensors, in each array, for the correct identification of the S waves. This approach did not work effectively in all cases because the S/N ratio of some of the channels was not high enough to allow for the correct time lag to be determined.
- applying AIC on autoregressive series (forward and backward) instead of using wavelet decomposition to correctly pick the P- and S- wave onset times (Oye and Roth, 2003). This approach worked effectively only in high S/N ratios.
- applying AIC on the raw rata instead of using wavelet decomposition to correctly pick the P- and S- wave onset times. This approach correctly identified more P-wave picks than *AIC-wav* but the error between manual picks and algorithm picked picks was greater.
- using bandstop recursive filter 48-52Hz (EM interference in Switzerland due to the mains frequency is at 50Hz) before following the rest of the workflow. A really small subset of the events was picked in this case. I expected that to happen given that bandstop filtering is not that effective when signal and noise bandwidths overlap.
- using denoising or soft thresholding as in Zhang (2003) which decreases the distortion of the P-wave but affects its amplitude in a lesser extent before following the rest of the workflow. Results in this case were similar as in the bandstop filter, as many of the visually identified events were discarded by the algorithm.

4.6.2.3 Quality assessment of *AIC-wav* onset time identifications

Quality check (through visual checks) performed in the identified onset times of *AIC-wav* workflow, indicated that although *AIC-wav* worked better than all other used workflows, it does not identify the onset times optimally in all occasions. In Figure 45 we can visualise different cases that the *AIC-wav* methodology (i) works well in picking the P- and S- onset times; (ii) works well in picking only the P-wave; and (iii) fails to correctly identify the phases of the signal. In (iii) the algorithm failed to correctly identify the S wave onset time of the horizontals hence the vertical was also affected by this. In reality the signal has lost the P wave and the picked as S wave should have been picked a bit later – this one that is picked is actually a misidentification of the S-wave as a P-wave arrival. This misidentification of the first arrival has also been observed by Ge (2005); his analysis showed that in a sample of 434 picked events, S-wave arrivals account for the 41% of the total picks while outliers account for the 10%. Potential misidentifications like the aforementioned would introduce significant and systematical errors in further analysis. Overall, *AIC-wav* workflow allows more confidence in the automatically picked P-waves rather than the S-waves.

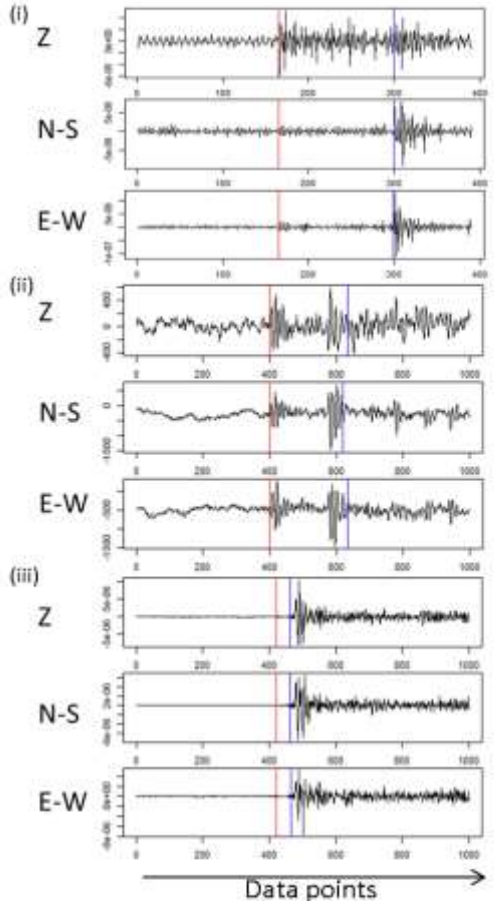


Figure 45: Cases the *AIC-wav* methodology (i) works well in picking the P- and S- onset times; (ii) works well in picking only the P-wave; and (iii) fails to correctly identify the phases of the signal. In all (i), (ii) and (iii) the three. The axis in all waveforms is in data points while y-axis is in velocities (m/s). The range of the x axis is not common for all graphs so that the waveform shape is visible in all cases.

4.6.3 Detections

With regards to the events detections, it should be considered that part of the seismic response may be attributed to an annually recurring hydrologically related causal mechanism. An annual cycle of detected seismicity has also been reported by Wolf et al. (1997) for networks in British Columbia and Alaska and was attributed solely to the seasonal variations. In particular, the researchers observed two pulses, one between the months of May and July, and another one between late August and October. Statistical analysis demonstrated that there is a clear difference between the statistics for the background seismicity, the period of lake drainage and the peaks due to seasonal variations. Regional seismicity was also checked for seasonal variation and it was found that it also exhibits some, but the pattern is not as clear as with the detected seismicity by the NpD algorithm. This is to be expected, since larger events are likely to be deeper (e.g. Mousavi et al., (2016)) and hence less affected by variations in the surface water and groundwater pressure heads.

It has been reported in the literature that seismicity is affected by the rate of water level increase, duration of loading, maximum levels reached, and the period for which the high levels are retained (e.g. Gupta et al. (1972b)) and that it occurs when the reservoir levels reach new maximum levels (e.g. by Simpson and Negmatullaev (1981)). It is therefore possible that the unloading / reloading does not have a direct effect on the granitic rocks which would potentially need a stress change higher than historic stress changes to respond in turn by demonstrating seismic activity / discharge of energy. Since in this case study I only had two cycles of loading and reloading and during which the reservoir loading reached similar height levels with similar reload rate, the aforementioned statements could not be verified.

As per Gupta (2002), the seismicity induced by the lake level changes is likely to continue, year after year, or after a few years gap, depending upon the state of stress at the seismogenic fault and the level of stimulus provided. Hence the effect of these 2 cycles of Raeterichsboden lake’s drainage and refilling may affect the surrounding rocks the later years during which there would not be any monitoring to capture them. As it may well be the case that a seismic response captured by my sensors is the result of previous unloading and reloading of the reservoir, expressed with a time delay.

Peaks in event numbers may be attributed to reservoir drainage. Also, the event numbers recorded by the Southern array may be attributed to the lake levels changes of both Raeterichsboden and Grimsel lakes given the array's location. To be able to confirm what is the triggering mechanism for the microseismicity the events will need to be located. If the events are caused by the unloading / reloading of the lake(s) then their hypocentres are expected to be either underneath the lake(s) or at their immediate vicinity. It is not however clear which kind of mechanism controls the seismicity; i.e. it can be either a rapid response to the reservoir loading or a delayed response to the unloading, according to Simpson et al. (1988). These authors proposed that rapid response is generated as a result of the instantaneous effect of loading (or unloading) and the delayed effect of pore-pressure diffusion. Talwani (1997) mentioned that the seismicity observed after impoundment is related to both mechanisms. Later, Talwani et al. (2007) mentioned that pore-pressure diffusion is primarily responsible for the build-up of fluid pressures and the onset of seismicity.

4.7 Summary

To conclude, in this Chapter I have detailed the researched literature review and methodology followed for detection and onset time determination of potential microseismicity. A novel research contribution presented in the chapter includes the NpD algorithm which was also published as Kinali et al. (2018) and is an open source algorithm for the detection of signals. The algorithm works well under not well-constrained conditions and without requirement of a priori knowledge about the expected signal frequencies and amplitudes. The efficiency of the NpD algorithm has been proven through the comparison with other well-known approaches.

The full data set of passive seismic data acquired between 1st November 2014 and 12th August 2017 was scanned using the NpD algorithm resulting in more than 241,000 detections. Some observations with reference to the acquired detection dataset as compared to the lakes water levels variations are summarised as follows: (i) the numbers of events detected by each of the individual arrays are broadly consistent, (ii) an anomalously high number of events are detected by the North and South arrays in early 2015, which may be as a result of reservoir drainage and refilling, (iii) the number of events appears to peak annually in the summer months (possibly related to the increase in glacial meltwater and consequent increase in reservoir levels, e.g. Johnson et al. (2017) and Roth et al. (1992), and (iv) a peak in event numbers may be associated with reservoir drainage.

It was also discussed that it is possible that (i) the unloading / reloading of Raeterichsboden lake does not have a direct effect on the granitic rocks which would potentially need a stress change higher than historic stress changes (see also Gupta et al. (1972b)) to respond, (ii) the effect of these 2 cycles of Raeterichsboden lake's drainage and refilling may induce delayed microseismicity response (Gupta, 2002), in the years that monitoring would no longer be in place, (iii) a seismic response captured by the sensors during this application may be the result of previous unloading and reloading of the reservoir, expressed with a time delay, (iv) the Southern array may be capturing weak events due to the lake water levels changes of both Raeterichsboden and Grimsel lakes because of its location, and (v) it is not clear which mechanism controls the seismicity; i.e. is it a rapid response to the reservoir loading or a delayed response to the unloading (see also Simpson et al. (1988)).

For the onset time picking of this application, I decided to proceed with the manual picking of the phases only for periods of interest. This involved an extensive amount of pattern recognition and great effort to pick the first breaks in a consistent and, as possible, objective way. For the visual inspection of the data records within the time periods of interest RTQT_View from REFTEK was used and the data in counts were filtered with a recursive Butterworth bandstop filter for the frequencies between 48 and 52 Hz (to remove the Swiss EM caused by power lines). The manual picking of P and S waves proved to be challenging because of the 250Hz sampling rate which proved to not be adequately high to differentiate the arrival time differences of certain signals between seismometers. The chosen time periods were the most probable dates for a microseismicity reaction to a stress change.

Finally, novelty in the present chapter lies in using, comparing suitability and adjusting a number of well-known onset time determination approaches from the literature, resulting to proposing a methodology (i.e. AIC-wav) that works well for high SNRs and allows more confidence in the P-waves' onset time determination but requires further work.

Having detected accurately the picks of the events of interest, the next chapter is focusing on their localisation and their source parameters determination.

Chapter 5 Analysis of recorded microseismic data: Seismic Source Location

5.1 Introduction

The precision of earthquake locations, as per Pavlis (1992), is governed by the source-receiver configuration (e.g., Eisner et al. (2010)), the complexity of the velocity model (e.g., Blias and Grechka (2013)), and accuracy of the arrival-time picking. Wuestefeld et al.(2018) add to these the choice of the location approach. While we have already discussed about the source-receiver configuration and the onset time determination, the rest of the parameters are going to be analysed in the present chapter as follows: Section 5.2 discusses about the velocity profile used in the analysis, Sections 5.3 and 5.4 include the literature review regarding localisation methods and source parameters considered to result in the methodology used in the present study (Sections 5.5 and 5.6 respectively).

5.2 Velocity profile

The local geology in seismic monitoring is reflected through the chosen velocity model. Wave velocities are attributed to different geological materials, i.e. a typical granite may have velocities $v_p = 4.5 - 6km/s$ and $v_s = 2.5 - 3.3km/s$ (Bourbie et al., 1992). In general, velocities are higher in rock formations than in soft soils, in saturated than dry ones and increase with effective pressure. The pressure dependence results from the closing of cracks, flaws, and grain boundaries, which elastically stiffens the rock mineral frame.

Every velocity model has intrinsic uncertainty and errors associated with event location will be passed on to any subsequent interpretation, such as the source mechanism (Usher et al., 2013). Hence the choice of the velocity model is of some importance, which in turn poses some questions to the researcher: can the subsurface reliably be approximated by a 1D velocity model, or is a 3D model required? How many layers are appropriate? Is anisotropy required (e.g., Wuestefeld et al., 2011), and to what degree?

Taking into account previous seismic explorations in the GTS area and the velocity profiles researchers used in these occasions I chose an initial homogeneous 1D model. The v_p and v_s values used in the GTS literature showed small variations: Majer et al. (1990) used 5.4 and 2.9km/s respectively and Blumling et al.(1992) 5.3 and 3.06km/s respectively. Then a sensitivity analysis was performed, using selected earthquakes, detected and located by the Swiss Seismological Service, and locating them using various combinations of v_p and v_s values. The average v_p and v_s values (i.e. $v_p = 5.4km/s$ and $v_s = 2.9km/s$), that minimised the residual between known and calculated locations, were the velocity values used for the location of the events in the present study.

The chosen values are within the reported range of wave velocities for a typical granite, i.e. $v_p = 4.5 - 6km/s$ and $v_s = 2.5 - 3.3km/s$ (Bourbie et al., 1992). A layered model was not deemed necessary in the present experiment because, according to Angus et al.(2014), in surface microseismic monitoring the influence of vertical velocity variation is less problematic than for downhole monitoring. Also, Usher et al.(2013) reported that differences between plausible velocity models result in 0.0035 seconds uncertainty in the P-waves arrival times picking and 0.0085 seconds in S-waves, which is equivalent to circa 20m in location difference. Since the present research is not a hydrofracking experiment, but instead the monitoring covers an area of kilometres, such an accuracy uncertainty was deemed acceptable.

5.3 Localisation literature review

The location problem consists in correctly estimating four unknown parameters that best fit a set of arrival times of seismic waves at different receivers: the hypocentre coordinates of the starting point of the rupture process (x,y,z) that caused the seismic event and the corresponding origin time (t). The arrival time function for each sensor can be described in mathematical form with the formula:

$$\int_{s_i} \frac{1}{v} ds = t_i - t_0 \xRightarrow{homogeneous\ medium} \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} = v(t_i - t), \qquad (eq.12)$$

where x_i, y_i, z_i are the coordinates of each triggered sensor$i$,t_i the triggered arrival time and v the propagation velocity of the medium.

The formula for the homogeneous medium is the simplest form of the arrival time function and depending on the velocity model (i.e. 3-D, anisotropic medium) the equation can become much more complicated. More information on the velocity determination approach used in the present study is provided in Chapter 5.2.

There are a number of methods that attempt to find the global minimum of the arrival time function. The most commonly used method to solve the system is to minimize the sum of the square differences of the equation. Thus, the location problem is transformed into optimizing an objective function, i.e. determining the smallest misfit between the theoretical and observed arrival times.

The literature review that follows is by no means extensive but rather a general overview of commonly used approaches. It serves the purpose of showing the logic behind the selection of the approach followed for the present research. The presented methods are categorised into non-iterative and iterative methods.

5.3.1 Non-iterative location methods

Non-iterative methods address the aforementioned non-linear problem of hypocentral determination employing a direct analytical linear approach. Their main limitation is their inflexibility to adapt to different velocity models; i.e. they can only operate with a constant velocity model for all sensors, which could lead to large localisation errors (see also Chapter 5.2).

5.3.1.1 Single station location

A simple approximation to the solution of the seismic location problem can be provided using a single three-component station (Bormann, 2012). Assuming known v_p and v_s getting the measurements of the t_p and t_s from the waveform, we can determine the angle defined from the North, the station and the epicentre counting clockwise (backazimuth). The epicentral distance can now be estimated as $D = \left(t_s - t_p\right)v_p v_s/(v_p - v_s)$ and the origin time can be calculated as $t_o = t_p - D/v_p$. The single station location is a purely epicentre determination method, i.e. we can only estimate the epicentre and not the source depth.

5.3.1.2 The circle / sphere and the hyperbola method

Another simple, graphical location approach that uses recordings from at least three stations is the circle method (Stamps and Smalley, 2006). After calculating the epicentral distance (D) as indicated in Chapter 5.3.1.1 for all sensors, circles are drawn with centres each station and radii the epicentral distances. The seismic location is bounded within the common area of all three circles with a best estimation (X) the crossing of chords which connect the crossing points of circle pairs. For a homogeneous velocity model the hypocentre (H; focal depth) can be given using the epicentral distances as radii of spheres in three dimensions; i.e. focal depth can be calculated from the equation: $H = \sqrt{D^2 - X^2}$ (Yfantis et al., 2014). In the case of a non-homogeneous model, the seismic waves would be refracted at the boundaries of the different geological materials and the spheres would be deformed, rendering this method valid only for homogeneous velocity models.

In the same logic lies another location method (Pujol and Smalley, 1990; Pujol, 2004), which fits quadratic surfaces (hyperbolas) instead of spheres to the observed arrival times. The method needs two-station combinations to operate and a minimum of four stations to yield a solution, creating 6 hyperboloids in total (i.e. 1+2+3 two-station combinations). The definition of the hyperbola is given by: $d_1 - d_2 = v(t_1 - t_2)$, where the known parameters on the right side of the equation are the medium velocity and the P-wave arrival times recorded at two seismometers, while the unknown parameters at the left side are the epicentral distances of the sensors. The centre of the area created from the intersection of all hyperboloids coincides with the event epicentre.

5.3.1.3 Inglada's method

Inglada's method is based on the graphical sphere method but attempts to solve the location problem through a direct analytical approach. The concept is that the seismic source will lie in a single point of the intersection of 4 spheres; i.e. the wavefront of four sensors (Inglada, 1928). Arrival time data and a constant velocity model are used as input data in *eq. 12*. Only a minimum number of sensors is used in this method, equal to the number of the unknowns in the arrival time equation (i.e. four) and because of this requirement no optimization method can be applied (Ge, 2003a).

5.3.1.4 USBM

Another mathematical solution to the location problem is given with USBM method, developed in the 1970s by researchers in the United States Bureau of Mines (e.g. Leighton and Duvall, 1972). The algorithm calculates the distance for n+1 (i.e. five in the case of a homogeneous model) sensors, where n is the number of the unknowns in the arrival time equation, and solves the non-linear system by linearizing it with a set of mathematical tricks; i.e. the origin time unknown parameter is eliminated by subtracting the distance of one of the sensors from the rest, then squaring and subtracting the distance of another sensor out of the rest of them etc.

5.3.1.5 Arrival order approaches

This type of approaches is extensively used and is based on the assumption that the traveltime of an event is monotonic with distance, hence the order that the signal arrives in the different sensors is a valid equivalent of epicentral distance (Ge, 2003a). These approaches provide a better idea of the relative seismic source location based on the first tiggered station; i.e. out of two sensors receiving the signal, the one that is triggered first is the one closer to the source of that signal. The source location can be further constrained in smaller zones since the procedure can be repeated depending on the number of triggered sensors. To better estimate the radial direction of the signal, a ratio of the observed arrival time difference between two sensors over a theoretical limit (computed as the ratio of the inter-sensors distance over the velocity) is used. If the ratio is closer to zero, the source tends to be equidistant from the sensors while if it is closer to one, it is closer the first sensor.

5.3.2 Iterative location methods

Iterative location methods are trial and error processes which assume an initial location and calculate an initial source to sensors set of travel times based on a known velocity model. The travel times are then compared to the recorded ones. With each iteration the misfit gets smaller until it reaches to a user defined threshold. Iterative methods can be further divided into:

5.3.2.1 Derivative approaches

This category hosts one of the most widely used location algorithms: Geiger's approach (1910), that solves the non-linear source determination problem by minimizing residuals between observed and predicted travel times of various phases. Geiger’s method needs as input data an initial solution, observed and arrival times and partial derivatives of the arrival time function. The partial derivatives are Geiger’s way of assessing the non-linear behaviour of arrival time functions and are used as a correction vector. In derivative methods, the final solution is reached through a continuous updating process of the initial trial solution, i.e. by adding the correction vector derived from each iteration to the next one. The least squares solution to a system of equations is the location the problem converges to. It sometimes occurs that the system does not converge which is a sign of instability of the associated mathematical system which traces back to a poor network geometry for locating the specific event.

Another derivative method, similar to Geiger’s is Thurber’s method (Thurber, 1985) with the only difference relying on Thurber using the second degree of Taylor polynomials (i.e. the first and second partial derivatives instead of just the first ones) as the correction vector. Hence, Thurber’s solution is geometrically represented by the extreme of a quadratic function.

5.3.2.2 Sequential or grid-search approaches

In these approaches (e.g. Saltogianni and Stiros (2012a,b, 2013, 2015)) the user defines a search volume and grid spacing. For each of the grid points the residual between theoretical observed arrival times is calculated (Wuestefeld et al., 2018). A finer grid would give a more accurate solution. An alternative to this is to start with a coarser grid and continue with the further refinement of the monitoring area in smaller ones (nested grid) as the search continues (Lomax et al., 2000). Another alternative to the typical grid search approach is the directed grid search (Sundhararajan et al., 1998). The source location search starts with a grid space of 8 nodes and the grid moves towards the node with the smallest gradient of the objective function. Grid searches are considered to have reached a global minimum when either a user defined maximum number of iterations or a user defined threshold residual is reached.

5.3.2.3 Simplex algorithm

Simplex algorithm is a curve fitting technique, developed by Wolfe (1965) and introduced into source location application by Prugger and Gendzwill (1988). The method first associates each point in space to its error (between observed and theoretical travel times) creating an error space and then searches for the point with the minimum error. The first step is to set an initial Simplex figure: a geometric figure with one more vertex than the dimension of space it is used in; i.e. for a two-dimensional space that would be a triangle and so on. Next, the errors for the vertices are calculated (usually L2 or L1 norm are used for this), and then the Simplex figure moves into space (through reflection, expansion, contraction and shrinkage). A series of error comparisons between all vertices begins, leading through a series of substitutions to a status of accepted convergence (Ge, 2003b).

5.3.3 Localisation discussion

Location approaches combine advantages and shortcomings. The single station location is an epicentre determination method that offers the advantage to locate an event using a single 3-D station. At the same time however, the method uses, as input data, a sensitive and errorprovoking parameter, amplitude values, which depend greatly on material properties (attenuation) and local geological conditions (site amplification). Also, the S-waves arrival time determination is difficult because S-waves usually overlap with the P-wave coda (Helmstetter and Garambois, 2010). This phenomenon is attributed to the near-field effect; i.e. when the wavelengths are large relative to the source–receiver distance, the P-wave, S-wave and surface wave phases can be intertwined (O'Brien et al., 2011). The circle / sphere location method and the hyperbola method are strictly for homogeneous media location methods. According to Chen et al.(2016), Inglada's method and USBM are not suited for microseismic monitoring since in the first case only a minimum number of sensors is used and no optimization method can be applied to the algorithm (because of the requirement of the algorithm for equal number of sensors and equations), while the latter is unstable in cases of large picking errors (LPEs). Using optimisation and more than the minimum number of sensors for seismic inversion allows for errors and uncertainties minimisation. Arrival order methods are simple and easy to use but are based on the assumption that all sensors are triggered by the same type of wave which may not be the case. They also do not estimate pin-point locations, which is the focus of microseismic applications (with acceptable errors that vary depending on the scope of the investigation), but rather a more general source location of the seismic wave. They are a good approximation however in cases when sufficient sensors for pin-point location are unavailable (Ge, 2003a) or for validation of estimated locations by other algorithms.

Derivative approaches effectiveness is greatly dependent on the sensor array geometry. If the initial set up of the sensors is not optimal, that may have great impact on the mathematical model stability which in turn causes divergence. In the grid-search approaches there is a trade-off between computation time and location accuracy, which depends on the grid resolution. In the nested grid approach in particular there is the danger that the initial coarse grid model calculations might miss the global minimum and in that case the algorithm will yield as a solution a local minimum. Simplex algorithm is flexible to deal with complicated velocity models (i.e. multiple velocity layers). In addition to this arrival time functions used in the method can be established during the data processing while in derivative methods they need to be established prior to the analysis.

Non-iterative methods are simple and easy to use in that they require little interaction from users. A common problem with these methods is the assumption of a single velocity, which severely restricts their applications. Iterative methods, because of their flexibility in dealing with different velocity profiles and arrival time functions, approach more realistically a great majority of source location problems. As is derived from the above, selecting an appropriate location method involves weighing the advantages and disadvantages of approaches in terms of accuracy requirements and data limitations. With these remarks in consideration, the location problem shifts in the accuracy of arrival times determination and a velocity model representative of the under study area.

5.4 Seismic source parameters literature review

5.4.1 Magnitude

Magnitude is one of the most fundamental parameters used to quantify an earthquake and is basically a logarithmic measure of its size based on instrumental measurements. Magnitudes are commonly derived from ground motion amplitudes and periods or from signal duration measured on instrumental records.

Currently there is not a uniform scale of measurement because of differences in instrumentation, data reduction methods, magnitude formula, station distribution, etc. Relations between different magnitude scales have been proposed by various researchers in order to maintain continuity and uniformity of the results. Most of the, proposed over the years, magnitude scales are empirical and lacking in that they do not relate magnitude to a physical model. The exception is moment magnitude, introduced by Hanks and Kanamori (1979), which can be used to bridge waveform amplitudes to the seismic moment.

Richter (1935) introduced the first magnitude scale to seismology. He developed the local magnitude scale M_L for describing the relative strengths of earthquakes in California, and related the amplitude of a waveform recorded with a particular instrument (a Wood-Anderson seismograph) at a given distance from an event to the strength of the event (Baig and Urbancic, 2010). The term local refers to the fact that the scale is tailored to a particular region with slightly different calibration curves (to account for differences in seismic attenuation, geology, etc.) when applied to other regions.

Gutenburg and Richter (1936) used the amplitude of teleseismic 20s surface waves to derive a magnitude scale for crustal earthquakes (surface-wave magnitude scale M_S). To arrive at a relatively standardized magnitude measurement, different calibrations needed to be employed. M_L and M_S scales agree at M_L=6 because these magnitudes events were used for calibration. Since M_L is determined from the maximum amplitude on a Wood-Anderson seismogram, the predominant period of the waves used is usually 0.1-3 sec. In contrast, M_S is usually determined from the maximum amplitude of surface waves with a period of about 20 sec.

Later, Gutenberg (1945b) introduced a body-wave magnitude, m_B, which is computed from the amplitude and the period of seismic body waves. In this scale, the maximum amplitude of a wave group corresponding to various seismic phases such as P, PP, and S were used for the determination of the magnitude. For this measurement, various types of seismographs, including short- and long-period mechanical and some electro-magnetic instruments, and periods of waves ranging from 0.5 to 12 sec were used. The two scales, m_B and M_S, agree over the magnitude range of the events used for the calibration (i.e. 6.5-7 M_S).

Since these initial magnitudes determinations, a number of similar magnitude scales have been introduced in the literature and can be summarized as:

$$Magnitude = log10\ (Amplitude) + Correction\ Factor \qquad (eq.\ 13)$$

where the correction factor depends on distance and sometimes the period of the waveform. A study conducted by Bormann et al. (2009) discussed the interrelations between different magnitude scales to earthquake data recorded in China. There are a couple problems with such magnitude scales: (i) they are empirical, i.e. there is no tie to a physical model which means that a single magnitude cannot be explicitly related to specific fault parameters; (ii) they are saturated for the large earthquakes.

Kanamori (1977) and Hanks and Kanamori (1979) developed the moment magnitude scale to address these shortcomings. The seismic moment M_0, is based on a model assuming shear displacement of a planar fault, and is the product of the shear modulus, μ, the average slip on the fault, $\bar{d}$, and the area of the fault, A. By measuring the energy, E, in the waveforms, they related this quantity to the seismic moment through the approximate relation:

$$M_0 = \mu dA \approx \frac{E}{20000} \qquad \textit{(eq. 14)}$$

and then developed a scale from this measurement that roughly matched the unsaturated part of the magnitude scale. Although this scale was developed specifically for large earthquakes, its range of applicability extends all the way down into the microseismic realm.

5.4.2 Corner frequency

Because of the band-limited nature of seismic signals, the recording system is an important consideration in any discussion of magnitude. The corner frequency f_c of an event is empirically related to the magnitude of the event. The relationship arises from the relationship of seismic moment to the area of a fault surface: a larger fault surface area gives rise to a larger wavelength and therefore lower frequency signals. The corner frequency can thus be viewed as the characteristic or natural frequency of the event.

Microseisms in the magnitude range of -3 to 0, give rise to corner frequencies in the range of approximately 50 to 500 Hz as they occur on fractures with dimensions of 10s of centimetres to a few metres. Microseisms with magnitudes of 0 to 5, occur on larger faults (10s to 100s of metres) leading to low corner frequencies in the range of 1 to 50 Hz. The largest earthquakes may cause slips on fault planes on the order of 1000 km, resulting in very low corner frequencies (in the scale of mHz).

5.4.3 Source spectrum and model-fitting approaches

The signal from three component sensors can be inverted for seismic moment by assuming a model for the source shape and relating that to a rupture area and an amount of slip. Simple models of fault slip such as the Brune model (Brune, 1970) or the Haskell model (Haskell, 1964) that assume a circular and a rectangular fault respectively, provide relationships between spectral (corner frequency f_c, low-frequency plateau Ω_o) and source parameters (slip area and seismic moment).

Details of theoretical source spectra depend on model-dependent constants and are thus different for circular and rectangular faults but also for different models of circular faults. Brune (1970) stipulates that frequency response of the signal is flat until the corner frequency f_c is reached, at which point, the amplitude falls off as f^{-2}, although other models yield high frequency asymptotes of $f^{-2.5}$ or f^{-3} (unilateral rupture, e.g. Haskell, 1964). Real fault ruptures may have any other shape and can only be approximated by these models. Generally small earthquakes are approximated by circular fault models (Bormann, 2012).

The chosen fitting algorithm solves for Ω_o, fc and Q simultaneously and therefore allows for further investigations into source parameters. Figure 46 explains the extrapolation of the first two elements in a simple manner: (a) and (b) refer to the time domain while (c) to the frequency domain (fast Fourier transform FFT). The moment function $M_0(t)$ is the increase in moment caused by an earthquake slip along a fault. To estimate M_0 in the time domain, the seismogram must be converted to displacement and integrated over the length of the pulse (b). In the frequency domain (c) and at low frequencies the spectral amplitude becomes constant and is proportional to seismic moment ($\sim M_0$). At high frequencies the spectra show a decay that falls off as f^{-2} to f^{-3}. Plotted on a log-log scale the spectrum can be approximated by two straight lines. The corner frequency f_c is the break-point of the asymptotes (straight-line Bode plot). The low-frequency plateau Ω_o is estimated and used directly in the calculation of seismic moment as this relates to the area of slip on the fracture plane (the area under the displacement curve).

The spectral characteristics of recorded seismic waves, however, are strongly influenced by the attenuation properties of the medium as the signal travels from the hypocentre to the geophone. In order to obtain reliable estimates of the source spectrum, the estimation of the attenuation quality factor, Q is therefore needed (Eaton, 2011). In general, as depth increases, the rock becomes harder and more rigid. Both Vp and Vs increase, Vp/Vs decreases, and there is less attenuation (higher Q factor; Xu and Stewart (2006)). To estimate Q we can fit models built with different attenuation factors and the slope of the best fitting straight line on a log-log plot versus frequency can be used (Xu and Stewart, 2006).

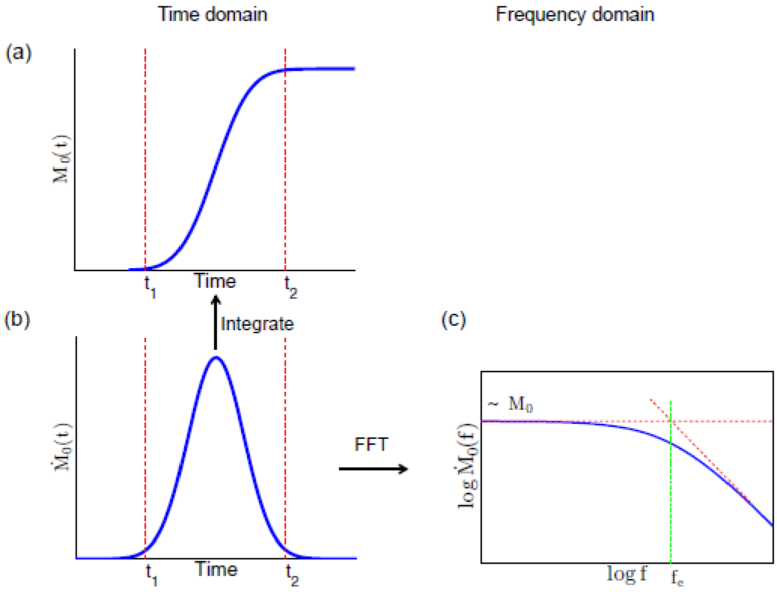


Figure 46: Figure reproduced from Stork et al. (2014) demonstrating the extrapolation of the seismic moment and corner frequency parameters from the fitting of the displacement spectrum of a seismic event.

Also entering the equation are geometrical spreading 1/r, wave-speed v (either v_p or v_s), the density of the rock ρ, and a factor R for the radiation pattern.

$$M_0 = \frac{4\pi\rho v^3 r\Omega_o}{R} = \mu\bar{d}A \qquad \textit{(eq. 15)}$$

Boore and Boatwright (1984) developed a rough estimate for R if one cannot adequately determine the radiation pattern of the moment tensor, 0.52 for P waves and 0.63 for S waves. The right side of the equation relates to parameters of the faulting process: μ is the shear modulus of the rock, $\bar{d}$ is the average displacement of the fault, and A is the area of the fault. The product $\bar{d}$A is termed geometric moment and can be determined directly from the zero spectral asymptote of the source in a seismogram, respectively by integrating over the area underneath the resituated broadband displacement P and/or S waveforms recorded in the far-field of the source. The value for the rigidity μ, however, has to be assumed.

With seismic moment, measured in Nm, we can then determine moment magnitude from the following formula (Havskov and Ottemoller, 2010):

$M_w = 2/3\, log_{10} M_0 - 6.1$ *(eq. 16)*

In contrast to the empirical relations discussed earlier, this definition links magnitude to the properties of the fault. The constants in the above relation ensure that moment magnitude falls in a range that is comparable to other magnitude scales and makes it a logical choice for use with microseisms.

5.4.4 Instrument response

The sensor's response consists of a set of zeros which can be thought of as the number of integrations, poles which shape the corners of the response, and a constant gain factor to convert from a voltage output to a ground motion (V / (m/s)). The second stage is a cascade of filters to convert from an analog signal to a digitized sample. Typically digitizers will initially record the signal with a digitization on the order of 36,000 samples per second. A series of low pass filters and decimations then convert that analog signal to the output sampling rate. This cascade of filters has its own set of poles, zeros, and scalar constants. The full response is then a product of all filter pole-zero-constant and the sensor pole-zero-constant. Haney et al. (2012) described a general method for causal instrument correction, applicable to data from a wide range of seismometers. Their Matlab codes are available online as rm_instrum_resp.m (https://github.com/ChrisBail/PSPicker).

5.5 Location determination in the present study

The TOPological INVersion (TOPINV) algorithm (Saltogianni and Stiros, 2012a,b, 2013, 2015) uses a deterministic search in an n-dimensional grid and does not assume any constraints a priori.

The inverse problem is defined as:

$f(x) - l = v,$ *(eq. 17)*

where v is an unknown error, $l = \{l_1, .., l_m\}$ are the m observations (i.e. difference between time of arrival of the S- minus the P- wave) and vector $x = \{x_1, .., x_3\}$ the unknown parameters, namely in this case Easting, Northing and depth of the hypocenter.

Saltogianni and Stiros (2013) transformed *eq.1* into:

$$|f - l| < k\sigma, \qquad (eq.\,18)$$

where σ is the standard deviation of the observations l and k, the only unknown parameter in this search process, is an optimization factor. For any k, a cluster of grid points, all of which satisfy all of the m inequalities (*eq.18*), is identified using a Boolean approach. The process is repeated until the optimal value of k, that satisfies certain user defined criteria, is obtained. An optimal subset of grid points is then identified as calculated solutions. The optimal estimator of the algorithm is computed from the first and second statistical moments of the coordinates of the potential solutions and is accompanied by its variance-covariance matrix.

The accuracy of the solution depends on the geophysical model; i.e. the accuracy of the user defined velocities, the quality of the observations, and the density of the grid. Saltogianni and Stiros (2015) found the algorithm works well even in low SNR, however it is a slow algorithm, limited by the number of grid points. A potential alternative to overcome this issue is the use of nested grids (D'Auria et al., 2006).

The TOPINV grid-search algorithm was used to locate events for both P- and S- wave times and only for the detected events for which at least 5 out of the 8 seismometers of the two surface arrays were triggered. In total, out of 168 events that their phases were re-picked, 70 microseisms were located using the TOPINV algorithm (successfully converged). This small percentage of convergence is attributed to the dimensions of the grid used for the analysis, i.e. the sources of the re-picked events would have been outside of the limits of the grid and therefore the scope of my research (not RIS). The location estimates had an average fit error, i.e. the sum of the absolute residuals, of less than 143 m. In Figure 47 we can see the locations of the events and in Figure 48 a histogram with the error values of the locations. The events depicted with red balloons are the ones that occurred in 2014 while with yellow balloons in 2016. We notice that the majority of events have errors of 0.05 to 0.15km. Figure 49 demonstrates the in-depth spatial distribution of the events. All events are distributed between -2.9 km and 7.8 km asl which roughly translates (based on Google earth vertical ground profiles) to 282m above ground surface up till 10km below ground surface.

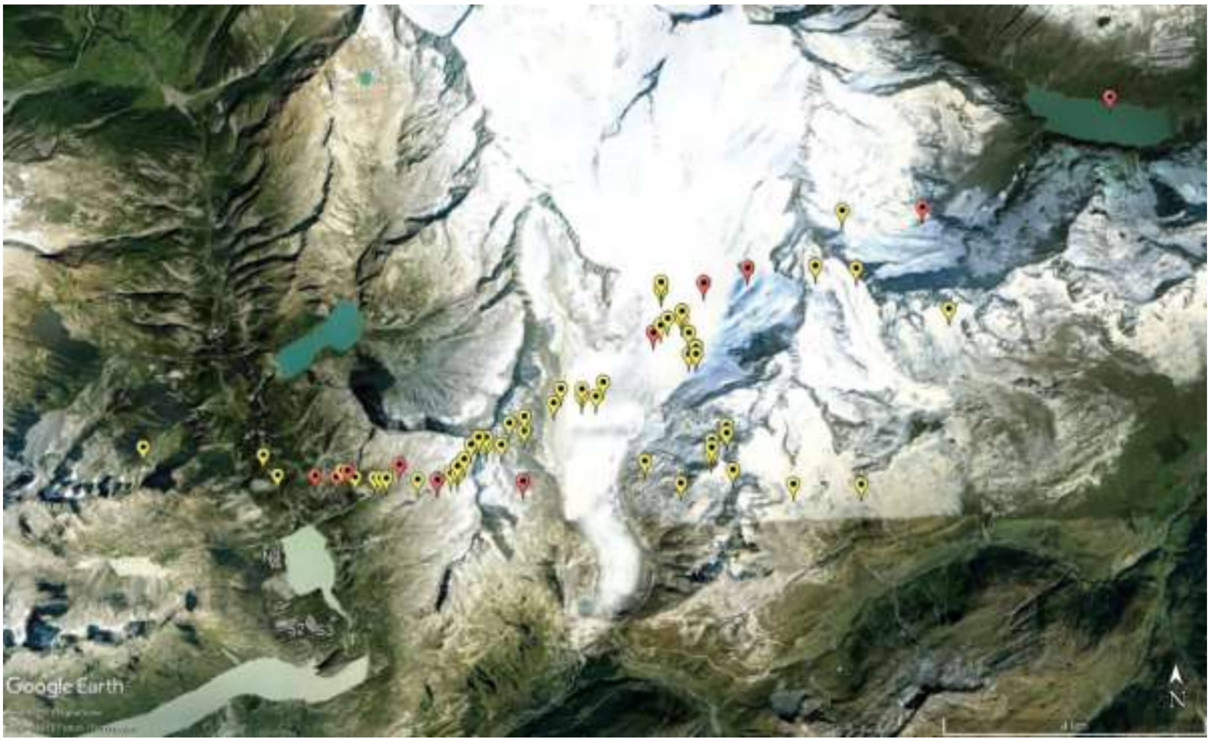


Figure 47: Locations of all manually repicked P- and S- waves events. Events that correspond to 2014 and 2016 are shown with red and yellow balloons respectively.

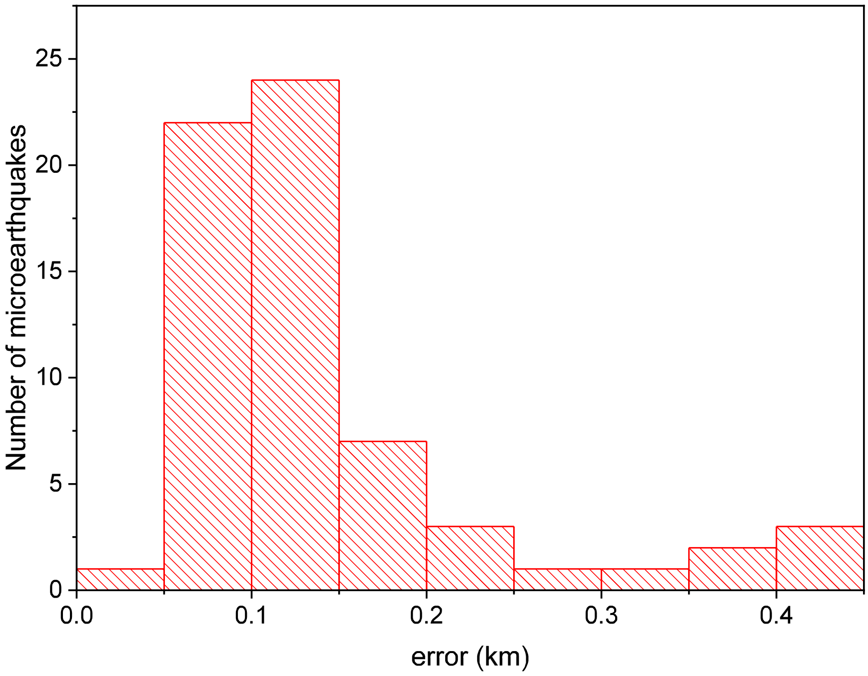


Figure 48: Histogram of error values for located events (sum of absolut residuals)

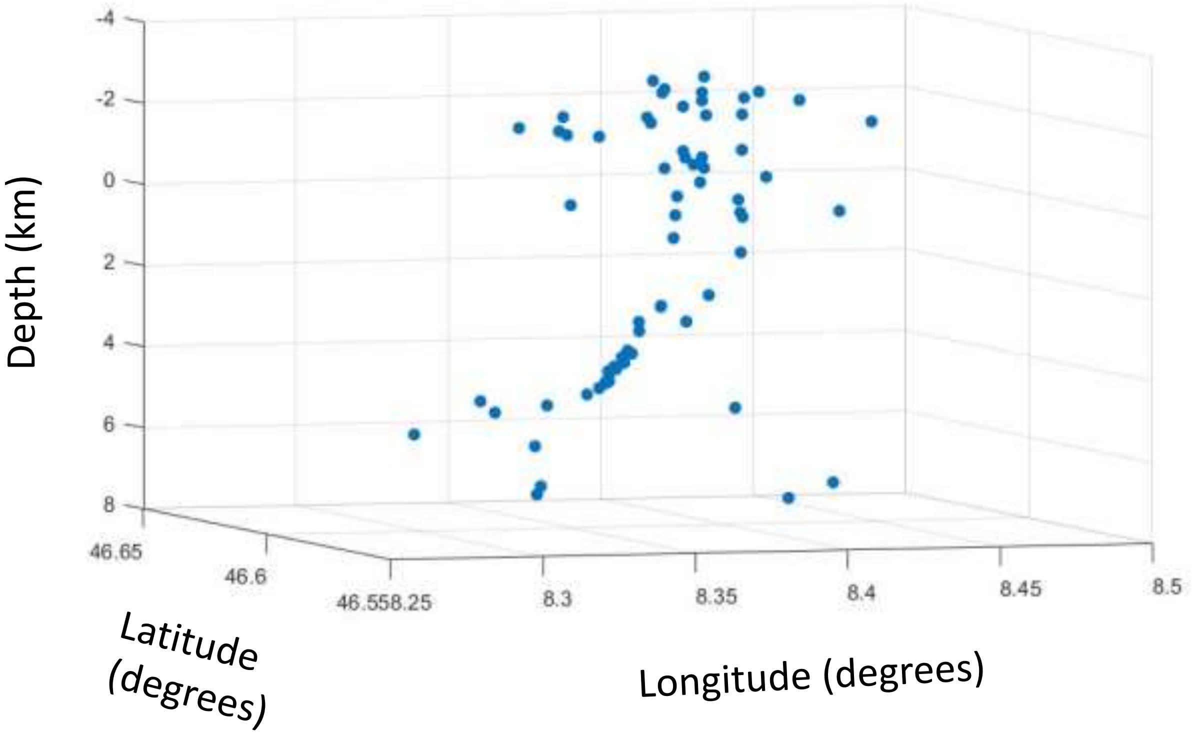


Figure 49: Location of manually picked events

I then further explored the periods before the times that drops in PH levels were identified in laboratory experiments (Stillings, 2020). In particular, I examined the full hours of data, for the North Array sensors only, for the following 3-day periods:

- 24.11.2014 from 13:00 to 26.11.2014 13:00
- 05.02.2016 from 12:00 to 07.02.2016 12:00
- 08.02.2016 from 10:00 to 10.02.2016 10:00
- 26.02.2016 from 07:00 to 28.02.2016 07:00
- 06.03.2016 from 10:00 to 08.03.2016 10:00
- 11.03.2016 from 13:00 to 13.03.2016 13:00
- 18.03.2016 from 11:00 to 20.03.2016 11:00

As now in this check, I visually scrutinised the waveforms, I found cases of events that were picked by the NpD algorithm but I was not able to visually identify the phases, as the one shown in Figure 50. The event's phases can be visually picked in the North array but only for one of the seismometers of the South array.

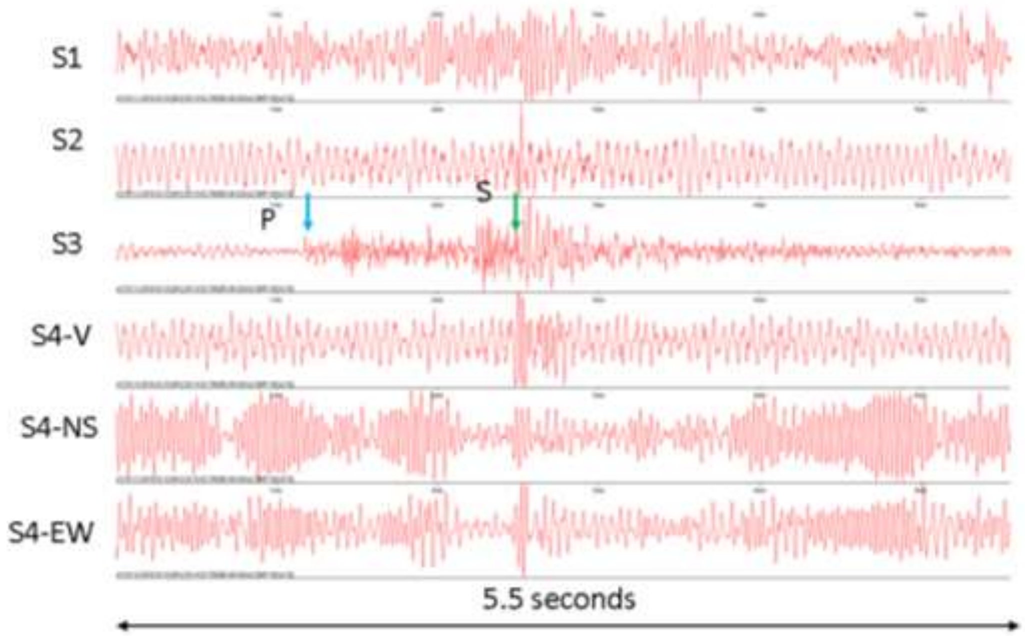


Figure 50: Event that its phases could not be visually picked, although detected by NpD and *AIC-wav* algorithms.

Further details on this analysis is provided in Chapter 5.7.2.

5.6 Source parameters in the present study

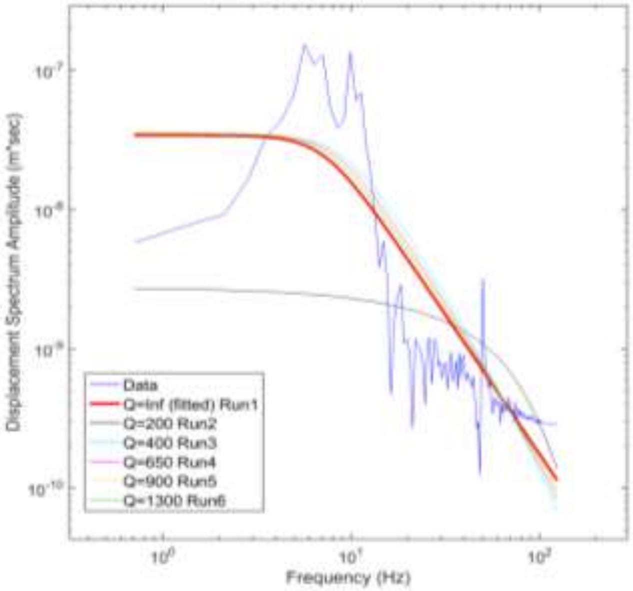
The workflow followed in the present study to determine the source parameters of the events is presented as follows. An appropriate time window is chosen to perform M_0 calculations. The start of the time window is the arrival time for the P- or S-waves and the end of the time window is at 2+ cycles of the P- or S-waves, while the signal is visible above the noise level. This is to maximize the event energy included in the calculation while minimizing the contribution of other sources of energy (Stork et al., 2014).

The calculated spectrum is then corrected through deconvolution with the complex instrument response function (response of the velocity seismometer and anti-aliasing filter of the recorder; (Haney et al., 2012)). A hanning taper is applied to both the start and end of the raw seismogram prior to instrument correction. The length of the taper is the smaller of the following two values: 10 percent of the entire length of the data or the longest period desired.

For each component, velocity spectra for signal and noise are computed by taking the Fourier transform of the windowed trace, normalized such that absolute units are preserved. Displacement spectra for individual components are then computed by dividing velocity spectra by iω. Because the receivers are at the surface, a free-surface correction (factor F) is also applied to the data (Aki and Richards 2002). Furthermore, the spectrum is corrected for attenuation, $\exp(i\omega t/2Q_p)$. Brune source parameters for all analysed events were computed using a procedure similar to Pytharouli et al. (2011).

In particular, the value of Q_p value was estimated through sensitivity analysis (see Figure 51), i.e. for different displacement spectrums the Brune fitting curves using various values for Q_p were estimated. Q_P for values around 650 (purple model fit) seemed to provide consistently better model fits hence that was the value kept for the rest of the analysis. The parameters that were used for the estimation of the attenuation quality factor can be seen on the right hand side of the figure. Greater attenuation values result in too low inferred corner frequencies; however, uncertainties in Q have insignificant effects on the magnitude estimations, which depend only on the low-frequency displacement asymptote.

Using the attenuation factor that was derived from the aforementioned sensitivity analysis, an estimate of the low-frequency plateau is obtained by taking the difference between the average signal and noise amplitude. Corner frequency is determined by finding the optimum fit between observed and modelled displacement spectra. The methodology by Pytharouli et al., (2011) uses Nelder-Mead algorithm to achieve the best fit between observed and modelled spectra. An example of a model fitting of an event is given in Figure 52, while in the insert the Brune source model parameters are summarised.



PARAMETERS USED FOR Q_P ESTIMATION:

Radiation coefficient for P waves R = 0.52 (Boore and Boatwright, 1984);

Average crustal density Ro = 2800kg/m^3 (Edwards et al., 2010);

Hypocentral distance hypod= 5983.318m;

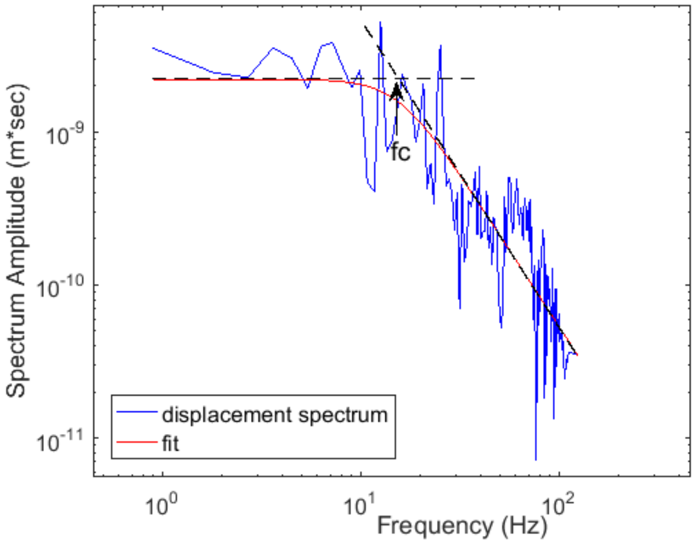
P-wave velocity = 5400 m/s;

Pi = 3.141592;

Free surface amplification F = 2 (Edwards et al., 2010) for normally incident SH waves.

$f_c, M_o and\ \Omega_0$ were estimated from the model optimization.

Figure 51: Sensitivity analysis for the estimation of Q_p



Radiation coefficient for P waves R = 0.52;

Average crustal density Ro = 2800kg/m^3 (Edwards et al., 2010);

P-wave velocity = 5400 m/s;

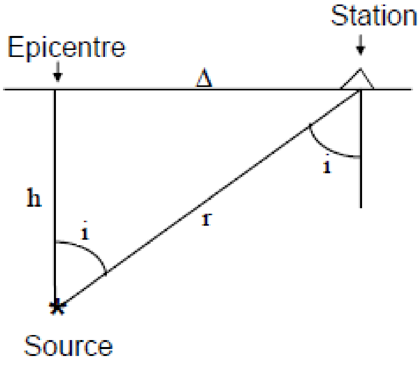
Free surface amplification F = 2;

$f_c, M_o and\ \Omega_0$ were estimated from the model optimization.

Q_P = 650.

Figure 52: Example of source-spectrum model fitting for an event. Blue curve shows the displacement spectrum for the P-wave arrival, after application of a Butterworth windowing function. The red curve shows the best-fitting Brune source model, with parameters summarized in the inset box.

The parameters M_o and f_c are the basic spectral data from which the source parameters are estimated. Event and material data required for further calculations are the epicentral distance Δ, the source depth h, the rock density ρ, the P-wave velocity v_P, and the averaged radiation pattern R for P-waves. Respective values are given under Figure 53 below. Using these and the equations listed in Table 10 below, the required source parameters can then be calculated.



p-wave velocity	v_p = 5.4km/s
source depth	h
epicentral distance	Δ
hypocentral distance	r = $\sqrt{(h^2 + \Delta^2)}$
incidence angle	i = arc cos(h/r)
free surface amplification	S_a for p-waves
averaged radiation pattern	R = 0.64 for p-waves

Figure 53: Schematic of geometric parameters of a seismic event

Table 10: Overview of formulas of the estimated source parameters

ESTIMATED PARAMETER	**EQUATION**	**REFERENCE**
Seismic moment M_0	$M_0 \; = \; 4\,\pi\, r_o\, v_P^3\, \Omega_0\, hypod/(F\, R)$ (in N*m)	(Keilis-Borok and Monin, 1959)
Moment magnitude M_w	$M_w = 2/3 \log_{10} M_0 - 6.06$	
Source / fault radius	$r_o \; = \; 2.34 v_P/(2\,\pi\, f_c)$ (in m)	(Brune, 1970)
Stress drop (circular rupture)	$\Delta\sigma = \frac{7M_0}{16 r_o^3} * 10^{-6}$ (in MPa)	(Brune, 1970)
Source duration	$T_d = 1/f_c$ (in sec)	
Area of rupture plane	Area = $\pi {r_o}^2$ (in m^2)	
Fault slip	Displacement = $M_0/(33e^9 * Area)$ (in m)	
Fault length	Length = $2r_o$ (in m)	

Out of a total of 70 manually re-picked and located events, I was able to analyse further and estimate the source parameters of only 36. An overview of the inferred source parameters for the 36 analysed events of the North array and the 24 of the South array is presented in Table 11 and Table 12 respectively. All of the events of the South array are a subpart of the analysed events from the North array.

For visualisation purposes the results of the North array are also shown in the form of bar charts, in particular in Figure 54(a) we can see the mean moment magnitudes, in (b) the displacements and (c) the size of the slip patch areas, all of the above accompanied by their standard deviations (lighter shaded bars). The mean of the estimates from all receivers (per array) is reported in the tables and figures as a way to provide increased confidence in estimates. As we can see in Figure 54(a), there are some events with standard deviations greater than actual values.

Hypocentral distances sampled by this set of events span a range from ~3.31 to ~34.85 km (considering both arrays). Inferred corner frequencies exhibit a high degree of scatter as they fall within the range $8.78 \leq f_c \leq 109.58$ Hz with corresponding low frequencies of 1.13E-8 and 1.35E-9Hz. Calculated seismic moments and moment magnitudes fall within the range of $7.81E+8 \leq M_0 \leq 4.11E+12$ and $-0.23 \leq M_w \leq 2.41$ respectively; this range reflects a sampling bias towards larger magnitudes, due to the focus of detection for this experiment. Fault lengths range from 25.45m to 469.35m, source radii range from 12.73m to 234.68m while the areas of the patches that slipped vary from 177,721 to 514.43m^2.

In Figure 55 the magnitudes of the events whose mechanisms were estimated from both arrays are demonstrated as spheres on Google Earth map. The spheres are coloured in 3 different colours according to the events magnitudes. For events whose moment magnitudes are less than 1, between 1 and 1.5, and above 1.5 the colours red, yellow and green are assigned accordingly. The moment magnitudes are the averaged values between the two arrays. It is interesting to see the concentration of the lowest magnitudes events around a specific area.

The located events are presented as part of a seismic catalogue along with seismic events within a 50km radius around the sensors arrays mid-point and for the duration of the project, i.e. Oct 2014 – Aug 2017. The seismic events were extracted from the Swiss Seismological Service catalogue that includes known earthquakes in Switzerland and its neighbouring countries (http://www.seismo.ethz.ch/en/earthquakes.xml). The seismic catalogue can be found in Appendix B. It can be observed that this research's detected events don't form part of the Swiss catalogue and constitute an addition. This strengthens the case with regards to the detected events being reservoir induced and not of tectonic origin.

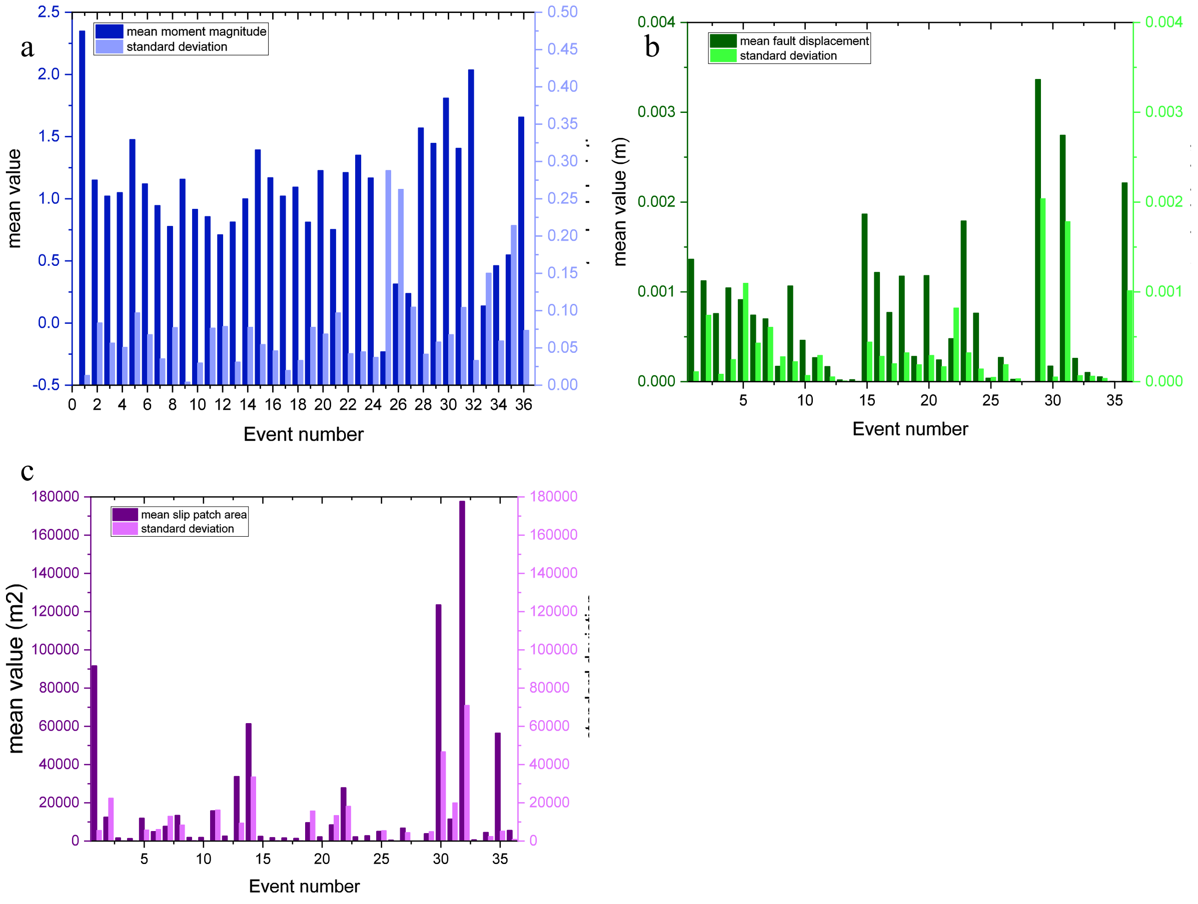
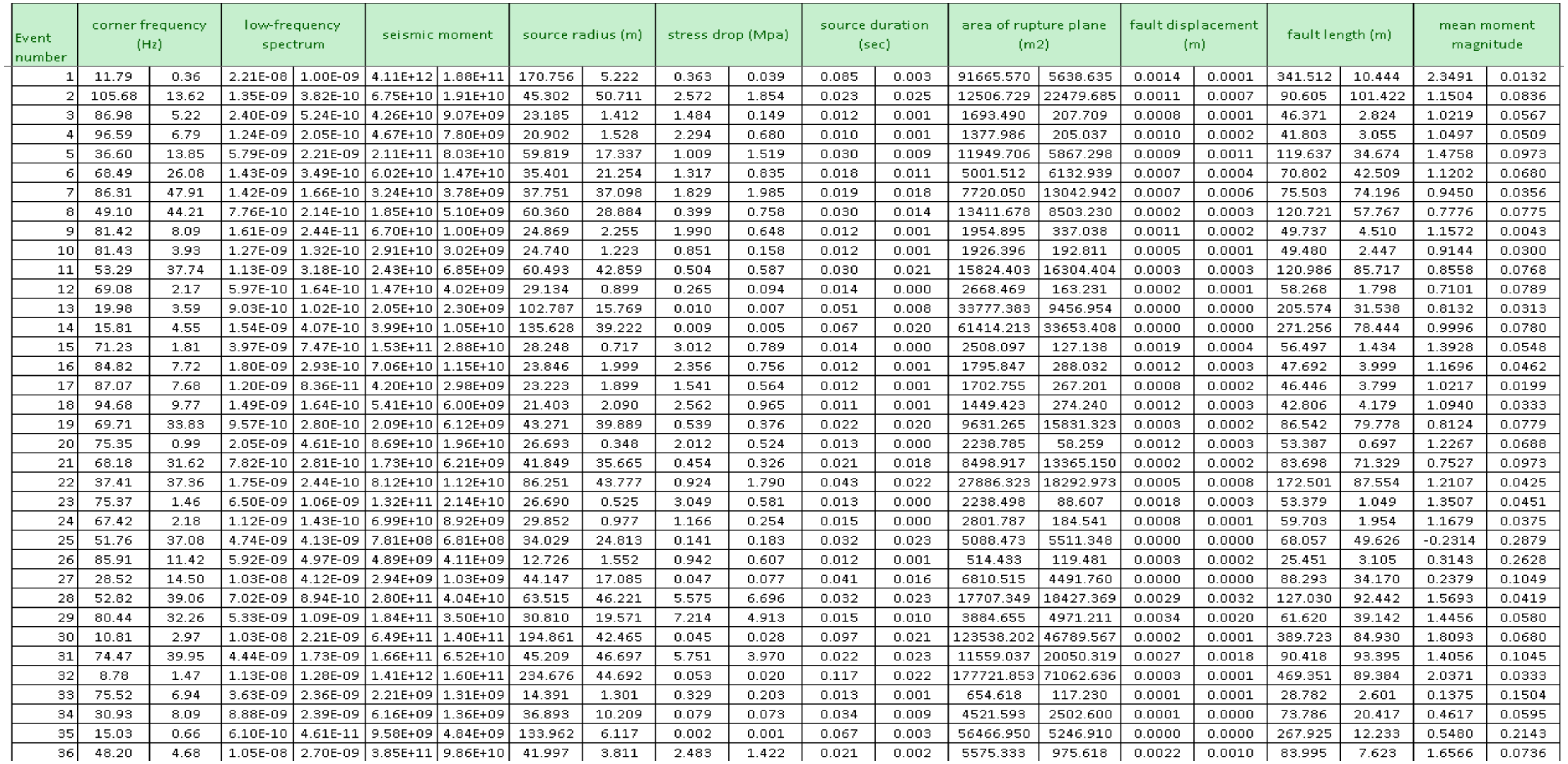


Figure 54: Moment magnitude (a), fault slip (b) and slip patch area (c) estimates for 36 events (Mean estimates of all sensors from the North array). In darker shaded bars are shown the estimates while in lighter the standard deviations

Table 11: Overview of the source parameters (mean values for all North array sensors)



Event number	corner frequency (Hz)		low-frequency spectrum		seismic moment		source radius (m)		stress drop (Mpa)		source duration (sec)		area of rupture plane (m2)		fault displacement (m)		fault length (m)		mean moment magnitude	
1	11.79	0.36	2.21E-08	1.00E-09	4.11E+12	1.88E+11	170.756	5.222	0.363	0.039	0.085	0.003	91665.570	5638.635	0.0014	0.0001	341.512	10.444	2.3491	0.0132
2	105.68	13.62	1.35E-09	3.82E-10	6.75E+10	1.91E+10	45.302	50.711	2.572	1.854	0.023	0.025	12506.729	22479.685	0.0011	0.0007	90.605	101.422	1.1504	0.0836
3	86.98	5.22	2.40E-09	5.24E-10	4.26E+10	9.07E+09	23.185	1.412	1.484	0.149	0.012	0.001	1693.490	207.709	0.0008	0.0001	46.371	2.824	1.0219	0.0567
4	96.59	6.79	1.24E-09	2.05E-10	4.67E+10	7.80E+09	20.902	1.528	2.294	0.680	0.010	0.001	1377.986	205.037	0.0010	0.0002	41.803	3.055	1.0497	0.0509
5	36.60	13.85	5.79E-09	2.21E-09	2.11E+11	8.03E+10	59.819	17.337	1.009	1.519	0.030	0.009	11949.706	5867.298	0.0009	0.0011	119.637	34.674	1.4758	0.0973
6	68.49	26.08	1.43E-09	3.49E-10	6.02E+10	1.47E+10	35.401	21.254	1.317	0.835	0.018	0.011	5001.512	6132.939	0.0007	0.0004	70.802	42.509	1.1202	0.0680
7	86.31	47.91	1.42E-09	1.66E-10	3.24E+10	3.78E+09	37.751	37.098	1.829	1.985	0.019	0.018	7720.050	13042.942	0.0007	0.0006	75.503	74.196	0.9450	0.0356
8	49.10	44.21	7.76E-10	2.14E-10	1.85E+10	5.10E+09	60.360	28.884	0.399	0.758	0.030	0.014	13411.678	8503.230	0.0002	0.0003	120.721	57.767	0.7776	0.0775
9	81.42	8.09	1.61E-09	2.44E-11	6.70E+10	1.00E+09	24.869	2.255	1.990	0.648	0.012	0.001	1954.895	337.038	0.0011	0.0002	49.737	4.510	1.1572	0.0043
10	81.43	3.93	1.27E-09	1.32E-10	2.91E+10	3.02E+09	24.740	1.223	0.851	0.158	0.012	0.001	1926.396	192.811	0.0005	0.0001	49.480	2.447	0.9144	0.0300
11	53.29	37.74	1.13E-09	3.18E-10	2.43E+10	6.85E+09	60.493	42.859	0.504	0.587	0.030	0.021	15824.403	16304.404	0.0003	0.0003	120.986	85.717	0.8558	0.0768
12	69.08	2.17	5.97E-10	1.64E-10	1.47E+10	4.02E+09	29.134	0.899	0.265	0.094	0.014	0.000	2668.469	163.231	0.0002	0.0001	58.268	1.798	0.7101	0.0789
13	19.98	3.59	9.03E-10	1.02E-10	2.05E+10	2.30E+09	102.787	15.769	0.010	0.007	0.051	0.008	33777.383	9456.954	0.0000	0.0000	205.574	31.538	0.8132	0.0313
14	15.81	4.55	1.54E-09	4.07E-10	3.99E+10	1.05E+10	135.628	39.222	0.009	0.005	0.067	0.020	61414.213	33653.408	0.0000	0.0000	271.256	78.444	0.9996	0.0780
15	71.23	1.81	3.97E-09	7.47E-10	1.53E+11	2.88E+10	28.248	0.717	3.012	0.789	0.014	0.000	2508.097	127.138	0.0019	0.0004	56.497	1.434	1.3928	0.0548
16	84.82	7.72	1.80E-09	2.93E-10	7.06E+10	1.15E+10	23.846	1.999	2.356	0.756	0.012	0.001	1795.847	288.032	0.0012	0.0003	47.692	3.999	1.1696	0.0462
17	87.07	7.68	1.20E-09	8.36E-11	4.20E+10	2.98E+09	23.223	1.899	1.541	0.564	0.012	0.001	1702.755	267.201	0.0008	0.0002	46.446	3.799	1.0217	0.0199
18	94.68	9.77	1.49E-09	1.64E-10	5.41E+10	6.00E+09	21.403	2.090	2.562	0.965	0.011	0.001	1449.423	274.240	0.0012	0.0003	42.806	4.179	1.0940	0.0333
19	69.71	33.83	9.57E-10	2.80E-10	2.09E+10	6.12E+09	43.271	39.889	0.539	0.376	0.022	0.020	9631.265	15831.323	0.0003	0.0002	86.542	79.778	0.8124	0.0779
20	75.35	0.99	2.05E-09	4.61E-10	8.69E+10	1.96E+10	26.693	0.348	2.012	0.524	0.013	0.000	2238.785	58.259	0.0012	0.0003	53.387	0.697	1.2267	0.0688
21	68.18	31.62	7.82E-10	2.81E-10	1.73E+10	6.21E+09	41.849	35.665	0.454	0.326	0.021	0.018	8498.917	13365.150	0.0002	0.0002	83.698	71.329	0.7527	0.0973
22	37.41	37.36	1.75E-09	2.44E-10	8.12E+10	1.12E+10	86.251	43.777	0.924	1.790	0.043	0.022	27886.323	18292.973	0.0005	0.0008	172.501	87.554	1.2107	0.0425
23	75.37	1.46	6.50E-09	1.06E-09	1.32E+11	2.14E+10	26.690	0.525	3.049	0.581	0.013	0.000	2238.498	88.607	0.0018	0.0003	53.379	1.049	1.3507	0.0451
24	67.42	2.18	1.12E-09	1.43E-10	6.99E+10	8.92E+09	29.852	0.977	1.166	0.254	0.015	0.000	2801.787	184.541	0.0008	0.0001	59.703	1.954	1.1679	0.0375
25	51.76	37.08	4.74E-09	4.13E-09	7.81E+08	6.81E+08	34.029	24.813	0.141	0.183	0.032	0.023	5088.473	5511.348	0.0000	0.0000	68.057	49.626	-0.2314	0.2879
26	85.91	11.42	5.92E-09	4.97E-09	4.89E+09	4.11E+09	12.726	1.552	0.942	0.607	0.012	0.001	514.433	119.481	0.0003	0.0002	25.451	3.105	0.3143	0.2628
27	28.52	14.50	1.03E-08	4.12E-09	2.94E+09	1.03E+09	44.147	17.085	0.047	0.077	0.041	0.016	6810.515	4491.760	0.0000	0.0000	88.293	34.170	0.2379	0.1049
28	52.82	39.06	7.02E-09	8.94E-10	2.80E+11	4.04E+10	63.515	46.221	5.575	6.696	0.032	0.023	17707.349	18427.369	0.0029	0.0032	127.030	92.442	1.5693	0.0419
29	80.44	32.26	5.33E-09	1.09E-09	1.84E+11	3.50E+10	30.810	19.571	7.214	4.913	0.015	0.010	3884.655	4971.211	0.0034	0.0020	61.620	39.142	1.4456	0.0580
30	10.81	2.97	1.03E-08	2.21E-09	6.49E+11	1.40E+11	194.861	42.465	0.045	0.028	0.097	0.021	123538.202	46789.567	0.0002	0.0001	389.723	84.930	1.8093	0.0680
31	74.47	39.95	4.44E-09	1.73E-09	1.66E+11	6.52E+10	45.209	46.697	5.751	3.970	0.022	0.023	11559.037	20050.319	0.0027	0.0018	90.418	93.395	1.4056	0.1045
32	8.78	1.47	1.13E-08	1.28E-09	1.41E+12	1.60E+11	234.676	44.692	0.053	0.020	0.117	0.022	177721.853	71062.636	0.0003	0.0001	469.351	89.384	2.0371	0.0333
33	75.52	6.94	3.63E-09	2.36E-09	2.21E+09	1.31E+09	14.391	1.301	0.329	0.203	0.013	0.001	654.618	117.230	0.0001	0.0001	28.782	2.601	0.1375	0.1504
34	30.93	8.09	8.88E-09	2.39E-09	6.16E+09	1.36E+09	36.893	10.209	0.079	0.073	0.034	0.009	4521.593	2502.600	0.0001	0.0000	73.786	20.417	0.4617	0.0595
35	15.03	0.66	6.10E-10	4.61E-11	9.58E+09	4.84E+09	133.962	6.117	0.002	0.001	0.067	0.003	56466.950	5246.910	0.0000	0.0000	267.925	12.233	0.5480	0.2143
36	48.20	4.68	1.05E-08	2.70E-09	3.85E+11	9.86E+10	41.997	3.811	2.483	1.422	0.021	0.002	5575.333	975.618	0.0022	0.0010	83.995	7.623	1.6566	0.0736

Table 12: Overview of the source parameters (mean values for all South array sensors)

Event number	corner frequency (Hz)		low-frequency spectrum		seismic moment		source radius (m)		stress drop (Mpa)		source duration (sec)		area of rupture plane (m2)		fault displacement (m)		fault length (m)		mean moment magnitude	
1	10.69	0.22	2.75E-08	9.66E-10	5.09E+12	1.79E+11	188.158	3.869	0.335	0.017	0.094	0.002	111258.193	4574.189	0.0014	0.0000	376.316	7.738	2.4113	0.0103
2	27.07	5.68	1.01E-09	3.06E-10	5.05E+10	1.53E+10	76.579	14.743	0.052	0.017	0.038	0.007	18935.404	6918.492	0.0001	0.0000	153.158	29.486	1.0637	0.1002
3	66.10	25.53	2.20E-09	3.44E-10	4.54E+10	7.30E+09	37.147	23.245	0.954	0.638	0.018	0.012	5608.179	7119.686	0.0005	0.0003	74.293	46.491	1.0418	0.0450
4	19.20	4.53	2.18E-09	1.27E-09	8.35E+10	4.87E+10	110.315	31.791	0.028	0.010	0.055	0.016	40612.917	24794.584	0.0001	0.0000	220.631	63.582	1.1897	0.1486
5	48.03	14.93	1.31E-08	8.09E-09	4.79E+11	2.94E+11	45.095	13.880	4.712	5.456	0.022	0.007	6842.483	3929.933	0.0035	0.0036	90.190	27.760	1.6840	0.1837
6	57.55	69.60	1.24E-09	1.09E-10	5.18E+10	4.58E+09	69.853	39.053	2.887	5.705	0.035	0.019	18922.914	13042.920	0.0008	0.0015	139.707	78.106	1.0820	0.0253
7	56.62	47.58	1.75E-09	5.21E-10	3.99E+10	1.19E+10	75.491	63.437	0.764	0.892	0.038	0.032	27385.282	30098.643	0.0004	0.0004	150.981	126.873	0.9971	0.0919
8	18.41	4.27	1.12E-09	3.21E-10	2.70E+10	7.71E+09	113.745	25.946	0.009	0.004	0.057	0.013	42232.060	18529.714	0.0000	0.0000	227.491	51.891	0.8854	0.0841
9	109.58	25.36	1.91E-09	8.85E-10	8.00E+10	3.71E+10	18.977	3.596	6.176	4.723	0.009	0.002	1161.775	390.799	0.0023	0.0013	37.953	7.191	1.1870	0.1282
10	74.86	36.60	1.26E-09	2.48E-10	2.90E+10	5.68E+09	39.542	35.276	1.029	0.759	0.020	0.018	7844.134	12641.443	0.0005	0.0003	79.085	70.551	0.9101	0.0620
11	37.19	37.01	1.34E-09	2.92E-10	2.87E+10	6.26E+09	87.441	46.257	0.316	0.613	0.043	0.023	29061.663	21123.822	0.0002	0.0003	174.881	92.514	0.9065	0.0616
12	67.42	23.04	6.26E-10	1.91E-10	1.55E+10	4.71E+09	34.100	16.839	0.329	0.205	0.017	0.008	4321.062	4471.425	0.0002	0.0001	68.199	33.678	0.7226	0.0909
13	35.45	35.83	1.19E-09	1.94E-10	2.73E+10	4.44E+09	93.999	51.756	0.247	0.477	0.047	0.026	34070.106	26336.827	0.0001	0.0002	187.998	103.513	0.8950	0.0452
14	9.91	1.89	2.78E-09	1.15E-09	7.23E+10	2.99E+10	208.556	40.076	0.003	0.001	0.104	0.020	140429.634	53049.525	0.0000	0.0000	417.112	80.152	1.1590	0.1279
15	75.67	2.19	2.66E-09	6.04E-10	1.03E+11	2.34E+10	26.593	0.769	2.418	0.637	0.013	0.000	2223.116	128.571	0.0014	0.0004	53.186	1.538	1.2758	0.0699
16	77.07	4.50	3.43E-09	9.17E-10	1.35E+11	3.60E+10	26.161	1.510	3.290	0.731	0.013	0.001	2155.486	247.002	0.0019	0.0004	52.322	3.020	1.3532	0.0719
17	76.51	1.87	3.01E-09	4.88E-10	1.06E+11	1.72E+10	26.295	0.636	2.571	0.603	0.013	0.000	2173.177	104.695	0.0015	0.0003	52.591	1.272	1.2869	0.0443
18	57.89	81.55	1.85E-09	4.83E-10	6.73E+10	1.76E+10	91.605	54.749	5.130	10.224	0.046	0.027	33424.850	23548.352	0.0013	0.0025	183.209	109.497	1.1503	0.0844
19	77.71	42.82	9.92E-10	2.61E-10	2.17E+10	5.69E+09	51.164	61.560	0.826	0.562	0.025	0.031	17152.973	31686.280	0.0004	0.0003	102.327	123.120	0.8240	0.0718
20	74.80	39.66	1.98E-09	2.55E-10	8.48E+10	1.09E+10	38.935	32.637	3.030	3.177	0.019	0.016	7272.117	11184.456	0.0014	0.0011	77.869	65.274	1.2238	0.0374
21	41.23	46.43	9.16E-10	3.31E-10	2.04E+10	7.38E+09	92.753	57.871	0.325	0.637	0.046	0.029	34918.562	32708.754	0.0001	0.0002	185.506	115.743	0.8006	0.0962
23	76.84	2.95	4.27E-09	5.68E-10	8.74E+10	1.17E+10	26.201	0.983	2.148	0.427	0.013	0.000	2158.921	159.859	0.0012	0.0002	52.402	1.965	1.2324	0.0387
24	51.93	12.67	1.07E-09	1.15E-10	6.59E+10	7.02E+09	40.918	12.123	0.586	0.387	0.020	0.006	5606.126	3513.432	0.0005	0.0002	81.835	24.246	1.1515	0.0312
36	44.67	6.28	8.60E-09	8.53E-10	3.03E+11	3.08E+10	45.63	5.83	1.57	0.89	0.02	0.00	6621.11	1606.16	0.00	0.00	91.26	11.67	1.59	0.03

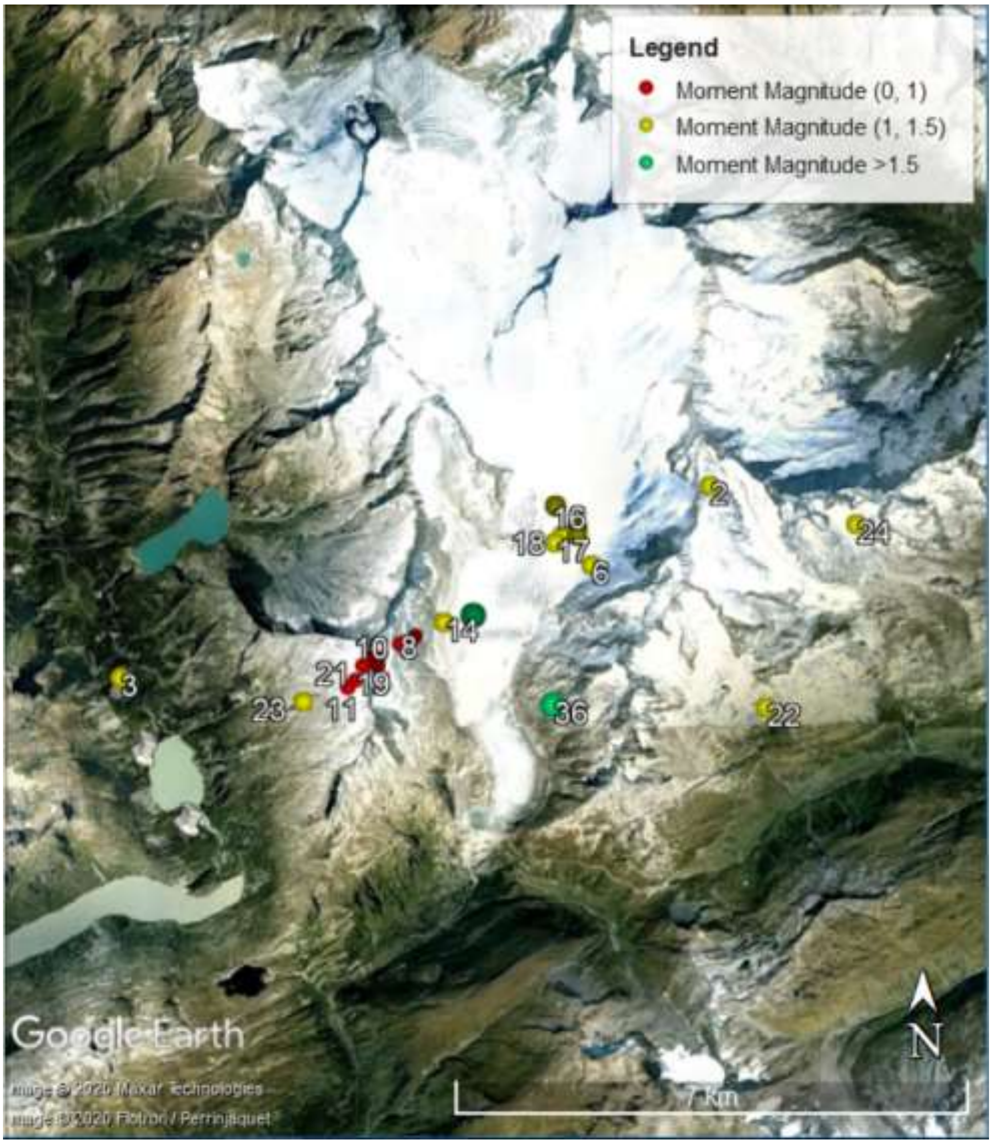


Figure 55: Moment magnitudes of events whose source parameters were estimated from both arrays

5.7 Discussion

5.7.1 Localisation of events

The time periods that were explored for RIS (after the 2 unloading periods of Raeterichsboden Lake) were expected to be the most probable for a microseismicity reaction to a stress change to happen provided the response was rapid. In case that microseismicity occurred as a result of the delayed effect of pore-pressure diffusion then it would be expected to occur later on, timewise.

By simply observing the fractures reported by Schneeberger et al. (2017) and Stillings (2020) it is clear that the detected events could be easily accommodated by these fractures. An approximate plane has been drawn that fits a number of the located events and is shown in Figure 56 with an orange dotted line as an extension to the mapped fractures to the events. The heading of the plane approximation is at 50°. Four of the events that fit to the plane occurred in DOY 75 of 2016 and in two consecutive hours. Their inferred corner frequencies fall within the range of $33.75 \leq f_c \leq 78.15$Hz, while the calculated seismic moments and moment magnitudes fall within the range of $1.89E+10 \leq M_0 \leq 2.9E+10$ and $0.77 \leq M_w \leq 0.91$. Fault lengths range from 64m to 174m, source radii range from 32m to 87m while the areas of the patches that slipped vary from 4,885 to 27,821m^2.

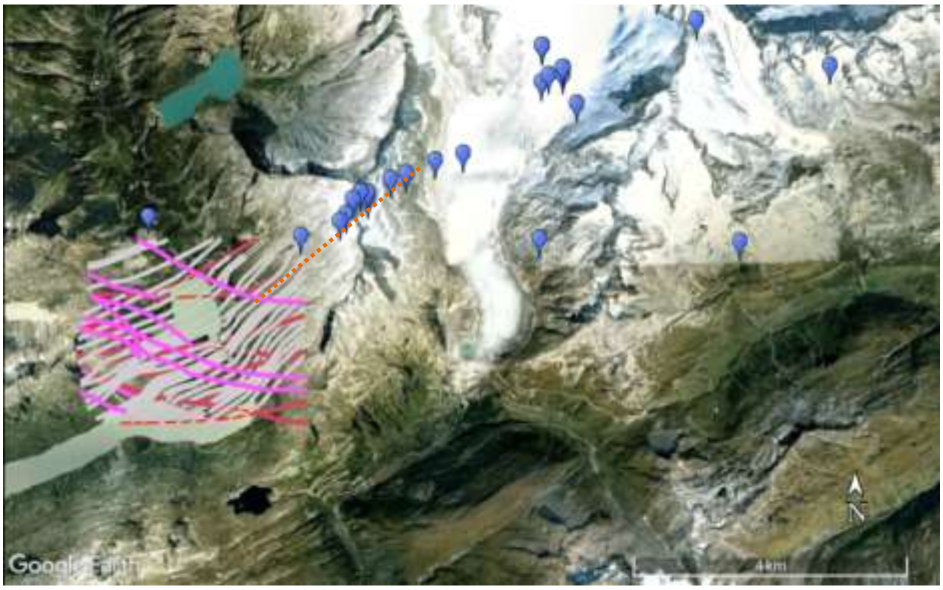


Figure 56: Plane fitting for a number of the located events presented as an orange dotted line. The plane agrees with the fractures reported by Schneeberger et al. (2017) and Stillings (2020).

It has been reported that events with bandwidths outside of the flat frequency response are distorted by the instrumental response and yield unreliable magnitude estimates if they are detected at all. Since my instruments flat frequency is 80Hz, and I have indeed provided source estimates for events with corner frequencies above this frequency, the updated list with events that yield reliable magnitude estimates is limited to events No 1, 5, 6, 8, 11, 12, 13, 14, 15, 19, 20, 21, 22, 23, 24, 36.

The solutions of the seismic inversion are out of the network, as expected from the shadowing analysis presented in Section 3.7, i.e. because of the tunnels acting as a barrier and shadowing the areas to the west of and between the arrays the recordability of seismicity occurring in these areas was evidently diminished. It has to be noted as well that it is possible that events located towards the edges of the grid are actually outside the grid at further distances and that the algorithm confused the global minimum and yielded as a solution a local minimum, hence the well resolved areas are those within a grid volume lesser than the one used for the localisation by a couple of grid points. This exclusion would event 1 as a not RIS event as it is located at a distance of 34km away, towards the north, from the north arrays' sensors.

5.7.2 Comparison with geochemistry

In order to contribute to the LASMO project (see also Figure 4) and specifically to explore the relationship between geochemical findings and microseismicity I ran some additional analysis in specific periods that alkalinity drops (during November 2014 and February 2016; Stillings (2020)) were detected during groundwater sampling in the GTS. In particular, I visually inspected 48 hours of data prior to the each recorded pH drop occurrence. Only those events that simultaneously satisfied certain criteria were considered: events that demonstrated impulsive P-wave arrivals; events that were detected by either all 3D channels or by four 1D channels; events whose P- and S- wave arrival times had a maximum of 0.3s time difference. While the first two criteria were imposed to ensure the events can be properly located, the latter was to limit down events within a pre-chosen area of 2km radius around the GTS. Any events with greater S-P time differences were assumed to have occurred further than 2km away and were not expected to be liable for pH changes in the GTS. In total, ten events that satisfied the aforementioned criteria were located (three in 2014 and seven in 2016).

The detected events are shown, as yellow circles, in Figure 57. Out of these, events (1) and (4)-(8) were located using P- and S- wave arrivals from all vertical components of the North array only. These events had less than 0.5 sec durations and were part of a 'sequence' of short duration events (inter-event distances of less than 1 sec) of 13 seconds total duration (Stillings, 2020). Events (9)-(10) were located using the P- and S- wave arrivals from all seismometers of both arrays. All of the above events were located using the algorithm proposed by Saltogianni and Stiros (2013). Origin times were estimated by Wadati plots using the P- and S-p arrival times. The remaining two events ((2) and (3)) were located using single-station location (Bormann, 2012) and the recordings of the 3D seismometer only.

The v_p/v_s ratio was calculated, using the arrival times of p and s phases for all events that were detected by at least 4 seismometers, and demonstrated a variance of up till 2%. Errors in location because of the P-wave picking inaccuracy (i.e. 2 sampling points ~±0.008 sec) reach up to approximately 270 m (assuming v_p=4.5km/sec and a maximum epicentral distance of 1 km). Errors due to the location method used reach up to 50m.

The local magnitude relationship for Switzerland was used to estimate local magnitudes for the events, as suggested by Fäh et al. (2011) and the resulting local magnitudes ranged between -1.2 and 1.04. The source dimensions were calculated as per Brune (1970). Corner frequencies for events (1)-(8) were within the range 13 and 20 Hz, corresponding to rupture patches of radii between 21 and 38 m, while events (9)-(10) had corner frequencies 96 and 100Hz, respectively, corresponding to rupture patches of radii 8 and 9m respectively. The estimated slip patch fracture surface areas for these events are 1,300 - 6,350 m^2 (Stillings, 2020).

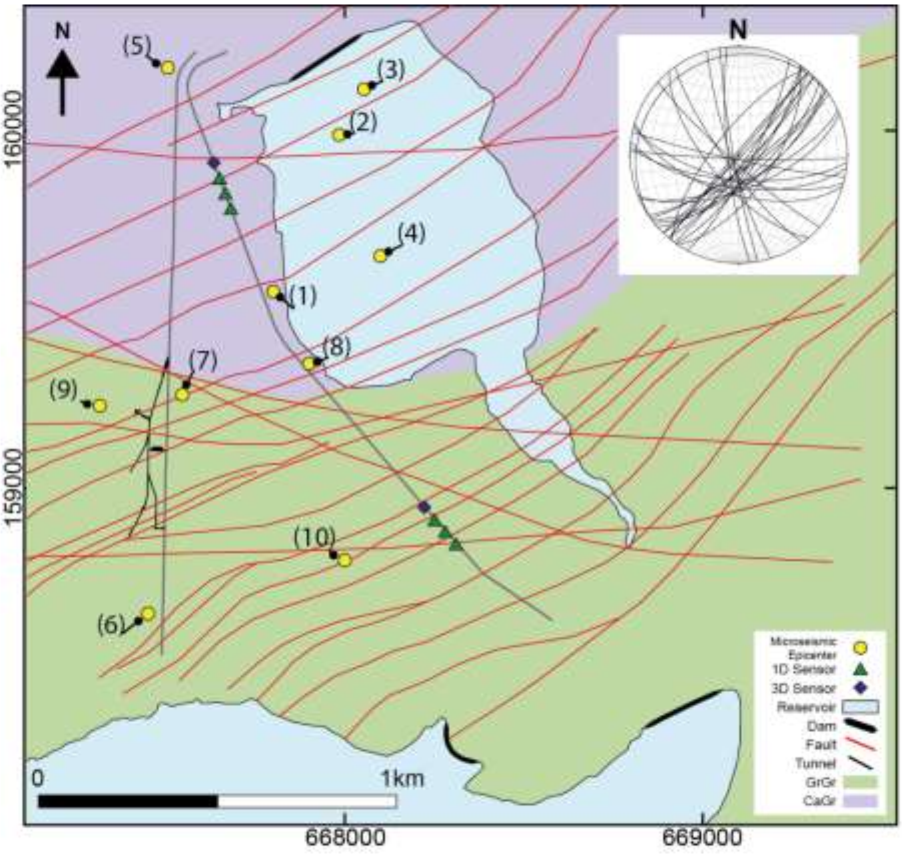


Figure 57: Figure adapted by Stillings (2020), representing microseismic tunnels (grey lines), fault traces (red lines) and the lithological contact at the surface between Aar Granite (purple) and Grimsel Granodiorite (green). The epicentres of the detected events are numbered and shown in yellow circles. The stereonet, at the top right corner of the figure, shows orientations of open fractures in the GTS and other tunnels.

Fracture sets previously mapped in the area (Schneeberger et al., 2017), striking NE-SW and NW-SE, have sufficient length to host seismic events of these magnitudes and are also visible within the GTS cutting the tunnels and boreholes (stereonet in Figure 57; Stillings (2020)).

By observing the events on this map we can see that most of the locations are either on mapped fractures or in their close vicinity. All the located events are very small and likely hosted on fractures (rather than major faults). Hence, it is not possible to match events to individual geological features as fracture density in the area is large. Analysis of the stereonets of fracture orientations indicates it is likely that they occur on the set of SW-NE oriented fractures that are observed to extend beneath the reservoir and are several kilometres in length.

Overall, these experiments demonstrated a causal relationship between the formation of new cracks seismically induced and the short-lived pH drops observed in GTS groundwater Stillings (2020). In other words, when a microseismic event occurs, a pH drop occurs in the groundwater within this slip patch and is propagated by a pressure rise through the fracture network to the sampled boreholes.

5.7.3 Partially efficient automated locations of events

5.7.3.1 Automated location using VFOM and AIC-wav

As part of the search for an automated way to locate the events of interest, by using the AIC-wav picks and not the manually picked events, I explored the virtual field optimization method (VFOM; Li et al.(2015)) as a location algorithm. VFOM optimizes a continuous and virtually established objective function and searches for the intersection of hyperboloids, determined by sensor pairs, in the 3D-space. The method claims to eliminate location errors caused by large picking errors (LPEs) with the assumption that all pairs of hyperboloids, except for those related to the LPEs, will intersect at the source. Many factors can cause LPEs, for instance a 50 Hz power line interference can lead to LPEs both in P-wave picking and S-wave picking, especially for arrivals with low SNR (Li et al., 2015).

VFOM introduces a function called closeness basis (CB) to describe the proximity of a point to all pairs of hyperboloids. Following the method's assumption, the sum of all CBs at the source should be greater than at other positions; this is the location determination. VFOM employs the Quasi-Newton algorithm (Métivier et al., 2014), a two gradient-optimization method, and the iterative process is repeated for various initial values.

User-defined input parameters include a homogeneous single-layered velocity profile, the number of sensors that need to be triggered for an event to be considered as valid, the estimated picking errors ranges that the input P- and S-wave arrivals contain (dtp and dts respectively), and the confidence level for the general picking quality and array display (lev). Because the VFOM is based on the information of sensor pair, i.e., the double difference, the minimum number of sensors that can be used for acceptable event location is 5. However, in this case no LPEs are allowed in the data. If the number of sensors is set to 6, one sensor containing LPEs may not influence the result significantly. Depending on the picking quality and array display lev parameter may range from values close to 1 (optimal) to 0 (worse cases).

There are two iteration-stopping criteria that can be used i.e., the SC-A and SC-B. SC-A is a measure of the picking quality, obtains more stable and accurate results in the 40% of the cases with errors in the scale of 20m while refuses to yield results if unreliable. SC-B ensures VFOM always yields a solution, with the majority of the errors lying again in the scale of 20m but with some reaching more than 50m.

The researchers tested the method using synthetic and real-field mining data that suffer from systematic LPEs, caused from uncertainties in the velocity profile and the arrival picking times. The method was checked against velocity uncertainties of 2km/s range, systematic errors of 4ms and LPEs of 0.2s with different probabilities. Overall, VFOM yielded more accurate and stable solutions against traditional location methods using arrival times with different probabilities of LPEs, both in the case that it was applied to P- and S-wave arrivals but also when it was applied to P-wave arrivals only. VFOM showed similar sensitivity in cases when the S-wave was mistakenly identified as a P-wave. With these optimistic findings in consideration VFOM was deemed as an optimal location algorithm to combine with the picks of the events from *AIC-wav* algorithm.

VFOM algorithm runs in Wolfram Mathematica programming language and was kindly provided by Wang, Z. Additional codes were written by me in MATLAB to automatically feed all *AIC-wav* results (.tsv files) into the particular input format VFOM required (see also **Appendix A**). The optimal VFOM parameters were set to fit the needs of my data: at least three sets of sensors had to be considered for an event to be considered valid, the picking errors for P- and S- wave were estimated to be in the range of ± 0.01s, and lev parameter was set equal to 0.2. SC-A stopping criterion was used, to ensure accuracy of locations.

Locations for all three years were estimated using VFOM. By way of example Figure 58 demonstrates on Google Earth all the locations as estimated using VFOM algorithm, for the events of year 2016 whose P- and S- wave picks were determined using *AIC-wav*. The events are denoted with balloons of different colours depending on their depth, ranging from the shallowest shown in red (-1.34 to -1.29km a.s.l.) to the deepest in blue (-1.89 to -1.84km a.s.l.). The fractures reported by Schneeberger et al. (2017) in Figure 7 are also drawn in the map. Fractures of different orientations are denoted with different colours. The events seem to nicely fit the fractures reported by Schneeberger et al. (2017).

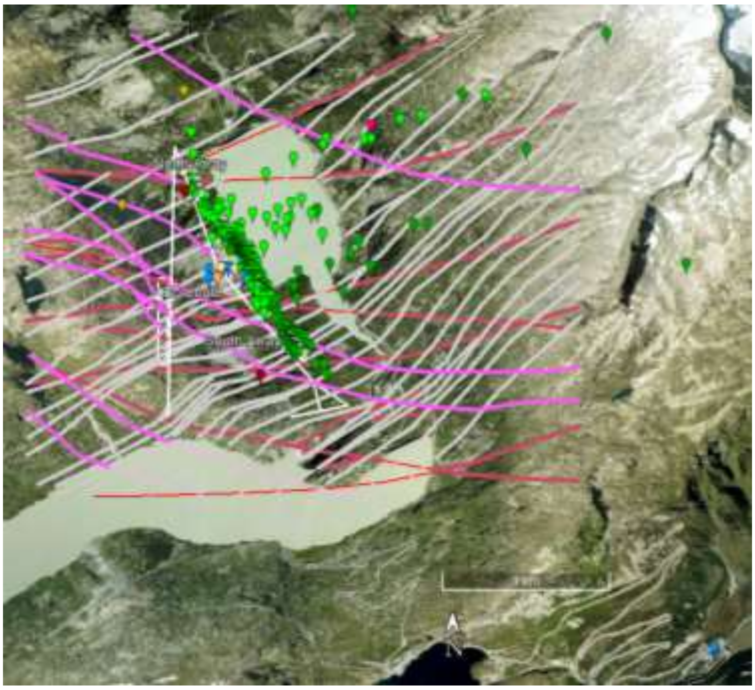


Figure 58: Locations obtained with VFOM algorithm using all P- and S-wave arrival picks as determined from *AIC-wav* workflow for the year 2016. The events are denoted with balloons of different colours as an indication of their depth, ranging from the shallowest in red (-1.34 to -1.29km a.s.l.) to the deepest in blue (-1.89 to -1.84km a.s.l.). The events seem to nicely fit the fractures reported by Schneeberger et al. (2017).

After the events were located, using the aforementioned parameters, a trial and error location process followed to test the sensitivity of the locations. During this test the values of the estimated picking errors for P- and S- wave were adjusted and values from ±0.1s to ±0.005s were used, while for lev parameter I checked values in the range of (0.2,0.8) with a step of 0.1. VFOM was also used with the option of using P-waves only. All combinations of parameters resulted more or less in the same image: a cloud of located events around Gerstenegg tunnel and more or less events further away, some of which coincided with fractures.

By way of example Figure 59 shows comparisons of the sensitivity of the VFOM-obtained locations of the 2016 events, using only P-waves, using different combinations of the input parameters, i.e. the number of sensors, estimated picking errors of the P- wave arrivals (dtp) and confidence level (lev). In particular, the figure demonstrates locations for the parameters (i) (8 sensors, 0.2, 0.01) shown with pins and (8 sensors, 0.2, 0.2) with balloons and (ii) (8 sensors, 0.2, 0.01) shown with pins and (6 sensors, 0.2, 0.01) with rhombi. The locations results seem quite stable.

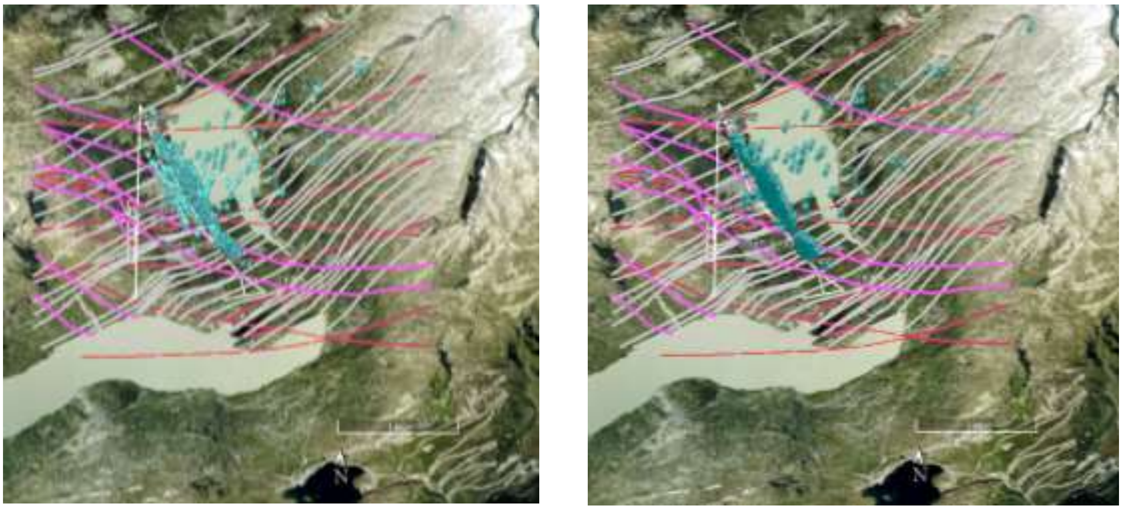


Figure 59: Comparison of locations for the 2016 events obtained with VFOM with different combinations of the input parameters (no of sensors, lev, dtp): (i) locations for the parameters (8 sensors, 0.2, 0.01) are shown with pins and (8 sensors, 0.2, 0.2) with balloons and (ii) locations for the parameters (8 sensors, 0.2, 0.01) are shown with pins and (6 sensors, 0.2, 0.01) with rhombi.

This consistency gave birth to doubts: either the input data were indeed stable which would mean by extension that there were not LPEs in my dataset (a scenario dismissed by default only by taking into account the 50Hz EM interference) or VFOM was not yielding unbiased locations. Indeed, if we examine one of the located events shown in Figure 60 (the first one of this series of events) we witness a discrepancy. According to the waveforms on the left-hand side of the figure, the signals are first recorded by the North array (top six waveforms) and then by the South which yields the expectation that their origin would be towards the North array (Chapter 5.3.1.5 Arrival order approaches). The arrays are indicated on the right hand side of the figure with yellow dots. The located event however is located really close to the South array (position shown with a pin and a yellow arrow) which constitutes a paradox. The illustrated example was not the only inconsistency found within the located data and that was the reason why my confidence in the algorithm was shaken.

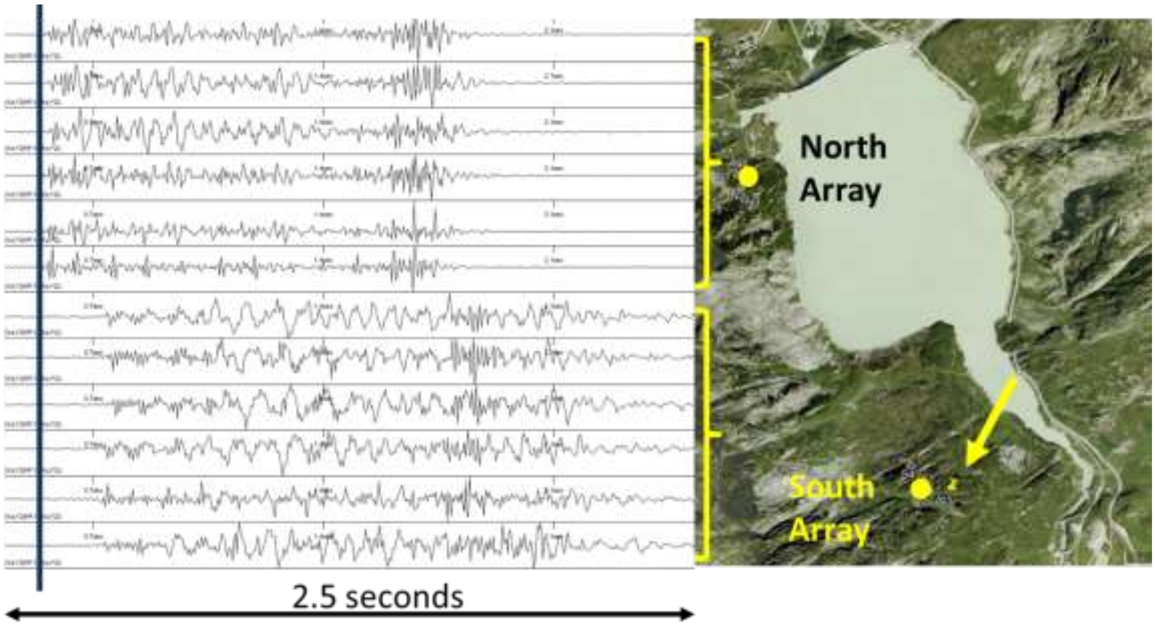


Figure 60: Mutliple events 26 Nov 2016, ~08:13:01. Signals arrive first at the North Array, then at the South Array. Calculated location from VFOM is not consistent with this.

5.7.3.2 Automated location using TOPINV and AIC-wav

As an alternative approach, I checked one of the LPEs intolerant versions of the commonly used grid search algorithm (i.e. TOPINV), with the AIC-wav picks, even if the LPEs yielded problematic locations, if these were of small scale, they could be tolerated. TOPINV algorithm was applied for the *AIC-wav* picked events in a grid with spacing of 0.05km in all three dimensions and using only the P-wave time arrivals of all 8 stations. The process was then repeated using both phases time arrival information. The 3-dimensional grid of points chosen is assumed that surrounds the real solution. Then, for each grid point and all triggered sensors, an error tolerance of 0.05km is assigned to the corresponding measurement. The algorithm is repeated for all grid points, and the common section of the sets of grid points is defined. From this subset of grid points, the estimate with the minimum RMS is chosen as the estimate of the optimal solution. A number of *AIC-wav* outputs were located using both approaches. The aim of this check was to see whether the issue of the LPEs that arose from the use of *AIC-wav*, identified mainly in the S-wave arrivals, could be overcome by the use of TOPINV algorithm using just the P-wave arrivals. One of the comparison visualisation figures produced can be visualised in Figure 61.

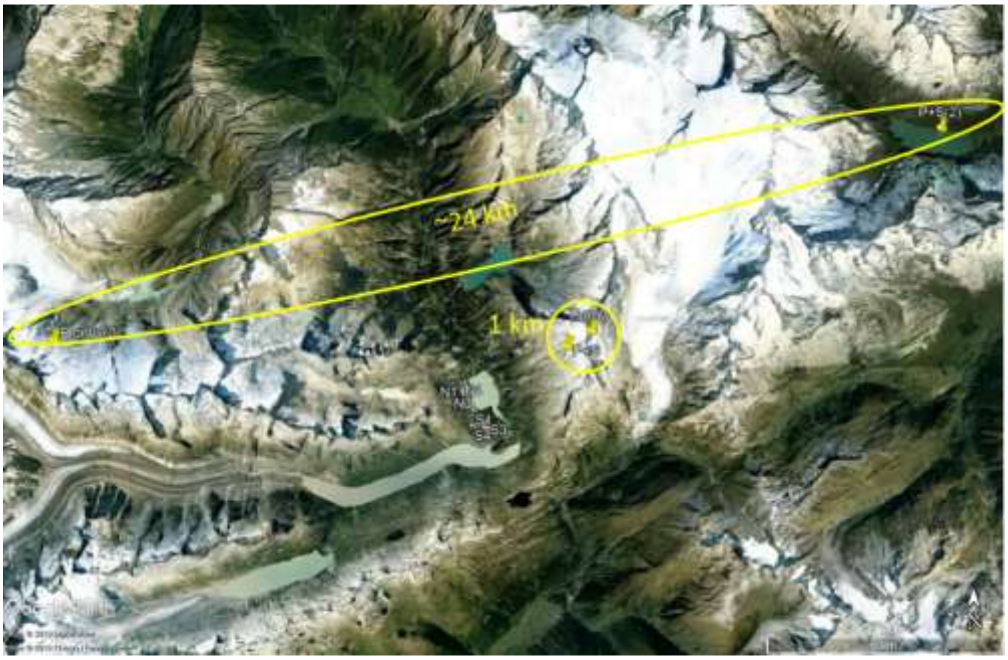


Figure 61: Comparison of locations estimated by TOPINV algorithm using just P-wave arrivals as input and both phases.

Figure 61 clearly illustrates that there can be great discrepancies in the outputs of TOPINV algorithm when the input data are just the P-wave phases as compared to when both phases arrivals are used. Particularly in this figure we can see two events locations as located by the algorithm. In the first case there is 1km distance between the locations estimated using just the P-wave and both phases, while in the second case there is a 24km discrepancy. Such an error in location estimations does not provide confidence.

5.7.3.3 Automated location using clustering, VFOM and AIC-wav

As another alternative, and again with the purpose of using the automatically picked AIC-wav picks, I decided to use clustering (similar to De Meersman et al. (2009)) and see if the LPEs still influence in a major way the results. For this analysis, I used the GISMO toolbox (Thompson and Reyes, 2018) a free, available online, seismic data analysis toolbox for MATLAB. I had to write the necessary algorithms in MATLAB for conversion of my data in the input format that GISMO required, i.e. waveform objects. I chose an adequately high inter-cluster correlation index (0.80, shown with a red vertical line in Figure 62(a)) and cross correlated all waveforms, i.e. 0.6s duration cropped traces around the *AIC-wav* detected events. In Figure 62(a) we can see the hierarchical cluster tree relationship between traces for the South array sensors and for all years. The cluster tree relationship was then trimmed into discrete clusters of events (15 in total). The events per cluster and over time can be seen in Figure 62(b).

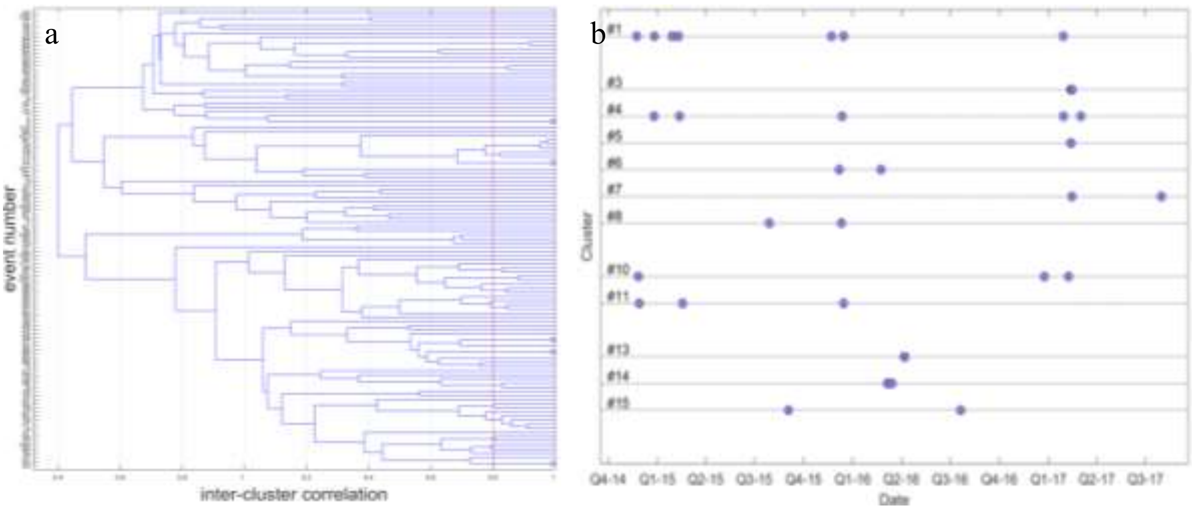


Figure 62: (a) Inter-cluster correlation index used for clustering of the events, and (b) spread of events per clusters over time

In Figure 63(a) we can see the stack of the traces for each cluster as well as the number of events that were included per cluster. All traces from the largest cluster (#1, 7 events in total) are plotted together with the stack of the traces in Figure 63(b). Finally, in Figure 64 we can visualise the spatial distribution of the clustered events. We can see that all events are below the ground surface (shown with a mesh grid) and within the boundaries of the sensors of the North and South array.

At first glance, the clustering of events seemed to provide a solution to the LPEs. However, quality control over the events included within the clusters proved that previously recognised as actual microseismic events were ignored, i.e. they were not included within any cluster. Signals within clusters were in fact either part of longer signals (clusters 1, 4, 6, 10 and 11 were actually precursors of landslides), or symmetrical signals that could either be attributed to mechanical glitches or just noise.

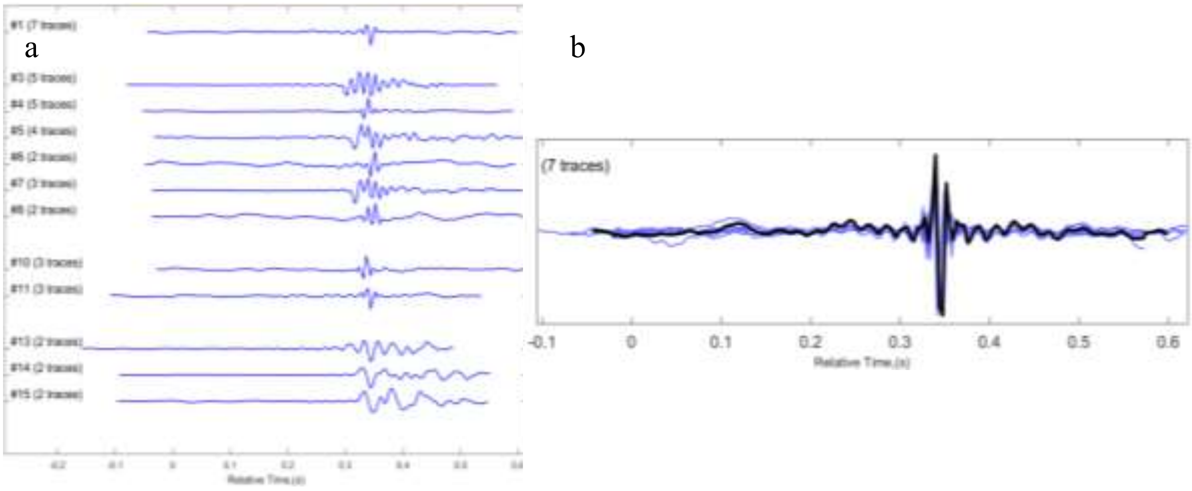


Figure 63: (a) Stack of the traces and the number of events per cluster, and (b) all traces from the biggest cluster (#1) plotted together and aligned.

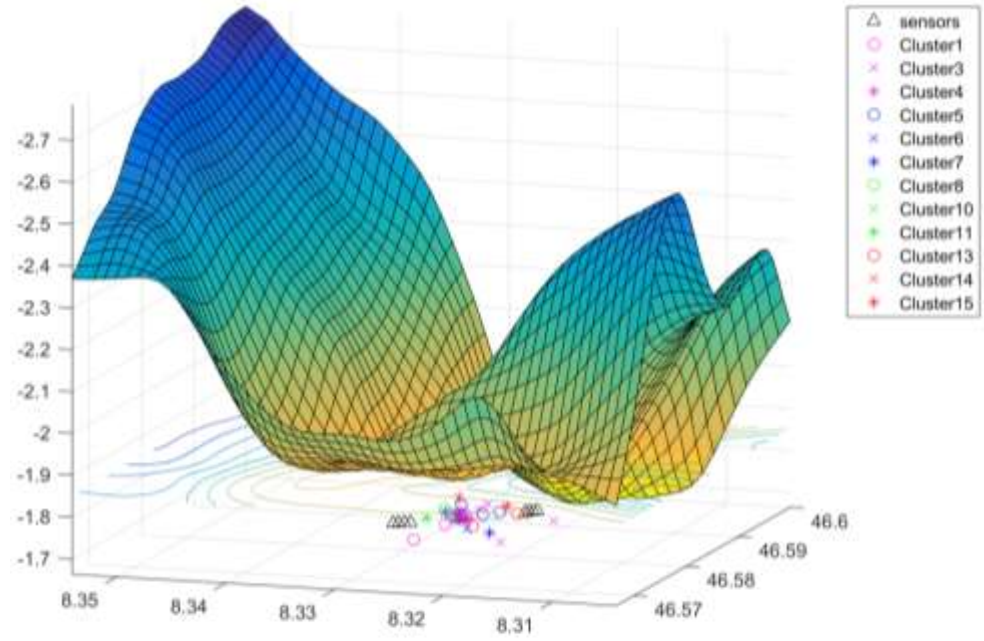


Figure 64: 3D representation of events. In the figure we can also see the sensors and with a mesh grid surface the ground surface.

5.7.4 Parameters that influence the source mechanisms estimates

There are a number of factors that influence the source parameters estimates.

- M_w estimates may vary depending on the method chosen to compute M_o. According to Stork et al. (2014) who compared five different methods: time domain estimate, Brune model estimate, Boatwright model estimate, directly measured Ω_0 with spectra not corrected for Q and with spectra corrected for Q; can demonstrate discrepancies of up to 0.6 units. The greatest differences were reported between the time and frequency domain estimates of M_w when the latter combined correction for attenuation. Q-corrected spectra M_w estimates were found to be larger than the values estimated from uncorrected spectra.
- Another parameter that affects the source parameters estimates is the radiation pattern correction. Stork et al. (2014) compared the M_w values estimated using R calculated from the focal mechanism solutions given in Rutledge et al. (2004) to the ones using the average values for the radiation pattern (0.44 for P-waves and 0.60 for S-waves) and witnessed that differences can reach up till 1.1 magnitude units. The greater discrepancies were reported in cases of larger events that S-waves were used for their M_w estimates. This effect occurs because the larger events are reported by a greater number of more distributed sensors. Some of these sensors are closer to a node in the S-wave radiation pattern than the P-wave radiation pattern. This highlights that the radiation pattern correction is particularly important if the receiver is situated near a node in the radiation pattern because R changes very rapidly with a small change in fault plane solution or event location. To account for this discrepancy it is advisable that averaged values of P- and S- phases should be used and these should be averaged over all sensors available for an event.
- Another discrepancy occurs when estimates are based on measured energy from a P- rather than an S- phase. In particular, S-wave estimates are observed to be larger than P-wave estimates. The maximum difference between the P- and S-phase estimates can be up till 0.8 units. This discrepancy between P- and S-phase estimates arises largely because of the aforementioned radiation pattern coefficients errors. The estimates from S-waves are expected to be more reliable because usually the SNR for S-waves is

higher, although, depending on the proximity of the receiver to the event, the S-wave arrival could be affected by scattered P energy.

- The length of the window used for estimating the values of M_w is another factor that, if inappropriately chosen, can cause errors in the estimates. M_w values were estimated using time windows picked individually for each recording. Stork et al. (2014) found that the P- and S- phase estimates may have outliers of up to 0.4 units, which occur because of the impulsive nature of the arrivals. This is because longer time windows do not contribute significant extra event energy but instead cause a reduction in the SNR of the time window used for M_w estimations.
- The P- and S- arrival time separation is another parameter that restricts the analysis. That is, at a sampling rate of 250Hz the minimum ts-tp that could be analysed was 0.04s, allowing for 10 data samples within that window. In cases of time windows with less than 20 sampling points, the length of the window chosen to perform M_o calculations is not sufficient to maximise the event energy while minimising the contribution of other sources of energy, rendering the estimation of the frequency spectrum liable to underestimation (Stork et al., 2014). In general, P-phase values are more affected by this parameter than S-phases, which is expected because the S-phase time windows are longer. P-phase estimates at a sampling rate of 250 Hz may be up to 0.3 units smaller than those at 1000 Hz.
- Before a magnitude estimate is even attempted, pre-existing work may affect the accuracy of any M_w estimates. Any errors in the velocity model, the rock density and the source location will propagate into the M_w estimates. For instance a 10% error in the velocity, distance and density values would cause a 22% error in Mo estimates. An error of such degree in a 10^8 Nm M_o could affect M_w estimate by 0.1-0.2 magnitude units (Stork et al., 2014). Apart from the geophysical parameters, the quality of the signals could also affect the estimates. However the work of Stork et al. (2014), having examined ratios of 1.5, 3, 5, 10 SNRs showed that there was not a discernible influence in M_w estimations (<0.1 units) provided noise can be separated from the signal.

Regarding the velocity profile, based on the findings of Angus et al.(2014) and Usher et al.(2013) who claimed that the influence of vertical velocity variation is less problematic in surface microseismic monitoring than in downhole monitoring, and that differences between velocity models result in location differences of circa 20m, respectively, I decided to use a single layered velocity model in my case study. Since the present research is not a hydrofracking experiment, but covers instead an area of kilometres, I figured that such an accuracy uncertainty would be deemed acceptable. Hence, I chose a homogeneous 1-D velocity model based on velocity data from previous experiments conducted in the GTS area and optimised my selection by minimising the residuals between theoretical and actual locations of documented tectonic earthquakes' locations.

Although the velocity errors are not expected to affect on their own greatly the location and source parameters results, the cumulative effect of errors caused by the combination of onset time picking accuracy, velocity profile, location technique, source – receiver configuration, anisotropy may end up in large variances from the true conditions. The location estimates had an average fit error, i.e. the sum of the absolute residuals, of less than 143 m. For the purposes of my experiments that needed to cover a large area such errors were accepted as satisfactory.

- Finally, uncertainties in the estimates can occur because of the assumptions of the analysis. For instance Brune source model assumes that the events have pure double-couple mechanisms which may not be the case for my events. Also, magnitudes were calculated using spectra corrected for frequency - independent Q. In fact, it is likely that Q is frequency dependent (e.g., Campillo et al., 1985; Oye et al., 2005), but this is not thought to significantly affect the results.

5.8 Summary

To conclude, in this Chapter I have presented the literature review researched for localisation and source parameters determination of detected microseismicity and the methodology that was followed.

In particular, taking into account previous seismic explorations in the GTS area and the velocity profiles researchers used in these occasions I chose a homogeneous 1D model, performed some sensitivity analysis using earthquakes reported by SED and recorded by my sensors, and resulted in accepting $v_p = 5.4km/s$ and $v_s = 2.9km/s$ as the velocity values to be used for the localisation process.

I then chose to use TOPINV algorithm to locate the manual re-picks of the onset phases of my data (Section 4.5). The TOPINV grid-search algorithm was used to locate events for both P- and S- wave times and only for the detected events for which at least 5 out of the 8 seismometers of the two surface arrays were triggered. In total, out of 168 events that their phases were re-picked, 70 microseisms were located using the TOPINV algorithm (successfully converged). This small percentage of convergence is attributed to the dimensions of the grid used for the analysis, i.e. the sources of the re-picked events would have been outside of the limits of the grid and therefore the scope of my research (not RIS). The location estimates had an average fit error, i.e. the sum of the absolute residuals, of less than 143 m.

The source parameters of the located events were calculated assuming the Brune source model (pure double-couple mechanisms). As an overview, the hypocentral distances sampled by this set of events span a range from ∼3.31 to ∼34.85 km (considering both arrays). Inferred corner frequencies exhibit a high degree of scatter as they fall within the range $8.78 \leq f_c \leq 109.58$ Hz with corresponding low frequencies of 1.13E-8 and 1.35E-9Hz. Calculated seismic moments and moment magnitudes fall within the range of $7.81E+8 \leq M_0 \leq 4.11E+12$ and $-0.23 \leq M_w \leq 2.41$ respectively; this range reflects a sampling bias towards larger magnitudes, due to the focus of detection for this experiment. Fault lengths range from 25.45m to 469.35m, source radii range from 12.73m to 234.68m while the areas of the patches that slipped vary from 177,721 to 514.43m^2.

Novelty in the present chapter lies in using, comparing suitability and adjusting a number of well-known location determination approaches from the literature and resulting in the production of a seismic catalogue related to the dynamic response of the rock mass to reservoir drainage (Appendix B).

Chapter 6 Discussion and further work

The present chapter is an overview of the research undertaken, highlighting the novelty contributions and key findings on the various topics researched within this thesis, as well as further work.

6.1 Research overview

In **chapter 2**, I introduced the chosen small-scale field analogue to glacial retreat: the drainage and refilling of Raeterichsboden lake in the Swiss Alps (subsection 2.2). The reason why RIS has been chosen as an analogue to glacial retreat induced seismicity was explained as follows: both RIS and glacial retreat induced seismicity have been documented to follow water level variations, in a similar way, because both phenomena cause the same two processes of short-term stress adjustment: the direct effect of loading and pore water pressure changes (subsection 2.3); hence the short-term response to ice thickness fluctuations is similar to water reservoir induced fluctuations. In particular in this project, the planned construction works of KWO, which led to the unloading of Raeterichsboden lake twice in a period of two years, as well as the conveniently located underground GTS tunnels, rendered Raeterichsboden lake as an ideal reservoir to monitor for RIS. The stress changes in a reservoir have been reported to be followed by a rapid rock response, i.e. low magnitude, shallow seismic events, occurring below or in the immediate vicinity of the reservoir area. The pore water pressure changes, on the other hand, are followed by a delayed rock response and may extend well beyond the reservoir confines ($\geq$ 10km), not show an immediate correlation with major changes in reservoir level and be associated with large magnitude earthquakes, at deeper depths (Simpson et al., 1988).

However, a glacier may consist of a 1km thick ice load, while the reservoir load head of Raeterichsboden lake is 74m. The stress release of 0.74MPa (74m) only triggered small microseismic events ($-0.23 \leq M_w \leq 2.41$). In the case of a 10MPa load change, i.e. glacial unloading, it may be the case that large magnitude events could be triggered as well as more frequent microseismic events ($M_L \geq 2.5$; Sauber and Molnia (2004)). However, glaciers melt slowly during the warm months and then refreeze during the cold ones, so a rapid unloading of 10MPa is highly unlikely.

The faults in the Grimsel area have been examined by various researchers but here the findings of Raphael Schneeberger were taken into consideration (Section 2.6; Schneeberger et al., 2017). As an overview, the area is governed by granitic rocks with 3 main groups of fractures, A, B, and C, dipping southeast, south and southwest respectively. The majority of the faults are in the south of the Raeterichboden lake and GTS tunnels. Not all fractures that are contained within these groups are of the same size, i.e. group A consists of 8 major faults (dip<80°) and 23 relay structures, while group B from 6 major and 7 relay, and finally group C from 6 major and 32 relay. However, as per Costain et al. (1987), in most cases RIS is attributed to a small stress increase triggering the release of large pre-stress, i.e. a pore pressure stress increase triggering failure along a pre-existing fault already tectonically stressed close to failure, rather than because of a great stress increase. Hence, all fractures are susceptible to microseismicity, while only the size of the fracture slip will inform the magnitude of the event. In the Grimsel area, the transfer of stresses originating from an induced microseismic event can be facilitated through water conduits at large distances since the region is dominated by granitic rocks with an interconnected network of fractures of lengths of a few hundred metres. This may not be the case if the host rock is impermeable for instance, and in which case any transfer of stresses in a wider region would not be permitted, hence the energy dissipation would have to be in the immediate vicinity of the reservoir.

In **chapter 3**, the monitoring approach was presented and the reasons why it was chosen as such were explained. In particular: microseismic monitoring was selected as the optimum approach for a number of reasons: (i) it allows for the recording of very weak events of energy dissipation, i.e. microseismicity; (ii) it is a continuous monitoring method, therefore appropriate for a 3.5years continuous monitoring application; (iii) data can be transmitted in real-time from the monitoring site to the University premises, i.e. staff on-site are required only for the seismic sensors installation and regular maintenance; (iv) there is a great availability of automated software routines; (v) seismometers can be easily and rapidly deployed at a close proximity to the reservoir to record events; (vi) seismometers are easily retrievable; and finally (vii) the necessary equipment was readily available for this application.

The optimal seismic network configuration depends on the type of sensors, the size of the seismic network, and the geometry; all of the above are determined based on the scope of the application. The seismic network configuration influences greatly the localisation of events, hence the researcher should invest time in choosing a seismic configuration that suits their project's needs. There are always going to be limitations in the seismic network design, for instance, in this particular application, the available instrumentation consisted of two surface arrays (of three 1-D sensors and one 3-D sensor each) and a borehole 3-D sensor, while the availability of locations for sensors deployment was limited by external constraints such as accessibility and jurisdiction. With reference to the type of sensors, given that in this application the goal was to record microseismicity, the readily available sensors owned by the University of Strathclyde were deemed satisfactory, as their sensitivity (80Hz) targets weak motion. With reference to the size of the network, the minimum amount of records / readings for the source location determination is four (Joswig et al., 1992), hence the number of sensors (15 channels in total) was considered adequate for microseismic monitoring.

Surface arrays and downhole sensors have different advantages / disadvantages (e.g. see Eisner et al., (2010)) and it is common practice to combine their use in the field to counterpart their shortcomings. Borehole sensors / arrays suffer from lateral positioning accuracy while array sensors from low SNR because of the environmental noise that exists in the surface and the signal attenuation (mostly of the s-wave). Conversely, surface arrays demonstrate high lateral positioning accuracy and borehole sensors / arrays higher depth resolution accuracy. As there was a specific budget for this application, there was not the availability to purchase further instrumentation hence the two arrays and the borehole sensor had to be positioned in a triangular relevant positioning to achieve a good hypocentral and lateral coverage of seismic locations. The best estimates of the seismic locations would be for events originating from within this triangular area (see also 5.3.1.2). Because of the borehole sensor's malfunction however the triangular positioning was no longer feasible and the experiments proceeded with a final number of eight sensors (six 1Ds and two 3Ds; i.e. a total number of twelve channels).

The pre-processing of the acquired data comprised of the transformation of raw data to a format that could be read and processed using MATLAB (i.e. the majority of the codes were written in this language, incl. NpD algorithm) R studio (*AIC-wav*) and Mathematica (VFOM), the conversion of the data from counts in velocities, the formation of data series, the corrections of time lags in the start times of the hourly records, the conversion of the timeseries from cell to structure arrays.

The preliminary visual observation of the acquired data which was undertaken to allow for an understanding of the noise sources that influenced the recordings and check the effectiveness of the microseismic monitoring network is also presented in Chapter 3. Various reoccurring noise signals of different frequencies and amplitudes were evident in the recordings. The sources of noise that were frequent in the area were related to hydroelectric activity, to construction activities, hydrofracking experiments, large seismic events and landslides. The identification of small tectonic earthquakes, which were also found in the Swiss Seismic Network database, confirmed that the sensors were able to successfully record small seismic events even at long distances.

The network design analysis (again presented in Chapter 3) demonstrated that the Gerstenegg tunnel is shadowing the area on its west and between the two arrays, while the main access tunnel on its west (GTS area). Both tunnels act as a barrier, rendering all events occurring in the areas to the west of and between the deployed arrays more difficult to be recorded. No shadowing was observed on the east side of Gerstenegg tunnel. The tunnels shadowing is expected to affect the recording and localisation of events. Further analysis indicated that were there additional sensors arrays deployed in the main access tunnel they would allow for better recordability in the area between Gerstenegg and the main access tunnel. The magnitude threshold analysis demonstrated that magnitudes of up till -3M_w are detectable only in the close proximity of each array (0.4km) which coincides with the area that suffers from the Gerstenegg tunnel's shadowing. At longer distances the detectability falls to magnitudes of -2.30M_w (0.8km) and -1.58 (after 1km).

In **Chapter 4** the researched literature review and methodology followed for detection and onset time determination of potential microseismicity was presented. In particular, to detect weak seismic signals in not well-constrained conditions and without requirement of a priori knowledge about the expected signal frequencies and amplitudes, I developed the NpD algorithm. The main advantage of my approach is that it does not require any pre-filtering of the data as would be the case for detection of weak signals with most other methodologies. As a result, pre-filtering could remove information from the recordings, discarding it as noise, especially in cases of low SNR data. Avoiding pre-filtering altogether, minimizes the possibility of information loss in these low SNR recordings.

Another advantage of the NpD algorithm is that it is suitable for non-stationary background noise since the upper bound to the spectral amplitude of background noise, above which an event is detected, varies over both space and time; significant differences were observed in hourly noise characteristics between sensors 1km apart. The approach is also equally effective with non-parametric data i.e. an assumption of normality is not required.

The efficiency of the NpD algorithm has been proven through the comparison with other well-known approaches (STA/LTA and PSD picker). However, although considerable effort was put to avoid having multiple NpD peaks for the same signal, i.e. consecutive events in consecutive PSD time segments were "cleaned up" and only the first arrival from the consecutives was considered a trigger, it was observed, that in the cases of landslides for instance this was ineffective. The initial crack is perceived as a separate event to the main body of the landslide and NpD triggers multiple peaks for an individual event in this occasion. The NpD algorithm is a powerful microseismicity detection tool but its output does not include accurate onset times for the detected events. Its accuracy depends on the duration of the individual time segments to which each recording is divided. For windows of duration 0.5 seconds, such as those used in this case study, it means that the onset time is within a 0.5 second frame centred around the estimated NpD time of the 'event'.

The full data set of passive seismic data acquired between 1st November 2014 and 12th August 2017 was scanned using the NpD algorithm resulting in more than 241,000 detections. Some observations with reference to the acquired detection dataset as compared to the lakes water levels variations are summarised as follows: (i) the numbers of events detected by each of the individual arrays are broadly consistent, (ii) an anomalously high number of events are detected by the North and South arrays in early 2015, which may be as a result of reservoir drainage and refilling, (iii) the number of events appears to peak annually in the summer months (possibly related to the increase in glacial meltwater and consequent increase in reservoir levels, e.g. Johnson et al. (2017) and Roth et al. (1992), and (iv) a peak in event numbers may be associated with reservoir drainage.

It was also discussed that it is possible that (i) the unloading / reloading of Raeterichsboden lake does not have a direct effect on the granitic rocks which would potentially need a stress change higher than historic stress changes (see also Gupta et al. (1972b)) to respond, (ii) the effect of these 2 cycles of Raeterichsboden lake's drainage and refilling may induce delayed microseismicity response (Gupta, 2002), in the years that monitoring would no longer be in place, (iii) a seismic response captured by the sensors during this application may be the result of previous unloading and reloading of the reservoir, expressed with a time delay, (iv) the South array may be capturing weak events due to the lake water levels changes of both Raeterichsboden and Grimsel lakes because of its location, and (v) it is not clear which mechanism controls the seismicity; i.e. is it a rapid response to the reservoir loading or a delayed response to the unloading (see also Simpson et al. (1988)).

For the onset time picking of this application, I decided to proceed with the manual picking of the phases only for periods of interest. This involved an extensive amount of pattern recognition and great effort to pick the first breaks in a consistent and, as possible, objective way. For the visual inspection of the data records within the time periods of interest RTQT_View from REFTEK was used and the data in counts were filtered with a recursive Butterworth bandstop filter for the frequencies between 48 and 52 Hz (to remove the Swiss EM caused by power lines). The manual picking of P and S waves proved to be challenging because of the 250Hz sampling rate which proved to not be adequately high to differentiate the arrival time differences of certain signals between seismometers. The chosen time periods were either dates following the unloading of Raeterichsboden Lake, or two days before and after PH measurements peaks, or within the hours of Czech stress peaks. These dates were expected to be the most probable for a microseismicity reaction to a stress change.

In **Chapter 5** the literature review researched for localisation and source parameters determination of detected microseismicity and the methodology that was followed were presented.

In particular, taking into account previous seismic explorations in the GTS area and the velocity profiles researchers used in these occasions, I chose a homogeneous 1D model, performed some sensitivity analysis using earthquakes reported by SED and recorded by my sensors, and resulted in accepting $v_p = 5.4km/s$ and $v_s = 2.9km/s$ as the velocity values to be used for the localisation process.

I then used TOPINV algorithm to locate the manual re-picks of the onset phases of my data. The TOPINV grid-search algorithm was used to locate events for both P- and S- wave times and only for the detected events for which at least 5 out of the 8 seismometers of the two surface arrays were triggered. In total, out of 168 events that their phases were re-picked, 70 microseisms were located using the TOPINV algorithm (successfully converged). This small percentage of convergence is attributed to the dimensions of the grid used for the analysis, i.e. the sources of the re-picked events would have been outside of the limits of the grid and therefore the scope of my research (not RIS). The location estimates had an average fit error, i.e. the sum of the absolute residuals, of less than 143 m.

The source parameters of the located events were calculated assuming the Brune source model (pure double-couple mechanisms). As an overview, the hypocentral distances sampled by this set of events span a range from ∼3.31 to ∼34.85 km (considering both arrays). Inferred corner frequencies exhibit a high degree of scatter as they fall within the range $8.78 \leq f_c \leq 109.58$ Hz with corresponding low frequencies of 1.13E-8 and 1.35E-9Hz. Calculated seismic moments and moment magnitudes fall within the range of $7.81E+8 \leq M_0 \leq 4.11E+12$ and $-0.23 \leq M_w \leq 2.41$ respectively; this range reflects a sampling bias towards larger magnitudes, due to the focus of detection for this experiment. Fault lengths range from 25.45m to 469.35m, source radii range from 12.73m to 234.68m while the areas of the patches that slipped vary from 177,721 to 514.43m^2.

Overall, the seismicity appears to respond to the variations in the volume of the reservoir with a small time delay. This observation agrees with Talwani (1997), who states that in most cases seismic activity is induced by positive lake-level changes, and with Gupta et al. (1972b), who observed that in different case studies seismicity increases significantly within a short time after water loading. The detected events are characterized by a pattern different (through visual observation of their waveforms) from the known tectonic seismic events in the region and are located at distances less than 10km away from the reservoir. These observations favour a hypothesis for non-tectonic origin; hence the located events suggest that they can be reservoir induced. The fact that the detected events are not reported within the Swiss Seismic Database (see also seismic catalogue within Appendix B) strengthens this theory with reference to RIS.

6.2 Safety case for a radioactive waste disposal facility

In terms of the safety case consideration of a deep geologic disposal facility primary implications can be summarised as follows:

- Overall, glacial loading, through its reservoir analogue was found to induce seismicity, albeit not all of the seismic signals were possible to locate or characterise. Research showed that whilst reservoir induced seismicity occurred due to load changes, the magnitudes of events were very small ($-0.23 \leq M_w \leq 2.41$) and hence unlikely to have a significant effect on the groundwater velocities and radionuclide transport calculations, as part of a safety case for a GDF. Further evidence to support this conclusion is that despite the occurrence of these frequent small microseismic events, neither pressure changes nor changes in major/minor ion chemistry have been observed in the GTS boreholes during, in-between or after the reservoir drainage or refilling periods (Stillings, 2020). Indeed, it may be the case that whilst the reservoir unloading increased the frequency of occurrence of microseismic events for a short time period in the region surrounding the reservoir, these events would have naturally occurred anyway and that reservoir unloading simply focussed them into a shorter time window. To test this hypothesis, however, a longer background microseismic monitoring period is required.

- The research has shown that relatively small changes in load (0.74MPa) and in groundwater water pressure can trigger frequent small microseismic events. Consequently, it is possible that GDF construction may have a similar effect (since it will change water flows and the rock mass load in some locations). The research here demonstrates that it is possible to identify and locate these events, even within quite seismically noisy environments. Hence, it would be advantageous to implement microseismic monitoring during both site investigation (in order to obtain baseline event data for low magnitude events) and subsequently during construction. Ideally, monitoring would be deployed using a dense network of surface and borehole-based seismometers, which would allow events to be accurately located and potentially, the orientations and extent of fractures that host such events to be determined (Pytharouli et al. 2011; Urbancic and Mountjoy (2011)). Microseismic analysis could then still be applied to understand how microseismicity evolves after and during the GDF development. Use of shadow space modelling would inform on network configuration and optimal network geometry.

6.3 Novelty / key contributions of research

The key contributions of my research as these were discussed within the thesis are:

- A small-scale test site of a reservoir is chosen as an analogue for exploring the seismic hazard in radioactive deep geologic disposal facilities due to glacial retreats; as described in **Chapter 2**.
- Although sensor deployment design and sensor data cleaning with noise characterisation for micro-seismicity monitoring is not an unmapped area, in this PhD this is explored in detail and for an extended period of several years; as described in **Chapter 3**.

- A new algorithm (NpD) for detecting potential seismic signals under not well-constrained conditions and without requirement of a priori knowledge about the expected signal frequencies and amplitudes was proposed in **Chapter 4**.
- The NpD algorithm and the data set are open source.
- Although there are various onset time picking algorithms suggested by researchers (e.g. Irving‘s method, Akazawa’s method, JER-AIC, e.t.c. (Akram and Eaton, 2016)) the novelty of this research lies by using, comparing suitability and adjustment of a number of well-known approaches from the literature in conjunction for real field data and for microseismic events detection purposes (**Chapter 4**).
- A seismic catalogue related to the dynamic response of the rock mass to reservoir drainage was presented in **Chapter 5 (Appendix B)**.

6.4 Further work

As already discussed, although considerable effort was put to avoid having multiple NpD peaks for the same signal, this was not possible in all occasions (e.g. in the cases of landslides). This could be a topic to be investigated in the future in order to minimise detections of non microseismic events; a potential way to deal with this is to explore classification algorithms (e.g. machine learning, neural networks, etc.).

Also, the NpD algorithm is not an onset time determination algorithm with its accuracy depending on the size of user selected windows. Potential further work ideas on this include: (i) combination of NpD algorithm with other automated picking algorithms, such as autoregressive techniques (Kong, 1997; Leonard and Kennett, 1999; Oye and Roth, 2003), (ii) check the performance of the algorithm by incorporating wavelet decomposition instead of Fourier transforms, and (iii) record timeseries with a high sampling rate which would allow for smaller windows hence greater accuracy in the detections and downsample for data storing reasons.

AIC-wav is an algorithm developed for the automated accurate picking of the P and S arrival times and was developed in R studio programming language; a hybrid approach, with a simple formulation and parameter settings, that combines single and multilevel based existing methodologies. The *AIC-wav* worked exceptionally well in cases that the signal was visually distinct to noise. However, in cases when the signal has lost the P wave, the algorithm misidentifies the S-wave as a P-wave arrival. This misidentification of the first arrival has also been observed by Ge (2005). Overall, *AIC-wav* workflow allows more confidence in the automatically picked P-waves rather than the S-waves and can constitute further work to resolve the S-wave misidentifications. Provided this is resolved, *AIC-wav* algorithm will be providing automated onset times as efficiently as an observer's manual picks within a one data per second range accuracy.

Both algorithms' effectiveness when used with other than seismic signals can also explored as future work. It is believed that they are going to be performing well in other signal processing science fields, for instance to capture ECG signals, or for monitoring of wind turbines.

Chapter 7 Conclusions

Resulting in an effective solution for the geologic radioactive waste disposal with no environmental risk is a paramount issue for the future reliance on nuclear power. Although geological disposal is considered as the best option, there are still elements of risk to be addressed, such as glacial retreat which could impact the safety performance of the GDF. Damage kinematics has been reported to be influenced by the discontinuity geometry and the relative glacier position.

Glacial stress changes are expected to occur whilst the waste is still radioactive. Due to the long-lived nature of radionuclides there is a need to consider the effect these stresses would have, being exerted on the repository and the surrounding rocks. For this thesis I explored the response of granitic rock, as a potential host rock to a radioactive disposal facility, to glacial retreat. According to Zhang et al. (2019) a consensus on how fluids trigger earthquakes has not been reached and although numerous papers explain the role of fluids, the long-term dynamic mechanism that fluid plays in stress transformation and strength changes of the fault zone remain inexplicable, mainly because actual fault zones are complex, and the fluid - stress mechanism needs to be investigated for different fault zones.

My findings from this PhD in terms of the implications for a GDF are summarised as follows:

<u>A reservoir is a suitable small scale field analogue to represent glacier retreat</u>: To investigate the response of granitic rock to glacial retreat I used the draining and refilling of Raeterichsboden lake as a small scale analogue to 2 glacial cycles during 2 consecutive years. There was an opportunity to monitor the lake level changes of this particular reservoir and their effect on the surrounding granitic rocks because of the convenient positioning of the GTS

which allowed for the deployment of monitoring equipment. The analogy glacier / reservoir was found appropriate for researching the safety considerations of a GDF. For the detection of such events the NpD algorithm can be used, modified with a classification algorithm (e.g. machine learning or neural networks).

Further experiments can be undertaken to establish the long-term microseismic effects that a couple or annually repeated load stress / relief cycles may cause to the GDF's barriers / surroundings since the two cycles of unloading / reloading may have repercussions in the seismicity in the following years that the experiments and the monitoring took place.

Microseismic monitoring is an efficient tool to monitor the effects of glacial retreat to the host rock: Microseismic monitoring was chosen as the optimum tool for my research since it is proven to detect and locate extremely weak geological processes. The seismic sensors are easily deployed (and retrieved) and record continuous data remotely for an extended time period. Two microseismic arrays and a borehole sensor were deployed for my case study to cover a triangular coverage area of approximately 340m^2. Specific consideration should be given prior to the deployment for the network geometry; the tunnels shadowing and expected magnitude threshold should be estimated to allow for the optimal network configuration.

I would also recommend that the sensors of each array are not placed in one straight line (should the site allow) but in star shape or grid designs. Also, a borehole array would be preferred to a single borehole sensor to complete the network design. And the sensors should surround the area that is to be monitored (should the site allow). Borehole sensors / arrays suffer from lateral positioning accuracy while array sensors from low SNR because of the environmental noise and the signal attenuation (mostly of the S-wave). Conversely, surface arrays demonstrate high lateral positioning accuracy and borehole sensors / arrays higher depth resolution accuracy. Hence a combination of the two types of arrays would complement the advantages and minimise the disadvantages of each.

Another note that I would like to point out is in regards to the sampling frequency. Although consideration was given in this study to accurately account for the frequency content of the signals that I was expecting to record and eliminated the potential for alias frequencies by selecting a sampling rate at least twice the Nyquist frequency (sampling theorem), this seemed to be rather a small sampling rate for the needs of my case study. In fact I should have oversampled at a rate of 8 or even 10 times the Nyquist frequency to allow for extraction of information from the reconstructed waveform (such as P- and S- wave arrival times). Oversampling is capable of improving resolution and SNR however increases dramatically the data storage requirements and the algorithms analysis speed (more time to process).

Assessment of the developed methodology for the analysis of the acquired data: MATLAB and R studio codes were written for the detection, onset time determination and location of the events. The newly proposed detection algorithm NpD was also published in Bulletin of Seismological Society of America. NpD algorithm can successfully detect even low SNR events but can further be improved by applying a classification algorithm, fiddling with the sampling rate or using wavelet decomposition instead of Fourier transforms.

The onset time determination *AIC-wav* needs further refinement to be able to capture both phases of signals in an accurate manner; currently, it allows more confidence in the automatically picked P-waves rather than the S-waves. This occurs because of the misidentification of the S-wave as a P-wave arrival in cases that the P-wave is lost in the coda, a misidentification also reported by Ge (2005) to occur in 41% of the total picks. Multiple workflows were tried out to be able to automatically accurately pick P- and S-waves but that ended up fruitless and led to the location of manually selected events.

Owing to the plethora of detection / onset time determination / location algorithms, limitations of time and equipment, it has not, by any means, been possible to test all aspects of the methodology developed by researchers whose work was explored during my literature review and I cannot deny that other methodologies, for instance neural network techniques, might have been more effective in identifying accurately signals.

Overall, I believe that I have successfully addressed the objectives of my thesis and have added one more pebble of knowledge in this fluid-stress complex dynamic relationship research. To summarise, the key contributions of the thesis as these were discussed and proven within my thesis are:

- A small-scale test site of a reservoir is chosen as an analogue for exploring the seismic hazard in radioactive deep geologic disposal facilities due to glacial retreats; as described in Chapter 2.
- Although sensor deployment design and sensor data cleaning with noise characterisation for micro-seismicity monitoring is not an unmapped area, in this PhD this is explored in detail and for an extended period of several years; as described in Chapter 3.
- A new algorithm (NpD) for detecting potential seismic signals under not well-constrained conditions and without requirement of a priori knowledge about the expected signal frequencies and amplitudes was proposed in Chapter 4.
- The NpD algorithm and the data set are open source.
- Although there are various onset time picking algorithms suggested by researchers (e.g. Irving‘s method, Akazawa’s method, JER-AIC, e.t.c. (Akram and Eaton, 2016)) the novelty of this research lies by using, comparing suitability and adjustment of a number of well-known approaches from the literature in conjunction for real field data and for microseismic events detection purposes (Chapter 4).
- A seismic catalogue related to the dynamic response of the rock mass to reservoir drainage was presented in Chapter 5.

References

Ahmed, A., Sharma, A., 2007. Wavelet based automatic phase picking algorithm for 3-component broadband seismological data 9, 15–24.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Akram, J., Eaton, D., St. Onge, A., 2013. Automatic event-detection and time-picking algorithms for downhole microseismic data processing. https://doi.org/10.3997/2214-4609.20142372

Akram, J., Eaton, D.W., 2016. A review and appraisal of arrival-time picking methods for downhole microseismic data. Geophysics 81, KS71–KS91. https://doi.org/10.1190/geo2014-0500.1

Alt, R.C., Zoback, M.D., 2017. In Situ Stress and Active Faulting in Oklahoma. Bull. Seismol. Soc. Am. 107, 216–228. https://doi.org/10.1785/0120160156

Anant, K.S., Dowla, F.U., 1997. Wavelet transform methods for phase identification in three-component seismograms. Bull. Seismol. Soc. Am. 87, 1598–1612.

Angus, D.A., Aljaafari, A., Usher, P., Verdon, J.P., 2014. Seismic waveforms and velocity model heterogeneity: Towards a full-waveform microseismic location algorithm. J. Appl. Geophys. 111, 228–233. https://doi.org/10.1016/j.jappgeo.2014.10.013

Bagaini, C., 2005. Performance of time-delay estimators. GEOPHYSICS 70, V109–V120. https://doi.org/10.1190/1.1996909

Baig, A., Urbancic, T., 2010. Magnitude Determination, Event Detectability, and Assessing the Effectiveness of Microseismic Monitoring Programs in Petroleum Applications.

Barnett, V., Lewis, T., 1994. Outliers in statistical data. Wiley.

Baumberger, R., 2015. Quantification of lineaments: Link between internal 3D structure and surface evolution of the Hasli valley (Aar massif, Central Alps, Switzerland).

Bense, V.F., Person, M.A., 2008. Transient hydrodynamics within intercratonic sedimentary basins during glacial cycles. J. Geophys. Res. 113, F04005. https://doi.org/10.1029/2007JF000969

Blias, E., Grechka, V., 2013. Analytic solutions to the joint estimation of microseismic event locations and effective velocity model. GEOPHYSICS 78, KS51–KS61. https://doi.org/10.1190/geo2012-0517.1

Blumling, P., Cosma, C., Korn, M., Gelbke, C., Cassell, B., Montan Technologie, D., 1992. Technical report 90-07 Geophysical methods for the detection of discontinuities ahead of a tunnel face.

Boore, D.M., Boatwright, J., 1984. Average body-wave radiation coefficients. Bull. Seismol. Soc. Am. 74, 1615–

1621.

Bormann, P. (Ed.), 2012. New Manual of Seismological Observatory Practice (NMSOP-2). GFZ German Research Centre for Geosciences, Potsdam. https://doi.org/10.2312/GFZ.NMSOP-2

Boulton, G.S., Caban, P., Hulton, N., Karnbranslehantering, S., 1999. R-99-73 Simulations of the Scandinavian ice sheet and its subsurface conditions.

Bourbie, T., Coussy, O., Zinszner, B., Junger, M.C., 1992. Acoustics of Porous Media. J. Acoust. Soc. Am. 91, 3080–3080. https://doi.org/10.1121/1.402899

Brune, J.N., 1970. Tectonic stress and the spectra of seismic shear waves from earthquakes. J. Geophys. Res. 75, 4997–5009. https://doi.org/10.1029/JB075i026p04997

Campillo, M., Plantet, J.-L., Bouchon, M., 1985. Frequency dependent attenuation in the crust beneath central France from Lg waves: data analysis and numerical modeling. Bull. Seismol. Soc. Am. 75, 1395–1411.

Chambers, K., Kendall, J.-M., Brandsberg-Dahl, S., Rueda, J., 2010. Testing the ability of surface arrays to monitor microseismic activity. Geophys. Prospect. 58, 821–830. https://doi.org/10.1111/j.1365-2478.2010.00893.x

Chan, T., Christiansson, R., Boulton, G.S., Ericsson, L.O., Hartikainen, J., Jensen, M.R., Mas Ivars, D., Stanchell, F.W., Vistrand, P., Wallroth, T., 2005. DECOVALEX III BMT3/BENCHPAR WP4: The thermo-hydro-mechanical responses to a glacial cycle and their potential implications for deep geological disposal of nuclear fuel waste in a fractured crystalline rock mass. Int. J. Rock Mech. Min. Sci. 42, 805–827. https://doi.org/10.1016/J.IJRMMS.2005.03.017

Chapman, N.A., Mckinley, I.G., Smellie, J.A.T., 1984. EIR-Bericht Nr. 545 The Potential of Natural Analogues in Assessing Systems for Deep Disposal of High-Level Radioactive Waste.

Chen, Q., Liu, X., Li, J., Wang, M., Liu, P., 2016. Two-Dimensional Far Field Source Locating Method with Nonprior Velocity. Math. Probl. Eng. 2016, 1–14. https://doi.org/10.1155/2016/8540303

Corbet, T.F., Bethke, C.M., 1992. Disequilibrium fluid pressures and groundwater flow in the western Canada sedimentary basin. J. Geophys. Res. 97, 7203. https://doi.org/10.1029/91JB02993

Costain, J.K., Bollinger, G.A., Speer, J.A., 1987. Hydroselsmicity: A Hypothesis for The Role of Water in the Generation of Intraplate Seismicity. Seismol. Res. Lett. 58, 41–64. https://doi.org/10.1785/gssrl.58.3.41

D'Auria, L., Marotta, E., Martini, M., Ricciolino, P., 2006. Seismic and acoustic detection of a bolide airburst in the Gulf of Naples (southern Italy). J. Geophys. Res 111, 10307. https://doi.org/10.1029/2005JB004254

Dai, H., MacBeth, C., 1995. Automatic picking of seismic arrivals in local earthquake data using an artificial neural network. Geophys. J. Int. 120, 758–774. https://doi.org/10.1111/j.1365-246X.1995.tb01851.x

Das, I., Zoback, M.D., 2011. Long-period, long-duration seismic events during hydraulic fracture stimulation of a shale gas reservoir. Lead. Edge 30, 778–786. https://doi.org/10.1190/1.3609093

De Meersman, K., Kendall, J.-M., Van Der Baan, M., 2009. Case History The 1998 Valhall microseismic data set: An integrated study of relocated sources, seismic multiplets, and S-wave splitting. https://doi.org/10.1190/1.3205028

Eaton, D.W., 2011. Q Determination, Corner Frequency and Spectral Characteristics of Microseismicity Induced by Hydraulic Fracturing, in: SEG Technical Program Expanded Abstracts 2011. Society of Exploration Geophysicists, pp. 1555–1559.

Edwards, B., Allmann, B., Fäh, D., Clinton, J., 2010. Automatic computation of moment magnitudes for small earthquakes and the scaling of local to moment magnitude. Geophys. J. Int 183, 407–420.

https://doi.org/10.1111/j.1365-246X.2010.04743.x

Eisner, L., Hulsey, B.J., Duncan, P., Jurick, D., Werner, H., Keller, W., 2010. Comparison of surface and borehole locations of induced seismicity. Geophys. Prospect. 58, 809–820. https://doi.org/10.1111/j.1365-2478.2010.00867.x

Fäh, D., Giardini, D., Kästli, P., Deichmann, N., Gisler, M., Schwarz-Zanetti, G., Alvarez-Rubio, S., Sellami, S., Edwards, B., Allmann, B., 2011. ECOS-09 earthquake catalogue of Switzerland release 2011 report and database. Public catalogue, 17. 4. 2011. Swiss Seismological Service ETH Zurich, Risk.

Freiberger, W.F., 1963. An approximate method in signal detection. Q. Appl. Math. 20, 373–378. https://doi.org/10.1090/qam/139498

Ge, M., 2005. Efficient mine microseismic monitoring. https://doi.org/10.1016/j.coal.2005.03.004

Ge, M., 2003a. Analysis of source location algorithms Part I: Overview and non-iterative methods. J. Acoust. Emiss. 21, 14–28.

Ge, M., 2003b. Analysis of source location algorithms Part II: Iterative methods, J. Acoustic Emission.

Geiger, L., 1910. Herdbestimmung bei Erdbeben aus den Ankunftszeiten. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen Math. 331–349.

Geller, R.J., Mueller, C.S., 1980. Four similar earthquakes in central California. Geophys. Res. Lett. 7, 821–824. https://doi.org/10.1029/GL007i010p00821

Gibbons, S.J., Ringdal, F., 2006. The detection of low magnitude seismic events using array-based waveform correlation. Geophys. J. Int. 165, 149–166. https://doi.org/10.1111/j.1365-246X.2006.02865.x

Gibowicz, S.J., Kijko, A., 1994. An Introduction to Mining Geophysics.

Goertz-Allmann, B.P., Kühn, D., Oye, V., Bohloli, B., Aker, E., 2014. Combining microseismic and geomechanical observations to interpret storage integrity at the In Salah CCS site. Geophys. J. Int. 198, 447–461. https://doi.org/10.1093/gji/ggu010

Goforth, T., Herrin, E., 1981. An automatic seismic signal detection algorithm based on the Walsh Transform, Bunetm of the Seismological Socmty of America.

Got, J.-L., Mourot, P., Grangeon, J., 2010. Pre-failure behaviour of an unstable limestone cliff from displacement and seismic data. Nat. Hazards Earth Syst. Sci. 10, 819–829. https://doi.org/10.5194/nhess-10-819-2010

Grämiger, L.M., Moore, J.R., Gischig, V.S., Ivy-Ochs, S., Loew, S., 2017. Beyond debuttressing: Mechanics of paraglacial rock slope damage during repeat glacial cycles. J. Geophys. Res. Earth Surf. 122, 1004–1036. https://doi.org/10.1002/2016JF003967

Gupta, H.K., 2002. A review of recent studies of triggered earthquakes by artificial water reservoirs with special emphasis on earthquakes in Koyna, India. Earth-Science Rev. 58, 279–310. https://doi.org/10.1016/S0012-8252(02)00063-6

Gupta, H.K., 1992. Reservoir-induced earthquakes. Elsevier.

Gupta, H.K., Rastogi, B.K., Narain, H., 1972a. Some discriminatory characteristics of earthquakes near the Kariba, Kremasta, and Koyna artificial lakes. Bull. Seismol. Soc. Am. 62, 4–7.

Gupta, H.K., Rastogi, B.K., Narain, H., 1972b. Common features of the reservoir-associated seismic activities. Bull. Seismol. Soc. Am. 62, 481–492.

References

Hafez, A.G., Khan, M.T.A., Kohda, T., 2010. Clear P-wave arrival of weak events and automatic onset determination using wavelet filter banks. Digit. Signal Process. 20, 715–723. https://doi.org/10.1016/J.DSP.2009.10.002

Han, L., Wong, J., Bancroft, J., 2009. Time picking and random noise reduction on microseismic data 21, 1–13.

Haney, M.M., Power, J., West, M., Michaels, P., 2012. Causal Instrument Corrections for Short-Period and Broadband Seismometers. Seismol. Res. Lett. 83, 834–845. https://doi.org/10.1785/0220120031

Haskell, N.A., 1964. Total energy and energy spectral density of elastic wave radiation from propagating faults. Bull. Seismol. Soc. Am. 54, 1811–1841.

Havlova, V., Martin, A.J., Landa, J., Sus, F., Eikenberg, J., Siitari-Kauppi, M., Soler, J., Miksova, J., 2013. Long-term diffusion experiment LTD phase 1: evaluation of results and modelling., in: Migration. Brighton.

Havskov, J., Ottemoller, L., 2010. Routine Data Processing in Earthquake Seismology. Springer. https://doi.org/10.1007/978-90-481-8697-6_1

Hedin, A., Kautsky, U., 2000. SR 97: POST-CLOSURE SAFETY FOR A KBS3 TYPE DEEP REPOSITORY FOR SPENT NUCLEAR FUEL.

Helmstetter, A., Garambois, S., 2010. Seismic monitoring of Séchilienne rockslide (French Alps): Analysis of seismic signals and their correlation with rainfalls. J. Geophys. Res. 115, F03016. https://doi.org/10.1029/2009JF001532

Huang, C.-J., Yin, H.-Y., Chen, C.-Y., Yeh, C.-H., Wang, C.-L., 2007. Ground vibrations produced by rock motions and debris flows. J. Geophys. Res. 112, F02014. https://doi.org/10.1029/2005JF000437

Husen, S., Bachmann, C., Giardini, D., 2007. Locally triggered seismicity in the central Swiss Alps following the large rainfall event of August 2005. Geophys. J. Int. 171, 1126–1134. https://doi.org/10.1111/j.1365-246X.2007.03561.x

Husen, S., Wiemer, S., Smith, R.B., 2004. Remotely Triggered Seismicity in the Yellowstone National Park Region by the 2002 Mw 7.9 Denali Fault Earthquake, Alaska. Bull. Seismol. Soc. Am. 94, S317–S331. https://doi.org/10.1785/0120040617

IAEA, 2001. Monitoring of geological repositories for high level radioactive waste.

Inglada, V., 1928. The calculation of the stove coordinates of a Nahbebens. Gerlands Beitrage Zur Geophys. 19, 73–98.

International Atomic Energy Agency, 2018. Storage and Disposal of Radioactive Waste [WWW Document]. URL https://www.world-nuclear.org/information-library/nuclear-fuel-cycle/nuclear-waste/storage-and-disposal-of-radioactive-waste.aspx (accessed 3.2.19).

Irving, J.D., Knoll, M.D., Knight, R.J., 2007. Improving crosshole radar velocity tomograms: A new approach to incorporating high-angle traveltime data. GEOPHYSICS 72, J31–J41. https://doi.org/10.1190/1.2742813

Itasca Consulting Group, I., 2015. InSite | Itasca Consulting Group [WWW Document]. URL http://www.itascacg.com/software/insite-0 (accessed 11.25.15).

Jiménez, M.-J., García-Fernández, M., 2000. Occurrence of shallow earthquakes following periods of intense rainfall in Tenerife, Canary Islands. J. Volcanol. Geotherm. Res. 103, 463–468. https://doi.org/10.1016/S0377-0273(00)00237-7

Johnson, C.W., Fu, Y., Bürgmann, R., 2017. Seasonal water storage, stress modulation, and California seismicity. Science (80-.). 356, 1161–1164. https://doi.org/10.1126/science.aak9547

Joswig, M., 1990. Pattern recognition for earthquake detection, Bulletin of the Seismological Society of America.

Joswig, M., Camelbeeck, T., Flick, T., Ducarme, B., 1992. System architecture of seismic networks and its implications to network automatization. Cah. Cent. Eur. Geodyn. Seism 5, 75–84.

Keilis-Borok, V.I., Monin, A.S., 1959. Magnetoelastic Waves and the Boundary of the Earth's Core. Izv. Akad. Nauk SSSR, Ser. Geofiz 11, 1529–1541.

Keusen, H.R., Ganguin, J., Schuler, P., Buletti, M., 1989. Felslabor Grimsel: Geologie. Nagra technical report NTB 87-14. Baden, Switzerland.

Khandelwal, M., Singh, T.N., 2006. Prediction of blast induced ground vibrations and frequency in opencast mine: A neural network approach. J. Sound Vib. 289, 711–725. https://doi.org/10.1016/j.jsv.2005.02.044

Kinali, M., Pytharouli, S., Lunn, R.J., Shipton, Z.K., Stillings, M., Lord, R., Thompson, S., 2018. Detection of weak seismic signals in noisy environments from unfiltered, continuous passive seismic recordings. Bull. Seismol. Soc. Am. 108, 2993–3004. https://doi.org/10.1785/0120170358

Kong, X., 1997. Forward and backward autoregressive modeling of EEG, in: Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. "Magnificent Milestones and Emerging Opportunities in Medical Engineering" (Cat. No.97CH36136). IEEE, pp. 1215–1217. https://doi.org/10.1109/IEMBS.1997.756582

Kraft, T., Wassermann, J., Schmedes, E., Igel, H., 2006. Meteorological triggering of earthquake swarms at Mt. Hochstaufen, SE-Germany. Tectonophysics 424, 245–258. https://doi.org/10.1016/J.TECTO.2006.03.044

Küperkoch, L., Meier, T., Lee, J., Friederich, W., Working Group, E., 2010. Automated determination of P-phase arrival times at regional and local distances using higher order statistics. Geophys. J. Int. 181, 1159–1170. https://doi.org/10.1111/j.1365-246X.2010.04570.x

LASMO [WWW Document], n.d. URL http://www.grimsel.com/gts-phase-vi/lasmo/lasmo-introduction (accessed 3.3.19).

LASMO Team, 2016. Project Report NPB 16-04. GTS Phase VI – LASMO Project: Annual Report 2015.

Lee, W.H.K., Stewart, S.W., 1981. Principles and applications of microearthquake networks. Academic Press.

Lei, X., Ma, S., Chen, W., Pang, C., Zeng, J., Jiang, B., 2013. A detailed view of the injection-induced seismicity in a natural gas reservoir in Zigong, southwestern Sichuan Basin, China. J. Geophys. Res. Solid Earth 118, 4296–4311. https://doi.org/10.1002/jgrb.50310

Lei, X., Yu, G., Ma, S., Wen, X., Wang, Q., 2008. Earthquakes induced by water injection at ~3 km depth within the Rongchang gas field, Chongqing, China. J. Geophys. Res. 113, B10310. https://doi.org/10.1029/2008JB005604

Leighton, F., Duvall, W.I., 1972. Least squares method for improving rock noise source location techniques (No. BM-RI-7626). Washington, DC (USA).

Leonard, M., 2000. Comparison of manual and automatic onset time picking. Bull. Seismol. Soc. Am. 90, 1384–1390. https://doi.org/10.1785/0120000026

Leonard, M., Kennett, B.L.N., 1999. Multi-component autoregressive techniques for the analysis of seismograms. Phys. Earth Planet. Inter. 113, 247–263. https://doi.org/10.1016/S0031-9201(99)00054-0

Levy, C., Jongmans, D., Baillet, L., 2011. Analysis of seismic signals recorded on a prone-to-fall rock column (Vercors massif, French Alps). Geophys. J. Int. 186, 296–310. https://doi.org/10.1111/j.1365-

246X.2011.05046.x

Li, X.-B., Wang, Z.-W., Dong, L.-J., 2015. Locating single-point sources from arrival times containing large picking errors (LPEs): the virtual field optimization method (VFOM) OPEN. https://doi.org/10.1038/srep19205

Lomax, A., Virieux, J., Volant, P., Berge-Thierry, C., 2000. Probabilistic earthquake location in 3D and layered models., in: In Advances in Seismic Event Location. pp. 101–134.

Lorraine, W., Rowe, C.A., Horner, R.B., 1997. Periodic seismicity near Mt. Ogden on the Alaska-British Columbia border: A case for hydrologically triggered earthquakes? Bull. Seismol. Soc. Am. 87, 1473–1483.

Maeda N., 1985. A method for reading and checking phase times in autoprocessing system of seismic wave data. Zisin 38, 365–379.

Majer, E.L., Myer, L.R., Peterson, J.E., K., K., Long, J.C.S., Martel, S.J., Bliimling, P., Vomvoris, S., 1990. Joint Seismic, Hydrogeological, and Geomechanical Investigations of a Fracture Zone in the Publication Date.

Massey, F.J., 1951. The Kolmogorov-Smirnov test for goodness of fit. J. Am. Stat. Assoc. 46, 68–78. https://doi.org/10.1080/01621459.1951.10500769

Maxwell, S., 2011. Microseismic hydraulic fracture imaging: The path toward optimizing shale gas production. Lead. Edge 30, 340–346. https://doi.org/10.1190/1.3567266

Métivier, L., Brossier, R., Operto, S., Virieux, J., 2014. Applying Gauss-Newton and Exact Newton method to Full Waveform Inversion Second-order adjoint state methods for Full Waveform Inversion.

Mooney, M., Walter, B., Steele, J., Cano, D., n.d. Influence of Geological Conditions on Measured TBM Vibration Frequency.

Moran, S.C., Matoza, R.S., Garcés, M.A., Hedlin, M.A.H., Bowers, D., Scott, W.E., Sherrod, D.R., Vallance, J.W., 2008. Seismic and acoustic recordings of an unusually large rockfall at Mount St. Helens, Washington. Geophys. Res. Lett. 35, L19302. https://doi.org/10.1029/2008GL035176

Moriya, H., 2008. Precise arrival time detection of polarized seismic waves using the spectral matrix. Geophys. Prospect. 56, 667–676. https://doi.org/10.1111/j.1365-2478.2008.00713.x

Mousavi, S.M., Stephen, P., Langston, C., Samei, B., 2016. Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression. Geophys. J. Int.

NAGRA, BGR, GSF, 1985. Technical Report 85-46 Grimsel Test Site: Overview and Test Programs.

Neuzil, C.E., 2012. Hydromechanical effects of continental glaciation on groundwater systems. Geofluids 12, 22–37. https://doi.org/10.1111/j.1468-8123.2011.00347.x

O'Brien, G.S., Lokmer, I., De Barros, L., Bean, C.J., Saccorotti, G., Metaxian, J.-P., Patane, D., 2011. Time reverse location of seismic long-period events recorded on Mt Etna. Geophys. J. Int. 184, 452–462. https://doi.org/10.1111/j.1365-246X.2010.04851.x

Obert, L., Duvall, W.I., 1967. Rock mechanics and the design of structures in rock. Wiley, New York.

Oye, V., Bungum, H., Roth, M., 2005. Source Parameters and Scaling Relations for Mining-Related Seismicity within the Pyhäsalmi Ore Mine, Finland. Bull. Seismol. Soc. Am. 95, 1011–1026. https://doi.org/10.1785/0120040170

Oye, V., Roth, M., 2003. Automated seismic event location for hydrocarbon reservoirs. Comput. Geosci. 29, 851–863. https://doi.org/10.1016/S0098-3004(03)00088-8

Pavlis, G.L., 1992. Appraising relative earthquake location errors. Bull. Seismol. Soc. Am. 82, 836–859.

Phillips, W.S., 2000. Precise Microearthquake Locations and Fluid Flow in the Geothermal Reservoir at Soultz-sous-Forêts, France, Bulletin of the Seismological Society of America.

Prejean, S.G., Hill, D.P., Brodsky, E.E., Hough, S.E., Johnston, M.J.S., Malone, S.D., Oppenheimer, D.H., Pitt, A.M., Richards-Dinger, K.B., 2004. Remotely Triggered Seismicity on the United States West Coast following the Mw 7.9 Denali Fault Earthquake. Bull. Seismol. Soc. Am. 94, S348–S359. https://doi.org/10.1785/0120040610

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical recipes : the art of scientific computing. Cambridge University Press.

Prugger, A.F., Gendzwill, D.J., 1988. Microearthquake location: a nonlinear approach that makes use of a simplex stepping procedure. Bull. Seismol. Soc. Am. 78, 799–815.

public.geo.admin.ch, swisstopo, n.d. Maps of Switzerland - Swiss Confederation - map.geo.admin.ch [WWW Document]. URL https://map.geo.admin.ch (accessed 11.26.15).

Pujol, J., 2004. Earthquake Location Tuton'al: Graphical Approach andApproximate Epicentrai Location Techniques. Seismol. Res. Lett. 75, 63–74.

Pujol, J., Smalley, R., 1990. A preliminary earthquake location method based on a hyperbolic approximation to travel times. Bull. Seismol. Soc. Am. 80, 1629–1642.

Pytharouli, S.I., Lunn, R.J., Shipton, Z.K., Kirkpatrick, J.D., do Nascimento, A.F., 2011. Microseismicity illuminates open fractures in the shallow crust. Geophys. Res. Lett. 38, n/a-n/a. https://doi.org/10.1029/2010GL045875

Razali, N.M., Wah, Y.B., 2011. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. J. Stat. Model. Anal. 2, 13–14.

Ripper, B., 2017. How much radioactive waste is there? [WWW Document]. URL https://nda.blog.gov.uk/2017/04/03/how-much-radioactive-waste-is-there/ (accessed 3.2.19).

Roeloffs, E.A., 1988. Hydrologic precursors to earthquakes: A review. Pure Appl. Geophys. PAGEOPH 126, 177–209. https://doi.org/10.1007/BF00878996

Roth, P., Pavoni, N., Deichmann, N., 1992. Seismotectonics of the eastern Swiss Alps and evidence for precipitation-induced variations of seismic activity. Tectonophysics 207, 183–197. https://doi.org/10.1016/0040-1951(92)90477-N

Saar, M.O., Manga, M., 2003. Seismicity induced by seasonal groundwater recharge at Mt. Hood, Oregon. Earth Planet. Sci. Lett. 214, 605–618. https://doi.org/10.1016/S0012-821X(03)00418-7

Sabbione, J.I., Velis, D., 2010. Automatic first-breaks picking: New strategies and algorithms. GEOPHYSICS 75, V67–V76. https://doi.org/10.1190/1.3463703

Saltogianni, V., Stiros, S., 2015. A Two-Fault Model of the 2003 Leucas (Aegean Arc) Earthquake Based on Topological Inversion of GPS Data. Bull. Seismol. Soc. Am. 105, 2510–2520. https://doi.org/10.1785/0120140355

Saltogianni, V., Stiros, S.C., 2013. Topological inversion in geodesy-based, non-linear problems in geophysics. Comput. Geosci. 52, 379–388. https://doi.org/10.1016/j.cageo.2012.11.010

Saltogianni, V., Stiros, S.C., 2012. Modeling the Mogi magma source centre of the Santorini (Thera) volcano, Aegean Sea, Greece, 1994-1999, based on a numerical-topological approach. Stud. Geophys. Geod 56, 1037–1062. https://doi.org/10.1007/s11200-012-0408-z

Sarout, J., Esteban, L., Josh, M., 2012. Arbeitsbericht NAB 12-47: Laboratory Characterisation of Cores: Petrophysics and rock mechanics / physics.

Sauber, J.M., Molnia, B.F., 2004. Glacier ice mass fluctuations and fault instability in tectonically active Southern Alaska, in: Global and Planetary Change. pp. 279–293. https://doi.org/10.1016/j.gloplacha.2003.11.012

Schneeberger, R.., 2017. Interplay in 3D between faults and water flow paths in crystalline bedrock (Grimsel, Switzerland). Universität Bern.

Schneeberger, R., Berger, A., Herwegh, M., Eugster, A., Kober, F., Spillmann, T., Blechschmidt, I., 2016. GTS Phase VI - LASMO: Geology and structures of the GTS and Grimsel region. Nagra Arbeitsbericht NAB 16-27. Wettingen, Switzerland.

Schneeberger, R., De La Varga, M., Egli, D., Berger, A., Kober, F., Wellmann, F., Herwegh, M., 2017. Methods and uncertainty estimations of 3-D structural modelling in crystalline rocks: a case study. Solid Earth 8, 987–1002. https://doi.org/10.5194/se-8-987-2017

Senfaute, G., Duperret, A., Lawrence, J.A., 2009. Micro-seismic precursory cracks prior to rock-fall on coastal chalk cliffs: a case study at Mesnil-Val, Normandie, NW France. Nat. Hazards Earth Syst. Sci. 9, 1625–1641. https://doi.org/10.5194/nhess-9-1625-2009

Shelke, S., 2017. Vibration Analysis of Pelton Wheel Turbine by Theoretical, Numerical & Experimental Approach. Int. J. Eng. 1396–1400.

Simpson, D.W., 1976. Seismicity changes associated with reservoir loading. Eng. Geol. 10, 123–150. https://doi.org/10.1016/0013-7952(76)90016-8

Simpson, D.W., Leith, W.S., Scholz, C.H., 1988. TWO TYPES OF RESERVOIR-INDUCED SEISMICITY, Bulletin of the Seismological Society of America.

Simpson, D.W., Negmatullaev, S.K., 1981. Induced seismicity at Nurek Reservoir, Tadjikistan, USSR. Bull. Seismol. Soc. Am. 71, 1561–1586.

SKI, 1996. SKI SITE-94. Deep repository performance assessment project. SKI Report 96:36, Swedish Nuclear Power Inspectorate, Stockholm.

Spillmann, T., Maurer, H., Green, A.G., Heincke, B., Willenberg, H., Husen, S., 2007. Microseismic investigation of an unstable mountain slope in the Swiss Alps. J. Geophys. Res. 112, B07301. https://doi.org/10.1029/2006JB004723

Stamps, D.S., Smalley, R.J., 2006. Strings and Things for Locating Earthquakes. Seismol. Res. Lett. 77, 677–683.

Stillings, M., 2020. Exploring subsurface groundwater and geochemical rock interactions during drainage of a surface water reservoir in Switzerland.

Stork, A.L., Verdon, J.P., Kendall, J.-M., 2014. The robustness of seismic moment and magnitudes estimated using spectral analysis. Geophys. Prospect. 62, 862–878. https://doi.org/10.1111/1365-2478.12134

Sundhararajan, S., Pahwa, A., Krishnaswami, P., 1998. A Comparative Analysis of Genetic Algorithms and Directed Grid Search for Parametric Optimization. Eng. Comput. 14, 197–205.

Talbot, C.., 1999. Ice ages and nuclear waste isolation. Eng. Geol. 52, 177–192. https://doi.org/10.1016/S0013-7952(99)00005-8

Talwani, P., 1997. On the Nature of Reservoir-induced Seismicity. pure Appl. Geophys. 150, 473–492.

https://doi.org/10.1007/s000240050089

Tarasov, L.., Peltier, R.W., 1997. Terminating the 100 kyr ice age cycle. J. Geophys. Res. 102, 21.665-21,693.

Thompson, G., Reyes, C., 2018. GISMO - a seismic data analysis toolbox for MATLAB [software package] [WWW Document]. https://doi.org/http://doi.org/10.5281/zenodo.1404723

Thurber, C.H., 1985. Nonlinear earthquake location: Theory and examples. Bull. Seismol. Soc. Am. 75, 779–790.

Torgoev, A., Lamair, L., Torgoev, I., Havenith, H.-B., 2013. A review of recent case studies of landslides investigated in the Tien Shan using microseismic and other geophysical methods, in: Earthquake-Induced Landslides. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 285–294. https://doi.org/10.1007/978-3-642-32238-9_29

Trnkoczy, A., 1999. Understanding and parameter setting of STA/LTA trigger algorithm. https://doi.org/10.2312/GFZ.NMSOP-2_IS_8.1

Tronicke, J., 2007. The influence of figh frequency uncorrelated noise on first-break arrival times and crosshole traveltime tomography. J. Environ. Eng. Geophys. 12, 173–184. https://doi.org/10.2113/JEEG12.2.173

Tshibangu K, J.P., Descamps, F., 2011. Engineering Challenges in the Geological Disposal of Radioactive Waste and Carbon Dioxide, in: Advances in Global Change Research. Springer International Publishing, pp. 185–213. https://doi.org/10.1007/978-90-481-8712-6_7

Usher, P.J., Angus, D.A., Verdon, J.P., 2013. Influence of a velocity model and source frequency on microseismic waveforms: some implications for microseismic locations. Geophys. Prospect. 61, 334–345. https://doi.org/10.1111/j.1365-2478.2012.01120.x

Vaezi, Y., Van Der Baan, M., 2015. Comparison of the STA/LTA and power spectral density methods for microseismic event detection. Geophys. J. Int. Geophys. J. Int 203, 1896–1908. https://doi.org/10.1093/gji/ggv419

Vaezi, Y., Van Der Baan, M., 2014. Analysis of instrument self-noise and microseismic event detection using power spectral density estimates. Geophys. J. Int. Geophys. J. Int 197, 1076–1089. https://doi.org/10.1093/gji/ggu036

Vouillomaz, J., 2009. Strain localization along shear zones in the Juchlistock area. University of Bern, Switzerland.

Walter, Marco, Joswig, Manfred, Walter, M, Schwaderer, U., Joswig, M, 2012. Seismic monitoring of precursory fracture signals from a destructive rockfall in the Vorarlberg Alps, Austria. Hazards Earth Syst. Sci 12, 3545–3555. https://doi.org/10.5194/nhess-12-3545-2012

Wang, J., Teng, T.-L., 1995. Artificial Neural Network-Based Seismic Detector, Bulletin of the Seismological Society of America.

Wehrens, P., 2015. Structural evolution in the Aar Massif (Haslital transect): Implications for mid-crustal deformation. University of Bern, Switzerland.

Wicki, T., 2011. 3D-shear zone pattern in the Grimsel area: ductile to brittle deformation in granitic rocks. University of Bern, Switzerland.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., Trujillo, J., 1998. A comparison of select trigger algorithms for automated global seismic phase and event detection. Bull. Seismol. Soc. Am. 88, 95–106.

References

Wolf, L.W., Rowe, C.A., Homer, R.B., 1997. Periodic Seismicity near Mt. Ogden on the Alaska-British Columbia Border: A Case for Hydrologically Triggered Earthquakes?, Bulletin of the Seismological Society of America.

Wolfe, P., 1965. The Composite Simplex Algorithm. Soc. Ind. Appl. Math. 7, 42–54.

Wuestefeld, A., Greve, S.M., Näsholm, S.P., Oye, V., 2018. Benchmarking earthquake location algorithms: A synthetic comparison. Geophysics 83, 35–47. https://doi.org/10.1190/GEO2017-0317.1

Wust-Bloch, H.G., 2010. Characterizing and Locating Very Weak (−2.2 ≥ M L ≥ −3.4) Induced Seismicity in Unstable Sandstone Cliffs by Nanoseismic Monitoring. Pure Appl. Geophys. 167, 153–167. https://doi.org/10.1007/s00024-009-0011-6

Xu, C., Stewart, R., 2006. Estimating seismic attenuation (Q) from VSP data, in: 2006 CSPG/CSEG/CWLS Joint Convention.

Yfantis, G., Carvajal, H.E.M., Pytharouli, S., Lunn, R.J., 2014. Spectral characteristics of landslide induced seismicity: experimental validation based on the use of an up-scaled sheer box. Am. Geophys. Union, Fall Meet. 2014, Abstr. id. NH43A-3821.

Young, R.P., Martin, C.D., 1993. Potential role of acoustic emission/microseismicity investigations in the site characterization and performance monitoring of nuclear waste repositories. Int. J. Rock Mech. Min. Sci. Geomech. Abstr. 30, 797–803. https://doi.org/10.1016/0148-9062(93)90025-9

Zhao, Y., Takano, K., 1999. An Artificial Neural Network Approach for Broadband Seismic Phase Picking, Bulletin of the Seismological Society of America.

Zhou, R., Huang, L., Rutledge, J., 2010. Microseismic event location for monitoring CO2 injection using double-difference tomography. Lead. Edge 29, 208–214. https://doi.org/10.1190/1.3304826

Ziegler, Martin, Loew, S., Bahat, D., 2014. Growth of exfoliation joints and near-surface stress orientations inferred from fractographic markings observed in the upper Aar valley (Swiss Alps). Tectonophysics 626, 1–20. https://doi.org/10.1016/J.TECTO.2014.03.017

Ziegler, M., Loew, S., Moore, J.R., 2013. Distribution and inferred age of exfoliation joints in the Aar Granite of the central Swiss Alps and relationship to Quaternary landscape evolution. Geomorphology 201, 344–362. https://doi.org/10.1016/J.GEOMORPH.2013.07.010

Ziegler, M., Loew, S., Valley, B., Amann, F., 2014. Near-surface principal stress trajectories in Alpine valleys (Grimsel region, Switzerland) as inferred from exfoliation fracture plumose structure axes and numerical modeling, in: 48th US Rock Mechanics / Geomechanics Symposium. Minneapolis, MN, USA.

Zoback, M.D., Harjes, H.P., 1997. Injection-induced earthquakes and crustal stress at 9 km depth at the KTB deep drilling site. J. Geophys. Res.-Solid Earth 102(B8), 18477–18 491.

Appendix A

In this Appendix, the codes written for this research are listed. The command in subchapter A.1 runs in CMD (i.e. command prompt), in A.6 runs in R studio while in the rest of the subchapters operate in Matlab. All mentioned codes (subchapter A.2 onwards) can be found in Strathcloud Sharefile in the link provided: https://strathcloud.sharefile.eu/d-sfd42fe4480d40b48.

A.1 Conversion from raw data to ASCII

Using RT_ASC.exe, run in command prompt the command: for /r %v in (1*) do RT_ASC.exe %v, and all folders containing raw files are converted in ASCII files. Each raw file generates as many ASCII files as the sensors that were attached to the data logger; i.e. six in the case of BD57, three in the case of BD961, per hour of data.

A.2 Conversion from ASCII to .mat files

The algorithms, written in MATLAB to convert, in an automated way, ASCII (.atr files) to MATLAB variables include the main script, called main.m, which calls functions importfile.m and parser.m. All .m files need to be within the same folder as the ASCII files. Function parser.m uses the following theoretical background and manufacturer provided info. The raw numerical data from the A/D chips are measured in counts. The A/D volts per count (Bit Weight) is measured by REFTEK and stored in each A/D board when built. To convert the waveforms amplitude from counts to ground velocity we need to use the formula: Ground velocity = [Amplitude * Bit Weight] / [Gain * Sensitivity], where bit weight, gain and sensitivity vary depending on the data logger (see also Table 1 for values).

A.3 Correction of files start time and length

A script, called fixStarts1.mat, calls functions ts_split.m and ts_append.m and fixes the files that, because of some malfunction, are not 1-hr duration files that start from the beginning of the hour. The script splits three consecutive hours in bits (ts_split.mat):

a. 1st hour: from the beginning of the file until 60 minutes
b. 1st hour: after the 60 minutes (beginning of 2nd hour) until the end of the file
c. 2nd hour: from the beginning of the file until 60 minutes
d. 2nd hour: after the 60 minutes (beginning of 3rd hour) until the end of the file
e. and so on

and concatenates every two bits (ts_append.mat); i.e. b and c make a full hour file starting from the beginning of the hour. These malfunctions were observed only in the South array recordings (LOC1).

A.4 Reconfiguration of .mat files from cell to structure arrays

This is a script to automatically convert all .mat files from cell arrays to structure arrays to increase functionality and enable a more easy manipulation of the data (reconfig.mat). Both cell and structure arrays allow you to store data of different types and sizes. Structure arrays contain data in fields that you access by name while cell arrays by numeric indexing. The further complicated codes became the simpler the data storing tools I used had to become, hence I converted all my data in structure arrays.

A.5 NpD algorithm

Appendix A.5 has been published as Kinali et al. (2018).

The NpD algorithm is going to be distributed as an open-source detection algorithm. The algorithm steps (Step 1 and 2) have been automatized in the form of a code that runs in Matlab environment. The raw seismic data are converted from ASCII format to MATLAB files using simple algorithms. In this step the files are named: sensor_year_DOY_hour_min_ sec_μs_ channel, where sensor can be either LOC1, LOC2 or BH, DOY is the day of the year, the μsec have an accuracy of four digits and channel can be CH1:6. Then the mat files are pre-processed before fed into the algorithm: the counts are converted to ground velocity within the passband. Faulty files are dismissed (e.g. files that due to electrical malfunction of the sensors recorded some minutes instead of a full hour data record) during this step. The data are filtered with just a band-stop recursive Butterworth filter at 48-52Hz to remove the mains electromagnetic interference which is prevalent. No further filtering has been applied. The mat files are also demeaned and fed into the algorithm as structure arrays. Each structure array contains four fields: data (900000 data points), date (character array in the form of 'dd-mmm-yyyy HH:MM:SS.mmmm' which indicates the beginning of the file), sensor (e.g. 'LOC2') and channel (e.g. 'CH1').

The output of the code contains the variable 'FinalRslts' which is a structure array with 3 fields: names (character array in the form of 'DOY_HH'), times (the times from the beginning of the hour the potential events are detected, in sec), timesForXcel (the times from the beginning of the hour the potential events are detected, in MM:SS:mmm). The variable 'listingTotal' is another useful output variable of the code listing the full names of the files checked from the code. The output variables 'Step1_all_values' contains two column cells: the second column encloses the file checked while the first the values of misfits and corresponding times of all data points during the first step of the algorithm. The output variables 'Step1_above_threshold' follows the logic of 'Step1_all_values' only this time the first column cells enclose the values of misfits and corresponding times of only the data points that successfully passed the first step of the algorithm. The output variable 'PredictedEventsIndivChannel' follows the previous logic and contains all values of misfits and corresponding times of only the data points that successfully passed the second step of the algorithm. This variable is different from the 'FinalRslts' because the former refers to individual channels (the voting scheme has not yet been applied), neither has the consecutive events cleaning.

NpD algorithm includes the script Script_final.m which calls functions: function1_08_2017.m and function2_08_2017.m, meta_analysisFinal.m, and removalOfExcessPeaks.m.

A.6 *AIC-wav* algorithm

As presented in Chapter 4, *AIC-wav* algorithm takes small windows around the NpD potential events times and uses Daubechies wavelet transform filter of length 2 with 2 levels of decomposition and the Akaike information criterion (AIC). The code runs in R studio programming language and includes script P_wave.R which calls util.R. The outputs are in the form of .tsv files.

A.7 Conversion for VFOM

As discussed in Chapter 5, VFOM algorithms which run in Wolfram Mathematica programming language were kindly provided by Wang, Z but needed the potential events to be fed into the codes in a particular input format. For this a script in Python was written (script.py.), which uses information in the filename and creates a folder tree within which we will store the data files as per VFOM requirements. Once the folders are created then a code named CompareTSV_VFOM_code_2014.m which uses readTSVFile.m compares the 2 arrays TSV outputs from *AIC-wav* and gets the common potential picks.

A.8 Conversion for GISMO

To conduct the analysis presented in Chapter 5, I used the seismic data analysis toolbox for the cluster analysis in MATLAB (GISMO). GISMO requires the data to be input as waveform objects and for that the necessary conversion algorithms were written: openWaveform.m or fullWaveform.m.

A.9 TopInv conversion

The TOPological INVersion (TOPINV) algorithm (Saltogianni and Stiros, 2012a,b, 2013, 2015) was translated in Matlab and adapted for the purposes of the current research and the matlab function for that can be found in Sharefile (locode.m).

A.10 Source parameters

The script modelfit.m. which calls functions Brunewright.m and fitfun2outputfcn.m and fitfun2.m are also included in Sharefile. modelfit was written by Stella Pytharouli in Matlab and fits Brune (1970) and Boatwright (1980) model using the Nelder/Mead algorithm and was modified for the needs of this project. The script calculates the model values for the values of the frequency and plots the spectrum and the fit. It then computes the source parameters and creates a matrix with their values. Brunewright.m was also written by Stella Pytharouli while fitfun2outputfcn.m and fitfun2.m are freely available by the Mathworks.Inc.

Appendix B

Year	UTC (Month/Day HH:MM:SS)	Magnitude	Location	Depth	Latitude	Longitude
2014	Nov 13 16:51:17	0.80	'Unterschaechen UR'	1.00	46.84	8.85
2014	Dec 06 22:26:04	0.20	'Goeschenen UR'	2.10	46.68	8.49
2014	Dec 07 09:34:27	0.30	'Simplon Dorf VS'	1.00	46.18	8.03
2014	Dec 13 23:13:33	0.90	'Brig-Glis VS'	7.40	46.30	7.98
2014	Dec 16 13:28:17	1.00	'Goeschenen UR'	5.40	46.67	8.50
2014	Dec 17 00:43:17	0.40	'Goeschenen UR'	5.50	46.68	8.50
2015	Jan 04 02:37:01	0.60	'Gadmen BE'	2.00	46.68	8.34
2015	Jan 16 04:54:06	0.70	'Reckingen VS'	9.00	46.51	8.28
2015	Jan 17 15:18:56	0.50	'Altdorf UR'	6.00	46.89	8.66
2015	Feb 12 17:11:28	0.80	'Muerren BE'	3.00	46.53	7.99
2015	Feb 23 00:30:30	0.80	'Fusio TI'	8.50	46.42	8.54
2015	Mar 01 22:00:06	1.30	'Engelberg OW'	-0.60	46.85	8.38
2015	Mar 04 13:01:06	0.50	'Olivone TI'	5.80	46.53	8.89
2015	Mar 12 22:45:59	1.80	'Altdorf UR'	8.20	46.90	8.60
2015	Mar 19 04:49:44	0.50	'Reckingen VS'	2.00	46.44	8.28
2015	Mar 19 21:01:34	0.80	'Bosco/Gurin TI'	5.40	46.34	8.54
2015	Mar 29 22:59:38	0.60	'Brienz BE'	3.00	46.72	8.08
2015	Mar 29 23:43:22	0.90	'Brienz BE'	3.00	46.72	8.08
2015	Mar 31 01:47:44	1.50	'Altdorf UR'	7.00	46.88	8.60
2015	Apr 18 03:37:51	0.40	'Unterschaechen UR'	5.70	46.90	8.78
2015	Apr 20 06:59:11	0.80	'Brig-Glis VS'	7.20	46.31	8.00

Appendix B

Year	UTC (Month/Day HH:MM:SS)	Magnitude	Location	Depth	Latitude	Longitude
2015	May 01 20:06:27	0.70	'Simplon Dorf VS'	5.10	46.25	8.10
2015	May 06 00:47:27	1.70	'Reckingen VS'	7.90	46.46	8.18
2015	May 07 07:24:13	1.20	'Amsteg UR'	9.10	46.81	8.64
2015	May 08 19:17:12	1.00	'Amsteg UR'	6.20	46.80	8.63
2015	May 16 11:54:01	0.90	'Reckingen VS'	2.00	46.55	8.20
2015	May 25 03:17:14	0.60	'Brig-Glis VS'	2.70	46.30	7.98
2015	Jun 04 11:33:06	1.20	'Amsteg UR'	8.50	46.80	8.64
2015	Jun 08 11:57:12	0.30	'Oberwald VS'	7.10	46.61	8.34
2015	Jun 13 01:04:10	1.00	'Bosco/Gurin TI'	2.00	46.35	8.41
2015	Jun 13 22:48:44	0.80	'Oberwald VS'	5.30	46.61	8.31
2015	Jun 13 22:48:45	0.80	'Oberwald VS'	5.50	46.61	8.32
2015	Jun 16 00:57:53	0.70	'Reckingen VS'	7.50	46.51	8.22
2015	Jun 18 15:06:48	0.50	'Oberwald VS'	7.10	46.62	8.35
2015	Jun 18 15:09:01	0.70	'Oberwald VS'	6.00	46.63	8.36
2015	Jun 18 15:41:34	0.30	'Oberwald VS'	7.10	46.62	8.35
2015	Jul 12 08:07:34	0.80	'Reckingen VS'	3.80	46.45	8.36
2015	Jul 12 22:07:03	1.10	'Unterschaechen UR'	7.40	46.84	8.80
2015	Jul 13 16:18:20	1.50	'Cevio TI'	12.10	46.23	8.64
2015	Jul 19 11:39:39	0.60	'Spiez BE'	4.60	46.66	7.69
2015	Jul 20 22:01:37	0.80	'Oberwald VS'	8.40	46.57	8.40
2015	Jul 29 17:12:58	1.00	'Unterschaechen UR'	9.60	46.84	8.85
2015	Jul 29 18:06:20	1.20	'Bosco/Gurin TI'	7.50	46.40	8.53
2015	Aug 18 07:10:10	2.10	'Soerenberg LU'	7.70	46.83	8.08
2015	Aug 18 17:04:12	0.90	'Muerren BE'	1.00	46.52	7.80
2015	Aug 25 19:40:35	0.80	'Binn VS'	3.30	46.28	8.20
2015	Aug 26 16:25:41	1.50	'Oberwald VS'	5.30	46.58	8.36
2015	Aug 26 21:02:52	0.80	'Reckingen VS'	2.00	46.42	8.29
2015	Sep 05 20:47:52	1.20	'Brione (Verzasca) TI'	3.00	46.36	8.81
2015	Sep 07 17:13:41	1.60	'Brig-Glis VS'	9.20	46.36	8.06
2015	Sep 10 06:51:25	0.70	'Visp VS'	7.30	46.36	7.89

Year	UTC (Month/Day HH:MM:SS)	Magnitude	Location	Depth	Latitude	Longitude
2015	Sep 15 06:15:53	1.10	'Bosco/Gurin TI'	5.10	46.38	8.41
2015	Sep 18 16:07:58	1.00	'Reckingen VS'	3.60	46.54	8.16
2015	Sep 27 21:00:11	0.40	'Bosco/Gurin TI'	3.00	46.36	8.41
2015	Oct 02 13:44:27	1.00	'Unterschaechen UR'	3.20	46.83	8.84
2015	Oct 05 11:27:07	0.80	'Unterschaechen UR'	2.10	46.84	8.80
2015	Oct 05 14:41:22	0.80	'Binn VS'	7.30	46.35	8.15
2015	Oct 06 16:44:47	0.80	'Fusio TI'	10.20	46.41	8.55
2015	Oct 08 19:34:53	0.50	'Binn VS'	2.00	46.36	8.26
2015	Oct 10 01:14:37	1.00	'Binn VS'	2.00	46.28	8.17
2015	Oct 17 20:11:06	0.50	'Visp VS'	5.60	46.28	7.95
2015	Oct 23 18:44:53	0.90	'Bosco/Gurin TI'	9.30	46.35	8.47
2015	Nov 01 18:06:06	1.00	'Fiesch VS'	7.00	46.35	8.10
2015	Nov 04 21:18:19	1.10	'Reckingen VS'	5.80	46.45	8.36
2015	Nov 12 16:59:01	0.60	'Goeschenen UR'	5.60	46.68	8.50
2015	Nov 16 00:55:25	0.70	'Goeschenen UR'	2.70	46.68	8.52
2015	Nov 18 07:16:53	0.80	'Goeschenen UR'	2.00	46.68	8.51
2015	Nov 19 06:41:47	0.70	'Meiringen BE'	2.00	46.71	8.18
2015	Nov 22 06:19:41	0.10	'Goeschenen UR'	5.90	46.68	8.49
2015	Nov 22 06:33:23	0.50	'Goeschenen UR'	5.60	46.68	8.50
2015	Nov 24 18:49:28	0.40	'Goeschenen UR'	5.20	46.67	8.50
2015	Nov 26 23:56:54	0.40	'Brig-Glis VS'	7.30	46.29	7.95
2015	Dec 03 03:56:39	0.70	'Binn VS'	6.30	46.37	8.17
2015	Dec 06 11:59:42	-0.10	'Realp UR'	9.20	46.62	8.42
2015	Dec 09 00:36:18	0.10	'Goppenstein VS'	5.00	46.38	7.81
2015	Dec 09 11:12:04	0.30	'Goppenstein VS'	6.80	46.38	7.82
2015	Dec 15 04:55:00	-0.10	'Oberwald VS'	1.20	46.61	8.37
2015	Dec 15 20:32:31	0.40	'Muerren BE'	3.00	46.54	7.91
2015	Dec 17 03:49:45	2.00	'Brienz BE'	5.80	46.75	8.07

Appendix B

Year	UTC (Month/Day HH:MM:SS)	Magnitude	Location	Depth	Latitude	Longitude
2015	Dec 18 10:00:38	1.60	'Oberwald VS'	10.50	46.51	8.36
2015	Dec 20 07:39:34	1.50	'Brienz BE'	4.80	46.75	8.07
2015	Dec 23 18:18:06	0.80	'Unterschaechen UR'	2.00	46.84	8.80
2015	Dec 23 21:13:23	-0.10	'Realp UR'	7.20	46.59	8.45
2015	Dec 24 15:25:58	0.80	'Bosco/Gurin TI'	2.10	46.35	8.45
2015	Dec 26 00:32:17	0.40	'Unterschaechen UR'	2.00	46.83	8.81
2016	Dec 18 07:51:08	0.40	'Realp UR'	5.00	46.65	8.49
2016	Dec 18 01:32:12	0.80	'Realp UR'	6.10	46.65	8.48
2016	Dec 17 23:38:21	0.30	'Realp UR'	5.60	46.65	8.48
2016	Dec 17 17:30:26	0.30	'Realp UR'	7.00	46.65	8.47
2016	Dec 17 17:22:38	0.50	'Kandersteg BE'	3.00	46.53	7.73
2016	Dec 11 17:27:18	-0.10	'Oberwald VS'	7.10	46.60	8.40
2016	Dec 02 16:14:42	0.90	'Fiesch VS'	7.40	46.37	8.06
2016	Nov 28 08:01:48	0.50	'Goeschenen UR'	4.90	46.68	8.50
2016	Nov 11 12:29:20	1.20	'Brione (Verzasca) TI'	3.70	46.33	8.76
2016	Nov 03 21:19:41	0.20	'Goeschenen UR'	4.50	46.68	8.50
2016	Nov 03 20:02:36	0.60	'Unterschaechen UR'	2.00	46.82	8.77
2016	Oct 29 05:16:41	0.80	'Faido TI'	6.10	46.58	8.77
2016	Oct 26 04:45:49	0.40	'Brig-Glis VS'	7.60	46.30	7.95
2016	Oct 16 19:46:13	0.60	'Goeschenen UR'	3.50	46.68	8.50
2016	Oct 11 07:08:50	0.90	'Unterschaechen UR'	2.00	46.84	8.85
2016	Oct 08 02:31:59	2.60	'Fusio TI'	4.00	46.40	8.55
2016	Oct 03 07:24:53	0.90	'Goeschenen UR'	5.90	46.68	8.49
2016	Oct 03 06:43:43	3.20	'Goeschenen UR'	4.90	46.67	8.51
2016	Sep 29 02:17:50	0.40	'Goeschenen UR'	4.10	46.68	8.50
2016	Sep 26 02:06:02	0.40	'Goeschenen UR'	3.00	46.68	8.50
2016	Sep 25 04:23:03	0.80	'Goeschenen UR'	3.00	46.68	8.50
2016	Sep 23 12:58:07	0.90	'Goeschenen UR'	3.00	46.68	8.50
2016	Sep 18 21:14:05	0.70	'Visp VS'	6.30	46.27	7.94
2016	Sep 18 21:12:30	0.90	'Bosco/Gurin TI'	9.50	46.41	8.53

Year	UTC (Month/Day HH:MM:SS)	Magnitude	Location	Depth	Latitude	Longitude
2016	Sep 14 13:54:30	1.10	'Domodossola I'	10.90	46.18	8.31
2016	Sep 12 00:31:56	0.90	'Muerren BE'	-0.20	46.51	7.79
2016	Sep 06 00:46:16	1.00	'Muerren BE'	3.40	46.49	7.87
2016	Aug 19 01:59:14	0.80	'Grindelwald BE'	3.00	46.62	8.09
2016	Aug 12 16:51:49	1.00	'Beckenried NW'	10.70	47.00	8.47
2016	Aug 10 13:50:46	0.60	'Fusio TI'	9.30	46.43	8.55
2016	Aug 10 11:15:37	-0.30	'Oberwald VS'	-1.90	46.57	8.35
2016	Aug 05 23:11:54	1.70	'Brunnen SZ'	2.30	46.95	8.66
2016	Aug 03 14:48:40	1.90	'Fusio TI'	8.60	46.41	8.54
2016	Jul 31 03:47:18	1.10	'Beckenried NW'	1.80	46.90	8.44
2016	Jul 19 04:11:11	0.70	'Goppenstein VS'	7.70	46.38	7.82
2016	Jul 18 23:04:34	0.70	'Binn VS'	5.30	46.36	8.22
2016	Jul 12 16:26:38	0.80	'Grindelwald BE'	0.40	46.58	8.02
2016	Jul 09 04:24:55	0.60	'Frutigen BE'	4.40	46.60	7.68
2016	Jun 17 12:06:45	2.70	'Frutigen BE'	7.30	46.60	7.70
2016	Jun 10 22:29:59	0.90	'Sarnen OW'	0.30	46.90	8.35
2016	May 23 10:47:27	0.30	'Brig-Glis VS'	1.90	46.35	8.05
2016	May 14 18:29:01	0.40	'Goeschenen UR'	3.90	46.68	8.50
2016	May 14 18:04:45	0.40	'Goeschenen UR'	3.40	46.68	8.50
2016	May 14 16:40:10	0.50	'Goeschenen UR'	4.60	46.68	8.50
2016	May 04 23:36:19	0.80	'Visp VS'	2.70	46.26	7.91
2016	May 01 09:36:48	1.20	'Brione (Verzasca) TI'	2.10	46.38	8.84
2016	Apr 30 14:18:54	1.20	'Biasca TI'	8.50	46.37	8.88
2016	Apr 28 18:04:07	0.30	'Brig-Glis VS'	4.30	46.31	8.02
2016	Apr 21 12:23:57	1.70	'Amsteg UR'	4.50	46.79	8.72
2016	Apr 15 00:50:01	1.00	'Brig-Glis VS'	7.10	46.31	8.08
2016	Apr 12 06:36:03	1.00	'Unterschaechen UR'	2.20	46.83	8.80
2016	Mar 31 23:51:44	0.20	'Goeschenen UR'	5.90	46.68	8.50

Appendix B

Year	UTC (Month/Day HH:MM:SS)	Magnitude	Location	Depth	Latitude	Longitude
2016	Mar 30 19:39:19	2.10	'Binn VS'	5.90	46.40	8.21
2016	Mar 28 19:43:26	0.70	'Oberwald VS'	6.90	46.61	8.36
2016	Mar 27 20:02:50	1.40	'Visp VS'	6.60	46.27	7.94
2016	Mar 15 19:06:04	0.70	'Oberwald VS'	7.60	46.60	8.36
2016	Mar 15 18:56:06	-0.60	'Oberwald VS'	6.50	46.59	8.35
2016	Mar 15 18:55:48	0.00	'Oberwald VS'	7.50	46.61	8.36
2016	Mar 15 18:44:06	0.00	'Oberwald VS'	7.50	46.61	8.36
2016	Feb 28 11:49:13	0.10	'Unterschaechen UR'	2.00	46.85	8.81
2016	Feb 28 04:55:54	1.40	'Unterschaechen UR'	5.50	46.84	8.80
2016	Feb 27 12:08:23	0.40	'Reckingen VS'	7.30	46.43	8.38
2016	Feb 27 01:10:46	1.50	'Unterschaechen UR'	2.00	46.83	8.80
2016	Feb 24 23:54:51	1.40	'Reckingen VS'	5.50	46.44	8.38
2016	Feb 20 06:22:07	1.00	'Faido TI'	2.00	46.56	8.77
2016	Feb 18 03:03:02	1.70	'Fusio TI'	8.80	46.41	8.55
2016	Feb 17 02:58:14	0.30	'Graechen VS'	2.00	46.23	7.92
2016	Feb 09 03:07:55	0.80	'Muerren BE'	2.00	46.54	7.92
2016	Feb 04 22:39:30	0.50	'Oberwald VS'	8.50	46.53	8.28
2016	Feb 02 21:48:48	0.10	'Bosco/Gurin TI'	7.40	46.36	8.46
2016	Feb 02 07:50:54	1.50	'Bosco/Gurin TI'	5.00	46.35	8.45
2016	Jan 31 05:44:09	0.80	'Binn VS'	2.00	46.33	8.19
2016	Jan 23 20:23:12	0.40	'Reckingen VS'	2.10	46.44	8.33
2016	Jan 21 19:47:20	1.30	'Muerren BE'	3.30	46.54	7.91
2016	Jan 21 18:48:29	0.10	'Oberwald VS'	6.30	46.60	8.34
2016	Jan 15 20:34:05	0.70	'Grindelwald BE'	2.00	46.63	8.09
2016	Jan 15 01:14:17	0.80	'Visp VS'	7.70	46.27	7.94
2016	Jan 11 06:40:29	1.00	'Brig-Glis VS'	4.90	46.31	8.03
2016	Jan 10 07:35:55	1.40	'Binn VS'	7.60	46.28	8.14
2016	Jan 02 11:48:46	0.30	'Goeschenen UR'	3.00	46.68	8.50
2016	Jan 02 10:42:44	3.10	'Muerren BE'	6.90	46.53	8.00
2016	Feb 28 04:56:02	2.38	Event 1	1.86	46.75	8.32

Year	UTC (Month/Day HH:MM:SS)	Magnitude	Location	Depth	Latitude	Longitude
2016	Mar 07 21:16:23	1.11	Event 2	1.38	46.62	8.42
2016	Mar 08 11:22:47	1.03	Event 3	6.50	46.60	8.32
2016	Mar 08 11:36:16	1.12	Event 4	2.29	46.62	8.40
2016	Mar 14 12:13:13	1.58	Event 5	-0.28	46.61	8.38
2016	Mar 15 18:37:35	1.10	Event 6	0.50	46.61	8.40
2016	Mar 15 18:41:03	0.97	Event 7	7.14	46.60	8.36
2016	Mar 15 18:44:36	0.83	Event 8	6.73	46.60	8.37
2016	Mar 15 18:45:52	1.17	Event 9	0.71	46.62	8.40
2016	Mar 15 18:45:29	0.91	Event 10	7.08	46.60	8.36
2016	Mar 15 18:48:45	0.88	Event 11	7.28	46.60	8.36
2016	Mar 15 18:48:53	0.72	Event 12	6.45	46.60	8.37
2016	Mar 15 18:49:43	0.85	Event 13	7.22	46.60	8.36
2016	Mar 15 18:54:16	1.08	Event 14	5.85	46.60	8.38
2016	Mar 15 18:56:08	1.33	Event 15	1.95	46.62	8.40
2016	Mar 15 19:07:27	1.26	Event 16	1.81	46.62	8.40
2016	Mar 15 19:22:29	1.15	Event 17	3.33	46.61	8.40
2016	Mar 15 19:23:00	1.12	Event 18	2.90	46.62	8.40
2016	Mar 15 19:52:46	0.82	Event 19	7.28	46.60	8.36
2016	Mar 15 20:18:08	1.23	Event 20	0.34	46.62	8.40
2016	Mar 16 01:41:59	0.78	Event 21	7.28	46.60	8.36
2016	Mar 16 15:38:45	1.26	Event 22	2.70	46.59	8.43
2016	Mar 21 10:42:57	1.29	Event 23	7.29	46.59	8.35
2016	Apr 06 15:18:00	1.16	Event 24	9.97	46.62	8.45
2017	Aug 13 23:25:56	0.60	'Brig-Glis VS'	3.30	46.28	7.95
2017	Aug 12 15:08:04	-0.20	'Oberwald VS'	1.50	46.56	8.24
2017	Aug 09 20:26:55	0.50	'Realp UR'	7.80	46.63	8.46
2017	Jul 30 05:59:35	1.50	'Reckingen VS'	7.20	46.47	8.17
2017	Jul 19 23:13:25	1.20	'Oberwald VS'	2.90	46.59	8.36

Appendix B

Year	UTC (Month/Day HH:MM:SS)	Magnitude	Location	Depth	Latitude	Longitude
2017	Jul 18 04:39:47	0.70	'Goeschenen UR'	4.70	46.67	8.50
2017	Jul 17 01:25:39	0.40	'Goeschenen UR'	3.70	46.68	8.50
2017	Jul 08 03:43:26	0.70	'Muerren BE'	6.30	46.47	7.83
2017	Jul 07 07:12:46	1.20	'Stansstad NW'	-0.90	46.96	8.32
2017	Jun 20 01:52:10	0.90	'Kandersteg BE'	0.30	46.49	7.79
2017	Jun 15 17:01:29	0.60	'Oberwald VS'	7.80	46.59	8.33
2017	Jun 03 01:09:43	0.30	'Kandersteg BE'	-0.50	46.47	7.72
2017	May 22 13:34:04	0.40	'Goeschenen UR'	5.50	46.68	8.49
2017	May 13 10:29:55	0.40	'Muerren BE'	6.00	46.47	7.83
2017	May 09 22:46:33	-0.10	'Goppenstein VS'	6.70	46.41	7.85
2017	Apr 30 21:10:45	0.60	'Visp VS'	6.10	46.30	7.84
2017	Apr 29 13:01:42	0.30	'Goeschenen UR'	5.40	46.66	8.65
2017	Apr 29 00:44:09	0.90	'Visp VS'	6.70	46.30	7.84
2017	Apr 29 00:32:27	0.20	'Visp VS'	6.30	46.30	7.84
2017	Apr 29 00:21:54	1.00	'Visp VS'	6.40	46.31	7.84
2017	Apr 12 20:36:13	0.70	'Fiesch VS'	8.10	46.37	8.07
2017	Apr 12 02:59:31	1.20	'Goeschenen UR'	7.00	46.67	8.51
2017	Apr 11 09:11:26	1.20	'Brione (Verzasca) TI'	2.00	46.37	8.81
2017	Apr 07 19:21:59	0.30	'Gadmen BE'	2.70	46.78	8.31
2017	Apr 07 14:32:23	1.00	'Brione (Verzasca) TI'	-0.40	46.32	8.78
2017	Apr 06 23:56:46	0.50	'Realp UR'	8.50	46.61	8.45
2017	Mar 20 10:34:53	0.70	'Domodossola I'	2.20	46.17	8.31
2017	Mar 17 17:02:59	0.70	'Binn VS'	2.00	46.30	8.12
2017	Mar 14 23:21:27	0.30	'Brienz BE'	5.80	46.79	8.05
2017	Mar 14 03:07:12	0.80	'Faido TI'	2.20	46.57	8.75
2017	Mar 07 01:53:49	0.60	'Grindelwald BE'	3.00	46.52	8.07
2017	Mar 06 12:17:54	1.30	'Reckingen VS'	6.00	46.54	8.18
2017	Mar 03 22:16:51	1.00	'Reckingen VS'	5.40	46.49	8.12
2017	Feb 28 03:25:23	0.70	'Brione (Verzasca) TI'	9.30	46.32	8.86
2017	Feb 19 12:17:00	0.60	'Fiesch VS'	8.10	46.44	8.09

Year	UTC (Month/Day HH:MM:SS)	Magnitude	Location	Depth	Latitude	Longitude
2017	Feb 17 09:01:07	1.20	'Muerren BE'	3.00	46.56	7.96
2017	Feb 15 15:39:08	1.40	'Frutigen BE'	6.50	46.60	7.69
2017	Feb 15 05:34:03	0.00	'Goeschenen UR'	3.20	46.68	8.50
2017	Feb 13 12:02:03	0.10	'Oberwald VS'	7.50	46.60	8.40
2017	Feb 13 10:55:19	0.10	'Oberwald VS'	7.20	46.59	8.39
2017	Feb 13 10:46:21	0.40	'Oberwald VS'	7.90	46.59	8.40
2017	Feb 13 10:46:13	0.00	'Oberwald VS'	7.10	46.60	8.39
2017	Feb 13 10:43:00	0.00	'Oberwald VS'	6.80	46.60	8.39
2017	Feb 12 10:10:44	-0.20	'Oberwald VS'	7.30	46.59	8.38
2017	Feb 11 21:10:11	-0.20	'Oberwald VS'	7.00	46.60	8.39
2017	Feb 09 10:10:16	1.90	'Brienz BE'	6.10	46.78	8.06
2017	Feb 02 00:05:18	0.40	'Goeschenen UR'	5.50	46.68	8.50
2017	Jan 27 11:20:52	1.40	'Cevio TI'	13.30	46.25	8.63
2017	Jan 20 01:08:38	0.40	'Kandersteg BE'	2.00	46.53	7.74
2017	Jan 19 18:08:03	1.30	'Kandersteg BE'	2.00	46.53	7.74
2017	Jan 19 11:21:16	0.70	'Domodossola I'	9.80	46.19	8.34
2017	Jan 13 09:04:04	0.80	'Goeschenen UR'	4.00	46.68	8.50
2017	Jan 13 06:13:13	2.50	'Goeschenen UR'	5.00	46.67	8.51
2017	Jan 12 08:41:46	0.40	'Goeschenen UR'	4.10	46.68	8.50
2017	Jan 11 16:28:38	0.00	'Goeschenen UR'	3.80	46.68	8.51
2017	Jan 09 09:54:51	0.50	'Goeschenen UR'	5.70	46.68	8.51
2017	Jan 07 03:52:17	0.70	'Binn VS'	6.60	46.37	8.17
2017	Jan 04 18:18:32	0.60	'Unterschaechen UR'	3.90	46.88	8.76
2017	Jan 04 17:01:36	1.90	'Unterschaechen UR'	6.80	46.89	8.76
2017	Jan 01 08:36:21	0.50	'Realp UR'	7.60	46.60	8.44