



Automatic Detection and Location of Microseismic Events from Sparse Network and Its Application to Post-mining Monitoring

D. NAMJESNIK,¹  J. KINSCHER,¹ Y. GUNZBURGER,² N. POIATA,^{3,4} P. DOMINIQUE,⁵ P. BERNARD,³ and I. CONTRUCCI¹

Abstract—In post-mining regions with seismic hazard, timely decision making for risk management faces the challenge of quick and reliable detection and location of seismic events. As a response to the increasing density of monitoring stations, generating large volumes of seismic data, automatic, full waveform-based methods have been developed in recent years in global seismology. Such methods often cannot be directly applied to post-mining monitoring with a limited station coverage, as it is the case when temporarily networks are installed as an emergency response. In this paper we propose a new methodology that bridges this gap and enables the application of a full waveform, backprojection based method (BackTrackBB) to data of sparse network. The methodology was successfully tested on an abandoned and flooded underground coalmine in South-eastern France. Steps preceding BackTrackBB application were implemented in order to remove the coherent noise that otherwise results in numerous false detections. First results indicate that seismic activity in the study area is controlled by water level variation within former room-and-pillar mine works and fault segments (re)activation below them.

Keywords: Post-mining seismicity, full-waveform analysis, automatic detection and location, sparse network.

1. Introduction

1.1. Challenges of Microseismic Monitoring in Post-mining Conditions

With the rising number of worldwide mining closures in the last few decades, potential hazard of abandoned mining districts in post-mining period started raising concerns of governments due to potential socio-economic impact. Therefore, many countries developed regulations for mine closure processes and management of post-mining hazards, especially in urban areas, where hazard identification and location of exposed zones needs to be as precise as possible. Potential consequences affecting people or infrastructure can persist long time and include modifications of the underground water flow, surface instabilities, toxic gas emission or discharge of potentially dangerous chemicals into environment. Post-mining hazard assessment and mapping methods are based on the nature of phenomena, and have as objective to identify potentially dangerous areas, determine the possible intensity of the foreseeable phenomena and predict the probability of its occurrence in observed zone (Didier et al., 2008).

Surface instability hazard that can affect areas with underground mining works depends on the geological context and exploitation techniques. Often implemented in most sensitive zones, microseismic monitoring is proven to be a valuable hazard assessment tool (Kinscher et al., 2017) for measurement of the failure initiation in rocks and expected ground surface instabilities such as subsidence (sudden collapse or continuous subsidence). This is especially important in the areas where risk cannot be reduced with other options such as backfilling, which is rarely used due to its high cost (Contrucci

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00024-021-02773-4>.

¹ Institut National de l'Environnement Industriel et des Risques (Ineris), Direction Ground and Underground Risk, Mines Nancy, Campus ARTEM, BP 14234, 54 042 Nancy Cédex, France. E-mail: dalija.namjesnik@gmail.com

² Université de Lorraine–CNRS, GeoRessources Lab., Mines Nancy, Campus ARTEM, BP 14234, 54 042 Nancy Cédex, France.

³ Institut de Physique du Globe de Paris–Université de Paris, Seismology, UMR 7154 CNRS, 75005 Paris, France.

⁴ National Institute for Earth Physics, Magurele, Romania.

⁵ Bureau de Recherches Géologiques et Minières (BRGM), Risk and Prevention Division, 3, Av. Claude-Guillemin, BP 36009, 45060 Orléans Cedex 2, France.

et al., 2008, 2011, 2019; Couffin et al., 2003; Didier, 2008).

Hazard assessment and mapping often faces challenges as mining areas can be large, and, as mines are sometimes abandoned several decades ago, data such as technical notes or maps are frequently lost (e.g. in Lorraine region in France, Bennani et al., 2003; Contrucci et al., 2019). Consequently, the determination of zones with priority for seismic monitoring can be difficult. Often seismic networks in such large mining areas are rather sparse (due to the large extent of the potential risk areas, as well as for economic reasons) and very often consist of only a single antenna, such as borehole sensors, positioned at a place identified as having the highest risk. These single station networks do not have high performance in terms of seismic event location accuracy and are clearly focused on event detection. When seismicity appears in any part of these large mining areas, complementary temporal mobile seismic monitoring networks could be installed to improve location accuracy and better understand the origin of seismic sources. Temporary networks mostly comprise a very low number of stations (as in case of Gardanne mine) which are sometimes only single-component geophones. Automatic processing of recorded seismic data for such sparse “task force” networks is challenging. Therefore, an automatic detection and location method applicable to seismic monitoring networks of limited station coverage composed of one-component geophones would be of great value especially in emergency situations where real-time operational monitoring is required.

1.2. Recent Developments of Automatic Detection and Location Methods

Standard approaches for detection and location in microseismic monitoring are generally adapted from global seismology, typically relying on manual or automatic P- and S-wave phase picking, sometimes combined with evaluation of polarization angles (e.g. Abdul-Wahed et al., 2001; Contrucci et al., 2010; Oye & Roth, 2003).

The continuous increase in seismic data availability and quality due to the increased numbers of monitoring stations and networks in global

earthquake monitoring has led to the development of various automatic detection and location methods with better detection capacities, contributing to a more thorough analysis of seismicity, due to detection of lower magnitude events leading to the lowering catalogue’s magnitude of completeness, and developing better constrains on location, with decreased location uncertainties or errors.

Waveform-based location methods that have emerged in recent years provide an alternative to the conventional phase-picking based techniques. They bypass phase picking and identification, focusing on the information provided from full waveform analysis. Waveform-based methods comprise partial waveform stacking (e.g. Li et al., 2018), time reverse imaging (e.g. Li & van der Baan, 2016), wavefront tomography (e.g. Diekmann et al., 2019), and full waveform inversion (e.g. Willacy et al., 2019), adapted from migration or stacking techniques in exploration seismology. These methods have been rapidly gaining popularity due to their simplicity and minimal necessary a priori constrains (Cesca & Grigoli, 2015), proving to provide robust and effective source location results at various scales and allowing to detect and locate events with low signal-to-noise ratio (SNR) (Li et al., 2020). Among the recently developed automatic methods, Back-TrackBB (further-BTBB; Poiata et al., 2016), a methodology for detection and space-time location of seismic sources based on multiscale, frequency-selective coherence of the wavefield characteristics recorded by dense large-scale seismic networks and local antennas, showed significant improvement in detection and location performance for different environments. The detection capacity of such coherence-based full waveform methods as BTBB often shows improvement by a factor 10 compared to standard triggering-based approaches (e.g., Aden-Antoniow et al., 2020; López-Comino et al., 2017). The method consists of signal transformation by constructing characteristic functions (CFs), and further the station-pair time-delay functions estimation, exploiting information on the coherency of the CFs between station-pairs. Finally, detection and location are performed on time-series of 3D spatial likelihood images, created from spatial mapping and stacking of station-pair time- delay estimate functions.

These new, automatic full waveform approaches have been so far applied to different industrial activities including (active) mining environment. However, they are still not standard in operational microseismic monitoring (e.g. Gharti et al., 2010; Gibbons & Ringdal, 2006; Grigoli et al., 2013; Palgunadi et al., 2019). Detailed overview of developments for the full waveform-based detection and location methods, related challenges and recent applications can be found in Cesca and Grigoli (2015) and Li et al. (2020). The strong potential of these methods in mining environments was recently demonstrated by the work of Palgunadi et al. (2019), where the main challenges were related to the presence of a wide range of seismic noise sources, connected to the ongoing mining activity, and a high sampling rate of recorded data (several kHz). In this study, detection capacity of BTBB demonstrated improvement of 50 times compared to previously available catalogue, what improve reliability of rock burst hazard assessment.

1.3. Seismic Activity at the Gardanne Former Coalmine and Connection with Hydrogeological Conditions

The Gardanne underground coalmine, in South-eastern France, was industrially exploited since 18th century until final closure in 2003 (Fig. 1). Zones of high risk of ground movements, related to mine workings instability, were defined by GEODERIS (public interest group, providing assistance and technical expertise on post-mining to the French government) as a part of post-mining management following the mine closure (Lefebvre & Dommagnet, 2010) and monitored since 2008 with permanent seismic monitoring network of Ineris (French National Institute for Industrial Environment and Risks) (Figs. 1, 2, 3).

After mine closure and stopping of the mine pumping, a gradual rise of underground water level led to progressive flooding of the residual voids,

starting from the deeper parts of mine workings, on the west (around -1100 m NgF¹), and progressing towards the east (around - 10 m NgF). Pumping activity, controlled by pumps set in a former mine shaft Gérard well (position shown in Fig. 3b) and managed by BRGM (the French Geological Survey), started again in 2010, to avoid any overflow of mineralized water at the surface. Since then, water level around - 10 m NgF is relatively stable by pumping, with annual fluctuations of the order about twenty meters (Dominique, 2016b). In addition to underground waters, the influence of effective rain on the rise of the water levels has also been noticed (Dheilly & Brigati, 2015).

Following the flooding front arrival to the eastern part of the mine in 2010 (indicated by the water level rise in Gérard well), seismicity started to occur there (Figs. 1, 2, 3a).

Corresponding seismic events were, unexpectedly, located outside of the previously identified risk zones, monitored with Ineris permanent network. Since then, swarming seismic activity has been re-appearing periodically in the same area approximately every 2 years. The strongest events with local magnitude close to 2 are occasionally felt by local population, which has led to rising concerns regarding seismic hazard and risk (Dominique, 2016a; Kinscher et al., 2018; Matrullo et al., 2015).

In 2013 the BRGM installed an additional temporary monitoring network in the seismic swarm area in order to investigate the origin of seismicity (Fig. 3b). The network, consisted of 2 to 5 stations within observed period (the configuration changes with time as shown in details in Fig. 1), all equipped with identical three-component accelerometers, installed on the surface and irregularly spaced at approximately 1–2 km one from another, across the approximately 10 km² seismically active zone (Fig. 3b). Data are recorded both in continuous and triggered mode.

Prior analysis of seismic data revealed apparent connection of seismic activity with seasonal changes of underground water level in the mine, as well as with pumping rates at Gérard well (which are changing during time) (Dominique, 2016b), as provided by piezometric measurement at the well, whose location is indicated in Fig. 3b.

¹ m NgF is a unit referring to The General Levelling of France (nivellement général de la France) with the 'zero level' determined by the tide gauge at Marseille. In this area 0 m NgF corresponds to approximately 400 m of depth below surface.

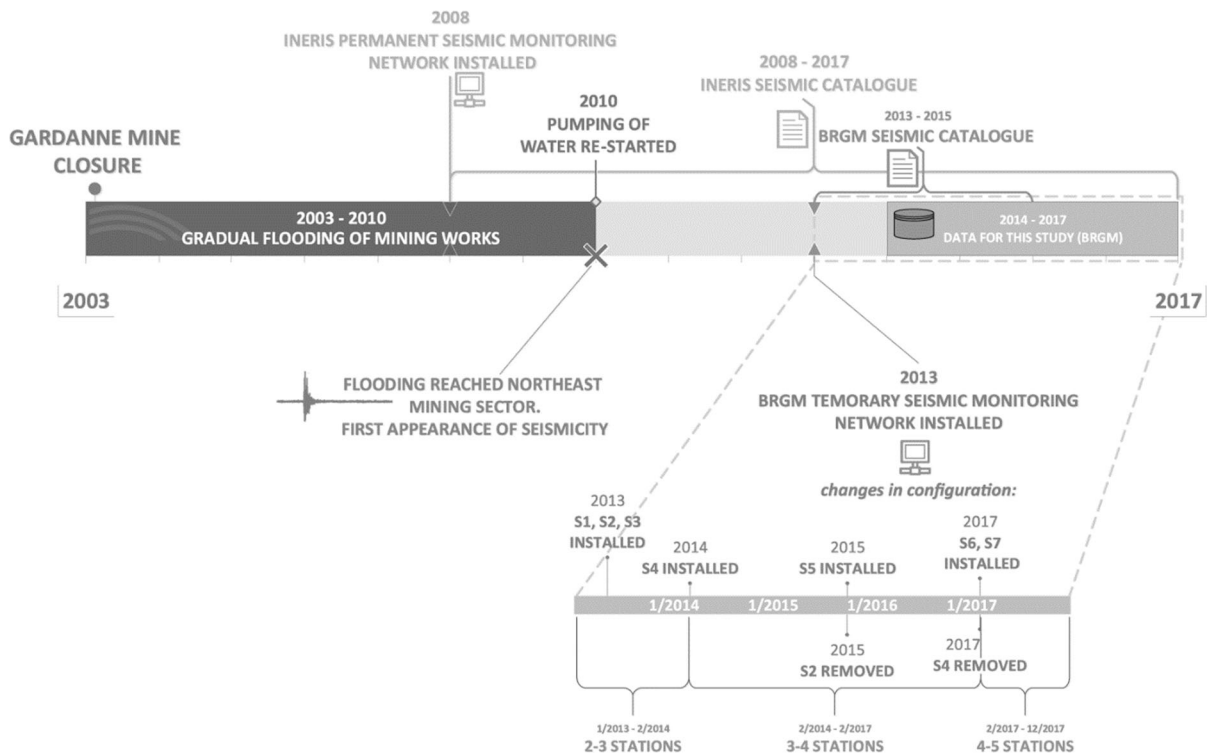


Figure 1

Timeline of important events in the study area of Gardanne mine, and available data for this study: starting from the mine closure in 2003, followed by turning off the pumps and gradual flooding of mining works. Following the Geoderis risk analysis, Ineris installed in 2008 permanent monitoring network in areas classified as having risk of ground instabilities. Flooding front reaches east part of mine in 2010 and soon after first seismic activity starts appearing in north east part of mine, outside of monitoring zones. Pumping of water restarted in 2010 to control the flooding front. Following again increase of seismic activity in 2012, BRGM installed temporary (sparse) monitoring network during 2013/2014. Zoomed in period 2013–2017 shows BRGM network station configuration changes in time. Data available for this study cover period 2014–2017. Ineris seismic catalogue is available for period starting 2008 until today, while BRGM seismic catalogue covers 2 years: 2014–2015

The presence of larger number of seismic multiplet families, which could indicate nearby interacting fault segments or repetitive ruptures on identical segments (seismic repeaters) resulting from surrounding aseismic creep, was also discovered (Kinscher, 2017; Namjesnik et al., 2019). Source mechanism analysis of the events indicated existence of normal faults located below mining works, striking NW–SE in coherence with the orientation of preexisting faults (Kinscher, 2017).

1.4. Objectives of this Study

For a better understanding of the geomechanical processes governing the seismicity and an eventual re-assessment of seismic hazard, accurate detection

and precise location of microseismic events are crucial. An automated detection and location process is necessary, due to the large volume of available data, with a main challenge of insufficient density of the seismic stations.

In this paper we successfully adapted and applied for the first time an automatic full waveform-based detection and location method BTBB, initially developed for dense seismic networks, to the continuous data recorded by a sparse seismic monitoring network in post-mining setting of Gardanne mine (example shown in Fig. 4).

To our knowledge this is one of the few attempts of applying the coherence-based full waveform methods to data of sparse networks, or application in post-mining microseismic monitoring, as it is the

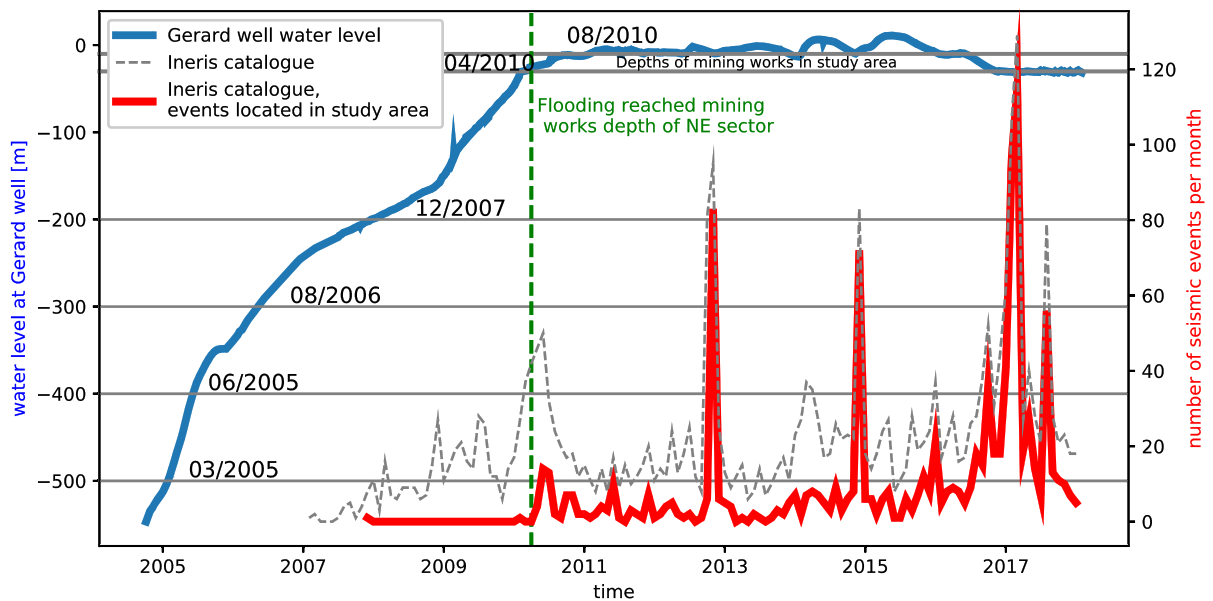


Figure 2

Gradual flooding of mine works and first appearance of seismicity as flooding front reaches mine works level at northeast area in 2010. Water level at Gérard well GW (blue line), corresponds to flooding front progress through mine works shown in Fig. 3a, together with time distribution of seismic events of Ineris catalogue in period 2008–2017 (all events shown as dotted line, events located in study area shown as full red line)

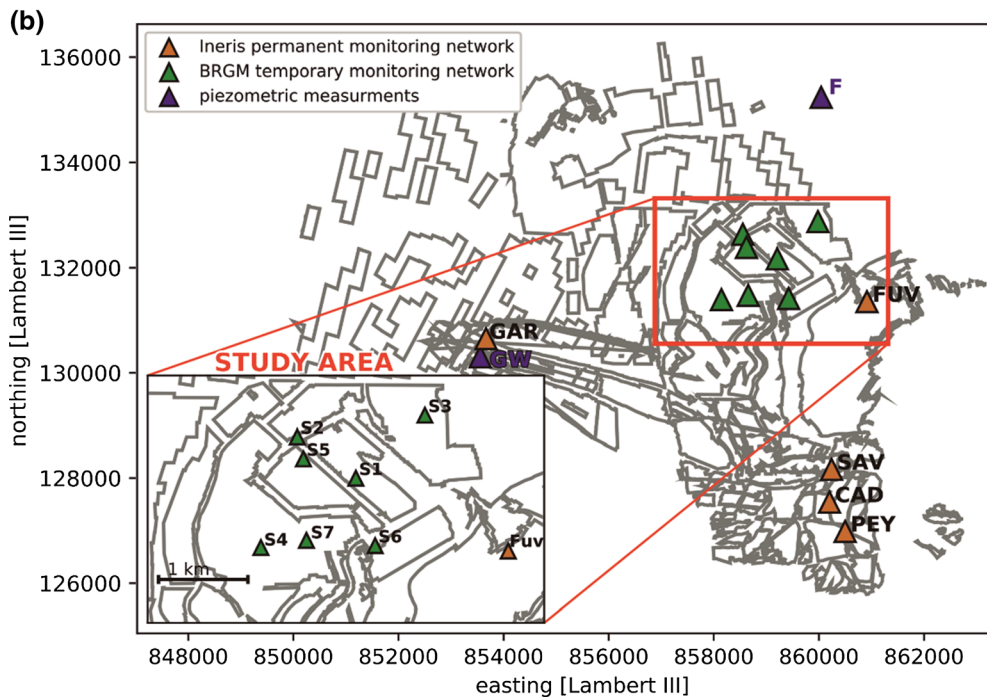
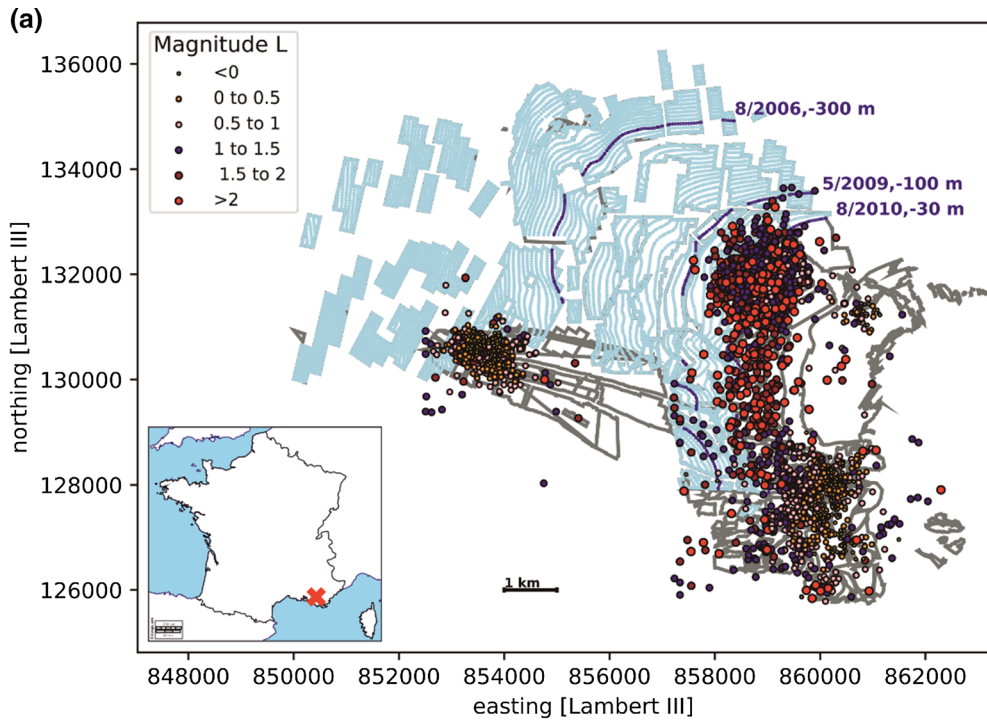
case in Gardanne, since the waveform-based methods are multi-station approach, exploiting the array-coherence of the recorded wavefield or its properties. Similarly to our application of BTBB to sparse network, another waveform method has been previously applied to the Pohang Earthquake sequence (Grigoli et al., 2018) with an almost equally sparse network.

Detection and location criteria of BTBB, which is designed to exploit the coherency of signals characteristics across the stations in order to detect and extract seismic event, represented a critical issue in our study. Due to similar coherency values of some noise sources and microseismic events in our data, it was not always possible to make a distinction between false and true event detection, using only the BTBB detection parameters such as MaxStack (maximum value of 3D likelihood source location function) and RMS (between the theoretical and observed time delay estimates), due to our inability of setting discrimination threshold values for those. As it can be observed from Fig. 5a, which compares the values of MaxStack vs RMS for all detections by

BTBB obtained from continuous data of December 2014 and for events in BRGM catalogue in the same period, the separation is not evident. Figure 5b shows the example of a trace where MaxStack and RMS have similar values in two 2-s windows (in grey), one with a clearly identifiable (high signal-to-noise ratio) seismic event, while the other corresponding to the incoherent noise. Therefore, the processing of 4 years of continuous seismic data of Gardanne mine network required the development of additional pre-BTBB processing steps to overcome this issue.

The developed methodology represents a novel approach for noise removal from continuous data, which also allows to significantly reduce the volume of managed data. A location quality-based classification system, followed by clustering analysis, was designed for the events in derived catalogue providing a clearer image of the active geological structures and allowing a preliminary interpretation of possible triggering mechanisms for seismicity.

This approach has a potential for implementation as (near) real-time operational seismic monitoring in the Gardanne site, as well as in others sites with



sparse networks. It can be particularly useful when monitoring network comprises only one-component instruments, as is often is the case for the mobile or

temporary deployments of task-force missions during the periods of stronger seismic activity.

◀Figure 3

Flooded abandoned Gardanne mine with observed seismicity and monitoring networks. **a** Flooded part of mine works shown in blue. Seismicity shown as a function of magnitude by coloured dots, observed by Ineris monitoring network in period 2008-June 2020. Flooding front progress through mine works is indicated by dark blue lines and corresponding dates of arrival. Gardanne mine location indicated by red cross mark on the inset map of France. URL source of original map: https://d-maps.com/carte.php?num_car=2812. **b** Gardanne mine area with permanent monitoring network installed in 2010 (Ineris) in risk area (orange triangles), BRGM temporary monitoring network installed since 2013 in study area (shown in zoom, green triangles), with shown stations active within period 2013–2017, and two available piezometric measurements: Gérard well (GW, purple triangle), considered as indication of water level in mining works, and Fuveau Regagnas well (F, purple triangle), indication of efficiency rain in the area

2. Processing of Gardanne Seismic Data

2.1. Data used in the Study

Data processed in this study are continuous recordings from the BRGM temporary seismic network, from period 2014–2017 (Figs. 1 and 3b). Figure 1 provides a detailed timeline of the BRGM station configuration changes during this period: from 3 initially available stations (installed in 2013) to 5 stations starting from the beginning of 2017. Shorter periods, during which some stations are not

operational (due to malfunction for example) are not shown.

The BRGM's seismic event catalogue (Dominique, 2016b) covers the period from June 2013 to 31st December 2015. In this catalogue, locations were determined based on standard techniques that include manual picking of P- and S- phases and iterative location approach of minimizing the difference between the observed and predicted arrival times at a number of seismic stations. Catalogue lists 756 events with magnitude range $-1.3 \leq M_L \leq 1.7$, among which 606 are located within the study area. The vast majority of events were located using the data recorded at the nearest stations. Estimated hypocentral depths are mostly several hundred meters beneath mine works (Dominique, 2016b). The data recorded during years 2016 and 2017 were unprocessed prior to this study.

In addition to BRGM's catalogue, Ineris's seismic event catalogue is available for the period 2008-June 2020. It contains 2–3 times less events than BRGM catalogue, for the overlapping period of 2014–2015, due to the larger distance of Ineris stations (Fig. 3b) from the study area. However, as it covers also the period of 2016–2017, analyzed in this study, when BRGM catalogue is not available, we use it for qualitative evaluation of our results.

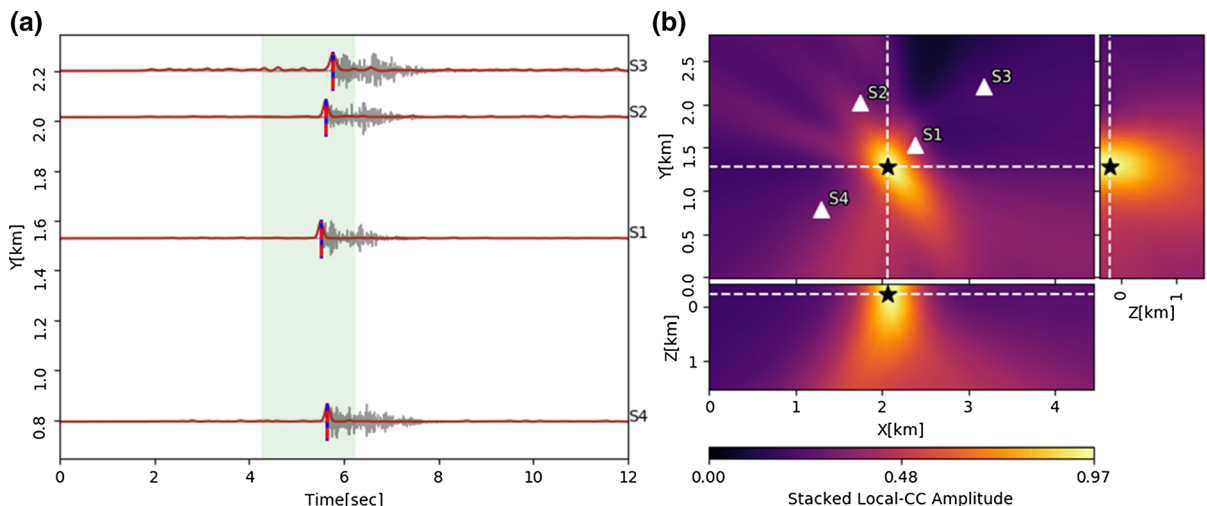


Figure 4

Example of a successful detection and location of seismic event as result of application of BTBB on extracted 12 s windows of vertical components of four available stations in study area, containing a seismic event **a** available traces (grey lines) with final broad-band multi-band frequency characteristic functions (brown lines). Vertical blue and dashed red bars indicate picked and theoretical P-wave arrival times from estimated location. **b** A horizontal and two vertical sections through the maximum of imaging function corresponding to determined location (black star) of an earthquake identified in the BRGM catalogue. White triangles correspond to station locations

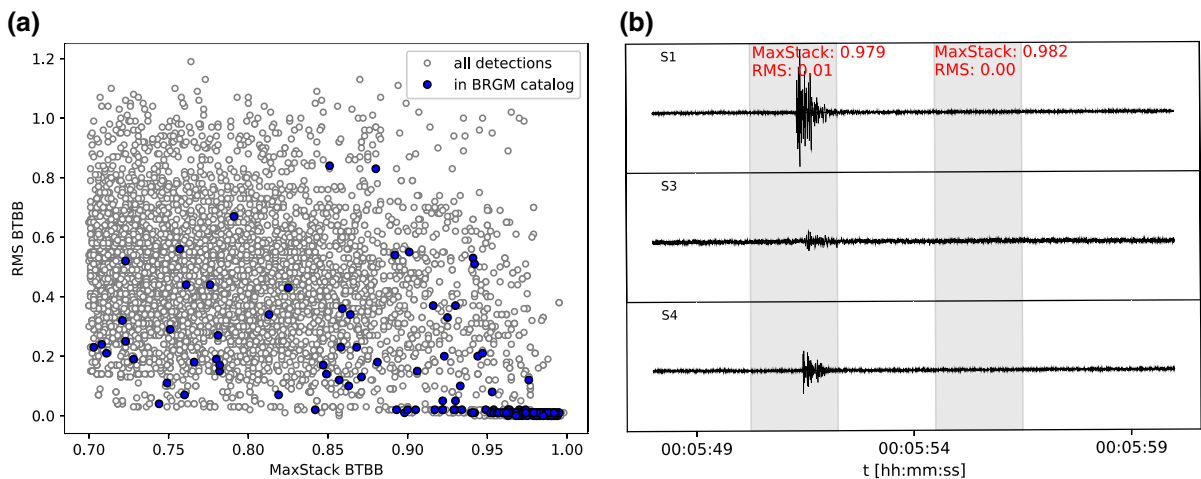


Figure 5

Challenges of discriminating between detected coherent noise and detected events when applying BTBB directly to continuous data of December 2014 testing dataset of 4 active stations, **a** parameters MaxStack vs RMS of all detections (white dots) and seismic events identified from BRGM catalogue (blue dots). **b** Extracted 12 s window of vertical component recordings for three available stations, corresponding to the detections shown in **a** with similar values of parameters in detection windows (grey areas), for seismic event and noise

In order to exploit the benefits of BTBB method automatic processing and refine the location accuracy, we developed pre-processing detection and location steps with implemented noise removal criteria, focused on reduction of noise in the data and minimizing the number of false detections. As shown in Fig. 6, the processing scheme is divided into detection and location steps, followed by a quality assessment scheme. This methodology follows a principle similar to the two-folded approach presented in (Palgunadi et al. 2019).

The first step comprises a noise robust, modified STA/LTA detection, applied directly to the time-continuous data in a moving time-window. Following it, the location step consists of an amplitude-ratio based location algorithm, followed by a relocation with BTBB method. An additional gain of this approach, besides noise-minimized detection and location of pre-BTBB processing is the reduction of a large volume of continuous data to a more manageable data volume. Lastly, a classification scheme based on event location quality assessment was designed allowing to distinguish between high and low quality events and locations, consequently simplifying the interpretation of results. The final catalogue comprises the epicentral location, the local magnitude, the seismic moment and moment

magnitude as well as a class assigned to each event based on an assessed quality of location.

The time-continuous data of December 2014 were used for method parametrization of detection and location steps. This month corresponds to increased seismic activity, for which BRGM catalogue that lists 214 events in the observed area is available for comparison. Four stations were active at that time and the network remained unchanged during this month. Developed methodology is further applied to full testing the data set recorded during the period of 2014–2017. Out of these data, period of 2014–2015 was used for development of the quality evaluation system, due to availability of the BRGM catalogue. Details of each step of processing scheme are described in the following paragraphs.

2.1.1 Step 1: Event Detection

In step 1, with the objective to detect potential events in the continuous dataset while simultaneously minimizing false detections, we introduce parameters that represent noise removal criteria within standardly used for triggering signal characterization method known as “short term average over long term average trigger” (STA/LTA) (Allen, 1982; Withers et al., 1998).

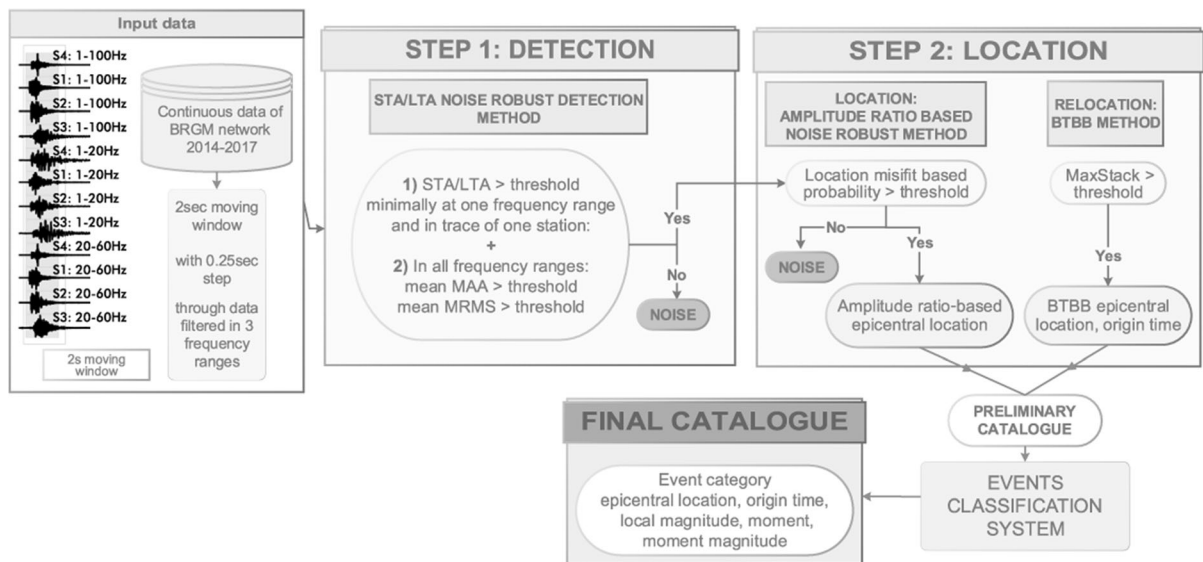


Figure 6

Scheme of the detection and location methodology developed in this study, comprising detection and location steps and event classification system based on location accuracy assessment. Maximum value of parameters such as absolute amplitude of STA/LTA (MAA) and maximum value of root-mean-square of STA/LTA (MRMS) are defined in Eqs. (1) and (2) (see Sect. 2.1.1 for details), parameter MaxStack represents maximum value of 3D likelihood source location calculated within BTBB method (see Sect. 2.1.2 for details). Final outcome is seismic event catalogue with event category, epicentral location, time of origin, local magnitude, seismic moment and moment magnitude

Only the vertical components have been used in analysis, to allow a wider application of developed approach to other post-mining environments to the temporary monitoring networks consisting of one-component geophones.

By visually inspecting the spectral content of known events in testing data set we identify dominant frequency range of the targeted events and divide our further analysis into three frequency bands: 1–20 Hz, 20–60 Hz and 1–100 Hz. Use of the multiple frequency bands allows a similar detection capacity for events with different, due to the source size and location, spectral content.

The filtered records (Fig. 7a) are transformed into characteristic function (CF) in each frequency band separately, using a recursive short-term average over long-term average (STA/LTA) algorithm (Withers et al., 1998) (Fig. 7b). Both STA and LTA window lengths are adjusted based on the frequency range. As a general rule, we decrease both STA and LTA windows with the increase of frequency. The ratio of the LTA and STA window lengths is always kept as 10 (e.g., Palgunadi et al., 2019).

To examine if an event is recorded on at least one station within the moving time-window, we check if the maximum of STA/LTA of at least one signal of any station in one of the frequency bands reaches the predetermined STA/LTA trigger value. Decision on trigger value of the STA/LTA CF is based on testing several different STA/LTA trigger levels over the entire testing dataset, taking into consideration the number of events of BRGM catalogue that were triggered versus total number of triggers. As we can observe in the example in Fig. 7b, this condition is reached in two different moving time-windows, for both noise source and event source.

To reduce false triggers, and further examine signal characteristics, we introduce two additional parameters: the maximum value of absolute STA/LTA CF amplitude (MAA) and the maximum value of root-mean-square STA/LTA CF (MRMS). The values of these parameters are calculated in 2 s moving time-window for CFs of each station, in all 3 frequency bands f_i separately. In general, the moving time-window length needs to be chosen based on the largest distance between the stations in the network to

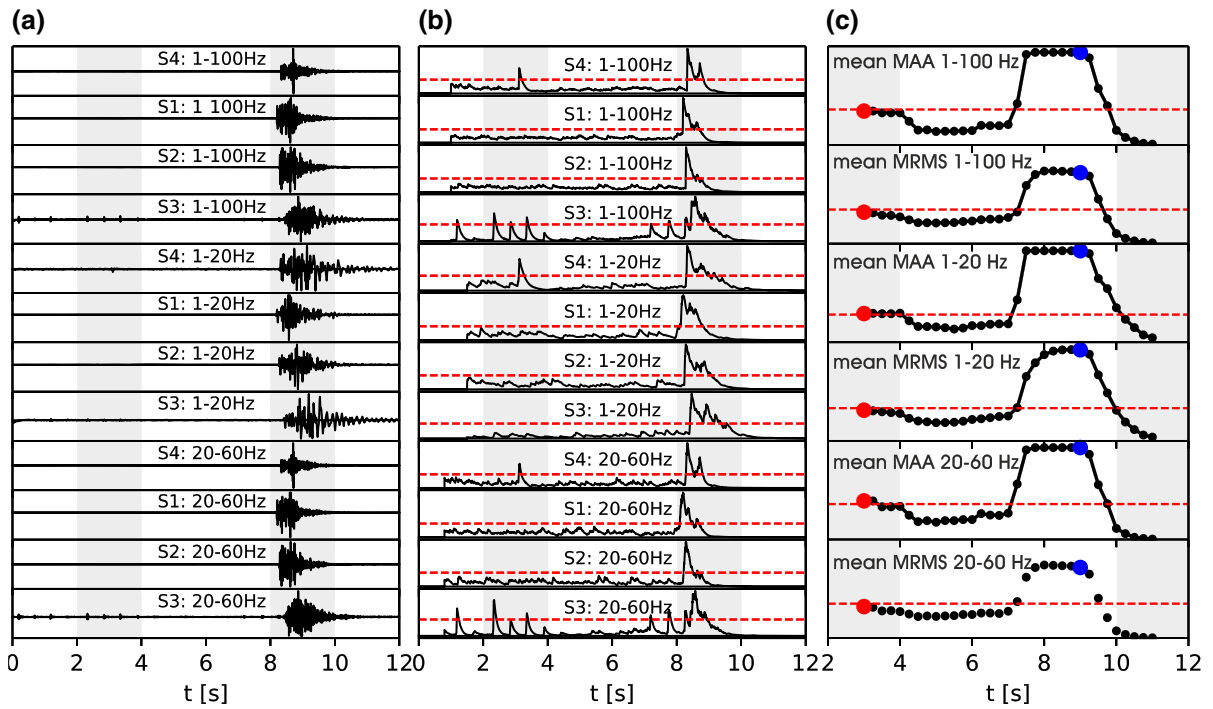


Figure 7

Example of step 1 processing for extracted 12 s window containing an identified seismic event from BRGM catalogue. **a** Vertical traces for each station filtered in three frequency bands. **b** STA/LTA CF of each filtered trace. Red line indicates pre-set triggering threshold level. Threshold is reached for noise (window 2–4 s) and for the earthquake (window 8–10 s). **c** Noise removal criteria parameters calculated in 2 s moving window (black dots, red and blue dots) with set threshold values (red dashed line). Noise check for windows whose CF was above threshold in **b**: values below threshold for noise (window 2–4 s, red dot), values above threshold for earthquake (window 8–10 s, blue dot)

account for the maximum expected time delay between the signals of all stations. Our choice of window length corresponds to approximately 1.5 of the estimated maximum expected time delay of the farthest station pair in our network (S3 and S4, with distance approximately 2.3 km between them) for S wave velocity of 1.8 km s^{-1} .

Finally, the criteria parameters to separate noise and event detections are derived as mean values across all stations and for each frequency range separately:

$$\text{mean MAA}(f_i) = \frac{1}{n} \sum_{\text{stations}} \max \left| \frac{\text{STA}(f_i)}{\text{LTA}(f_i)} \right| \quad (1)$$

$$\text{mean MRMS}(f_i) = \frac{1}{n} \sum_{\text{stations}} \max \left(\text{RMS} \left(\frac{\text{STA}(f_i)}{\text{LTA}(f_i)} \right) \right) \quad (2)$$

Choice of parameters MAA and MRMS for the noise removal criteria is based on the observation that seismic events in most cases have a higher energy and a longer duration recorded at multiple stations compared to noise, implying that the RMS envelope of STA/LTA function and the maximum of STA/LTA function also have distinctly high values for events, compared to the values for noise. It can be observed on the example in Fig. 7c, where points present values of parameters in Eqs. (1) and (2), calculated for each moving time-window.

Threshold values of parameters defined in Eqs. (1) and (2), for each frequency band f_i separately, are determined based on the values calculated separately for all previously known events of the testing dataset, already in BRGM catalogue (213 events), and for preceding noise time windows. As we can see in Fig. 8, in each frequency band, a representative range of values for both noise and

events can be distinguished. Nevertheless, due to the existing significant overlap in parameter values, decision on threshold selection is subjectively adjusted based on the desired outcome (similarly to setting of the STA/LTA trigger threshold, which can be either more noise-free dataset but with potential loss of weaker seismic events, or a dataset where loss of seismic events is prevented but higher number of false triggers remains).

The analysis of triggered 2 s time-windows in multiple frequency bands enables indirect the assessment of its spectral content. As the seismic events tend to have a wider spectral content compared to the more monochromatic noise sources, we declare the trigger as an event if values of parameters (1) and (2) reach the thresholds in all three frequency bands. Triggers are assessed as noise in case of a narrow-band detection and an inability to reach a threshold in all 3 frequency bands and are, therefore, removed from the dataset.

Values of parameters (1) and (2) for selected example of detected noise and event (shown in Fig. 7 and Fig. 8), confirm the accuracy of our noise removal criteria.

Further applying the described noise robust detection procedure to our testing dataset (1 month

of continuous data of December 2014) resulted in 603 potential event detections, with 94% of previously known events of the BRGM catalogue being correctly identified. Due to roughly three more times of detections in this step, compared to the 213 events in the BRGM catalogue within the observed area for the same period, as well as due to the unlikely possibility that such increase results from previously undetected events (even if we account for some detected but non located events due to low magnitude (Dominique, 2016b) or human error during manual processing), we performed a visual inspection of corresponding waveforms and confirmed that the detection still contains some unwanted triggers. Therefore, we further de-noise our detections and create a preliminary seismic catalogue in the next step, by introducing a second noise removal criteria based on the amplitude-ratio based location approach.

2.1.2 Step 2: Location

Amplitude-ratio based location approaches, derived from attenuation law (Battaglia & Aki, 2003), allow estimation of event location by determining amplitude ratio of station pairs (Taisne et al., 2011) and minimizing the error between expected and observed

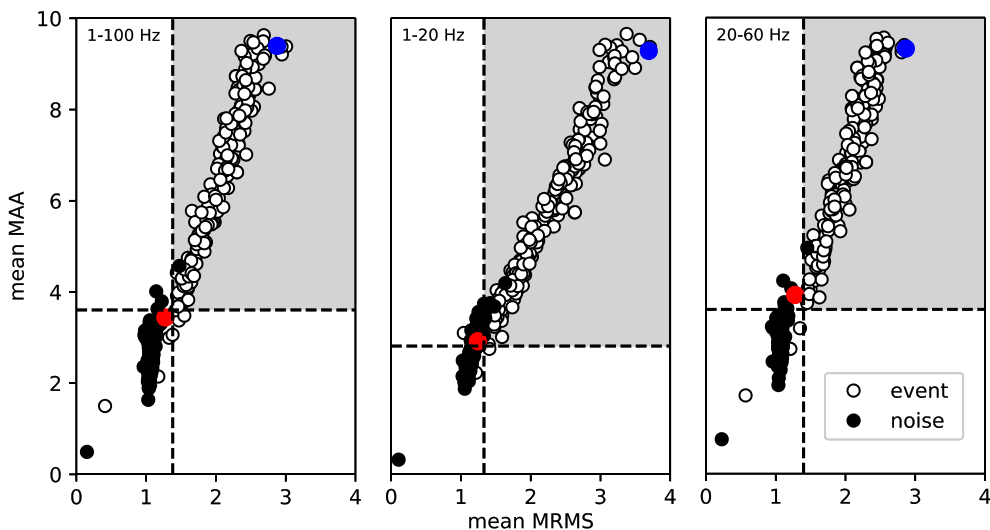


Figure 8

Setting threshold values in each frequency band for first noise removal criteria based on comparison of MAA and MRMS parameters of seismic events identified from BRGM catalogue and surrounding noise. Blue and red points correspond to detections marked by red and blue dots in Fig. 7c

values. The importance of this location approach, which is part of our methodology, is that it provides the mean to recognize and remove remaining triggered noise in the dataset resulting from previous step.

The main idea which lies behind the noise removal criteria that we introduce here is the significant difference in the values between the interstation amplitude ratios related to occurrences of noise and seismic event. These two types of signals manifest different attenuation behaviour due to differences in source location, as seismic events are located at depth and often not in direct proximity to one particular station, in contrary to noise sources which are mostly at the surface and close to one station only. Consequently, the noise sources will result in significantly higher amplitude ratio values between the closest station and the other stations, compared to amplitude ratio values of seismic event sources.

To determine the location of events, we follow the ideas of amplitude ratio based location method, previously applied to studies of salt solution mining microseismicity (Kinscher et al., 2015, 2016), and to mining induced microseismicity (Palgunadi et al. 2019). However, differences in environment of the present study allowed us to introduce modifications and somewhat simplify approach, following (Kinscher et al., 2017).

As events of testing dataset are concentrated in smaller area, in this approach we do not calculate site effect. Calibration with seismicity focused in one zone would be misleading and would improve location only for that area, while introducing the bias in other zones distant from the one used for calibration. Reducing number of parameters allows us to avoid wrong calibration.

We describe the attenuation law exclusively with a geometrical spreading, ignoring contributions from anelastic attenuation:

$$\log_{10} \frac{A_i}{A_j} = n \log_{10} \frac{r_j}{r_i} \quad (3)$$

where A_i and A_j are the maximum amplitudes of signals at stations i and j , r_i and r_j are the source-station distances of the same station pair. We assumed a body wave geometric spreading behaviour

with $n = 2$ (Lay & Wallace, 1995) which provided a good fit for observed amplitude ratios as a function of the distance ratios obtained from manual location of the BRGM catalogue (Appendix A1, Fig. 20).

Indeed, we confirm that the fit is equally good for all frequency bands, justifying our choice of frequency independent attenuation model.

The location of a detected potential event is determined based on the L1 norm misfit between observed amplitude ratios and theoretical amplitude ratios of each station pair and at the potential source point l on the defined grid.

Following expression (3), the theoretical (expected) values of amplitude ratios $A_{i,j,l}^{\text{theoretical}}$ for each station pair are estimated as a function of the inverse ratio of hypocentral distances, on a grid of predetermined (potential) sources:

$$A_{i,j,l}^{\text{theoretical}} = 2 \cdot \log_{10} \frac{r_{jl}}{r_{il}} \text{ for station pair } (i,j) \text{ at grid point } l \quad (4)$$

Grid of potential sources of each station pair is simplified and presented as a 2D plane fixed to the mine layer depths, while stations are distributed on the surface. Size of the grid was set to 4440 m (EW) \times 2770 m (NS), which frames the area of expected locations, based on locations of events in BRGM catalogue. Spacing between grid points (potential sources) was set to 50 m, which represents the horizontal resolution of this location method.

The observed amplitude ratio $A_{i,j}^{\text{observed}}(f_k)$ of a potential event for station pair (i,j) is determined as a ratio of the maximum absolute amplitudes, recorded at station i and j respectively, and filtered in the frequency band f_k :

$$A_{i,j}^{\text{observed}}(f_k) = \log_{10} \frac{\max |A_i(f_k)|}{\max |A_j(f_k)|}, \text{ for station pair } (i,j) \quad (5)$$

Observed amplitude ratios (Eq. (5)) for each station pair i and j are determined in a 2 s time-window corresponding to the trigger of the previous step.

For every point l on a defined grid of potential sources we determine the misfit between observed amplitude ratios and theoretical amplitude ratios of each station pair, expressed in a form of a PDF, for a

mean of the values observed over all station pairs, summing over the three frequency bands:

$$P(l) = \frac{1}{n} \sum_{f_k} \sum_{\text{stat}} e^{-|A_{i,j}^{\text{theoretical}} - A_{i,j}^{\text{observed}}(f_k)|/2}, \text{ at grid point } l \text{ pairs } i, j \quad (6)$$

where n represents number of station pairs. Location of the event is determined as the grid point l with the maximum value of $P(l)$. An example of event and noise time-window located with this approach is shown in Fig. 9.

The described amplitude-ratio based location approach was applied to all triggered time-windows from December 2014 that passed the detection procedure with noise removal criteria in the first step.

As expected, due to differences in amplitude ratio values for noise and events, the histogram plot of maximum values of $P(l)$ for all triggers (Fig. 10) reveals bi-modal distribution of values, interpreted as two normal distributions: the one on the left, with lower probability values, corresponding to a distribution of noise and the one on the right, with higher probability values, corresponding to distribution of seismic events, confirmed also by identifying the maximum of $P(l)$ values for events from BRGM catalogue, as shown in Fig. 10. We exploit this observation to define the second noise removal

criteria, by identifying the threshold value of $P(l)$ between events and noise. The threshold value in our case was determined as $P(l) = 2$ (Fig. 10) which allowed the elimination of the remaining false triggers from the dataset.

Events passing second noise criteria were further relocated by BTBB method, which consists of two main steps: signal processing, and detection and location. In the signal processing step, the raw waveforms are transformed to multi-band frequency characteristic functions (CF), represented by kurtosis, i.e., the fourth central moment higher-order statistics function (HOS). In the second step, extracted time series of CFs are used for determining event location by exploiting their coherency across all stations, and the estimated location is associated to the maximum of 3D imaging functions, based on stack of station-pair time-delay estimates projected to a grid of theoretical time differences of arrivals, for assumed velocity model (Poiata et al. 2016, 2018).

We configurate BTBB to calculate CFs in 50 logarithmically spaced frequency bands covering the same frequency range from 1 to 100 Hz as in part 1 of our analysis. The theoretical P-wave travel times necessary for second detection and location step are calculated using the Grid2Time routine of NonLin-Loc program (Lomax, 2005, 2008), over the grid covering the same horizontal extent as that for the

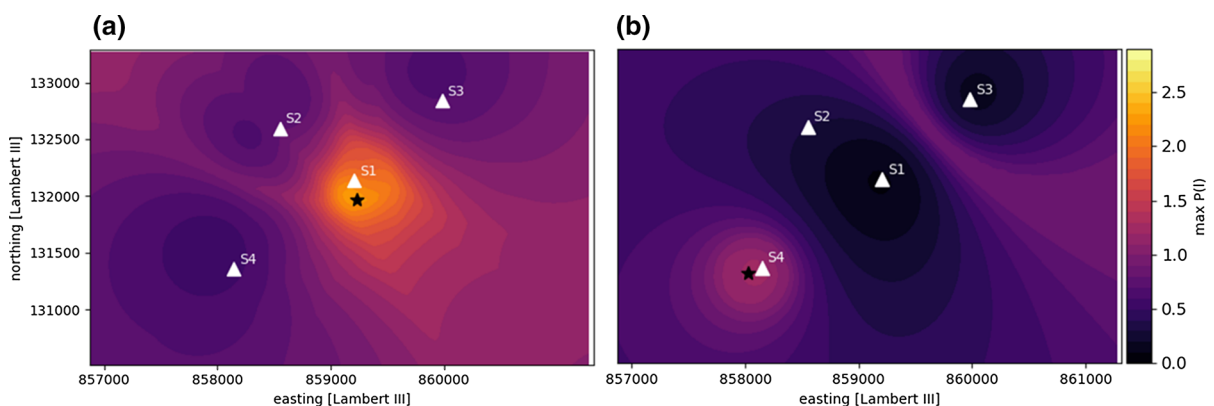


Figure 9

Example of amplitude ratio-based location method applied to the two detections passing first noise removal criteria. **a** determined location of event with $P(l)$ value above second noise removal criteria threshold **b** determined location of noise with $P(l)$ value below second noise removal criteria threshold

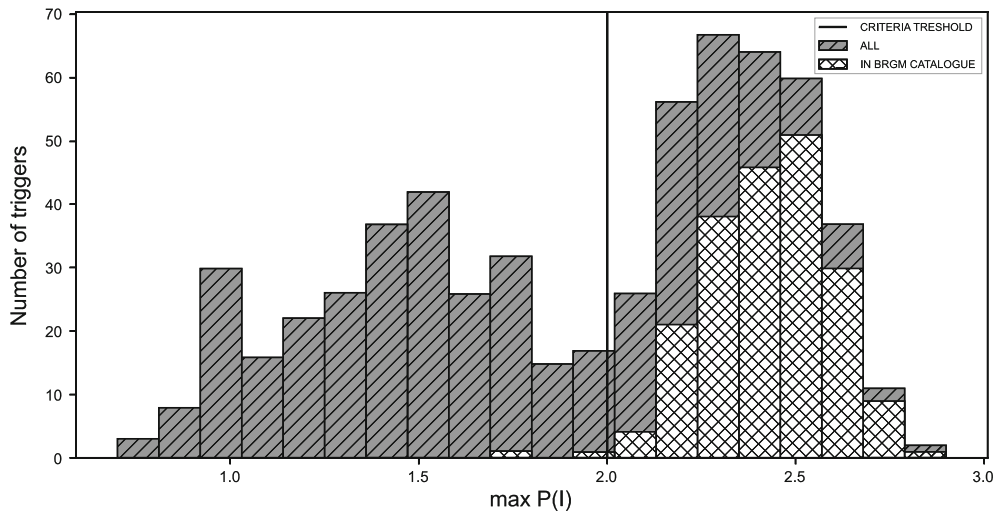


Figure 10

Bi-modal distribution of misfit probability $P(l)$ values between observed and theoretical amplitude ratios, for all detections during December 2014 that passed first noise removal criteria and are located by amplitude ratio-based method (grey bars). $P(l)$ values of seismic events from BRGM catalogue (white bars) allow setting threshold value for second noise removal criteria (black line)

amplitude ratio location approach, however with a denser, 10 m, spacing and depths up to 1.5 km, with a constant homogeneous velocity model $V_p = 4.1 \text{ km s}^{-1}$ (Kinscher, 2017). Minimum value of MaxStack parameter was set to 0.7, which presents the detection threshold criteria.

Comparing the events that were detected and located with our pre-BTBB steps, with the events of the BRGM catalogue for period of December 2014, we observe that 94% of the BRGM catalogued events were detected and pass the first criteria, and within those, 93% pass the second noise removal criteria. As expected, a small number of events from BRGM catalogue was left undetected (13 out of 213), due to values of their noise removal criteria parameters that was below the chosen thresholds, as can be seen in Figs. 8 and 10, or due to their origin outside of our defined grid. Finally, 314 events were located by the amplitude ratio-based method, out of which 200 are BRGM catalogued events, while 114 events represent new detections. The number of relocated events is somewhat reduced, not only due to noise in the data but also due to the data availability that was limited sometimes to only 2 stations. BTBB successfully relocated 177 events, 166 of which are in the BRGM

catalogue. Relocated with BTBB events have additional information of origin time.

Following the described steps, the developed methodology was further applied to the entire testing dataset, the continuous data of 2014–2017. Parametrization (such as threshold values, sliding window size, criteria values, grid size etc.) were kept unchanged. As described in Sect. 2.1, the network configuration varies in time as well as space during these 4 years and, depending on the observed period, consists of 2 to 5 stations, distributed over 7 different locations (Figs. 1, 3b).

Even though BTBB determines location based on a 3-D likelihood map, preliminary tests showed large errors in depth estimation, in comparison to BRGM catalogue. Due to the low number of stations, all of which are placed on the surface, dependence on assumed velocity model and use of P-phases only, we do not have a good depth resolution. Therefore, we disregard depth estimations and limit our catalogue to epicentral location only.

The reinforcement of monitoring network within the observed area with additional instrumentation during 2019, will enable a necessary application of additional approaches to allow for constraining of

source depths. One possible approach under consideration is the re-location of events with new developments of BTBB which implements S-waves into the analysis (Aden-Antoniow et al., 2020). However, due to the previously discovered presence of P-SV and SV-P converted phases within our data, probably caused by a layer of mine workings (Kinscher, 2017), this approach could prove to be challenging. New investigations will need to focus on avoiding misinterpretation of these phases as direct S waves.

Local magnitudes, moment magnitudes and seismic moments were calculated as well for every detected event. More details on this step are provided in [Appendix A2](#).

2.2. Event Classification and Location Quality Assessment

In order to assess the quality of the new seismic catalogue in terms of location accuracy, resulting location of 398 events in period 2014–2015, common to BRGM catalogue, are compared to BRGM location, using it as a reference. During this period, our catalogue contains total of 2154 detections, out of which 1691 are new events, not appearing in the BRGM catalogue. Within the same period, the BRGM catalogue overall comprises 606 events, out of which 131 are undetected with our approach.

As we can observe in [Fig. 11](#), the BTBB location shows qualitatively a better consistency with BRGM catalogue (especially for stronger events) than the amplitude-ratio based locations, which are more spread out.

As shown in [Poiata et al. \(2016\)](#), BTBB requires minimum three stations to determine the event location, meaning that, in our case, minimum three records need to have identifiable P-phase arrival, with high enough signal to noise ratio on CF's. Even though all our detections have records of minimum 3 stations, events are often buried in noise and not visible on every available record, which can lead to false locations.

In order to identify these low-quality events, we define event visibility criteria with STA/LTA function with threshold of 2, and we apply it to all available records for each potential event.

Further, as shown on [Fig. 12](#), we determine location errors as a function of event visibility. The location errors are based on the comparison of both amplitude-based location and BTBB location with matching locations of the BRMG catalogue, and defined as the Euclidean distance between them.

As expected, BTBB location error increases for low visibility events, ([Fig. 12a](#)) while it provides a better accuracy for events visible on a minimum 3 stations, compared to amplitude ratio based location ([Fig. 12b](#)). Amplitude-ratio based location, as expected, provides more robust location quality for events visible on 2 stations only.

We define events visible on minimum 3 stations as events of A and B class, with high quality BTBB locations. Events that are visible on less than 3 stations ($n = 26$) are defined as C class events, and amplitude ratio location is assigned to them, as a better quality location.

We further assess the location quality of BTBB for high quality events visible on 3 or 4 stations (A and B class), as a function of local magnitude and visibility, as shown in [Fig. 13](#). This allows to discriminate between the highest quality events visible on minimum 4 stations with local magnitudes equal or higher than zero, which we define as class A events ($n = 66$), and events of somewhat less accurate location, which we define as class B events ($n = 306$).

Based on the described observations, we designed an event classification scheme, shown in [Fig. 14](#).

3. Results of 4 Years of Continuous Data Analysis

As previously mentioned, the methodology was applied to the full available dataset of 4-year (2014–2017) continuous data from BRGM monitoring network. Finally obtained catalogue for this period comprises 4705 events.

The developed classification scheme allowed us to categorize all events in a new catalogue based on location quality. Out of 4705 events in the new catalogue, 364 were classified as A class events, 1624 as B class events and 2717 as C class events.

To evaluate the method performance, we first compare the temporal distributions and the number of

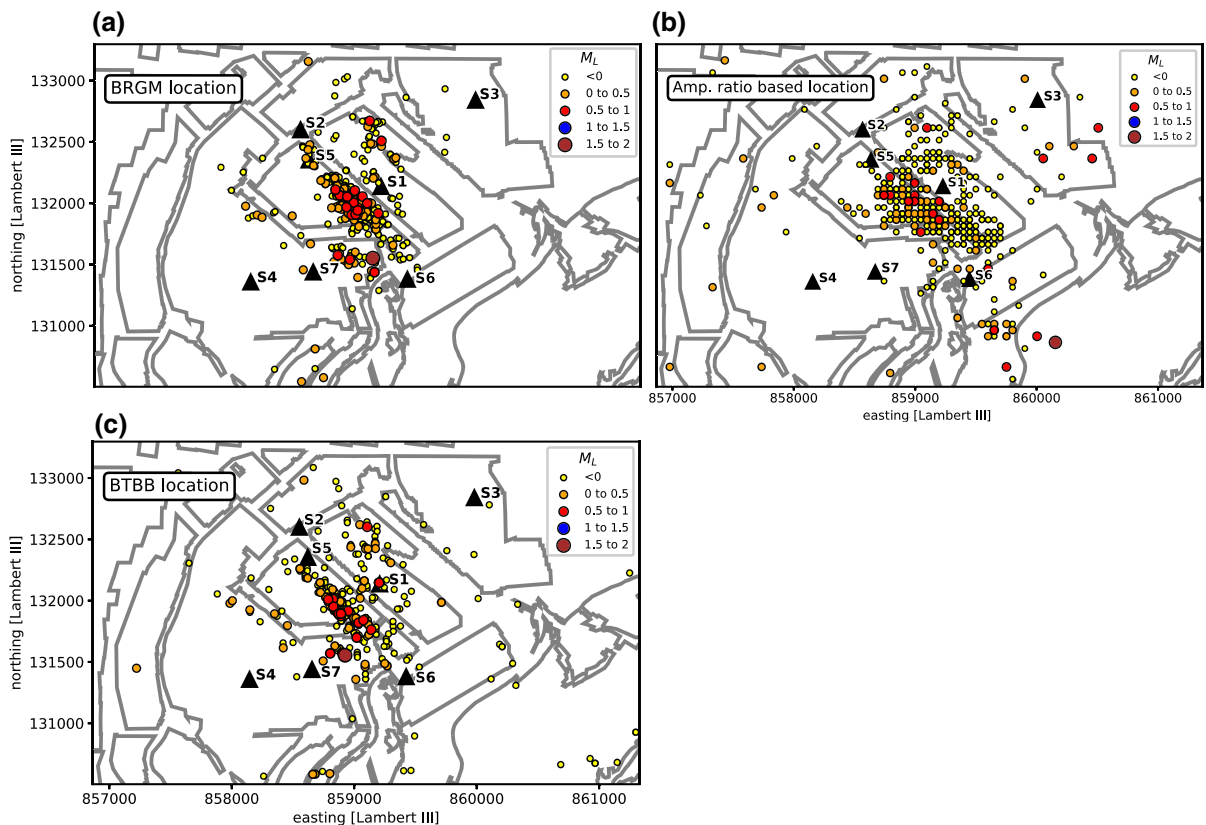


Figure 11

Comparison of locations determined in this study with 372 matching seismic event locations identified in BRGM catalogue during period of 2014–2015. **a** BRGM location **b** matching events located by amplitude ratio based method and **c** matching events relocated with BTBB

detected events of the new catalogue with previously available catalogues: the BRGM catalogue, available for period 2014–2015 and comprising 606 events within the studied area, and the Ineris catalogue for the entire period 2014–2017, comprising 790 events within the studied area.

As we can see on Fig. 15, for the two periods of increased seismic activity (late 2014 and late 2016 to early 2017) all catalogues are consistent, and the high quality A and B class events of the new catalogue provide a good match with both Ineris and BRGM catalogues, while indicating increased number of detections. However, low quality C class events of the new catalogue indicate two additional seismically active periods (middle of 2014 and middle of 2015). As these apparent increases of seismic activity have not been previously observed, it broadly implies that

C class events correspond to a mixture of small magnitude earthquakes just above the noise level with strong uncertainties in location as well as some remaining noise sources.

The spatial distribution of events of the new catalogue of 2014–2015, separately for each class, is shown in Fig. 16. As we can observe, a clustering of the highest quality events of class A are clearly visible. The class B events, even though showing somewhat more diffuse image, also indicate clustering. The spatial distribution of the C events is very diffused. These observations confirm the good functioning of our estimated location errors and classification approach. With our developed classification system, we are thus able to reliably assess the quality of the detected events and use it as the basis of

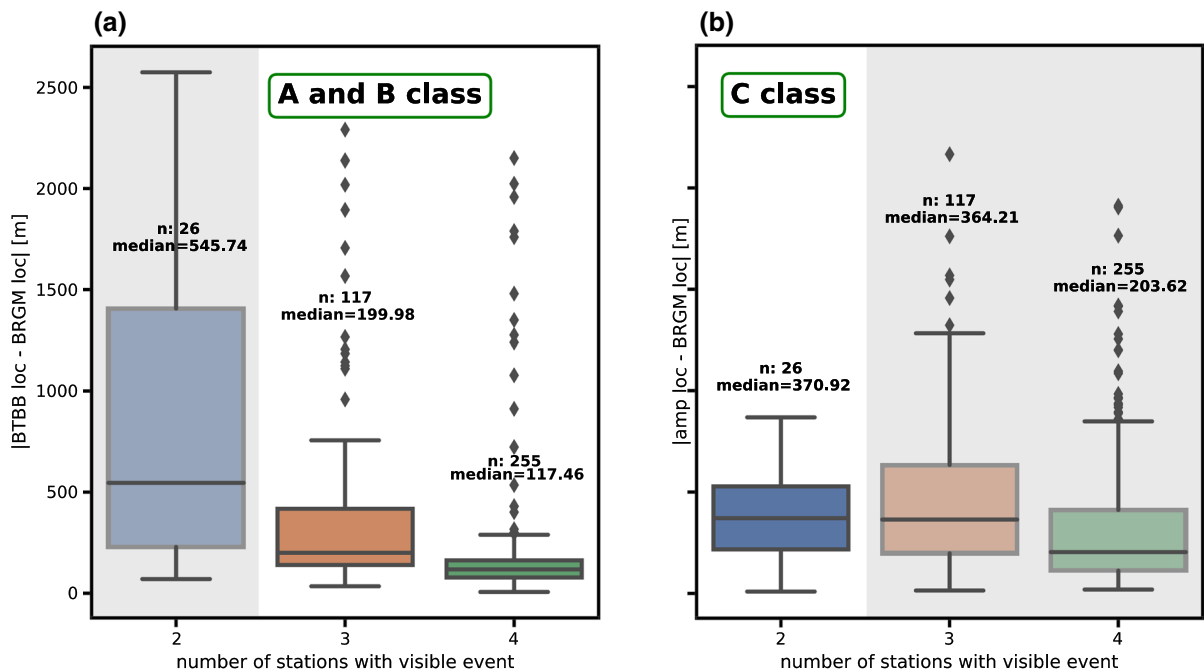


Figure 12

Identification of highest quality events based on location error as a function of event visibility. **a** Error of BTBB locations, **b** error of amplitude ratio-based locations. Events with high visibility (minimum 3 stations), are classified as A and B class with BTBB location providing better quality than amplitude ratio-based method (white area in **a**) and grey area in **b**). For events with lower visibility (C class events) BTBB error distribution exceeds 2.5 km in some cases (grey area in **a**), while amplitude ratio-based location quality is more robust (white area in **b**). The box shows the distance between the quartiles, with the median marked as a line, and the 'whiskers' show the extremes. Outliers, shown in the graph as separate points (diamonds), are the observations whose distance from the edge of the box (i.e. the quartile) is more than 1.5 times the length of the interquartile range. For each box, median and number of events in the corresponding category (n) are shown

subsequent interpretation of the new seismic catalogue (see following section).

4. Discussion

4.1. Connection of Seismic Activity with Water Level Variations in the Mine

As previously mentioned, Fig. 16 illustrates clear space clustering of events from the new catalogue, most clearly visible for the class A events. These observations are consistent with previous findings, as clustering of events was identified in the BRGM catalogue (Dominique, 2016b).

In order to separate, classify and discriminate clusters in an objective manner, we apply a K-means clustering analysis to the class A events, allowing us

to avoid subjectivity in assigning events to each cluster visually (for details please see Appendix A1). Figure 17 shows 6 identified clusters with corresponding centroids in A class events (a) and class B events (b) assigned to same clusters.

We observe also that clusters are spatially aligned with the direction of mine workings, which were oriented according to known discontinuities. Based on these findings we associate clusters to fault segments (Fig. 17c). Most seismically active clusters 1 and 5 are identified as part of one fault segment, clusters 2 and 4 as a part of second fault segment, while the position of clusters 3 and 6 correlate with position of previously known faults in area. This is in agreement with previous findings (Kinscher, 2017; Matrullo, 2015) showing that the seismicity origin is beneath the mine workings, with source mechanisms showing normal faulting.

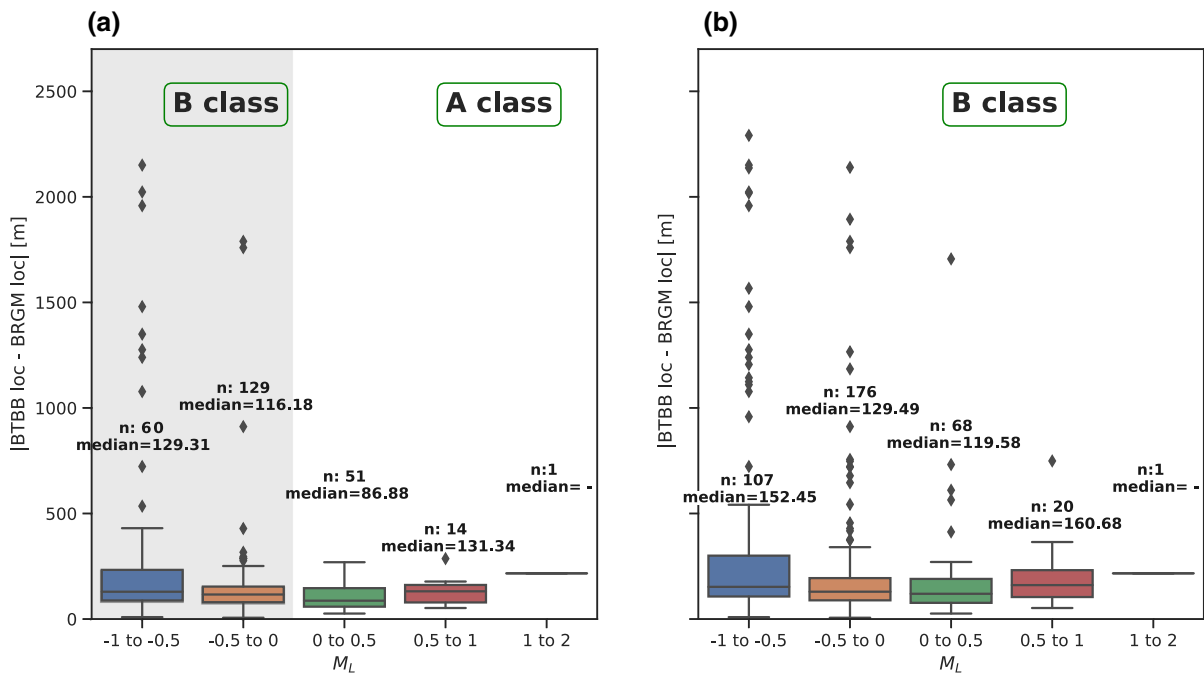


Figure 13

Error distribution of BTBB locations of class A and class B events as a function of event visibility and local magnitude distribution. **a** Error distribution for events visible on minimum 4 stations **b** error distribution for events visible on minimum three stations. Highest quality events (class A) are determined as events with visibility on minimum 4 stations and local magnitude equal or above zero. Remaining events (grey area in **a**) and white area in **b** are classified as B class. The box shows the distance between the quartiles, with the median marked as a line, and the 'whiskers' show the extremes. Outliers, shown in the graph as separate points (diamonds), are the observations whose distance from the edge of the box (i.e. the quartile) is more than 1.5 times the length of the interquartile range. For each box, median and number of events in the corresponding category (n) are shown

As mentioned in the Introduction section of this paper, apparent connection of seismicity with hydrogeological conditions has been previously observed in the studied area of the mine (Figs. 2, 3a).

The seismic activity of class A and class B events, as well as of each of the identified potential faults (clusters) separately is shown in Fig. 18, in comparison with hydrological data from the Fuveau Regagnas well (location shown in Fig. 3b).

We consider hydrological data of this well a good indication of the amount of efficiency rain due the shallow depth of the well, as it does not reach depths of mine works or known aquifers. We observe that strongest seismic activity of most clusters correlates with rainfall periods (Fig. 18), suggesting that rainfall presents a significant factor in the triggering and activation of faulting.

Indeed, seismic triggering from rainfall is known from other sites (Hainzl et al., 2006; Husen et al., 2007; Ogasawara et al., 2002). In general, fluid induced seismicity due to industrial activity is widely known phenomena, where increased fluid pore pressure causes reduction of fault strength which can potentially lead to fault rupture. Seasonal re-charges of aquifers and precipitation have also been correlated with seismic activity in several cases (Hainzl et al., 2006; Saar & Manga, 2003), and it was demonstrated that faults can be driven so close-to-failure that even tiny pressure variations associated with precipitation can trigger earthquakes at depths of few kilometres. However, the strongest activity of cluster 6 in the northwest is observed in a dry period, indicating the involvement of the other factors for the triggering mechanism of seismicity.

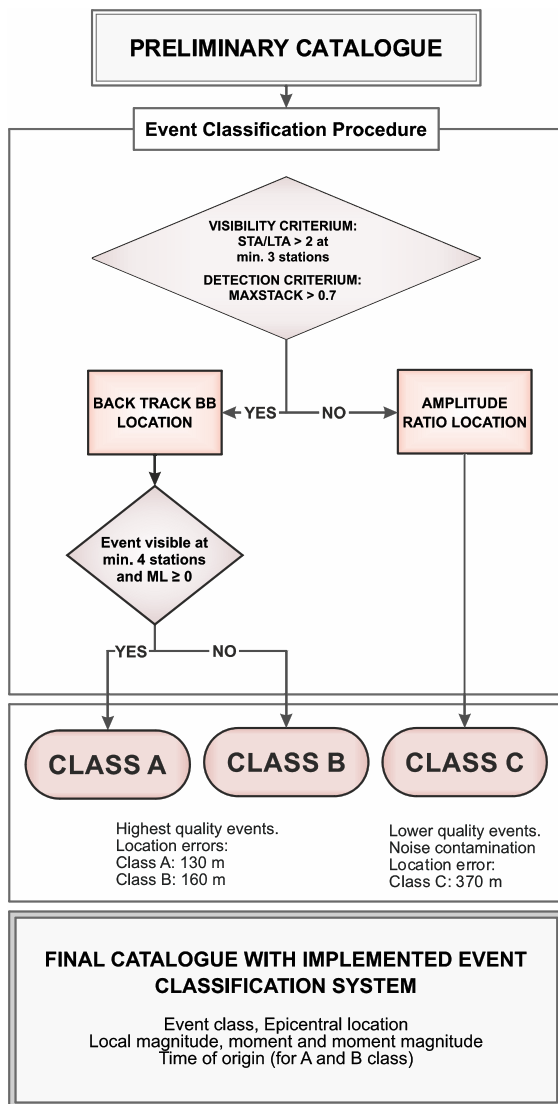


Figure 14

Event classification scheme based on location quality evaluation, applied to located events from new catalogue of 2014–2017

Spatio-temporal distribution of seismicity, and its connection to flooding front migration within the two key periods can be observed on Fig. 19. In the period 2014–2015, flooding front fluctuations were confined to shallower depths of mine works, and strongest seismic activity is observed in the south east (cluster 1, 2 and 3). Period of 2016–2017, where flooding front withdrawal can be observed, is characterized by still active clusters in the southeast with the increased seismic activity observed in the northwest (clusters 4,

5 and 6). As we can see, the strongest seismic activity seem to follow the withdrawal of the flooding front (dark blue and light blue lines).

However, when looking at the sequence of end 2016–early 2017 (Fig. 18 and 19), after a first moderate activation phase everywhere (with stronger cluster 4, green and 3, blue), there is a clear order of activation/migration from north to south, starting with cluster 6 (orange), followed by activity of cluster 5 (red), then cluster 1 (yellow) and cluster 4 (green), thus avoiding the cluster 3 (blue) activated just before.

The few weeks of delay between these activations might either reflect the time scale for build-up of stress redistribution from one primary structure to the next activated one, finally reaching its failure strength (for instance elastic stress build up due to the slow slip on a fault segment), or/and the time scale for decreasing the strength of the target structure (for instance through pore pressure diffusion and increase).

As going into further detailed analysis of clusters activity is outside of the scope of this paper, evaluation of our hypothesis resulting from this study and answering newly raised questions of triggering mechanism of seismicity during the dry season will be addressed in ongoing and future research.

The influence of the station configuration changes on this observation of migration of strongest seismic activity was analysed (details in Appendix A4), with the conclusion that it does not have any effect.

Seismic activity seems to be rather controlled by hydrological conditions in the mine works as well.

The flooding front position, indicating water levels in mine works, is influenced by seasonal aquifer level changes and controlled by pumping. Clusters in the northwest part are more active when pumping rate increased (period indicated in Fig. 18), which consequently lowered the water level in the mine works.

Focusing on the temporal distribution of each active cluster separately (Fig. 18), we see that most of the clusters have a repetitive nature, remaining active over a much longer period than the seismic crises itself, with events through the entire observed period of 4 years, which is in agreement with the repetitive nature of previously observed multiplet

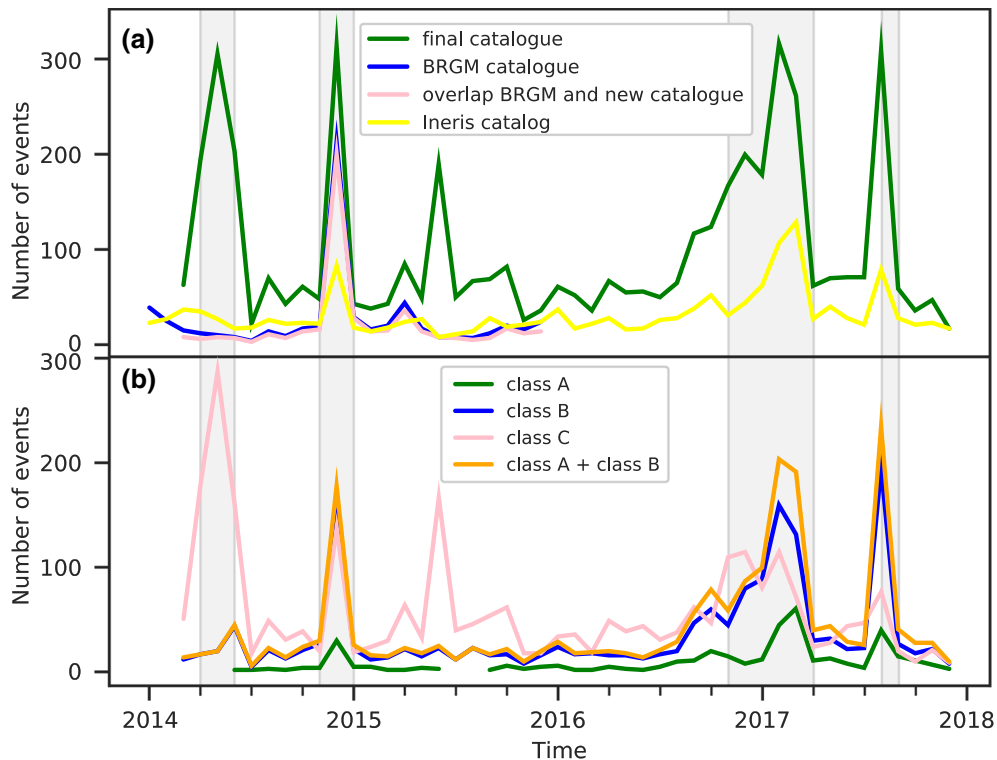


Figure 15

Temporal distribution of events in new catalogue 2014–2017 **a** comparison with temporal distribution of events in BRGM catalogue, available for period 2014–2015, and events of Ineris catalogue for period 2014–2017, both limited to study area boundaries. **b** Temporal distribution of events of new catalogue, separated in classes based on developed classification scheme

families (Namjesnik et al. 2019). This suggests that the source of seismicity originates from the repeated rupture of faults (or faults segments), rather than from ground instabilities caused by a failure initiated at the mine level. Progressive mine collapsing, which is often characterized as “the domino effect”, is characterized as spatially progressive and non-repetitive, in contradiction with what we observe here. Furthermore, the observed values of magnitude appear rather strong for a collapse of a mine of 2 m height.

4.2. Outlook

The exact mechanism behind the apparent coupling of seismicity and water level changes, in particular (un)loading effects of the water columns, pore-pressure changes, stress redistribution from

(a) seismic faulting in the mine, remain beyond the scope of the present study. It will be addressed by a recently started analysis, based on an improved seismic network. This will allow for source mechanism analysis and source depth information extraction as well as better understanding of water circulation in the mine based on installation of further piezometers in the later stages.

As the monitoring network in the study area has been enhanced since 2018 and 2019, now comprising total of thirteen instruments, in the future study we will exploit the newly recorded data of 2019 to better constrain seismic source mechanisms and source depths of identified faults segments. To evaluate the hypothesis of connection to the hydrogeological system, we will exploit the newly enhanced hydrogeological monitoring data covering the same period, as two new piezometers were installed in pre-existing

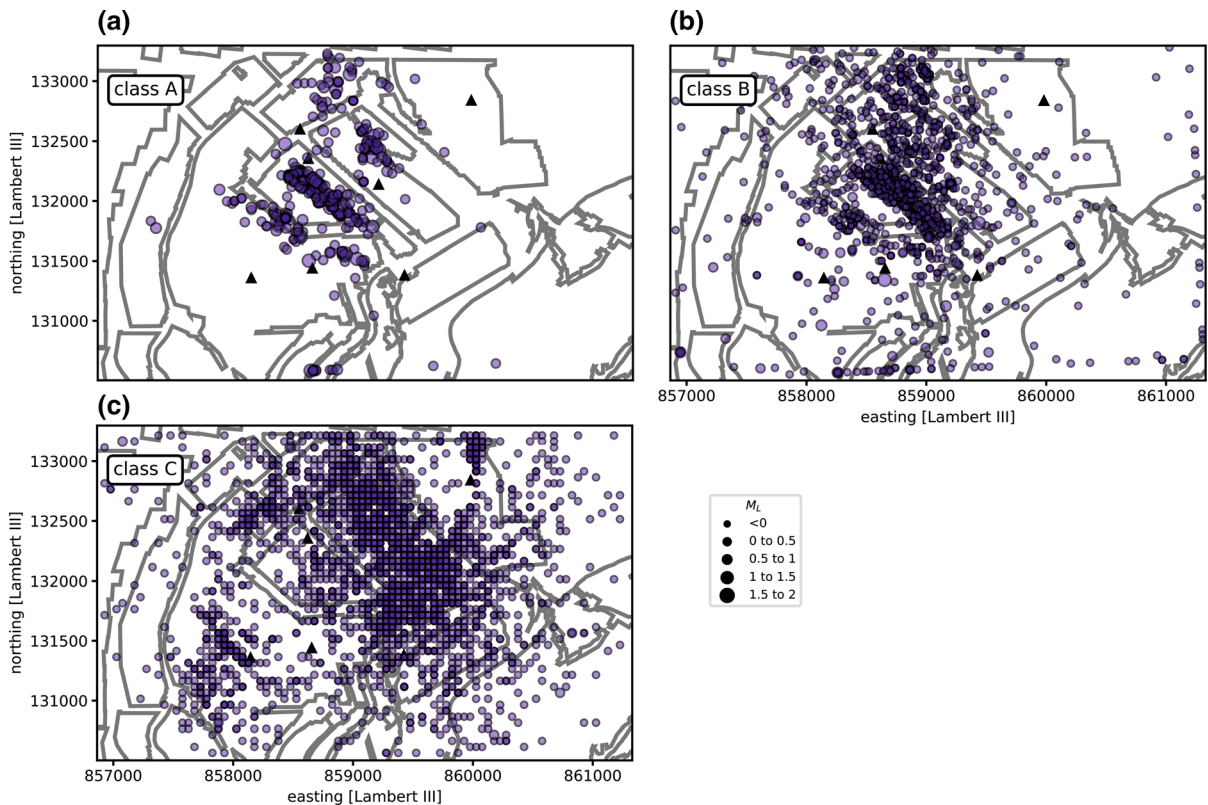


Figure 16

New catalogue 2014–2017 with events categorized and separated based on developed location quality based classification scheme **a** class A, 364 events, **b** class B, 1624 events and **c** class C, 2717 potential events

wells in summer of 2018. The hypothesis of physical mechanisms for seismicity triggering due to water level changes in mining works will be evaluated with simple geomechanical modelling.

More detailed analysis of the seismic activity within each fault segment will focus on exploiting the previous discovery of a large number of multiplet families (Namjesnik et al. 2019), located within the same clusters that were identified in this study. The precise spatial dimension of each cluster (fault segment), and the identification of whether they represent separate faults or one continuous fault segment rupturing during longer period, as well as whether a large dynamic rupture is possible or rather slow creep with breaking of asperities, will allow us to estimate the potential highest magnitude of future events and to assess the potential hazard in the observed area.

5. Conclusion

In this paper we presented a first successful adaptation of the automatic waveform-based BTBB method to post-mining setting of the Gardanne mine, monitored with sparse seismic network. As the direct application of BTBB to data with limited number of stations was challenging, we developed a novel methodology that bypasses these limitations making use of two noise removal criteria.

Furthermore, we designed an event classification system based on the evaluation of location quality, which demonstrated a satisfactory performance and enabled us to observe and classify the clustering of highest quality (A class) events.

We show preliminary interpretations of spatio-temporal cluster analysis and correlation with hydrogeological data, which indicate that seismic activity in Gardanne mine is controlled by

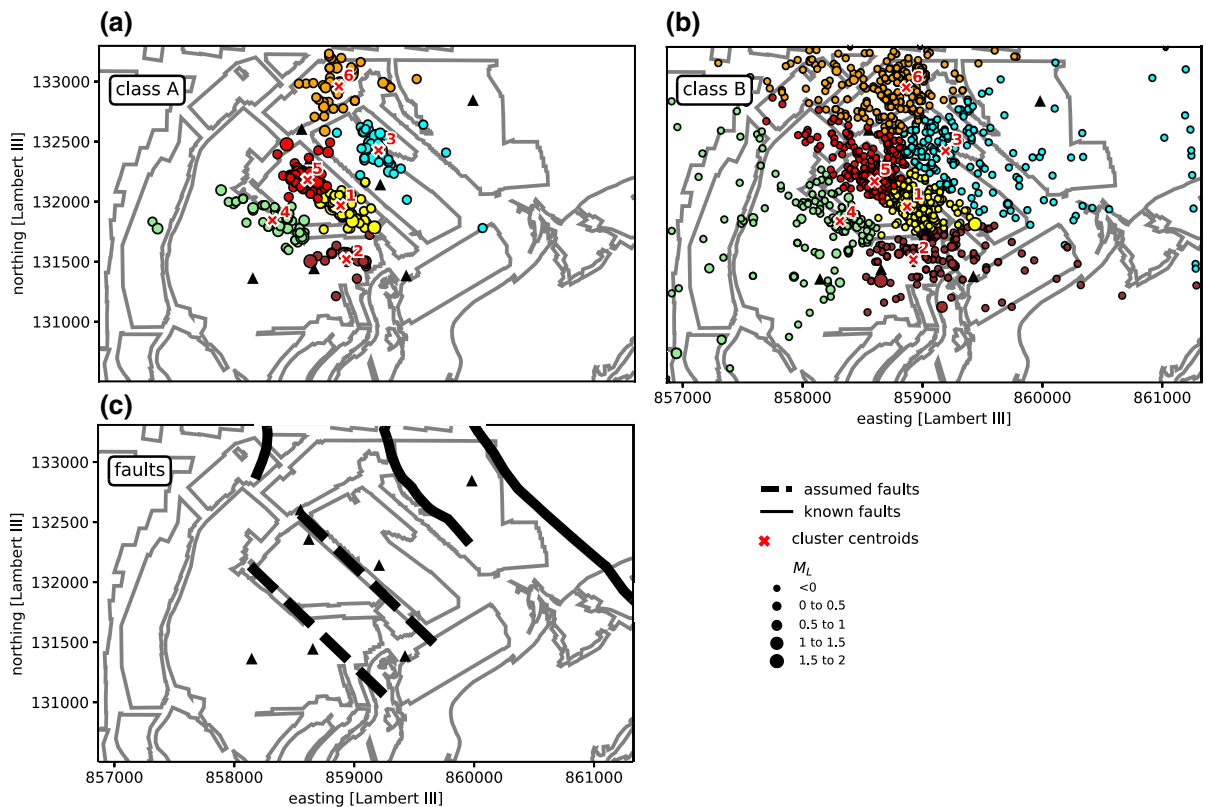


Figure 17

Results of cluster analysis of highest quality events of new catalogue 2014–2017 **a** K-means applied to A class events **b** class B events assigned to same clusters and **c** assumed fault locations and known faults locations

hydrological conditions in mine works which acts as an anthropogenic aquifer, influenced by rainfall and pumping. The interpretation of the results indicates that the origin of seismic activity is on fault segments, rather than due to mining works collapse, which is in agreement with prior hypothesis (Kinscher, 2017).

In conclusion, the methodology presented here offers a solution for automatizing detection and location in operational microseismic monitoring, especially, but not limited, to post-mining settings, where sparse temporary monitoring networks with one-component geophones are very often the only available tool for hazard assessment, and where timely decisions based on accurate seismic event locations are of most importance.

Acknowledgements

We thank the editor Ulrich Achauer as well as Francesco Grigoli and another anonymous reviewer for their constructive and thoughtful comments and suggestions which helped improve this manuscript. Part of the data analysis was implemented using ObsPy (Beyreuther et al. 2010, Megies et al. 2011) and Seaborn (Waskom et al. 2018). Figures were prepared using Matplotlib (Hunter, 2007) and Seaborn.

Funding

This work has been conducted in the framework of the scientific collaboration of Ineris and the BRGM and has been partially supported by the SERA project, grant agreement No.730900, EU H2020.

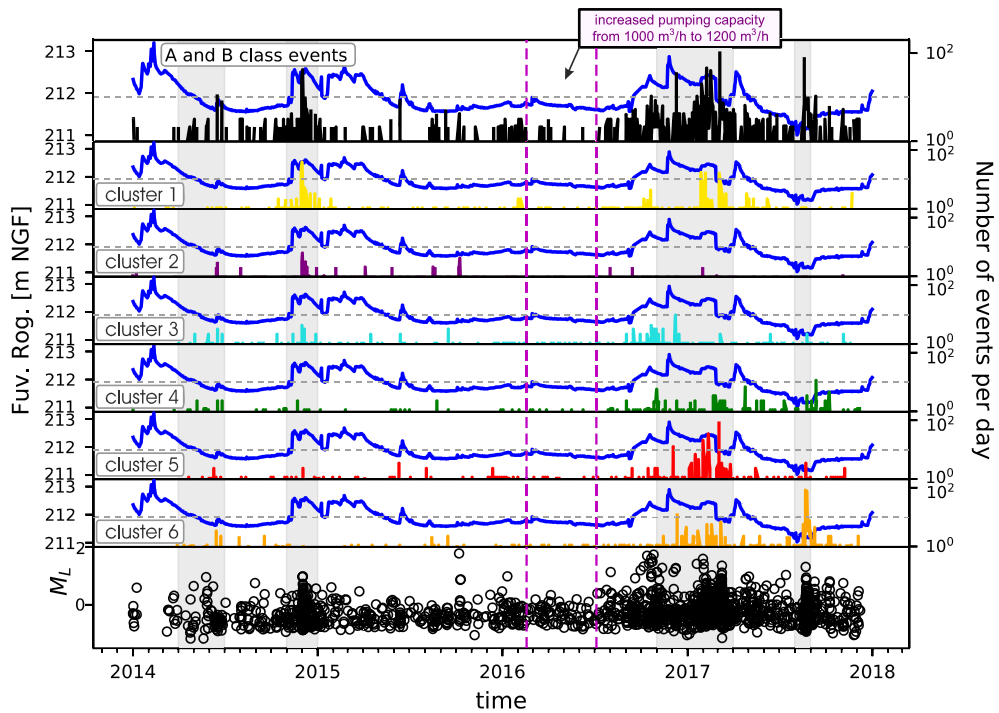


Figure 18

Temporal distribution of seismic activity for A and B class events per day (top figure), and each cluster separately (figures below), compared with water levels at Fuveau Regagnas well (F on Fig. 2b) presenting efficient rainfall (blue line). Bottom figure shows local magnitude of corresponding events

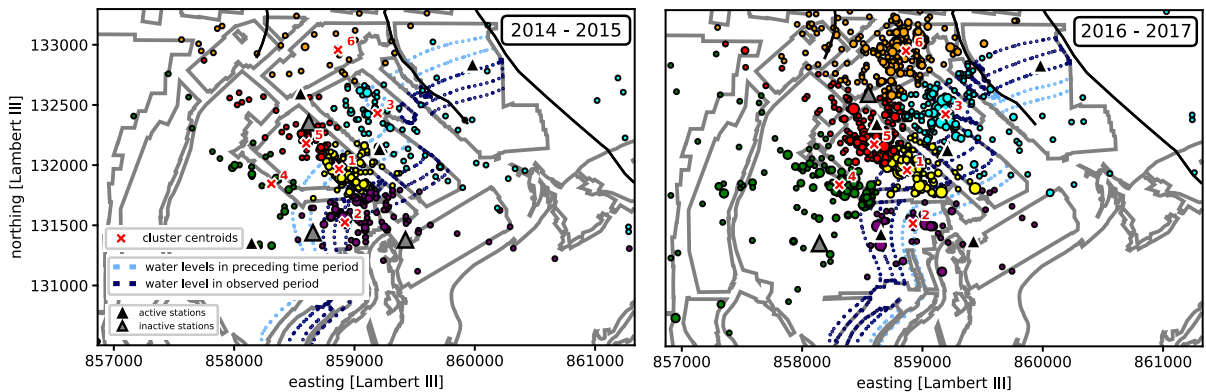


Figure 19

Spatio-temporal distribution of seismicity within two periods **a** 2014–2015, **b** 2016–2017. Station configuration changes are indicated: active station within observed period are shown with black triangles, while grey triangle are non active stations. Flooding front migration is shown with front position: dark blue indicates levels in this sector in observed period., while light blue indicates flood front position in preceding period

Availability of data and material (data transparency)

Data used in this article are properties of Ineris and BRGM.

Code availability (software application or custom code)

The code corresponding to new development will be available upon request. The software for the full-waveform detection and location BackTrackBB is available from Git-Hub (<http://backtrackbb.github.io>).

Declarations

Conflict of interest We declare that we have no competing interests.

Appendix A1: Attenuation Relationship

We calibrate expression (3) assuming $n = 2$, the typical geometrical spreading coefficient used to describe body wave propagation. This choice presents a good approximative fit for observed values as evidently shown in Fig. 20, where testing dataset of 215 catalogued events of December 2014 was used.

We observe also that the frequency range does not affect significantly the values of logarithms of amplitude ratios of each event plotted versus logarithms of ratios of their inverse source-station distances, indication that the frequency dependent terms from attenuation law could be omitted.

We expect the effect of the mine layer to be relatively small since the analysed wavelength ($V/(4*f)$) (for amplitudes in the frequency range of 1–100 Hz) are generally larger than the thickness of the mine layer. Nonetheless, scattering effects are clearly expected to be present and have an influence on the recorded amplitudes, which however is below the uncertainty of ± 1 for estimated amplitude ratios used here. However, for the purpose of calculating amplitude ratios in frequency range of 1–100 Hz, based on observations, we consider only the loss of seismic wave energy due to the geometric spreading.

Appendix A2: Local Magnitude, Seismic Moment and Moment Magnitude Determination

To complete the information in the new catalogue, a local magnitude was determined for each detected event as well, as a product of the absolute amplitude

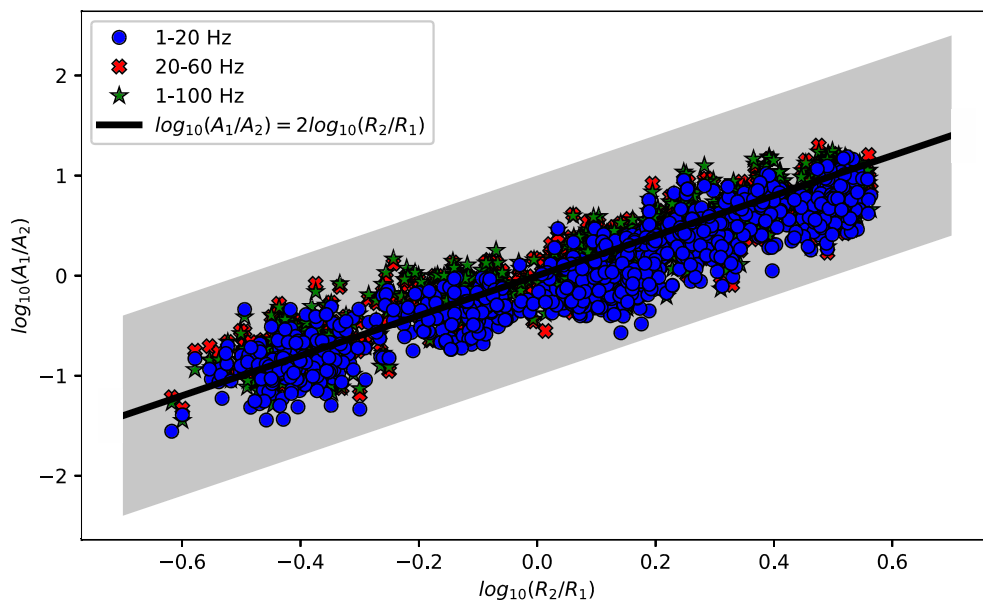


Figure 20

Logarithm of amplitude ratios without frequency dependent terms, shown in three different frequency ranges. Grey area corresponds to intercept of ± 1

of an event in frequency range 1–20 Hz, taking into account all events listed in the new catalogue, and source-station distance r averaged over number of the available stations.

To calculate the source-station distance, the BTBB location is taken for events successfully relocated in part 2, otherwise it is taken as the amplitude ratio-based location of part 1 of the analysis. The local magnitude is expressed as follows:

$$M_L = \log_{10}(\text{mean}[(\max|A| \times r)]) \quad (7)$$

and the determined values are in range – 1.38 to 1.7.

The relationship between moment magnitude M_w and local magnitude M_L for the study region was investigated and estimated empirically based on data of 48 events of December 2014 for which values of M_w were available from a precedent study and M_L determined in this study based on Eq. (7):

$$M_w = 0.68 \times M_L + 0.57 \quad (8)$$

This expression was further applied to calculate the moment magnitude for all detected events. The seismic moment M_0 was determined based on the relation (Hanks & Kanamori, 1979):

$$M_w = \frac{2}{3} \log M_0 - 6.1 \quad (9)$$

Appendix A3: K-Means Clustering

In order to avoid subjectivity due to the difficulty of accurately assigning events to clusters “by eye”, due to absence of other information on seismic source such as focal mechanism, faults locations etc., we apply the K-means clustering approach (Hartigan, 1975; Hartigan & Wong, 1979).

K-means is a clustering algorithm with Euclidean square distance metric, with hard partitioning algorithm based on a central point (centroid) where each event is allocated to only one particular cluster, in an iterative process where the data are initially randomly partitioned, the mean position of each group calculated, and then the data partitioned again by allocating each event to its nearest mean cluster position (Weatherill & Burton, 2009). K-means has

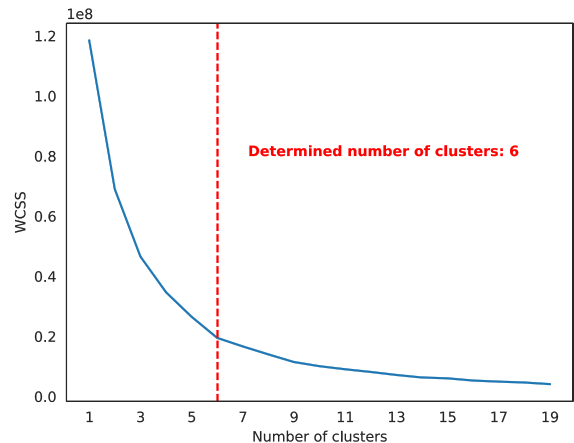


Figure 21

Elbow method used to determine optimal number of clusters (k)

been most widely used for cluster analysis in data mining and analysis field as an unsupervised machine learning technique, but has several recent applications on earthquake data as well (Novianti et al., 2017; Rehman et al., 2014; Weatherill & Burton, 2009). The best K value (which corresponds to the number of centroids and number of clusters) in correspondence to our observation in Fig. 16 was estimated as 6 by the Elbow method as shown in Fig. 21.

Appendix A4: Influence of Station Configuration Changes on the Location of Events and Apparent Migration of Seismicity

Monitoring network configuration could potentially have biasing effect in locating event. Due to station changes in time during observed period (timeline of changes shown in Fig. 1) it is necessary to examine its influence on apparent spatio-temporal migration of seismicity between two seismic crises (Fig. 22).

As it is not possible to conduct test by determining location of one event in the different configuration settings, we examine the influence of station configuration change on event locations by focusing on seismic clouds locations of A class events during three periods of seismic crises, in between which the station configuration had changes in settings.

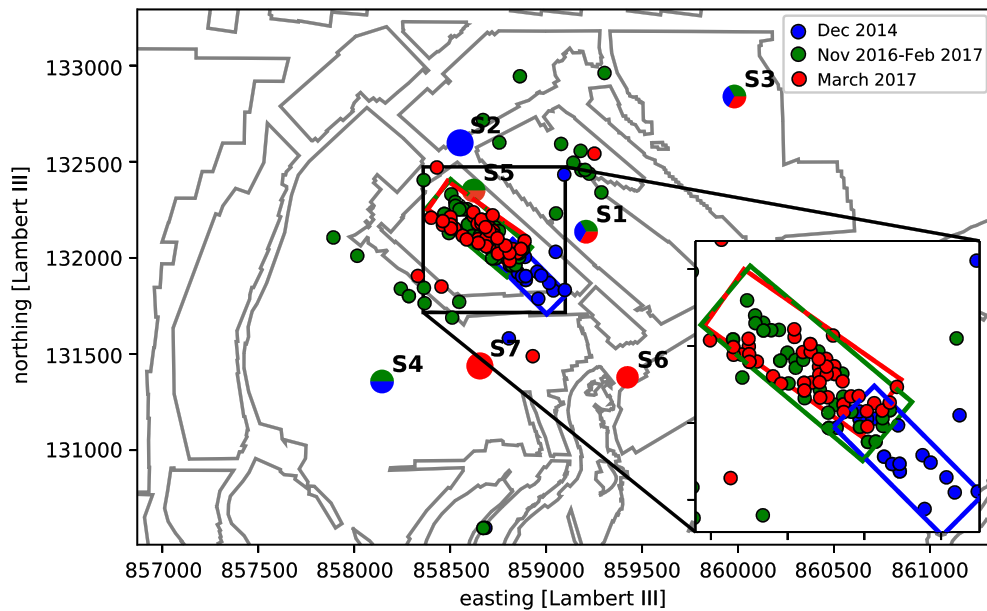


Figure 22

Station configuration change during periods of three seismic crises and its influence on spatio-temporal migration of seismic cloud in active area. Colors of seismic events, seismic stations and seismic cloud boundaries are corresponding to colors of observed periods shown in label. Monitoring stations are marked with name of stations and slightly bigger circles than seismic events, colored with 1 to 3 colors indicating their active period(s). Seismic clouds area are bounded with rectangle corresponding to its approximate area

First, we focus on a seismic crisis which lasted from late 2016 until March 2017. Within this period, the station configuration had a major change prior to the seismic activity in March, as station S4 was removed and two new stations (S6 and S7) were installed. Therefore, we separately observe the seismic cloud locations during one configuration setting (November 2016–February 2017) and during a new configuration (March 2017). As we can observe on Fig. 21, the location of the events during the entire crisis period does not change albeit the station configuration changes, there is no migration of seismicity. Hence, we conclude that this configuration change did not have any effect on determining of location of seismic events.

However, a migration is noticeable between the crisis of December 2014 and the crisis of November 2016–March 2017. To analyse the influence of the station configuration on this migration, we observe that the configuration of the crisis of December 2014 has almost the same station configuration as the first part of the observed crisis of November 2016–

February 2017, as the only change is the station transfer from S2 to S5 location.

Based on these observations, we conclude that station configuration changes do not have any meaningful influence on the location of events of class A and that migration of seismic cloud between two crises is not an artefact due to the change of station configuration.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Abdul-Wahed, M. K., Senfaute, G., & Piguat, J. P. (2001). Source location estimation using single station three-component seismic date. *Rock mechanics : a challenge for society : proceedings of the ISRM Regional Symposium Eurock 2001, Espoo, Finland, 4-7 June 2001*. pp. 783–788. ineris-00972215. Retrieved from <https://hal-ineris.archives-ouvertes.fr/ineris-00972215>. Accessed 1 May 2020

- Aden-Antóniow, F., Satriano, C., Bernard, P., Poiata, N., Aissaoui, E. M., Vilotte, J. P., & Frank, W. B. (2020). Statistical analysis of the preparatory phase of the Mw 8.1 Iquique earthquake Chile. *Journal of Geophysical Research*, *125*, 1–14. <https://doi.org/10.1029/2019JB019337>
- Allen, R. (1982). Automatic phase pickers: Their present use and future prospects. *Bulletin of the Seismological Society of America*, *72*, S225–S242.
- Battaglia, J., & Aki, K. (2003). Location of seismic events and eruptive fissures on the Piton de la Fournaise volcano using seismic amplitudes. *Journal of Geophysical Research*, *108*, 2364. <https://doi.org/10.1029/2002JB002193>
- Bennani, M., Josien, J., & Bigarre, P. (2003). Surveillance des risques d'effondrement dans l'après mine, besoins, méthodes, Colloque International Après-mine 2003, Nancy, France, 2003. ineris-00972401. Retrieved from <https://hal-ineris.archives-ouvertes.fr/ineris-00972401>. Accessed 1 May 2020
- Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A python toolbox for seismology. *Seismological Research Letters*, *81*, 530–533. <https://doi.org/10.1785/gssrl.81.3.530>
- Cesca, S., & Grigoli, F. (2015). Full waveform seismological advances for microseismic monitoring. *Advances in Geophysics*, *56*, 169–228. <https://doi.org/10.1016/bs.agph.2014.12.002>
- Contrucci, I., Balland, C., Kinscher, J., Bennani, M., Bigarré, P., & Bernard, P. (2019). Aseismic mining subsidence in an abandoned mine: Influence Factors And Consequences For Post-Mining Risk Management. *Pure and Applied Geophysics*, *176*, 801–825. <https://doi.org/10.1007/s00024-018-2015-6>
- Contrucci, I., Klein, E., Bigarre, P., Lizeur, A., & Lomax, A. (2008). Early-warning microseismic systems applied to the management of post-mining large-scale ground failures: calibration by a geophysical field experiment of blast swarms. *Geophysical Research Abstracts*, *10*, EGU2008-A-09300.
- Contrucci, I., Klein, E., Bigarré, P., Lizeur, A., Lomax, A., & Bennani, M. (2010). Management of post-mining large-scale ground failures: Blast swarms field experiment for calibration of permanent microseismic early-warning systems. *Pure and Applied Geophysics*, *167*, 43–62. <https://doi.org/10.1007/s00024-009-0005-4>
- Contrucci, I., Klein, E., Cao, N.-T., Daupley, X., & Bigarre, P. (2011). Multi-parameter monitoring of a solution mining cavern collapse: First insight of precursors. *Comptes Rendus Geoscience*, *343*, 1–10. <https://doi.org/10.1016/j.crte.2010.10.007>
- Couffin, S., Bigarre, P., Bennani, M., & Josien, J. P. (2003). Permanent real time microseismic monitoring of abandoned mines for public safety. Field Measurements in Geomechanics. In *Proceedings of the 6th International Symposium*, Oslo, Norway, 23–26 September 2003, pp. 437–444.
- Dheilly, A., & Brigati, B. (2015). Gestion du réservoir minier de Gardanne, Bouches du Rhône: Approches environnementales et hydrauliques du pompage et de son rejet dans la Mer Méditerranée. Exploit. minières passées présentes Impacts environnementaux sociétaux Collection EDYTEM, 71–84. Retrieved from <https://hal-brgm.archives-ouvertes.fr/hal-01100951>. Accessed 1 May 2020.
- Didier, C. (2008). The French experience of post-mining management. Symp. Post-Mining 2008, Feb 2008, Nancy, Fr. ineris-00973291 HAL, Vol. 45, ASGA. Vandoeuvre-lès-Nancy. Retrieved from <https://hal-ineris.archives-ouvertes.fr/ineris-00973291>. Accessed 1 May 2020.
- Didier, C., der Merwe, J. N., Bétournay, M., Mainz, M., Kotyrba, A., Aydan, Ö., Josien, J. P., et al. (2008). *Mine closure and post-mining management. International State-of-the-art report, International Commission on Mine Closure*. ISRM.
- Diekmann, L., Schwarz, B., Bauer, A., & Gajewski, D. (2019). Source localization and joint velocity model building using wavefront attributes. *Geophysical Journal International*, *219*(2), 995–1007. <https://doi.org/10.1093/gji/ggz342>
- Dominique, P. (2016a). Bassin houiller de Provence (13). Résultats issus du dispositif microsismique de surveillance complémentaire La crise de décembre 2014. Rapport BRGM/ RP-65561-FR, 25 p., 10 fig. BRGM.
- Dominique, P. (2016b). Bassin houiller de Provence (13). Bilan de l'activité sismique 2015 - Secteur de Fuveau ouest. Rapport BRGM/RP-66203-FR, 64 p., 32 fig., 2 ann. BRGM.
- Lefebvre, O., & Dommanget, A. (2010). Bassin de lignite de Provence Anciennes concessions détenues par les Charbonnages de France Définition et cartographie préliminaire de l'aléa. Edition par commune, GEODERIS S 2010/37DE-10PAC 2210 NOM. Geoderis.
- Gharti, H. N., Oye, V., Roth, M., & Kuhn, D. (2010). Automated microearthquake location using envelope stacking and robust global optimization. *Geophysics*, *75*(4), MA27–MA46. <https://doi.org/10.1190/1.3432784>
- Gibbons, S. J., & Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation. *Geophysical Journal International*, *165*, 149–166. <https://doi.org/10.1111/j.1365-246X.2006.02865.x>
- Grigoli, F., Cesca, S., Rinaldi, A. P., Manconi, A., López-Comino, J. A., Clinton, J. F., Westaway, R., et al. (2018). The November 2017 Mw5.5 Pohang earthquake: A possible case of induced seismicity in South Korea. *Science*, *360*, 1003–1006. <https://doi.org/10.1126/science.aat2010>
- Grigoli, F., Cesca, S., Vassallo, M., & Dahm, T. (2013). Automated seismic event location by travel-time stacking: An application to mining induced seismicity. *Seismological Research Letters*, *84*(4), 666–677. <https://doi.org/10.1785/0220120191>
- Hainzl, S., Kraft, T., Wassermann, J., Igel, H., & Schmedes, E. (2006). Evidence for rainfall-triggered earthquake activity. *Geophysical Research Letters*, *33*, L19303. <https://doi.org/10.1029/2006GL027642>
- Hanks, T. C., & Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research*, *84*, 2348. <https://doi.org/10.1029/JB084iB05p02348>
- Hartigan, J. A. (1975). *Clustering algorithms*. John Wiley & Sons Inc.
- Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-means clustering algorithm. *Applied Statistics*, *28*, 100–108. <https://doi.org/10.2307/2346830>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, *9*(3), 90–95.
- Husen, S., Bachmann, C., & Giardini, D. (2007). Locally triggered seismicity in the central Swiss Alps following the large rainfall event of August 2005. *Geophysical Journal International*, *171*, 1126–1134. <https://doi.org/10.1111/j.1365-246X.2007.03561.x>
- Kinscher J., Bernard, P., Contrucci, I., Mangeney, A., Pigué, J. P., Bigarre, P. (2015). Location of microseismic swarms induced by salt solution mining. *Geophysical Journal International*, *200*(1), 337–362. <https://doi.org/10.1093/gji/ggu396>
- Kinscher, J., Namjesnik, D., Contrucci, I., Dominique, P., Survey, F. G., & Klein, E. (2018). Relevance of seismic risk assessment

- in abandoned mining districts: the case of the Gardanne coal mine, Provence, France. Mine closure 2018: Proceedings of the 12th International Conference on Mine Closure. Technical University Bergakademie Freiberg, pp. 615–624.
- Kinscher, J. (2017). Origine de la sismicité après mine du bassin houiller de Gardanne Programme EAT-DRS06-Rapport de fin d'opération 4 sismicité induite par les activités industrielles et extractives. Rapport INERIS DRS-17-164272-00556A. Ineris.
- Kinscher, J., Coccia, S., Daupley, X., & Bigarre, P. (2017). Microseismic monitoring of caving and collapsing events in solution mines. In *9th International Symposium on Rockbursts and Seismicity in Mines*, Santiago, Chile, November 15–17, 2017, 117–122. University of Chile, 2017.
- Kinscher, J., Cesca, S., Bernard, P., Contrucci, I., Mangeney, A., Piguet, J. P., & Bigarre, P. (2016). Resolving source mechanisms of microseismic swarms induced by solution mining. *Geophysical Journal International*, 206, 696–715. <https://doi.org/10.1093/gji/ggw163>
- Lay, T., & Wallace, T. C. (1995). *Modern global seismology*. Elsevier.
- Li, L., Becker, D., Chen, H., Wang, X., & Gajewski, D. (2018). A systematic analysis of correlation-based seismic location methods. *Geophysical Journal International*, 212, 659–678. <https://doi.org/10.1093/gji/ggx436>
- Li, L., Tan, J., Schwarz, B., Staněk, F., Poiata, N., Shi, P., Diekmann, L., et al. (2020). Recent advances and challenges of waveform-based seismic location methods at multiple scales. *Reviews of Geophysics*, 58(1), e2019RG000667. <https://doi.org/10.1029/2019RG000667>
- Li, Z., & van der Baan, M. (2016). Microseismic event localization by acoustic time reversal extrapolation. *Geophysics*, 81(3), 1M1–Z27. <https://doi.org/10.1190/GEO2015-0300.1>
- Lomax, A. (2005). A Reanalysis of the hypocentral location and related observations for the great 1906 California earthquake. *Bulletin of the Seismological Society of America*, 95, 861–877. <https://doi.org/10.1785/0120040141>
- Lomax, A. (2008). Location of the Focus and Tectonics of the Focal Region of the California Earthquake of 18 April 1906. *Bulletin of the Seismological Society of America*, 98, 846–860.
- López-Comino, J. A., Cesca, S., Heimann, S., Grigoli, F., Milkereit, C., Dahm, T., & Zang, A. (2017). Characterization of hydraulic fractures growth during the Äspö hard rock laboratory experiment (Sweden). *Rock Mechanics and Rock Engineering*, 50, 2985–3001. <https://doi.org/10.1007/s00603-017-1285-0>
- Matrullo, E. (2015). Induced micro-seismicity by flooding of the abandoned Gardanne coal field (Provence, France): analysis and interpretation, Rapport INERIS DRS-15-151231-10877A. Ineris.
- Matrullo, E., Contrucci, I., Dominique, P., Bennani, M., Aochi, H., Kinscher, J., Bernard, P., et al. (2015). Analysis and interpretation of induced micro-seismicity by flooding of the Gardanne Coal Basin (Provence–Southern France). *EAGE Conference and Exhibition-Workshops*. <https://doi.org/10.3997/2214-4609.201413521>
- Megies, T., Beyreuther, M., Barsch, R., Krischer, L., & Wassermann, J. (2011). ObsPy: What can it do for data centers and observatories? *Annales Geophysicae*, 54, 47–58. <https://doi.org/10.4401/ag-4838>
- Namjesnik, D., Kinscher, J., Dominique, P., Gunzburger, Y., Contrucci, I., & Klein, E. (2019). Seismic multiplets induced by flooding of an abandoned coal mine and their implication for seismic hazard assessment. *Geophysical Research Abstracts*, Vol. 21, EGU2019-18685-1, 2019.
- Noviantî, P., Setyorini, D., & Rafflesia, U. (2017). K-means cluster analysis in earthquake epicenter clustering. *International Journal of Advances in Intelligent Informatics*, 3, 81–89. <https://doi.org/10.26555/ijain.v3i2.100>
- Ogasawara, H., Fujimori, K., Koizumi, N., Hirano, N., Fujiwara, S., Otsuka, S., Nakao, S., et al. (2002). Microseismicity induced by heavy rainfall around flooded vertical ore veins. *The Mechanism of Induced Seismicity*. https://doi.org/10.1007/978-3-0348-8179-1_4
- Oye, V., & Roth, M. (2003). Automated seismic event location for hydrocarbon reservoirs. *Computers and Geosciences*, 29, 851–863. [https://doi.org/10.1016/S0098-3004\(03\)00088-8](https://doi.org/10.1016/S0098-3004(03)00088-8)
- Palgunadi, K. H., Poiata, N., Kinscher, J., Bernard, P., De Santis, F., & Contrucci, I. (2019). Methodology for full waveform near real-time automatic detection and localization of microseismic events using high (8 kHz) sampling rate records in mines: Application to the Garpenberg mine (Sweden). *Seismological Research Letters*, 91(1), 399–414. <https://doi.org/10.1785/0220190074>
- Poiata, N., Satriano, C., Vilotte, J.-P., Bernard, P., & Obara, K. (2016). Multiband array detection and location of seismic sources recorded by dense seismic networks. *Geophysical Journal International*, 205, 1548–1573. <https://doi.org/10.1093/gji/ggw071>
- Poiata, N., Vilotte, J.-P., Bernard, P., Satriano, C., & Obara, K. (2018). Imaging different components of a tectonic tremor sequence in southwestern Japan using an automatic statistical detection and location method. *Geophysical Journal International*, 213, 2193–2213. <https://doi.org/10.1093/gji/ggy070>
- Rehman, K., Burton, P. W., & Weatherill, G. A. (2014). K-means cluster analysis and seismicity partitioning for Pakistan. *Journal of Seismology*, 18, 401–419. <https://doi.org/10.1007/s10950-013-9415-y>
- Saar, M. O., & Manga, M. (2003). Seismicity induced by seasonal groundwater recharge at Mt. Hood Oregon. *Earth and Planetary Science Letters*, 214, 605–618. [https://doi.org/10.1016/S0012-821X\(03\)00418-7](https://doi.org/10.1016/S0012-821X(03)00418-7)
- Taisne, B., Brenguier, F., Shapiro, N., & Ferrazzini, V. (2011). Imaging the dynamics of magma propagation using radiated seismic intensity. *Geophysical Research Letters*, 38, 2–6. <https://doi.org/10.1029/2010GL046068>
- Waskom, M., Botvinnik, O., O’Kane, D., Hobson, P., Ostblom, J., Lukauskas, S., Gempferline, D. C., et al. (2018). mwaskom/sea-born: v0.9.0 (July 2018). (Version v0.9.0). Zenodo. <https://doi.org/10.5281/zenodo.1313201>.
- Weatherill, G., & Burton, P. W. (2009). Delineation of shallow seismic source zones using K-means cluster analysis, with application to the Aegean region. *Geophysical Journal International*, 176, 565–588. <https://doi.org/10.1111/j.1365-246X.2008.03997.x>
- Willacy, C., Van Dedem, E., Minisini, S., Li, J., Blokland, J. W., Das, I., & Droujinine, A. (2019). Full-waveform event location

and moment tensor inversion for induced seismicity. *Geophysics*, 84(2), KS39–KS57. <https://doi.org/10.1190/geo2018-0212.1>

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. *Bulletin of the Seismological Society of America*, 88, 95–106.

(Received August 31, 2020, revised May 22, 2021, accepted May 26, 2021, Published online June 10, 2021)