Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

Earth Surface Dynamics Discussions
Open Access
EGU

Systematic Identification of External Influences in Multi-Year Micro-Seismic Recordings Using Convolutional Neural Networks

Matthias Meyer[1], Samuel Weber[2], Jan Beutel[1], and Lothar Thiele[1]

[1]Computer Engineering and Networks Laboratory, ETH Zurich, Zurich, Switzerland
[2]Department of Geography, University of Zurich, Zurich, Switzerland

Correspondence: Matthias Meyer (matthias.meyer@tik.ee.ethz.ch)

Abstract. Natural hazards, e.g. due to slope instabilities, are a significant risk for the population of mountainous regions. Monitoring of micro-seismic signals can be used for process analysis and risk assessment. However, these signals are subject to external influences, e.g anthropogenic or natural noise. Successful analysis depends strongly on the capability to cope with such external influences. For correct slope characterization it is thus important to be able to identify, quantify and take these influences into account.

In long-term monitoring scenarios manual identification is infeasible due to large data quantities demanding accurate automated analysis methods. In this work we present a systematic strategy to identify multiple external influences, characterize their impact on micro-seismic analysis and develop methods for automated identification. We apply the developed strategy to a real-word, multi-sensor, multi-year micro-seismic monitoring experiment on the Matterhorn Hörnliridge (CH). We present a convolutional neural network for micro-seismic data to detect external influences originating in mountaineers, a major unwanted influence, with an error rate of less than 1 %, which is 3x lower than comparable algorithms. Moreover, we present an ensemble classifier for the same task obtaining an error rate of 0.79 % and an F1 score of 0.9383 by using images and micro-seismic data. Applying the classifiers to the experiment data reveals that approximately 1/4 of events detected with an event detector are not due to seismic activity but due to anthropogenic mountaineering influences and that time periods with mountaineer activity have a 9x higher event rate. Due to these findings we argue that a systematic identification of external influences, like presented in this paper, is a prerequisite for a qualitative analysis.

1 Introduction

Rock-slope failures and landslides pose a significant risk for humans and infrastructure (Glade et al., 2005). Identifying and monitoring potential slope instabilities is therefore of great importance. Micro-seismic analysis can be used to assess the stability and characteristics of slopes in various environments (Spillmann et al., 2007) and as a result is beneficial for two main applications: (i) process analysis and scientific understanding and (ii) natural hazard warning systems. With the advent of

low-power wireless sensor networks, long-term monitoring with continuous in-situ monitoring of acoustic emission and micro-seismic signals has become possible even in remote areas (Girard et al., 2012) and the feasibility of the two aforementioned application areas have been demonstrated. Especially if deployed in larger quantities, the data volumes originating from such high-rate sensors necessitate an automated analysis workflow since manual treatment of the sensor data is infeasible.

In practice, micro-seismic signals are influenced by anthropogenic or extrinsic ambient noise (Eibl et al., 2017; van Herwijnen and Schweizer, 2011), leading to biased assessments. One way to account for such external influences is to manually identify their sources in the recordings (van Herwijnen and Schweizer, 2011). This procedure, however, is not feasible for continuous, autonomous monitoring due to scaling issues. Another approach is to focus and limit analysis to decisively chosen time periods known not to be influenced by e.g. anthropogenic noise (Occhiena et al., 2012) or limit to field sites far away from possible sources of uncontrolled (man-made) interference. In practice both the temporal limitation as well as the spatial limitation pose severe restrictions resulting in a very unattractive overall solution. Research applications can benefit from close proximity to man-made infrastructure since set up and maintenance of monitoring infrastructure is facilitated. Applications in natural hazard early warning must not be restricted to special time-periods only. Moreover, they are specifically required to be usable close to inhabited areas with an increasing likelihood for human interference on the signals recorded. As a conclusion it is a requirement that external influences can be taken into account with an automated workflow, including pre-processing, cleaning and analysis of micro-seismic data.

A method commonly used to assess seismic activity is to calculate the number of seismic events per time interval for the time period of interest (Amitrano et al., 2005; Senfaute et al., 2009). The correct detection of seismic events is thus of importance for a good analysis. Due to its simplicity, a popular filtering technique for event detection is to use short-term/long-term average triggering (STA/LTA) (Withers et al., 1998). This is often used in the analysis of unstable slopes (Colombero et al., 2018; Levy et al., 2011), is available in commercial data loggers (Geometrics, 2018) and can be used to detect footsteps (Anchal et al., 2018). Due to it's inherent simplicity, STA/LTA cannot reliably discriminate seismic activity from external (unwanted) influence factors such as noise from human beings, wind, rain or hail without manually supervising and intervening with the detection process. As a result the blind application of STA/LTA will inevitably lead to the false estimation of relevant intrinsic seismic events if significant external influences, e.g. wind, are present (Allen, 1978).

There exist several algorithmic approaches to mitigate the problem of external influences by increasing the sensitivity of seismic event detection. Auto-correlation and cross-correlation methods (Brown et al., 2008; Gibbons and Ringdal, 2006) use seismic event examples to find similar events, failing if events differ in "shape" or if the medium is very inhomogeneous (Weber et al., 2018). The most recent advanced methods are based on machine learning techniques (Reynen and Audet, 2017). The use of neural networks (Kislov and Gravirov, 2017; Perol et al., 2018) shows promising results with the drawback that large datasets containing ground truth (verified seismic events) are required to train these networks. In earthquake research these large database of known seismic events exist, but are difficult to obtain in slope stability analysis where effects are strictly local to a given field site, inhomogeneities are commonly found on a small scale and each field site differs in its characteristics with respect to signal attenuation and impulse response. Thus, obtaining a database of known events is required for each new field site which requires expert knowledge and is generally known to be an arduous time consuming task.

By using additional sensors like weather stations, cameras or microphones and external knowledge such as helicopter flight plans or mountain hut occupancy it is possible to semi-automatically label non-seismic events, e.g. helicopters, footsteps or wind without the need for expert knowledge. Such "external" information sources can be used to establish an annotated data subset that allows to train an algorithm that is able to identify unwanted external influence factors in the complete dataset. Following this pre-processing step, seismic activity can be assessed by using the simple STA/LTA method mentioned earlier or more complex approaches taking into account the external influences identified in the earlier step. The efficient labeling of multiple external influence sources in large seismic datasets enables a broad set of algorithms to be applied subsequently. The concept of this method is generic and can be applied to many signal source/influence pairs.

We focus on an application to analyze micro-seismic signals originating from steep, fractured bedrock permafrost that is used as a driving example to illustrate the method and quantify its benefit on a real world example. Our case study is based on a multi-sensor, multi-year micro-seismic monitoring experiment on the Matterhorn Hörnligrat (CH) which is affected by mountaineer and wind influences.

In this paper we

- propose a strategy to identify and deal with unwanted external influences in multi-sensor, multi-year experiments.
- propose a convolutional neural network for acoustic event detection using micro-seismic data which detects mountaineers with an error rate of 0.96 % and a F1 score of 0.9167.
- present an ensemble classifier which detects mountaineers on images and micro-seismic data with an error rate of 0.79 % and a F1 score of 0.9383.
- compare the suitability of multiple algorithms for mountaineer detection using micro-seismic data and images and show that our approach shows a 3x lower error rate than other algorithms.
- exemplify our strategy for the case of micro-seismic event detection on real-world data from measurements in steep, fractured bedrock permafrost and show that time periods with mountaineer activity have a approximately 9x higher event rate and that 25% of all detected events are of non-seismic nature due to mountaineer interference
- argue that due these findings extensive identification of external influences is a prerequisite for qualitative analysis.

In the following, the concept of identifying external influences will be presented in Sect. 2 using a case study. The strategy to select and train a set of classifiers will be described in Sect. 3. Finally, the results are presented in Sect. 4 and the advantages and disadvantages of the presented method are discussed in Sect. 5.

2 Identification of External Influences

As noted earlier, measurements in a real-world scenario do not only contain signals of the phenomena of interest but are also affected by external influences as depicted in Fig. 1. For quantitative analysis it is crucial to account for such external influences,

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

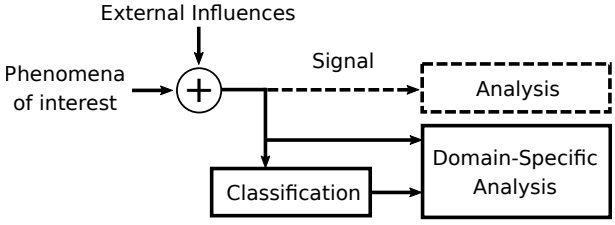

Figure 1. Model of the sensing method. In a real world measurement the phenomena of interest is superimposed by external influences. In a first approach the resulting signal can be directly analyzed by appropriate methods, which could be constrained by the external influences. In a second approach, which we suggest in this paper, classification of external influences can provide further information which can be utilized by domain-specific analysis methods.

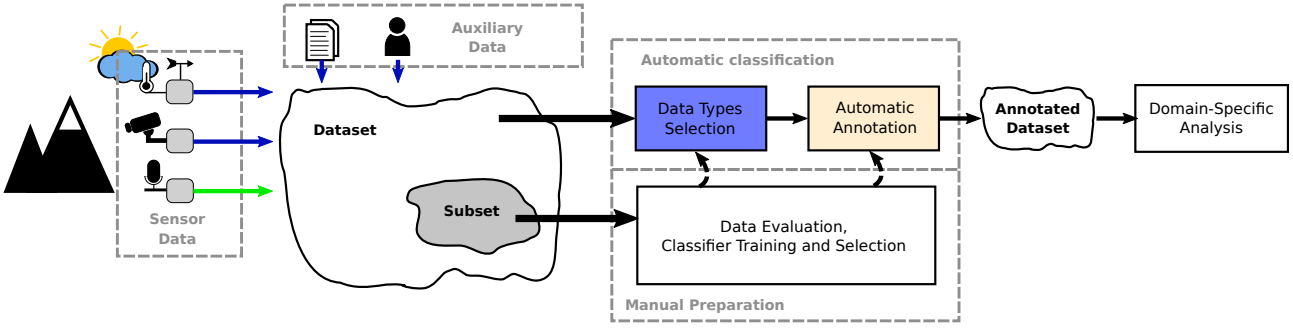

Figure 2. Illustration of the presented concept to enable domain-specific analysis on an existing measurement setup. Primary signal (light green) and secondary data (dark blue) are combined to form a dataset on which automatic classification can be performed. Secondary data is a combination of sensor and auxiliary data. The resulting annotated dataset can be used for domain-specific analysis. In a prior step, the correct settings and tools for automatic classification are determined in a manual preparation phase based on a subset of the dataset.

possibly also differentiating between different types of external influences present in a given signal. It is important to point out that external influences are not to be seen as something strictly negative that ought to be mitigated and removed for as much as possible. External influences can serve as context further describing the circumstances at a given point in time and space. Therefore a simple removal of all external influences is not desirable and we advocate that external influences are classified in a pre-processing step as presented in this paper. Using this extra information, tailored, domain specific analysis methods can be used on the signal benefiting from the additional information where applicable. An example for domain specific analysis is the estimation of the STA/LTA event rate for time periods when mountaineers are present and when they are not, which will be used in this paper to exemplify our method. The approach described is flexible, modular and extensible and does not constitute an information loss but rather an information gain constituting a significant advantage over competing approaches.

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

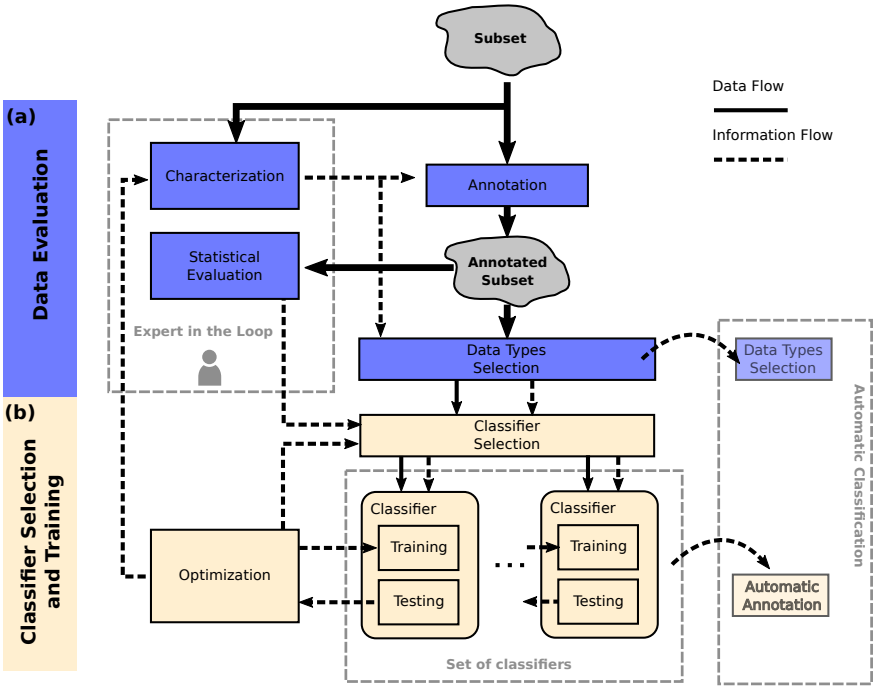

Figure 3. The manual preparation phase is subdivided into data evaluation and classifier selection and training. First, the data subset is characterized and annotated. This information can be used to do a statistical evaluation and select data types which are useful for classification. Domain experts are not required for the labor intensive task of annotation. The classifiers are selected, trained and optimized in a feedback loop until the best set of classifiers is found.

2.1 Method

In the following we provide an overview how to identify and quantify external influence sources using a systematic combination of manual, semi-automatic and automated steps. Care is taken to reduce work requiring a human in the loop and to use automated methods as much as possible. Figure 2 illustrates the concept. In a first step the available data sources are assessed. Given a measurement setup consisting of multiple sensors, one or more sensor signals are specified as primary signals (e.g. micro-seismic signals) targeted by a subsequent domain-specific analysis method. Additionally, secondary data are chosen which can support the classification of external influences. These secondary data can be either other sensor signals, e.g. images or weather information, or auxiliary data, e.g. local observations or helicopter flight data. All data sources are combined into a dataset. However, the resulting dataset is not yet annotated which is required to perform domain-specific analysis. Two key challenges need to be addressed before an annotated dataset can be established by automatic classification: (i) suitable data types need to be selected for classification since not every data type can be used to continuously classify every external influence (e.g. wind sensors are not designed to capture the sound of footsteps; flight data cannot be obtained for each time step) and (ii) one or a set of suitable, good-performing classifiers have to be found for each external influence. These challenges

can be addressed in a manual preparation phase, which includes dataset evaluation as well as classifier training, evaluation and selection.

2.2 Manual preparation

A ground truth is often needed for state-of-the-art classifiers (e.g. neural networks), which requires manual annotation of data points. In a realistic setting the dataset comprises large data quantities which is impractical for manual annotation. To reduce the amount of manual labor only a subset of the dataset is selected and used in a manual preparation phase, which consists of data evaluation and classifier training and selection as depicted in Fig. 3. Data evaluation can be subdivided into four parts: (i) characterization of external influences in the primary signal (i.e. the relation between primary and secondary signals), (ii) annotating the subset based on the primary and secondary signals, (iii) selecting the data types suitable for classification and (iv) performing a first statistical evaluation with the annotated dataset, which facilitates the selection of a classifier. Much of the preparatory work in creating this ground truth dataset is of manual nature and varies for different secondary data types depending on their specifics. But the meticulous care does pay off. By separately treating different categories of external influences and secondary data it is possible to evaluate in detail the impact of every factor in solitude or in combination, including detailed statistics. Moreover, the presented strategy reduces the work requiring an expert. Characterization and statistical evaluation are the only steps where domain expertise is required while it is not required for the time and labor intensive annotation process.

The classifier selection and training phase presumes the availability of a variety of classifiers for different input data types. The classifiers do not perform equally well on the given task with the given subset. Therefore classifiers have to be selected based on their suitability for classification given the task and the data. A selection of classifiers is therefore trained and tested with the annotated subset and optimized for best performance. The classifier selection, training and optimization is repeated until a sufficiently good set of classifiers has been found. These classifiers can then be used for application in the automatic classification process.

In the following, the previously explained method will be exemplified for wind and mountaineer detection using microseismic, wind and image data from a real-world experiment.

2.3 Case Study

The data used in this paper originate from a multi-sensor and multi-year experiment (Weber et al., 2018) focusing on slope stability in high-alpine permafrost rock walls and understanding the underlying processes. Specifically, the sensor data is collected at the site of the 2003 rockfall event on the Matterhorn Hörnligrat, (Zermatt, Switzerland) at 3500 m a.s.l. where an ensemble of different sensors monitors the rockfall scar and surrounding environment over the past ten years. Relevant for this work are data from a three-axial seismometer (Lennartz LE-3Dlite MkIII), images from a remote controlled high-resolution camera (Nikon D300, 24 mm fixed focus), rock surface temperature measurements, net radiation measurements and ambient weather conditions, specifically wind speed from a co-located local weather station (Vaisala WXT520).

In (Weber et al., 2018) a seismometer is used to assess the seismic activity by using an STA/LTA event detector. The seismometer is chosen as the primary sensor and STA/LTA triggering is used as the reference method to assess seismic activity.

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

Seismic data is recorded locally using a nanometrics Centaur data logger and transferred daily by means of directional WLAN connectivity. The data is processed on-demand by STA/LTA triggering. The high-resolution camera's (Keller et al., 2009) field of view covers the immediate surroundings of the seismic sensor location as well as some backdrop areas further away on the mountain ridge. Figure 4 shows an overview of the field site including the location of the seismometer and an example image acquired with the camera. The standard image size is 1424x2144 pixels captured every 4 minutes. The Vaisala WXT520 weather data as well as the rock surface temperature are transmitted using a custom wireless sensor network infrastructure with a sampling rate once every 2 minutes.

Care has been taken to prevent significant data gaps but given the circumstances on such a demanding high-alpine field site certain outages of single sensors, e.g. due to power failures or also maintenance could not be prevented. Nevertheless this dataset constitutes an extensive and close-to-complete dataset.

The case study was affected by external influences, especially mountaineers and wind. This reduced the set of possible analysis tools. In following, we use this case study to exemplify the method presented in the previous sections.

2.4 Data Evaluation

As explained in Sect. 2.2, a prerequisite for automatic classification is the analysis of the given dataset to specify the requirements of the classifier. This sections explains the required steps of subset creation, characterization, annotation, statistical evaluation and the selection of the data type for classification. The subset is created by collecting continuous data from the sensors and additional data, which helps to characterize the external influences. In the presented case the additional data is non-continuous and consist of local observations, pre-processed STA/LTA triggers from (Weber et al., 2018), accommodation occupancy of a nearby hut and a non-exhaustive list of helicopter flight data from a duration of approximately 7 weeks provided by a local helicopter company. Before an annotated subset can be created the collected data must be characterized for its usefulness in the annotations process, i.e. which data type can be used to annotate which external influence. In the following the steps for creation of an annotated subset, as illustrated in Fig. 3 (a) are explained.

2.4.1 Characterization

The seismometer data consists of sounds originating from different sources. This section will discuss commonly captured sources in a real world measurement setup. The definition of an event can be ambiguous. Geoscientific studies often aim at identifying events related to rupture or fracture in rock and/or ice originating for example from thermal stresses, pressure variations or earthquakes (Amitrano et al., 2012). We will refer to them as events or seismic events. In audio classification literature an event is a specific sound which identifies a certain source, e.g. footsteps identifying a human. We will refer to them as acoustic events. While a seismic event is mostly regarded as a short-duration impulse, acoustic events can be a combination of different sounds of varying duration and spectral properties.

In this study we will consider multiple acoustic events, which are the influences of mountaineers, helicopters, wind and rockfalls. Additionally, we will analyze the time frames where none of the before mentioned acoustic events happen. The mountaineer impact will be characterized on a long-term scale by correlating with hut accommodation occupancy (see Fig. 9)

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

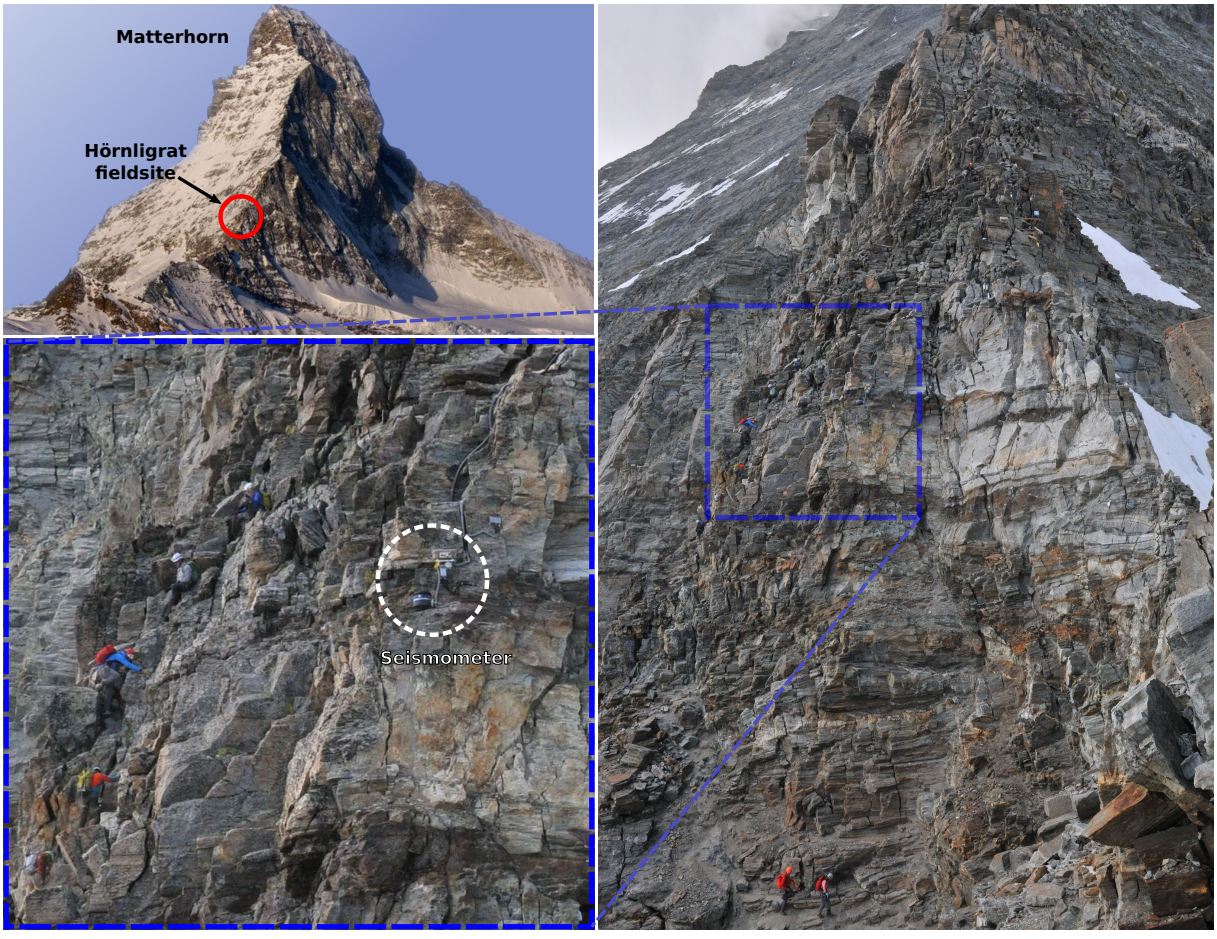

Figure 4. The field site is located on the Matterhorn Hörnligrat at 3500 m a.s.l. which is denoted with a red circle. The photograph on the right is taken by a remotely controlled DSLR camera on the field site at 2016-08-04T12:00:11. The seismometer of interest (white circle) is located on a rock instability which is close to a frequently used climbing route

and on a short-scale by person identification on images. Helicopter examples are identified by using flight data and local observations. By analyzing spectrograms one can get an intuition what mountaineers or helicopters "look" like, which facilitates the manual annotation process. In Fig. 5 different recordings from the field site are illustrated, which have been picked by using the additional information described at the beginning of this section. For six different examples the time domain signal, it's corresponding spectrogram and STA/LTA triggers are depicted. The settings for the detector are the same for all the subsequent plots. It becomes apparent in Fig. 5 (b)-(c) that anthropogenic noise, such as mountaineers walking by or helicopters, can have a strong influence on seismic recordings. Moreover, it becomes apparent that it might be feasible to assess acoustic events on a larger time frame. Mountaineers for example show characteristic patterns of increasing or decreasing loudness and helicopters have distinct spectral patterns, which could be beneficial to classify these events. Additionally, the images captured on site

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

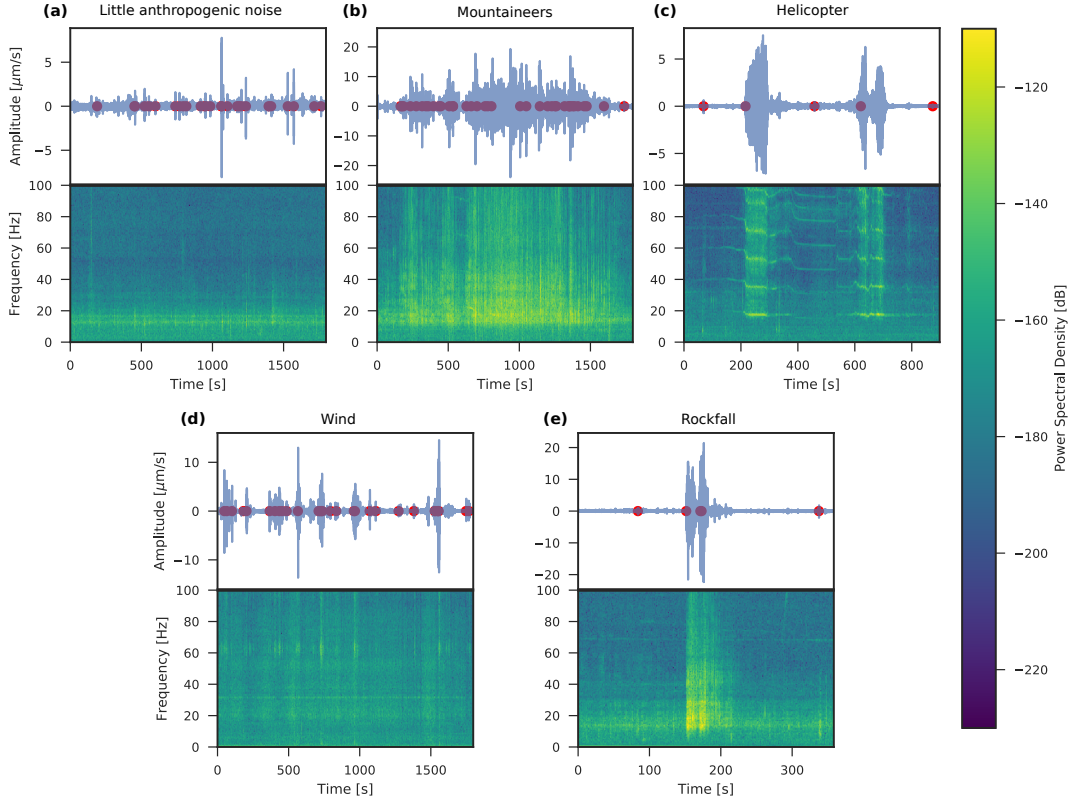

Figure 5. Micro-seismic signals and the impact of external influences: (a) During a period of little anthropogenic noise the seismic activity is dominant. (b) In the spectrogram the influence of mountaineers become apparent. (c) helicopter in close spatial proximity to the seismometer (d) wind influence on the signal (e) a rockfall in close proximity to the seismometer

show when a mountaineer is present (see Fig. 4), but due to fog, lens flares or snow on the lens the visibility can be limited. Another limiting factor is the temporal resolution of one image every 4 minutes. Mountaineers could move through the visible are in between two images.

The wind sensor can directly be used to identify the impact on the seismic sensor. By manually analyzing the correlation between tremor amplitude and wind speed (e.g. in Fig. 6) it can be deduced that wind speeds above approximately 30 km/h have a visible influence on the tremor amplitude. Tremor amplitude is the frequency-selective, median, absolute ground velocity and has been calculated for the frequency range of 1-60 Hz according to (Bartholomaus et al., 2015).

Rockfalls can best be identified by local observations since the camera captures only a small fraction of the receptive range of the seismometers. Figure 5 (e) shows an example of a rockfall. The number of rockfall observations and rockfalls caught on camera are however very limited.

The subplot 5 (a) shows a time frame from a period which is not strongly influenced by external influences. Repetitive seismic events are visible and trigger the event detector.

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

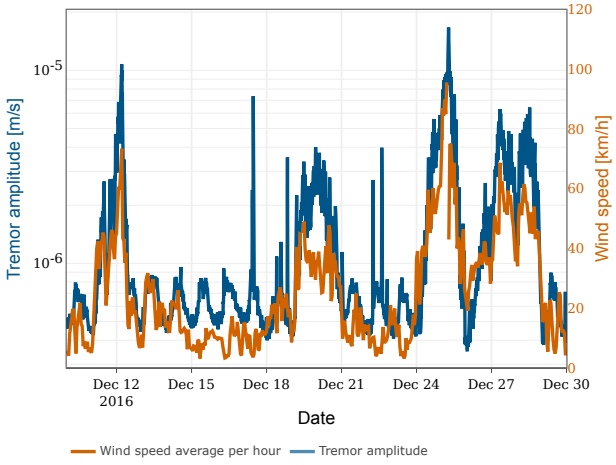

Figure 6. Impact of wind (light orange) on the seismic signal. The tremor amplitude (dark blue) is calculated according to (Bartholomaus et al., 2015). The effect of wind wind speed on tremor amplitude becomes apparent for wind speeds above approximately 30 km/h. Note the different scales on the y-axes.

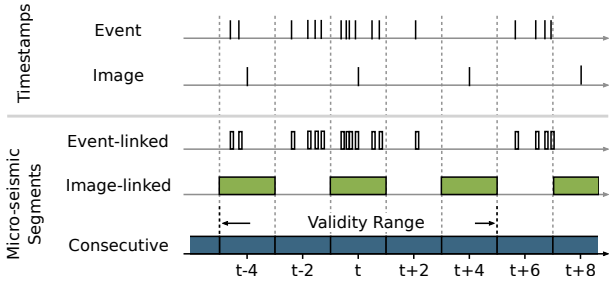

Figure 7. Illustration of micro-seismic segmentation. Event-linked segments are 10 second segments starting on event timestamps. Image-linked segments are two minute segments centered around an image timestamp. Consecutive segments are 2 minute segments sequentially extracted from the continuous micro-seismic signal.

2.4.2 Annotation

For annotation and evaluation the continuous micro-seismic signals are segmented. Figure 7 provides an overview of the three segmentation types, event-linked segments, image-linked segments and consecutive segments. Image-linked segments are extracted due the fact that a meaningful relation between seismic information and photos is only given in close temporal proximity, therefore images and micro-seismic data are linked in the following way: For each image a 2 minute micro-seismic segment is extracted from the continuous micro-seismic signal. The micro-seismic segment's start time is set to 1 minute before the image timestamp. Event-linked segments are extracted based on the STA/LTA triggers from (Weber et al., 2018). For each trigger 10 seconds following the timestamp of the trigger are extracted from the micro-seismic signal. Consecutive segments are 2 minute segments sequentially extracted from the continuous micro-seismic signal.

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

Only the image-linked segments are used during annotation, their label can however be transferred to other segmentation types by assigning the image-linked label to overlapping event-linked or consecutive segments. Image-linked and event-linked segments are used during data evaluation and classifier training. Consecutive segments are used for automatic classification on the complete dataset. Here, falsely classified segments are reduced by assigning each segment a validity range. A segment classified as mountaineer is only considered correct if the distance to the next (or previous) mountaineer is less than 5 minutes. This is based on an estimation of how long the mountaineers are typically in the audible range of the seismometer.

For mountaineer classification the required label is *mountaineer* but additional labels will be annotated which could be beneficial for classifier training and statistical analysis. These labels are *helicopters*, *rockfalls*, *wind*, *low visibility* (if the lens is partially obscured), and *lens flares*. The *wind* label applies to segments where the wind speed is higher than 30 km/h, which is the lower bound for noticeable wind impact as resulted from Sect. 2.4.1.

Figure 8 depicts the availability of image-linked segments per week during the relevant time frame. A fraction of the data is manually labeled by the authors, which is illustrated in Fig. 8. Two sets are created, a training set containing 5579 samples from the year 2016, and a test set containing 1260 data samples from 2017. The test set has been sampled randomly to avoid any human prejudgment. For each day in 2017 four samples have been chosen randomly, which are then labeled and added to the test set. The training set has been specifically sampled to include enough training data for each category. This means for example that more mountaineers samples come from the summer period where the climbing route is most frequently used. The number of verified rockfalls and helicopters is non-representative and although helicopters can be manually identified from spectrograms the significance of these annotations is not given due the limited ground truth from the secondary source. Therefore, for the rest of this study we will focus on mountaineers for qualitative evaluation. For statistical evaluation we will however use the manually annotated helicopter and rockfall samples to exclude them from the analysis. The labels for all categories slightly differ for micro-seismic data and images since the type of events which can be registered by each sensor differ. This means for instance that not every classifier uses all labels for training (e.g. a micro-seismic classifier cannot detect a lens flare). It also means that for the same time instance one label might apply to the image but not to the image-linked micro-seismic segment (e.g. mountaineers are audible but the image is partially obscured and the mountaineer is not visible). This becomes relevant in Sect. 3.2.4 when multiple classifiers are used for ensemble classification.

2.4.3 Data Types Selection

After the influences have been characterized the data type need to be selected which best describe each influence. The wind sensor delivers a continuous data stream and a direct measure of the external influence. In contrast, mountaineers, helicopters and rockfalls cannot directly be identified. A data type including information about these external influences needs to be selected. Local observations, accommodation occupancy and flight data can be discarded for the use as classifier input since the data cannot be continuously collected. According to Sect. 2.4.1 it seems possible to identify mountaineers, helicopters and rockfalls from micro-seismic data. Moreover, mountaineers can also be identified from images. As a consequence, the data types selected to perform classification are micro-seismic data, images and wind data.

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

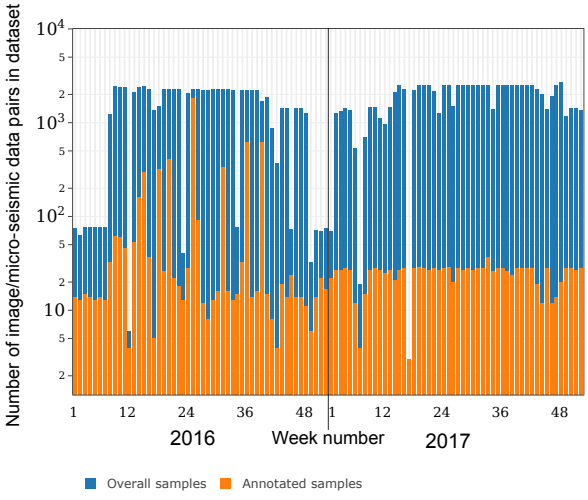

Figure 8. Number of image/micro-seismic data pairs in the dataset (dark blue) and in the annotated subset (light orange) displayed over the week number of the year 2016 and 2017. Note the logarithmic scale on the y-axis

2.4.4 Statistical Evaluation

The annotated test set from Sect. 2.4.2 is used for a statistical evaluation involving the impact of annotated external influences on micro-seismic analysis. The test set (2017) is chosen since wind data is not available for the whole training set due to a malfunction of the weather station in 2016. The experiment from (Weber et al., 2018) provides STA/LTA event triggers for 2016 and 2017. Table 1 shows statistics for several categories, which are 3 external influences and one category where none of the 3 external influences are annotated (declared as "None"). For each category, the total duration of all annotated segments is given and how many events per hour are triggered. It becomes apparent that mountaineers have the biggest impact on the event analysis. 95.26 events per hour are detected on average during time periods with mountaineer activity, while during all other time periods the average ranges from 10.6 to 13.12 events per hour. This finding supports our choice to mainly focus on mountaineers in this paper.

3 Classifier Selection and Training

The following section describes the classifier pre-selection, training, testing and how the classifiers are used to annotate the whole data stream as illustrated in Fig. 3 (b).

3.1 Classifier Selection

Multiple classifiers are available for the previously selected data types: wind data, images and micro-seismic data.

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

	None	Mountaineer	Wind	Helicopter
duration (hours)	28.87	1.9	6.6	3.73
mean number of events per hour	10.6	95.26	11.21	13.12

Table 1. Statistics of the annotated test set (2017) per annotation category. "None" represents the category when none of the other categories apply. Given are the total duration of annotated segments per category and the mean number of STA/LTA events per category.

For wind data a simple threshold classifier can be used, which indicates wind influences based on the wind speed. For simplicity the classifier labels time periods with wind speed above 30 km/h as *wind*. For images a convolutional neural network is selected to classify the presence of mountaineers in the image. The image classifier architecture is selected from the large pool of available image classifiers. For micro-seismic data, three different classifiers will be pre-selected: (i) a footstep detector based on manually selected features (standard deviation, kurtosis and frequency band energies) using a linear support vector machine (LSVM) similar to the detector used in (Anchal et al., 2018), (ii) a seismic event classifier adopted from (Perol et al., 2018) and (iii) an acoustic event classifier. We reimplemented the first two algorithms based on the information from the respective papers. The third is a major contribution in this paper and has been specifically designed for acoustic event classification on micro-seismic data.

The proposed convolutional neural network (CNN) for acoustic event classification on micro-seismic signals uses a time-frequency signal representation as input and consists of 2D convolutional layers. The time-domain signal, sampled at 1 kHz, is first offset-compensated and then transformed with a Short-Time Fourier Transformation (STFT). Subsequently, the STFT output is further processed by selecting the frequency range from 2 to 250 Hz and subdividing it into 64 linearly-spaced bands. The network consists of multiple convolutional, batch normalization and dropout layers, as depicted in Table 2. Except for the first convolutional layer, all convolutional layers are followed by batch normalization and Rectified Linear Units (ReLU) activation. Finally, a set of global average pooling layer, dropout and sigmoid activation reduces the features to one value representing the probability of a mountaineer. In total the network has 30,243 parameters. In this architecture multiple measures have been taken to minimize overfitting: the network is all-convolutional (Springenberg et al., 2014), batch normalization (Ioffe and Szegedy, 2015) and dropout (Srivastava et al., 2014) are used and the size of the network is small compared to recent audio classification networks (Hershey et al., 2016).

3.2 Training and Testing

We will evaluate the micro-seismic algorithms in two scenarios in Sect. 4.1. In this section, we describe the training and test setup for the two scenarios as well as for image and ensemble classification. The first scenario classifies event-linked segments. The second scenario compares the classifiers on image-linked segments, since the characterization from Sect. 2.4.1 suggested that using a longer temporal input window could lead to a better classification because it can capture more characteristics of a mountaineer sound. Training is performed with the annotated subset from Sect. 2.4 and a random 10 % of the training set are used as validation set, which is never used during training. For the acoustic and seismic event classifiers the number of epochs

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

Layer	stride	output channels
Conv2D + BatchNorm + Linear	1	32
Conv2D + BatchNorm + ReLU	2	32
Dropout	-	32
Conv2D + BatchNorm + ReLU	2	32
Dropout	-	32
Conv2D + BatchNorm + ReLU	1	32
Conv2D + BatchNorm + ReLU	1	32
Dropout	-	32
Conv2D + BatchNorm + ReLU	1	1
Global Average Pooling	1	1
Dropout	-	1
Sigmoid Activation	-	1

Table 2. Structure of the proposed acoustic event classifier.

has been fixed to 100 and for the image classifier to 20. After each epoch the F1 score of the validation set is calculated and based on it the best performing network version is selected. The F1 score is defined as

$$\text{F1 score} = \frac{2 \cdot \text{true positive}}{2 \cdot \text{true positive} + \text{false negative} + \text{false positive}} \tag{1}$$

The threshold for the network's output is determined by running a parameter search with the validation set's F1 score as metric. Training was performed in batches of 32 samples with the ADAM (Kingma and Ba, 2014) optimizer and cross-entropy loss. The Keras (Chollet and others, 2015) framework with a TensorFlow backend (Abadi et al., 2015) was used to implement and train the network. The authors of the seismic event classifier (Perol et al., 2018) provide TensorFlow source code, but to keep the training procedure the same it was re-implemented with the Keras framework. The footstep detector is trained with scikit-learn (Pedregosa et al., 2011). Out of 10 runs the footstep detector which performed best on the validation set was selected. Testing is performed on the test set which is independent of the training set and has not been used during training. The metrics error rate and F1 score are calculated.

Since neural networks are initialized with random values, and thus the classifier performance can vary, it is common to do multiple iterations of training and testing to get the best performing classifier instance. We train and test 10 iterations and select the best classifier instance of each classifier type to evaluate and compare their performances in Sect. 4.

The input of the micro-seismic classifiers must be variable to be able to perform classification on event-linked segments and image-linked segments. Due to the principle of convolutional layers, the CNN architectures are independent of the input size and therefore no architectural changes have to be performed. The footstep detector's input features are averaged over time by design and are thus also time-invariant.

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

3.2.1 Event-linked Segments Experiment

Literature suggests that STA/LTA cannot distinguish seismic events from other noise events (Allen, 1978). Therefore the first micro-seismic experiment investigates if the presented algorithms can distinguish events induced by mountaineers from other events in the signal. The event-linked segments are used for training and evaluation. The results will be discussed in Sect. 4.1.

3.2.2 Image-linked Segments Experiment

In the second micro-seismic experiment the image-linked segments will be used. Each classifier is trained and evaluated on the image-linked segments. The training parameters for training the classifiers on image-linked segments are as before but additionally data augmentation is used to minimize overfitting. Data augmentation includes random circular shift and random cropping on the time axis. Moreover, to account for the uneven distribution in the dataset, it is made sure that during training the convolutional neural networks see one example of a mountaineer every batch. The learning rate is set to 0.0001. The classifiers are then evaluated on the image-linked segments.

To be able to compare the results of the classifiers trained on image-linked segments to the classifiers trained on event-linked segments (Sect. 3.2.1), the classifiers from Sect. 3.2.1 will be evaluated on the image-linked segments as well. The metrics can be calculated with the following assumption: If any of the event-linked segments which are overlapping with an image-linked segment are classified as mountaineer, the image-linked segment is considered to be classified as mountaineer as well.

The results will be discussed in Sect. 4.1.

3.2.3 Image Classification

Since convolutional neural networks are a predominant technique for image classification, a variety of network architectures have been developed. For this study, the MobileNet (Howard et al., 2017) architecture is used. The number of labeled images is small in comparison to the network size (approx. 3.2M parameters) and training the network on the Matterhorn images will lead to overfitting on the small dataset. To reduce overfitting a MobileNet implementation which has been pre-trained on ImageNet (Deng et al., 2009), a large-scale image dataset, will be used. Retraining is required since ImageNet has a different application focus than this study. The climbing route, containing the subject of interest, only covers a tiny fraction of the image and rescaling the image to 224x224, the input size of the MobileNet, would lead to vanishingly mountaineers (compare Fig. 4). However, the image size cannot be chosen arbitrarily large since a larger input requires more memory and results in a larger runtime. To overcome this problem the image has been rescaled to 448x672 pixels and although the input size differs from the pretrained version network retraining still benefits from pre-trained weights. Data augmentation is used to minimize overfitting. For data augmentation each image is slightly zoomed in and shifted in width and height. The network has been trained to detect 5 different categories which are relevant for a good accuracy of the mountaineer classifier. These categories consist of mountaineer, low visibility (if the lens is partially obscured), lens flare, snowy (if the seismometer is covered in snow) and bad weather (as far as it can be deduced from the image). In this paper only the metrics for mountaineers are of interest for the evaluation and the metrics for the other labels are discarded in the following.

3.2.4 Ensemble

In certain cases, a sensor cannot identify a mountaineer although there is one, for example the seismometers cannot detect when the mountaineer is not moving or the camera does not capture the mountaineer if the visibility is low. The usage of multiple classifiers can be beneficial in these cases. In our case micro-seismic and image classifier will be jointly used for mountaineer prediction. Since micro-seismic labels and image labels are slightly different, as discussed in Sect. 2.4.2, the ground truth labels must be combined. For a given category, a sample is labeled true if any of micro-seismic or image labels are true (logical disjunction). After individual prediction by each classifier the outputs of the classifiers are combined similarly and can be compared to the ground truth.

3.3 Optimization

Due to potential human errors during data labeling, the training set has to be regarded as a weakly-labeled dataset. Such datasets can lead to a worse classifier performance. To overcome this issue a human-in-the-loop approach is followed where a preliminary set of classifier is trained on the training set. In the next step, each sample of the dataset is classified. This procedure produces a number of true positives, false positives and true negatives. These are then manually relabeled and the labels for the dataset are updated based on human review. The procedure is repeated multiple times. This does however not completely avoid the possibility of falsely labeled samples in the dataset, since the algorithm might not find all human-labeled false negatives, but it increases the accuracy significantly.

3.4 Automatic Annotation

In Sect. 4.1 it will be shown that the best set of classifiers are the ensemble of image classifier and MicroseismicCNN. Therefore, the trained image classifiers and MicroseismicCNN classifier are used to annotate the whole time period of collected data to quantitatively assess the impact of mountaineers. The image classifier and the MicroseismicCNN will be used to classify all the images and micro-seismic data, respectively. The consecutive segments and images are used for prediction. To avoid false positives we assume that a mountaineer requires a certain amount of time to pass by the seismometer as illustrated in Fig. 7, therefore a mountaineer annotation is only considered valid if its minimum distance to the next (or previous) mountaineer annotation is less than 5 minutes. Subsequently, events within time periods classified as mountaineer are removed and the event count per hour is calculated.

4 Results

In the following the results of the different classifiers experiments described in Sect. 3.2 will be presented to determine the best set of classifiers. Furthermore, in Sect. 4.2 and Sect. 4.3 results of the automatic annotation process (Sect. 3.4) will used to evaluate the impact of external influences on the whole dataset.

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

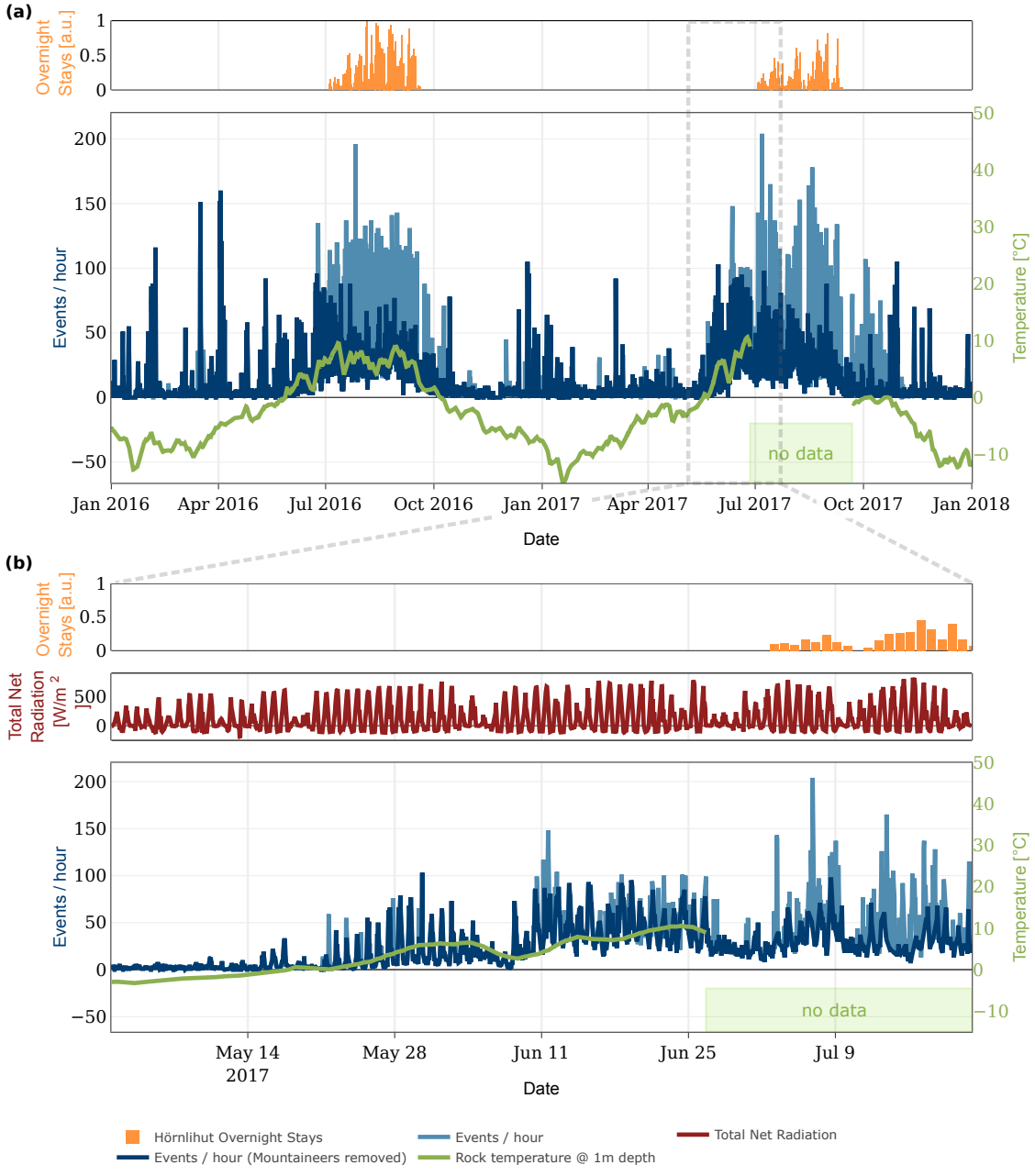

Figure 9. Event count, hut occupancy and rock temperature over time. (a) For the years 2016/2017 and (b) for a selected period during defreezing of the rock. The event rate from (Weber et al., 2018) is illustrated in light blue and the rate after removal of mountaineer induced events in dark blue. The strong variations in event count correlate with the presence of mountaineer, hut occupancy and in (b) with the total net radiation. The impact of mountaineers is significant after July 9th and event detection analysis becomes unreliable.

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

	Error Rate	F1 Score
Event-linked Segments		
Footstep Detector (Events)	0.1702	0.7692
Seismic Event Classifier (Events)	0.1250	0.8291
MicroseismicCNN (Events)	**0.0641**	**0.9062**
Image-linked Segments		
Footstep Detector (Events)	0.0706	0.5389
Seismic Event Classifier (Events)	0.0540	0.6047
MicroseismicCNN (Events)	0.0309	0.731
Footstep Detector	0.0952	0.52
Seismic Event Classifier	0.0313	0.7383
MicroseismicCNN	**0.0096**	**0.9167**
Image Classifier	0.0088	0.9134
Ensemble	**0.0079**	**0.9383**

Table 3. Results of the different classifiers. The addition "(Events)" labels the classifier versions trained on event-linked segments

	None	Mountaineer	Wind
duration (hours)	6832.3	296.53	1364.2
mean number of events per hour	11.76	105.9	9.09

Table 4. Statistics of the automatically annotated dataset (2017 only) per annotation category. "None" represents the category when none of the other categories apply. Given are the total duration of annotated segments per category and the mean number of STA/LTA events per category.

4.1 Classifier Evaluation

The results of the classifier experiments from Sect. 3.2 are listed in Table 3. The table shows that the footstep detector is the worst at differentiating events. Both convolutional neural networks score a lower error rate of 0.0096 (MicroseismicCNN) and 0.0313 (Seismic Event Classifier). For the given dataset our proposed MicroseismicCNN network outperforms the seismic event classifier, in both the event-linked segment experiment as well as the image-linked segment experiment. The MicroseismicCNN using a longer input window (trained on image-linked segments) is comparable to classification on images and outperforms the classifier trained on event-linked segments. When combining image and micro-seismic classifiers the best results can be achieved.

4.2 External Influences

In Sect. 2.4.1 and 2.4.4 we have shown that external influences can have a strong impact on the quality of analysis. From Table 4 it can be deduced that the average number of triggered events per minute is high for times when the signal is influenced by

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

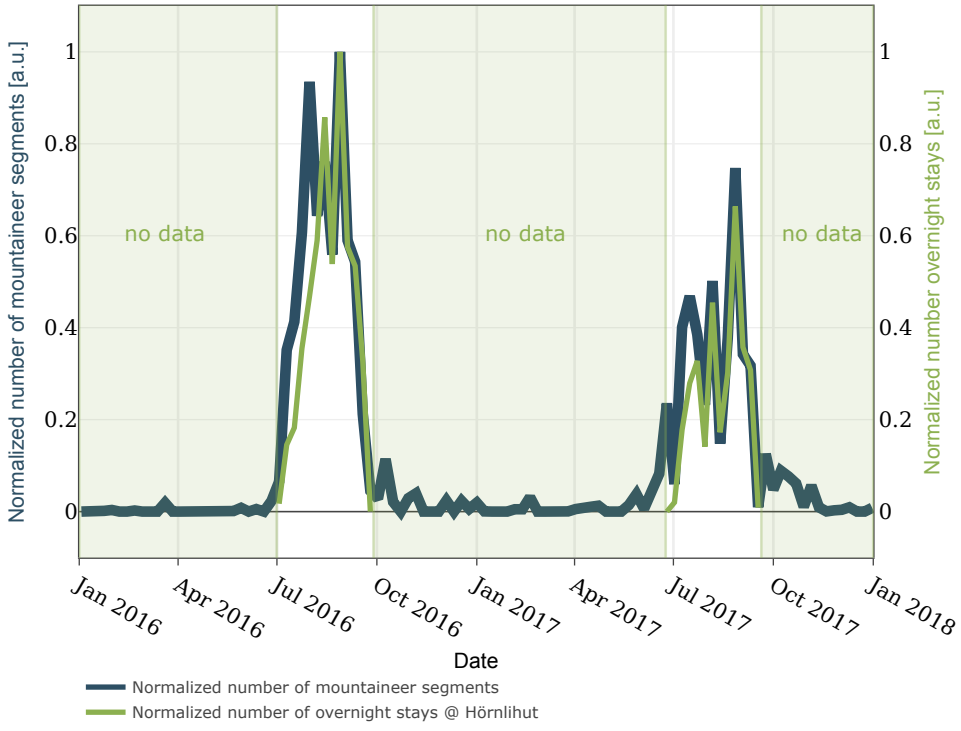

Figure 10. Correlation of mountaineer activity and hut occupancy. The normalized number of mountaineer segments per week and the normalized number of overnight stays at the Hörnlihut per week plotted over time.

mountaineers (105.9 events/hour), which is an approximately 9x increase in comparison to periods without annotated external influences. This shows that mountaineers have a strong impact on the analysis and a high activity detected by the event trigger does not correspond to a high seismic activity, thus relying only on this kind of event detection may lead to a false interpretation. The effect of wind influences on event rate is not as clear as the influence of mountaineers. The values in Table 4 indicate a decrease of events per hour during wind periods, which will be briefly discussed in Sect. 5.2.

4.3 Automatic Annotation in a Real-World Scenario

The results of applying the ensemble classifier to the whole dataset is visualized for two time periods in Fig. 9. The figure depicts the event count per hour before and after removing periods of mountaineer activity, as well as the rock temperature, the overnight stays at the Hörnlihut and the total net radiation. From Fig. 9 (a) is becomes apparent that the mountaineer activity is mainly present during summer and autumn. An increase is also visible during increasing hut overnight stays. During winter and spring only few mountaineers are detected but some activity peaks remain. By manually review we were able to discard mountaineers as cause for most of these peaks, however further investigation is needed to explain their occurence.

Figure 9 (b) focuses on the defreezing period. The zero-crossing of the rock temperature has a significant impact on the event count variability. A daily pattern becomes visible starting around the zero-crossing. Since few mountaineers are detected in May these can be discarded as the main influence for these patterns. The total net radiation however indicates an influences of solar radiation on the event count. Further in-depth analysis is needed but this examples shows the benefits of a domain-specific analysis, since the additional information gives an intuition of relevant processes and their interdependences. After July 9th, the impact of mountaineers is significant and the event detection analysis becomes unreliable. Different evaluation methods are required to mitigate the influence of mountaineers during these periods. Figure 10 depicts that mountaineer predictions and hut occupancy correlate well, which indicates that the classifiers work well. The discrepancy in the first period of each summer needs further investigation. With the annotations for the whole timespan it can be estimated that from all events detected in (Weber et al., 2018), approximately 25% originate in time periods with mountaineer activity and should therefore not be regarded as seismic events.

5 Discussion

5.1 Labeling

The previous section has shown that an understanding of the collected data is necessary for a significant analysis. The creation of such an annotated subset, despite being time and labor consuming, is therefore an overhead which is outweighted by the benefits of a better analysis. For data annotation two main approaches can be followed, annotating the phenomena of interest (positive examples) or annotating the external influences (negative examples). For positive examples, which are used in (Yuan et al., 2018; Ruano et al., 2014; Kislov and Gravirov, 2017), we would be restricted to find seismic events in the sensor being influenced by compounding factors and thus no ground truth information except for experience by professionals can be used. Therefore, the strategy presented in this work creates an annotated dataset using negative examples, because they can be more ascertainable for human annotators allowing the annotation process to be outsourced. Also, additional sensors can provide a direct measure of possible influences.

5.2 Multi-Sensor Classification

In Sect. 3 multiple classifiers for different sensors have been presented. The advantages of classifying on micro-seismic signals are that continuous detection is possible and that no additional sensor is required. The classification accuracy of our presented convolutional neural network and the image classifier are comparable. Image classification has however the disadvantage of a low time resolution with a capture frequency of at maximum 15 images per hour. Continuous video recording would close the gap, but the complexity of the image classifier and the size of it's input results in a high processing time, which makes it unfeasible for video. The main advantage of images is that they can be used as additional independent sensor to augment and verify micro-seismic recordings. First, images can be used for annotations and second they can be used in an ensemble classifier to increase the overall accuracy. The different modalities strengthen the overall meaningfulness and make the classifier more

robust. Table 4 shows that during wind segments less events are triggered than for periods without a category. A possible explanation could be that the micro-seismic activity is superimposed by broadband noise coming from wind. For these time segments a different parameter set or event detection algorithm could improve the analysis. Shielding the seismometer from the wind would probably reduce the wind influences significantly, but the common approach to embed it into the ground is difficult to implement in steep rock and hard to reach regions. Nonetheless we have shown that with the presented method we are able to quantify the impact of external influences on a long-term scale.

5.3 Outlook

This work has only focused on identifying external influences, what we have shown to be a prerequisite for micro-seismic analysis. Future work lies in finding and applying specific analytical methods, especially finding good parameter sets and algorithms for each context. Additionally, the classifier could be extended to include helicopters and rockfalls.

6 Conclusions

In this paper we have presented a strategy to evaluate the impact of external influences on a micro-seismic measurement by categorizing the data with the help of additional sensors and information. With this knowledge a method to classify mountaineers has been presented. We have shown how additional sensors can be beneficial to isolate the information of interest from unwanted external influences and provide a ground truth in a long-term monitoring setup. Moreover, we have presented a mountaineer detector, implemented with a convolutional neural network, which scores an error rate of only 0.96 % (F1 score: 0.9167) on micro-seismic signals and a mountaineer detector ensemble which scores an error rate of 0.79 % (F1 score: 0.9383) on images and micro-seismic data. The classifiers outperform comparable algorithms. Their application to a real-word, multi-sensor, multi-year micro-seismic monitoring experiment showed that time periods with mountaineer activity have a approximately 9x higher event rate and that approximately 25% of all detected events are due to mountaineer interference. Finally, the findings of this paper show that an extensive, systematic identification of external influences is required for a qualitative analysis.

Code and data availability. The dataset is available under the DOI 10.5281/zenodo.1320835 and the accompanying code under the DOI 10.5281/zenodo.1321176.

Author contributions. Matthias Meyer, Samuel Weber, Jan Beutel and Lothar Thiele developed the concept. Matthias Meyer and Samuel Weber developed the code and maintained the data together with Jan Beutel. Matthias Meyer prepared and performed the experiments, as well as the visualizations. Matthias Meyer prepared the manuscript with contributions from all co-authors.

Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2018-60
Manuscript under review for journal Earth Surf. Dynam.
Discussion started: 6 August 2018

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The work presented in this paper is part of the X-Sense 2 project. It was financed by http://www.nano-tera.ch (ref. no. 530659). We would like to thank Tonio Gsell and the rest of the PermaSense team for technical support. We acknowledgement Kurt Lauber for providing us with hut occupancy data and the Air Zermatt helicopter company for providing us with helicopter flight data. We thank Lukas Cavigelli for insightful discussions.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015.

Allen, R. V.: Automatic Earthquake Recognition and Timing from Single Traces, Bulletin of the Seismological Society of America, 68, 1521–1532, 1978.

Amitrano, D., Grasso, J. R., and Senfaute, G.: Seismic Precursory Patterns before a Cliff Collapse and Critical Point Phenomena, Geophysical Research Letters, 32, L08 314, https://doi.org/10.1029/2004GL022270, 2005.

Amitrano, D., Gruber, S., and Girard, L.: Evidence of Frost-Cracking Inferred from Acoustic Emissions in a High-Alpine Rock-Wall, Earth and Planetary Science Letters, 341-344, 86–93, https://doi.org/10.1016/j.epsl.2012.06.014, 2012.

Anchal, S., Mukhopadhyay, B., and Kar, S.: UREDT: Unsupervised Learning Based Real-Time Footfall Event Detection Technique in Seismic Signal, IEEE Sensors Letters, 2, 1–4, https://doi.org/10.1109/LSENS.2017.2787611, 2018.

Bartholomaus, T. C., Amundson, J. M., Walter, J. I., O'Neel, S., West, M. E., and Larsen, C. F.: Subglacial Discharge at Tidewater Glaciers Revealed by Seismic Tremor, Geophysical Research Letters, 42, 6391–6398, https://doi.org/10.1002/2015GL064590, 2015.

Brown, J. R., Beroza, G. C., and Shelly, D. R.: An Autocorrelation Method to Detect Low Frequency Earthquakes within Tremor, Geophysical Research Letters, 35, https://doi.org/10.1029/2008GL034560, 2008.

Chollet, F. and others: Keras, 2015.

Colombero, C., Comina, C., Vinciguerra, S., and Benson, P. M.: Microseismicity of an Unstable Rock Mass: From Field Monitoring to Laboratory Testing, Journal of Geophysical Research: Solid Earth, https://doi.org/10.1002/2017JB014612, 2018.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L.: ImageNet: A Large-Scale Hierarchical Image Database, in: CVPR09, 2009.

Eibl, E. P. S., Lokmer, I., Bean, C. J., and Akerlie, E.: Helicopter Location and Tracking Using Seismometer Recordings, Geophysical Journal International, 209, 901–908, https://doi.org/10.1093/gji/ggx048, 2017.

Geometrics: Geode Exploration Seismograph Specification Sheet, version GeodeDS_v1 (0518), 2018.

Gibbons, S. J. and Ringdal, F.: 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, 2006.

Girard, L., Beutel, J., Gruber, S., Hunziker, J., Lim, R., and Weber, S.: A Custom Acoustic Emission Monitoring System for Harsh Environments: Application to Freezing-Induced Damage in Alpine Rock Walls, Geoscientific Instrumentation, Methods and Data Systems, 1, 155–167, 2012.

Glade, T., Anderson, M., and Crozier, M. J.: Landslide Hazard and Risk, John Wiley & Sons, Ltd, Chichester, West Sussex, England, https://doi.org/10.1002/9780470012659, 2005.

Hershey, S., Chaudhuri, S., Ellis, D. P. W., Gemmeke, J. F., Jansen, A., Moore, R. C., Plakal, M., Platt, D., Saurous, R. A., Seybold, B., Slaney, M., Weiss, R. J., and Wilson, K.: CNN Architectures for Large-Scale Audio Classification, arXiv: 1609.09430, 2016.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H.: MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, arXiv:1704.04861 [cs], 2017.

Ioffe, S. and Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, arXiv:1502.03167 [cs], 2015.

Keller, M., Yuecel, M., and Beutel, J.: High Resolution Imaging for Environmental Monitoring Applications, in: International Snow Science Workshop 2009: Programme and Abstracts, pp. 197–201, Davos, Switzerland, 2009.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, arXiv:1412.6980 [cs], 2014.

Kislov, K. V. and Gravirov, V. V.: Use of Artificial Neural Networks for Classification of Noisy Seismic Signals, Seismic Instruments, 53, 87–101, https://doi.org/10.3103/S0747923917010054, 2017.

Levy, C., Jongmans, D., and Baillet, L.: Analysis of Seismic Signals Recorded on a Prone-to-Fall Rock Column (Vercors Massif, French Alps), Geophysical Journal International, 186, 296–310, https://doi.org/10.1111/j.1365-246X.2011.05046.x, 2011.

Occhiena, C., Coviello, V., Arattano, M., Chiarle, M., Morra di Cella, U., Pirulli, M., Pogliotti, P., and Scavia, C.: Analysis of Microseismic Signals and Temperature Recordings for Rock Slope Stability Investigations in High Mountain Areas, Nat. Hazards Earth Syst. Sci., 12, 2283–2298, https://doi.org/10.5194/nhess-12-2283-2012, 2012.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-Learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Perol, T., Gharbi, M., and Denolle, M.: Convolutional Neural Network for Earthquake Detection and Location, Science Advances, 4, e1700 578, https://doi.org/10.1126/sciadv.1700578, 2018.

Reynen, A. and Audet, P.: Supervised Machine Learning on a Network Scale: Application to Seismic Event Classification and Detection, Geophysical Journal International, 210, 1394–1409, https://doi.org/10.1093/gji/ggx238, 2017.

Ruano, A. E., Madureira, G., Barros, O., Khosravani, H. R., Ruano, M. G., and Ferreira, P. M.: Seismic Detection Using Support Vector Machines, Neurocomputing, 135, 273–283, https://doi.org/10.1016/j.neucom.2013.12.020, 2014.

Senfaute, G., Duperret, A., and Lawrence, J. A.: Micro-Seismic Precursory Cracks Prior to Rock-Fall on Coastal Chalk Cliffs: A Case Study at Mesnil-Val, Normandie, NW France, Natural Hazards and Earth System Sciences, 9, 1625–1641, 2009.

Spillmann, T., Maurer, H., Green, A. G., Heincke, B., Willenberg, H., and Husen, S.: Microseismic Investigation of an Unstable Mountain Slope in the Swiss Alps, Journal of Geophysical Research, 112, https://doi.org/10.1029/2006JB004723, 2007.

Springenberg, J. T., Dosovitskiy, A., Brox, T., and Riedmiller, M.: Striving for Simplicity: The All Convolutional Net, arXiv:1412.6806 [cs], 2014.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R.: Dropout: A Simple Way to Prevent Neural Networks from Overfitting, J. Mach. Learn. Res., 15, 1929–1958, 2014.

van Herwijnen, A. and Schweizer, J.: Monitoring Avalanche Activity Using a Seismic Sensor, Cold Regions Science and Technology, 69, 165–176, https://doi.org/10.1016/j.coldregions.2011.06.008, 2011.

Weber, S., Faillettaz, J., Meyer, M., Beutel, J., and Vieli, A.: Acoustic and Microseismic Characterization in Steep Bedrock Permafrost on Matterhorn (CH), Journal of Geophysical Research: Earth Surface, 123, 1363–1385, https://doi.org/10.1029/2018JF004615, 2018.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., and Trujillo, J.: A Comparison of Select Trigger Algorithms for Automated Global Seismic Phase and Event Detection, Bulletin of the Seismological Society of America, 88, 95–106, 1998.

Yuan, S., Liu, J., Wang, S., Wang, T., and Shi, P.: Seismic Waveform Classification and First-Break Picking Using Convolution Neural Networks, IEEE Geoscience and Remote Sensing Letters, 15, 272–276, https://doi.org/10.1109/LGRS.2017.2785834, 2018.