Geophysical Journal International

Geophys. J. Int. (2013) **192,** 425–439 doi: 10.1093/gji/ggs036

Classifying seismic waveforms from scratch: a case study in the alpine environment

C. Hammer,[1] M. Ohrnberger[1] and D. Fäh[2]

[1] *Institute of Earth and Environmental Science, University of Potsdam, Karl-Liebknecht-Str.* 24-25, *D*-14476 *Potsdam, Germany. E-mail: hammer@geo.uni-potsdam.de*

[2] *Swiss Seismological Service, Institute of Geophysics, ETH Zentrum,* CH-8092 *Zürich, Switzerland*

Accepted 2012 October 17. Received 2012 October 17; in original form 2011 December 19

SUMMARY

Nowadays, an increasing amount of seismic data is collected by daily observatory routines. The basic step for successfully analyzing those data is the correct detection of various event types. However, the visually scanning process is a time-consuming task. Applying standard techniques for detection like the STA/LTA trigger still requires the manual control for classification. Here, we present a useful alternative. The incoming data stream is scanned automatically for events of interest. A stochastic classifier, called hidden Markov model, is learned for each class of interest enabling the recognition of highly variable waveforms. In contrast to other automatic techniques as neural networks or support vector machines the algorithm allows to start the classification from scratch as soon as interesting events are identified. Neither the tedious process of collecting training samples nor a time-consuming configuration of the classifier is required. An approach originally introduced for the volcanic task force action allows to learn classifier properties from a single waveform example and some hours of background recording. Besides a reduction of required workload this also enables to detect very rare events. Especially the latter feature provides a milestone point for the use of seismic devices in alpine warning systems. Furthermore, the system offers the opportunity to flag new signal classes that have not been defined before. We demonstrate the application of the classification system using a data set from the Swiss Seismological Survey achieving very high recognition rates. In detail we document all refinements of the classifier providing a step-by-step guide for the fast set up of a well-working classification system.

Key words: Time series analysis; Neural networks, fuzzy logic; Seismic monitoring and test-ban treaty verification; Early warning; Probability distributions.

1 INTRODUCTION

In the seismic observatory practice seismologists seek to separate transient signals from background noise one is not interested in. Besides the detection of seismic events we are often interested in classifying different types of seismic signals. In studies of natural seismicity artificial seismic events (e.g. quarry blasts) are a major source of error and thus we are interested in identifying those signals in order to exclude them from further investigations (Habermann 1987; Horasan *et al.* 2009). The successful identification of suspicious seismic events is one of the key issues in routine data processing. Another field of application are monitoring systems. Automatic classification systems are already in use in volcano seismology (e.g. Ohrnberger 2001; Langer *et al.* 2006) as different seismic signals are associated with different stages of volcano activity (e.g. McNutt 1996, 2002). Another promising area of application are alpine alarm systems, where seismic sensors are increasingly being used as monitoring and warning systems for mass movements (e.g Arattano 1999; Marchi *et al.* 2002; Rice *et al.* 2002).

Most observatories endeavor to identify transient signals during their daily data analysis. The careful analysis of waveforms by an experienced seismologist can give key insights into event parameters allowing to classify the observed signals with a high level of confidence. However, such a detailed inspection by an analyst is a time-consuming process, needs experience and may suffer from the subjective view of the observer. For these reasons automatic systems provide a valuable alternative: consistent and objective results are provided in short time enabling to scan large volumes of data while minimizing the workload for the observatory staff.

Traditionally the detection of seismic events is carried out using a [short-term average/long-term average (STA/LTA)] trigger as discussed by Withers *et al.* (1998). In most implementations a valid event detection is declared if a minimum number of stations is triggered (e.g. Baer & Kradolfer 1987; Ruud & Husebye 1992). After event detection a second analyst control is needed for classifying the detection into different event types. However, this common approach may cause several problems. Often just a few triggered stations are available. The reasons for this may be, on the one hand,

technical problems of individual stations or, on the other hand, low magnitudes of small local events inhibiting their confirmation in networks. Thus, seismic events should be at best identified on their signature at a single station. Therefore, an approach enabling the reliable identification of seismic events at a single station would be advantageous for many monitoring purposes. Furthermore, successfully classifying seismic events at a single station provides the basis for any well-working multistation approach. Improving single station statistics will improve results of any subsequent analysis (e.g. coincidence trigger).

Alternatively, the classification can also be carried out by automatic systems, which are mostly trained in a supervised fashion from a large pre-classified data set (e.g. Del Pezzo *et al.* 2003; Esposito *et al.* 2006; Beyreuther *et al.* 2008; Curilem *et al.* 2009; Kuyuk *et al.* 2011). However, it is a well-known fact that deficiencies of the training data can cause serious problems for a well-working application. An insufficient amount of training data often leads to improper class descriptions as available data may not cover the center body and/or the range of the complete distribution. That strongly limits the application of automatic classification systems. Especially when observing rare events or setting up a short-term monitoring project a large number of training data is generally not available, for example even though the use of seismic devices in alpine warning systems may be encouraging (e.g. Arattano 1999) actual applications show unsatisfactory results. Interesting events (i.e. rockfall, debris flow, etc.) cannot be identified reliably in an automatic fashion based on their seismic signature (e.g. NPRA 1994; Leprettre *et al.* 1996; Bessason *et al.* 2007). The variability of waveforms belonging to the same signal class (i.e. regarding the signal length) is not captured by the classifier leading to a large number of missed events and several spurious detections.

For this reason, we will focus in this study on the setup of an automatic classification system that requires a minimum amount of training data while enabling to recognize highly variable event patterns. We use a new training approach that was originally introduced for monitoring active volcanoes (Hammer *et al.* 2012) and which alleviates the issue of sparse training data. As soon as interesting events have been identified the classification system can be built-up from scratch allowing the continuous data stream to be scanned immediately for corresponding events. Due to a minimum amount of preparation time the algorithm provides a valuable tool for both daily observatory practice as well as many short-term monitoring applications. In order to demonstrate its usage and capabilities the method will be applied to a data set recorded in the Swiss Alps. We classify events according to their underlying source processes in earthquakes, quarry blasts and rockfalls. At first the data set is described in more detail, followed by the description of the set up of the automatic classification system including several steps to further improve the system accuracy. Finally, the results and consequently its usefulness for other monitoring systems are discussed in detail.

2 DATA SET

We apply the new tool to a data set recorded on a high-gain broadband network operated by the Swiss Seismological Service (SED). Seismicity in Switzerland is moderate (e.g Baer *et al.* 2007; Deichmann *et al.* 2008) with mostly small events with magnitudes below M_L 3.0 and a maximum focal depth of 30 km (Deichmann & Baer 1990). Additionally to tectonic events a large number of quarry blasts is recorded due to ongoing construction purposes. Less often rockfalls are recorded by stations in the alpine environment.

Out of the wealth of the continuous data recordings at the SED waveform archive we selected recordings of 3 min each of which containing an event embedded in noise. In this way a quasi-continuous data stream is simulated, which we call event-based data set (EBDS) in the following. We used single station data recorded from 2002 to 2010 at station FUSIO only. The station was equipped with a three component STS-2 sensor. The sampling frequency was 120 Hz. For details on the instrumentation see Baer *et al.* (2001). As events in different regions show a great variability of signal characteristics due to strong lateral heterogeneities of the crust (Fäh & Koch 2002) only events close to station FUSIO, that is events between 8.40° and 9.15°E and 46.15°and 46.60°N, were considered in a first application (Fig. 1).

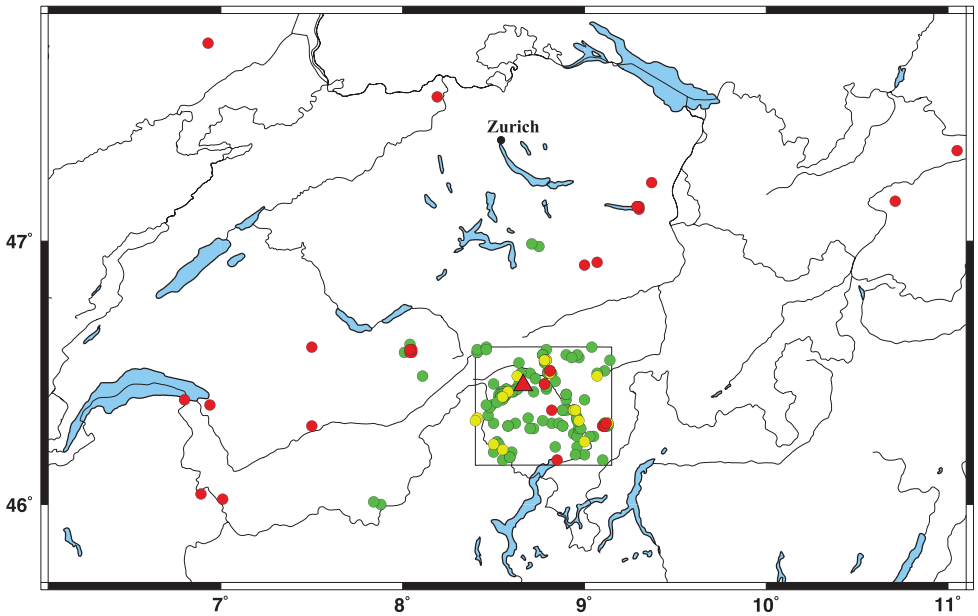

Figure 1. Distribution of events used as input for the classification system. The source area of the event-based data set is marked by a rectangle. Classification results are colour coded: ⟨green⟩ manual SED classification and automatic classification do not differ, ⟨yellow⟩ manual SED classification and automatic classification differ, ⟨red⟩ missed. Station FUSIO is shown by a red triangle.

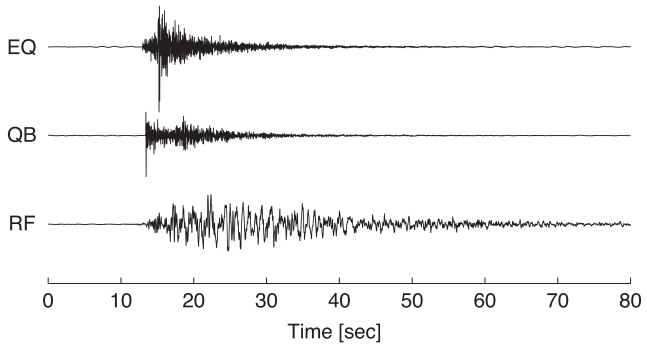

Figure 2. Example waveforms of the signal classes earthquake ⟨EQ⟩, quarry blasts ⟨QB⟩ and Rockfall ⟨RF⟩, recorded at station FUSIO, Switzerland.

The data set contains 159 earthquakes, 3 rockfalls, 46 quarry blasts and was pre-classified by the classification routine currently carried out at SED. Seismic transients are automatically detected by a classical STA/LTA trigger. An event detection is raised if a minimum of four stations is triggered. After an event is reported by the automatic system the signal is manually revised and classified. In case it can be positively confirmed the event enters the 'Earthquake Catalog of Switzerland' (ECOS). The expert classification is based on the appearance of the waveforms on the seismograms (Fig. 2). Earthquakes within the EBDS (i.e. within the source area between 8.40° and 9.15°E and 46.15° and 46.60°N) show magnitudes below M_L 2.5 and a maximum focal depth of 20 km leading to durations of less than 20 s. Corresponding spectrograms are dominated by frequencies above 5 Hz. The seismograms of quarry blasts are poorer in high frequencies than corresponding earthquakes in the investigated data set. The lack of high frequencies may be related to strong attenuation in the near surface layer and/or path effects due to very shallow focal depths. The long cigar shaped waveform of a rockfall signal shows durations of several tens of seconds and very low frequencies. Table 1 provides a summary of typical signal characteristics used in expert classification. The completely automatic classification procedure introduced in this paper is based on features extracted from the seismogram. Signal characteristics used for recognizing earthquakes, quarry blasts and rockfalls completely automatically comprise above expert criteria and and are discussed in detail in Section 3.

The data set above is restricted in space and time and thus may provide a biased estimate of the true system accuracy. In order to obtain a more realistic estimate of the true error we carried out further testing by classifying a data stream of 12 hr about every 3 month [hereafter referred as quasi-continuous data set (QCDS)]. As rockfalls are very rare in the EBDS we cannot provide confidence limits of the system accuracy on a statistically basis and therefore time periods were selected such as to close this gap. In every third month a period of 12 hr containing (if possible) a rockfall was selected for the QCDS. In case no rockfall is listed within this month, 12 hr containing earthquakes and/or quarry blasts were used. Within the QCDS, 11 rockfalls, 15 earthquakes and 3 quarry blasts took place all over Switzerland and neighboring regions (Fig. 1). None of the events is duplicated in both data sets (EBDS and QCDS) and thus both data sets can be considered as completely independent. The set up of the classification system is demonstrated first on the EBDS. The results of classifying the QCDS are discussed in Section 4.

3 REALIZATION OF A CLASSIFICATION SYSTEM FOR CONTINUOUS SEISMIC DATA

A pattern recognition system can be represented by three main elements (marked in grey in Fig. 3). In the feature generation step, a set of wavefield parameters is calculated from the observed data. Thus, the continuous seismogram is 'translated' into a sequence of parameter vectors, which provide a compressed signal representation. Based on those extracted features individual classifiers are constructed in the so-called training or learning phase. In supervised learning approaches classifier properties are learned from a prelabelled training data set. If class labels are missing or are to costly to obtain an unsupervised (i.e. cluster) method has to be used. After the classifier has been trained the classification of unseen signals can start. In order to further improve the system accuracy various iterations of the three-step approach may be needed for setting up a well-working classification system. For example we start the classification using a subset of available features that is chosen on *a priori* knowledge. After first events are recognized we may further adjust the feature subset to actual class properties. Updating classifiers may be needed due to dynamic changes in the observations reflecting for instance evolving source processes in active volcanic systems. Consequently, a new adaption of the models to the actual pattern would be required. This may be achieved by re-training the models using recently detected events. We describe the three main individual steps in the next paragraphs.

Feature Generation In order to enable the discrimination of different seismic signal portions a total set of 30 features is extracted from the seismogram. This compressed signal representation involves complex trace attributes (instantaneous bandwidth, instantaneous frequency, normalized envelope, centroid time, cepstral coefficients, half-octave-bands), spectral characteristics (predominant frequency, bandwith, central frequency) and polarization attributes (planarity, rectilinearity, largest eigenvalue). For details on particular features and references see Hammer *et al.* (2012). All features are computed in a sliding window of 3 s length in order to resolve frequencies down to 0.3 Hz. For the step size between successive computation of these short-term features we chose a value of 0.05 s. By calculating the above features for each 3 s window, the raw waveform is replaced by the time series of a feature vector in which each entry corresponds to one feature. Window length and step size have to be chosen appropriate for the actual application. Both values determine the temporal resolution of the corresponding event class. Large values smooth the extracted feature pattern while small values allow to capture subtle differences in signal properties. However, depicting very small structures may result in class patterns that

Table 1. Characteristics of considered signal classes of events within the pre-defined source area between 8.40° and 9.15°E and 46.15° and 46.60°N.

Name	Dominant frequency	Duration	Waveform characteristics
Earthquake	>5 Hz	<20 s	Impulsive *P* and *S* waves
Quarry blast	2−20 Hz	<30 s	Low *S*/*P* amplitude ratio
Rockfall	1−10 Hz	Up to several tens of seconds	Cigar shaped emergent signal, no separate *P* and *S* wave

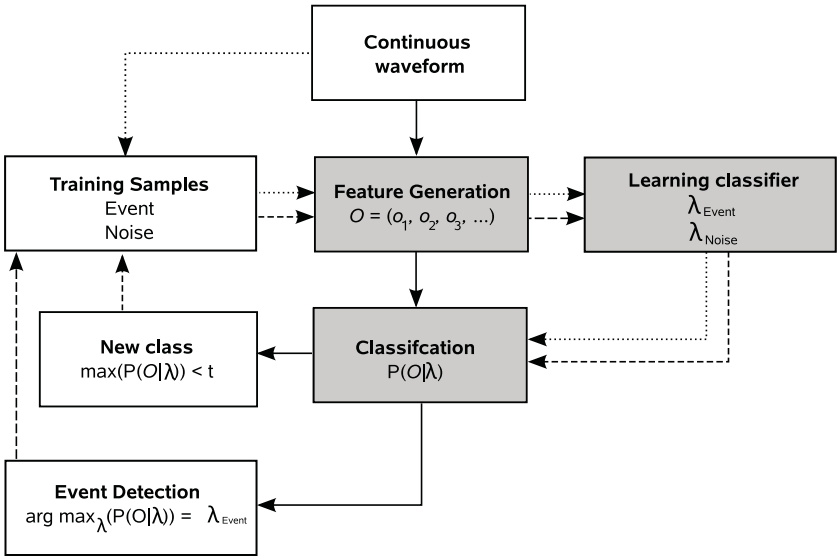

Figure 3. Elements of a pattern recognition system. Main elements are marked by grey boxes. The processing for an incoming data stream is processed as shown by the black arrows. In the training phase a model λ is learned for each class. In the classification task $P(O|\lambda)$ is evaluated for each λ. An event is detected if $P(O|\lambda_{\mathrm{Event}}) > P(O|\lambda_{\mathrm{Noise}})$. A new class is detected if the likelihood of all classes is lower than a threshold **t**. The training and re-training procedures are indicated by dotted arrows (λ_{Noise}) and dashed arrows (λ_{Event}).

go into too much detail, decreasing the classifier accuracy for the corresponding class. In this study we tested a step size of 0.25 s and 0.05 s and found 0.05 s to be more appropriate to discriminate earthquakes and quarry blasts (Table 3).

Classifier Design In this study we make use of a statistical classifier, called hidden Markov models (HMMs). HMMs are a special case of directed graphical models making use of the characteristic temporal structure of seismic signals. Observations (i.e. extracted features) are modelled by a sequence of multidimensional probability distributions whose characteristics (i.e. means, covariances) are learned from pre-labelled training data. For a detailed description of probabilistic graphical models and their possible variations the reader may refer to Koller & Friedman (2009). In order to exploit its advantage with respect to time and objectivity we apply a novel training procedure suggested by Hammer *et al.* (2012), which was motivated by a work of Wilcox & Bush (1992) in the speech recognition realm. Originally, this approach was introduced for classifying volcano seismic signals. We test its applicability in the current context, as, like in volcanic environments, the training sample collection may pose a problem. The algorithm was developed in order to reduce the dependence on previously acquired data bases and classification schemes. The clever trick consists in taking advantage of all the information available. General wavefield properties are modelled from lots of unlabelled continuous data streams. Then, this widespread background model is used to adjust event model descriptions from a single example waveform. Following the trajectory of features extracted from the reference event waveform in the overall feature space (derived from the background model) allows to construct corresponding HMMs from scratch. For details on the procedure see Hammer *et al.* (2012). Thus, there is no need for collecting many reference events for training the system. The positively confirmed samples, obtained in the first classification run, can then be used to re-train the models. This 'learning-while-recording' approach allows for automatic gathering of larger training data sets including the advantage of sorting the events objectively in different classes.

Classification After a model for each class of interest has been constructed we can start to classify an unknown data stream. This is done by asking for each of the available HMMs the question: 'What is the likelihood that the observation sequence has been generated by the HMM describing class *X*?' The winning model can then be found by the principle of maximum likelihood. Consequently, for an incoming continuous data stream each time frame can be assigned to a particular signal type, which can be a specific seismic event or noise. In this way no preceding trigger is required. Two classifiers are running in parallel (Fig. 4). One containing all classes and one containing the noise only. The class whose model provides the highest score (i.e. the highest likelihood) is determined in a sliding window of 15 s length with 10 s overlap between adjacent frames. Within each window the noise classifier calculates the log likelihood that the contained signal has been generated by the noise model. The noise is modelled without any temporal structure (ergodic process) and can be repeated arbitrarily within the detection window. The second classifier evaluates the log likelihood of an event embedded in this window. The event can appear somewhere in the window and can be preceded and/or followed by noise (Fig. 4). After the final step of comparing likelihoods of both classifiers the detection window is moved by 5 s (100 frames) and the procedure starts again. An example for a successful classification is given in Fig. 5. It is shown that the log likelihood of the earthquake class increases when the earthquake appears on the continuous recording while the log likelihoods of all other classes decrease at the same time. An event detection is raised as soon as the log likelihood of an event class is larger than the log likelihood of noise.

When evaluating the classifier performance, we can distinguish two different ways for looking at the problem. First, we can consider the task as a classification problem. In this case, we classify the unknown data stream into one of multiple classes (here noise, earthquakes and quarry blasts). Results corresponding to the classification problem are called classification or recognition rates in the following. Alternatively, we may consider the task as a detection problem. In this case, the classification problem is reduced to a two-class problem comprising only the classes noise and non-noise

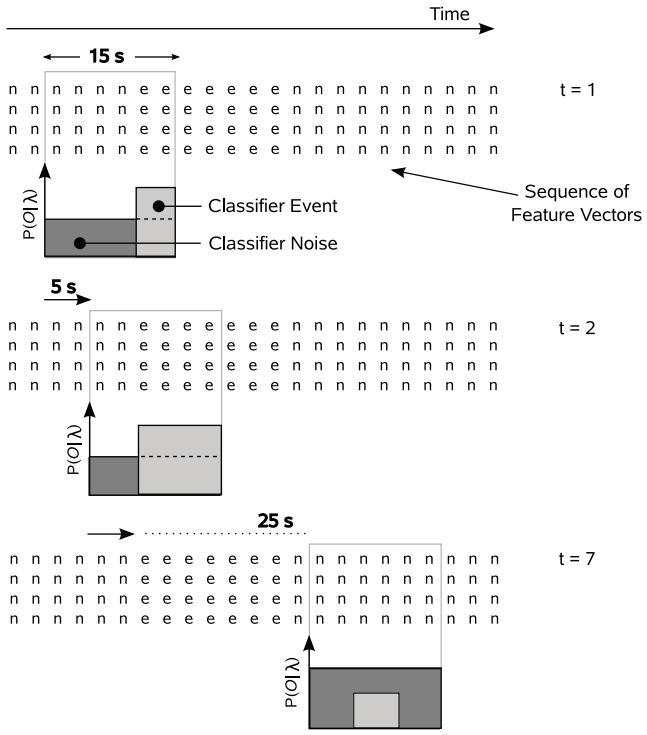

Figure 4. Sketch of the classification procedure. A window (marked by the grey rectangle) is shifted along the sequence of extracted feature vectors. Feature vectors are coded: ⟨*n*⟩ noise present; ⟨*e*⟩ event present. Actually feature vectors are sampled at a rate of 20 Hz (successive time windows at 0.05 s). Within the window two classifiers are evaluated. The first (displayed in dark grey) gives the likelihood that the window is completely filled with noise. The second (displayed in light grey) calculates the likelihood of an event embedded somewhere in this window, which can be preceded/followed by noise. Note, that only the likelihood of the event is shown here, likelihoods of preceding/following noise have been omitted for the sake of clarity. After both classifiers are evaluated the window is shifted by 5 s. While at t_1 and t_2 an event detection would be raised ($P(O|\lambda_{\text{Event}}) > P(O|\lambda_{\text{Noise}})$) no event would be detected at t_7 ($P(O|\lambda_{\text{Event}}) < P(O|\lambda_{\text{Noise}})$) .

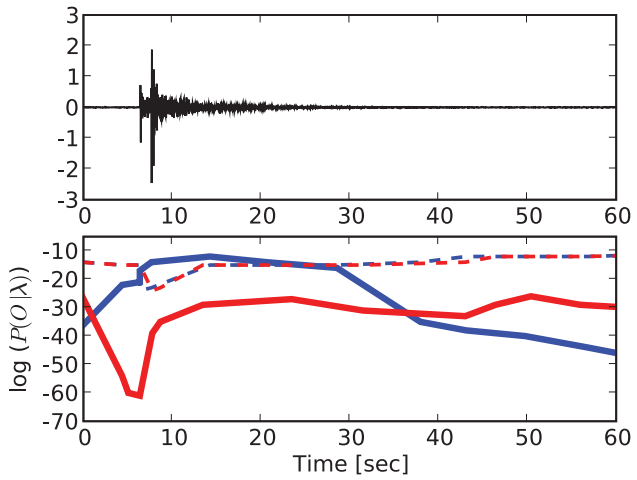

Figure 5. Results of the continuous classification are shown. The top row shows a seismogram recorded at station FUSIO. On the bottom row the HMM log likelihood for model ⟨earthquake⟩ is displayed in blue ($P(O|\lambda_{EQ})$), the log likelihood for model ⟨quarry blast⟩ is shown in red ($P(O|\lambda_{QB})$). The log likelihood of the in parallel running noise classifier is marked by dashed lines. Each time step the winning class can be calculated. An event detection is declared if the log likelihood of an event class is larger than the log likelihood of noise. Between 7 and 22 s an earthquake is detected as the corresponding model achieves larger log likelihoods than the noise and quarry blast model.

signal (here earthquakes and quarry blasts). Hence, the number of detected events is the sum of correctly classified events and confused events. Corresponding results are called detection rates in the following.

For volcano induced signals the algorithm as implemented by Hammer *et al.* (2012) performs very well. Still, we describe in the next sections how to further simplify and automate the existing training procedure. The set up of the final well-working system including all improvements is presented as a step-by-step guide comprising the two main parts

- Initial training of the system (baseline system),
- Refinement of classifier (improved system).

3.1 Initial training

The setup of the automatic classification system is started with choosing a segment of several hours from the continuous data stream as unlabelled training data. Additionally, we select one reference waveform for each signal event class of interest as training example. In this study the unlabelled data set consists of 12 hr continuous recording from January 2002 covering both day and night time. The reference events are taken from the same time period. In the following we consider signals related to earthquakes and quarry blasts only. We deliberately choose not to train a classifier for rockfall signals as to demonstrate the system behaviour to unknown event classes (Section 3.2.2).

The initial training procedure used by Hammer *et al.* (2012) can be summarized in three steps. First, features described in Section 3 are extracted in a sliding window. Second, those parameters are used to learn multi-Gaussian mixture densities which are then used to describe the overall background wavefield. In the last step, a HMM for each event class of interest is constructed from a single reference waveform and the overall background model. Although this algorithm performs very well, there is still room for improvement. First, the discrimination of signal classes can be optimized by feature selection. Second, the number of components in the Gaussian mixture model can be determined in an automatic fashion further minimizing the required workload. Both steps are described in detail in the next sections.

3.1.1 Feature selection

In Hammer *et al.* (2012) a fixed feature set has been used for the description of different signal classes. It consists of 30 individual short-term estimates of the wavefield. Given the large dimension of the feature space the HMM classifier requires to learn a large number of parameters. For each additional feature another set of variables [consisting of a mean and corresponding covariances for each mixture component (see Section 3.1.2)] is added to the problem. Thus, in order to keep the number of parameters as low as possible and to maximize the discriminative power between different classes only the most suitable features are desired as input for the recognition system.

Although the single reference waveform does not represent all members of the corresponding class it may help to decide for a more appropriate feature subset. In literature several automatic

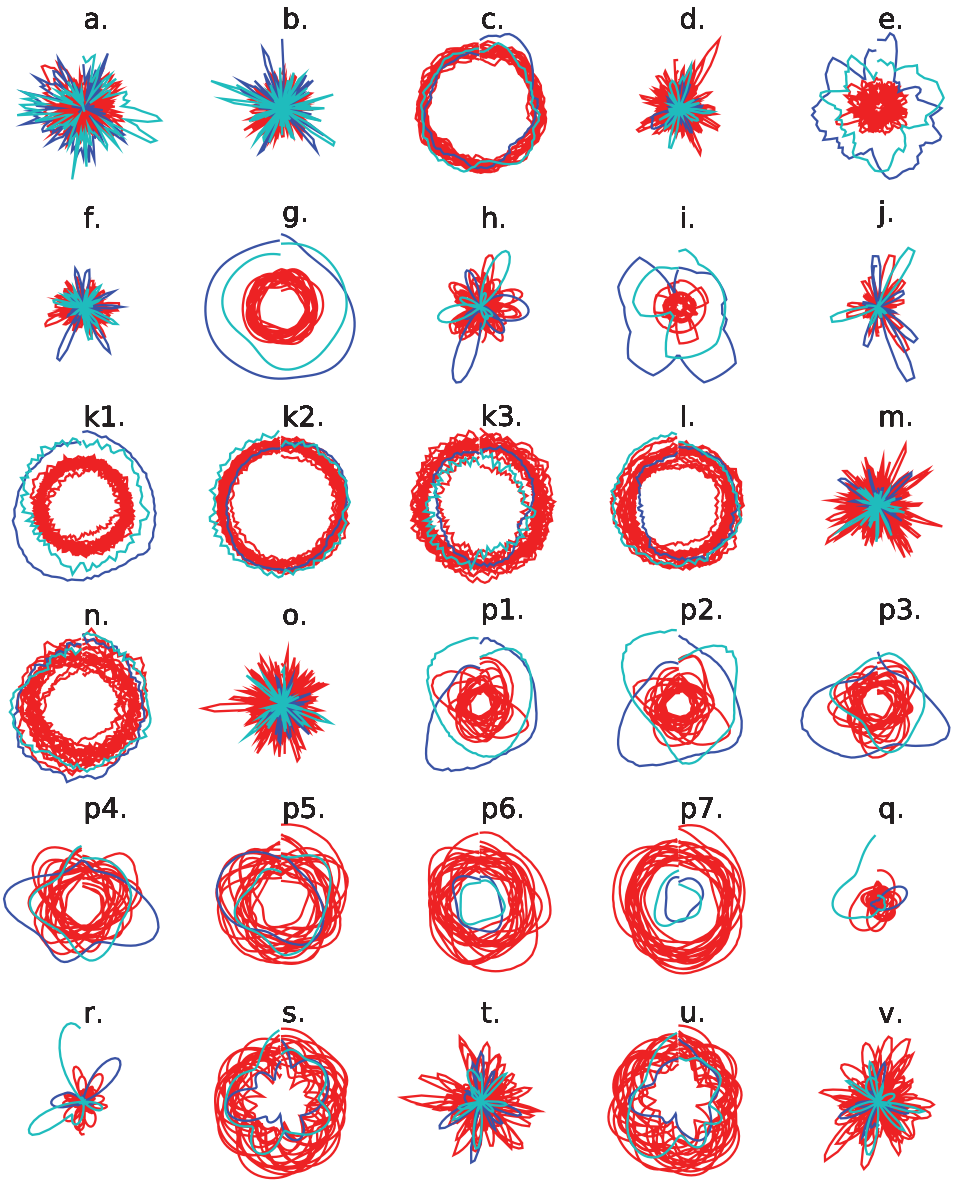

Figure 6. Normalized features are shown. Signal classes are colour coded: ⟨blue⟩ earthquake, ⟨cyan⟩ quarry blast, ⟨red⟩ noise. The time axis is wrapped around the circle. From top left to bottom right: (a) norm. envelope $\hat{A}$; (b) its time derivative; (c) centroid time; (d) its time derivative; (e) bandwidth; (f) its time derivative; (g) central frequency; (h) its time derivative; (i) dominant frequency; (j) time derivative; (k1)–(k3) three cepstral coefficients; (l) inst. frequency; (m) its time derivative; (n) inst. bandwidth; (r) its time derivative; (p1)–(p7) seven half octave bands; (q) largest eigenvalue; (r) its time derivative; (s) rectilinearity; (t) its time derivative; (u) planarity; (v) its time derivative.

procedures have been suggested for this task. Often the Karhunen-Loève transform is used to reduce the dimensionality of the problem as it sorts the features according to their degree of information. Therefore, small components may be dropped with minimal loss of information. However, the physical meaning of the new features generated by linear combinations is not clear. Another approach has been suggested by Köhler *et al.* (2009), who applied a self-organizing map approach in order to asses the in-between feature correlations. However, a sufficiently number of reference signals is necessary in order to cluster the features in groups with similar information content.

In the approach presented here we have only limited access to typical class characteristics (single reference waveform). For that reason, the most appropriate feature subset was chosen by visual inspection of all features because in doing so we are still able to include prior knowledge about relevant criteria to distinguish different signal types. This can be carried out by means of polygon plots as shown in Fig. 6. While for detection purposes it would be enough to use features where event classes are clearly separated from the noise data, for classification purposes also different event types have to be clearly separated. Based on Fig. 6 we decided for the bandwidth, the central frequency as well as its time derivative, the dominant frequency, the cepstral coefficients, the half-octave bands and the largest eigenvalue of the polarization ellipsoid. By using only one reference waveform in this procedure we neglect the existing variability of signals within a given class. For that reason, the feature selection may be re-evaluated after initial successful classification of events then further improving the

Table 2. The coefficient J for 8,16 and 32 mixture components.

No of mixture components	8	16	32
$J = \frac{\text{trace}\{S_W+S_B\}}{\text{trace}\{S_W\}}$	1.37	1.65	1.44

recognition accuracy. In order to demonstrate the gain when using only a subset of available features we also present the obtained results when using all 30 parameters.

3.1.2 Number of Gaussians in overall background model

Following Hammer *et al.* (2012) a HMM for the background is estimated from an unlabelled data stream. For this purpose the feature distribution observed in the unlabelled data stream is modelled by a Gaussian mixture density (i.e. a weighted sum of M Gaussian densities). This Gaussian mixture distribution is then used to describe the chosen reference events. The number M of mixture components, which is appropriate to describe the overall wavefield characteristics has to be chosen in advance and is related to the capability to describe the reference events. If the number of mixture components is too small the center body and/or range of the feature distribution may be modelled improperly resulting in poorly modelled reference patterns. However, if we decide for too many mixture components, the number of parameters is increased and the model tends to overfitting. We therefore need an objective and useful criterion in order to decide which number of Gaussians is most appropriate to describe the overall feature space.

A simple and effective method is provided by scatter matrices. In classification approaches scatter matrices are in general used as a measure for class separability (Theodoridis & Koutroumbas 2006). If we assume each mixture component as representative of one cluster (i.e. a class) in the overall feature space we can compare the within-cluster scatter matrices, S_W, and the between-cluster scatter matrices, S_B. Then, the coefficient

$$J = \frac{\text{trace}\{S_W + S_B\}}{\text{trace}\{S_W\}} \tag{1}$$

takes large values if the data are well clustered around their mean within each mixture component, and the different mixture components are well separated. In this case, the selected number of Gaussians would 'resolve' the feature space just enough to discriminate different event classes. Hence, we can decide for an appropriate number of Gaussians by evaluating J for a different number of mixture components. In this study we tested 8, 16 and 32 Gaussians for modelling the background wavefield and found 16 to be best discriminated as J takes the largest value (Table 2).

3.1.3 Classification results for baseline system

After having fixed the parameters of our baseline system as discussed before we start to learn individual event classifiers according to Hammer *et al.* (2012). These draft models are then used to scan the incoming data stream for corresponding events. The incoming data stream consists here in the EBDS, classification results for the QCDS are discussed in Section 4. In the following we make use of the problem descriptions introduced in Section 3 for discussing the results (i.e. detection and classification problem). In the first run 69 per cent of all the events are detected (i.e. recognized independent of type) and 61 per cent of all events are classified correctly (i.e. classified according to manual *a priori* label) (Table 3 and Fig. 7). The difference of 8 per cent corresponds to events where the automatic label and the manual *a priori* classification differ (so-called false type detections). Therefore, it is also called confusion error. Although the models are based on one reference waveform only, most of the events are recognized correctly. However, a larger number of spurious detections is disturbing the overall performance of the HMM classifier. Considering their durations of a few seconds (Fig. 8) this problem can be avoided by requiring a minimum detection length for a valid event recognition. An appropriate threshold depends on the investigated data set. In this study we choose a minimum detection length of 5 s as shortest events show durations of approximately 5 s. Detections with durations less than 5 s are discarded. By applying this rule, the number of false alarms decreases significantly from 52 to 13 while recognition rates for classified events, missed events and confused events do not change. Increasing the minimal detection length to 6 s decreases the number of classified events to 58 per cent. For that reason, we retain a minimum detection length of 5 s in the subsequent sections.

Alternatively to using only a subset of features we can use all features. However, in this case recognition accuracy decreases to 57 per cent. The description of individual event classes is less appropriate as shown by the larger number of missed events. Furthermore, the discriminative power between classes seems slightly reduced as there is one more misclassified event in this configuration.

Both configurations show a large number of missed events (Table 3 and Fig. 7). However, it is no surprise that the draft models derived from one single training example are to not able to capture the center body and/or range of existing waveform variabilities within a class. A similar, but even more restrictive approach provide cross-correlation-based techniques. Analogous to the proposed method only one reference waveform is needed per class. The reference waveform is shifted along the continuous seismic signal. Each time the cross correlation coefficient exceeds a given threshold (here 0.5 and 0.6) an event is detected. Similar to above, only events with durations of 5 s or more are accepted. However, no variations of the waveforms are allowed within a class. Consequently, the cross-correlation performs significantly less than our approach (Table 4). Furthermore, the HMM-based procedure can be easily optimzed. By further adjusting the models to individual event classes we are able to significantly improve the recognition accuracy. This can be done in an partially automatic way and is explained in the next sections.

Table 3. Summary of classification results using draft models. The second last column shows the results for extracting the features in 3 s windows with a time step of 0.25 s between successive windows. The last column shows results using all 30 features.

	# Earthquakes	# Blasts	# Total	# Total minDur_5	# Total, $\Delta_t = 0.25$s minDur_5	# Total, fv_{30} minDur_5
Correct class.	96 of 158	27 of 45	123 (61 per cent)	123 (61 per cent)	89 (44 per cent)	115 (57 per cent)
False type	7	10	17 (8 per cent)	17 (8 per cent)	47 (23 per cent)	18 (9 per cent)
Missed	55	8	63 (31 per cent)	63 (31 per cent)	67 (33 per cent)	70 (34 per cent)
False alarm	43	9	52	13	16	13

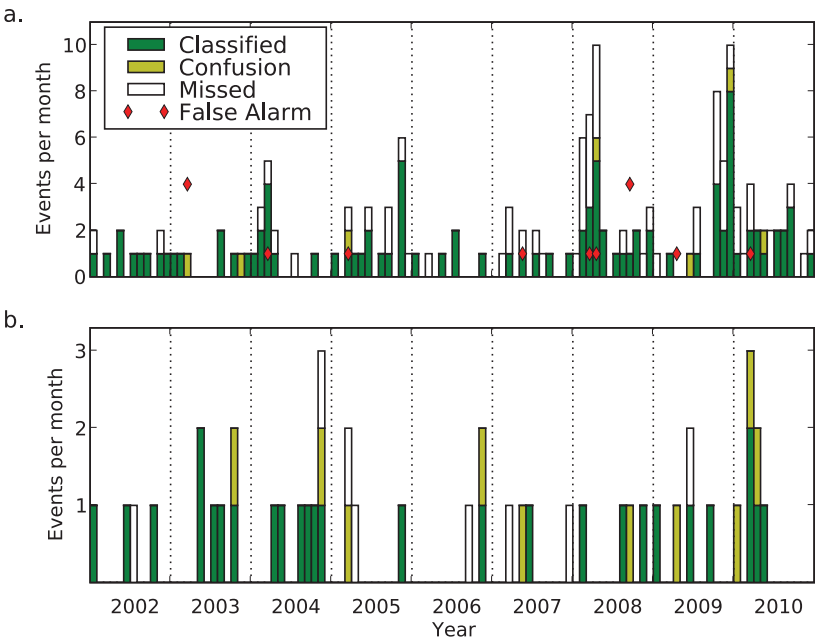

Figure 7. The classification results for classifying earthquakes (a.) and quarry blasts (b.) are shown. A minimum detection length of 5 s is required for reducing the high false alarm rate. The term 'Classified' refers to events were manual SED classification and automatic classification do not differ. The term 'Confusion' refers to events were manual SED classification and automatic classification differ (false type).

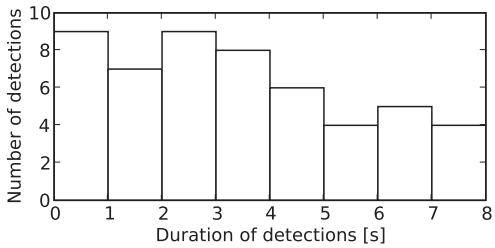

Figure 8. Histogram of detection lengths of spurious detections are shown.

3.2 Refinement of classifier

In order to capture existing waveform variabilities we re-estimate the classifier properties. Finally, we expand the system to enable the detection of unknown signal classes.

3.2.1 Adaption of noise and event models

The model parameters can be re-estimated from the positively confirmed samples and thus a better adaption to the actual event pattern is ensured (Hammer *et al.* 2012). For the investigated data set the training data set is increased by 20 to 30 events per year (Fig. 7). Thus, we re-train the model not only once but on a regular basis in order to take all possible variations of class properties into account. However, generally this approach should be handled with caution. Typical class characteristics may eventually become more diffuse ('blurred') resulting in a reduction of the discrimination power between classes due to a higher degree of overlapping between individual classes. Several options are conceivable to circumvent this negative effect of repeated re-training. One possibility is to 'forget' the earliest training samples. Alternatively, the iterative training cycle is not run until convergence but is stopped after few iterations on the most recent events. By doing so the model will steer towards the actual class pattern (Riggelsen & Ohrnberger 2012). Re-estimating the event model parameters from positively confirmed samples increases detection rates from 69 per cent (Table 3) to 95 per cent (Table 5). The number of spurious detections is decreased by 6 to 7 false alarm for the complete time period and there are almost as many confused events as before (8 per cent and 11 per cent).

Besides a better adaption of event models to actual class patterns also adjustment of the noise model is highly recommended. Regular daily and seasonal variations (e.g. Sheen *et al.* 2009; Hillers & Ben-Zion 2011) of the noise characteristics play an important role. Furthermore, noise characteristics may also change due to environmental modifications such as construction works. Events in the EBDS are embedded in noise. Consequently, we decide to adapt the noise model regularly when classifying the EBDS. We re-train the noise model at the beginning of each month using 12 hr of continuous recording. When running a continuous classification over a long time span the continuous updating of the noise model has been promising (Riggelsen & Ohrnberger 2012). Re-training the noise model shows its biggest impact on the number of spurious

Table 4. Summary of classification results using cross-correlation. A cross correlation coefficient of 0.5 and 0.6 is required to declare an event detection.

	# Earthquakes	# Blasts	# Total
Detection if cross-correlation coefficient > 0.5			
Correct class.	37 of 158	10 of 45	47 of 203 (23 per cent)
Detection if cross-correlation coefficient > 0.6			
Correct class.	27 of 158	8 of 45	35 of 203 (17 per cent)

Table 5. Summary of final classification results. $\hat{\lambda}_{\text{Event}}$ and $\hat{\lambda}_{\text{Noise}}$ indicate re-trained models. The last column shows the results using all 30 features.

	# Earthquakes	# Blasts	# Rockfalls	# Total $\hat{\lambda}_{\text{Event}}$	# Total $\hat{\lambda}_{\text{Event}}$, $\hat{\lambda}_{\text{Noise}}$	# Total, fv_{30} $\hat{\lambda}_{\text{Event}}$, $\hat{\lambda}_{\text{Noise}}$
Correct class.	138 of 158	35 of 45	3 of 3	173 (84 per cent)	176 (86 per cent)	161 (78 per cent)
False type	16	7	0	23 (11 per cent)	23 (11 per cent)	27 (13 per cent)
Missed	4	3	0	10 (5 per cent)	7 (3 per cent)	18 (9 per cent)
False alarm	0	1	0	7	1	3

detections, which is decreased to 1 for the complete time period. Additionally, detections rates are increased to 97 per cent while the number of confused events does not change. Results are summarized in Table 5.

3.2.2 Detection and processing of unknown signal classes

In the observatory practice there is a large chance that at some point the classifier is confronted with an unknown signal that is not part of the defined classification scheme. In case we want to detect such unseen patterns the classification task is reduced to a detection problem, asking the question, 'Does the observed sequence belong to any of the defined signal classes, or does it represent a novel type?'. Given a trained HMM, the sample likelihood of an observed sequence with respect to the model can be computed for each time step. In case the signal is not described properly by any of the available models, the log likelihoods for all classes should be low (Fig. 9). Therefore, it is in principle possible to detect such novel event types by monitoring the likelihood of all models. A threshold on the log likelihood is then needed to discriminate against being a member of the given set of HMMs or being a new signal type.

In the field of 'novelty detection' (Markou & Singh 2003) the threshold for abnormal behaviour is mostly chosen as the minimum likelihood among all available training sequences (Yeung & Ding 2003) or with the help of extreme value theory (Strachan & Clifton 2009). However, both approaches are biased by the available test data. Alternatively, the 'principle of indifference' (in Bayesian framework also called simplest non-informative prior) provides a suitable framework for this task as already suggested by Ohrnberger (2001). If we assume all observations as equally likely the model does not depend on the observations and can be seen as a baseline in the detection process. A model generating a signal with less likelihood than the uniform model cannot be regarded as a proper description of the observation sequence. Thus, a threshold is calculated as explained in the following. The continuous uniform distribution is defined as a probability function such that each observation within the interval ($o_{\min}$, $o_{\max}$) is equally probable (Frank & Althoen 1994). In the D-dimensional feature space the probability density function of the continuous uniform distribution is defined as

$$f(\mathbf{o}) = \prod_{d=1}^{D} \frac{1}{o_{d,max} - o_{d,min}}. \quad (2)$$

In this study, appropriate values for $o_{\min}$ and $o_{\max}$ are taken from the 99 per cent confidence interval of the overall output distribution. $f(\mathbf{o})$ is only defined for observation within the intervall ($o_{\min}$, $o_{\max}$). In case an observation lies outside this range the value of $o_{\min}$ or $o_{\max}$ is replaced with the corresponding observed value. Alternatively, one can use a class specific threshold. The threshold is re-calculated each time the models are changed (i.e. re-estimated). If the log likelihood of all available HMMs is lower than the log likelihood of the corresponding uniform model, the signal is assumed as not belonging to any of the defined signal classes. The presence of a new signal type has been considered as valid, if all HMMs provide a likelihood lower than the threshold for at least 5 s. As soon as an unknown signal type has been confirmed manually, a corresponding HMM can be build immediately in order to detect those signals in the future.

By applying the given criteria to the event-based data set three segments have been found that cannot be assigned to any of the defined signal classes. All detected segments correspond to rockfall events, which have not been defined before in our proposed classification task. As soon as the first rockfall has been identified a corresponding HMM has been trained and integrated in the system. The two following rockfalls are then classified correctly by the recently added HMM.

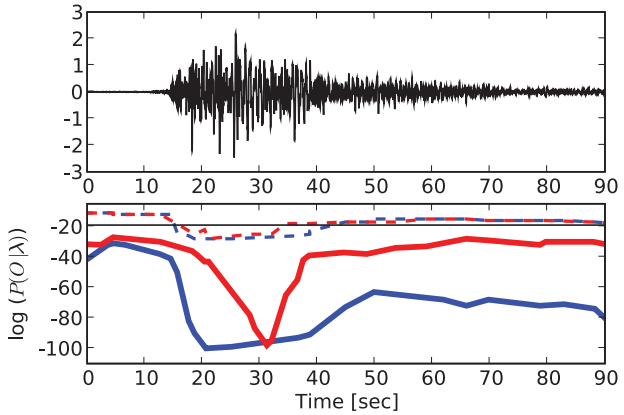

Figure 9. Recognition of unknown signals are shown. The top row shows a rockfall signal recorded at station FUSIO. On the bottom row the log likelihoods of the different classes (earthquake ⟨blue⟩, quarry blasts ⟨red⟩, solid lines) and the in parallel running noise classifier (dashed lines) are plotted. For an unknown signal class all log likelihoods are reduced at the same time (i.e. at about 15 s). If the log likelihood is lower than a threshold (horizontal black line) the corresponding model can not be assumed as a proper description of the given observation sequence.

4 DISCUSSION OF RESULTS

In this study 97 per cent of all events within the considered source area of 8.40°–9.15°E and 46.15°–46.60°N are detected in the automatic classification process (Fig. 10, Table 5) by the improved system (after re-training). While none of the earthquakes or quarry blasts is classified as rockfall the confusion error between earthquakes and quarries is about 11 per cent.

In the evaluation process the automatically assigned class is compared to the manual set reference. The human analyst may be trained to achieve highly consistent classification results but there still remains a subjective impact. Classification results of one observer might not be comparable to results given by another expert. Thus, part of the misclassified events might be

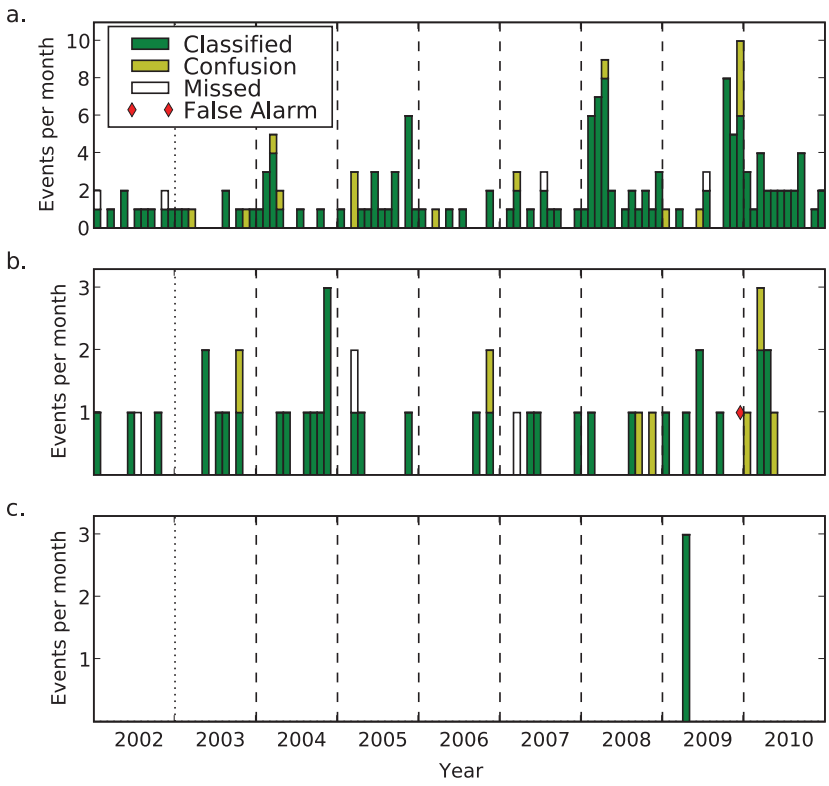

Figure 10. The classification results for classifying earthquakes (a), quarry blasts (b) and rockfalls (c) are shown. A minimum detection length of 5 s is required. The term 'Classified' refers to events were manual SED classification and automatic classification do not differ. The term 'Confusion' refers to events were manual SED classification and automatic classification differ (false type). The system was re-trained at the beginning of 2004, 2005, 2006, 2007, 2008, 2009 and 2010 (dashed lines) using all previously detected events.

related to the *a priori* classification, which is difficult if quarry blasts and tectonic events occur in the same area and earthquakes are shallow.

In other studies the most promising discriminants for earthquakes and quarry blasts have been spectral characteristics such as the spectrogram, the frequency content of certain wave groups or various types of spectral ratios (e.g. Wuster 1993; Gitterman *et al.* 1998; Fäh & Koch 2002; Del Pezzo *et al.* 2003). In this study the spectrogram information is captured in five half-octave bands. In combination with central frequency and dominant frequency the features implicitly include discriminants such as spectral peaks of certain wave groups or the S/P spectral amplitude ratio. Consistently with results of other studies (e.g. Plafcan *et al.* 1997; Ursino *et al.* 2001; Allmann *et al.* 2008) the seismograms of quarry blasts are poorer in high frequencies than corresponding earthquakes in the investigated data set (Fig. 6). Besides spectral characteristics time domain properties as amplitude ratios between different wave groups have been successfully discriminated explosions from earthquakes (e.g. Baumgardt & Young 1990; Wuster 1993). Additional to waveform characteristics also the origin time has been suggested as a powerful criterion for discriminating earthquakes and quarry blasts Wiemer & Baer (2000). However, in some regions there is no direct correlation between origin time and signal type (Kuyuk *et al.* 2011).

Based on these facts we carefully re-revised all misclassified events with the help of N. Deichmann. Sixteen events *a priori* labelled as earthquakes were automatically recognized as quarry blasts. Eleven out of 16 took place during night. To our knowledge, there have not been any blasts outside working hours in Switzerland. For that reason we assume the *a priori* identification as earthquake as correct. Two out of the remaining five events were located close to Faido and are assumed to be induced earthquakes related to the Gotthard Base Tunnel (Baer *et al.* 2005). Another event was part of the Val Bavona sequence (Deichmann *et al.* 2009). The two remaining events were located in Val Formazza close to several quarry sites. Supported by emergent S-phases and dominating frequencies below 15 Hz (Fig. 11) we conclude that both events were incorrectly labelled as earthquakes and show signals caused by quarry blasts instead.

In summary 14 earthquakes have been misclassified as quarry blasts. The difficult discrimination between both signal types due to strong wavefield similarities has been reported by several authors (e.g. Ursino *et al.* 2001; Kuyuk *et al.* 2011). Ten out of 14 misclassified events are localized at a depth of less than 3 km. Assuming the focal depth (i.e. its influence on the frequency content) as a discriminative feature one reason for these confusions may be the small focal depth. The shallow events are dominated by low frequencies due to the strong attenuation of the near surface layers similar to quarry blasts. Propagation effects in the vicinity of the source appear to mask differences between the seismic signals of earthquakes and quarry blasts. Similar, Plafcan *et al.* (1997) demonstrated that the crustal structure and path effects determine the seismic character of earthquakes and explosions to a greater extent than the corresponding source mechanism.

Seven of the events manually labelled as quarry blasts (in total 45) were automatically classified as earthquakes. All of them were located close to active quarries. However, rather unusual for quarry

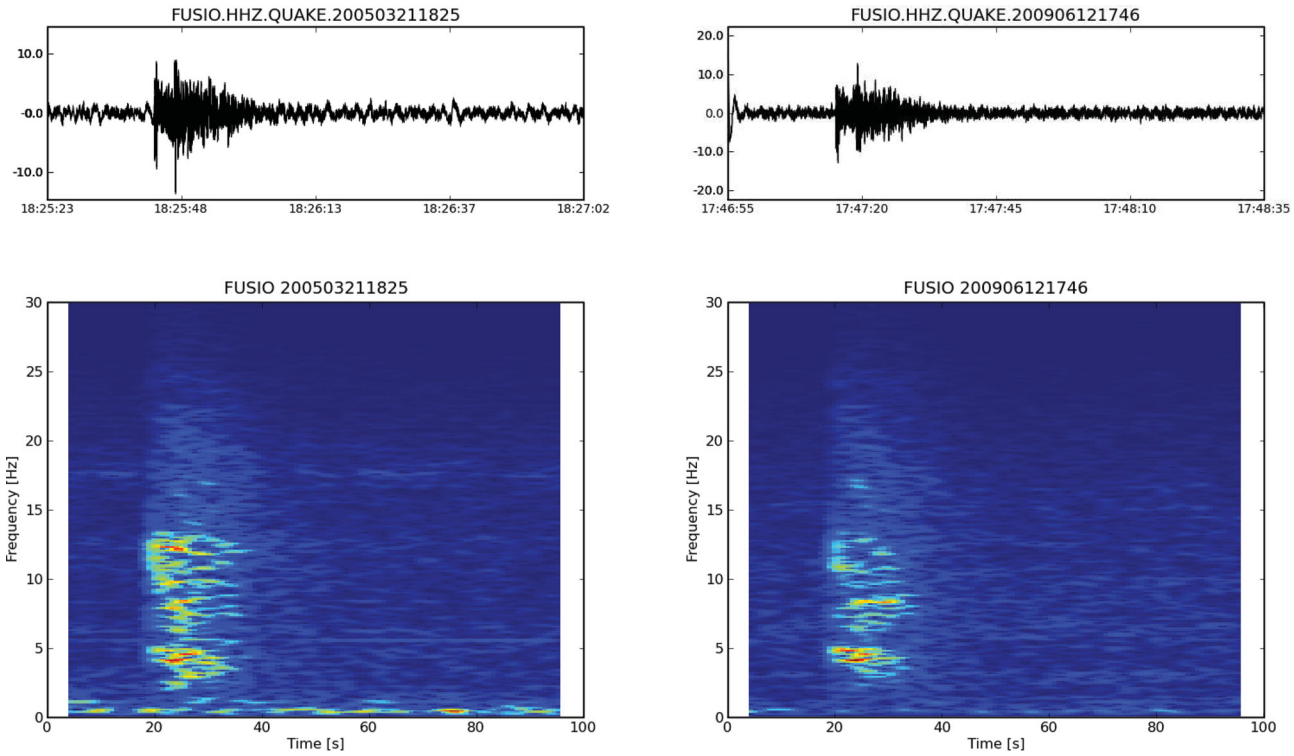

Figure 11. Waveforms and spectrograms of two events *a priori* labelled as earthquakes are shown. Both were labelled as quarry blast in the automatic processing. After carefully re-vising both events were re-classified as quarry blast according to the automatic label.

blasts one event shows clear P- and S-phases and is dominated by frequencies around 15 Hz (Fig. 12). As this event occurred in the same area as another earthquake sequence (sequence of Lodrino) we tend to think that the automatic classification is correct.

Additionally we investigated the missed events in more detail In year 2002 and 2003 we use only the draft models in the classification process as at the beginning of 2004 a sufficiently large number of training samples (i.e. detected events) is available for re-estimation. Therefore, the models do not perfectly match the corresponding classes resulting in three missed events (two earthquakes and one quarry blast) in 2002. However, when re-classifying the signals from 2002 to 2003 using the re-trained models, also previously missed events are recognized correctly. It should be noted that when doing this there is no longer independence between training and evaluation data set. All other missed events (i.e. later than 2003) are characterized by a very low signal to noise ratio (<2), which is determined by comparing maximum amplitudes of event and preceding noise (within 10 s before event onset). This indicates that re-training enhances the classification accuracy and the robustness of the detection process.

For comparison a data stream of 12 hr about every 3 month (QCDS) was classified using the current models, respectively. For rockfall signals the model from 2009 has been used for the whole time period as in the EBDS the first rockfall occurred in 2009 and the model has been derived from this event. Five out of 11 rockfalls are recognized correctly. Four out of the remaining six are buried in noise and are hardly visible neither in time nor in frequency domain at station FUSIO. In the daily processing at the SED they were detected at other stations that were situated closer to the epicenter and therefore the rockfalls are reported in the bulletin. The last two missed rockfalls took place in a source region further east. Their transients are visible in the records from station FUSIO. However,

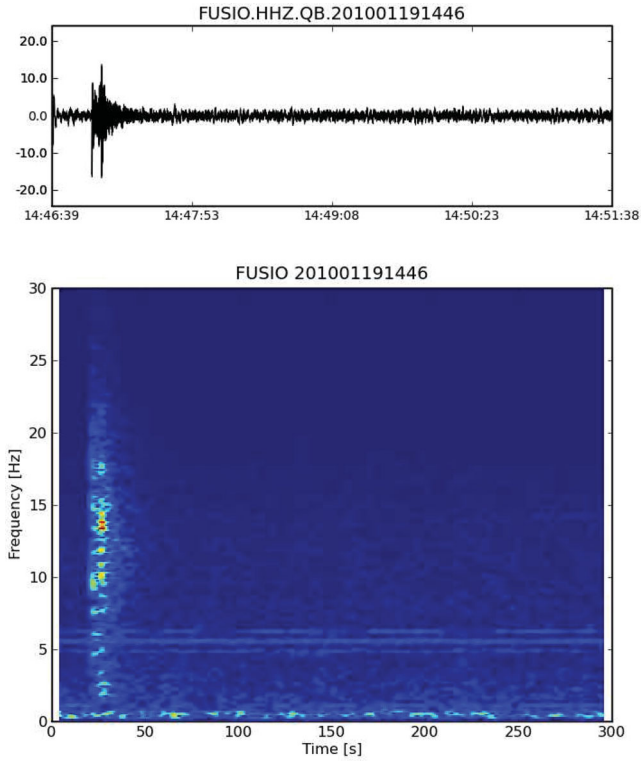

Figure 12. Waveform and spectrogram of an event *a priori* labelled as quarry blast is shown. The event was labelled as earthquake in the automatic processing. After carefully re-vising the event was re-classified as earthquake according to the automatic label.

both waveforms cannot be described appropriately by the corresponding models as signals are dominated by characteristics not captured by the rockfall HMM. This observation also applies to earthquakes and quarry blasts in the QCDS. Events close to the source area considered before are detected and recognized correctly (3 out of 18) while events further away cannot be described appropriately by the corresponding models. This dependence of recognition capability and source-receiver geometry is caused by two reasons. First, Fäh & Koch (2002) report on very inhomogeneous crustal structures in Switzerland and the corresponding variability of seismogram signatures. Thus, signal characteristics change in dependence of the source area. The second influence of the recognition performance arises from different travel paths. With larger epicentral distances the spectral content of individual events changes due to the increasing depletion of higher frequencies. Hence, it is not surprising that detection rates decrease with increasing source-receiver distance. A way to overcome both problems may be to incorporate events with different source-receiver geometries in the training data set. Alternatively, this dependence opens up the possibility to classify events according to their epicentral distance or source area with one model per defined distance range or source region (Beyreuther & Wassermann 2008).

In addition to the listed events the algorithm reports several spurious detections in both data sets (EBDS and QCDS). We carefully revised those segments and here one of the main improvements in using the suggested approach shows up. Confirmed by visual inspection of corresponding waveforms five additional earthquakes and three rockfalls have been found by the algorithm. Waveforms are shown in the Appendix. Due to a missing trigger detection they have not been recognized before. We suspect two reasons being responsible for this: First, STA/LTA trigger show a reduced performance if events fall within the 'shadow zone' that is a short time period after passing an energy transient, of the trigger (Withers *et al.* 1998). Second, small local events are missed as only too few stations are triggered. Given the approach taken in this paper multiple events with short inter event times are not merged and small local events can be detected by their appearance on only a single station (here FUSIO).

Swarm-like occurrences of events are a problem for most detection algorithms. In this study short time segments between consecutive signals are correctly recognized as noise due to the construction of the classifier. The most probable class is determined in a sliding window of 15 s. Considering the time shift of 5 s (10 s overlap) it is in principle possible to detect a new event every 5 s. An event is assumed to be embedded in noise. As the noise class is modelled as a signal without specific temporal structure (Hammer *et al.* 2012) a single time instance (one 0.05 s interval), preceding or succeeding an event, can be labelled as noise (Fig. 4). However, no overlapping events are contained in the investigated data set. Whenever the signals are not separated by short segments of noise their detection may still pose a problem. In this case, an improvement for the recognition accuracy may be obtained by allowing a different structure within the detection window. If events can be repeated arbitrarily within the window the events do not need to be separated by noise for a successful detection. In order to recognize two merged events of the same class the output of the classifier must be inspected in detail. In case the current observation is assigned to an earlier (already passed) state of the event model we expect a second event of the same class merged with the first one. Another alternative is reducing the length of the reference pattern and the length of the corresponding classifier window. By detecting only the onset of an event another event occurring shortly would not interfere the first detection. However, later phases of this first event may superimpose with the onset of the second event preventing its successful recognition. The recognition performance regarding closely spaced events is important for various applications (e.g. aftershock monitoring) and will be tested in future work.

After re-evaluating the results, three false alarms remain in the event-based data set. In the continuous 12-hr recordings six false alarms have been declared throughout the years, resulting in approximately one false alarm every third day.

5 CONCLUSION

Often in observatory practice the classification is still carried out manually by a human observer. This visual data screening process is a tedious and time-consuming task that one would like to automatize as much as possible. Although the suggested automatic algorithm for scanning the continuous data stream cannot completely replace a final manual revision the work load can be strongly reduced.

In the considered source area 97 per cent of the events are detected and (if assuming the manually re-vised classification as correct) 87 per cent are classified correctly. Eight additional events have been found that have been missed in the manual review process. As only short sections of the continuous data set have been processed in this study we expect more missed events in a fully continuous data set to exist. In HMMs the detection and classification is done in one step as each time segment is assigned to a specific signal class, which can be an defined event type, noise or an unknown signal class. Consequently, HMMs are applicable to single station data and can be used as a probabilistic earthquake detector (e.g. Beyreuther & Wassermann 2011; Beyreuther *et al.* 2012) or, as demonstrated here, to identify different types of signals. The recognition accuracy mainly depends on the appropriate representation of individual signal classes through the chosen features. For example if the used feature subset is less sensitive to path effects events may be classified correctly independent of the source-receiver geometry. Alternatively, the models may be extended by desired signal characteristics by including corresponding events (e.g. events with different source-receiver geometries, other ranges of magnitude) manually in the re-training data set.

Requiring a minimum amount of preparation time and workload the method has several advantages over classical techniques. Especially for rare events the algorithm provides a turning point as the procedure does not require a large number of training samples. Until now the use of seismic warning systems for mass movements was mainly limited by the incorrect identification of the signal source (NPRA 1994; Bessason *et al.* 2007; Arattano & Marchi 2008). By using the suggested approach this problem can be overcome leading to an automatic detection of rockfalls, avalanches or debris flows. In this study 8 out of 10 rockfalls clearly visible at station FUSIO were recognized correctly. Thus, we conclude that the suggested algorithm offers the opportunity to successfully use seismic sensors in alpine warning systems. Besides the early warning context the seismic monitoring of mass movements is often used to estimate their properties (Norris 1994; Surinach *et al.* 2000), understand the influence of external triggering factors on their dynamics (Helmstetter & Garambois 2010) or to identify possible precursors (Amitrano *et al.* 2005). However, in those studies the classification is done manually, which takes up valuable working time and might reflect the subjective view of the analyzer. With the possibility of setting up an automatic procedure the researcher would be relieved and consistent and time-invariant results would be provided for further processing.

The detection of new classes, carried out in using a threshold criteria to flag a poor match between the incoming signal and all defined signal classes, also benefits from the independence of previously collected training data. As soon as the new event type has been confirmed a corresponding classifier can be constructed allowing to scan the incoming data stream for corresponding events straightaway.

A drawback of the usual STA/LTA trigger are so-called 'shadow-zones' that inhibit the detection of consecutive events. This behaviour can be reduced when using the suggested approach as the noise is modelled as a signal with variable length. Furthermore, the system can be expanded to detect merged events not separated by noise.

The problem of changing class characteristics due to different source areas or epicentral distances is not critical to the overall system performance when applying the suggested technique. Instead it gives the opportunity to classify signals according to their source area or epicentral distance range as learning a new model requires minimum effort.

The system performance has been estimated by comparing the automatic classification with the manual *a priori* classification. However, the human observer might not be considered as completely error free (compare Section 4). Although an experienced seismologist can be regarded as one of the most powerful recognition systems, the visual inspection expresses a subjective view of the pattern. The problem is enhanced by the fact that there is not necessarily a clear separation between the waveforms of different signal classes. Even though the separation in different event types is justified by the true underlying source processes (earthquake, quarry blast, rockfall) the recorded waveforms become blurred by propagation effects. The spectral content of individual events may vary with propagation path, as paths with greater attenuation produce a depletion of higher frequencies. Thus, signal characteristics change in dependence of the source area. The situation is made more difficult as often information on the blasting times as well as the number and exact locations of blasts are missing (personnel communication N. Deichmann). Therefore, the obtained classification rates can only be seen as rough estimate of the true recognition accuracy.

ACKNOWLEDGMENTS

This work was partially funded by the German Ministry for Education and Research (BMBF), GEOTECHNOLOGIEN grant 03G0646F and by the project SwissExperiment, funded by the Competence Center for Environment and Sustainability of the ETH Domain (CCES). We thank N. Deichmann for providing the data for this study. His comments and help regarding the characteristics of individual events greatly improved the manuscripts quality.

REFERENCES

Allmann, B.R., Shearer, P.M. & Hauksson, E., 2008. Spectral discrimination between quarry blasts and earthquakes in southern California, *Bull. seism. Soc. Am.,* **98**(4), 2073–2079.

Amitrano, D., Grasso, J. & Senfaute, G., 2005. Seismic precursory patterns before a cliff collapse and critical point phenomena, *Geophys. Res. Lett.,* **32**(8).

Arattano, M., 1999. On the use of seismic detectors as monitoring and warning systems for debris flows, *Natural Hazards,* **20**(2-3), 197–213.

Arattano, M. & Marchi, L., 2008. Systems and sensors for debris-flow monitoring and warning, *Sensors,* **8**(4), 2436–2452.

Baer, M. & Kradolfer, U., 1987. An automatic phase picker for local and teleseismic events, *Bull. seism. Soc. Am.,* **77**(4), 1437–1445.

Baer, M. *et al.*, 2001. Earthquakes in Switzerland and surrounding regions during 2000, *Eclogae Geologicae Helvetiae,* **94**(2), 253–264.

Baer, M. *et al.*, 2005. Earthquakes in Switzerland and surrounding regions during 2004, *Eclogae Geologicae Helvetiae,* **98**(3), 407–418.

Baer, M. *et al.*, 2007. Earthquakes in Switzerland and surrounding regions during 2006, *Swiss J. Geosci.,* **100**(3), 517–528.

Baumgardt, D. & Young, G., 1990. Regional seismic wave-form discriminants and case-based event identification using regional aarrays, *Bull. seism. Soc. Am.,* **80**(6, Part b), 1874–1892.

Bessason, B., Eiriksson, G., Thorarinsson, O., Thorarinsson, A. & Einarsson, S., 2007. Automatic detection of avalanches and debris flows by seismic methods, *J. Glaciol.,* **53**(182), 461–472.

Beyreuther, M. & Wassermann, J., 2008. Continuous earthquake detection and classification using discrete Hidden Markov Models, *Geophys. J. Int.,* **175**(3), 1055–1066.

Beyreuther, M. & Wassermann, J., 2011. Hidden semi-Markov model based earthquake classification system using weighted finite-state transducers, *Nonlin. Processes Geophys.,* **18,** 81–89.

Beyreuther, M., Hammer, C., Wassermann, J., Ohrnberger, M. & Megies, T., 2012. Constructing a hidden Markov model based earthquake detector: application to induced seismicity, *Geophys. J. Int.,* **189**(1), 602–610.

Beyreuther, M., Carniel, R. & Wassermann, J., 2008. Continuous hidden Markov models: application to automatic earthquake detection and classification at Las Canadas caldera, Tenerife, *J. Volc. Geotherm. Res.,* **176**(4), 513–518.

Curilem, G., Vergara, J., Fuentealba, G., Acuna, G. & Chacon, M., 2009. Classification of seismic signals at Villarrica volcano (Chile) using neural networks and genetic algorithms, *J. Volc. Geotherm. Res.,* **180**(1), 1–8.

Deichmann, N. & Baer, M., 1990. *Earthquake focal depths below the Alps and northern Alpine foreland of Switzerland*. In The European Geotraverse: Integrative studies pp. 277–288, eds Freeman, P., Giese, P. & Mueller, S., European Science Foundation, Strasbourg.

Deichmann, N. *et al.*, 2008. Earthquakes in Switzerland and surrounding regions during 2007, *Swiss J. Geosci.,* **101**(3), 659–667.

Deichmann, N. *et al.*, 2009. Earthquakes in Switzerland and surrounding regions during 2008, *Swiss J. Geosci.,* **102**(3), 505–514.

Del Pezzo, E., Esposito, A., Giudicepietro, F., Marinaro, M., Martini, M. & Scarpetta, S., 2003. Discrimination of earthquakes and underwater explosions using neural networks, *Bull. seism. Soc. Am.,* **93**(1), 215–223.

Esposito, A.M., Giudicepietro, F., Scarpetta, S., D'Auria, L., Marinaro, M. & Martini, M., 2006. Automatic discrimination among landslide, explosion-quake, and microtremor seismic signals at Stromboli volcano using neural networks, *Bull. seism. Soc. Am.,* **96**(4), 1230–1240.

Fäh, D. & Koch, K., 2002. Discrimination between earthquakes and, chemical explosions by multivariate statistical analysis: a case study for Switzerland, *Bull. seism. Soc. Am.,* **92**(5), 1795–1805.

Frank, H. & Althoen, S.C., 1994. *Statistics: Concepts and Applications,* Cambridge University Press.

Gitterman, Y., Pinsky, V. & Shapira, A., 1998. Spectral classification methods in monitoring small local events by the Israel seismic network, *J. Seismol.,* **2**(3), 237–256.

Habermann, R., 1987. Man-made changes of seismicity rates, *Bull. seism. Soc. Am.,* **77**(1), 141–159.

Hammer, C., Beyreuther, M. & Ohrnberger, M., 2012. A seismic event spotting system for volcano fast response systems, *Bull. seism. Soc. Am.,* **102**(3), 948–960.

Helmstetter, A. & Garambois, S., 2010. Seismic monitoring of Sechilienne rockslide (French Alps): analysis of seismic signals and their correlation with rainfalls, *J. geophys. Res.,* **115,** F03016, doi:10.1029/2009JF001532.

Hillers, G. & Ben-Zion, Y., 2011. Seasonal variations of observed noise amplitudes at 2-18 Hz in southern California, *Geophys. J. Int.,* **184**(2), 860–868.

Horasan, G., Guney, A.B., Kusmezer, A., Bekler, F., Ogutcu, Z. & Musaoglu, N., 2009. Contamination of seismicity catalogs by quarry blasts: an example from Istanbul and its vicinity, northwestern Turkey, *J. Asian Earth Sci.,* **34**(1), 90–99.

Köhler, A., Ohrnberger, M. & Scherbaum, F., 2009. Unsupervised feature selection and general pattern discovery using self-organizing maps for gaining insights into the nature of seismic wavefields, *Comput. Geosci.,* **35**(9), 1757–1767.

Koller, D. & Friedman, N., 2009. *Probabilistic Graphical Models: Principles and Techniques,* MIT Press.

Kuyuk, H.S., Yildirim, E., Dogan, E. & Horasan, G., 2011. An unsupervised learning algorithm: application to the discrimination of seismic events and quarry blasts in the vicinity of Istanbul, *Natural Hazards Earth Syst. Sci.,* **11**(1), 93–100.

Langer, H., Falsaperla, S., Powell, T. & Thompson, G., 2006. Automatic classification and a-posteriori analysis of seismic event identification at Soufriere Hills volcano, Montserrat, *J. Volc. Geotherm. Res.,* **153**(1-2), 1–10.

Leprettre, B., Navarre, J. & Taillefer, A., 1996. First results from a pre-operational system for automatic detection and recognition of seismic signals associated with avalanches, *J. Glaciol.,* **42**(141), 352–363.

Marchi, L., Arattano, M. & Deganutti, A., 2002. Ten years of debris-flow monitoring in the Moscardo Torrent (Italian Alps), *Geomorphology,* **46**(1-2), 1–17.

Markou, M. & Singh, S., 2003. Novelty detection: a review - part 1: statistical approaches, *Signal Processing,* **83**(12), 2481–2497.

McNutt, S., 2002. Chapter 25: Volcano seismology and monitoring for eruptions, in *International Handbook on Earthquake and Engineering Seismology,* no. 81A in Int. Geophys. Ser., eds Kanamori, H., Jennings, P. & Lee, W., Academic Press, San Diego, CA.

McNutt, S.R., 1996. Seismic monitoring and eruption forecasting of volcanoes: a review of the state-of-the-art and case histories, *Monitoring and Mitigation of Volcano Hazards,* pp. 99–146, Springer-Verlag, Berlin, Federal Republic of Germany.

Norris, R., 1994. Seismicity of rockfalls and avalanches at 3 cascade range volcanoes - implications for seismic detection of hazardous mass movements, *Bull. seism. Soc. Am.,* **84**(6), 1925–1939.

NPRA, N.P.R.A., 1994. *Snow Engineering for Roads: About Snow Avalanches and Drifting Snow.,* Serial no 172, Public Roads Administration, Directorate of Public Roads. Oslo, Norway.

Ohrnberger, M., 2001. Continuous automatic classification of seismic signals of volcanic origin at Mt. Merapi, Java, Indonesia, *Ph.D. thesis,* Universität Potsdam.

Plafcan, D., Sandvol, E., Seber, D., Barazangi, M., Ibenbrahim, A. & Cherkaoui, T., 1997. Regional discrimination of chemical explosions and earthquakes: a case study in Morocco, *Bull. seism. Soc. Am.,* **87**(5), 1126–1139.

Rice, R., Decker, R., Jensen, N., Patterson, R., Singer, S., Sullivan, C. & Wells, L., 2002. Avalanche hazard reduction for transportation corridors using real-time detection and alarms, *Cold Regions Sci. Technol.,* **34**(1), 31–42.

Riggelsen, C. & Ohrnberger, M., 2012. A machine learning approach for improving the detection capabilities at ctbto/ims 3c seismic stations, *Pure appl. Geophys*., **2,** PAGEOPH, doi:10.1007/s00024-012-0592-3.

Ruud, B. & Husebye, E., 1992. A new 3-component detector and automatic single-station bulletin production, *Bull. seism. Soc. Am.,* **82**(1), 221–237.

Sheen, D.-H., Shin, J.S. & Kang, T.-S., 2009. Seismic noise level variation in South Korea, *Geosci. J.,* **13**(2), 183–190.

Strachan, I. & Clifton, D., 2009. A hidden markov model for condition monitoring of a manufacturing drilling process, *IET Condition Monitoring, Dublin, Ireland,* pp. 803–814.

Surinach, E., Sabot, F., Furdada, G. & Vilaplana, J., 2000. Study of seismic signals of artificially released snow avalanches for monitoring purposes, *Phys. Chem. Earth Part B-Hydrology Oceans Atmosphere,* **25**(9), 721–727, General Assembly of the European-Geophysical-Society, The Hague, Netherlands, Apr 22, 2000.

Theodoridis, S. & Koutroumbas, K., 2006. *Pattern Recognition,* 3rd edn, Academic Press, Inc., Orlando, FL, USA.

Ursino, A., Langer, H., Scarfi, L., Di Grazia, G. & Gresta, S., 2001. Discrimination of quarry blasts from tectonic microearthquakes in the Hyblean Plateau (Southeastern Sicily), *Annali di Geofisica,* **44**(4), 703–722.

Wiemer, S. & Baer, M., 2000. Mapping and removing quarry blast events from seismicity catalogs, *Bull. seism. Soc. Am.,* **90**(2), 525–530.

Wilcox, L. & Bush, M., 1992. Training and search algorithms for an interactive wordspotting system, *IEEE International Conference on Acoustics, Speech, and Signal Processing,* **2,** 97–100.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S. & Trujillo, J., 1998. A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bull. seism. Soc. Am.,* **88**(1), 95–106.

Wuster, J., 1993. Discrimination of chemical explosions and earthquakes in central-europe - a case-study, *Bull. seism. Soc. Am.,* **83**(4), 1184–1212.

Yeung, D. & Ding, Y., 2003. Host-based intrusion detection using dynamic and static behavioral models, *Pattern Recognition,* **36**(1), 229–243.

APPENDIX: ADDITIONAL EVENTS

Please note: OUP is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

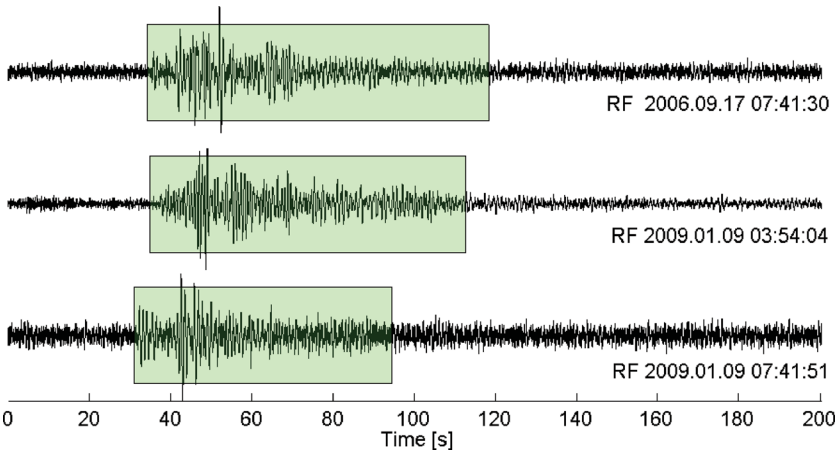

Figure A1. Signals recognized as rockfalls by the automatic system. Detected segments are marked by green windows. Events are not listed in the catalog. Detection times (i.e. start of green window) are given in GMT for each signal.

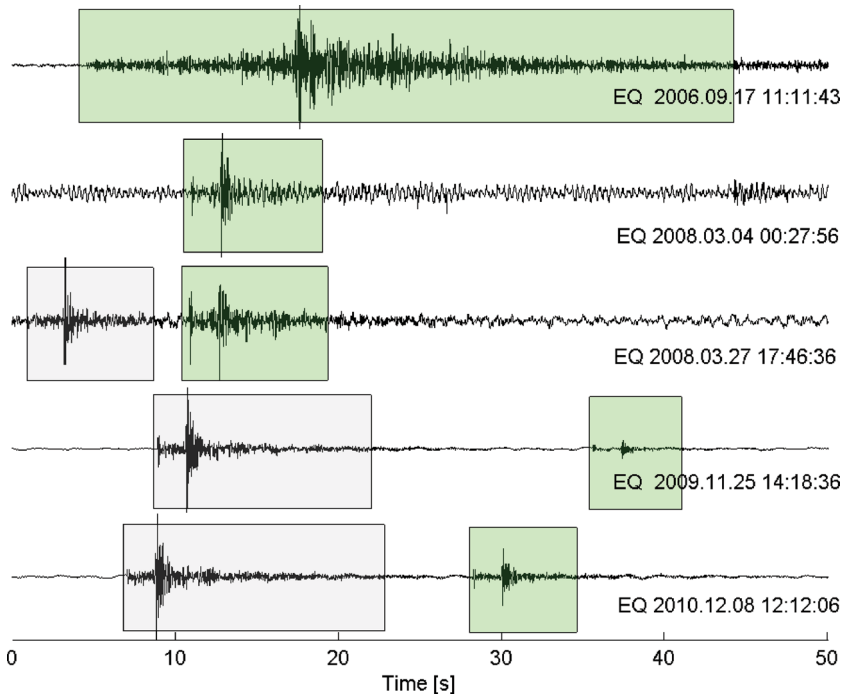

Figure A2. Signals recognized as earthquakes by the automatic system. Detected segments are marked by coloured windows. Events listed in the catalog: grey detection window. Events not listed in the catalog: green detection window. Detection times (i.e. start of green window) are given in GMT for each signal.