

# Automatic detection of avalanches using a combined array classification and localization

Matthias Heck<sup>1</sup>, Alec van Herwijnen<sup>1</sup>, Conny Hammer<sup>2</sup>, Manuel Hobiger<sup>2</sup>, Jürg Schweizer<sup>1</sup>, and Donat Fäh<sup>2</sup>

<sup>1</sup>WSL Institute for Snow and Avalanche Research SLF, Davos

<sup>2</sup>Swiss Seismological Service SED, ETH Zurich, Zurich

**Correspondence:** Matthias Heck (matthias.heck@slf.ch)

## Abstract.

We ~~use~~used a seismic monitoring system to automatically determine the avalanche activity at a remote field site near Davos, Switzerland. By using a recently developed approach based on hidden Markov models (HMMs), a machine learning algorithm, we were able to automatically identify avalanches in continuous seismic data by providing ~~as little as one a~~a single training event. Furthermore, we implemented an operational method to provide near real-time classification results. For the 2016-2017 winter period 117 events were ~~then~~then automatically identified. False classified events such as airplanes and local earthquakes were filtered using a new approach containing two additional classification steps. In a first step, we implemented a second HMM based classifier at a second array 14 km away to automatically identify airplanes and earthquakes. By cross-checking the results of both arrays we reduced the ~~amount~~number of false classifications by about 50%. In a second step, we used multiple signal classifications (MUSIC), an array processing technique, to determine the direction of the source. Although avalanche events have a moving source character, only small changes of the source direction are common ~~for snow avalanches~~ whereas false classifications ~~had large changes in the source direction and were therefore~~ showed large changes and thus were dismissed. From the 117 initially detected events during the ~~4-month~~4-month period we were able to identify 90 false classifications based on these two additional steps. The ~~obtained~~ avalanche activity based on the remaining 27 avalanche events was in line with visual observations performed in the ~~area~~region of Davos.

## 1 Introduction

During the winter seasons, snow avalanches are a common threat in mountain regions. Avalanche warning services therefore inform the public of the current avalanche danger. To assess the ~~danger~~level, warning services rely on information about the snowpack, amount of new snow, weather conditions and avalanche activity (McClung and Schaerer, 2006). Whereas the first three parameters can be measured or modeled, avalanche activity data are often hard to obtain, especially during snow storms or at night. Monitoring systems have therefore been developed to estimate the avalanche activity for a certain region.

Snow avalanches, like any other mass movement, generate seismic and infrasound waves (e.g. van Herwijnen and Schweizer, 2011b; Suriñach et al., 2005; Marchetti et al., 2015). Seismic signals of avalanches show some common characteristics, ~~such~~

as including a spindle shaped envelope of the time series (Nishimura and Izumi, 1997) and a typical frequency content between 2 and 30 Hz (Schaerer and Salway, 1980; Suriñach et al., 2001). Several classification approaches were therefore developed to automatically detect avalanches in seismic data. Lepretre et al. (1996) used a fuzzy logic approach to distinguish between different types of signals. Besson et al. (2007) used a nearest neighbor approach to classify new recorded events. ~~Although they relied on a 10-year database of avalanche events, the identification was rather poor.~~ Using this approach, they were able to detect 65% of all confirmed avalanches. Rubin et al. (2012) compared 12 machine learning algorithms, 10 of which were able to detect at least 90% of all ~~known manually identified~~ avalanches, however, at the cost of very high false alarm rates. Hammer et al. (2017) recently used hidden Markov models (HMMs), an advanced machine learning algorithm, to automatically detect large avalanches released during the winter of 1998-1999 in seismic data recorded by a single broadband station maintained by the Swiss Seismological Service (SED). Using this approach, they were able to identify 43 avalanches during a 5-day period within a radius of 30km of the station. Heck et al. (2018a) also used the HMM approach to automatically detect avalanches, however, in data recorded during the winter season 2009-2010 by a seismic array consisting of seven less sensitive vertical geophones. They obtained the best results for the automatic detection by combining the classification results of all sensors and requiring a minimal event duration for the detections.

Apart from using seismic signals for the automatic detection of avalanches, several studies focused on the use of infrasound signals. Localization parameters determined using cross-correlation techniques were used to automatically identify avalanches in continuous data sets (Scott et al., 2007; Marchetti et al., 2015; Thüring et al., 2015). By comparing the back-azimuth with the directions of known avalanche paths, possible avalanche events were identified (Marchetti et al., 2015). Thüring et al. (2015) used a similar approach for the automatic detection, but relied on support vector machines (SVM), a machine learning algorithm.

In addition to the automatic detection of avalanches, Lacroix et al. (2012) and Heck et al. (2018b) used seismic array processing techniques to locate the source of the avalanche. Lacroix et al. (2012) implemented a beam-forming approach and were able to assign recorded avalanches to three known avalanche paths. Heck et al. (2018b) compared a beam-forming method with a multiple signal classification (MUSIC) approach (Schmidt, 1986) and obtained better results with the latter and they subsequently applied this method to avalanches monitored during a two-day period in March 2017. ~~However, they also observed~~ Based on these results they concluded that their seismic array mostly recorded infrasound due to the limited distance between ~~their~~ the sensors. Nevertheless, they were able to reconstruct the avalanche path of several recorded events.

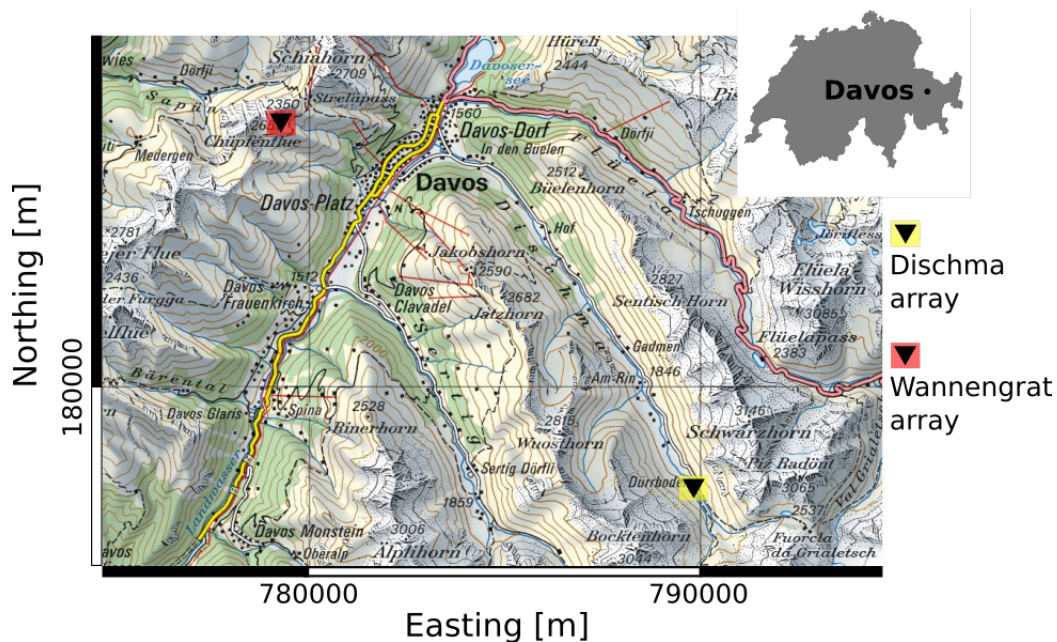
~~In several studies, avalanches were automatically identified in infrasonic data by using localization parameters determined using cross-correlation techniques (Scott et al., 2007; Marchetti et al., 2015; Thüring et al., 2015). By comparing the back-azimuth with the directions of known avalanche paths, possible avalanche events were identified (Marchetti et al., 2015). Thüring et al. (2015) used a similar approach for the automatic detection, but relied on support vector machines (SVM), a machine learning algorithm.~~ Lacroix et al. (2012) and Heck et al. (2018b) both used less sensitive vertical component geophones for the seismic monitoring resulting in an avalanche detection distance of approximately 3 km.

Our aim is to automatically identify avalanches in continuous data recorded during the winter period 2016-2017 using the same machine learning techniques based on hidden Markov models as used by Heck et al. (2018a). To reduce the false alarm rate we first use an additional classification performed at a second array 14km away to dismiss events recorded almost

simultaneously at both arrays such as earthquakes and airplanes. In a second step, we analyze the median back-azimuth path of the detections using the MUSIC method as performed by Heck et al. (2018b) and dismiss all events with a randomly distributed back-azimuth. We perform-performed the classification and localization of the events with the data recorded at a-the seismic array located in the Dischma Valley above Davos, Switzerland during the winter season 2017-and-combine-it-with-the  
 5 2016-2017 (yellow square in Figure 1). These results were then combined with data obtained at the Wannengrat array, which is located 14km to the northwest of the Dischma field site (red square in Figure 1).

## 2 Field site and instrumentation

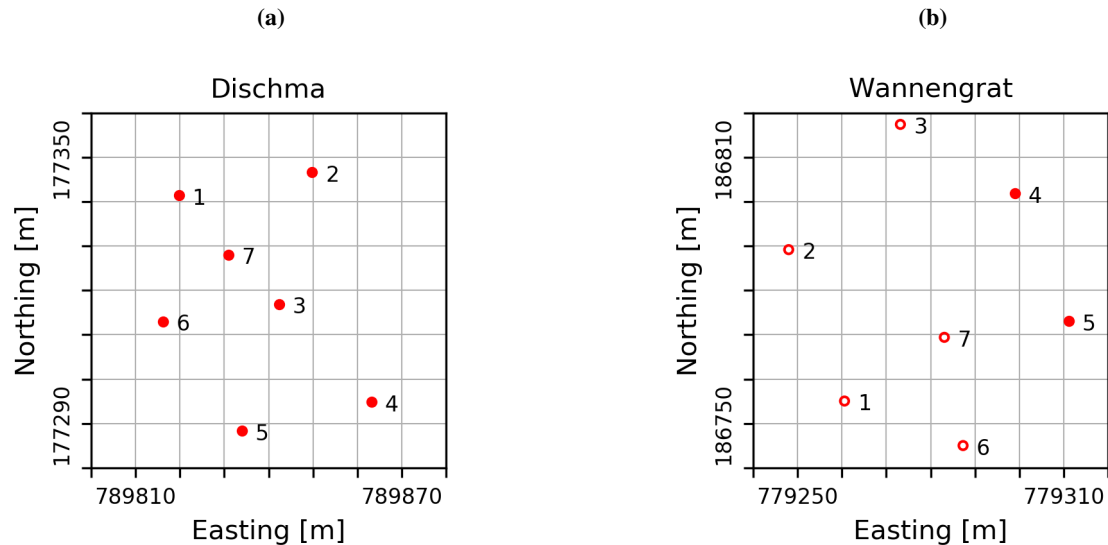
Prior to the 2016-2017 winter season, we installed two seismic arrays above Davos, Switzerland, similar to the systems described by van Herwijnen and Schweizer (2011a). The first array was deployed at the Dischma field site (yellow square in  
 10 Figure 1), 14km away from Davos at the end of a tributary valley (Heck et al., 2018b). The field site is a flat meadow at an elevation of 2000m a.s.l. surrounded by mountain peaks which rise up to 3000m. The second array was deployed at the Wannengrat field site above Davos at 2500m a.s.l. (red square in Figure 1). This field site is surrounded by several avalanche starting zones (van Herwijnen and Schweizer, 2011a).



**Figure 1.** Map of the area of Davos, Switzerland. The two arrays are indicated by a black triangle on colored ground. Red represents the Wannengrat array, yellow the Dischma array. Reproduced by permission of swisstopo (JA100118).

Both arrays consisted of a 300m long string with 7 vertical component geophones with an eigenfrequency of 4.5 Hz. The  
 15 sensors of the Dischma array were buried 50 cm deep into the ground whereas the sensors at the Wannengrat field site were

attached to rocks using an anchor. For each array the sensors were circularly arranged (Figure 2 a) and b)). The maximum distance between two sensors at the Dischma and Wannengrat field site was 64m and 74m, respectively, and the average distance was 36m at the Dischma array and 45m at the Wannengrat array.



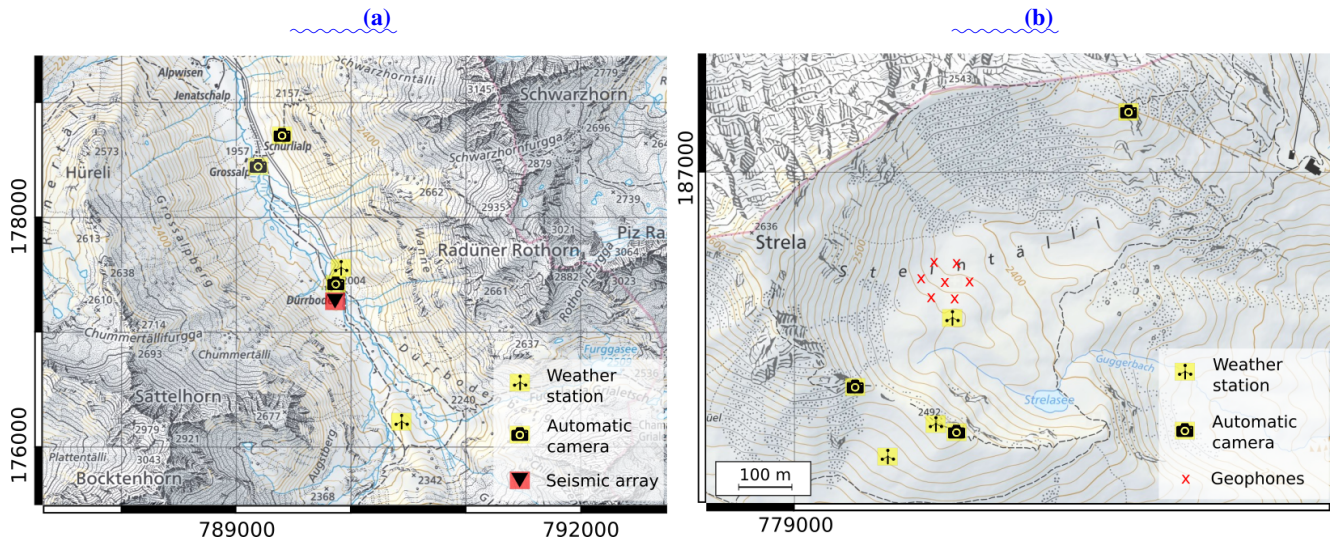
**Figure 2.** Setup of sensor arrays a) Dischma, b) Wannengrat. The open red circles indicate positions of not working sensors during the winter 2017.

The instrumentation and data logging systems were identical for both arrays. Data were continuously sampled at a rate of 500 Hz. However, due to technical problems only two sensors of the Wannengrat array recorded data throughout the entire winter (4 and 5 in Figure 2 b)). Both field sites were equipped with several automatic weather stations (3 at Dischma, 4 at Wannengrat) as well as automatic cameras (8 at Dischma, 5 at Wannengrat). The automatic cameras visually monitored the surrounding slopes and images were recorded every 10 minutes throughout the winter. ~~In some cases, these images helped (Figure 3). As already shown for avalanche activity periods in the winter season 2009-2010 by van Herwijnen and Schweizer (2011b), those images can help to identify and confirm seismic events produced by avalanches in the data (van Herwijnen and Schweizer, 2011b).~~ In addition to the automatic cameras, we also performed field surveys shortly after a period of high avalanche activity to identify avalanches and map their outlines (Figure 4) (Heck et al., 2018b).

### 3 Methods

#### 3.1 Data pre-processing

The continuous seismic data mostly consist of noise. Since ~~this for the current application~~ noise is of little interest, we applied a simple threshold based event detector to reduce the amount of data (Heck et al., 2018a). For a window  $i$  with a length of 1024 samples a mean absolute amplitude  $A_i$  was determined. When  $A_i \geq 5\bar{A}$ , with  $\bar{A}$  the daily mean amplitude, the data within the



**Figure 3.** Detailed map of the field sites a) Dischma, b) Wannengrat. On the map the positions of the seismic sensors, automatic cameras and weather stations are indicated. Due to scaling not every weather station is shown on the map.

window were processed. If the amplitude threshold for the following window was also reached, data were concatenated. Furthermore, a section of  $t = 60$ s was cut before and after the window to ensure that the onset and coda of each event was incorporated. Doing so, data were reduced by 80% to several data windows of various lengths. In addition, we filtered the data using a 4th order Butterworth bandpass filter with corner frequencies of 1 and 50 Hz.

### 5 3.2 Classification of events

To automatically identify avalanches in the continuous seismic data we used hidden Markov models (HMMs). ~~This method was already successfully used on continuous seismic data collected at the Wannengrat field site during the 2009-2010 winter season (Heek et al., 2018a). HMMs are statistical classifiers, which (Rabiner, 1989). These statistical classifiers use a sequence of multivariate Gaussian probability distributions to model observations (i.e. the e.g. seismic time series or its features) by a sequence of multivariate Gaussian probability distributions. However, the classical approach relying on well-known). To determine the characteristics of the distributions (i.e. mean and covariance) a large number of training sets of known events, so called pre-labeled training sets as performed by Ohrnberger (2001) or Beyreuther et al. (2012) is not suited for the classification of rare seismic events like avalanches. We therefore, are required. For each different type of observation (e.g. avalanche, airplane or earthquake in the seismic data) a separate HMM is trained. By combining all HMMs the whole classification system with several classes is constructed. This classical approach, which relies on a large number of well-known pre-labeled training sets, was successfully used to automatic identify seismic events in continuous seismic data (Ohrnberger, 2001; Beyreuther et al., 2012). Avalanches, however, are rare events and it is nearly impossible and too time consuming to obtain a large training set. To circumvent this, we performed the classification based on a new an approach developed by Hammer et al. (2012) exploiting~~



**Figure 4.** Picture of the field site of Dischma facing to the south. It was taken on 15 March 2017 shortly after a period of high avalanche activity and several recent avalanches were observed.

the abundance of data containing mainly background signals to obtain general wave-field properties. ~~Using these properties,~~ From these properties a widespread background model can be learned ~~from the general properties.~~ A new event model ~~is (e.g. representing avalanches) is then~~ obtained by using the widespread background model to adjust the event model description by using only one training event. ~~The so obtained classifier therefore consists of the~~ In contrast to the classical HMM approach, the   
5 classification system of this approach consists of a background model and one event model for each ~~interesting implemented~~ event class. The classification process itself calculates the likelihood that an unknown data stream was generated by a specific event class for each individual HMM class (Hammer et al., 2012, 2013).

~~The classification performed by Heck et al. (2018a)~~ This new approach was successfully used on continuous seismic data collected at the Wannengrat field site during the 2009-2010 winter season by Heck et al. (2018a). Their classification, however,   
10 consisted of creating a new background model for each day and the resulting ~~HMMs were~~ classification system was used to classify the data of the same day. A near real-time classification, as would be required for operational purposes, is then not possible. To overcome this problem, here we implemented a classification process by learning the background model using

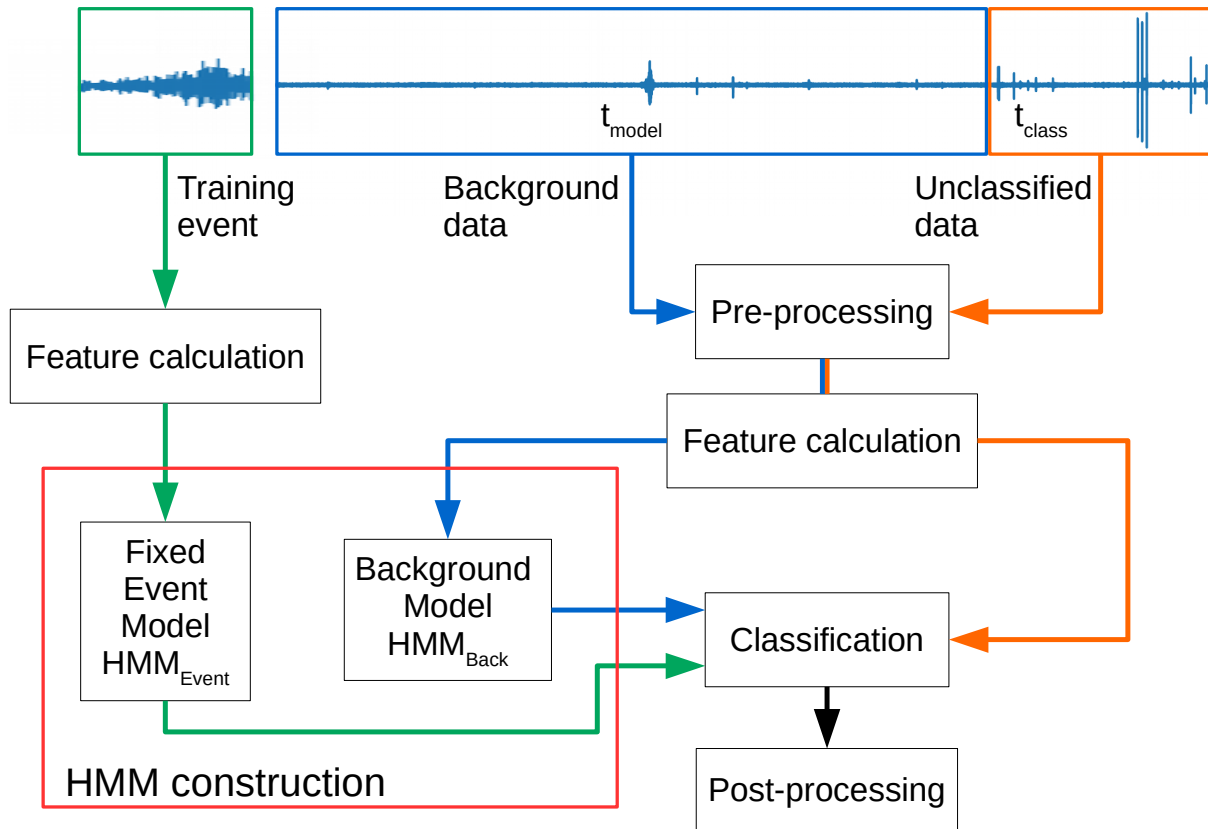
data from a different time window than the data we ~~want~~wanted to classify. To train the background model we used the pre-processed data taken from the window  $t_{\text{model}}$ , whereas the pre-processed data we want to classify are in the time window  $t_{\text{class}}$  (Figure 5). This so-called operational classification was performed by using a window length of  $t_{\text{model}} = 24$  h and  $t_{\text{class}} = 1$  h with the start time of  $t_{\text{class}}$  corresponding the end of  $t_{\text{model}}$ . ~~Once the classification for~~This means, that our background model  
5 is always determined by the pre-processed data of a 24-hour window. By choosing a length of 1 h for the window  $t_{\text{class}}$ , we  
were able to classify the pre-processed continuous seismic data of one hour during one step of the operational classification.  
Once one classification step, which is the classification of the window  $t_{\text{class}}$  is finished, both windows are shifted by one hour  
and the classification was executed for the shifted windows. The so performed classification takes  $\sim 6$  min for the classification  
of one day without the feature calculation. In contrast, the classification performed by Heck et al. (2018a) only took  $\sim 30$  s for  
10 one day. All calculations were performed on a computer with a regularly available 8-core processor and 12 GB ram running a  
standard Ubuntu Linux Distribution.

As input for the HMMs a compressed form of the data was used, so-called features. Features represent different aspects of the time series such as spectral, temporal or polarization characteristics. Since we used single component geophones, we only used the following spectral and temporal features similar to those used by Heck et al. (2018a).

- 15 – Central frequency (Barnes, 1993)
- Dominant frequency (Kramer, 1996)
- Instantaneous bandwidth (Barnes, 1993)
- Instantaneous frequency (Taner et al., 1979)
- Cepstral coefficients (Kanasewich, 1981)
- 20 – Half-octave bands (Joswig, 1994)

For the feature calculation, we used a sliding window of width  $w = 512$  samples and a step size of 0.05 s or 25 samples resulting in an overlap of 97%. ~~The~~We used in total 6 half-octave bands for the classification and the first half-octave band ~~has had~~ a central frequency of 3.9 Hz~~and a total number of 6 bands~~. Calculating the features from the pre-processed data takes  $\sim 15$  min for a complete day for all sensors. Since we shift the windows for 1 h after each step of the operational classification, the  
25 calculation of the features for a complete day needs to be performed only for the very first step. For the following steps it is  
sufficient to calculate only the features for the window  $t_{\text{class}}$  with a length of 1 h, which approximately takes less than  $\sim 2$  min  
for the calculation.

The classification process consists of five steps: pre-processing, feature calculation, HMM construction, classification and post-processing (Figure 5). First the data used to build the background model are selected from the time window  $t_{\text{model}}$  and  
30 the data to be classified are determined by the window  $t_{\text{class}}$ . The data from the selected time windows are pre-processed to reduce the amount of noise and then the features are calculated. In the HMM construction part, the features calculated from the data within  $t_{\text{model}}$  and from the data of the training event are used to construct a background model  $\text{HMM}_{\text{Back}}$  and an event



**Figure 5.** Flow chart of the classification process. Green lines show the construction of the event model. Blue lines show the construction of the background model. The orange lines show, how the data to be classified are processed.

model  $HMM_{Event}$ . The event model  $HMM_{Event}$  was learned using only one training event that is representative of avalanches at a specific field site (Heck et al., 2018a). It was determined once and is then applied for the entire winter season. In contrast, the background model  $HMM_{Back}$  was constructed-reconstructed every hour. Using both models, the pre-processed data within the window  $t_{class}$  was-were classified in the classification process. The features calculated for this window are therefore passed to the classifiers. Once the classification is performed, post-processing steps are applied (~~Heck et al., 2018a~~)as proposed by Heck et al. (2018a). The first step consisted of applying a duration threshold to the classified events. Each classified event having a duration shorter than 12s in duration was dismissed. The second step, the so-called voting-based classification, combines the results of all sensors. Only events that were classified by at least 5 sensors are considered as possible avalanches.

### 3.3 Combined array detection

10 Initial classification results performed on the data set of the winter season 2016-2017 revealed that although the total number of detected events was low, many detected events were very likely generated by airplanes or regional earthquakes (local magnitude



between 1.5 and 4 for earthquakes at local and regional distance triggered by at least 6 stations according to the earthquake catalog of the Swiss Seismological Service). In contrast to avalanches, which are recorded only at one array, these events are recorded at both arrays. We therefore used a combined array detection to remove earthquakes and airplanes from the detections. To perform the combined array detection, a second HMM was implemented to identify earthquakes and airplanes in the data recorded at the Wannengrat array ~~12km-away-at~~ 14 km distance from the Dischma array. Classification results were then combined to remove all events recorded simultaneously at both arrays.

### 3.4 Localization ~~as avalanche indicator~~ results to confirm avalanches

Heck et al. (2018b) determined the direction of several avalanches using a multiple signal classification algorithm called MUSIC and were able to locate the avalanche path of several avalanche events based on the data of a single array. The MUSIC algorithm determines the back-azimuth angle and the apparent velocity of the incoming wave-field for a small time window. This time window is shifted to provide a time series of back-azimuth and apparent velocity values. The MUSIC method is based on the covariance matrix taking the data of all sensors into account at once, whereas beam-forming methods rely on pair-wise cross-correlation (Schmidt, 1986; Rost and Thomas, 2002). MUSIC can resolve multiple sources more easily than beam-forming methods. Furthermore, the MUSIC method can be applied to small frequency bands and the different frequency contents of the wave-field can be analyzed. For further information on multiple signal classification the reader is referred to Schmidt (1986) and Hobiger et al. (2016).

Heck et al. (2018b) found, that due to the small distance between the sensors, the seismic array mostly resolved sonic wave-fields rather than seismic wave-fields to estimate the back-azimuth. Using a median smoothing filter they then calculated a so-called median back-azimuth path with time. In this study, we used these median back-azimuth paths obtained by the event localization to decide whether a classification was ~~a correct or a false alarm~~ associated with an avalanche. Specifically, we used a threshold value for the derivative of the median back-azimuth path. The assumption is that avalanche events have a smooth median back-azimuth path with little variations, whereas false detections have randomly distributed back-azimuths with large variations in time. By analyzing several avalanche events, especially the events ~~observed-identified~~ by Heck et al. (2018b), we observed small changes below  $10^\circ$  for the back-azimuth path. ~~Even for events passing close to the array we observed small changes below  $10^\circ$  in the median back-azimuth path.~~ Hence, we used a threshold value of  $10^\circ$  between two adjacent points of the median back-azimuth path for the event detection. ~~Similar to the minimum event duration used in the post-processing steps, we require a minimum duration of 20s for the~~ Even for events passing close to the array, we observed changes below  $10^\circ$  in the median back-azimuth path. ~~The minimum duration for this post-processing step is higher than for the minimal duration due to the windowing of~~ Heck et al. (2018a) suggested that a detected event should have a minimum duration of 12s to be considered as an avalanche. For the localization step, however, it was necessary to increase this duration because the window length used for the median smoothing filter was already 8s ~~of the data to obtain the median-long~~ (Heck et al., 2018b). To cover enough data points to use the minimal event duration as a reliable classification criterion, we therefore required a minimum length of 20s for the back-azimuth path (Heck et al., 2018b). They. Heck et al. (2018b) also showed that only the frequency content of the signal between 4.5 and 12.5 Hz contained information valuable for the localization performed at the used array. By further

analysis of the already localized events by Heck et al. (2018b), we observed, that a reduced frequency range provided similar results. Hence we reduced the number of analyzed frequency bands to ~~4~~four bands between 6 and 7.5 Hz and were able to speed up the calculation time. Nevertheless, the processing time for the localization is about three times real time on the same computer used for the classification as mentioned earlier.

## 5 4 Classification results

We performed the avalanche detection for the data recorded at the Dischma array during the winter season 2016-2017 and compared the results with the avalanche activity ~~which was~~ visually obtained by local observers and compiled by the avalanche warning service at the SLF. It has to be noted that this compilation is incomplete and covers an area ~~that can not be completely~~ much larger than that monitored with the Dischma array. Therefore, comparison with this compilation remains indicative.

### 10 4.1 Overview of the winter season

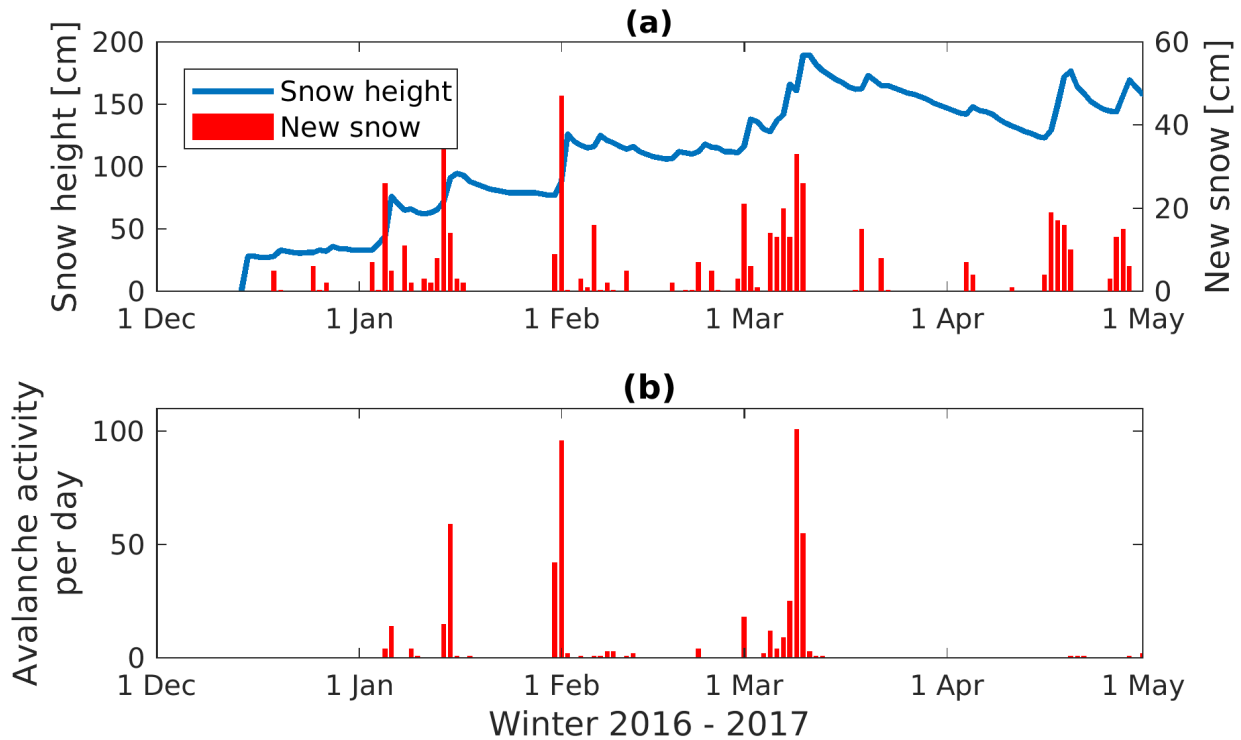
The winter period of 2016-2017 was relatively short and characterized by a below-average snow depth. First snowfalls were quite late in the season, in the middle of December, followed by four weeks without substantial precipitation and low temperatures. Due to the constant high temperature gradient within the snowpack, a poorly bonded layer of depth hoar was formed at the base of the snow cover.

15 During the winter season, three significant snowfall periods ~~were observed~~occurred; one in each month from January to March (increase of blue line in Figure 6). Each of these snowfalls was associated with considerable avalanche activity in the region of Davos (red bars in Figure 6).

In addition to these avalanche observations, we analyzed the pictures taken by the automatic cameras installed at our field sites. Surprisingly, ~~the~~ avalanche activity was low at the Dischma field site in January and February. During the early March snow storm the visibility was poor and only very few avalanches were identified on the images of the automatic camera. ~~Once~~ However, once the storm was over, the intensity of the avalanche cycle became clear as many avalanche deposits were visible on the images. Five days after the storm we mapped 24 avalanches within a 4 km radius of the Dischma field site (Heck et al., 2018b).

### 4.2 Classification performed at single array

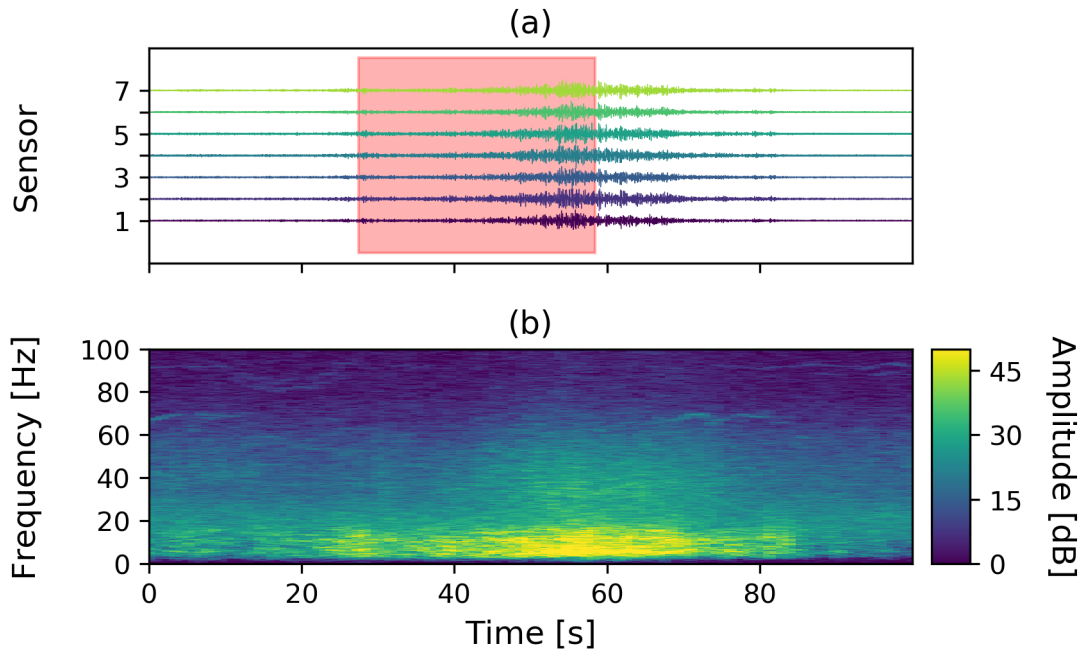
25 The main classification was performed at the Dischma field site for all seven sensors. Based on the visual inspection of the seismic data performed by Heck et al. (2018b), several avalanche events suitable as training events for the HMM were identified. However, they had mainly analyzed the period of high avalanche activity on 9 and 10 March 2017. Visually inspecting the entire winter season we identified 44 avalanche events. However, as already shown by Heck et al. (2018a), visually inspecting seismic data ~~also implicates~~ contains many uncertainties. An avalanche released on 9 March 2017 at 06:47 was already analyzed in detail by Heck et al. (2018b) and can unambiguously be classified as an avalanche. We therefore decided to use this event as our training event (Figure 7).



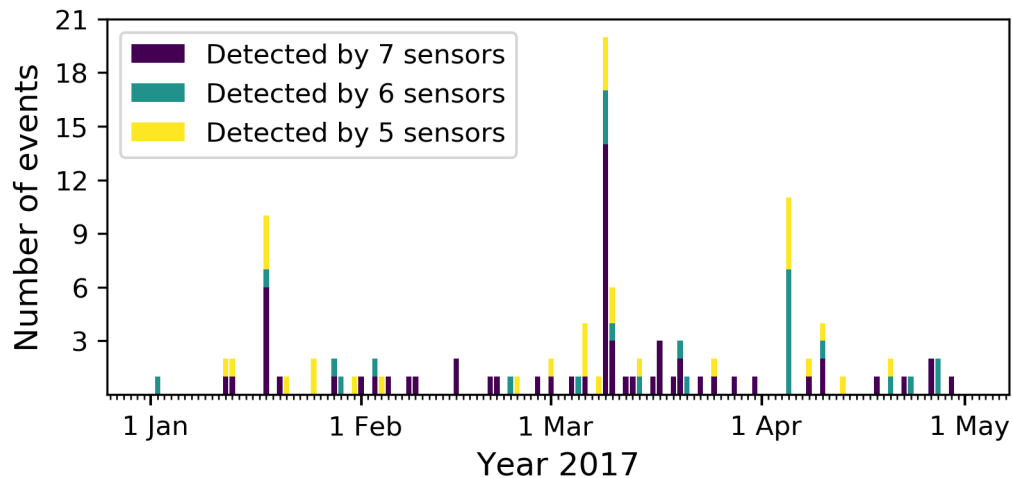
**Figure 6.** a) Snow height measured at the automatic weather station at Weissfluhjoch 12km to the northwest of the Dischma array at  $\sim 2600\text{m}$  a.s.l. for the winter season 2016-2017. Red bars are the height of new snow measured each day at 8:00 am. b) Number of avalanches observed per day in the region of Davos ( $\sim 175\text{ km}^2$ ).

Using the classifier trained with this event, we performed the classification for each single sensor of the array. In a next step, the results of the classification were post-processed; first all results of each sensor with a duration  $\leq 12\text{s}$  were dismissed (Heck et al., 2018a). Finally we dismissed all classifications ~~which that~~ were classified by less than 5 sensors.

The classification and the following post-processing was applied to the continuous data set recorded from 1 January to 30 April 2017. For this period, a total of 117 events were classified as avalanches. ~~Half A quarter~~ of the events were detected by ~~7 sensors whereas the classes with 5 and sensors, a quarter by 6 votes are each represented with 25% of the events. and about half by 7.~~ Most events were classified during the early March snow storm on 9 and 10 March 2017 (Figure 8). In addition a peak in the middle of January and beginning of April ~~as well as and~~ a cluster of events around the beginning of February ~~are was~~ visible. The peak in January as well as the cluster around the beginning of February correspond with the avalanche activity period visually recorded in the region of Davos (Figure 6). For the peak in April, however, no avalanches were observed in the surroundings of Davos. Furthermore, several single detections ~~are randomly were~~ distributed over the season showing no accordance with the visual avalanche observations. Therefore we visually inspected the time series and the corresponding



**Figure 7.** Avalanche released on 9 March 2017 at 06:47 used as training event for the classifier. a) time series for the 7 sensors and b) the corresponding spectrogram. The red area indicates the part of the time series used as training event. b) corresponding spectrogram of the seismic time series.



**Figure 8.** Classification results after post-processing (including voting-based classification) at the Dischma array. The colored bars indicate the number of classified events per day depending on the number of sensors: Violet bar indicates detections by 7 sensor, turquoise bars by 6 sensors and yellow bars by 5 sensors.

spectrograms of each of the 117 classifications and found that the HMM also classified various airplanes (Figure 9 a)) and regional earthquakes (Figure 9 b)) as avalanches.

Although these events can be distinguished from avalanches through visual inspection (e.g. the sharp onset visible for earthquakes), the classifier identified these events as belonging to the avalanche class, even when we used different training events or varied the setup for the classification (results not shown). This was most likely because earthquake and airplane signals were more similar to avalanches than to the background model, and consequently tagged as avalanches.

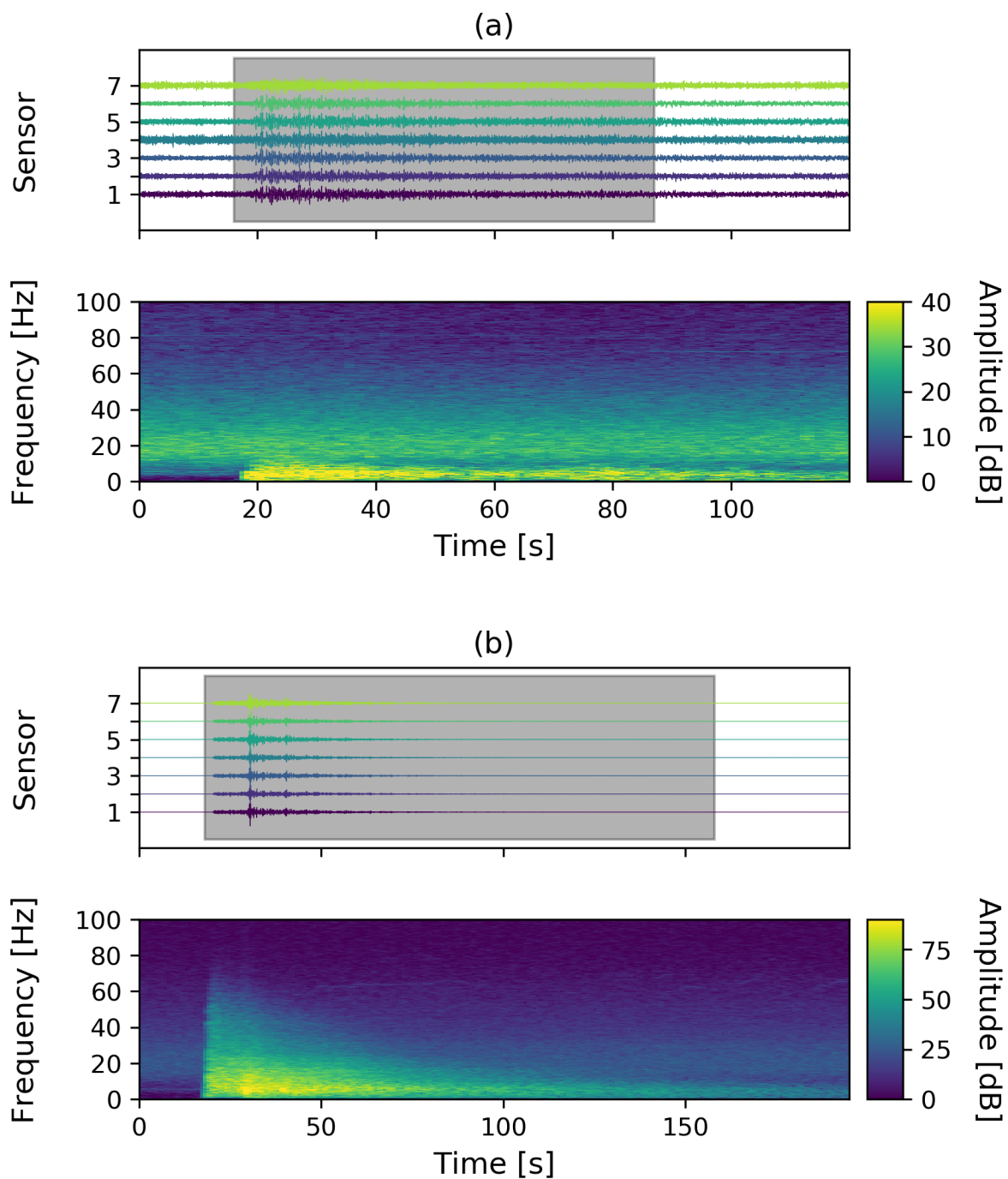
Analyzing the features also showed the similarities of the different types of events, especially the time dependent behavior. In the beginning of the event, the feature behavior of the events is different, however, at the end of the event a similar time dependent behavior is visible (Figure 10). Due to these similarities, airplane and earthquake events are more similar to avalanches than to noise resulting in false classifications of these events.

### 4.3 Classification performed at both arrays

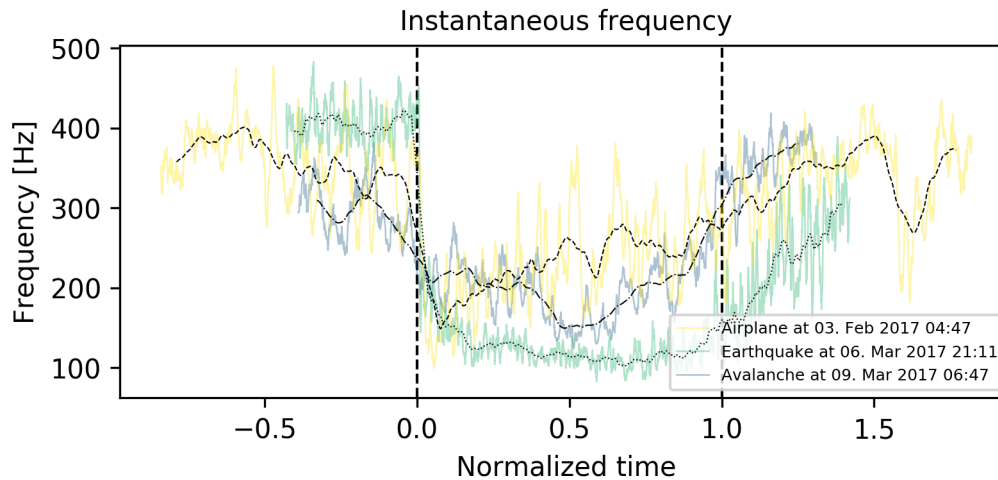
The majority of the misclassifications were produced by two types of events, i.e. airplanes and earthquakes. A comparison of several detected earthquake events with the earthquake catalog of the Swiss Seismological Service (SED) showed, that all compared earthquake events occurred within a range of 120 km. As the Wannengrat array was deployed only 14 km away, all observed earthquakes were likely to be detected simultaneously at both arrays. Moreover, Davos lies within an approaching corridor of the international airport Zürich and several commercial airplanes are passing by per pass by every hour at an altitude of at least 5 km. Similar to avalanches, airplanes also have a moving source character and due to the fast movement they are also observed almost simultaneously at both arrays. Indeed, a comparison of both time series revealed that earthquakes were recorded at both arrays at the same time whereas airplanes were recorded with a small delay of 20 to 30 s due to the movement of the source. The time series and spectrograms at Dischma and Wannengrat were very similar (Figure 11).

Avalanche signals, however, were only detected within a radius of 3 to 4 km of the array and were therefore only recorded at one array (Heck et al., 2018b). In order to eliminate classified events recorded at both arrays, we performed a second classification at the Wannengrat array. Due to similarities of the transient signals as mentioned earlier, a HMM trained with an airplane signal was capable to also detect earthquakes. A closer look at the classification results for the Wannengrat array revealed, that it was sufficient to only use the HMM trained with an airplane signal (not shown here). The number of detected events at this array varies strongly per day (blue bars in Figure 12).

The start times obtained by the secondary classification performed with the Wannengrat data were then compared with the classification results for the Dischma array. Overall, 53 of the 117 classifications were detected almost simultaneously at both arrays and we considered these events as airplanes or earthquakes (yellow bars in Figure 12). After the comparison of the classification results obtained by both arrays 64 potential avalanche events remained (turquoise bars in Figure 12). The distribution of these events is similar to visually observed avalanches (Figure 6), except for detections at the beginning of April and no detections at the beginning of February. These events were only detected using the automatic classification approach. Furthermore, due to the previously mentioned acquisition problem of the Wannengrat array, all events between 12 and 20



**Figure 9.** Time series and corresponding spectrograms for two false classifications, a) airplane, b) earthquake. The gray rectangular area indicates the part classified as event by the HMMs.

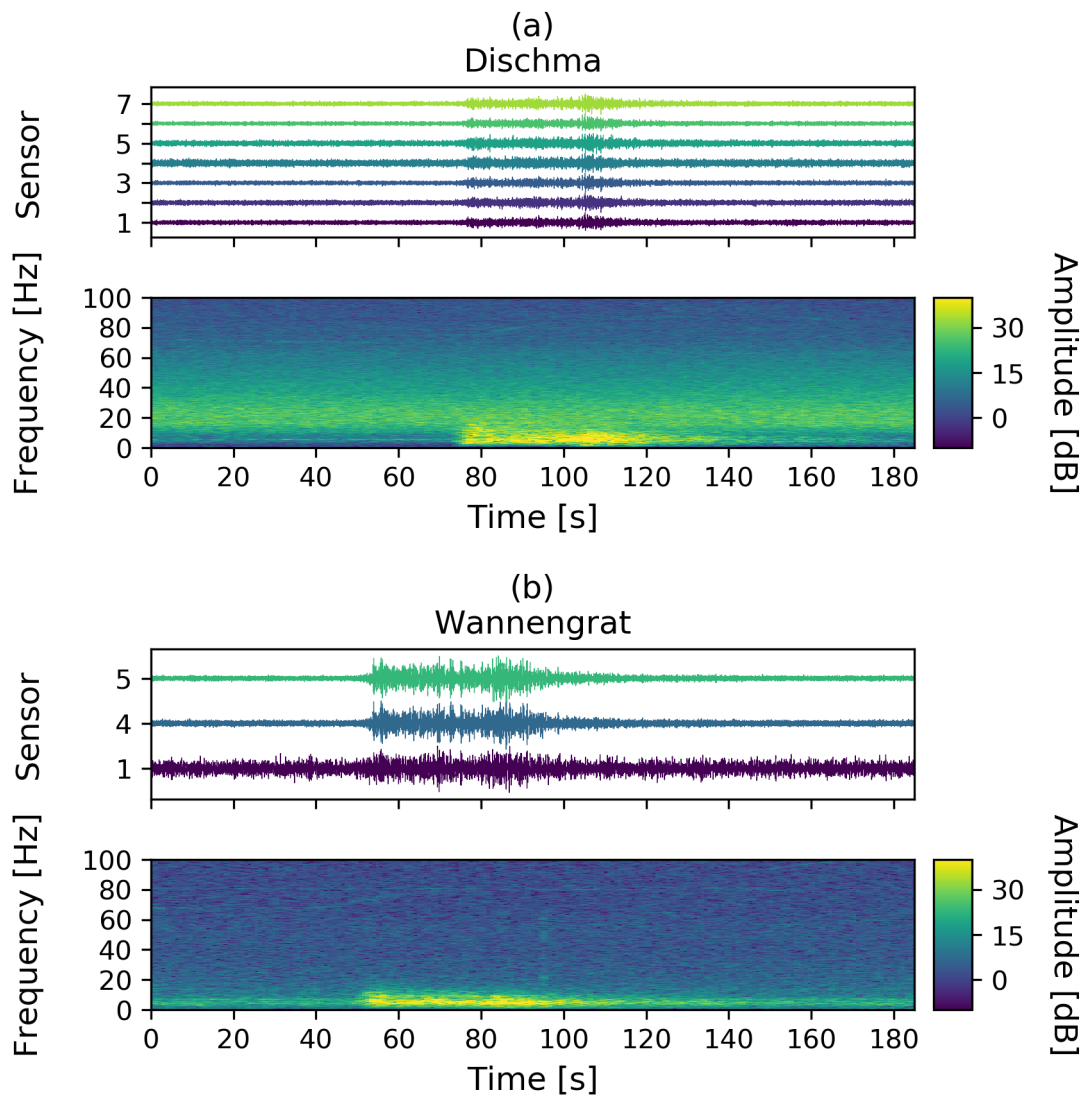


**Figure 10.** Feature instantaneous frequency for three different event types, yellow represents the signal produced by an airplane, green of an earthquake and blue of an avalanche. The black lines are the mean of the features. The dashed line at 0 indicates the start of the events and the dashed line at 1 the end of the event.

January 2017 were considered as avalanches as we had no further information from the second array. Hence, we expected some misclassifications among the remaining 64 avalanche events.

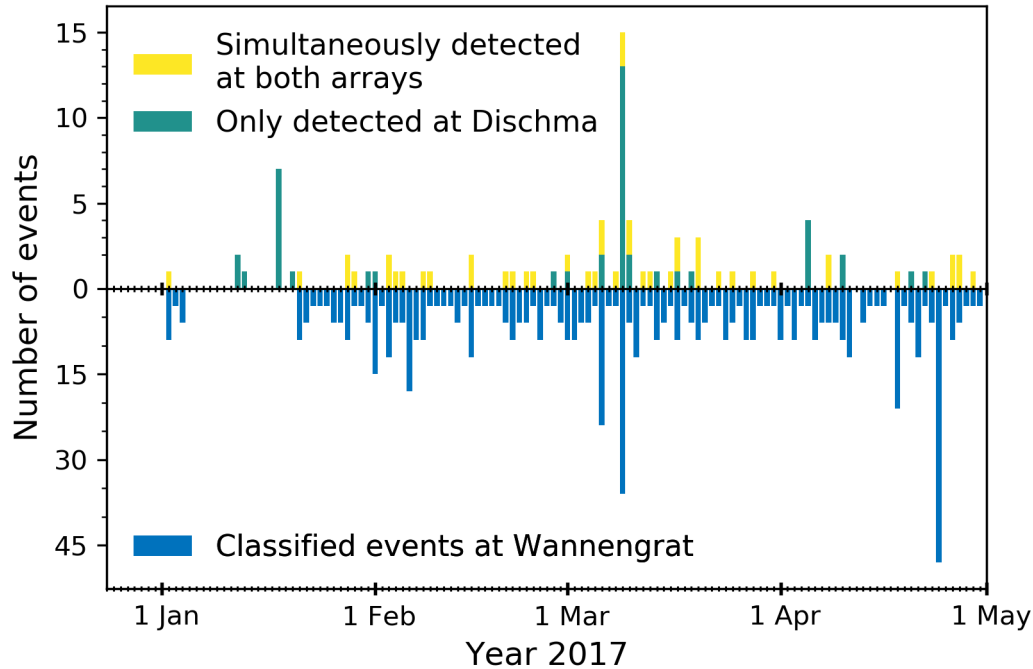
#### 4.4 Localization post-processing

In a last processing step, we applied the MUSIC method to the remaining 64 classified events to estimate the back-azimuth and to find a possible median back-azimuth path. The event used to train the HMMs had a duration of around 50s showing a median back-azimuth path with slight changes in the angle (straight black line in Figure 13 a)). Before and after the event, however, the back-azimuths were randomly distributed as would be expected for noise. For the training event, the derivative of the back-azimuth path has low values for the 50s part with a median back-azimuth path with small changes (Figure 13 b)). For this 50s long interval, changes below  $10^\circ$  were observed for the median back-azimuth path. Before and after the event, however, the changes are very high due to the randomly distributed back-azimuths. Further analyzed avalanche events also had changes of the back-azimuth below  $10^\circ$  and we therefore set the  $10^\circ$  as a maximum threshold value. Doing so, another 37 events were dismissed and only 27 avalanche events remained. 15 of the remaining avalanche events were observed during the 9 and 10 March 2017 and some events were detected during the other periods of considerable avalanche activity in February (Figure 6). Furthermore, another 10 single events were also confirmed, ~~but were randomly distributed over the season~~. For each of the 27 events we determined a mean back-azimuth, which is the mean direction the signals were coming from. The mean back-azimuths were all pointing towards the surrounding slopes where we expected avalanches to release (Figure 15). Events with a duration longer than 100s were detected coming either from the north-west or south-east.



**Figure 11.** Time series and corresponding spectrograms of an airplane detected at both arrays on 28 January 2017 at 9:17. a) shows the signal recorded at the Dischma array and b) at the Wannengrat array.



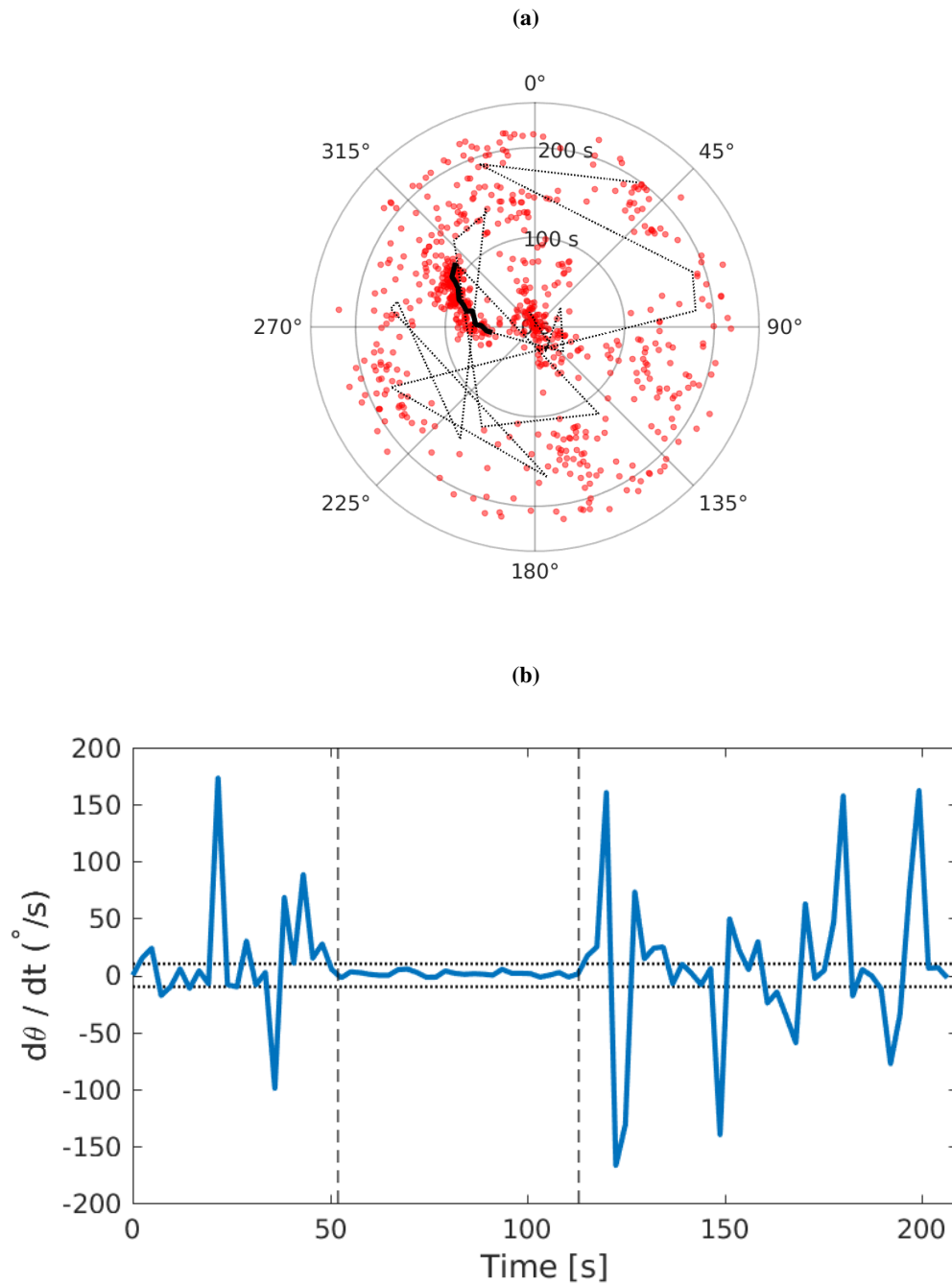


**Figure 12.** Yellow bars indicate the number of events detected at both arrays and turquoise bars only events recorded at the Dischma array. Blue bars are the number of airplanes and earthquakes detected at the Wannengrat array. Between 5 and 20 January no data were recorded at the Wannengrat array due to technical issues.

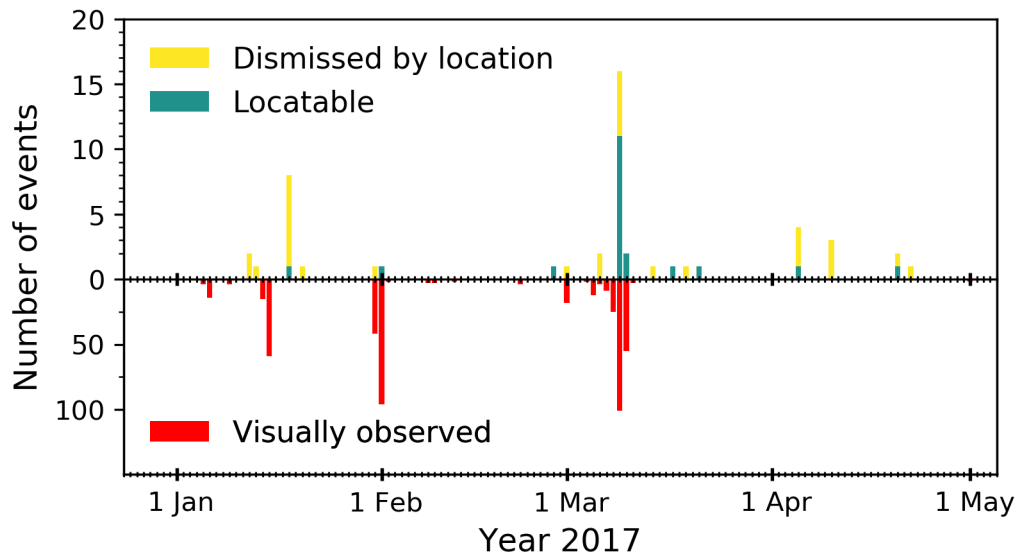
Apart from analyzing only the events remaining after the combined array classification, we also performed the localization post-processing for those 53 events we had dismissed. Based on the localization, 48 events were again dismissed, but 5 had a median back-azimuth path within the threshold value. Hence, by directly applying the localization step 32 avalanche events remained, but including at least 5 false detections (15%).

## 5 5 Discussion

We used hidden Markov models (HMMs), a machine learning algorithm, to automatically detect avalanches in data from seismic systems deployed above Davos, Switzerland. The approach builds on the work of Heck et al. (2018a), who adapted the HMM method developed by Hammer et al. (2017) to detect avalanches in continuous seismic data from a small aperture geophone array. Using their approach on our data resulted in automatic detections that still contained a large number of falsely classified events because only one event type (avalanche) and the background noise was used for the classification with HMM.



**Figure 13.** Localization results for [an](#) avalanche event recorded 9 March 2017 at 6:47-47. a) polar plot representation of the back-azimuth calculated using the MUSIC method. Red dots are the back-azimuth values for a single time window. The black line represents the median back-azimuth path. The solid part [of the line](#) has variations below the threshold [value for the derivative](#), whereas the dotted [shows line refers to](#) strong variations. b) derivative of median back-azimuth path. The dotted lines represents the threshold value of 10°. [The part between 52s and 113s corresponds to the solid line in a\).](#)

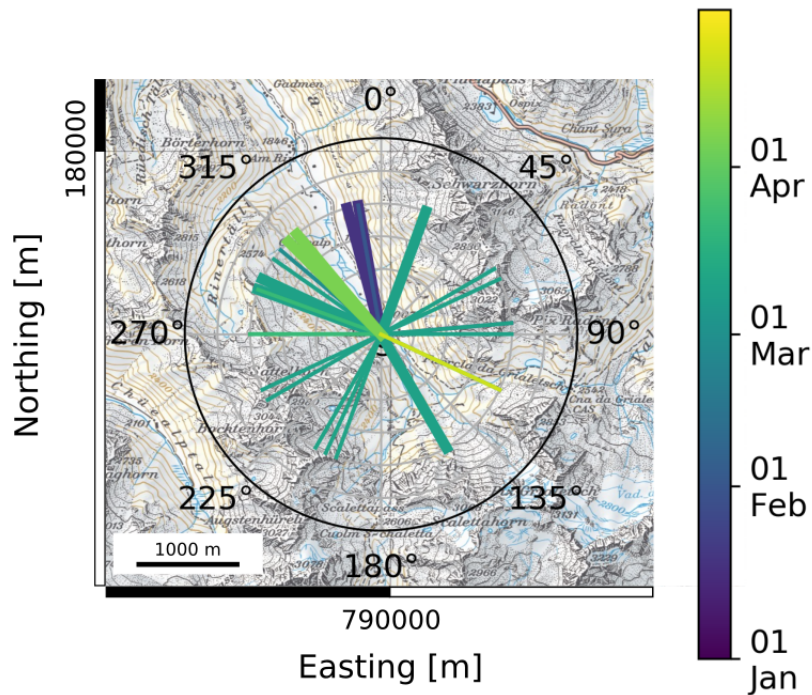


**Figure 14.** Turquoise bars are the number of events per day which are locatable and are considered as avalanches. Yellow bars are the number of events per day which were not locatable and therefore dismissed. Red bars are the number of avalanches visually observed in the area of Davos.

Earthquake and airplane signals have characteristics closer ~~the~~to avalanches than the background noise, and were therefore included in the detections.

By combining the classification results with a classification performed at a second array located 14km away, simultaneously recorded events such as local earthquakes and airplanes could be dismissed. In addition, we applied the multiple signal classification (MUSIC) method to estimate the back-azimuth of the detected events to eliminate false alarms. Overall, this work flow allowed us to automatically identify 27 events that were very likely generated by avalanches, as the temporal trend corresponded well with the avalanche activity for the region of Davos obtained through conventional visual field observations (Figure 6 and 14). It was not possible to confirm any event with visual observations since most avalanches released during periods of bad visibility. However, ~~based on the work of Heck et al. (2018b), we were able to confirm another 12 events identified~~ Heck et al. (2018b) manually identified 13 avalanches during 9 and 10 March ~~2017-2017~~, 12 of which were automatically identified with the approach presented here.

~~Recent work with HMMs by Heck et al. (2018a)~~ Apart from HMMs, several other machine learning techniques are suited to classify signals in seismic data. It is possible to use a convolutional neural network for earthquake detection and location (Perol et al., 2018) or to pick the P-wave arrival of seismic wave fields (Ross et al., 2018). Comparable to the classical HMM approach, these studies rely on large pre-labelled training data sets. Another approach is the so-called Random Forest classifier, which can be used to discriminate seismic waves (Li et al., 2018). Automatic classification approaches are also suitable to differentiate between earthquakes and quarry blasts (Hammer et al., 2013) or to characterize larger rockfalls (Dammeier et al., 2016).



**Figure 15.** Polar plot representation overlaid on a map section of the field site. The radius indicates the duration of the event, the angle represents the direction of the origin of the event, thin lines represent events with a duration  $< 60s$ , thick lines events with duration  $\geq 60s$ . The different colors of the lines represent the time of the year. Reproduced by permission of swisstopo (JA100118).

Further mass movements, such as landslides, can also be identified in the seismic data based on automatic classification approaches (Esposito et al., 2006; Hibert et al., 2014; Provost et al., 2016).

The automatic classification of avalanches yet remains a difficult task. Rubin et al. (2012) used several machine learning algorithms to identify avalanches in seismic data and compared the results obtained with the different approaches. With all methods a high probability of detection was achieved, but the number of false alarms was too high. A recent study by Heck et al. (2018a) showed that HMMs are a suitable tool to detect avalanches, but there is still a need for additional post-processing steps. The work presented here confirms that HMMs in combination with further post-processing steps provide reliable classification results.

In addition, Heck et al. (2018a) highlighted the difficulty in obtaining a reliable classifier trained on data from a geophone array very similar to the one used in this work. Their results showed that there were large differences in model performance between the sensors, with the number of detections per sensor ranging from about 150 to over 2000. This was attributed to local heterogeneities as the sensors were packed in a Styrofoam housing and inserted within the snowpack. Heck et al. (2018a) therefore suggested to deploy the sensors below the snow cover and either on or below the ground. In our deployment, the

geophones were buried about half a meter below the ground on a flat meadow. This approach was successful as it resulted in a much more consistent number of detections per sensor, ranging from 125 to 169. Clearly, the deployment strategy can have a substantial influence on the performance of the classifier.

In contrast to the classification approach used by Hammer et al. (2017), who ~~were able to use~~ used a fixed background model as they analyzed a relatively short period (5 days), we used an approach more suited for operational purposes. Indeed, for the operational set-up the background model was determined using 24 hours of data prior to the hourly data that were classified. In combination with the post-processing steps related to signal duration and number of ~~votes~~ sensors the events were detected as suggested by Heck et al. (2018a), the HMMs identified 117 possible avalanche events (Figure 8). Even though this approach identified the main avalanche cycle in March 2017 (compare Figure 6 and Figure 8), visual inspection of the classified events indicated that at least 50% of the events were false alarms produced by distant airplanes or regional earthquakes (Figure 9). Even by training a classifier with different feature combinations, changing the training event and/or the length of the training event, the classification results did not substantially change and airplanes and earthquakes were still classified as avalanches. This highlights the difficulty in training an accurate HMM for low energy signals generated by avalanches. We concluded that using HMMs to automatically identify avalanches in seismic data from our geophone array will inherently contain false detections, as the overall feature behavior from distant airplanes or regional earthquakes ~~were~~ was very similar to signals generated by avalanches (Figure 10).

To circumvent the problem of developing an optimal event classifier for one specific array, we made use of a second array at the Wannengrat. There we performed a second classification to automatically identify airplanes and earthquakes using an event model trained by an airplane event. Since transient signals produced by earthquakes, airplanes or avalanches have similarities, the results obtained for the second classification based on the airplane event model also falsely identified avalanches and earthquakes. Hence, a classification performed with only one event model was sufficient. The assumption for the second classification was that most falsely classified events were recorded at both arrays. Comparing the time series of detected events at both arrays allowed us to dismiss about 50% of the classified events (Figure 12). Identifying co-detections across arrays is therefore an efficient approach to reduce the number of false alarms. Although it was possible, that the classification results of the second array contained avalanche events, it was ~~really~~ unlikely that two avalanches released simultaneously at both field sites. Furthermore, avalanches were only recorded at one array since the distance between both arrays was about 14 km. In the future, a promising approach could be to reduce the distance to about 2 or 3 km, as this could also help improving the localization, while still allowing to dismiss false alarms improve the localization.

Although the combination of two arrays for the classification allowed us to reduce the number of false classification, some uncertainty remained about the origin of the identified events. In a final step, we therefore used the MUSIC method to estimate the median back-azimuth path, as suggested by Heck et al. (2018b), to further dismiss false detections. Similar approaches were suggested for the automatic detection of avalanches in infrasonic data by Marchetti et al. (2015) and Thüring et al. (2015). In those studies, the back-azimuth of continuous infrasound data was calculated on the fly using cross-correlation techniques, and only events with slight changes in back-azimuth over a predefined minimal duration were assumed as avalanche events. In contrast, here we only determined the back-azimuth for events automatically identified by the HMM with the MUSIC method,

as Heck et al. (2018b) showed that for our instrumentation pair-wise cross-correlation technique (beam forming) did not result in robust back-azimuth estimates. This last processing step further reduced the number of classified events to 27 (Figure 14). We also found out, that the MUSIC method would have been sufficient to determine the reliability of a detection as it was not possible to locate airplane or earthquake events with our array. After applying the localization based step to all detections, 32 events were identified as avalanches, 5 more events compared to the combined array and the localization based classification.

After applying the combined array classification and the MUSIC method to the data, 27 classification remained for the winter season 2017. Nearly all of ~~theses~~ ~~these~~ classifications occurred during periods of observed high avalanche activity (Figure 14). However, for the first two periods of high avalanche activity in January and February only few events were detected, whereas in the surroundings of Davos many events were observed. This may be due to fact that the Dischma site is located about 12km to the southeast of Davos and weather and snow conditions are sometimes different since major storms arrive ~~form~~ ~~from~~ the northwest. Indeed, based on the images from the automatic cameras, very little avalanche activity was observed in the area in January and February. Nevertheless, it was possible to reconstruct the avalanche activity period in March based on the automatic classification. Results from the localization showed that during the season avalanches released from many different slopes at the field site (Figure 15). This could be observed especially during the snow storm in March. A seismic monitoring system is therefore a suitable tool to monitor a wide area and not just one single slope. Although the detection range is with ~~3~~ ~~2~~ - ~~4km~~ ~~3km~~ rather small (Heck et al., 2018b) the seismic system in combination with an automatic classifier provides great potential to identify the major avalanche periods.

Although we were able to identify one major avalanche activity period in the winter season 2016 - 2017, the method presented here has its limitations. Based on the sensors used for the automatic monitoring, we identified avalanches within a range of 2 - 3km. However, by using more sensitive sensors, e.g. seismological broadband stations, the detection range of avalanches can be increased, even up to 30km for very large avalanches (run-out distance > 2 km (Hammer et al., 2017). However, it is difficult to deploy such sensors in mountain terrain, since these stations require existing infrastructure (e.g. electricity, storage room in a hut), which is typically not available at remote locations. In addition, the last post-processing step requires a second array. Hence low-power systems with less sensitivity proved to be the best solution. Furthermore, the limited power supply at the field sites also prevents performing first processing steps directly at the field sites and hence limits the possibility of near real-time analysis. However, it is possible to overcome this problem by designing special hardware for this particular task.

Based on the ~~here presented approach~~ approach presented here, a near real-time classification of the seismic data and hence a near real-time detection of avalanches seems possible. The computational times on a computer with a regularly available 8-core processor with 12GB ram are reasonably short and almost near real-time whereas the localization based on the MUSIC is very costly (three times real time). Although we decided to implement a combined array classification step to save computational time, directly localizing every detection is also possible. Since the amount of detections for the whole season was very low, a near real-time detection could be provided with or without the combined array classification. In future systems, pre-processing steps can be integrated in the data logging unit to reduce the amount of data while recording. Using a standard personal computer, feature calculation is performed near real-time for all sensors simultaneously as well as the HMM construction and the classification. However, a major obstacle of our method is the necessity of an adequate training event recorded at the

seismic array. Using training events recorded at different arrays ~~will not work~~ might be unreliable due to possible differences in the instrumentation and changes in the overall background noise or local heterogeneities in the local geology and in snow conditions. To set up the classification experts will ~~be still~~ still be needed to define correct and confirmed training events. Future research will ~~check~~ assess the possibility to use one training event for several seasons recorded at the same array.

## 5 6 Conclusions

During the winter season 2016-2017 we used a seismic array to continuously monitor avalanche activity in a remote area above Davos, Switzerland. By implementing an operational classification method based on hidden Markov models (HMMs), we ~~were able to detect~~ detected 117 events in the seismic data from January to April, which were likely produced by avalanches. Subsequent visual inspection revealed a false alarm rate of at least 50%. Most of the false detections were associated with  
10 airplanes or earthquakes. By implementing additional steps such as a combined array classification and the localization of the events based on multiple signal classifications (MUSIC), we improved the classification results by reducing the number of identified events to 27. Only using the localization to remove false detections resulted in at least 15% of false detections yet  
15 at a higher computational cost. Our results therefore show that dismissing false detections with a second array improves the overall classification accuracy. If a second avalanche monitoring array is in the vicinity, combing the results of both arrays will improve the classification results. In future experiments we want to reduce the distance between the arrays to some kilometers to improve the localization of avalanches ~~and the combined array processing~~.

*Acknowledgements.* M.H. was supported by a grant of the Swiss National Science Foundation (200021\_149329). We thank numerous colleagues from SLF for help with field work and maintaining the instrumentation.

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

## References

- Barnes, A. E.: Instantaneous spectral bandwidth and dominant frequency with applications to seismic reflection data, *Geophysics*, 58, 419–428, 1993.
- Besson, B., Eiriksson, G., Thorarinsson, O., Thorarinsson, A., and Einarsson, S.: Automatic detection of avalanches and debris flows by seismic methods, *J. Glaciol.*, 53, 461–472, <https://doi.org/10.3189/002214307783258468>, 2007.
- 5 Beyreuther, M., Hammer, C., Wassermann, J., Ohrnberger, M., and Megies, T.: Constructing a Hidden Markov Model based earthquake detector: application to induced seismicity, *Geophys. J. Int.*, 189, 602–610, <https://doi.org/10.1111/j.1365-246X.2012.05361.x>, 2012.
- Dammeier, F., Moore, J. R., Hammer, C., Haslinger, F., and Loew, S.: Automatic detection of alpine rockslides in continuous seismic data using hidden Markov models, *J. Geophys. Res. Earth Surf.*, 121, 351–371, <https://doi.org/10.1002/2015JF003647>, 2016.
- 10 Esposito, A. M., Giudicepietro, F., Scarpetta, S., D’Auria, L., Marinaro, M., and Martini, M.: Automatic discrimination among landslide, explosion-quake and microtremor seismic signals at Stromboli Volcano using neural networks, *Bull. Seismol. Soc. Am.*, 96, 1230, <https://doi.org/10.1785/0120050097>, +<http://dx.doi.org/10.1785/0120050097>, 2006.
- Hammer, C., Beyreuther, M., and Ohrnberger, M.: A seismic-event spotting system for Volcano fast-response systems, *Bull. Seismol. Soc. Am.*, 102, 948–960, <https://doi.org/10.1785/0120110167>, 2012.
- 15 Hammer, C., Ohrnberger, M., and Fäh, D.: Classifying seismic waveforms from scratch: a case study in the alpine environment, *Geophys. J. Int.*, 192, 425–439, 2013.
- Hammer, C., Fäh, D., and Ohrnberger, M.: Automatic detection of wet-snow avalanche seismic signals, *Nat. Hazards*, 86, 601–618, <https://doi.org/10.1007/s11069-016-2707-0>, 2017.
- Heck, M., Hammer, C., van Herwijnen, A., Schweizer, J., and Fäh, D.: Automatic detection of snow avalanches in continuous seismic data using hidden Markov models, *Nat. Hazards Earth Syst. Sci.*, 18, 383–396, <https://doi.org/10.5194/nhess-18-383-2018>, 2018a.
- 20 Heck, M., Hobiger, M., van Herwijnen, A., Schweizer, J., and Fäh, D.: Localization of seismic events produced by avalanches using multiple signal classifications, *Geophys. J. Int.*, <https://doi.org/10.1093/gji/ggy394>, <http://dx.doi.org/10.1093/gji/ggy394>, 2018b.
- Hibert, C., Mangeney, A., Grandjean, G., Baillard, C., Rivet, D., Shapiro, N. M., Satriano, C., Maggi, A., Boissier, P., Ferrazzini, V., and Crawford, W.: Automated identification, location, and volume estimation of rockfalls at Piton de la Fournaise volcano, *Journal of Geophysical Research: Earth Surface*, 119, 1082–1105, <https://doi.org/10.1002/2013JF002970>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JF002970>, 2014.
- 25 Hobiger, M., Cornou, C., Bard, P.-Y., Le Bihan, N., and Imperatori, W.: Analysis of seismic waves crossing the Santa Clara Valley using the three-component MUSIQUE array algorithm, *Geophys. J. Int.*, 207, 439–456, 2016.
- Joswig, M.: Knowledge-Based Seismogram Processing by Mental Images, *IEEE Trans. Syst. Man. Cybern.*, 24, 429 – 439, 1994.
- 30 Kanasewich, E. R.: *Time Sequence Analysis in Geophysics*, The University of Alberta Press, Edmonton, Alberta, Canada, 1981.
- Kramer, S.: *Geotechnical Earthquake Engineering*, Always learning, Prentice-Hall international series in Civil Engineering and engineering mechanics, Upper Saddle River, New Jersey, 1996.
- Lacroix, P., Grasso, J.-R., Roulle, J., Giraud, G., Goetz, D., Morin, S., and Helmstetter, A.: Monitoring of snow avalanches using a seismic array: Location, speed estimation, and relationships to meteorological variables, *J. Geophys. Res. Earth Surf.*, 117, F01034, <https://doi.org/10.1029/2011JF002106>, 2012.
- 35 Leprettre, B., Navarre, J., and Taillefer, A.: First results from a pre-operational system for automatic detection and recognition of seismic signals associated with avalanches, *J. Glaciol.*, 42, 352–363, <https://doi.org/10.1109/ICMLA.2012.12>, 1996.



- Li, Z., Meier, M.-A., Hauksson, E., Zhan, Z., and Andrews, J.: Machine Learning Seismic Wave Discrimination: Application to Earthquake Early Warning, *Geophysical Research Letters*, 45, 4773–4779, <https://doi.org/10.1029/2018GL077870>, 2018.
- Marchetti, E., Ripepe, M., Ulivieri, G., and Kogelnig, A.: Infrasound array criteria for automatic detection and front velocity estimation of snow avalanches: towards a real-time early-warning system, *Nat. Hazards Earth Syst. Sci.*, 15, 2545 – 2555, <https://doi.org/10.5194/nhess-15-2545-2015>, 2015.
- McClung, D. and Schaerer, P. A.: *The Avalanche Handbook*, The Mountaineers Books, 2006.
- Nishimura, K. and Izumi, K.: Seismic signals induced by snow avalanche flow, *Nat. Hazards*, 15, 89–100, <https://doi.org/10.1023/A:1007934815584>, 1997.
- Ohrnberger, M.: Continuous automatic classification of seismic signals of volcanic origin at Mt. Merapi, Java, Indonesia, PhD. thesis, 2001.
- Perol, T., Gharbi, M., and Denolle, M.: Convolutional neural network for earthquake detection and location, *Science Advances*, 4, <https://doi.org/10.1126/sciadv.1700578>, 2018.
- Provost, F., Hibert, C., and Malet, J.-P.: Automatic classification of endogenous landslide seismicity using the Random Forest supervised classifier, *Geophysical Research Letters*, 44, 113–120, <https://doi.org/10.1002/2016GL070709>, 2016.
- Rabiner, L.: A tutorial on Hidden Markov Models and selected application in speech recognition, *Proceedings of the IEEE*, 77, 257–286, <https://doi.org/10.1109/5.18626>, 1989.
- Ross, Z. E., Meier, M.-A., and Hauksson, E.: P Wave Arrival Picking and First-Motion Polarity Determination With Deep Learning, *Journal of Geophysical Research: Solid Earth*, 123, 5120–5129, <https://doi.org/10.1029/2017JB015251>, 2018.
- Rost, S. and Thomas, C.: Array seismology: methods and application, *Rev. Geophys.*, 40, 2.1–2.27, <https://doi.org/10.1029/2000RG000100>, 2002.
- Rubin, M., Camp, T., van Herwijnen, A., and Schweizer, J.: Automatically detecting avalanche events in passive seismic data, *IEEE International Conference on Machine Learning and Applications*, 1, 13–20, <https://doi.org/10.1109/ICMLA.2012.12>, 2012.
- Schaerer, P. A. and Salway, A. A.: Seismic and impact-pressure monitoring of flowing avalanches, *J. Glaciol.*, 26, 179–187, <https://doi.org/10.1017/S0022143000010716>, 1980.
- Schmidt, R.: Multiple emitter location and signal parameter estimation, *IEEE Trans. Antennas Propag.*, 34, 276–280, <https://doi.org/10.1109/TAP.1986.1143830>, 1986.
- Scott, E., Hayward, C., Kubichek, R., Hamann, J., Pierre, J., Comey, B., and Mendenhall, T.: Single and multiple sensor identification of avalanche-generated infrasound, *Cold Reg. Sci. Technol.*, 47, 159–170, <https://doi.org/10.1016/j.coldregions.2006.08.005>, 2007.
- Suriñach, E., Furdada, G., Sabot, F., Biescas, B., and Vilaplana, J.: On the characterization of seismic signals generated by snow avalanches for monitoring purposes, *Ann. Glaciol.*, 32, 268–274, <https://doi.org/10.3189/172756401781819634>, 2001.
- Suriñach, E., Vilajosana, I., Khazaradze, G., Biescas, B., Furdada, G., and Vilaplana, J.: Seismic detection and characterization of landslides and other mass movements, *Nat. Hazards Earth Syst. Sci.*, 5, 791–798, <https://doi.org/10.5194/nhess-5-791-2005>, 2005.
- Taner, M. T., Koehler, F., and Sheriff, R.: Complex seismic trace analysis, *Geophysics*, 44, 1041–1063, 1979.
- Thüring, T., Schoch, M., van Herwijnen, A., and Schweizer, J.: Robust snow avalanche detection using supervised machine learning with infrasonic sensor arrays, *Cold Reg. Sci. Tech.*, 111, 60 – 66, <https://doi.org/10.1016/j.coldregions.2014.12.014>, 2015.
- van Herwijnen, A. and Schweizer, J.: Seismic sensor array for monitoring an avalanche start zone: design, deployment and preliminary results, *J. Glaciol.*, 57, 257–264, <https://doi.org/10.3189/002214311796405933>, 2011a.
- van Herwijnen, A. and Schweizer, J.: Monitoring avalanche activity using a seismic sensor, *Cold Reg. Sci. Tech.*, 69, 165–176, <https://doi.org/10.1016/j.coldregions.2011.06.008>, 2011b.

## Reply to reviewer 1

We thank the reviewer for the constructive comments. Below we reply in detail.

P2, 5: what does “rather poor” mean? Can you quantify it or specify? Is this their conclusion or your interpretation?

Bessason et al. obtained a probability of detection around 65%. Seismic signals generated by rockfalls and debris flows were wrongly classified as avalanches and vice versa. Information about the false alarm rate was not given directly, but seems to have been high. (P.2 L.3)

*Bessason et al. 2007 used a nearest neighbor approach to classify newly recorded events. Using this approach, they were able to detect 65% of all confirmed avalanches.*

P2, 30: I think it is unclear what these arrays are. e.g. the one to locate avalanches and the one at 14 km distance and then you are talking about one in Dischma Valley and one at Wannengrat array. Are these the same arrays or different ones? Maybe the names or location of arrays should be introduced earlier and a link to figure 1 should be added?

Thanks for pointing this out. We rephrased this sentence and added a reference to Figure 1. (P.2 L.34)

*We performed the classification and localization of the events with the data recorded at the seismic array located in the Dischma Valley above Davos, Switzerland during the winter season 2017 (yellow square in Figure 1). These results were then combined with data obtained at the Wannengrat array, which is located 14 km to the northwest of the Dischma field site (red square in Figure 1).*

P2, 32: is this the winter season 2016/17?

Yes, we referred to the winter season 2016-2017. We changed the number in the text.

P3, 10: two “)” too much

Removed.

P3, 15: “)” too much

Removed.

P3, 15: where were these cameras and weather stations located?

We included a detailed view of both field sites including all different types of weather stations and automatic cameras.

P4, 2: In this sentence you describe that the cameras helped to identify avalanches in the winter of 2016/17. But then you cite a publication from 2011? Clearly this publication does not describe the winter 2016/17? Maybe rephrase.

We rephrased this sentence to clarify that van Herwijnen et al. performed a similar analysis in 2011 in the surroundings of Davos. (P.4 L.5)

*As already shown for avalanche activity periods in the winter season 2009-2010 by van Herwijnen et al. 2011b, those images can help to identify and confirm seismic events produced by avalanches.*

P4, 6: Is the amplitude in noise that stable in time, that it ok to use a fixed threshold like you do or did you change it in time?

As mentioned in the text, we use a threshold value that is 5 times higher than the daily mean amplitude, i.e. the noise amplitude threshold changes each day.

P4, 6: given a sampling rate of 500 Hz your time window is only 2 seconds long when selected. This sounds pretty short to me when looking for avalanches.

In this step we are looking at the seismic energy recorded at the sensors for a 2 second window. If the energy for this window is higher than the threshold, we look at the following windows. If the energy for these windows also reaches the threshold value, we concatenate the windows. As soon as the energy decreased below the threshold value, we cut an additional 60 s before and after the total length of the event to not dismiss the onset and the coda of the event.

We clarified this in the text. (P.4 L.13)

*For a window  $i$  with a length of 1024 samples a mean absolute amplitude  $A_i$  was determined. When  $A_i \leq 5 A^*$ , with  $A^*$  the daily mean amplitude, the data within the window were cut. If the amplitude threshold for the following window was also reached, data were concatenated. Furthermore, a section of  $t=60$  s was cut before and after the window to ensure that the onset and coda of each event was incorporated. Doing so, data were reduced by 80% to several data windows of various lengths.*

P5, 4 “Using these properties, a widespread background model can be learned from the general properties” I think this sentence sounds odd. Are you trying to built a model from information you derive from the general properties?

We rephrased the sentence. (P. 5 L.14)

*From these properties a widespread background model can be learned.*

P5, 4-7: I cannot follow how your method works in detail. Maybe the text should be rewritten with more reference to figure 3?

We rephrased the section of the brief introduction of the Hidden Markov Models. However, this topic is rather complex and we provided references. (P.5 L.4)

*To automatically identify avalanches in the continuous seismic data we used hidden Markov models (HMMs). These statistical classifiers use a sequence of multivariate Gaussian probability distributions to model observations (e.g. seismic time series). To determine the characteristics of the distributions (i.e. mean and covariance) a large number of training sets of known events, so called pre-labeled training sets, are required. For each different type of observation (e.g. avalanche, airplane or earthquake in the seismic data) a separate HMM is trained. By combining all HMMs the whole classification system with several classes is constructed. This classical approach, which relies on a large number of well-known pre-labeled training sets, was successfully used to automatic identify seismic events in continuous seismic data (Orhnberger 2001, Beyreuther et al. 2012).*

*Avalanches, however, are rare events and it is nearly impossible and too time consuming to obtain a large training set.*

*To circumvent this, we performed the classification based on an approach developed by Hammer et al. (2012) exploiting the abundance of data containing mainly background signals to obtain general wave-field properties. From these properties a widespread background model can be learned. A new event model (e.g. representing avalanches) is obtained by using the background model to adjust the event model description by using only one training event. In contrast to the classical HMM approach, the classification system of this approach consists of a background model and one event model for each implemented event [or observation]? class.*

*The classification process itself calculates the likelihood that an unknown data stream was generated by a specific event [or observation]? class for each individual HMM class (Hammer et al. 2012, Hammer et al. 2013).*

**P5, 14: are assuming that two events are separated by at least 24 hours? And if two events have a closer spacing in time they cannot be picked/ located?**

We do not assume that two events are separated by at least 24 hours. We just use the data within a 24-hour window to construct our background model and a one hour window is then classified. Once the classification is performed, we shift both windows one hour.

We added additional information to clarify this. (P.6 L.8)

*This means, that our background model is always determined by the pre-processed data of a 24-hour window. By choosing a length of 1 h for the window  $t_{class}$ , we were able to classify the pre-processed continuous seismic data of one hour during one step of the operational classification. Once one classification step, which is the classification of the window  $t_{class}$  is finished, both windows are shifted by one hour and the classification was executed for the shifted windows.*

**P5, 15: you state that the  $t_{class}$  window is 1 h long, but on p4, 6-8 you state that the chosen event is only 122 s long. Is there an error somewhere?**

$t_{class}$  is the length of the window we want to classify, which means that we want to identify all events within this window and define their origin using the HMM. The length of the identified events is independent of  $t_{class}$ .

**P5, 25: What is an instantaneous frequency?**

The instantaneous frequency is the time dependent change of phase.

Taner, M., F. Koehler, and R. Sheriff (1979). Complex seismic trace analysis, Geophysics 44, no. 6, 1041–1063. We now include this reference in the text.

**P5, 27: maybe you should explain cepstral coefficients?**

We added a reference for the cepstral coefficients. (P.7 L.7)

**P5, 30: what do you mean with “the first half-octave band has [. . .] a total number of 6 bands”?**

In total, we calculated 6 half-octave bands. The first half-octave band had a central frequency of 3.9 Hz. Depending on this frequency, the central frequency for the remaining half-octave bands are already determined. We rephrased this sentence. (P.7 L.14)

We used in total 6 half-octave bands for the classification and the first half-octave band had a central frequency of 3.9 Hz.

P6, 6: what if the event model is very unlike the avalanche signal you try to detect?

Then we would not be able to identify avalanches in the winter season. Therefore, it is crucial to use a training event which is representative of avalanches at this specific field site. We changed the wording in the text to emphasize the importance of the event model and included a reference to Heck et al. (2018). (P.7 L.24)

*The event model  $HMM_{Event}$  was learned using only one training event that is representative of avalanches at a specific field site (Heck et al. 2018a). It was determined once and then applied for the entire winter season.*

P6, 10: “Each classified event having a duration shorter than 12 s was dismissed” replace with “Each classified event shorter than 12 s in duration was dismissed”

We rephrased this sentence.

p7, 6: I am a bit surprised that you state that your second array at 14 km distance does not record the avalanche any more. After all you mentioned this array in the introduction that could detect avalanches up to a distance of 30 km (p2, 10).

The detection radius strongly depends on the size of the avalanche, and to a lesser degree on the sensors used to record the signals. It is possible to identify avalanche within a range of 30 km, however, only by using broadband seismometer and only for very large events (runout distances > 2000 m). These events are very uncommon and were not observed during our study period. The avalanches we monitored were substantially smaller avalanches and we used less sensitive geophones. For these avalanches, we recently showed that the detection radius is about 3 km (Heck et al., 2018b).

p7, 8: “12 km away” replace with “at 12 km distance”

We rephrased this sentence.

p7, 10: rephrase the heading as I find it pretty unspecific

We rephrased the heading. (P.9 L.1)

*Localization results to confirm avalanches*

p7, 12: what MUSIC code did you use? Where is it available?

The MUSIC code we used was developed by Dr. Manuel Hobiger from the Swiss Seismological Service. We will supply the scripts and data on an open access storage.

p7, 27: does this approach not exclude avalanches along other potentially longer or more curved paths?

We analyzed one event which approached the array to less than 100 m and therefore exhibited large changes in the back-azimuth. These changes were still well within the limits we used here. For more distant avalanches, changes in back-azimuth will be much smaller, even for more curved paths.

p7, 33: what is the “used array”?

The used array is also the Dischma array. We rephrased this sentence. (P.11 L.4)

... performed at the Dischma array.

p7, 34: “through further analysis” instead of “by further analysis”?  
Changed as suggested.

p8, 2: “to speed up the calculation time”: you “reduce the calculation time” or “speed up the calculation”

We rephrased the sentence.

p8, 2: so if I understand this correctly for a 2 minute long window it takes 6 minutes to process? So in order to do this in real time you need to skip time windows e.g. of “noise”  
That is correct. This is also one reason to pre-process the data.

p8, 15: figure 4a p8, 16: figure 4b p9, figure 4: maybe remove the legend in figure 4b as the information is already there as label of the y axis. Could you limit the yaxis at 110 or so in order to make the low numbers of avalanches in February more visible?

We changed the figures as suggested.

p9, 2: On p7, 30 you state that you minimum event length is 20s whereas here you state it is 12 s.

These are two different post-processing steps. In the first step, we applied a duration threshold to the events obtained from the HMM classification, as suggested by Heck et al. (2018). In the second part, however, we applied a duration threshold to the median back-azimuth path obtained by MUSIC. This median path was calculated using a sliding window of 8 seconds. To obtain reliable results, it was thus necessary to use a second duration threshold value. We clarified this better in the text. (P.9 L.20)

*Heck et al. 2018a suggested that a detected event should have a minimum duration of 12 s to be considered as an avalanche. For the localization step, however, it was necessary to increase this duration because the window length used for the median smoothing filter was already 8 s long (Heck et al. 2018b).*

*To cover enough data points to use the minimal event duration as a reliable classification criterion, we therefore required a minimum length of 20 s for the back-azimuth path.*

p9, 6: What do you mean with “classes with 5 and 6 votes” what votes?

This term refers to the events that are detected by 5 sensors or 6 sensors of the array. We rephrased this sentence and removed the terms votes and classes, to avoid confusion. (P.11 L.15)

*A quarter of the events were detected by 5 sensors, a quarter by 6 and about half by 7.*

p9, 10: It that a good thing or a bad thing that you detect avalanches that are not listed in figure 4? E.g. does this mean that there are avalanches missing in figure 4 that should have been listed or are there completely different avalanches recorded in different areas and the only common thing is the huge amount of snow in that time period?

In principle, it is entirely possible that we detect avalanches with our monitoring system that were not recorded by visual observations in the area of Davos. However, at this point in the analysis, it remains unclear if the events automatically identified with the HMM correspond to

avalanches or not. We therefore manually went through all the 117 detected events to evaluate if the signal characteristics correspond to those typically seen for avalanches.

p10, figure 5: move the sentence “the red area...” up to the description of subfigure a

We changed the caption. (P.11)

*Avalanche released on 9 March 2017 at 06:47 used as training event for the classifier. a) time series for the 7 sensors. The red area indicates the part of the time series used as training event. b) corresponding spectrogram of the seismic time series.*

p10, figure 6: what do you mean with vote in the legend? What is a vote in the context of avalanches?

As already mentioned earlier, this refers the number of sensors the signal was detected at. We changed this term, since it is confusing here and throughout the text and figures.

p11, 2: two “)” too much

We changed it.

p11, 21: I keep wondering why you detect the avalanches only up to 4 km distance and not 30 km distance as mentioned in the introduction.

This is due to the instrumentation used and the size of the avalanches that were monitored. We added some references in the introduction considering this small range of detection. (P.2 L.26)

*In the both studies of Lacroix et al. (2012) and Heck et al. (2018) less sensitive vertical component geophones were used for the seismic monitoring resulting in an avalanche detection of approximately 3 km.*

p11, 31: “except for detections at the beginning of April” consider replacing it with: “except for detections at the beginning of April and no detections at the beginning of February” Any ideas as of why these avalanches in February could not be detected?

During the winter 2016 – 2017 we had slightly different meteorological conditions at the Dischma field site compared to the surroundings of Davos. In particular, there was less snow in the back of the Dischma valley resulting in fewer avalanches early in the season. Based on field observations and the images from the automatic cameras, we determined that in early February there were no medium-sized to large avalanches in the vicinity of our array.

P12, figure 7: Does the gray area in figure 7 have the same meaning as the red area in figure 5? If yes I suggest to use the same color.

The red area in Figure 5 is used as the training event, so we manually defined this area. The gray area in Figure 7, however, shows the part of the time series, which was identified as an event by the hidden Markov model. Therefore these two areas have a different meaning.

P12, figure 7: I am surprised about the low frequency content of the airplane and the lack of overtones. How did you classify this as an airplane?

We agree with the reviewer that typically signals generated by airplanes either have clear overtones or at least a clear Doppler effect in the signal. However, based on our experience with seismic data from both arrays above Davos, we are confident that the signals shown in Figure 7 are generated by airplanes. We do not know yet why airplanes generate such signals, and we have not identified a clear pattern, which explains their presence. Perhaps it is related to the altitude or distance of the airplane, or perhaps specific atmospheric conditions are

responsible. Nevertheless, we have seen multiple signals like these recorded at both arrays and we are confident that these signals are generated by airplanes.

[P13, figure 8: What is the unit of the normalised time and how is it calculated? Do the events have the same length or did you just stretch/ squeeze them to fit in between 0 and 1?](#)

In Figure 8, time is normalized by the duration of each event. The normalized time therefore has no -units. 0 corresponds to the start of the event and 1 to the end of the event. The duration of the events can therefore not be determined from this plot as each event has a different duration. We chose this representation to better highlight similarities in the temporal behavior of the features for different events.

[p13, 6: one “\)” too much](#)

We changed it.

[p13, 8: one “\)” too much](#)

We changed it.

[p13, 11: Do you know what these 37 other avalanche like events might be? Maybe these are just avalanches along an unexpected path or longer paths?](#)

Based on a visual inspection of the seismic time series and spectrograms we assumed that we have several sources for these signals. For one part, some airplane signals were still present in the classification results after the combined array classification and signals produced during periods of strong winds.

[p13, 16: It sounds to me a bit like you remove events until you end up with back azimuths or locations you would like to get.](#)

We did not filter specific back-azimuth angles. We only require the derivative of the median back-azimuth path to fall within certain threshold values (i.e. the source of the signal should not vary too much within a certain period of time), independent of the values of the back-azimuth angle. The fact that the remaining events all had a mean back-azimuth pointing towards nearby avalanche slopes therefore confirms that this method is suited to identify signals likely generated by avalanches.

[p14, figure 9: How do you know that these are airplanes?](#)

During the last years we analyzed the continuous seismic data recorded during the last winter periods. These types of signals were often recorded almost simultaneously at both field sites at any time of the winter season, even in the beginning, when there was no snow at the field sites. We also compared the times of some signals with flight information and were able to identify the airplanes. See also comment above.

[p15, discussion: I find the discussion a bit repetitive with respect to the rest of the manuscript. Many points seem to have been made already in the rest of the text. Also my impression is that they barely refer to work of others in the discussion i.e. papers that are not lead by “Heck” or “Hammer”.](#)

We agree with the reviewer that some parts of the discussion were redundant. We therefore partly rewrote the discussion to avoid this. However, since we are not aware of many studies which tackle the automatic detection of mass movements (avalanches or other) in continuous seismic data, there are not many other studies which are relevant. (P.20 L.5)



*Apart from HMMs, several other machine learning techniques are suited to classify signals in seismic data. It is possible to use a convolutional neural network for earthquake detection and location (Perol et al., 2018) or to pick the P-wave arrival of seismic wave fields (Ross et al., 2018). Comparable to the classical HMM approach, these studies rely on large pre-labelled training data sets. Another approach is the so-called Random Forest classifier, which can be used to discriminate seismic waves (Li et al., 2018). Automatic classification approaches are also suitable to differentiate between earthquakes and quarry blasts (Hammer et al., 2013) or to characterize larger rockfalls (Dammeier et al., 2016). Further mass movements, such as landslides, can also be identified in the seismic data based on automatic classification approaches (Esposito et al., 2006; Hibert et al., 2014; Provost et al., 2016). The automatic classification of avalanches yet remains a difficult task. Rubin et al. (2012) used several machine learning algorithms to identify avalanches in seismic data and compared the results obtained with the different approaches. With all methods a high probability of detection was achieved, but the number of false alarms was too high. A recent study by Heck et al. (2018a) showed that HMMs are a suitable tool to detect avalanches, but there is still a need for additional post-processing steps. In the work presented here we confirm that HMMs in combination with further post-processing steps provide reliable classification results.*

p15, figure 10: change to that the legend is not overlapping the bar any more  
We changed the legend.

p16, figure 11: “for avalanche event” replace with “for an avalanche event”  
We rephrased it.

Figure 11a: I don’t understand to what part of the figure you refer to with “solid part”.  
Beneath what threshold?

The solid part refers to the solid line in the polar plot.

Figure 11b: is this really the derivative of the angle (y axis label) or derivative of the back-azimuth path (caption)? To me this figure seems to show the “angle” or “back azimuth” during, before and after the avalanche event with very stable back azimuths during the event and larger scatter afterwards.

This plot shows the approximate derivative of the angle (due to the small amount of data points). It was calculated using the angdiff function of matlab and the step size dt between the data points  $\text{angdiff}/dt$ . We changed the label to clarify this. (P.17)

p17, figure 12: so there are 100 visually observed avalanches in Davos but you could detect only 20? Were you too far away or was this recorded but not classified as event? Move the legend so that it does not overlap with the bars

Indeed the differences in the number of avalanches relate to the size of the area that is monitored. The visual observations are made for an area of about  $175 \text{ km}^2$ , while with the seismic system only avalanches within a radius of 3 km can be monitored ( $\sim 30 \text{ km}^2$ ). We changed the legend of the figure as suggested. (P.18)

p17, 1: “closer” replace with “closer to”?  
We changed it.

p17, 8-10: First you say that you could confirm no avalanche visually, but in the next sentence you state that “another 12” events were identified. Were they identified in a different way i.e. not visually or is there an error in the sentence?

The events were not visually identified, but by inspecting the seismic time series and analyzing the events in more detail. Based on the results, these 12 events were identified as avalanches.

We clarified this point in the manuscript. (P.20 L.3)

*However, Heck et al. 2018B manually identified 13 avalanches during 9 and 10 March 2017, 12 of which were automatically identified with the approach presented here.*

P18, figure 13: How do you know to what distance the duration of the event corresponds to?

The duration and back-azimuth of the event are automatically determined with our method. We do not have any information on the distance of the source. Perhaps the reviewer misunderstood the results shown in Figure 13 and we therefore changed this figure. In the polar plot the direction of the lines indicates the direction of the back-azimuth, the thickness of the lines indicates the duration of the event. Thin lines correspond to short events, thick lines to long events. The thicker the line, the longer the event. The color code shows the release time of the avalanche. (P.19)

P18, 4: number of votes: in my opinion it would be better to replace “vote” with something like “detections on sensors” or similar.

As already mentioned above, we changed this term, since it is confusing.

p18, 12: the overall feature behavior from distance airplanes... “was” not “were”

We changed it.

p19, 9: remove “really”. Based on the 5 events that were possible to locate, it is apparently possible to detect some avalanches on both arrays.

We changed it.

p19, 9: I am not sure I fully agree. It is not possible to record an avalanche at 14 km distance if it couples to the ground sufficiently or is large enough?

It might be possible to record avalanches at distances  $> 14$  km. However, as previously mentioned, this is only for large and catastrophic avalanches which were not observed during our study period.

P19, 10: “since distance” replace with “since the distance”

We changed it.

p19, 10: I am not sure where installing two arrays at 2-3 km distance would help. They would then pick up the same avalanches, and hence “events recorded at both arrays” are then not a valid criteria any more to find falsely classified earthquakes or airplanes.

We agree. Two arrays will mainly improve the localization.

p19, 11: “improving” replace with “improve”

We changed it.

p19, 22-24: Can you not locate airplanes and earthquakes with the array because the frequency content is different? So if the MUSIC method is perfectly suitable of detecting

avalanches, why should one go through the hassle of finding an exemplary event, the need of having two arrays and then removing a lot of false detections? Rather than using the output from the array method to detect events?

We performed several tests during this study considering the calculation time of the MUSIC method. Depending on the frequency range of the signals, the duration of the calculation varied. However, based on the actual code, it was not possible to apply the MUSIC method in near real-time. If we skipped the classification process at all, we would have to analyze all data remaining after the pre-processing step. It would take too much time to calculate the MUSIC results for all these events, even if we reduced the frequency range to a minimum. The classification process, however, has proven to be faster and allowing us to perform a near real-time classification.

p19, 26: typo in “theses”

We changed it.

p19, 30: typo “form”

We changed it.

p19, 32: “avalanches were released” instead of “avalanches released”?

We changed it.

p20, 5: Why is it that costly? Can the processing be sped up?

At the moment the algorithm is written in MATLAB. In this algorithm, the covariance matrix of small time windows is calculated. It might be possible to speed up the calculation process by using computers with more RAM, since matrix calculations depend on the size of RAM. Unfortunately, we are not computer engineers and our knowledge is too limited to answer the question.

p20, 14: “be still needed” replace with “still be needed”

We changed it.

p21, references. There are 11! referrals in the text to a not published paper (Heck et al. 2018b). Can the authors provide the manuscript in order to cross-check e.g. the content?

The paper is now accepted and accessible online: <https://doi.org/10.1093/gji/ggy394>

## Reply to reviewer 2

We thank the reviewer for the constructive comments. Below we reply in detail.

1.) I'd like the author to better address the possible technical limitations of their methods, in particular the field deployment and the near-real-time application of the classification methods based on two stages.

We adapted the Discussions section and added a more detailed description of the limitations of our method (P. 22 L.5).

*Although we were able to identify one major avalanche activity period in the winter season 2016-2017, the method presented here has its limitations. Based on the sensors used for the automatic monitoring, we identified avalanches within a range of 2 – 3 km. However, by using more sensitive sensors, e.g. seismological broadband stations, the detection range of avalanches can be increased, even up to 30 km for very large avalanches (run-out distance > 2 km) (Hammer et al. 2017). However, it is difficult to deploy such sensors in mountains terrain, since these stations require existing infrastructure (e.g. electricity, storage room in a hut), which is typically not available at remote locations. In addition, the last post-processing step requires a second array. Hence low energy systems with less sensitivity proved to be the best solution. Furthermore, the limited power supply at the field sites also prevents performing first processing steps directly at the field sites and hence limits the possibility of near real-time analysis. However, it is possible to overcome this problem by designing special hardware for this particular task.*

2.) The network geometry has a strong impact on the success rate of this latter criterion, could you add some details on that?

Unfortunately, the resolution of our seismic array was limited due to the instrumentation used. Lacroix et. al 2012, e.g., used a seismic array with larger distances between the sensors resulting in a better resolution for seismic waves. They could use beam-forming methods to calculate the source direction. We assume that with a larger interstation distance the resolution can be improved. With a larger array, a new comparison between beam-forming and MUSIC would be required.

3.) I have the impression that this second stage can be surely useful to recognize earthquakes but it probably needs a calibration for anthropic sources.

It depends on the type of anthropogenic source. Helicopters, for example, have a very characteristic spectrogram and Doppler effects can be observed at all times. For airplanes, this is a little bit more complicated. Due to the different types of airplanes (turbine or propeller) and the different flight altitudes, signals vary. But most of the time, a dominant frequency is visible as well as Doppler effects. However, the airplanes, which were considered as avalanches by our algorithm, did not have a dominant frequency. We expect that these airplanes were flying at a larger distance to the seismic array and due to the topography signals were naturally filtered.

We gained this knowledge over the past years in a long learning step. We agree that calibrations for anthropogenic sources may provide more reliable classification results.

4.) In addition, technical limitations in such extreme environments like high Alpine areas (e.g., data transmission) can be a possible trivial but concrete limitation for a real time application.

That is correct. The biggest problem we have is the computational power at the field sites. At the moment, we rely on Raspberry Pis which are mainly used for data storage and data transfer purposes. Data are transmitted and then mainly processed at our institute. A first improvement would be, to establish a fast and stable wireless link to the arrays to provide good data transfer. Second, better hardware with a low energy consumption would also be a great advantage. However, hardware capable to perform our calculations near real-time cannot run solely on solar power and batteries.

5.) [The application of the proposed methodology on another dataset gathered on another test site would be of great interest for the reader. For instance, is it possible to run the methods the other way round, testing it on the other array currently used for the second classification step?](#) Of course it would be possible to perform the classification task the other way round.

However, due to some technical issues, we only recorded with two sensors at the second field site. As we showed in a previous study (Heck et. al 2018a), best classification results with the HMMs are obtained by only considering events classified as an avalanche by at least five sensors. We could perform the classification with only two sensors, however, we expect too many false alarms. Furthermore, it is impossible to determine the source direction of the signals based on only two vertical geophones. Hence, the localization step could not be used to confirm or neglect detected events as avalanches.

6.) [Visual observations are used as validation, could the authors add some information about that? Which are the observation sources? How is compiled the avalanche catalog? If available, an image of one reference event could be also useful to show the test site.](#)

We have installed several automatic cameras at the field sites, still, we could only determine the exact release time for two avalanches. For the remaining events, especially for the main avalanche activity period in March, we narrowed the release times down to a 24-hour window. This was due to the fact, that most avalanches released during snow storms or at night when the visibility was bad.

In addition to the cameras, we relied on the avalanche data base compiled by the avalanche warning service in Davos. They monitor the avalanche activity for the region of Davos (~ 175 km<sup>2</sup>) and also use information from voluntary observers.

We included a picture shortly after the avalanche period in March taken during a field survey.

7.) [Figure 2, it would be useful to add a map with terrain information \(slope, morphology, etc\).](#)

Since the interstation distance between the sensors is too small, it is nearly impossible to show additional terrain information in these particular figures. Hence we added two additional figures, each showing additional information of the field sites also including the location of all instrumentation, at a lower scale (P.6).