Phenomenology of Avalanche Recordings From Distributed Acoustic Sensing

Patrick Paitz[1] , Nadja Lindner[2], Pascal Edme[2] , Pierre Huguenin[3], Michael Hohl[3], Betty Sovilla[3] , Fabian Walter[1], and Andreas Fichtner[2]

[1]WSL, Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf, Switzerland, [2]ETH Zurich, Department of Earth Sciences, Institute of Geophysics, Seismology and Wave Physics, Zürich, Switzerland, [3]SLF, WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland

Abstract Avalanches and other hazardous mass movements pose a danger to the population and critical infrastructure in alpine areas. Hence, understanding and continuously monitoring mass movements are crucial to mitigate their risk. We propose to use Distributed Acoustic Sensing (DAS) to measure strain rate along a fiber-optic cable to characterize ground deformation induced by avalanches. We recorded 12 snow avalanches of various dimensions at the Vallée de la Sionne test site in Switzerland, utilizing existing fiber-optic infrastructure and a DAS interrogation unit during the winter 2020/2021. By training a Bayesian Gaussian Mixture Model, we automatically characterize and classify avalanche-induced ground deformations using physical properties extracted from the frequency-wavenumber and frequency-velocity domain of the DAS recordings. The resulting model can estimate the probability of avalanches in the DAS data and is able to differentiate between the avalanche-generated seismic near-field, the seismo-acoustic far-field, and the mass movement propagating on top of the fiber. By analyzing the mass-movement propagation signals, we are able to identify group velocity packages within an avalanche that propagate faster than the phase velocity of the avalanche front, indicating complex internal structures. Importantly, we show that the seismo-acoustic far-field can be detected before the avalanche reaches the fiber-optic array, highlighting DAS as a potential research and early warning tool for hazardous mass movements.

Plain Language Summary Avalanches and other hazardous mass movements pose a danger to the population and critical infrastructure in alpine areas. Therefore, it is important to be able to reliably measure and detect these hazardous events. We show a successful example to measure and characterize avalanches recorded with a Distributed Acoustic Sensing device that measures deformation along a fiber optic cable. We apply unsupervised machine learning to our avalanche recordings and are able to identify consistent properties between 12 avalanches. Ultimately, our results indicate that DAS might be a useful tool for detecting hazardous mass movements.

1. Introduction

1.1. Motivation

Practically all mountainous regions worldwide are subject to some forms of rock falls, snow/rock/ice avalanches, debris flows and sediment-transporting floods. These rapid mass movements pose a significant hazard to both the population and infrastructure, with billions of dollars in financial damage and thousands of fatalities each year (Dilley, 2005; Emberson et al., 2020; Froude & Petley, 2018; Petley, 2012). According to the 2021 Intergovernmental Panel on Climate Change report, the "*magnitude of debris flows might increase* […] *and the debris-flow season may last longer in a warmer climate*" (Zhongming et al., 2021). This suggests that global warming will exacerbate the hazard potential of debris flows and various types of related mass movements. To early detect destructive events and mitigate their impact, extensive, reliable, and high-resolution monitoring and warning solutions are crucial. Seismic and acoustic instruments are increasingly popular for mass movement monitoring, since they record signatures of hazardous events even kilometers away from their occurrence without the need for a direct line of sight between source and sensor (Allstadt et al., 2018; Marchetti et al., 2019). The combination of unrivaled temporal resolution of seismic records and wide spatial sensitivity is a pivotal advantage over in situ measurements (Arattano & Marchi, 2008) and remote sensing approaches like radar technology (Leinss et al., 2020).

Investigation: Patrick Paitz, Nadja Lindner, Pascal Edme, Betty Sovilla, Fabian Walter, Andreas Fichtner
Methodology: Patrick Paitz, Pascal Edme, Betty Sovilla, Fabian Walter, Andreas Fichtner
Project Administration: Patrick Paitz, Pierre Huguenin, Betty Sovilla, Fabian Walter, Andreas Fichtner
Resources: Patrick Paitz, Pierre Huguenin, Michael Hohl, Betty Sovilla, Fabian Walter, Andreas Fichtner
Software: Patrick Paitz
Supervision: Patrick Paitz, Pascal Edme, Betty Sovilla, Fabian Walter, Andreas Fichtner
Validation: Patrick Paitz, Pascal Edme, Betty Sovilla, Fabian Walter, Andreas Fichtner
Visualization: Patrick Paitz, Pascal Edme, Pierre Huguenin
Writing – original draft: Patrick Paitz, Nadja Lindner, Betty Sovilla, Fabian Walter, Andreas Fichtner
Writing – review & editing: Patrick Paitz, Nadja Lindner, Pascal Edme, Betty Sovilla, Fabian Walter, Andreas Fichtner

Implementation of seismic measurements into operational mass movement warning has to fulfill two requirements. First, the seismic sensors have to be placed in various locations to maximize coverage of failure-prone terrain. Second, the detection algorithms have to handle real-time data streams and reliably recognize significant telltale signals in the presence of environmental and anthropogenic noise.

Seismic instrumentation has in recent years undergone rapid developments toward more portable sensors (Leinss et al., 2020). However, even in densely instrumented countries like Switzerland, sensor coverage is still insufficient to encompass significant amounts of unstable slopes. In view of snow avalanches, in particular, the typically harsh terrain in avalanche-prone regions tends to limit the spatial coverage and temporal resolution of most measurements, and of seismic arrays in particular (Pérez- Guillén et al., 2016). In the past, simple avalanche classification models have been created based on the frequency content of conventional seismometers (Suriñach et al., 2001, 2005, 2020). These models, however, do not incorporate the spatiotemporal properties of the avalanches and hence disregard an important aspect of the complex dynamic processes of hazardous mass movements.

In addition to instrumentational developments, recent improvements in machine learning algorithms show great promise for the automatic recognition of emergent and complicated mass movement seismograms (Chmiel et al., 2021). Yet further improvements are necessary to recognize events at sites where little or no training data are available and to identify signal characteristics, which reveal scientific insights into the dynamic characteristics of mass movements.

Here, we address these open challenges with the applications of Distributed Acoustic Sensing (DAS) for snow avalanche detection and characterization. DAS is a technique to measure strain or strain rate along a fiber-optic cable with submeter resolution and mHz to kHz frequency bandwidth and a well-studied instrument response— confirming that DAS measurements are directly related to true ground motion (Lindsey et al., 2020; Paitz et al., 2021). Unused fiber-optic infrastructure initially installed for communication purposes can thus be turned into countless seismic sensors increasing spatial coverage of seismic measurements. We leverage the dense seismic sensing of DAS with unsupervised algorithms to automatically recognize snow avalanches and their internal properties, offering new perspectives for monitoring and alarm systems.

1.2. Fiber-Optic Sensing in a Natural Hazard Context

The introduction of distributed fiber-optic sensing systems to geophysics marks a milestone. By turning a fiber-optic cable into a high resolution seismic measurement network, fiber-optic sensing technologies have opened up new possibilities in exploration geophysics and passive seismology (Lindsey & Martin, 2021; Zhan, 2019), especially in difficult terrain like glaciers and volcanoes (Klaasen et al., 2021; Walter et al., 2020) or on the bottom of the ocean (Lindsey et al., 2019; Williams et al., 2019). For more background information on fiber-optic sensing, the reader is referred to Hartog (2017).

In the context of natural hazards, Brillouin-based distributed fiber-optic sensing systems (BOTDA) have been utilized for landslide and deformation monitoring (Iten et al., 2008; Minardo et al., 2018), and coherent optical time-domain reflectometry was successfully used for ground motion and deformation measurements on landslides (Yu et al., 2018). The suitability of OTDR systems like the Silixa iDAS(TM) DAS unit (used in this study) for the recording of acoustic emission precursors in soil in a laboratory setting was also already established several years ago (Michlmayr et al., 2017). The study by Walter et al. (2020) used a similar DAS system to successfully monitor rockfalls and icequakes on a glacier. Early studies by Prokop et al. (2014) explored avalanche monitoring with fiber-optic sensing systems for avalanche detection and runout distance monitoring. Since the frequency response of the Silixa iDAS system is well studied and known (Lindsey et al., 2020; Paitz et al., 2021) for the frequency ranges expected for avalanche signals between 0.1 and 50 Hz (Pérez- Guillén et al., 2016), we believe it is suited for applications in avalanche dynamics and early warning applications. In an initial study, we showed that a simple frequency-based characterization of avalanches as discussed in Pérez-Guillén et al. (2016) can be extended to single and multichannel DAS applications (Fichtner et al., 2021; Lindner et al., 2021). Since mass-movements are 3D phenomena, this simple classification based on single measurement points does not catch the full spatiotemporal characteristics of an avalanche. Therefore, measurements with high spatial resolution, such as DAS, pose a great potential in avalanche dynamics research, since they can capture avalanche-induced ground motion as a function of both space and time.
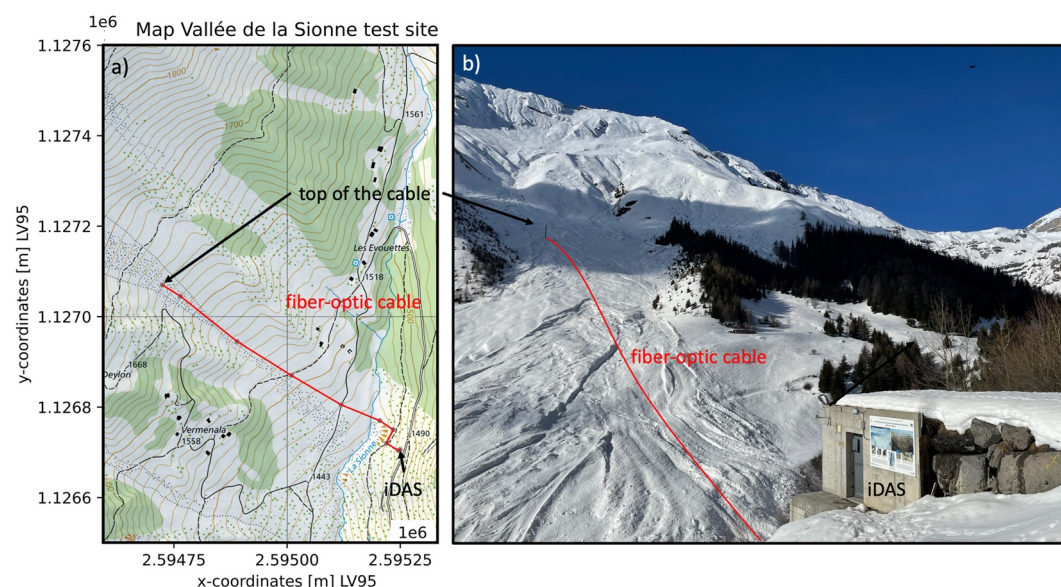
Figure 1. Test site overview. (a) Map view of the Vallée de la Sionne avalanche test site. Map from SwissTopo (2021). (b) Photograph of the test site area with the approximate fiber position indicated in red. The Distributed Acoustic Sensing interrogation unit was located in the bunker (building in the lower right corner).

2. Experiment Setup and Recorded Avalanches

We utilized a Silixa iDAS$^{(TM)}$ v2.4 interrogation unit on an existing fiber-optic cable at the Vallée de la Sionne avalanche test site in Switzerland from October 2020 to March 2021. A map and a photograph of the test site are shown in Figure 1. The test site has been operated by the WSL Institute for Snow and Avalanche Research SLF for over 20 years (Ammann, 1999). Several sensor points within the avalanche paths and runout zones feature seismic, pressure, and radar sensors and are connected via fiber-optic cables to a bunker at the valley bottom where data are stored and processed. The length of the interrogated fiber was around 800 m, and the interrogator was located in the bunker, positioned at the bottom of the path. The fiber crosses the La Sionne creek around 700 m from the topmost monitoring point. Over the entire array, the (single-mode) fiber is installed in a conduit that was excavated to a depth of less than a meter during the construction of the test site. This protects the fiber against avalanche damages. The sampling rate of the interrogator was 1 kHz at a spatial sampling of 2 m. Measurements of conventional seismic sensors in the vicinity of avalanches have been studied at this test site in the past. In an initial investigation by Lindner et al. (2021) and Fichtner et al. (2021), we found a good correspondence of single-channel DAS data with close by seismometers and showed that it is possible to characterize avalanches based on the cumulative energy within different frequency bands according to the classification of Pérez- Guillén et al. (2016). Since this simplified classification does not encode spatiotemporal information, we decided to compare the DAS data to the GEODAR data as opposed to single-station seismometers. The GEODAR is a pulsed doppler radar that measures reflected radar intensity relative to a background signal with a spatial resolution of 0.75 m and a temporal resolution of 111 Hz (Köhler, McElwaine, & Sovilla, 2018). The GEODAR recordings have been cropped to fit the same spatial extent as the DAS data. Whereas the GEODAR data correspond to snow-related phenomena with an areal coverage (particles larger than 5 cm, related to the dense layer and snow clustering), the DAS data correspond to subsurface strain along the fiber. Since the propagation of the avalanche, however, effects both GEODAR and DAS data to a similar extent, we are using the former as a ground truth confirmation for the occurrence and spatiotemporal distribution of an avalanche. It, however, must be mentioned that the observed physical quantities and measurement characteristics are different. It also must be mentioned that the simple fiber-layout in our experiment only measures strain-rate along a single line, projecting the complex 3D phenomenon of each avalanche onto the fiber. The geometry was dictated by the existing installation, and in the future, a more complex fiber-layout could be envisioned.

At the experimental site of the Vallee de la Sionne, the avalanche activity is monitored continuously throughout the winter season, thanks to seismic sensors located along the avalanche path recording the ground motion

generated by snow avalanche and other sources of vibrations, and serving as a trigger for all measuring devices situated in the path and inside the bunker, including the GEODAR, a Doppler radar and a camera. The trigger system includes two seismic stations with three-component seismometers (Ammann, 1999; Pérez- Guillén et al., 2016) and the data acquisition is triggered when the seismic signal exceeds a threshold. By comparing seismic triggers with GEODAR and Doppler radar data, as well as images collected by an automatic camera, it is possible to discern between the real avalanche signal and false alarms and thus build an extremely accurate and comprehensive catalog of avalanche events. This catalog, which to date contains more than 1,000 avalanches, is maintained by the SLF but is not officially published. During the data acquisition period, more than 20 events have been registered in the SLF database and were recorded with our DAS system. We cover examples of the entire spectrum of avalanches recorded throughout the measurement period. In this manuscript, we first discuss the key characteristics for one example (denoted "avalanche 3023"). This one was selected since it contains key features while still being not too complex. For additional events, the data are visualized in the Supporting Information S1. The recorded data of the observed avalanches cover large transitional powder snow avalanches with partial flow regime transitions and depositional regimes (avalanches 3009, 3022, and 3023) and smaller avalanches without a clearly distinguishable transition (avalanches 3016, 3020, and 3021). A summary of the measured avalanches is given in Table B1 in the appendix.

3. Seismogenesis of Mass Movements

Most avalanches considered in this study propagated on top of the fiber-optic cable. Therefore, the data contain a superposition of near-field observations of seismo-acoustic sources, as well as ground vibrations from sources that can potentially be further away. Note that we prefer the term "seismo-acoustic" over "seismic" as seismic records of avalanches may contain signatures of waves traveling through the air (Heck, Hobiger, et al., 2019). For particulate gravity currents like snow avalanches, various theoretical models for seismogenesis have been recently proposed, which form the basis for the following discussion. Nevertheless, air waves such as infrasound may also contain important information about avalanche volumes and dynamics and could influence our DAS records (Allstadt et al., 2018; Marchetti et al., 2021).

For multiphase flows (granular flow in dense flow regimes and turbulent flow in the aerial components) such as snow avalanches, different seismic source mechanisms have been considered in the past: (a) (quasi-)static deformation as a response to instantaneous weight and frictional shear forces (Wenner et al., 2022) and (b) (snow) particle-ground impacts (Tsai et al., 2012), (c) turbulent flow (Gimbert et al., 2014) and, by (d) abrupt stopping of mass movement due to friction (Tregaskis et al., 2022), and by (e) changes in traction due to mass deposition and erosion (Edwards & Gray, 2015). A schematic avalanche propagating over our fiber array is shown in Figure A1 in the appendix.

3.1. DAS Data

3.1.1. Time-Distance Domain

The raw data of avalanche 3023 are visualized in Figure 2. The total duration of the avalanche propagating over the array is about 2 minutes. The ground truth of the avalanche is confirmed by measurements of the GEODAR system (Keylock et al., 2014; Köhler, McElwaine, & Sovilla, 2018). The extent of the avalanche as measured by the GEODAR is highlighted in transparent blue colors in Figure 2. Different parts of the avalanche can be distinguished (where the numbers correspond to the features highlighted in the figure): (a) The earlier, faster part of the avalanche between 2.5 and 3 min time, and (b) a slower, later part of the avalanche between 3 and 4.5 min. What is also visible is that (5) there are signals arriving over the entire array before the avalanche front moves on top of the cable (at times before 2.75 min). Other observable features include (4) noisier channels (e.g., at 310 m), and (c) high amplitude and velocity events (nearly horizontal in the plot), spanning about 100 m in distance each. These events could be interpreted as (abrupt) stopping events, where parts of the avalanche abruptly decelerate, hence creating a high amplitude strain-change in the subsurface. Other potential explanations for these events include snowpack collapses. However, the fact that the strain-rate amplitudes of the mass-movement are significantly lower after these events makes the abrupt stop hypothesis more likely. It must be noted that for distances above 200 m along the fiber, there is no clear DAS signal visible in the raw and low-frequency data, whereas the GEODAR outline still records an avalanche signal there. A possible reason for that is the difference between the
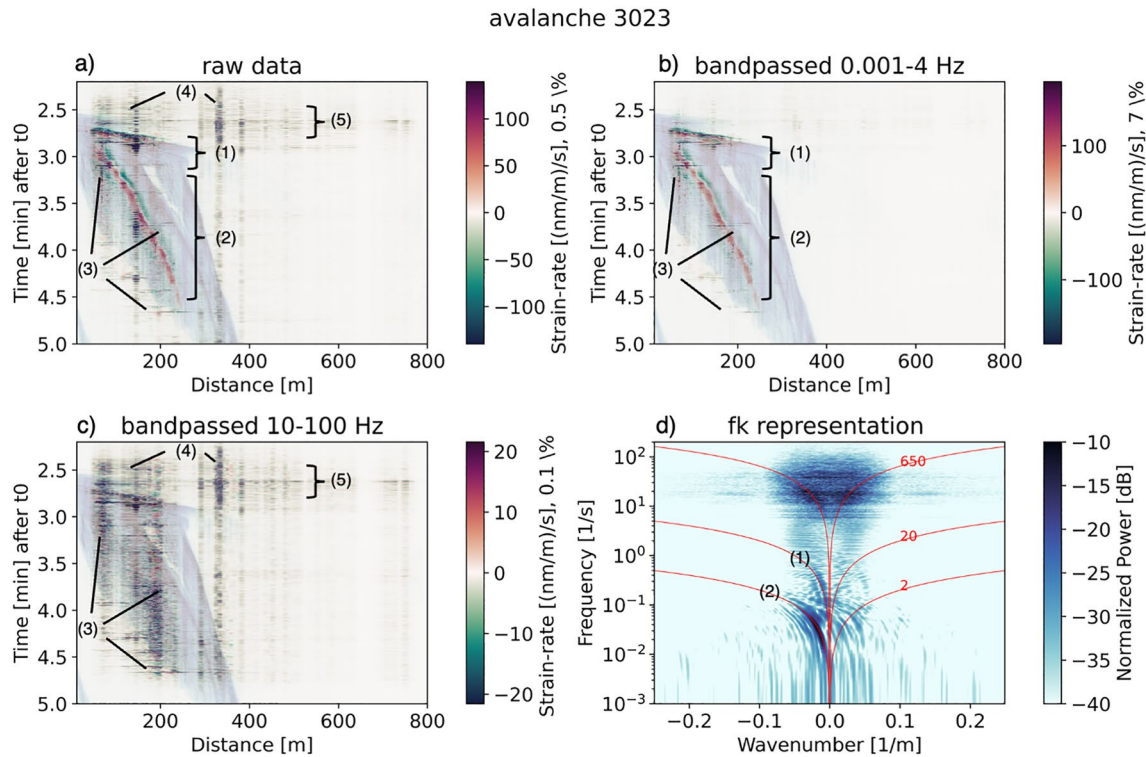
Figure 2. (a) Raw strain rate data of avalanche 3023. The amplitudes are clipped at 0.5% of the global maximum within the visualized window. (b) Band-pass filtered low-frequency strain rate data of avalanche 3023 (fourth order Butterworth band-pass filter from 0.001 to 4 Hz). The amplitudes are clipped at 7% of the global maximum within the visualized window. (c) Band-passed high-frequency strain rate data of avalanche 3023 (fourth-order Butterworth band-pass filter from 10 to 100 Hz). The amplitudes are clipped at 0.1% of the global maximum within the visualized window. For (a to c), the ground truth of the extent of the avalanche is highlighted in gray-blue shading of the measurements of the GEODAR system. (d) Frequency-wavenumber (*fk*) representation of the raw data from avalanche 3023. The red lines indicate apparent phase velocities along the array in m/s. Negative wavenumbers indicate energy propagating from the top to the bottom of the slope (downward), and positive wavenumbers indicate coherent energy propagating upward. The distance starts at the Northwest end of the cable (uphill) at 0 m and ranges down to the Southeast end of the cable in the bunker, where the interrogator was located (Figure 1).

two observed quantities. Whereas the GEODAR system measures the reflectivity of snow-related phenomena, the DAS data see the ground deformation due to seismic waves and gravitational mass-movements at the interface between the atmosphere and the ground. This would mean that for the given example, the GEODAR signal for distances >200 m along the fiber corresponds to predominant effects of low seismic energy and is hence not caught by the DAS system, whereas for shorter distances from the top, effects observed by DAS and GEODAR are directly coupled. The low-frequency time-offset strain rate data are also visualized in Panel b) of Figure 2. The two main parts of the avalanche (1) and (2), as well as the stopping mechanisms (3) are still visible, and the data look less noisy compared to the raw data in Figure 2a).

3.1.2. Frequency-Wavenumber Domain

Since the fiber is approximately straight with equal channel spacing over the entire array, it is straightforward to analyze the data in the frequency-wavenumber domain. The frequency-wavenumber representation of the data is defined as a 2-D Fourier Transform (over time and over space):

$$\dot{\epsilon}(f, k) = \int_x \int_t \dot{\epsilon}(t, x) \, e^{-2\pi i f t} \, e^{-2\pi i k x} \, dt \, dx, \tag{1}$$

for the strain rate $\dot{\epsilon}$, time t, distance x, frequency f, and wavenumber k (the inverse of the wavelength). This representation allows for analysis of the frequency and apparent phase velocity content of the data and can reveal dispersive behavior, that is, frequency dependence of the velocity of individual wave modes.

The frequency-wavenumber (*fk*) representation of the raw data of avalanche 3023 is visualized in Figure 2d. From the *fk* representation of the data, a clear separation can be observed between high-frequency seismic (10–30 Hz)

and low-frequency (0.01–10 Hz) signals. The two different parts of the avalanche observed in the raw data are also visible in the *fk* domain (features (1) and (2)) for frequencies below 1 Hz. The *fk* visualization associates these low-frequency (<1 Hz) downward propagating signals with phase velocities of between 5 and 20 m/s (1), and 2 m/s (2). In addition, high-velocity events are visible for frequencies between 5 and 30 Hz, propagating both upward and downward at speeds of about 650 m/s. Such apparent velocities and the omnidirectional propagation suggest that these events are seismic waves generated by the avalanche.

4. Signal Classification

In order to automatically identify and distinguish between the signals shown in Figure 2, we propose the use of unsupervised machine learning algorithms. In the past, unsupervised algorithms have proven useful in a geophysical context to extract subsets of signals with similar properties from large data sets (Grimm, 2021; Grimm & Poli, 2022; Martin et al., 2018). In a cryoseismic context, Grimm (2021) extracted different physical classes from continuous DAS recordings on a glacier, characterizing crevassing events, stick-slip icequakes and background noise in an automated way. Similarly, Grimm and Poli (2022) used spatial coherency features of an urban DAS data set from Grenoble (France) to identify spatiotemporally repeating events. Martin et al. (2018) utilized signal features from data segments that had been transformed using the continuous wavelet transform and minibatch-optimized K-means to find classes of coherent properties of the seismic wavefield from DAS data. Here, we propose to use unsupervised clustering to identify characteristic properties of avalanche recordings with DAS. Since the dimensionality of the DAS data is too high to perform clustering on raw data, we first extract representative features. This is a common first step in applied machine learning workflows, and the features have to be chosen problem-dependent (Alpaydin, 2020).

4.1. Feature Extraction

Automatic and (near) real-time processing for warning applications require signal feature extraction within small time- and space windows. These windows contain either avalanche signals and/or background noise, which includes natural (e.g., earthquake) and anthropogenic (e.g., road and air traffic) signals. We set the window sizes to 5 s in time and 50 m in space (with adjacent windows overlapping by 3 s and 30 m). The window size and overlap were chosen empirically such that coherent signals in time and space are detected, while small-scale changes are still captured. The chosen features introduced below aim to encode the apparent phase velocities along the fiber, frequency content, dominant propagation direction, and total strain rate energy within each window, and hence describe physical properties that are potentially important for avalanche characterization and discrimination from other signals like earthquakes.

An overview of the proposed features and their representation is visualized in Figure 3. In the first step, the raw data are windowed and transformed to the *fk* domain (see Equation 1). In the *fk* domain, the contents of the amplitude spectrum of velocity-frequency bins are analyzed, resulting in cumulative *fk* amplitudes $\dot{A}$ for each bin, where the dot indicates that the cumulative *fk* amplitudes are associated with a transformation of strain rate $\dot{\epsilon}$ rather than strain.

$$\dot{A}(v_1, v_2, f_1, f_2) = \sum_{f,k} |\dot{\epsilon}(f, k)|; \ \forall v = \frac{f}{k} \in \{v_1, v_2\} \cap f \in \{f_1, f_2\}, \tag{2}$$

where the local phase velocity *v* can be described in terms of frequency *f* and wavenumber *k*, following $v = f/k$. The frequency and velocity bins range from f_1 to f_2 and from v_1 to v_2, respectively. The above equation maps our data from the *fk* domain into a discrete velocity-frequency (*vf*) domain. In the next step, we find *n* numbers of local maxima in this domain and extract the corresponding frequency and apparent velocity of these maxima, for both positive and negative wavenumbers. This results in *n* local maxima for positive wavenumbers and negative wavenumbers:

$$M_n^+(v_{n^+}, f_{n^+}) = \max(\dot{A}(v, f)); \ \forall n^+ \in \{1 \dots n\} \cap k > 0 \tag{3}$$

$$M_n^-(v_{n^-}, f_{n^-}) = \max(\dot{A}(v, f)); \ \forall n^- \in \{1 \dots n\} \cap k < 0 \tag{4}$$

For each of these peaks $M_n(v_n, f_n)$, we compute the ratio $R_n(v_n, f_n)$ between the cumulative *fk* amplitudes $\dot{A}$ of positive and negative wavenumber, giving us an indication of a preferable directionality in propagation:

$$R_n^+(v_{n^+}, f_{n^+}) = \frac{\dot{A}(v_{n^+}, f_{n^+})}{\dot{A}(-v_{n^+}, f_{n^+})}; \ \forall n^+ \in \{1 \dots n\} \cap k > 0 \tag{5}$$
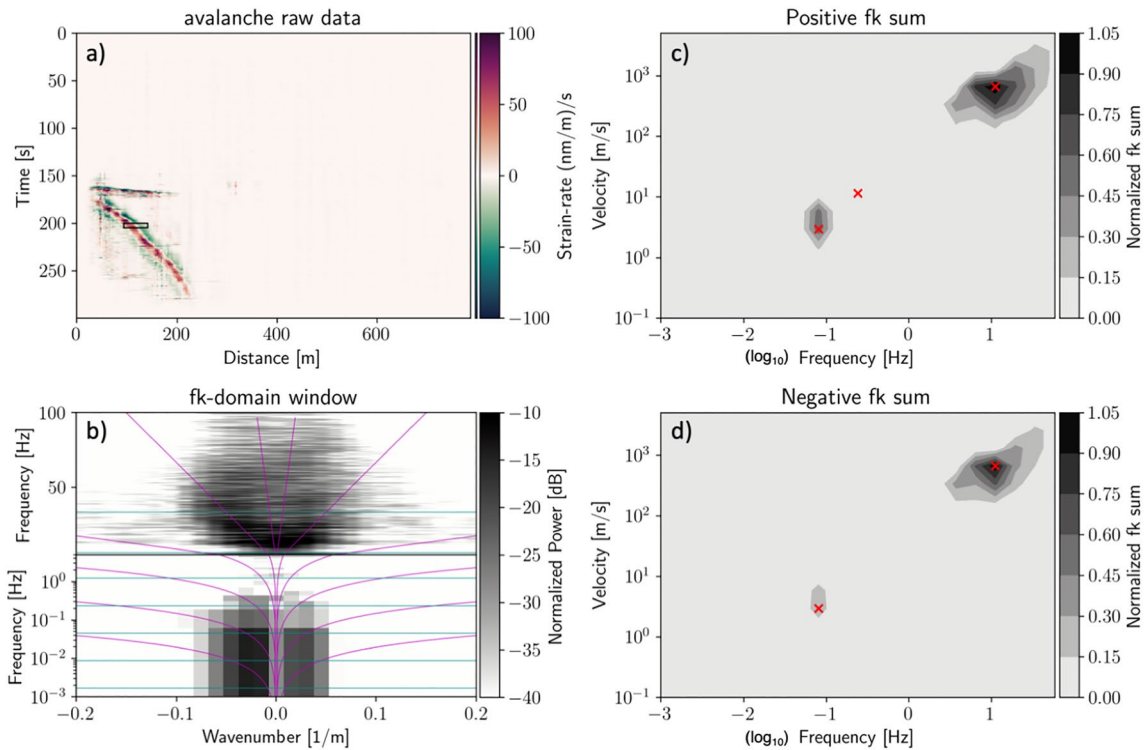
Figure 3. Example of the windowed algorithm proposed to extract signal properties used to detect and characterize avalanche data. (a) In a first step, the data are windowed in both time and space (indicated by the black rectangle). In our example, the length of the window is 50 m and 5 s. (b) The windowed data are then tapered (in both time and space) and transformed to the *fk* domain. The *fk* domain is then binned into frequency ranges (indicated by the horizontal blue lines) and velocity ranges (indicated by the magenta diagonal curves). For each bin, the absolute values of the *fk* amplitudes are then summed and normalized over the number of samples per bin. The *fk* domain amplitudes are displayed in dB relative to the maximum *fk* amplitudes. (c and d) The output of the bottom left is then arranged by frequency and velocity, and local maxima are extracted. This results in a number of *n* (in this example *n* = 3) extracted peaks, with corresponding frequency and velocity ranges, as well as summed *fk* amplitude values. The process is done for positive wavenumber values and for negative wavenumber values. For each peak, the ratio of positive to negative *fk* amplitude is also stored as a feature. For the bottom panel of (b) and the panels (c and d), the *y*-axes are in log-scale. The *x*-axes in panels (c and d) are also in log-scale.

$$R_n^-(v_{n^-}, f_{n^-}) = \frac{\dot{A}(-v_{n^-}, f_{n^-})}{\dot{A}(v_{n^-}, f_{n^-})}; \ \forall n^- \in \{1 \ldots n\} \cap k < 0 \tag{6}$$

In addition to the values of the peaks $M_n^{+|-}(v_n, f_n)$ in the *vf* domain, we extract information on the summed amplitude spectrum *S*, defined as follows:

$$S^+ = \sum_{f,k} |\dot{\epsilon}(f, k)|; \ \forall k > 0 \tag{7}$$

$$S^- = \sum_{f,k} |\dot{\epsilon}(f, k)|; \ \forall k < 0 \tag{8}$$

The last feature we extract is the ratio of cumulative amplitudes between positive and negative wavenumbers *C* within the specific time-space window:

$$C = \frac{S^+}{S^-}. \tag{9}$$

For each window, a total of 11 features are extracted per picked peak *n* as summarized in Table 1. This reduces the dimensionality for each window from 62,500 (2,500 samples for 25 channels) to 11 · *n*, which is <0.1% of the dimensionality of the raw data. In our case, *n* was chosen to be three in order to capture first, second, and third order effects within each window.

Table 1
Features Extracted in Each Sub-Window of the Distributed Acoustic Sensing Record That Are Used for the Characterization of the Data

Feature	Description
$M^+(v_{n^+}, f_{n^+})$	Local maxima in the vf domain for positive k
$M^-(v_{n^-}, f_{n^-})$	Local maxima in the vf domain for negative k
v_{n^+}, v_{n^-}	Apparent phase velocity of the local maximum M for pos. (+) and neg. (−) k
f_{n^+}, f_{n^-}	Frequency of the local maximum M for pos. (+) and neg. (−) k
$R^+(v_{n^+}, f_{n^+})$	Ratio between $M^+(v_{n^+}, f_{n^+})$ and corresponding vf amplitudes of neg. K
$R^-(v_{n^-}, f_{n^-})$	Ratio between $M^-(v_{n^-}, f_{n^-})$ and corresponding vf amplitudes of pos. K
S^+	Cumulative amplitude spectrum in the fk domain for positive k
S^-	Cumulative amplitude spectrum in the fk domain for negative k
C	Ratio between S^+ and S^-

4.2. Bayesian Gaussian Mixture Models: Unsupervised Clustering

The features in Table 1 are used in an unsupervised clustering algorithm to identify groups of data with similar properties. We chose a Bayesian Gaussian mixture model after evaluating the performance of various clustering algorithms including K-means, minibatch K-means, and affinity propagation (Bishop & Nasrabadi, 2006; Press et al., 2007).

Bayesian Gaussian mixture models are probabilistic models that fit training data onto a finite number of Gaussian distributions, utilizing the Expectation-Maximization algorithm (see Bishop and Nasrabadi (2006); Press et al. (2007) for details). This way, each feature set of the training data is associated to a cluster or class (one of the Gaussian distributions). Compared to other conventional clustering algorithms like K-Means, Gaussian mixture models have the advantage that a trained model can predict a class from data that it was not trained with. Furthermore, instead of associating one class to a set of features (hard assignment of data points to a cluster), the probability of the data for being a part of each subclass is estimated (soft assignment of data points to a class from the posterior probabilities for each class) (Bishop & Nasrabadi, 2006), which in our case becomes important where low-frequency low-velocity mass movements are within the same space time window as high-frequency seismic waves. We use the python package scikit-learn (Pedregosa et al., 2011) for the implementation of Bayesian Gaussian mixture models. Based on trial and error, we decided to set the number of classes to 10, initializing the weights, means, and covariances with the K-means algorithm, and using a Dirichlet process for the weight concentration prior. Increasing the number of classes resulted in a higher number of "noise" classes without improving the signal classification. Decreasing the number of classes led to inconsistent clustering results for multiple stages of training.

We train the model with 92,367 windows from 13 potential avalanche candidates, including windows containing environmental and anthropogenic noise, avalanches, and earthquake recordings. In total, 246,118 feature sets are used as input for the training (where for some of the windows, less than 3 feature sets could be extracted). After training, the model can be used to estimate the probability of each class for all recorded time windows.

4.3. Clustering Results

The results of the Bayesian Gaussian mixture model analysis are visualized in Figure 4 for avalanche 3023. They show the predicted classes with a probability higher than 0.3 for each window based on the trained model, together with a Gaussian kernel density estimation (Scott, 2015) for each class over both space and time. In addition to the normalized kernel density, their normalized cumulative values are also shown in panels b) and c) of Figure 4. We can observe that the entire avalanche is classified within the same class S2 and that class S1 emerges 30 s before the avalanche propagates over the fiber-optic array. In addition, classes S3 and S4 are also associated with the avalanche, following the time after class S2 emerges. From the comparison of the kernel densities of the clustering result to the normalized cumulative GEODAR intensity, we can see a correlation of classes S2 to S4 with the ground truth of the GEODAR data.
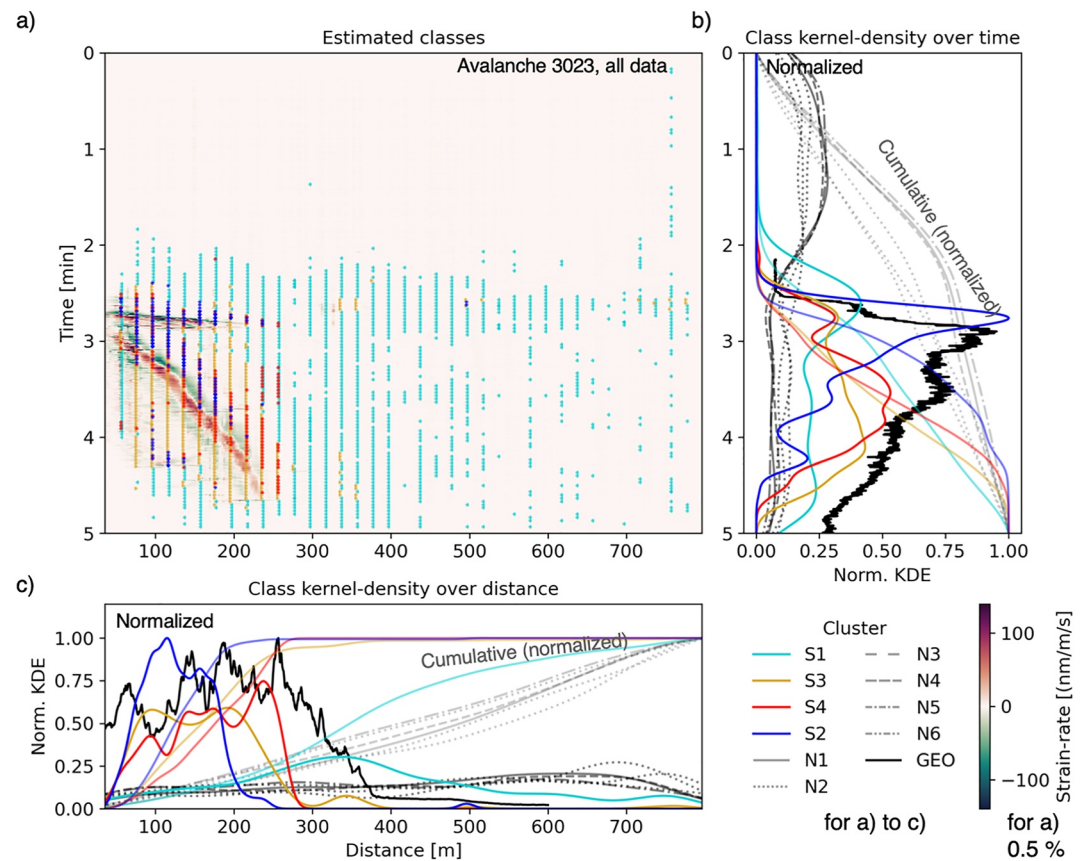
Figure 4. (a) Results of the predicted classes with a probability above 0.3 for each window of the Bayesian Gaussian mixture model clustering for avalanche 3023. The raw strain-rate data are plotted in the background (see colorbar at the bottom right, clipped to 0.5% of the total maximum amplitude), whereas the predicted classes are color-coded for each class. Multiple classes for any given time-window are possible and highlighted with different markers and colors. Six different "noise" classes could be identified which are not visualized here (N1 to N6). These classes most likely capture environmental and anthropogenic noise, as well as self-noise of the instrumentation. The signal classes of the Distributed Acoustic Sensing system are labeled S1 to S4. (b) Estimated probability density from a Gaussian kernel density over the samples within each class over time. The black line indicates the normalized GEODAR amplitudes for the given window denoted GEO in the labels. For a smoother display, only every 15th GEODAR datapoint is plotted in panel (b). (c) Estimated probability density from a Gaussian kernel over the samples within each class over space. The black line indicates the normalized cumulative GEODAR amplitudes for the given window. The transparent lines in (b and c) in the background give the normalized cumulative Gaussian kernel density corresponding to the normalized kernel densities. The corresponding figures for the other events mentioned in this manuscript are available in the Electronic Supplement to this manuscript.

5. Discussion: Physical Interpretation of the Classes

Because the only available ground-truth data with spatial extent in the form of GEODAR data to highlight the physical extent of the avalanches, it is difficult to verify the hypotheses presented in this chapter. The provided physical interpretation given in this section should be seen as a highly speculative first attempt to illuminate the DAS avalanche recordings. Nevertheless, our interpretations are backed up by physical evidence from the properties of the different classes.

Since Bayesian Gaussian mixture models give us a probability distribution of any feature set being part of one specific class, we can predict the classes of our data to extract apparent phase velocity ranges, frequency content, and dominant propagation direction of the measured signals. The clustering results for avalanche 3023 as well as all investigated avalanches are visualized in Figure 5 together with the mean feature values and standard deviations. The results are consistent over all events and we show details of the space-time dependent classification of the DAS signals for Avalanche 3023 in Figure 4.
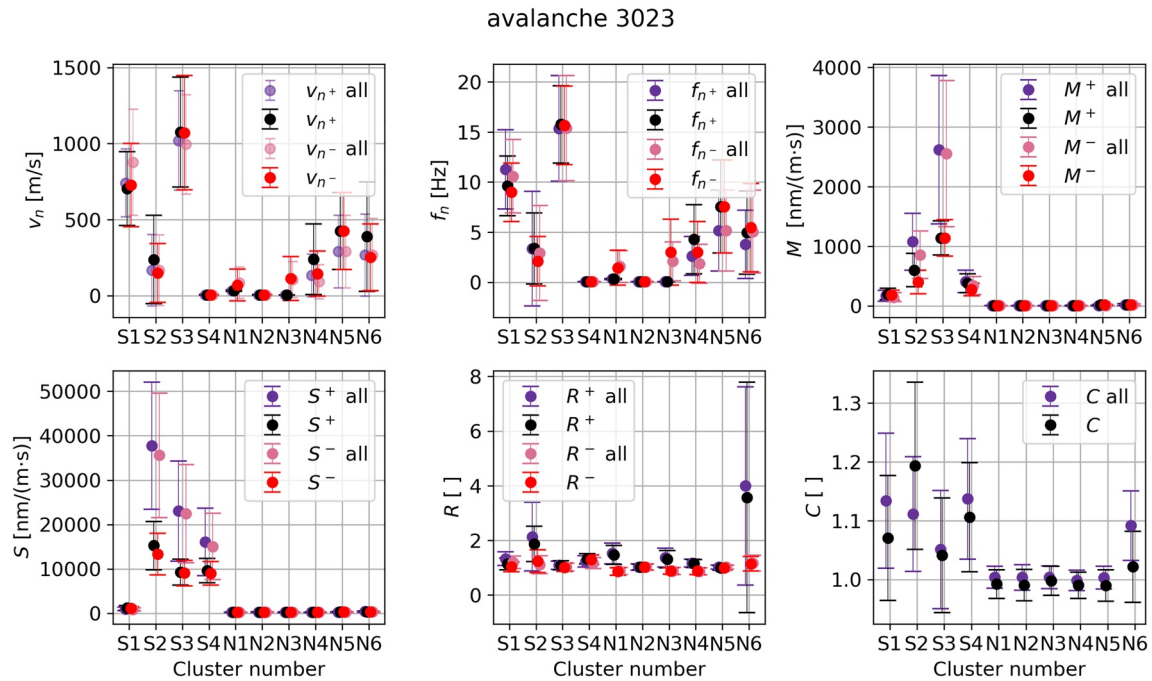
Figure 5. Mean feature values (dots) and $\frac{1}{2}$ standard deviation (error bars) for each clustered feature set. Black color indicates positive wavenumber values (downward propagating energy, superscript $^+$) and red indicates negative wavenumbers (upward propagating energy, superscript $^-$) along the fiber for avalanche 3023. The purple and light red colors indicate the features for the entire training data set, including all events that were used in the clustering process for positive and negative wavenumbers, respectively.

Figure 5 shows that whereas the classes denoted as noise exhibit consistently low values in fk amplitudes (features M and S), they can be distinguished from each other in terms of apparent phase velocities and frequency ranges. Classes N1 and N2 consist of windows with dominant frequencies of ≤ 0.5 Hz. The noise classes N5 and N6 contain mainly frequencies above 5 Hz at apparent phase velocities higher than the ones from class N1 to N3 (>350 m/s).

Class S3 is associated with the highest overall fk amplitudes at a mean frequency of around 16 Hz and a mean velocity of 1075 m/s, propagating both uphill and downhill along the cable (mean $C \approx 1.05$). This class can be interpreted as the seismic near-field, as this class is dominant on channels during the time the avalanche is on top of the cable. These values are realistic for seismic waves in sediment layers (Boore & Joyner, 1997; Michel et al., 2014), especially considering that near-field signals consist of multiple seismic phases including P-waves. These are faster than Rayleigh phases, which are the dominant far-field response to vertical particle-ground impacts (Sánchez-Sesma et al., 2011).

Class S2 has a significantly lower mean frequency (3.4 Hz) and apparent velocity (238 m/s). This class has the highest mean C value (1.2), meaning that the dominant propagation direction of these signals is downhill. It could be interpreted as the mass-movement class. The apparent velocities are exceptionally low but could be explained by slow Biot's waves propagating through the pore space of snow within the avalanche and/or the underlying substrate (Capelli et al., 2016).

Based on the apparent velocity content (mean 703 m/s) of class S1, we interpret it as the seismo-acoustic far-field generated by the avalanches. This apparent velocity is reasonable for surface and S-waves in generic rock sites (Boore & Joyner, 1997). The frequency content (mean 9.6 Hz) of this class is comparable to the frequency range of seismic waves generated by avalanches observed in the literature (Pérez- Guillén et al., 2016; Van Herwijnen & Schweizer, 2011). The probability of class S1 increases already 30 s before the avalanche arrives at the cable. With the above mean frequency, mean velocity and typical avalanche speeds of several meters per second, this likely corresponds to a time when the avalanche is several wavelengths away from the cable, which supports our hypothesis of seismic far-field waves.

Class S4 contains frequencies below 1 Hz, which we interpret as quasi-static ground deformation in response to the instantaneous avalanche weight and frictional shear forces. Waves of such quasi-static ground deformation
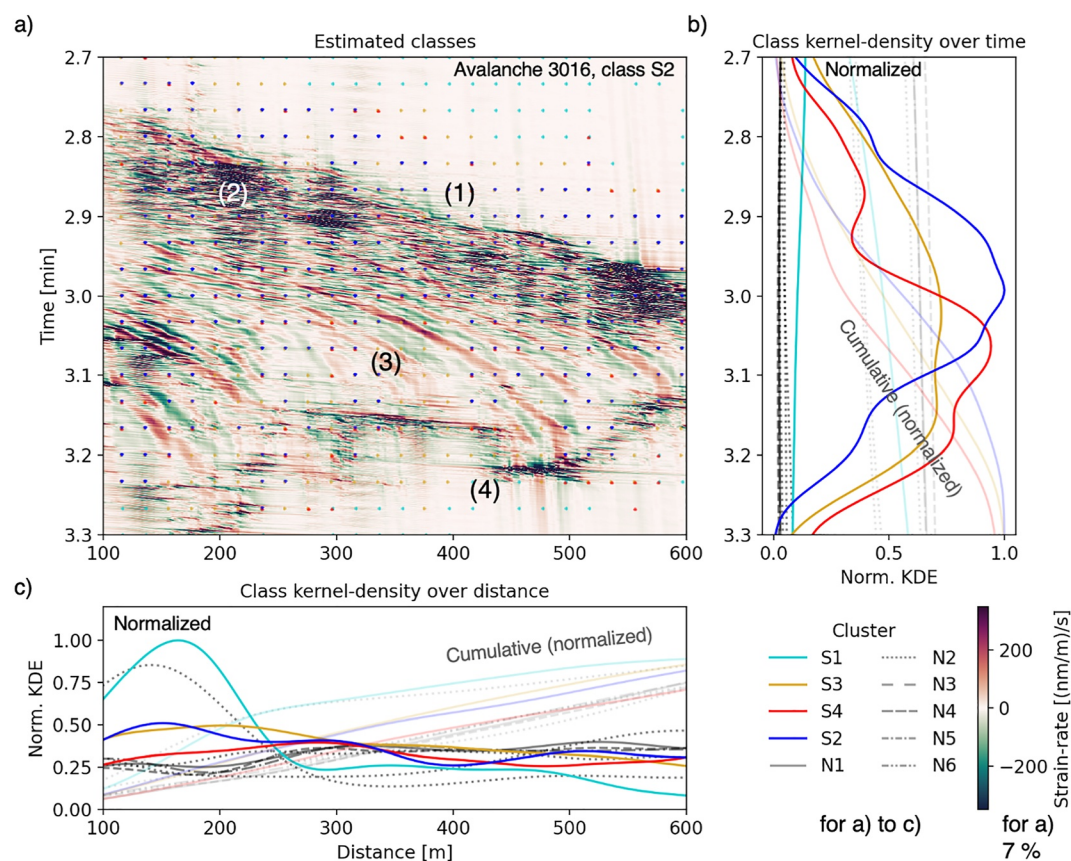
Figure 6. (a) Distributed Acoustic Sensing strain-rate data of avalanche 3016 in the frequency range from 0.001 to 5 Hz for apparent velocities between 1 and 250 m/s (corresponding to the mass-movement class S2, see colorbar at the bottom right, clipped to 7% of the total maximum amplitude). We can observe that the internal structure of the avalanche is more complex than for avalanche 3023. The group velocity of the avalanche front (1) is approximately 40 m/s, whereas the phase velocity inside the avalanche (2) is around 160 m/s. The later part of the avalanche lacks these high-velocity arrivals and instead consists of events that decrease from 40 m/s until they stop (3). In the deposit area, other events (4) also seem to propagate at apparent phase velocities of >160 m/s downhill. (b) Estimated probability density from a Gaussian kernel density over the samples within each class over time. (c) Estimated probability density from a Gaussian kernel over the samples within each class over space. The transparent lines in the background of (b and c) give the normalized cumulative Gaussian kernel density corresponding to the normalized kernel densities.

can result from flow depth, velocity, or particle concentration perturbations traveling within the avalanche body (Viroulet et al., 2018) and erosion-deposition mechanisms (Edwards & Gray, 2015). This explains the apparent phase velocities (<6 m/s) associated with avalanche motion rather than seismic energy propagation of classes S1-S3. The motion of the avalanche front and major secondary surges induce wavelets with periods of 10 seconds or longer (Figure 2b, e.g.,), which are not resolvable with our 5 s time window sizes. Nevertheless, class S4 seems to capture the highest frequencies of these signals. C values of larger than 1.1 indicate that S4 signals propagate preferentially downhill. Uphill propagation, however, is also possible and can be explained by shock wave dispersion (Liu & Mei, 1994).

So far we have discussed the avalanche signal in terms of dominant classes. However, this binary (dominant/nondominant) characterization is not always justified since several classes may reach similar probabilities at the same location in space and time. This is particularly apparent for Avalanche 3016, whose DAS signals and kernel densities have a simple appearance since the avalanche consists of only one surge (Figure 6). During times when Avalanche 3016 covers the cable, further signal details are visible and might be interpreted as the avalanche front (a), internal roll waves (b) or erosion-deposition waves (c) and stopping phases (d) (Razis et al., 2014; Tregaskis et al., 2022; Viroulet et al., 2018), traveling at different and variable speeds (different event move-outs in Figure 6). In addition to class S4, classes S2 and S3 can also be expected to characterize these internal dynamic

processes within the avalanche as like class S4, they dominate during times when the avalanche locates above the cable. In fact, for Avalanche 3016, kernel densities of classes S2, S3, and S4 increase and decrease parallel to each other and have comparable peaks (Figure 6 panel b)).

From this figure, we can also observe that the avalanche front propagates at a relatively high apparent group velocity of about 40 m/s (labeled (1)). Internal apparent phase velocities of up to 160 m/s are present in the earlier part of the avalanche (2). We can also observe the transition from class S2 toward S4 (panel b). This may be related to the transition from the erosive and intermittent flow-regime characteristic of the avalanche front toward the depositional flow regime at the avalanche tail (3) after around 3 min. The internal velocities of up to 160 m/s within the avalanche that are higher than the front propagation speed of around 40 m/s suggest that internal phases may be processes related to roll-waves activity taking place at the surface of the denser basal layer (Razis et al., 2014; Viroulet et al., 2018).

The above discussion shows that although different classes can be associated with characteristic ranges of frequencies and propagation velocities, their relative occurrences can be utilized to describe a range of avalanche dynamics. The arrival of the front with its instantaneous increase in local weight will induce low-frequency, quasi-static elastic deformation (class S4). This adds to weight variations resulting from snow entrainment, which for powder snow avalanches can take the form of violent eruptions (Carroll et al., 2012; Sovilla et al., 2006) and other internal phases or flow-depth variations like roll waves (Razis et al., 2014; Viroulet et al., 2018) to produce further low-frequency signals. At the avalanche front, we also expect the turbulent and suspended mass movement to couple into the ground (classes 1 and 3). Although we are not aware of theoretical descriptions of the seismogenesis of turbulent and laminar air-snow mixtures, an equivalent mechanism has been proposed for river flow (Gimbert et al., 2014). This leads to a mixing of signals from classes 1, 3, and 4. Similarly, temperature-dependent sintering produces macroscopic granules (Steinkogler et al., 2015) whose ground impacts generate high-frequency seismic signals (Tsai et al., 2012) constituting classes 1 and 3. It is not clear if these mechanisms also generate the slow seismic phases of class 2, which is a predominant signal when all avalanches override the cable. The existence of a potential Biot phase (Capelli et al., 2016) is possible but not the only explanation. The records of the local M 1.2 earthquake from Sanetschpass about 10 km from our recording site testify to the nonuniqueness of physical class meaning (see event 3036 in the Supporting Information S1, occurring on 03 March 2021, 00:38:14 UTC): The DAS earthquake records lack class S4, which is expected as the earthquake does not generate slowly propagating signals corresponding to the quasi-static elastic ground deformation induced by an avalanche. On the other hand, the earthquake records are predominantly classified into classes S1 and S3. Our interpretation of S3 as near-field seismic signals at frequencies resolvable within 5 s time windows is questionable since the earthquake located 10 km away from the cable. The absence of class S2 suggests that this class is indeed characteristic for mass movements, even though an explanation of its rather slow seismic propagation speeds remains elusive.

The frequency content of our signal classes associated to the avalanches falls within the same range as for existing studies investigating the frequency content of avalanche recordings from conventional seismometers as, for example, in Pérez- Guillén et al. (2016), and validates our automatic classification and interpretation. The cumulative kernel densities introduced can furthermore be seen as spatiotemporal extension of the conventional avalanche characterization based on frequency-content in the existing literature, such as in Pérez- Guillén et al. (2016), Suriñach et al. (2005, 2020).

To summarize, there exist distinct signal classes, which are shared among all of the recorded avalanches (classes S1 to S4). Class S1 is interpreted as the seismo-acoustic far-field that arrives at the cable before the mass movement itself. Classes S2–S4 are associated with low-frequency, quasi-static ground deformation, near-field ground shaking and other yet-to-be-confirmed signals generated as the avalanches override the cable. Although these interpretations may differ for other seismic sources like earthquakes, the signal classification seems to be characteristic for all of our avalanches and could be used for automatic detection.

5.1. Internal Avalanche Characteristics

We identified that class S2 is most likely related to the physical properties of the mass-movement inside the avalanche. We can use this to further analyze the internal structure of the avalanche propagating over the fiber-optic array. Since avalanche 3023 does not have a complex internal avalanche structure, but a well separated

slow and fast part, we will look at a different avalanche. We chose avalanche 3016 for this analysis, since it is a single surge avalanche with a complex internal structure. From the clustering analysis, we know the frequency and apparent phase-velocity ranges of the mass-movement class S2. Hence, we focus on frequencies <5 Hz and apparent phase velocities <250 m/s (only downward propagating). The DAS data of a zoomed-in version of avalanche 3016 within this frequency and velocity range are displayed in Figure 6.

Accordingly, we explain the low apparent phase velocities (<6 m/s) with the avalanche's bulk mass motion, which may vary between avalanche events or even between surges of a single avalanche (Figure 2). Hence, in terms of avalanche processes, the classes S2 and S4 could be associated with different flow regimes within the avalanche—where the earlier class S2 may correspond to the high-energy processes occurring at the front of the avalanche, such as entrainment of snow or impacts of the (turbulent) suspensions with the ground, and class S4 may be related to the signal generated by the following dense basal layer and its deposition (Köhler, McElwaine, & Sovilla, 2018).

Another observation is that the presence of some mass movement classes by themselves require additional analysis. A regional M 1.5 earthquake from Sanetschpass (see event 3036 in the Supporting Information S1) is predominantly classified into classes S1 and S3. The classification of S1 and S3 during the regional earthquake also indicate that additional signals might be included in these classes. The absence of class S2 indicates that there is no mass movement present during the earthquake, but this nonuniqueness of the classification needs to be kept in mind for potential automatic classifications in the future.

6. Perspectives for Snow Avalanche Monitoring

DAS enables the distributed measurement of ground deformation in response to avalanche flow with high temporal and spatial resolution. There exist specific reasons why this could be a game changer for avalanche monitoring and warning applications. Our analysis shows that nonsupervised classification of DAS recordings containing both noise and avalanche signatures are capable of separating the two. Although this method has to be applied to longer (multiple months) DAS records to evaluate its accuracy, the signal classes shared among all recorded avalanches suggests that automatic detections are feasible. The consistent detections of class S1 signals tens of seconds prior to the avalanche arrivals at the cable are particularly encouraging: interrogating preinstalled communication cables seems to be sensitive enough to detect avalanche seismograms remotely. This is important for the more realistic case where fiber-optic infrastructure locates parallel to pass roads or train lines, which are threatened by avalanche hazards in the lateral slopes and couloirs. For such cases, snow avalanches will cross rather than propagate along fiber optic cables and longitudinal wave propagation as presented here will be reduced. We can nevertheless expect to measure the seismic phases of classes S1, S2, and S3, which could be used to distinguish between powder snow avalanches containing a turbulent flow and pure dense snow avalanches (Köhler, Fischer, et al., 2018). How exactly such flow regime distinction manifests itself in the recognition of class S1, S2, and S3 signals remains to be seen. It may be necessary to further support classification with transition probabilities between states as has been done in previous application of machine learning algorithms to avalanche seismograms (Hammer et al., 2017; Heck et al., 2018). Finally, we stress the advantage of our Gaussian mixture models allowing for different states to coexist at the same time rather than identifying one single dominant state. Future classification could thus be improved with relative probabilities so that the "state mix" describes different parts and kinds of avalanches. Even though this could in theory be done with the presented data set, more data from different installations and test sites should be considered for this to quantify the generalizability of our approach.

Our processing leverages signal coherence over a set of spatially distributed seismic sensors. Equivalent signal processing has already been used in the past for seismic signals of avalanches in the form of array methods (Heck, Van Herwijnen et al., 2019; Lacroix & Helmstetter, 2011). In the present case, the unprecedented amount of seismic sensing locations was combined with unsupervised machine learning to automatically classify signals. To this, future applications could add waveform features (Chmiel et al., 2021) and image processing to further improve classification accuracy (Thrastarson et al., 2021). In any case, we do not expect user-defined threshold rules to perform better than our machine learning scheme since such methods cannot distinguish between signals with similar seismic amplitudes and frequency content.

6.1. Extension to Other Mass Movement Phenomena

Whereas this study focused on snow avalanches, the proposed DAS observation and signal classification could also be applied to other granular media like debris flows, rock-ice avalanches and smaller slope failures. Given

seismogenesis of water turbulence (Gimbert et al., 2014), flood waves could be monitored and detected, as well. Our signals show that DAS is able to detect complex internal processes originating from avalanches propagating over the cable, most likely related to internal avalanche characteristics, which could be manifestations of roll waves or shock waves. These internal waves are general features of open surface flows (Liu & Mei, 1994). Shock waves propagate as flow depth perturbations with larger waves traveling faster, which allows them to grow by "swallowing" smaller waves. The successive merging explains pulsing behavior of granular flows and frontal flow depths, which are much larger and thus more destructive than expected for steady flow (Razis et al., 2014; Viroulet et al., 2018; Zanuttigh & Lamberti, 2007). The detection of internal waves, which our DAS measurements provide, could therefore be a tool for better understanding and predicting maximum flow depths of granular flows and floods.

7. Challenges

Since the apparent phase velocity of an event propagating along the fiber is strongly dependent on the incidence angle of the event, it needs to be treated as a site- and event-specific property. Hence, the transferability of the trained Gaussian mixture model from the given test site to other installations might be limited. Nevertheless, the physical intuition we gained from the analyzed avalanches can be transferred to other sites and cable layouts.

A big challenge for monitoring hazardous alpine mass movements with DAS is the existence and access to fiber-optic infrastructure. Whereas the Vallée de la Sionne test site had accessible infrastructure existing for decades, sites like this are rare. The distribution of existing fibers in rural alpine areas, especially those that are subject to frequent hazardous mass movements, can be limited. In addition to fiber access, stable long-term power access for the interrogation unit can be challenging. If no fiber-optic infrastructure for data transfer exists, real-time monitoring and immediate early warning require a stable data transfer from the interrogation unit to the local responsible authorities. In case the fiber-optic infrastructure is required to be installed, the fiber needs to be protected against mass movement induced damages. For avalanche monitoring this means that the cable needs to be trenched deep enough and be protected against erosion processes. We believe that the DAS technology can not only improve our understanding of hazardous mass movements from a scientific point of view by highlighting interactions between the mass movement and the subsurface but can also improve seismic hazard monitoring and early warning solutions in the near future.

8. Conclusions and Outlook

We have shown that the DAS technology is capable to measure avalanches propagating toward and on top of a fiber-optic cable. The avalanche signals measured from such a system include the seismo-acoustic near- and far-field as well as various mass movement regimes. By combining DAS with Bayesian Gaussian mixture models, we are able to extract key avalanche characteristics and their developments over both space and time. Significant importance for the classification are both the frequency content and the apparent phase velocities of the data within local time-space windows.

DAS adds new observations to the toolbox of mass movement research. With high-resolution recordings, DAS delivers data from the interface of the avalanche with the (sub)surface of the Earth. We observed indications of roll waves. In the future, it can be envisioned that the Froude number could be calculated from the apparent velocities if the depth of the flow is known (similar to Pérez- Guillén et al. (2016)). Further research in amplitude calibration of DAS systems for mass movements is required but a site-specific flow regime characterization based on DAS recordings and physical properties of strain rate measurements can be envisioned in the future.

The incorporation of subsurface strain (rate) as observed with DAS into numerical avalanche simulation tools could increase the usability of DAS data in the field of avalanche dynamics research even further.

Appendix A: Schematic Avalanche Propagating Over DAS Array

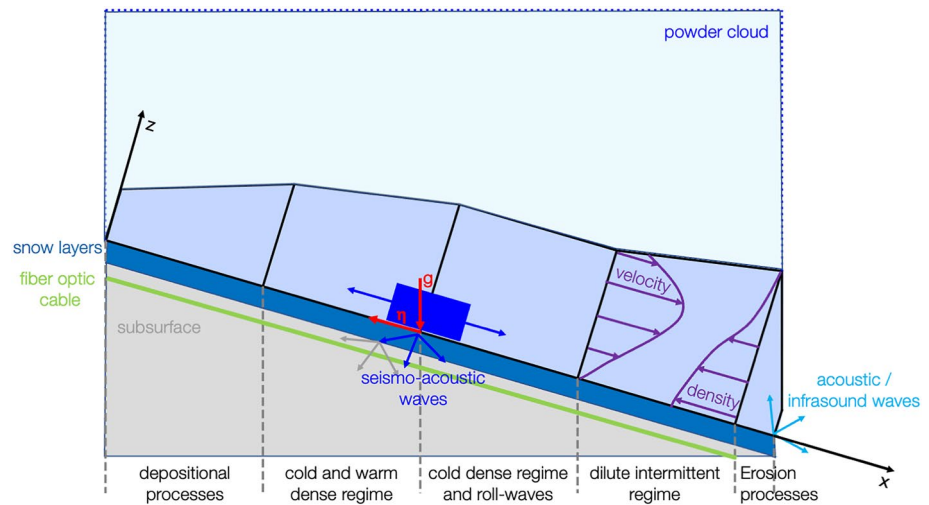See Figure A1 for a schematic avalanche propagation over a DAS array.

Figure A1. Schematic avalanche processes generating seismo-acoustic signals. The velocity and density profile vary with depth (z). The propagation of the avalanche downhill (x) results in acoustic infrasound waves in the air, as well as seismo-acoustic waves due to sliding friction, depositional mechanisms, and interaction of the mass movement with the topography. In this experiment, the fiber-optic cable is located in a conduit in the subsurface, over which the avalanches propagate. Figure schematically after Pérez- Guillén et al. (2016) and Sovilla et al. (2015).

Appendix B: Background Information on Avalanches Discussed in This Manuscript

See Table B1 for an overview of the avalanches investigated within this manuscript.

Table B1
Avalanche Characteristics of the Events Discussed in This Manuscript

Number	Date of occurrence	Duration	Characteristics
Avalanche 3005	2021-01-02	? minutes	- Very small avalanche
			- Did not reach the cable
			- No surges visible
Avalanche 3009	2021-01-15	4 min	- Large transitional powder snow avalanche
			- Fast and dilute component
			- 3 main surges
			- Depositional tail
Avalanche 3016	2021-01-25	1 min	- Powder snow avalanche
			- No dense transition
			- Depositional tail
Avalanche 3020	2021-01-27	0.5 min	- Small powder snow avalanche
			- No dense transition
			- One surge
Avalanche 3021	2021-01-28	1 min	- Dense avalanche
			- Depositional tail
			- One big surge
Avalanche 3022	2021-01-28	2.5 min	- Large transitional powder snow avalanche
			- Depositional tail

Table B1
Continued

Number	Date of occurrence	Duration	Characteristics
Avalanche 3023	2021-01-28	3 min	- Large powder snow avalanche
			- Partial transition
			- Fast and slow part
			- Well separated
			- Long tail
Avalanche 3024	2021-01-28	? minutes	- Large transitional powder snow avalanche
			- Not over array
Avalanche 3026	2021-01-31	? minutes	- No avalanche visible
Avalanche 3027	2021-02-01	? minutes	- Small one-surge powder snow avalanche
			- Not over array
Avalanche 3028	2021-02-01	? minutes	- Large multisurge snow avalanche
			- Not over array
			- Too far away
Avalanche 3030	2021-02-08	? minutes	- No avalanche visible
Event 3036	2021-03-03	<0.3 min	- Regional earthquake M1.2 at Sanetschpass
			- No avalanche

Note. The duration indicates the time the avalanches are propagating on top of the fiber-optic array.

Appendix C: Additional Plots

Additional plots of the avalanche candidates in Table B1 with their corresponding predicted classes from the clustering algorithm can be found in the Supporting Information S1.

Data Availability Statement

To facilitate open research, the data used in this manuscript are available to the public. This include the DAS data, metadata, extracted features from the DAS data, and the trained Bayesian Gaussian Mixture Model. To reproduce the figures in the manuscript as well as in the Supporting Information S1, a jupyter notebook is provided. All processing was performed in Python (v 3.9.13), utilizing SciPy (Virtanen et al., 2020) NumPy (Harris et al., 2020), scikit-learn (Pedregosa et al., 2011), pandas (pandas development team, 2020), cmocean (Thyng et al., 2016) and jupyter (Kluyver et al., 2016). The DAS data, the extracted features as well as the trained Bayesian Gaussian Mixture Model, together with instructions (in a jupyter notebook (Kluyver et al., 2016)) to replicate the figures in the manuscript are available on Zenodo (https://doi.org/10.5281/zenodo.7385433) (Paitz et al., 2022) and are associated to the Snow and Avalanche Dynamics Community (Zenodo, 2022). The data set with the corresponding code examples is currently under restricted access but will be made public upon publication. The reviewers as well as the editor may request access during the review stage of this manuscript.

References

Allstadt, K. E., Matoza, R. S., Lockhart, A. B., Moran, S. C., Caplan-Auerbach, J., Haney, M. M., et al. (2018). Seismic and acoustic signatures of surficial mass movements at volcanoes. *Journal of Volcanology and Geothermal Research*, *364*, 76–106. https://doi.org/10.1016/j.jvolgeores.2018.09.007

Alpaydin, E. (2020). *Introduction to machine learning*. MIT Press.

Ammann, W. J. (1999). A new Swiss test-site for avalanche experiments in the vallée de la sionne/valais. *Cold Regions Science and Technology*, *30*(1–3), 3–11. https://doi.org/10.1016/s0165-232x(99)00010-5

Arattano, M., & Marchi, L. (2008). Systems and sensors for debris-flow monitoring and warning. *Sensors*, *8*(4), 2436–2452. https://doi.org/10.3390/s8042436

Bishop, C. M., & Nasrabadi, N. M. (2006). Pattern recognition and machine learning. *Spring*, *4*(4).

Boore, D. M., & Joyner, W. B. (1997). Site amplifications for generic rock sites. *Bulletin of the Seismological Society of America*, *87*(2), 327–341. https://doi.org/10.1785/bssa0870020327

Capelli, A., Kapil, J. C., Reiweger, I., Or, D., & Schweizer, J. (2016). Speed and attenuation of acoustic waves in snow: Laboratory experiments and modeling with biot's theory. *Cold Regions Science and Technology*, *125*, 1–11. https://doi.org/10.1016/j.coldregions.2016.01.004

Carroll, C., Turnbull, B., & Louge, M. (2012). Role of fluid density in shaping eruption currents driven by frontal particle blow-out. *Physics of Fluids*, *24*(6), 066603. https://doi.org/10.1063/1.4725538

Chmiel, M., Walter, F., Wenner, M., Zhang, Z., McArdell, B. W., & Hibert, C. (2021). Machine learning improves debris flow warning. *Geophysical Research Letters*, *48*(3), e2020GL090874. https://doi.org/10.1029/2020GL090874

Dilley, M. (2005). *Natural disaster hotspots: A global risk analysis* (Vol. 5). World Bank Publications.

Edwards, A., & Gray, J. (2015). Erosion–deposition waves in shallow granular free-surface flows. *Journal of Fluid Mechanics*, *762*, 35–67. https://doi.org/10.1017/jfm.2014.643

Emberson, R., Kirschbaum, D., & Stanley, T. (2020). New global characterization of landslide exposure. *Natural Hazards and Earth System Sciences*, *20*(12), 3413–3424. https://doi.org/10.5194/nhess-20-3413-2020

Fichtner, A., Edme, P., Paitz, P., Lindner, N., Hohl, M., Huguenin, P., et al. (2021). Observing avalanche dynamics with distributed acoustic sensing.

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, *18*(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Gimbert, F., Tsai, V. C., & Lamb, M. P. (2014). A physical model for seismic noise generation by turbulent flow in rivers. *Journal of Geophysical Research: Earth Surface*, *119*(10), 2209–2238. https://doi.org/10.1002/2014jf003201

Grimm, J. (2021). Cryoseismic event analysis on distributed strain recordings leveraging statistical learning methods. MSc Thesis, ETH Zürich.

Grimm, J., & Poli, P. (2022). *Making sense of urban das data with clustering of coherence-based array features (Tech. Rep.). EGU2022*. Copernicus Meetings.

Hammer, C., Fäh, D., & Ohrnberger, M. (2017). Automatic detection of wet-snow avalanche seismic signals. *Natural Hazards*, *86*(2), 601–618. https://doi.org/10.1007/s11069-016-2707-0

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., et al. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hartog, A. H. (2017). *An introduction to distributed optical fibre sensors*. CRC Press.

Heck, M., Hammer, C., van Herwijnen, A., Schweizer, J., & Fäh, D. (2018). Automatic detection of snow avalanches in continuous seismic data using hidden Markov models. *Natural Hazards and Earth System Sciences*, *18*(1), 383–396. https://doi.org/10.5194/nhess-18-383-2018

Heck, M., Hobiger, M., Van Herwijnen, A., Schweizer, J., & Fäh, D. (2019). Localization of seismic events produced by avalanches using multiple signal classification. *Geophysical Journal International*, *216*(1), 201–217.

Heck, M., Van Herwijnen, A., Hammer, C., Hobiger, M., Schweizer, J., & Fäh, D. (2019). Automatic detection of avalanches combining array classification and localization. *Earth Surface Dynamics*, *7*(2), 491–503. https://doi.org/10.5194/esurf-7-491-2019

Iten, M., Puzrin, A. M., & Schmid, A. (2008). Landslide monitoring using a road-embedded optical fiber sensor. In W. Ecke, K. J. Peters, & N. G. Meyendorf (Eds.), *Smart sensor phenomena, technology, networks, and systems 2008* (Vol. 6933, pp. 328–336). SPIE. https://doi.org/10.1117/12.774515

Keylock, C., Ash, M., Vriend, N., Brennan, P., McEwaine, J., & Sovilla, B. (2014). Looking inside an avalanche using a novel radar system. *Geology Today*, *30*(1), 21–25. https://doi.org/10.1111/gto.12033

Klaasen, S., Paitz, P., Lindner, N., Dettmer, J., & Fichtner, A. (2021). Distributed acoustic sensing in volcano-glacial environments—Mount meager, British Columbia. *Journal of Geophysical Research: Solid Earth*, *126*(11), e2021JB022358. https://doi.org/10.1029/2021JB022358

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., et al. (2016). Jupyter notebooks - A publishing format for reproducible computational workflows. In F. Loizides & B. Scmidt (Eds.), *Positioning and power in academic publishing: Players, agents and agendas* (pp. 87–90). IOS Press. Retrieved from https://eprints.soton.ac.uk/403913/

Köhler, A., Fischer, J.-T., Scandroglio, R., Bavay, M., McElwaine, J., & Sovilla, B. (2018). Cold-to-warm flow regime transition in snow avalanches. *The Cryosphere*, *12*(12), 3759–3774. https://doi.org/10.5194/tc-12-3759-2018

Köhler, A., McElwaine, J., & Sovilla, B. (2018). Geodar data and the flow regimes of snow avalanches. *Journal of Geophysical Research: Earth Surface*, *123*(6), 1272–1294. https://doi.org/10.1002/2017jf004375

Lacroix, P., & Helmstetter, A. (2011). Location of seismic signals associated with microearthquakes and rockfalls on the séchilienne landslide, French alps. *Bulletin of the Seismological Society of America*, *101*(1), 341–353. https://doi.org/10.1785/0120100110

Leinss, S., Wicki, R., Holenstein, S., Baffelli, S., & Bühler, Y. (2020). Snow avalanche detection and mapping in multitemporal and multiorbital radar images from terrasar-x and sentinel-1. *Natural Hazards and Earth System Sciences*, *20*(6), 1783–1803. https://doi.org/10.5194/nhess-20-1783-2020

Lindner, N., Edme, P., Paitz, P., Hohl, M., Sovilla, B., Huguenin, P., et al. (2021). Observing avalanche dynamics with distributed acoustic sensing. In *Agu fall meeting abstracts* (Vol. 2021, p. NS12A–05).

Lindsey, N. J., Dawe, T. C., & Ajo-Franklin, J. B. (2019). Illuminating seafloor faults and ocean dynamics with dark fiber distributed acoustic sensing. *Science*, *366*(6469), 1103–1107. https://doi.org/10.1126/science.aay5881

Lindsey, N. J., & Martin, E. R. (2021). Fiber-optic seismology. *Annual Review of Earth and Planetary Sciences*, *49*(1), 309–336. https://doi.org/10.1146/annurev-earth-072420-065213

Lindsey, N. J., Rademacher, H., & Ajo-Franklin, J. B. (2020). On the broadband instrument response of fiber-optic das arrays. *Journal of Geophysical Research: Solid Earth*, *125*(2), e2019JB018145. https://doi.org/10.1029/2019jb018145

Liu, K.-f., & Mei, C. C. (1994). Roll waves on a layer of a muddy fluid flowing down a gentle slope—A bingham model. *Physics of Fluids*, *6*(8), 2577–2590. https://doi.org/10.1063/1.868148

Marchetti, E., Walter, F., Barfucci, G., Genco, R., Wenner, M., Ripepe, M., et al. (2019). Infrasound array analysis of debris flow activity and implication for early warning. *Journal of Geophysical Research: Earth Surface*, *124*(2), 567–587. https://doi.org/10.1029/2018JF004785

Marchetti, E., Walter, F., & Meier, L. (2021). Broadband infrasound signal of a collapsing hanging glacier. *Geophysical Research Letters*, *48*(16), e2021GL093579. https://doi.org/10.1029/2021gl093579

Martin, E. R., Huot, F., Ma, Y., Cieplicki, R., Cole, S., Karrenbach, M., & Biondi, B. L. (2018). A seismic shift in scalable acquisition demands new processing: Fiber-optic seismic signal retrieval in urban areas with unsupervised learning for coherent noise removal. *IEEE Signal Processing Magazine*, *35*(2), 31–40. https://doi.org/10.1109/msp.2017.2783381

Michel, C., Edwards, B., Poggi, V., Burjánek, J., Roten, D., Cauzzi, C., & Fäh, D. (2014). Assessment of site effects in alpine regions through systematic site characterization of seismic stations. *Bulletin of the Seismological Society of America*, *104*(6), 2809–2826. https://doi.org/10.1785/0120140097

Michlmayr, G., Chalari, A., Clarke, A., & Or, D. (2017). Fiber-optic high-resolution acoustic emission (ae) monitoring of slope failure. *Landslides*, *14*(3), 1139–1146. https://doi.org/10.1007/s10346-016-0776-5

Minardo, A., Catalano, E., Coscetta, A., Zeni, G., Zhang, L., Di Maio, C., et al. (2018). Distributed fiber optic sensors for the monitoring of a tunnel crossing a landslide. *Remote Sensing*, *10*(8), 1291. https://doi.org/10.3390/rs10081291

Paitz, P., Edme, P., Gräff, D., Walter, F., Doetsch, J., Chalari, A., et al. (2021). Empirical investigations of the instrument response for distributed acoustic sensing (das) across 17 octaves. *Bulletin of the Seismological Society of America*, *111*(1), 1–10. https://doi.org/10.1785/0120200185

Paitz, P., Lindner, N., Edme, P., Huguenin, P., Hohl, M., Sovilla, B., & Fichtner, A. (2022). Zenodo: Phenomenology of avalanche recordings from distributed acoustic sensing. [data and software]. Zenodo. https://doi.org/10.5281/zenodo.7385433

pandas development team, T. (2020). pandas-dev/pandas: Pandas. Zenodo. https://doi.org/10.5281/zenodo.3509134

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pérez- Guillén, C., Sovilla, B., Suriñach, E., Tapia, M., & Köhler, A. (2016). Deducing avalanche size and flow regimes from seismic measurements. *Cold Regions Science and Technology*, *121*, 25–41. https://doi.org/10.1016/j.coldregions.2015.10.004

Petley, D. (2012). Global patterns of loss of life from landslides. *Geology*, *40*(10), 927–930. https://doi.org/10.1130/g33217.1

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge University Press.

Prokop, A., Schön, P., Wirbel, A., & Jungmayr, M. (2014). Monitoring avalanche activity using distributed acoustic fiber optic sensing. In *Proceedings of the international snow science workshop* (pp. 129–133).

Razis, D., Edwards, A., Gray, J., & van der Weele, K. (2014). Arrested coarsening of granular roll waves. *Physics of Fluids*, *26*(12), 123305. https://doi.org/10.1063/1.4904520

Sánchez-Sesma, F. J., Weaver, R. L., Kawase, H., Matsushima, S., Luzón, F., & Campillo, M. (2011). Energy partitions among elastic waves for dynamic surface loads in a semi-infinite solid. *Bulletin of the Seismological Society of America*, *101*(4), 1704–1709. https://doi.org/10.1785/0120100196

Scott, D. W. (2015). *Multivariate density estimation: Theory, practice, and visualization*. John Wiley & Sons.

Sovilla, B., Burlando, P., & Bartelt, P. (2006). Field experiments and numerical modeling of mass entrainment in snow avalanches. *Journal of Geophysical Research*, *111*(F3), F03007. https://doi.org/10.1029/2005jf000391

Sovilla, B., McElwaine, J. N., & Louge, M. Y. (2015). The structure of powder snow avalanches. *Comptes Rendus Physique*, *16*(1), 97–104. https://doi.org/10.1016/j.crhy.2014.11.005

Steinkogler, W., Gaume, J., Löwe, H., Sovilla, B., & Lehning, M. (2015). Granulation of snow: From tumbler experiments to discrete element simulations. *Journal of Geophysical Research: Earth Surface*, *120*(6), 1107–1126. https://doi.org/10.1002/2014jf003294

Suriñach, E., Flores-Márquez, E. L., Roig-Lafon, P., Furdada, G., & Tapia, M. (2020). Estimation of avalanche development and frontal velocities based on the spectrogram of the seismic signals generated at the vallée de la sionne test site. *Geosciences*, *10*(3), 113. https://doi.org/10.3390/geosciences10030113

Suriñach, E., Furdada, G., Sabot, F., Biescas, B., & Vilaplana, J. (2001). On the characterization of seismic signals generated by snow avalanches for monitoring purposes. *Annals of Glaciology*, *32*, 268–274. https://doi.org/10.3189/172756401781819634

Suriñach, E., Vilajosana, I., Khazaradze, G., Biescas, B., Furdada, G., & Vilaplana, J. M. (2005). Seismic detection and characterization of landslides and other mass movements. *Natural Hazards and Earth System Sciences*, *5*(6), 791–798. https://doi.org/10.5194/nhess-5-791-2005

SwissTopo. (2021). Online tool for Swiss Raster images. Retrieved from https://www.swisstopo.admin.ch/de/geodata/maps/smr/smr10.html

Thrastarson, S., Torfason, R., Klaasen, S., Paitz, P., Cubuk Sabuncu, Y., Jónsdóttir, K., & Fichtner, A. (2021). Detecting seismic events with computer vision: Applications for fiber-optic sensing. *Earth and Space Science Open Archive*, *19*. https://doi.org/10.1002/essoar.10509693.1

Thyng, K. M., Greene, C. A., Hetland, R. D., Zimmerle, H. M., & DiMarco, S. F. (2016). True colors of oceanography: Guidelines for effective and accurate colormap selection. *Oceanography*, *29*(3), 9–13. https://doi.org/10.5670/oceanog.2016.66

Tregaskis, C., Johnson, C., Cui, X., & Gray, J. (2022). Subcritical and supercritical granular flow around an obstacle on a rough inclined plane. *Journal of Fluid Mechanics*, *933*, 933. https://doi.org/10.1017/jfm.2021.1074

Tsai, V. C., Minchew, B., Lamb, M. P., & Ampuero, J.-P. (2012). A physical model for seismic noise generation from sediment transport in rivers. *Geophysical Research Letters*, *39*(2), L02404. https://doi.org/10.1029/2011gl050255

Van Herwijnen, A., & Schweizer, J. (2011). Monitoring avalanche activity using a seismic sensor. *Cold Regions Science and Technology*, *69*(2–3), 165–176. https://doi.org/10.1016/j.coldregions.2011.06.008

Viroulet, S., Baker, J. L., Rocha, F. M., Johnson, C. G., Kokelaar, B. P., & Gray, J. M. N. T. (2018). The kinematics of bidisperse granular roll waves. *Journal of Fluid Mechanics*, *848*, 836–875. https://doi.org/10.1017/jfm.2018.348

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Walter, F., Gräff, D., Lindner, F., Paitz, P., Köpfli, M., Chmiel, M., & Fichtner, A. (2020). Distributed acoustic sensing of microseismic sources and wave propagation in glaciated terrain. *Nature Communications*, *11*(1), 1–10. https://doi.org/10.1038/s41467-020-15824-6

Wenner, M., Allstadt, K., Thelen, W., Lockhart, A., Hirschberg, J., McArdell, B. W., & Walter, F. (2022). Seismometer records of ground tilt induced by debris flows. *Bulletin of the Seismological Society of America*, *112*(5), 2376–2395. https://doi.org/10.1785/0120210271

Williams, E. F., Fernández-Ruiz, M. R., Magalhaes, R., Vanthillo, R., Zhan, Z., González-Herráez, M., & Martins, H. F. (2019). Distributed sensing of microseisms and teleseisms with submarine dark fibers. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-13262-7

Yu, Z., Dai, H., Zhang, Q., Zhang, M., Liu, L., Zhang, J., & Jin, X. (2018). High-resolution distributed strain sensing system for landslide monitoring. *Optik*, *158*, 91–96. https://doi.org/10.1016/j.ijleo.2017.12.013

Zanuttigh, B., & Lamberti, A. (2007). Instability and surge development in debris flows. *Reviews of Geophysics*, *45*(3), RG3006. https://doi.org/10.1029/2005rg000175

Zenodo. (2022). Snow avalanche dynamics community on zenodo. Author. Retrieved from https://zenodo.org/communities/avalanche-dynamics/

Zhan, Z. (2019). Distributed acoustic sensing turns fiber-optic cables into sensitive seismic Antennas. *Seismological Research Letters*, *91*(1), 1–15. https://doi.org/10.1785/0220190112

Zhongming, Z., Linong, L., Xiaona, Y., Wangqiang, Z., & Wei, L. e. a. (2021). Ar6 climate change 2021: The physical science basis. 2021 IPCC report.