**An Unsupervised Machine-Learning-Approach to Understanding Summer Seismicity at an Alpine Glacier**

*Theresa Sawia, Meredith Nettlesa, Ben Holtzmana, John Paisleyb, Fabian Walterc*

*aLamont Doherty Earth Observatory, Columbia University, bDepartment of Electrical Engineering, Columbia University, cETH Zürich Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie*

**Key points:**

* We apply multiple layers of unsupervised feature extraction to >2900 spectrograms of glacial icequakes and seismic noise recorded over two months at an Alpine glacier, then cluster the results to reveal characteristic, at times subtle, spectral and waveform attributes.
* We introduce a novel method of “reverse engineering” the features that were extracted during some of the unsupervised machine learning steps, allowing for direct physical interpretation of the dataset in terms of highly resolved spectral features.
* Based on our findings, we propose new timing constraints on the subglacial component of a glacial outburst flood, and suggest other seismic signatures of sub- and englacial hydrology.

**Abstract**

Continuous, long-term observations of glaciers and ice sheets are vital for understanding cryospheric dynamics, yet the sub- and englacial environments that control many aspects of ice behavior are inherently difficult to observe. Cryoseismology offers tools for monitoring interior, basal, and surface processes of glaciers and ice sheets, with seismic sources ranging from discrete “icequakes” to sustained, high-frequency seismic tremor. The varied, complex, and sometimes subtle nature of ice-derived seismic signals means there is much to discover in large cryoseismic datasets, but these same types of datasets pose challenges for more traditional seismic analysis techniques. Here, we use an unsupervised machine learning algorithm, now called “SpecUFEx” (Holtzman et al., 2018), on spectrograms of vertical component, single station seismograms from June-July 2007 at Gornergletscher, Switzerland, including 1411 previously cataloged icequakes (Walter, 2009), and 1521 minute-long samples of icequake-free “noise”. After extracting and clustering characteristic time-varying spectral features from the data, we discover that spectrograms vary systematically in time corresponding with glaciohydrologic events, including a glacial outburst flood and afternoon meltwater production. We introduce a novel method of “reverse engineering” the output of SpecUFEx to recover stationary approximations of the spectral features, which we interpret as driven by sub- and englacial hydrology. By using SpecUFEx, we are able to provide new insight about the timing of the glacial outburst flood and detect other, more subtle seismic fluid signatures of glacial hydrology. This work demonstrates a fully interpretable, unsupervised machine-learning approach for exploration of long-term continuous seismic data from a temperate glacier, with possible applications to other systems that host fluid-mediated seismicity.

**Introduction**

Rapid changes in the cryosphere could have substantial implications for our climate and oceans, yet the sub- and englacial conditions that control ice dynamics are inherently difficult to observe. Cryoseismology offers tools for monitoring surface, basal, and interior processes of glaciers and ice sheets over long periods of times and vast spatial scales. Decades of research have revealed a diverse range of cryoseismic sources, including impulsive “icequakes” caused by ice flow and deformation (see review by Podolsy and Walter, 2016) and more recently high-frequency (0.5-20 Hz) tremor caused by turbulent water flow and bedload transport in sub- and englacial channels (e.g., Röösli et al., 2014, Bartholomaus et al., 2015; Gimbert et al., 2016 Vore et al., 2019, Eibl et al.2020, Linder et al., 2020). Such signals are often inspected visually from spectrograms, but visual inspection could be a practical challenge if one needed to view, for example, 10,000 spectrograms from multiple stations in a network.

Gaining insight from the tremendous volume and variety of cryoseismic data poses a challenge for typical seismic analysis techniques, leading to growing interest in application of machine learning methods. Supervised machine learning algorithms — those trained on datasets that contain manually labeled data points — have shown remarkable success at detecting and classifying tectonic earthquakes (e.g., Ross et al 2018 and 2019; Zhu and Beroza, 2018; Mousavi et al., 2020; Park et al., 2020), including at fluid-rich volcanic or geothermal settings (Ren et al., 2020, and reviews by Malfante et al., 2018; Carniel and Guzman, 2020), whose fluid-mediated seismic behavior could be considered at times analogous to temperate glacial settings (both systems, for example, feature fluids flowing through conduits, whether subglacial water flow or magma through dikes). However, for the many cases where the labeled datasets are lacking or inexistent, unsupervised machine learning (UML) methods must be employed. UML models infer patterns in data without prior training on labeled data, and have been used to cluster similar groups of seismic data (e.g., Trugman and Shearer, 2017; Chamarczuk et al., 2019; Lamb et al., 2020; Seydoux et al., 2020; Steinmann et al., 2021), to characterize the data by reducing it into lower-dimension, statistically meaningful “features” (e.g., Yoon et al., 2015, Lindenbaum et al., 2018), or a combination of both (e.g., Sick et al., 2015; Holtzman et al., 2018; Ren et al., 2020; Jenkins et al., 2021). Unlike supervised machine learning, the output from UML can be abstract; difficult to interpret physically because no labels are predicted and there is no guarantee that the clusters or features are of scientific interest.

One unsupervised spectral feature extraction algorithm (now called “SpecUFEx,” Holtzman et al., 2018) has proven particularly well-suited for characterizing subtle, fluid- and thermal-driven effects in seismic data. SpecUFEx was shown to reveal time-varying spectral features of injection-induced earthquakes in a geothermal reservoir that corresponded with nearby fluid-injection levels (Holtzman et al., 2018). Here, we expand the post-SpecUFex workflow to include further dimensionality reduction prior to clustering, then reconstruction of the extracted features back into the spectral domain, thus allowing for direct physical interpretation of results. In anticipation of a deluge of cryoseismic data for which semi-automated, unsupervised machine-learning-based analyses could provide great practical benefits, we present here a case study describing the successes and limitations of using such an approach to explore two months of summertime-seismic activity at a temperate glacier in the Swiss Alps.

**Background**

Gorner Glacier is the second largest glacier in the Swiss Alps, spanning approximately 57 square kilometers at an average elevation of 2500m. Gorner Glacier converges with Grenzgletcher and they flow generally westward at a speed on the order of centimeters per day (Table 1, and Garcia et al., 2019). Proglacial discharge has been measured continuously for over 50 years by a hydro-electric company about 4.5 km away from Gornersee (e.g., Huss et al., 2007). An ice-marginal lake, Gorner See, forms every spring at the confluence of the two glaciers and by midsummer drains catastrophically in a “jökulhlaup”, a glacial outburst flood, releasing up to 60 million cubic meters per second of water in some combination of sub-, en-, and supra-glacial flow (Huss et al., 2007, Werder et al., 2009a, 2010). A timeline of reported events for the 2007 Gorner See glacial outburst flood is given in Table 1. Lake drainage began on July 4th, as Gorner See breached its shore and flowed over the surface of the glacier to a lake-proximal moulin. From July 7-15, the majority of the lake drained through an en- and subglacial conduits, concluding when the lake was mostly empty and the lake level meter subaerially exposed (Werder et al., 2009a - *but I need to ask Fabian/Jonny/Meredith about the hydrological timeline in this study, because I don’t understand some parts of it|||*). Although sediment transport data were not collected during the 2007 study period, historical (1973-1990) and more recent (2016, 2017) observations show high sediment concentration in proglacial runoff (~1kg/m^3), with peak sediment transport during lake drainages (e.g., Collins, 1989; Delaney et al., 2018). The annual outburst floods typically accelerate the flow of Gorner glacier (Sugiyama et al., 2007, 2008 & 2010), which can increase icequake activity (Riesen, 2007; Roux et al., 2010; Garcia et al., 2019). A burst of surficial icequakes marked the onset of the 2007 lake drainage, but no such increase in basal seismicity was observed (Walter et al., 2009). Meltwater runoff is another major hydrologic driver of seismicity at Gorner Glacier. Surface seismicity rates peak in the afternoon as meltwater-flow at the ice/bed interface raises the glacier to near flotation levels, allowing for reduced basal drag and increased ice deformation (Walter 2008 and 2009; Roux et al., 2010). Walter et al., (2009) detected 100s to 1000s of icequakes daily in June-July 2007, with the vast majority (>99%) caused by tensile openings of crevasses in the top twenty meters of the ice. A few tens of basal icequakes were also detected daily, typically in the early morning hours when subglacial water pressure is at its lowest. Icequakes are frequently co-occurring with persistent background glacial seismic “noise,” such as tremor, highlighting some of the complexities that we attempt to overcome by analyzing a carefully selected cryoseismic dataset using the expanded SpecUFEx workflow.

**Data**

Our seismic data come from an 8-instrument array of three-component Geospace GS-11D geophones deployed on-ice at Gorner Glacier in shallow boreholes between June-July, 2007. The instruments have an aperture of ~200 m, and are no farther than ~200 m from the shore of Gorner See at its maximum height prior to drainage. The seismometers sampled continuously at 1000 sps, and were leveled daily to account for the effect on differential melting of the ice surface beneath the instruments. For simplicity, only the vertical seismic velocity components are used in this study. We perform the same analysis on two separate data sets: one containing two-second-long icequake waveforms cataloged by Walter (2009), and the second containing minute-long records of icequake-free background noise that was compiled for this study. “Noise” in this case means a signal of unknown origin, or a signal that is non-impulsive at minute-long timescales (including some tremor). Because icequakes inevitably occur in the presence of noise, the two experiments help decipher the relative effects of the icequakes and noise on the unsupervised results. Additionally, exploring the noise on its own allows for more emphasis on seismic tremor that could be caused by water flow in and under the glacier.

Our first experiment is concerned with icequakes. Walter (2009) reports over 100,000 icequakes in June-July 2007, detected using a modified short term average/long term average (STA/LTA) filter. The STA/LTA filter is calculated by taking the ratio of two moving averages of the waveform signal; one short window and one long window (in our case, 3 ms and 50 ms, respectively). When the ratio exceeds a certain threshold, the STA/LTA detector is triggered. A subset (N= ~24,000) of those detected events were located with a grid-search inversion algorithm (Lee and Steward, 1981) using an ice-velocity gradient calculated from an active seismic experiment (citation needed |||). Errors in location range from ~1 m within the array footprint, to over 200 m with a mean error of 70 m. The icequakes are high frequency (~50 Hz), low magnitude (M<1), short-duration (~1s), and can be easily distinguished from basal events and background noise by their prominent Raleigh wave arrival.

We highlight results from a single station J8 from the eight-station network deployed during this study, [[and include results from other stations in the supplementary materials??]]. We limit our study to icequakes located in the top 50 meters of ice and within the array footprint where locations are most accurate. In other experiments we included more distal icequakes, but results appeared irrelevant for our interests. Our initial dataset is composed of two-second long waveforms of individual icequakes from the catalog, with the onset of the Raleigh wave onset time-shifted to the 0.5-second mark (Walter, 2009). We end up with a dataset of 1411 icequake events with locations and depths in Swiss Grid coordinates as shown in Figure 1, with their timing throughout the season and binned by hour shown in **Figure 2 b and d**, respectively.

The second experiment is performed on 1521, minute-long records of seismic “noise” that is free of icequakes, sampled every ten minutes during the same two-month time period as the icequakes (**Figure 2c**). We use the same STA/LTA filter as above to identify and discard records with prominent icequakes in them. We also discard records if icequakes are detected in the five-seconds preceding, or the two-seconds following the record of noise, to account for sustained icequake codas and possible precursory signals, respectively. The bar-plots of icequake and noise samples in Figures 2d and 2e show how the high rates of icequakes in the afternoon hours correspond with under-sampling of noise in that same time period. The pervasiveness of icequakes (at times >5 icequakes per minute) prevents a uniform sampling of minute-long, icequake-free noise records.

To understand how environmental factors are interacting with seismicity, we utilize local GPS, lake level, temperature and precipitation measurements collected simultaneously with the seismic data (**Figure 2a**; Sugiyama et al., 2007; Garcia et al., 2019). Four GPS stations are located between 100 m and 3 km down-glacier from Gorner See, sampling continuously every 2 minutes with a horizontal and vertical accuracy of +/-1.4 mm and +/-3 mm, respectively. We use the GPS displacement measurements from Garcia et al. (2019) relative to their location at the beginning of our study, correct the slope of the displacement over time by subtracting background glacier velocity reported at each GPS station (Table 1), then take an hourly rolling average for each station. This allows us to visualize details about the GPS time series (green lines, Figure 2), including diurnal periodic displacements as well as the larger surges mediated by temperature, precipitation, and the glacial outburst flood. Lake levels were measured by a pressure transducer sampling continuously every minute (thick blue solid dashed line, Figure 2a). Temperature (red solid line) and precipitation data (thin solid blue line) come from an automated weather station located on the northern margin of Gorner Glacier, 2.5 km from and 600 m higher than the seismic array (Garcia et al., 2019).

**Methods**

The workflow consists of using Fourier analysis to transform waveforms into spectrograms, analyzing these spectrograms with SpecUFEx to produce fingerprints, reducing the fingerprints using principal component analysis (PCA), clustering the fingerprints using k-means clustering, and evaluating the clusters using silhouette scores. We then look for empirical relationships between those clusters and physical/environmental properties of the glacier, and interpret the relationship between those properties and spectral and waveform patterns. We then demonstrate the ability to “reverse engineer” meaningful spectral patterns from the output of SpecUFEx to help with physical interpretation.

**Spectrogram generation:**

We calculate spectrograms of the seismic ground velocity measurments using a short-time Fourier transform (STFT). Table 2 lists the parameters used to generate the spectrograms, including STFT window length, overlap, and number of windows for the icequake or noise observations. The spectrograms are then filtered according to the parameters given in Table 2. The upper frequency constraint ,, is chosen to avoid the sharp spectral peaks in the noise data at around 90 Hz (see **Figure 3b,** top panel) and the prominent lobes in icequake data around 160 Hz (**Figure 3a**, top panel). Our preliminary experiments suggest that these features are instrument artifacts which can strongly impact results. The lower frequency bound, ,, is chosen to be near theoretical limit defined by the reciprocal of the STFT window length, while also keeping in mind that there must be enough STFT windows for robust analysis by the hidden Markov Model. HMM learns on reoccurring time-varying patterns, so having too few time steps will lead to under-fitting the model. We find that using at least 30 STFT windows is recommended. Following Holtzman et al. (2018), who first developed SpecUFEx, we normalize each spectrogram by its median, convert it to decibels, and force negative values, if they occur, to zero in order to comply with the non-negativity constraint of non-negative factorization (described below).

**SpecUFEx**

SpecUFEx applies two layers of unsupervised feature extraction to the spectrograms, beginning with nonnegative matrix factorization (NMF; **Figure 4a**). NMF reduces spectrograms into a basis “dictionary” of sparse and meaningful spectral patterns, and an activation coefficient matrix (ACM) that shows how those patterns vary over time (Lee and Seung, 1999). The NMF dictionary is shared for all spectrograms, but the ACMs are generated uniquely for each spectrogram. Unlike other unsupervised dimensionality reduction techniques (e.g., principal or independent component analysis, vector quantization) NMF is solely additive (since its components are nonnegative) leading to highly similar low-dimensionality representations of spectrograms. For example, **Figure 4 (first column in rows a and b)** shows the full-dimensionality spectrogram (Figure 4a) above its NMF-reduced dimensionality counterpart (the ACM, Figure 4b). Even as the frequency resolution of the spectrogram is reduced, its predominant features are retained in the ACM through time.

The ACMs are then used as input for a hidden Markov model (HMM, Figure 4b). The HMM assumes that the temporal evolution of the ACM’s spectral patterns depends on unobservable “hidden states,” where each hidden state is a combination of spectral patterns from the NMF dictionary that tend to cooccur in time. Those states are given in the “emissions matrix” (EB), which, like the NMF Dictionary, is shared for all events. The state transition matrices (STM, **Figure 4C**) show how those states vary in time, and are generated uniquely for each ACM. The final product, the fingerprints (**Figure 4D**), show the probability of each state transitioning to another (or itself) from one time step to the next. For further details of the SpecUFEx algorithm, see the appendix of Holtzman et al. (2018).

**Principal component analysis**

The fingerprints—the 255-dimensional representations of spectrograms—are very sparse in our case, so to improve statistical representation we further reduce their dimensionality prior to clustering using principal component analysis (PCA, i.e., Pearson, 1901; Lever et al., 2017). We choose the number of principal components so that at least 86% of the variance (about two standard deviations above the mean) of the original fingerprints is preserved, resulting in a 67-dimensional dataset for the icequake fingerprints and a 3-dimensional dataset for the noise fingerprints. In retrospect, reducing the fingerprints at this step does not drastically change clustering results for our particular dataset.

**Clustering**

We evaluate how well clustered each fingerprint is by using silhouette scores (Rousseeuw, 1987). Silhouette scores range from -1 to 1 with high positive scores indicating well-clustered fingerprints. The silhouette score, **S(i)**, of a given fingerprint (**i**)is defined as

,

where **a** is the ℓ norm of fingerprint i to every other fingerprint in its cluster, and **b** is the ℓ norm of fingerprint *i* to every fingerprint in the closest neighboring cluster. Negative silhouette scores imply that a fingerprint is on average more similar to fingerprints in a cluster to which it was not assigned, but only 84 of 1411 icequake fingerprints (< 6%) and 17 of the 1521 noise fingerprints (< 2%) have negative silhouette scores; too few to impact the overall clustering trends in this study. By evaluating the mean silhouette score (MSS) of a dataset, we can qualitatively evaluate how well-clustered the dataset is for a given number of clusters.

**Reconstructing spectral features**

We introduce a novel method for recovering the spectral features that were most important for the clustering assignments, allowing for direct physical interpretation of the unsupervised output. We first choose a “key state” for each cluster based on which element on the main diagonal of the fingerprint matrix has the highest value, then map that state back into the frequency domain. **Figure 11** shows the fingerprint with the highest silhouette score from each cluster with the key state circled and numbered in green. Qualitatively, we see that the fingerprints are more sparse for the noise experiment (right column) than the icequakes (left column). The noise fingerprints are composed primarily of one or two states on the main diagonal, suggesting that the entire spectrogram can be represented by one or two states occurring repeatedly through time. The icequake fingerprints are more complicated, with many off-diagonal elements in addition to “key states” on the diagonals.

**Results**:

Clustering analysis results in 3 clusters for the icequake data (maximum MSS = 0.11, **Figure 5c**) and 4 clusters for the noise data (maximum MSS = 0.63, **Figure 5d**). **Figure 5a** shows the fingerprints colored by their cluster assignment, projected onto their first three principal subspaces (for icequakes, ~28% of the original fingerprint variance is preserved in this image). The clusters are organized for clarity by randomly assigned colors, and we note that no similarities between the icequake and noise clusters are suggested by shared colors. To characterize the observations within each cluster, it is useful to work with a subset: the top 20 events in each cluster with the highest silhouette scores. These are the “representative events” from each cluster, marked with black crosses in **Figure 5a and b** (60 events for the icequakes, 80 for noise). Some of the fingerprints (and therefore their PCA reductions) are identical, resulting in overlapping crosses.

**Figure 6** shows the waveforms of the representative icequakes (left column) and spectra for the representative noise (right column). For visualization of the icequakes, we normalize the waveforms by their maximum amplitude, time-lag them to maximize their cross-correlation, stack them, and then visualize a 0.8 second window containing the majority of the waveform energy. These processing steps bring out subtle features that characterize waveforms in each of the clusters. For example, the S-phase of each event, indicated by the thick black arrows in Figure 6, takes on a distinct shape for each cluster. The first cluster (blue) has a relatively high amplitude S-wave compared to the other two clusters, whereas the third (orange) cluster appears to have a distinct indentation at the peak of the S-wave. Additionally, the first icequake cluster (blue) has a visibly lower signal-to-noise ratio than the other two clusters due to a constant background noise around 30-70 Hz (seen in representative spectrograms; not shown), and the blue and orange clusters both have remarkably coherent reverberations in their codas (thin black arrows, **Figure 6**).

We find that spectra -- not waveforms— are better for discerning distinguishing characteristics between clusters for the noise data. Instead of calculating the Fourier transform of the whole signal, we sum the STFT values (columns of the spectrograms) over the time dimension, in order to preserve the filtering and normalization used to generate the spectrograms. We attempted align and stack the noise waveforms to bring out coherent details between waveforms (as we did with the icequakes), but results were indistinguishable from the same method applied to random time series (as is expected from noise). By stacking the spectra, however, we notice distinguishing peaks between clusters (black arrows), as well as relative differences in amplitude and overall shape. The second (red) cluster, for example, retains energy in the higher frequencies compared to the other three clusters resulting in a distinctly flatter spectra.

For the icequakes, the highly similar waveforms for the second (red) and third (orange) clusters suggest a similar source *and* path, yet **Figure 7** shows that this cannot always be the case. The dots with black outlines are the 20 most representative events in each cluster. The representative red dots are concentrated on lake-marginal clusters, so one can suppose a somewhat similar path for those icequakes to station J8 (light green triangle). The blue and orange icequakes, however, are more dispersed, showing that that icequakes’ paths must not necessarily be similar for their fingerprints to be clustered together.

**Figure 8 (a, b)** describes the station-event distances for the 20 representative events in each cluster using “box-and-whisker plots”. The boxes enclose the 25th-75th percentiles of the distance measure, with a line indicating the 50th percentile (median). The whiskers extend 1.5 times the interquartile range either side from the median, and individual points beyond the whiskers are considered outliers. Depths vary little between clusters, although the station-distance of icequakes in the red cluster are on average greater than icequakes in the other clusters. **Figure 8** **(c-f)** shows box-and-whisker plots of all the waveforms and spectrograms within each cluster for icequakes (left column) and noise (right column). **Figure 8c,e** (“log10(RSAM)”) is based on the “real-time amplitude measurement” of the waveform,

,

where ***x(t)*** is the seismic ground velocity filtered through the same band-passed filter as applied to spectrograms during preprocessing. RSAM measures the overall energy of the waveform, and is frequently used in volcano-seismology as an indicator of volcanic activity (Endo & Murray, 1991). **Figure 8d,f** shows the top 20 representative waveforms’ mean spectral centroids (SC); the spectral center of mass of a waveform (Klapuri & Davy, 2007). To calculate the spectral centroid, each time frame of the spectrogram (i.e., a spectra) is treated as a distribution over *N* frequency bins from which a centroid can be found. SC(t) is calculated as:

,

where **s** is the magnitude spectrum, *n* indicates the frequency bin, and **f** is the array of frequencies for a given time frame of the spectrogram.

To understand the clusters’ connection to glacial behavior, we compare the timing of the observations in each cluster to ancillary geophysical data, including ambient temperature, on-ice GPS displacement, precipitation measurements and lake-level data. **Figure 9** shows the clusters binned by hour of the day (UTC) for the entire study period, along with the mean temperature difference from the daily mean temperature (dark red dashed line). There is little variation between the icequake clusters in the diurnal distribution of events. In contrast, several of the noise clusters display distinct distributions. The first (blue) and fourth (purple) noise clusters peak in the evening (04-22 UTC), with the purple peak preceding the blue by 2-4 hours. The second (red) cluster strongly resembles the timing distribution of the original dataset (top bar plot, in black). The third (orange) cluster has most of its noise events in the morning hours (00-12 UTC).

**Figure 10** shows the timing of clusters alongside the ancillary data shown in **Figure 2.** The acceleration of ice flow shown by the GPS data and the reported onset of the outburst flood (July 4th; earliest dashed black lines in Figure 10) both approximately coincide with the blue icequake cluster and the red noise cluster, although both of these clusters continue after the ice-flow acceleration has ceased. Beyond this, there is no strong seasonal trend in the clustering of icequake clusters. For the noise data, however, there is a transition in which clusters are active, which approximately coincides with the onset of the flood (summarized in Table 1). The blue, orange and purple clusters are all active in the pre-drainage season, but within hours of supraglacial lake drainage (July 4th), the purple noise cluster pauses. About a day later (late on July 5th, or early July 6th), the blue and orange clusters cease as well. About 99% (594/601) of the noise events in the red cluster begin on or after July 5th; two days prior to the reported sub- and englacial components of the flood. The purple and orange cluster resume on July 15th and 17th (respectively), but the blue cluster never resumes.

**Figure 12** shows the “reconstructed spectra” (dashed lines) from key states chosen from the fingerprints in Figure 11, along with “median spectra” (solid lines) from each cluster, calculated by taking the median value at each frequency bin for all spectra in a cluster. The reconstructed spectra are generated and visualized as such:

1) We choose a key state for each cluster based on which element in the fingerprint has the highest value.

2) We then multiply the HMM emissions matrix (“EB,” **Figure 4**) with the transpose of the NMF dictionary (“Dict,” **Figure 4)**, and use it to map the key states to a frequency vector (the “reconstructed spectra”).

3) To allow for clear comparison between the shapes of the reconstructed and median spectra, we scale them by their maximum value, that is, we divide all of the spectra by the maximum value among all spectra, and likewise divide all of the reconstructed spectra by the maximum value among all reconstructed spectra (*not* by cluster). We display the results in logarithmic scale so the lower-amplitude details at the uppermost frequencies are not lost in visualization.

Each reconstructed spectra in Figure 12 (dashed lines) is compared to the “actual” median-normalized spectra of the waveforms (solid lines). The reconstructed icequake spectra (upper panel Figure 12) do not closely resemble the actual spectra, in that the actual spectra are relatively flat across all frequencies whereas the reconstructed spectra each have numerous distinct local maxima and minima. The blue reconstructed icequake spectra (cluster 1), for example, has a broad peak in amplitude around 35-45 Hz, whereas the orange reconstructed spectra (cluster 3) has higher amplitude frequency content up to about 39 Hz.

The reconstructed noise spectra (lower panel Figure 12) are much more successful at recovering small details found in the actual, median-normalized spectra, capturing, for example, the small local peaks around 25 Hz for the blue, orange and purple clusters (clusters 1, 3 and 4, respectively). The reconstructed noise spectra also capture some general trends of the median-normalized noise spectra, including the flat frequency characteristic of the red (second) cluster. It is important to remember that since the spectrograms were median-normalized prior to SpecUFEx, the reconstructed spectra reflect that normalization. For example, the median spectra for the second noise cluster (red solid line, lower panel Figure 12) is in reality the highest amplitude signal across all frequencies. At higher frequencies the reconstructed noise spectra diverge significantly from the actual noise spectra, where the low-amplitude dips in the reconstructed spectra due to sparse NMF spectral patterns are exaggerated due to the logarithmic scaling. In summary, we are able to retrieve highly resolved details of the noise spectra — but not icequake spectra — by mapping the SpecUFEx HMM states back to the frequency domain

**Discussion**

We demonstrate how SpecUFEx, an unsupervised machine-learning algorithm, can discover characteristic time-varying spectral patterns in >4000 spectrograms recorded at a glacial system with active hydraulics. Unlike the initial SpecUFEx publication (Holtzman et al., 2018), we apply further processing steps to the SpecUFEx workflow, including PCA on the fingerprints, k-means clustering on the PCA projections, and clustering validation by silhouette scores. The PCA projections (Figure 5) show that the clusters form continuous branches connected to one center mass, and that, for the noise particularly, these branches seem related to changing hydrologic and dynamic conditions at the glacier (see Figure 10). We interpret the continuous nature of the branches as reflective of possible continuous subglacial parameters (e.g., conduit geometry, water flow speed, hydraulic pressure, bedload properties,…). The silhouette scores help us find define an optimal number of clusters for k-means clustering, and additionally allows uss to quantitatively select characteristic end-members of the different branches. Our method is able to explore the large- and fine-scale trends that exist in the dataset, revealing highly coherent detail between a small subset of representative cluster members while also displaying overall trends within the entire dataset; both important criteria for unraveling the physical origins of the signals.

The majority of the spectral energy in all of the noise clusters occur in the 5-40 Hz band; a signal which could be explained by subglacial turbulent water flow and saltation from sediment transport (citation|||||). The reported timing of the subglacial drainage (Werder et al., 2009a) was chosen based on the peak lake level before drainage, shown by the second dotted black line in **Figures 2, 10.** It is almost certain that subglacial drainage initiated prior to that, however, given that continuous meltwater input to the lake would delay the visible drop in lake level by minutes to hours. This leads us to conclude that the timing of onset the red noise cluster (July 5th or 6th) is a better indicator of subglacial drainage than surface lake level observations. Given that the onset of the second (red) noise cluster also coincides with glacial surge recorded by GPS (Figure 10), the source of the spectral features in the second (red) noise cluster could alternatively be from till deformation, slip at the ice-bed interface, or from minute fractures in the ice due to increased flow and deformation. Although we cannot rule these possibilities out, the fact that the noise cluster trends continue long past the end of the ice acceleration suggests that this noise is more likely hydrologically sourced.

Figure 9 shows how the diurnal trends of some of the noise clusters differ substantially from the dataset as a whole. Diurnal trends within the clusters are of interest because of the connection between diurnal temperature fluctuations (red dashed line), melt water production, and glacial velocity—all possible seismic sources —at this and other temperate glaciers. The purple cluster tends to be active 2-4 hours after the hottest time of day; a lag time that could be controlled by the storage capacity/residence time of the glacier [I’m blanking on the better term for this]. The peak in the orange cluster precedes the hottest time of day by about 6 hours, yet is inactive in the evening . The blue, orange and purple clusters are mostly inactive during the lake drainage period from July 7th-15th, but the orange and purple clusters return with a prominent diurnal trend after drainage ends, around July 15th. The orange and purple clusters’ timings could contain the seismic signatures of two separate, stable modes of subglacial meltwater runoff (purple: high and/or unimpeded runoff; orange: low and/or impeded runoff), which are interrupted by the catastrophic flooding period. Subglacial conduit networks that were deformed or destroyed during the flood could reform post-flood, eventually exhibiting the same stable modes of subglacial meltwater and leading to a return of the purple and orange clusters. Going beyond temporal correlations, we see in Figure 8e that representative waveforms from the purple cluster have an overall significantly higher amplitude than those from orange clusters, consistent with seismicity from relatively higher water flow. Additionally, we see in spectra evidence that the purple clusters is related to higher water flow,…. [[…. is something I would like to say but I am not sure how]]]

Some of the clusters identified by this method could have been found using more straight-forward methods. The noise cluster associated with the onset of lake drainage (red cluster), for example, could have been identified by applying K-means directly to the spectra or the features depicted in **Figure 8 c-d.** The other noise clusters, however, are less distinct. Performing K-means on the noise spectra, for example, does not produce the orange and purple clusters. Additionally, since the phases aren’t aligned between records in time, performing K-means directly on the waveforms or spectrograms would also fail to produce these results. Therefore, SpecUFEx is advantageous in objectively discriminating subtle spectral features of seismic data that elude simpler detection techniques.

One of the novel aspects of this study is the reconstruction of the output of the unsupervised classification methods back into the frequency domain, allowing for direct physical interpretation of the results. For the noise clusters, the reconstructed spectra in **Figure 12** recover highly resolved details of the median-normalized spectra in the 5-40 Hz range. We hypothesize that this is in part because the noise waveforms (and spectrograms) are essentially stationary time series, and can be well represented in fingerprints by a single hidden Markov state frequently transitioning to itself through time. The noise waveforms themselves do not correlate in time, and so would be impossible to identify using simpler methods like cross- or auto-correlation

For the icequakes, the reconstructed spectra in **Figure 12** fail to emulate the actual spectra, but can still be instructive for deciphering the clustering logic. For example, the first (blue) icequake cluster in **Figure 6** has a sustained background noise throughout the record. That noise is likely represented by state 12 in the icequake fingerprints (Figure 11), which has a broad peak in amplitude from 30-50 Hz. A signal in this frequency range could be caused by turbulent water flow and sediment saltation in subglacial conduits; a hypothesis that is supported by the timing of the blue icequake cluster with respect to lake drainage (Figure 10).

The fine-scale coherence between waveforms in **Figure 6** demonstrates how SpecUFEx could be a useful tool for identifying repeating earthquakes. Rather than auto-correlating a dataset or manually choosing templates for cross correlation, SpecUFEx can identify groups of highly similar event waveforms in a semi-automated, unsupervised fashion. Additionally, cross-correlation coefficients can be impacted by factors such as azimuth and epicentral distance (e.g., Gao et al., 2021), but we show here that SpecUFEx is insensitive to such variations (see Figure 7). Future studies will aim to understand what physical aspects of the source and/or path are causing the subtle differences between waveforms and spectra in **Figure 6,** particularly the features indicated with the black arrows, and how they may relate to the overall hydrology and dynamics at Gornergletscher.

**Conclusion:**

We use an unsupervised feature extraction algorithm, SpecUFEx, PCA, and k-means clustering to characterize two months of glacial seismic data based on characteristic time-varying spectral features in spectrograms of icequakes and seismic “noise”. We compare results to ancillary geophysical data (temperature, precipitation, GPS displacement and lake level), and utilize a novel method for “reverse engineering” the output of SpecUFEx back into the frequency domain. We discover systematic differences in waveform and spectral properties that are at times highly subtle and would elude simpler machine learning or seismic analysis techniques. Based on our analysis, we propose an updated timeline for the glacial outburst flood and identify other possible seismic signatures of subglacial conduit reorganization. This work demonstrates the value of unsupervised feature extraction as a seismic monitoring and discovery tool, particularly in systems where complex fluid-driven processes occur.

Works cited

1. Bartholomaus, T. C., Amundson, J. M., Walter, J. I., O’Neel, S., West, M. E., & Larsen, C. F. (2015). Subglacial discharge at tidewater glaciers revealed by seismic tremor. Geophysical Research Letters, 42(15), 6391–6398. https://doi.org/10.1002/2015GL064590
2. Bellman R. Dynamic programming. Science. 1966 Jul 1;153(3731):34-7. doi: 10.1126/science.153.3731.34. PMID: 17730601.
3. Bergen, K. J., Johnson, P. A., De Hoop, M. V., & Beroza, G. C. (2019, March 22). Machine learning for data-driven discovery in solid Earth geoscience. Science, Vol. 363. <https://doi.org/10.1126/science.aau0323>
4. Carniel, R., & Guzmán, S. R. (1989). Machine Learning in Volcanology: A Review. Volcanoes - Updates in Volcanology [Working Title] Dr. Karoly Nemeth, 32(tourism), 137–144.
5. Chamarczuk, M., Nishitsuji, Y., Malinowski, M., & Draganov, D. (2019). Unsupervised learning used in automatic detection and classification of ambient-noise recordings from a large-n array. Seismological Research Letters, 91(1), 370–389. <https://doi.org/10.1785/0220190063>
6. Collins, D. N. (1989). Seasonal development of subglacial drainage and suspended sediment delivery to melt waters beneath an alpine glacier. Annals of Glaciology, 13(May), 45–50. <https://doi.org/10.1017/s026030550000762x>
7. Delaney, I., Bauder, A., Werder, M. A., & Farinotti, D. (2018). Regional and Annual Variability in Subglacial Sediment Transport by Water for Two Glaciers in the Swiss Alps. Frontiers in Earth Science, 0, 175. <https://doi.org/10.3389/FEART.2018.00175>
8. Eibl, E. P. S., Bean, C. J., Einarsson, B., Pàlsson, F., & Vogfjörd, K. S. (2020). Seismic ground vibrations give advanced early-warning of subglacial floods. Nature Communications, 11(1). https://doi.org/10.1038/s41467-020-15744-5
9. Gao, D., Kao, H., & Wang, B. (2021). Misconception of Waveform Similarity in the Identification of Repeating Earthquakes. Geophysical Research Letters, 48(13). <https://doi.org/10.1029/2021GL092815>
10. Garcia, L., Luttrell, K., Kilb, D., & Walter, F. (2019). Joint geodetic and seismic analysis of surface crevassing near a seasonal glacier-dammed lake at Gornergletscher, Switzerland. Annals of Glaciology, 60(79), 1–13. <https://doi.org/10.1017/aog.2018.32>
11. Gimbert, F., Tsai, V. C., Amundson, J. M., Bartholomaus, T. C., & Walter, J. I. (2016). Subseasonal changes observed in subglacial channel pressure, size, and sediment transport. Geophysical Research Letters, 43(8), 3786–3794. https://doi.org/10.1002/2016GL068337
12. Huss, M., Bauder, A., Werder, M., Funk, M., & Hock, R. (2009). Glacier-dammed lake outburst events of Gornersee, Switzerland. Mitteilungen Der Versuchsanstalt Fur Wasserbau, Hydrologie Und Glaziologie an Der Eidgenossischen Technischen Hochschule Zurich, 213, 65–84.
13. Holtzman, B. K., Paté, A., Paisley, J., Waldhauser, F., & Repetto, D. (2018). Machine learning reveals cyclic changes in seismic source spectra in Geysers geothermal field. Science Advances, 4(5), eaao2929. <https://doi.org/10.1126/sciadv.aao2929>
14. Jenkins II, W. F., Gerstoft, P., Bianco, M. J., & Peter, D. (2021). Unsupervised Deep Clustering of Seismic Data : Monitoring the Ross Ice Shelf , Antarctica. ESSOAr Preprint Archive, 1–40. https://doi.org/10.1002/ESSOAR.10505894.1
15. Klapuri, A., & Davy, M. (Eds.). (2007). Signal processing methods for music transcription, chapter 5. Springer Science & Business Media.
16. Karl  Pearson  F.R.S. . (1901). LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11), 559–572. <https://doi.org/10.1080/14786440109462720>
17. Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., & Gerstoft, P. (2019). Machine learning in seismology: Turning data into insights. Seismological Research Letters, 90(1), 3–14. <https://doi.org/10.1785/0220180259>
18. Lamb, O., Lees, J., Marin, L. F., Lazo, J., Rivera, A., Shore, M., & Lee, S. (2020). Investigating potential icequakes at Llaima volcano, Chile. Volcanica, 3(1), 29–42. <https://doi.org/10.30909/vol.03.01.2942>
19. Lever, J., Krzywinski, M., & Altman, N. (2017). Points of Significance: Principal component analysis. In Nature Methods (Vol. 14, Issue 7, pp. 641–642). Nature Publishing Group. <https://doi.org/10.1038/nmeth.4346>
20. Lindner, F., Walter, F., Laske, G., & Gimbert, F. Glaciohydraulic seismic tremors on an Alpine glacier. 14(1). <https://doi.org/10.5194/tc-2019-155>
21. Malfante, M., Dalla Mura, M., Metaxian, J. P., Mars, J. I., Macedo, O., & Inza, A. (2018). Machine Learning for Volcano-Seismic Signals: Challenges and Perspectives. IEEE Signal Processing Magazine, 35(2), 20–30. <https://doi.org/10.1109/MSP.2017.2779166>
22. Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nature Communications, 11(1), 1–12. <https://doi.org/10.1038/s41467-020-17591-w>
23. Park, Y., Mousavi, S. M., Zhu, W., Ellsworth, W. L., & Beroza, G. C. (2020). Machine-Learning-Based Analysis of the Guy-Greenbrier, Arkansas Earthquakes: A Tale of Two Sequences. Geophysical Research Letters, 47(6), 1–8. https://doi.org/10.1029/2020GL087032
24. Podolskiy, E. A., & Walter, F. (2016). Cryoseismology. In Reviews of Geophysics (Vol. 54, Issue 4, pp. 708–758). Blackwell Publishing Ltd. <https://doi.org/10.1002/2016RG000526>
25. Ren, C. X., Peltier, A., Ferrazzini, V., Rouet-Leduc, B., Johnson, P. A., & Brenguier, F. (2020). SUPP Machine Learning Reveals the Seismic Signature of Eruptive Behavior at Piton de la Fournaise Volcano. Geophysical Research Letters, 47(3), 1–11. https://doi.org/10.1029/2019GL085523
26. Riesen, P., Sugiyama, S., & Funk, M. (2010). The influence of the presence and drainage of an ice-marginal lake on the flow of Gornergletscher, Switzerland. Journal of Glaciology, 56(196), 278–286. <https://doi.org/10.3189/002214310791968575>
27. Röösli, C., Walter, F., Husen, S., Andrews, L. C., Lüthi, M. P., Catania, G. A., & Kissling, E. (2014). Sustained seismic tremors and icequakes detected in the ablation zone of the Greenland ice sheet. Journal of Glaciology, 60(221), 563–575. https://doi.org/10.3189/2014joG13j210
28. Ross, Z. E., Meier, M. A., Hauksson, E., & Heaton, T. H. (2018). Generalized seismic phase detection with deep learning. Bulletin of the Seismological Society of America, 108(5), 2894–2901. <https://doi.org/10.1785/0120180080>
29. Ross, Z. E., Yue, Y., Meier, M. A., Hauksson, E., & Heaton, T. H. (2019). PhaseLink: A Deep Learning Approach to Seismic Phase Association. Journal of Geophysical Research: Solid Earth, 124(1), 856–869. <https://doi.org/10.1029/2018JB016674>
30. Roux, P. F., Walter, F., Riesen, P., Sugiyama, S., & Funk, M. (2010). Observation of surface seismic activity changes of an Alpine glacier during a glacier-dammed lake outburst. Journal of Geophysical Research: Earth Surface, 115(3), 1–13. [https://doi.org/10.1029/2009JF001521](https://doi.org/10.1029/2009JF001535)
31. Seydoux, L., Balestriero, R., Poli, P., Hoop, M. de, Campillo, M., & Baraniuk, R. (2020). Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning. Nature Communications, 11(1). https://doi.org/10.1038/s41467-020-17841-x
32. Sick, B., Guggenmos, M., & Joswig, M. (2015). Chances and limits of single-station seismic event clustering by unsupervised pattern recognition. Geophysical Journal International, 201(3), 1801–1813. <https://doi.org/10.1093/gji/ggv126>
33. Steinmann, R., Seydoux, L., Beauce, E., & Campillo, M. (2021). Hierarchical exploration of continuous seismograms with unsupervised learning. May, 1–26.
34. Sugiyama, S., Bauder, A., Weiss, P., & Funk, M. (2007). Reversal of ice motion during the outburst of a glacier-dammed lake on Gornergletscher, Switzerland. Journal of Glaciology, 53(181), 172–180. <https://doi.org/10.3189/172756507782202847>
35. 2008\_Sugiyama\_Triggering and drainage mechanisms of the 2004 glacier-dammed lake outburst in Gornergletscher, Switzerland.pdf. (n.d.).
36. Sugiyama, S., Bauder, A., Riesen, P., & Funk, M. (2010). Surface ice motion deviating toward the margins during speed-up events at Gornergletscher, Switzerland. Journal of Geophysical Research: Earth Surface, 115(3). <https://doi.org/10.1029/2009JF001509>
37. Trugman, D. T., & Shearer, P. M. (2017). GrowClust: A Hierarchical clustering algorithm for relative earthquake relocation, with application to the Spanish Springs and Sheldon, Nevada, earthquake sequences. Seismological Research Letters, 88(2), 379–391. https://doi.org/10.1785/0220160188
38. Vore, M. E., Bartholomaus, T. C., Winberry, J. P., Walter, J. I., & Amundson, J. M. (2019). Seismic Tremor Reveals Spatial Organization and Temporal Changes of Subglacial Water System. Journal of Geophysical Research: Earth Surface, 124(2), 427–446. https://doi.org/10.1029/2018JF004819
39. Walter, F., Deichmann, N., & Funk, M. (2008). Basal icequakes during changing subglacial water pressures beneath Gornergletscher, Switzerland. Journal of Glaciology, 54(186), 511–521. <https://doi.org/10.3189/002214308785837110>
40. Walter, F. (2009). SEISMIC ACTIVITY ON GORNERGLETSCHER DURING GORNERSEE OUTBURST FLOODS. Dissertation, 31(1–2), 14–27. https://doi.org/10.1007/BF03322148
41. Walter, F., Clinton, J. F., Deichmann, N., Dreger, D. S., Minson, S. E., & Funk, M. (2009). Moment tensor inversions of icequakes on Gornergletscher, Switzerland. Bulletin of the Seismological Society of America, 99(2 A), 852–870. <https://doi.org/10.1785/0120080110>
42. Walter, F., Dreger, D. S., Clinton, J. F., Deichmann, N., & Funk, M. (2010). Evidence for near-horizontal tensile faulting at the base of gornergletscher, a Swiss alpine glacier. Bulletin of the Seismological Society of America, 100(2), 458–472. <https://doi.org/10.1785/0120090083>
43. Werder, M. A., Loye, A., & Funk, M. (2009). Dye tracing a jökulhlaup: I. Subglacial water transit speed and water-storage mechanism. Journal of Glaciology, 55(193), 889–898. <https://doi.org/10.3189/002214309790152447>
44. Werder, M. A., Schuler, T. V., & Funk, M. (2010). Short term variations of tracer transit speed on alpine glaciers. Cryosphere, 4(3), 381–396. https://doi.org/10.5194/tc-4-381-2010
45. Yoon, C. E., O’Reilly, O., Bergen, K. J., & Beroza, G. C. (2015). Earthquake detection through computationally efficient similarity search. Science Advances, 1(11). https://doi.org/10.1126/sciadv.1501057