Geophysical Journal International
Article In Press
Acceptation date : October 2024
https://doi.org/10.1093/gji/ggae361
https://archimer.ifremer.fr/doc/00914/102608/

Self-supervised learning of seismological data reveals new eruptive sequences at the Mayotte submarine volcano

Rimpot Joachim [1, *], Hibert Clément [1, 2], Retailleau Lise [3, 4], Saurel Jean-Marie [3], Malet Jean-Philippe [1, 2], Forestier Germain [5], Weber Jonathan [5], Stangeland Tord S [6], Turquet Antoine [6], Pelleau Pascal [7]

[1] Institut Terre et Environnement de Strasbourg (ITES), CNRS UMR 7063, Université de Strasbourg , 5 rue René Descartes, F-67084 Strasbourg , france
[2] Ecole et Observatoire des Sciences de la Terre (EOST), CNRS UAR 830, Université de Strasbourg , 5 rue René Descartes, F-67084 Strasbourg , France
[3] Institut de Physique du Globe de Paris, CNRS, Université Paris Cité , 1 rue Jussieu, F-75005 Paris , France
[4] Observatoire volcanologique du Piton de la Fournaise, Institut de Physique du Globe de Paris , 14 RN3 - Km 27, F-97418 La Plaine des Cafres, La Réunion , france
[5] Institut de Recherche en Informatique, Mathématiques, Automatique et Signal (IRIMAS), UR 7499, University of Haute-Alsace , F-68100 Mulhouse , france
[6] NORSAR , Gunnar Randers Vei 15, 2007 Kjeller ,Norway
[7] IFREMER, Centre de Bretagne â€“ Unité Géosciences Marines , 1625 route de Sainte-Anne, 29280 Plouzané , France

* Corresponding author : Joachim Rimpot, email address : jrimpot@unistra.fr

Abstract :

Summary Continuous seismological observations provide valuable insights to deepen our understanding of geological processes and geohazards. We present a systematic analysis of two months of seismological records using an AI-based Self-Supervised Learning (SSL) approach revealing previously undetected seismic events whose physical causes remain unknown but that are all associated with the dynamics of the Mayotte submarine volcano. Our approach detects and classifies known and new event types, including two previously unknown eruptive sequences displaying properties similar to other sequences observed at underwater and aerial volcanoes. The clustering workflow identifies seismic events that would be difficult to observe using conventional classification approaches. Our findings contribute to the understanding of submarine eruptive processes and the rare documentation of such events. We further demonstrate the potential of SSL methods for the analysis of seismological records, providing a synoptic view and facilitating the discovery of rarely observed events. This approach has wide applications for the comprehensive exploration of diverse geophysical datasets.

Keywords : Volcano seismology, Computational seismology, Machine learning, Neural networks, fuzzy logic, Persistence, memory, correlations, clustering

1 INTRODUCTION

Analyzing seismic signals is a central area of research in Earth Sciences. However, the origin of a substantial proportion of these signals remains poorly understood. This knowledge gap is particularly critical for seismological stations deployed to monitor geohazards such as glaciers, landslides and volcanoes. Seismological instruments deployed in these areas record thousands of events in a broad spectrum of magnitudes. Cataloging these events, and exploring their spatial and temporal distribution allows deepening our understanding of the dynamics and physics of these geological processes.

One of the primary objectives of the seismologists is to construct catalogs of seismic events. However, real-time monitoring and retroactive exploration of seismological time series requires substantial human resources. Artificial Intelligence provides solutions for exploring and extracting information from continuous and massive seismological records. Mousavi & Beroza (2023) and Kubo et al. (2024) reviewed the recent advances in earthquake seismology using various machine learning approaches. They particularly explored the current solutions for earthquakes catalog creation, from the detection and picking of events to the earthquake location. The common approach for classifying seismic signals consists of training machine/deep learning models with large quantities of labeled data which is time-consuming and sensitive to expert knowledge. The amount of labeled samples needed to train

those algorithms is most of the time not enough when studying other sources than regional and global earthquakes. Although good performances are achieved in the creation of high-magnitude earthquake catalogs (Liu et al. 2020; Mousavi et al. 2020; Yang et al. 2024), the approaches are still not fully operational for constructing low-magnitude micro-seismic event catalogs. Creating micro-seismic catalogs further faces challenges in achieving consensus on nomenclature and a standardized labeling process, especially for unidentified seismic events. This is particularly true for volcanoes where seismic signals describing similar source mechanisms can look different from one volcano to another, and even for the same volcano over several time periods.

A large underwater eruption started in Mayotte (Comoros archipelago, Indian Ocean) in 2018, 50 km east from the island coast (Fig. 1a) (Cesca et al. 2020; Lemoine et al. 2020; Feuillet et al. 2021). The activity started on 2018 May 10 with a seismic crisis that culminated with a Mw 5.9 earthquake. It was followed by deflation and eastward displacement of the island and the occurrence of very-long period events. Because the eruption was submarine it could not be observed easily and remote observations have been employed to understand the volcano edification (Rinnert et al. 2019; Saurel et al. 2021, land surface seismometers and ocean bottom seismometers). Numerous earthquakes linked to the volcanic system have been recorded. These events include Volcano-Tectonic earthquakes (VT), Long-Period earthquakes (LP) and Very-Long Period earthquakes (Lavayssière et al. 2022; Retailleau et al. 2022b; Laurent, 2023, respectively). These events, apart from the early stages of activity and the first magma propagation towards the surface, have been located deeper than 20 km. The eruption also generated numerous hydro-acoustic signals (HA) that can be used to understand the volcanic activity (Saurel et al. 2022, e.g. monitor the eruption flows,). A wide range of HA signals are generated during eruptions but only a few submarine volcanoes have been monitored leading to a partial knowledge of the diversity of those signals (Chadwick, et al. 2019; Caplan-Auerbach et al. 2017; Wilcock et al. 2016, e.g.). Since the start of the eruption, several offshore prospecting campaigns have been conducted to take in-situ measurements and in particular to deploy Ocean Bottom Seismometers (OBS), leading to a comprehensive catalog for the period of October and November 2019.

Constructing manually this first event catalog was labor-intensive and subjective, as experts discovered new types of seismic sources while sequentially analyzing the data. To explore seismological records, unsupervised machine learning algorithms can be employed. Retailleau et al. (2022a) proposed a wrapper for earthquake monitoring based on a deep-neural-network automatic phase picker, which is used operationally for the Mayotte monitoring. Other developments in unsupervised algorithms showed promise in identifying seismic signals. Mousavi et al. (2019) employed deep embedded clustering to distinguish teleseismic events from local earthquakes, while Jenkins et al. (2021) used deep neural networks coupled with a Gaussian Mixture Model (GMM) to cluster impulsive seismic

events recorded at the Antarctica Ice Shelf. Following the same idea, Hu et al. (2024) proposed to explore detected events on the Dalk glacier, East Antarctica, with deep auto-encoders and GMM clustering on the resulting embedding. They highlighted three main event classes linked with the wind activity, the basal slip of the glacier and the thermal variations. Recently, Kinzel et al. (2024) proposed a Siamese-based workflow for clustering time series of pre-detected seismic events containing icequakes, earthquakes and spikes. The authors applied a centroïd-based clustering (k-means) directly on the feature space in order to capture meaningful learned information. These approaches rely on a priori detected events and, in the case of GMM and k-means, necessitate the selection of a specific number of clusters which might prevent the discovery of new and rare seismic events. Moreover, seismic signals have to be extracted from the seismological records with detection algorithms (Allen, 1982; Baillard et al. 2014; Zhu & Beroza, 2019; Mousavi et al. 2020, respectively, STA/LTA, PSPicker, PhaseNet and EQTransformer) which introduce selection bias in the catalogs as they limit the variability in both amplitude and type of the detected events (Yoon et al. 2015). To avoid a priori event detection and enable the analysis of complete seismic traces, Seydoux et al. (2020) proposed to transform the continuous seismic signal with deep scattering-network and perform clustering on the model outputs. By analyzing day-long windows of three-components seismic signals recorded at a single station, they recovered precursory signals of the Nuugaatsiaq (Greenland, June 2017). Using the same approach Steinmann et al. (2022) analyzed two day-long three-components seismological observations along the North Anatolian Fault (Turkey) leading to the clustering of events within which earthquakes were identified.

Due to the predominance of background noise and the relative rarity of the seismic signals of interest in continuous seismic records, designing an unsupervised classification approach to detect those signals is challenging. In this work, we propose a workflow (Fig. 2a) that combines Self-Supervised Learning (SSL) with dimension reduction and clustering to offer a complete and exhaustive exploration of the seismological data acquired during the MAYOBS6 and MAYOBS7 campaigns. SSL is the process of training models to produce meaningful representations using unlabeled data. In other words, SSL starts with finding correlations between the data in order to auto-label them thus allowing the creation of generalist models that can be further fine-tuned for many domain applications. Among the various SSL algorithms available (Grill *el al.*, 2020; Caron et al. 2020; Zheng et al. 2021, respectively, BYOL, SwAV or ReSSL), we used the Simple Siamese (SimSiam) network (Chen & He, 2021), which architecture is displayed in Fig. 2b. The SimSiam network is a robust deep learning method designed to maximize the similarity between two representations of the same image; it possesses the advantage of not requiring a large amount of data to be trained to recognize efficiently similar events

(Chen & He, 2021). This statistical property makes it more competitive than other unsupervised clustering algorithms for seismological applications.

2 DATA

During MAYOBS6, MAYOBS7 and MAYOBS8 campaigns (Guyavarch, 2019; Pelleau, 2019a; Pelleau, 2019b), in October and November 2019, short-period OBS were deployed close (1 km) to the active volcanic lava flows. They are composed of a 3-channels 4.5 Hz geophones and an hydrophone. The OBS were deployed in two stages: IF07C from 1 October 2019 to 25 October 2019 and IF07D from 26 October 2019 to 19 November 2019. They were both deployed in free fall from the ocean water surface making their exact location unknown and different up to a few hundred meters.

A group of seismological experts scanned manually the seismic signals recorded by the vertical components of the OBS for the two periods and produced a first seismic catalog (SefraN, Beauducel et al. 2020; Saurel et al. 2021). The goal was to explore the dataset and identify unusual events. The catalog is divided in four categories: VT events, HA events, single detected events (SDE) and events of unknown origin. The VT are high-frequency seismic signals with most of the signal energy concentrated in the 1 to 20 Hz frequency band, and which exhibit a characteristic exponential decrease of the amplitude after the maximum of the envelope (Fig. 1b). VT are thought to be linked to the deep volcanic system, such as magma propagation in the volcanic edifice (Lahr et al. 1994; Duputel, Lengliné & Ferrazzini 2019; Wilcock et al. 2016; Taisne et al. 2011). Their duration is usually short, ranging from 1 to 20 s. The HA seismic signals exhibit complex waveforms. They usually start with a very impulsive and high-frequency ($>$ 20 Hz) energy burst (Fig. 1c), followed by one or several phases with energy in the 1 to 20 Hz frequency band. Their origin is subject to discussion due to the depth of the volcano and the limited studies on the subject but these events might be caused by fast cooling lava flows (Saurel et al. 2022). SDE are short events only detected at a single station. Their origin is therefore uncertain however the highly energetic onset observed on the seismic and acoustic sensors could be explained by events occurring near the volcano surface (shallow VT) or by short HA sources at the surface.

The SimSiam based workflow uses images as inputs. To transform the seismological data into images, we extract the seismic signals over non-overlapping 30 s-duration windows. We have chosen this duration in order to fully capture most of the VT and HA events included in the first catalog. Images are constructed by combining a temporal and a spectral representation of the seismic signals. The representation in the temporal domain is the seismic signal recorded on the vertical component of the OBS, filtered between 1 and 30 Hz and detrended from the daily mean. The filtering allows removing high-frequency noise and increasing the signal-to-noise ratio for the seismic source of in-

terest. The representation in the spectral domain is a spectrogram which shows the evolution of the seismic signal frequency content with time and is computed from the non-filtered seismic traces to include frequencies greater than 30 Hz (Fig. 3a). The spectrograms are constructed using short time Fourier transforms, with windows of 0.12 s length, an overlap of 90 %, a sampling of 2048 points for the frequency vector and no smoothing. The dataset consists of a total of 138,712 images for a period of 50 days. Labeled events in the catalog represent a few part of these images. 2008 images correspond to labeled VT events, 453 images to HA events and 1173 images to SDE.

3 METHODS

3.1 Self-Supervised Learning approach

We propose a SSL clustering workflow (Fig. 2a) based on the SimSiam network (Chen & He, 2021) based on contrastive learning of two transformed views of the same input image (Fig. 2b). It aims to maximize their similarity, using their common intrinsic information.

The network is composed of several encoders f and a prediction layer h (Fig. 2b). Encoders are deep neural networks, composed of a succession of layers whose purpose is to extract relevant features from the input image thanks to filters organized in convolution layers. The prediction layer serves to adapt the output of the first encoder to better match the output of the second encoder, thereby maximizing the similarity between the two augmented views. The prediction layer plays a crucial role in ensuring that the representations learned are useful and informative. The encoders correspond to the ResNet-18 backbone (i.e. feature-extracting network) (He et al. 2016), a 18 layers deep convolutional neural network, and are initialized with random weights. The weights affect the filters' outputs and are updated at each epoch to optimize the network's performances at maximizing the similarity of the two augmented views.

The SimSiam network works with two parallel encoders. The first one is the input to the prediction layer; the second one applies a stop-gradient operation, consisting in blocking the back-propagation of the gradient (Fig. 2b). This approach avoids collapsing the model into a constant value with each image being described with the exact same feature values. The input images are first resized into square images of 256 x 256 pixels. Then, data augmentation is performed on each input image to create pairs of transformed images. The data augmentation consists of random rotation, horizontal and vertical flip and random changes of brightness, contrast, saturation and hue. The model does not aim to include any physics in these transformations; e.g. a flipped spectrogram remains understandable for seismologists even if it losses its physical sense. The idea behind data augmentation is to train a model being robust to variations and transformations in the inputs. In this way, the model focuses

on meaningful information, which are colors and invariant shapes on the images, useful for events discrimination. Then, we use a batch size of 64 pairs of augmented views to train the model. The negative cosine similarity (Chen & He, 2021), is used for measuring the dissimilarity between two vectors in a vector space. It is defined with the equation:

$$D(p_1, z_2) = -\frac{p_1}{||p_1||_2} \cdot \frac{z_2}{||z_2||_2}, \tag{1}$$

where $||\cdot||_2$ represents the l_2-norm. p_1 is defined as the output vector of the transformation of x_1 through f and h as $x_1 = h(f(x_1))$ while z_2 is defined as the output vector of the application of f on x_2 as $z_2 = f(x_2)$. Then, the symmetric loss is computed for each image as:

$$L = \frac{1}{2}D(p_1, stopgrad(z_2)) + \frac{1}{2}D(p_2, stopgrad(z_1)), \tag{2}$$

where $stopgrad(\cdot)$ represents the stop-gradient operation applied on one side of the network (Fig. 2b). Finally, the total loss is defined as the average of all the image loss, leading to a minimum possible value of -1.

Once trained, we use the encoder part of the network with the learned weights to encode all the original input images. The number of features is set to 512. This results in a group of 512 value vectors of features which is called a 512 dimensions embedding.

3.2 Dimension reduction and clustering

For the visual evaluation of the clustering, we use a dimension reduction techniques to transform the 512 dimensions embeddings into a two dimensions embeddings. We explored three algorithms: t-SNE (Van der Maaten & Hinton, 2008), Principal Component Analysis (PCA) and Uniform Manifold Approximation and Projection (McInnes, Healy & Melville 2018, UMAP). t-SNE aims to preserve the neighborhood of samples while reducing the dimension of the embedding. This allows neighbor samples in the 512 dimensions space to be described in two dimensions with close coordinates, preserving their neighborhood. PCA creates new features that are linear combinations of the original ones and that are uncorrelated with each others. In this way, PCA extracts information from the original embedding to represent the data in a lower dimension space and highlight trends specific to different groups of samples. UMAP is a non-linear dimension reduction method based on algebraic topology (McInnes, Healy & Melville 2018). In opposition to PCA, UMAP does not suggest linear relationships between features, which makes it possible to capture complex structures in high dimensions embeddings. It also aims to preserve the topology of the high dimension embedding, which results in a preservation

of local and global structures. Compared to t-SNE, UMAP performs better in the global structure preservation, giving more sense to sample distances in the two dimensions embeddings (McInnes, Healy & Melville 2018). However, with t-SNE, local structures are well preserved which makes easier the identification of clusters (Van der Maaten & Hinton, 2008).

While t-SNE, PCA and UMAP share a common parameter (*number of components*) driving the output size, they work differently and are controlled with distinct parameters. t-SNE is mainly dependent on the *perplexity* and the *early exaggeration* that will act on the preservation of local and global structures and the distance between clusters. PCA can be tuned with the *svd solver* parameters which defines the method used to compute the Singular Value Decomposition. Finally, UMAP uses a *number of neighbors* to adjust the focus on local or global structure preservation, a *metric* used to calculate the distance between the samples and a *number of epochs* used to optimize the embedding dimension reduction.

4 RESULTS

The procedure for model training consists in 15 epochs as it converges quickly to a loss value close to -1 with a collapse level below 0.1. We thus obtain the 512 dimensions embeddings of the encoded dataset.

4.1 Dimension reduction methods and clustering

The comparison of the dimension reduction techniques (Fig. 4) shows that t-SNE and UMAP provide a sparser two dimensions space than PCA, whether using a default or tuned parameterization. PCA has difficulties in two dimensions to separate samples from the stations IF07C and IF07D. t-SNE, PCA and UMAP parameterization are tuned in order to favor local structures and small dense groups of samples as these groups could represent consistent clusters . For t-SNE we use a *perplexity* of 40 and an *early exaggeration* of 1. PCA is tuned using *svd solver* = *arpack* and UMAP is optimized with *number of neighbors* = *10*, *metric* = *euclidean*, *number of epochs* = *100* and *init* = *random*. By comparing tuned parameterizations, we observe that t-SNE yields more scattered patches than UMAP for which the samples corresponding to the IF07D OBS remain grouped in a single patch (Fig. 4d, 4f). This is promising as it means that the t-SNE transformation is able to isolate groups of samples that can be considered afterward as clusters. t-SNE is also able to create patches consistent with the pre-existing knowledge by creating patches of almost only VT events. Thanks to its ability to preserve the neighborhood of the embedded samples, we thus use t-SNE as dimension reduction technique.

Cluster extraction is the last step of the workflow. Density-Based Spatial Clustering of Applica-

tions with Noise (Ester et al. 1996, DBSCAN) aims at clustering together points in the same neighborhood as long as they have a certain minimal number of samples with a distance lower than epsilon around them. If points are isolated, the algorithm assumes that the points are noise. Therefore, DBSCAN is a relevant technique to cluster the t-SNE patches delineations. DBSCAN is mainly sensitive to the *epsilon* and the *minimum number of samples* parameters. The values are respectively set to 0.51 and 10 as this combination allows the best match with the t-SNE delineations (Fig. 3).

4.2 Clusters analysis

Plots of input images based on their t-SNE transformed embeddings highlight similarities between groups of samples as displayed in Fig. 3(a). The processing yielded 95 clusters and 202 isolated samples (Fig. 3b). In Fig. 3(b), each point represents a 30 s duration window of seismic signal. The clusters are visually well defined with clear separations. The analysis identified clusters with 21,086 events for the largest cluster and 10 events for the smallest cluster. The median size of a cluster is 341 events. We obtained 7 clusters with more than 5000 samples and 36 clusters with less than 100 samples.

The manually labeled seismic events are grouped into a small number of clusters and are not randomly scattered in the t-SNE 2-dimensions embedding. This suggests that the clustering process is consistent (Fig. 3c). VT events are aggregated mainly in five clusters (clusters 3, 12, 39, 63, 94), while SDE in one (cluster 8). Some clusters agglomerate both VT events and SDE (clusters 0, 1, 6 and 59) suggesting proximity between these two types of seismic signals. The SDE are mainly observed on the IF07C records. The few labeled HA events are distributed in the clusters 51, 55-57, 61 and 65. As shown on Fig. 3(c), all the clusters containing known events include a large proportion of new events. We also observe that many clusters present the particularity of being composed by events only recorded at one of the two OBS station. This might be attributed to various factors, including their potential variations in distance from the seismic sources and the local deployment conditions. These conditions may include differences in seafloor properties beneath the stations or variations in tilt and orientation of the sensor. Nevertheless, several clusters (3, 4, 7, 9, 11, 12, 39, 41 and 53) group events recorded by the two stations over the whole time period.

Some clusters are correlated in time as shown by the number of events per hour for these clusters (Fig. 5). The temporal correlations between clusters seem to be organized around two sequences: the first sequence occurs on the 15 October 2019 and the second starts on the 31 October 2019. During the first sequence, clusters 3, 4, 11, 39 and 53 are observed. In the second sequence, cluster 53 is initially observed, followed by clusters 11, 12, 39, and 51 in succession. Cluster 53 exhibits seismic signals with all the features of eruptive volcanic tremors which are high energy continuous seismic

signals (hereinafter referred to as ”high-energy tremor”) that can last from a few dozens of minutes up to several months. Cluster 4, 51 and 68 also exhibit complex intertwined or continuous low-frequency (< 20 Hz) seismic signals (hereinafter referred to as ”tremor-like”), which form tremor-like sequences albeit with less energy in the highest frequency band than observed in cluster 53. Clusters 12, 39, and 64 encompass highly energetic events with duration exceeding 10 seconds and waveforms with multiple phases (hereinafter referred to as ”high-energy long-duration multiphase events”). Cluster 3 gathers seismic signals identified as highly energetic VT seismic signals.

In order to gain a deeper understanding of the physical processes underlying the clustered sequences depicted in Fig. 5, we explore days of data surrounding these sequences (Fig. 6). The groups called VT, HA, VT/HA and OTHER include all the clusters not considered in the Fig. 3(b) and were labeled following manual inspection of each cluster and a priori knowledge of the seismic events. The VT/HA group contains events whose sources have not been clearly identified because they combine seismic signals with VT or HA features, or a hybrid of these features.

With this new representation, we are able to provide a phenomenological description of the two eruptive sequences. The first sequence starts on 2019 October 14 with the increase in the occurrence of VT events. Subsequently, around noon on October 15, we observe the activation of the cluster 4 (tremor-like), followed by clusters 3 (VT-like), 11 (tremor-like) and 53 (high-energy tremor). This marks the beginning of a period in which VT, HA, and VT/HA events occur, with a notable increase in HA events. Finally, the sequence ends with the activation of clusters 3 and 4, along with the emergence of cluster 39 (high-energy long-duration multiphase events) followed by a drastic decrease in VT and HA activity. The second sequence exhibits a similar overall pattern. It initiates on October 31, marked by an increase in the occurrence of VT earthquakes and VT/HA events. Within this period, two short episodes of tremors are observed. The first is mainly composed of events from cluster 68, while the second (starting on 2019 November 2) is characterized by the activation of clusters 3, 4, 11, 12 (high-energy long-duration multiphase events), 39, and 53. Approximately 12 hours later, synchronous activity is observed among clusters 11, 12, 39, 51 (tremor-like signals), 64 (high-energy long-duration multiphase events), and 68 (tremor-like signals). Throughout this sequence, the clusters of tremor-like signals and high-energy long-duration multiphase events dominate the activity. The number of events per hour in clusters 11, 12, 39, 51 and 64 abruptly decreases after an episode dominated by events from clusters 4 and 39 on November 3, which is clearly discernible on the spectrogram (Fig. 6e). Following this episode, the activity transitions to being primarily dominated by VT events. The entire sequence concludes rapidly after noon on 2019 November 4.

5 DISCUSSION AND CONCLUSION

The proposed work aims at clustering continuous seismic data through a contrastive learning of images representing the spectrogram associated to 30 s length time series. Our choice of using images is motivated by the wish to reproduce the same approach as handmade seismic and micro-seismic catalog. It therefore justifies the choice of SSL approaches coming from the field of computer vision. Using images also provides access to a synoptic and comprehensive representation (Fig. 3a) of the dataset. Siamese networks for time series input have to be further explored as time series contain more information, which could be useful to achieve high clustering performance for more complex datasets, such as dense nodal network datasets. In fact, these datasets typically contain seismic information from several dozen stations. This dramatically increases the number of windows to be analyzed (several million) and increases the possibility of simultaneous events. Finally, some sites such as landslides present a high noise level, regrouping the faunal activity but also the human activity, resulting in difficulties to observe low-magnitude events that could be masked by the high-energy noise level. For these datasets, new representations of the input data need to be explored to take into account the spatial information.

The objective of the proposed clustering workflow (t-SNE + DBSCAN) is not to achieve a one-to-one correspondence between the number of clusters and the number of source mechanisms. Instead, it aims to obtain coherent clusters of noise and events, encompassing the complete dataset. Consequently, scientists are only required to identify several dozen clusters, as opposed to laboriously labeling thousands of individual events. It also offers the possibility of having sub-clusters of a pre-defined group of events.

Our SSL-based exploration of two months of seismological records allows discovering a wealth of new seismic events potentially associated with the volcanic activity of the Fani Maoré submarine volcano. In particular, we identified two sequences starting with the increasing occurrence of VT events and followed by complex seismic signals including tremors (Fig. 6). During the first sequence, we observed an increase in the number of HA events, potentially indicating shallow volcanic activity, which may result from lava-water interactions (Fig. 6c). The sequence concluded abruptly with low-amplitude tremor activity and the emergence of highly energetic, long-duration events, resembling to the dynamics seen in surface eruptions (McNutt and Nishimura, 2008), particularly those at effusive volcanoes like Piton de la Fournaise (Peltier et al. 2021). These eruptions typically involve seismo-volcanic crises reflecting deep magma migration, followed by tremors and degassing indicative of possible lava extrusion, and ultimately end with surface activity. Similar observations align with data from the 2015 eruption of Axial Seamount underwater volcano (Tepp & Dziak, 2021). The second sequence, starting on October 31, 2019, is more complex (Fig. 6f). It started with an in-

crease in VT and HA events, followed by tremors of varying energy levels close to those observed in the first sequence. The events marking the end of the first sequence reappeared. Approximately twelve hours later, clusters of high-energy, long-duration multiphase events and tremors are observed simultaneously. Interpretation of the 36 hours period between 2 November 2019, and 3 November 2019, presented challenges due to the difficulty in interpreting high-energy, long-duration multiphase signals. These events displayed characteristics of VT and HA signals but had longer durations and significantly higher amplitudes. Furthermore, these events were intertwined with tremors. Without detailed information on station site conditions, we cannot rule out the possibility of instrumental noise such as ocean currents, local vibrations and resonances. However, given the features of these seismic signals and their chronological appearance, we believe that these clusters are associated with ongoing eruptive activity.

Cui, Li & Huang (2021) studied a four-month period of seismic events in the vicinity of the Kilauea volcano that preceded a series of eruptions and collapse events. In their characterization of the seismicity, the authors consider a hybrid class whose frequency content is a mixture of long-period events and VTs. Since this study relies on a pre-computed catalog that includes the locations of the events, the authors hypothesize that the hybrid class is a mixture of fluid resonance and shear failure. The VT/HA events identified in the present article do not appear to be the same type of hybrid signals as those defined by Chouet et al. (2013), McNutt & Roman (2015) and Cui et al. (2021). A detailed study of the frequency content of the identified VT/HA events, coupled with the location, would definitely help to better interpret these complex signals and may have implications for their characterization in other volcano-seismological studies. Obtaining accurate locations of the events based on a single OBS is not possible considering the uncertainty of the OBS locations and their coupling with the seafloor. Including seismological stations at land will not change the fact that part of events are only detected in the OBS, which are deployed close to the volcano area. Investigating higher-energy events that could be recorded by land stations is therefore reserved for future studies.

The documentation of underwater volcanoes is crucial for understanding submarine eruptive processes and direct observations (e.g., Embley et al. 2014; Tepp & Dziak, 2021) are rarely available. The proposed SSL and clustering workflow highlights sequences of seismic events that would have been difficult to observe on the raw seismic signals and on the spectrograms as they were occurring in a long and complex sequence of intertwined seismic signals. Additionally, our analysis shows that the volcano activity over this period occurred intermittently in paroxismic sequences with complex dynamics. Further analysis will enable to understand and model the processes underlying the different clusters of observed signals and explore other periods of time to find similar or new seismic signals. The SSL workflow does not only provide a comprehensive and synoptic view of the dataset (Fig.

3a) but also facilitates the identification of rare seismic events, crucial for understanding the natural phenomena. The ability to capture both overall trends and short-lived anomalies is what makes these AI-based approaches compelling for seismological data processing and also for other geophysical time series exploration.

The scope of the SSL approach extends to various applications including the exploration of numerous historical and recent datasets, such as for studying other volcanoes, landslides, glaciers, faults and surface/underground aquifers that generate a wide and complex variety of seismic signals. The organizational and visualization capabilities of the approach enable a systematic exploration of continuous seismic data which also simplifies the analysis of dense instrumental arrays that demand extensive labeling efforts from experts. The approach might also be used to extract coherent noise periods for seismic studies based on background noise. The approach only requires expert inputs for labeling clusters rather than individual events which saves many human-labeling hours. Moreover, pertinent results can be obtained with a limited number of epochs. This allows similar studies to be conducted without requiring high computing resources. Lastly, as demonstrated in this study by considering several time periods and seismological stations, the SSL approach also paves the way for transfer learning for event class identification in seismology.

ACKNOWLEDGMENTS

This work is funded by Agence Nationale de la Recherche (ANR) - Project ANR-20-CE01-0006 ”HighLand: Landslides and Climate Change in highly Sensible Environments: Seismology, Earth Observation and Artificial Intelligence”. Since June 2019, the Mayotte monitoring observatory is funded through the REVOSIMA program (REVOSIMA, 2021) by French Ministry of Environment and Ecological Transition (MTECT) and the Ministry of Overseas Territories (MIOM), with the support of the Ministry of Research (MESR) and of the Centre national de la recherche scientifique (CNRS) and Institut National des Sciences de l’Univers (INSU). The authors acknowledge the High Performance Computing Center of the University of Strasbourg for supporting this work by providing scientific support and access to computing resources. Part of the computing resources were funded by the Equipex Equip@Meso project (Programme Investissements d’Avenir) and the CPER Alsacalcul/Big Data. The authors would like to thank Zefeng Li, the anonymous reviewer and the editor for helping to improve the quality and impact of the manuscript.

The authors declare that they have no competing interests.

Author contributions:

Conceptualization: SSL : JR, CH, TSS, AT Mayotte : CH, LR, JMS Methodology: JR, CH, JPM, GF, JW Investigation: CH, JR, LR, JMS, PP Visualization: JR, CH, LR, JMS Supervision : CH Writing—original draft: JR, CH, LR, JMS Writing—review and editing: JR, CH, LR, JMS, JPM, GF, JW, TSS, AT, PP

The data used to train and test the model as well as the scripts are available on the Zenodo repository https://doi.org/10.5281/zenodo.8189838.

REFERENCES

Allen, R., 1982. Automatic phase pickers: Their present use and future prospects. Bulletin of the Seismological Society of America, 72 (6B), S225–S242.

Baillard, C., Crawford, W. C., Ballu, V., Hibert, C., Mangeney, A., 2014. An automatic kurtosis-based p-and s-phase picker designed for local seismic networks. Bulletin of the Seismological Society of America, 104 (1), 394–409.

Beauducel, F. et al., 2020. Webobs: The volcano observatories missing link between research and real-time monitoring. Frontiers in Earth Science, 8. doi:10.3389/feart.2020.00048

Berthod, C. et al., 2022. Temporal magmatic evolution of the fani maoré submarine eruption 50 km east of mayotte revealed by in situ sampling and petrological monitoring. Comptes Rendus. Géoscience, 354 (S2), 1–29.

Caplan-Auerbach, J., Dziak, R., Haxel, J., Bohnenstiehl, D., Garcia, C., 2017. Explosive processes during the 2015 eruption of axial seamount, as recorded by seafloor hydrophones. Geochemistry, Geophysics, Geosystems, 18 (4), 1761–1774.

Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., Joulin, A., 2020. Unsupervised learning of visual features by contrasting cluster assignments. Advances in neural information processing systems, 33 , 9912–9924.

Cesca, S. et al., 2020. Drainage of a deep magma reservoir near mayotte inferred from seismicity and deformation. Nature Geoscience, 13 (1), 87–93.

Chadwick Jr, W. W., Rubin, K. H., Merle, S. G., Bobbitt, A. M., Kwasnitschka, T., Embley, R. W., 2019. Recent eruptions between 2012 and 2018 discovered at west mata submarine volcano (ne lau basin, sw pacific) and characterized by new ship, auv, and rov data. Frontiers in Marine Science, 495.

Chen, X., He, K., 2021. Exploring simple siamese representation learning. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 15750–15758.

Chouet, B. A., Matoza, R. S. 2013. A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. Journal of Volcanology and Geothermal Research, 252, 108–175. https://doi.org/10.1016/j.jvolgeores.2012.11.013

Cui, X., Li, Z., Huang, H. 2021. Subdivision of seismicity beneath the summit region of Kilauea volcano: Implications for the preparation process of the 2018 eruption. Geophysical Research Letters, 48, e2021GL094698. https://doi.org/10.1029/2021GL094698

Duputel, Z., Lengliné, O., Ferrazzini, V., 2019. Constraining spatiotemporal characteristics of magma migration at piton de la fournaise volcano from pre-eruptive seismicity. Geophysical Research Letters, 46 (1), 119–127. doi: 10.1029/2018gl080895

Embley, R. W. et al., 2014. Eruptive modes and hiatus of volcanism at west mata seamount, nelau basin: 1996–2012. Geochemistry, Geophysics, Geosystems, 15 (10), 4093–4115.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise, in KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, August 1996, pp. 226–231.

Feuillet, N. et al., 2021. Birth of a large volcanic edifice offshore mayotte via lithosphere-scale dyke intrusion. Nature Geoscience, 14 (10), 787–795. doi: 10.1038/s41561-021-00809-x

Grill, J.-B. et al., 2020. Bootstrap your own latent-a new approach to self-supervised learning. Advances in neural information processing systems, 33 , 21271–21284.

Guyavarch, P., 2019. MAYOBS6 cruise, RV LCT Ylang. doi: https://doi.org/10.17600/ 18001661

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 770–778.

Hu, Y., Li, Z., Fu, L., Liu, X., 2024. Environment-modulated glacial seismicity near Dålk Glacier in East Antarctica revealed by deep clustering. Journal of Geophysical Research: Earth Surface, 129, e2023JF007593. https://doi.org/10.1029/2023JF007593

Jenkins, W. F., Gerstoft, P., Bianco, M. J., Bromirski, P. D., 2021. Unsupervised deep clustering of seismic data: Monitoring the Ross Ice Shelf, Antarctica. Journal of Geophysical Research: Solid Earth, 126 (9), e2021JB021716.

Kinzel, L., Fromm, T., Schlindwein, V., and Maass, P., 2024. Unsupervised Deep Feature Learning for Icequake Discrimination at Neumayer Station, Antarctica. Seismological Research Letters 95, 1834–1848, doi: 10.1785/0220230078

Kubo, H., Naoi, M., Kano, M., 2024. Recent advances in earthquake seismology using machine learning. Earth Planets Space 76, 36 (2024). https://doi.org/10.1186/s40623-024-01982-0

Lahr, J. C., Chouet, B. A., Stephens, C. D., Power, J. A., Page, R. A., et al., 1994. Earthquake classification, location, and error analysis in a volcanic environment: Implications for the magmatic system of the 1989–1990 eruptions at redoubt volcano, alaska. Journal of Volcanology and Geothermal Research, 62 (1-4), 137–151.

Laurent, A., 2023. Détection, localisation et caractérisation des signaux très basses fréquences dans la crise sismique de Mayotte (2018-2022). PhD thesis, Univ. Paris-Cité, Institut de Physique du Globe de Paris.

Lavayssière, A., Crawford, W. C., Saurel, J.-M., Satriano, C., Feuillet, N., Jacques, E., Komorowski, J.-C., 2022. A new 1d velocity model and absolute locations image the mayotte seismo-volcanic region. Journal of Volcanology and Geothermal Research, 421 , 107440.

Lemoine, A. et al., 2020. The 2018–2019 seismo-volcanic crisis east of mayotte, comoros islands: seismicity and ground deformation markers of an exceptional submarine eruption. Geophysical Journal International , 223 (1), 22–44.

Liu, M., Zhang, M., Zhu, W., Ellsworth, W.L., Li, H., 2020. Rapid Characterization of the July 2019 Ridgecrest, California, Earthquake Sequence From Raw Seismic Data Using Machine-Learning Phase Picker. Geophysical Research Letters 47. https://doi.org/10.1029/2019gl086189

McInnes, L., Healy, J., Melville J., 2018. Umap: Uniform manifold approximation and projection for dimension reduction. *In: arXiv preprint arXiv:1802.03426.*

McNutt, S. R., Nishimura, T., 2008. Volcanic tremor during eruptions: temporal characteristics, scaling and constraints on conduit size and processes. Journal of Volcanology and Geothermal Research, 178 (1), 10–18.

McNutt, S. R., Roman, D. C. 2015. Chapter 59 - volcanic seismicity. In H. Sigurdsson (Ed.), The encyclopedia of volcanoes (2nd ed., pp. 1011–1034). Academic Press. https://doi.org/10.1016/B978-0-12-385938-9.00059-6

Mousavi, S.M., Beroza, G.C., 2023. Machine Learning in Earthquake Seismology. Annual Review of Earth and Planetary Sciences 51, 105–129.. https://doi.org/10.1146/annurev-earth-071822-100323

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), 3952.

Mousavi, S. M., Zhu, W., Ellsworth, W., Beroza, G., 2019. Unsupervised clustering of seismic signals using deep convolutional autoencoders. IEEE Geoscience and Remote Sensing Letters, 16 (11), 1693–1697.

Pelleau, P., 2019. MAYOBS7 cruise, RV LCT Ylang. doi: https://doi.org/10.17600/18001673

Pelleau, P., 2019. MAYOBS8 cruise, RV LCT Ylang. doi: https://doi.org/10.17600/18001684

Peltier, A. et al., 2021. Volcano crisis management at piton de la fournaise (la réunion) during the covid-19 lockdown. Seismological Society of America, 92 (1), 38–52.

Retailleau, L. et al., 2022. A Wrapper to Use a Machine-Learning-Based Algorithm for Earthquake Monitoring. Seismological Research Letters, 93(3), pp.1673-1682.

Retailleau, L. et al., 2022. Automatic detection for a comprehensive view of mayotte seismicity. Comptes Rendus. Géoscience, 354 (S2), 153–170.

Revosima, 2020. Bulletin de l'activité sismo-volcanique à mayotte. Retrieved from https://www.ipgp.fr/volcanoweb/mayotte/Bulletins-Communiques/Bulletins/2020/Revosima bull N18 08062020.pdf

REVOSIMA (Mayotte Volcanological And Seismological Monitoring Network), Institut de physique du globe de Paris (IPGP), Bureau de recherches géologiques et minières (BRGM), Institut français de recherche pour l'exploitation de la mer (IFREMER), Centre national de la recherche scientifique (CNRS), 2021. Data collection of the mayotte volcanological and seismological monitoring network (revosima). Institut de physique du globe de Paris (IPGP). doi: 10.18715/MAYOTTE.REVOSIMA

Rinnert, E., Lebas, E., Paquet, F., Jorry, S., Feuillet, N., Thinon, I., Fouquet, Y., 2019. Mayobs. doi:

10.18142/291

Saurel, J.-M. et al., 2021. Mayotte seismic crisis: building knowledge in near real-time by combining land and ocean-bottom seismometers, first results. Geophysical Journal International, 228 (2), 1281–1293.

Saurel, J.-M. et al., 2022. Combining hydro-acoustic sources and bathymetric differences to track the vent evolution of the mayotte eruption, mozambique channel. Frontiers in Earth Science, 10 , 983051.

Seydoux, L., Balestriero, R., Poli, P., Hoop, M. d., Campillo, M., Baraniuk, R., 2020. Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning. Nature communications, 11 (1), 3972.

Steinmann, R., Seydoux, L., Beaucé, E., Campillo, M., 2022. Hierarchical exploration of continuous seismograms with unsupervised learning. Journal of Geophysical Research: Solid Earth, 127 (1), e2021JB022455.

Taisne, B., Brenguier, F., Shapiro, N., Ferrazzini, V., 2011. Imaging the dynamics of magma propagation using radiated seismic intensity. Geophysical Research Letters, 38 (4).

Tepp, G., Dziak, R. P., 2021. The seismo-acoustics of submarine volcanic eruptions. Journal of Geophysical Research: Solid Earth, 126 (4), e2020JB020912.

Van der Maaten, L., Hinton, G., 2008. Visualizing data using t-sne. Journal of machine learning research, 9 (11).

Wilcock, W. S. et al., 2016. Seismic constraints on caldera dynamics from the 2015 axial seamount eruption. Science, 354 (6318), 1395–1399.

Yang, S., Xiao, Z., Wei, S., He, Y., Mon, C.T., Hou, G., Thant, M., Sein, K., Jiang, M., 2024. New Insights Into Active Faults Revealed by a Deep-Learning-Based Earthquake Catalog in Central Myanmar. Geophysical Research Letters 51. https://doi.org/10.1029/2023gl105159

Yoon, C. E., O'Reilly, O., Bergen, K. J., Beroza, G. C., 2015. Earthquake detection through computationally efficient similarity search. Science advances, 1 (11), e1501057.

Zheng, M., You, S., Wang, F., Qian, C., Zhang, C., Wang, X., Xu, C., 2021. Ressl: Relational self-supervised learning with weak augmentation. Advances in Neural Information Processing Systems, 34 , 2543–2555.

Zhu, W., Beroza, G. C., 2019. Phasenet: A deep-neural-network-based seismic arrival time picking method. Geophysical Journal International , 216 (1), 261–273.

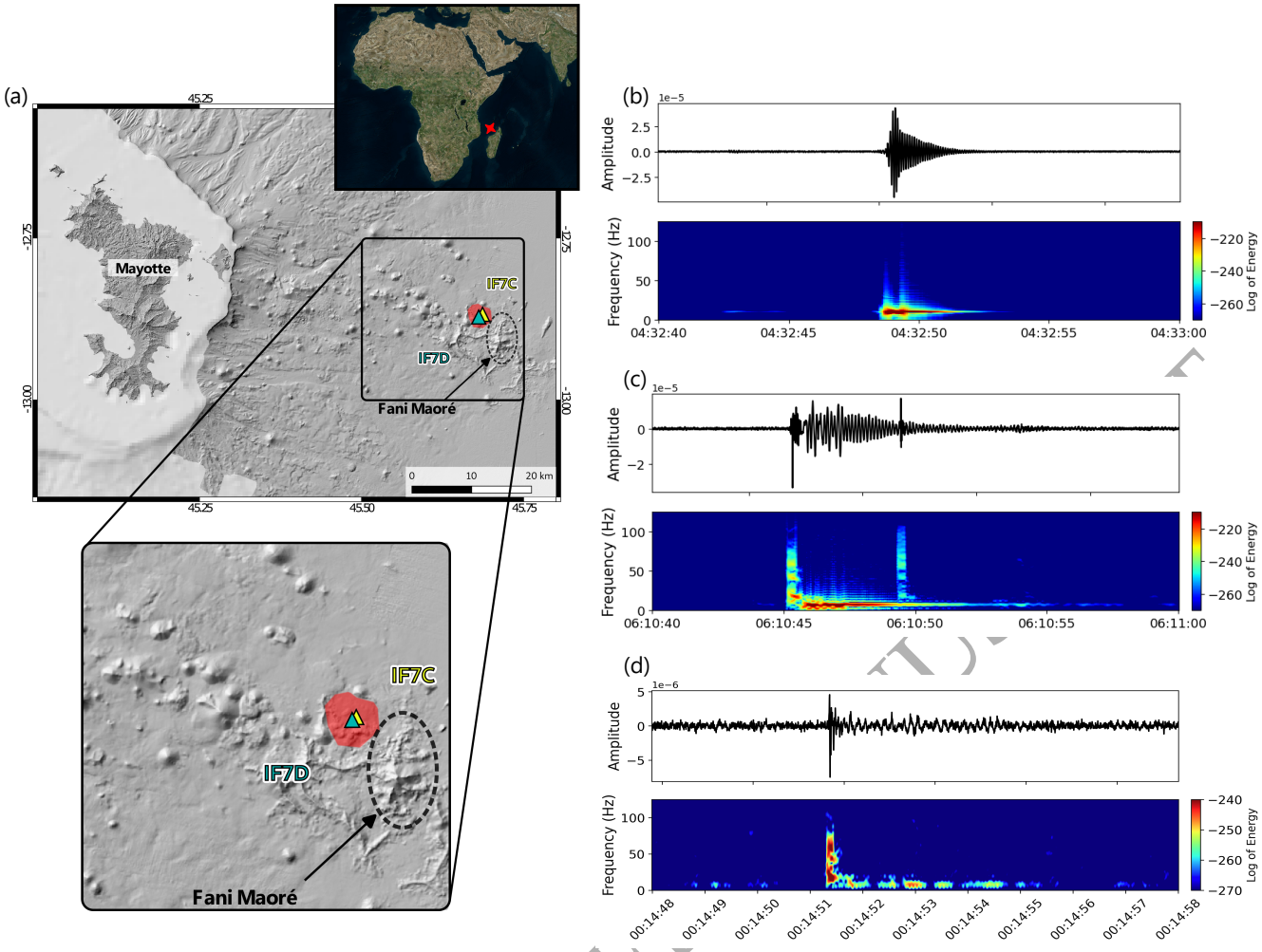

Figure 1. (a) Mayotte island and Fani Maoré submarine volcano. The red cross on the world map indicates the location of the Mayotte island. The blue and yellow triangles are the deployment positions of the seismological OBS stations IF07D and IF07C, respectively. The red patch represents the active lava flow area observed during the OBS deployment (Revosima, 2020; Berthod et al. 2022). (B) VT earthquake. (C) HA event. (D) SDE.

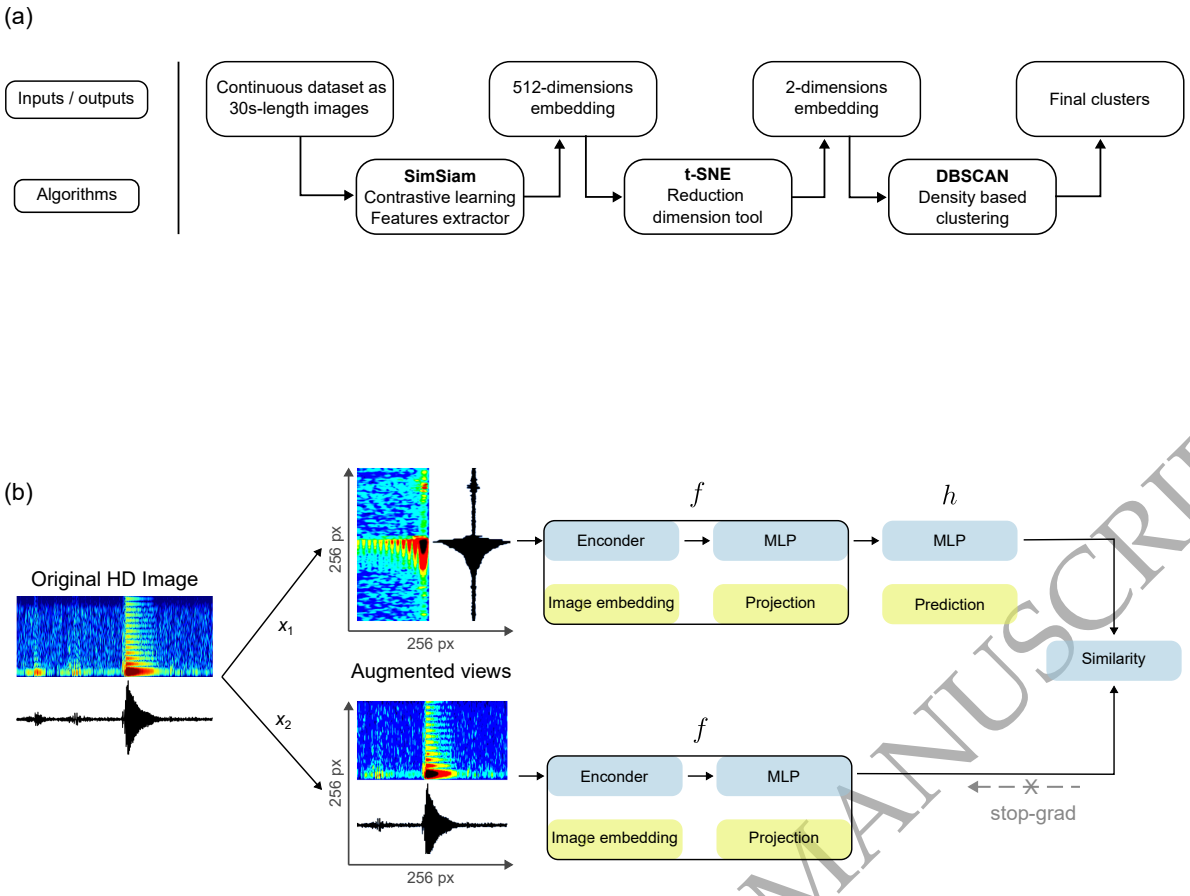

Figure 2. (a) Flowchart of the clustering workflow from the input images to the final clusters. The workflow uses three algorithms: SimSiam which is the SSL network; t-SNE (t-distributed stochastic neighbors embedding) as a dimension reduction tool and DBSCAN (density based spatial clustering of application with noise) for clustering application. (b) Architecture of the SimSiam network (after Chen & He (2021), figure 1). x_1 and x_2 represent the two augmented views of the original high dimension image. f is an encoder network, composed of an encoder and a projection multi layer perceptron. x_1 and x_2 are processed by the same encoder network f. The prediction MLP h is then applied on one side. On the other side, a stop-gradient operation is applied. Finally the model maximizes the similarity between both sides.

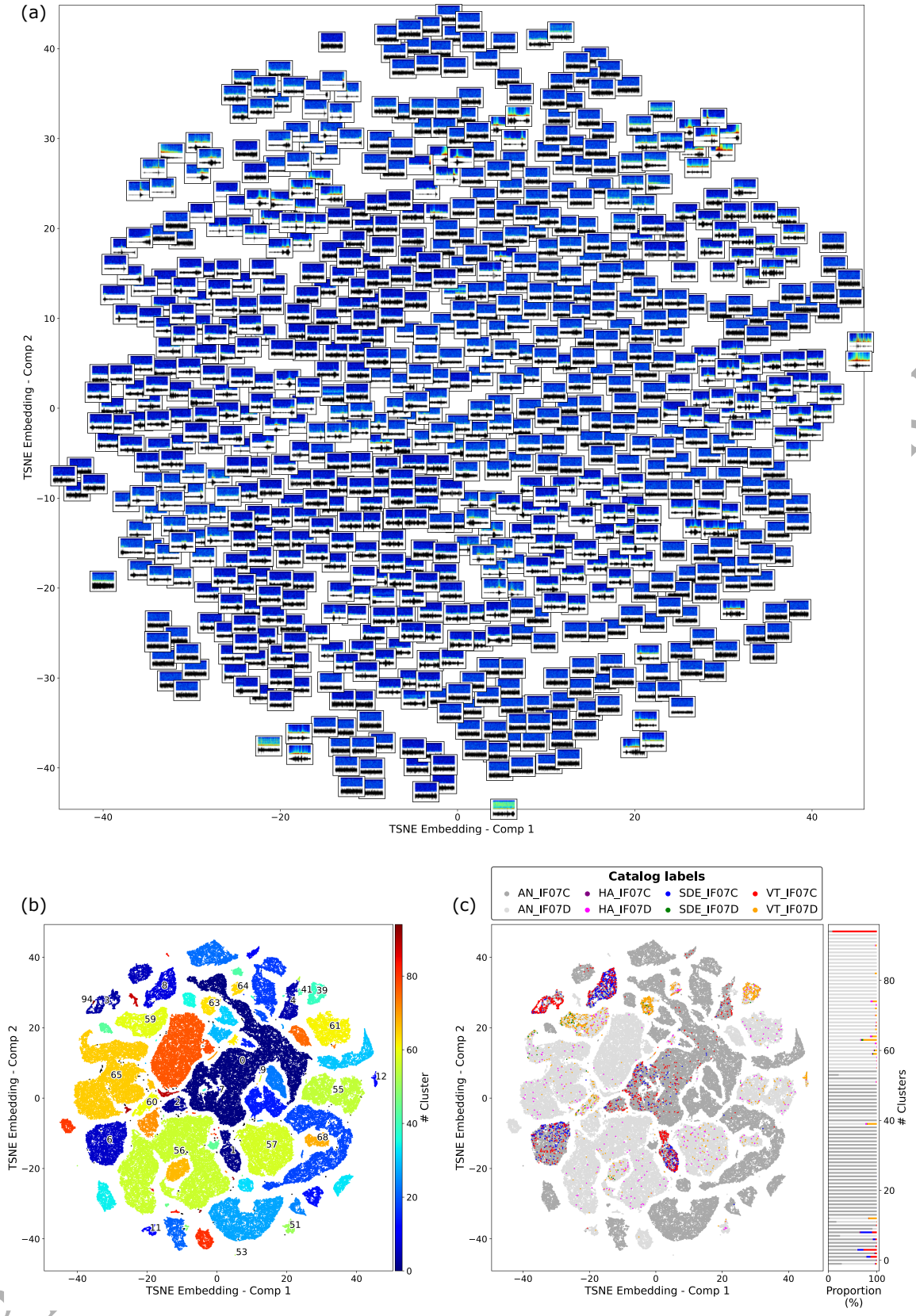

Figure 3. (a) t-SNE representation with few associated input images. This synoptic view of the dataset may be interesting for seismologists as it allows in a short time to visually inspect the dataset, in an organized matter. For the high-quality version, see Fig. S1 (available online). (b) t-SNE representation with the associated DBSCAN clusters. Black dots represent isolated points. Clusters of interest are marked with their associated number. (c) t-SNE representation with the associated labels of the events present in the REVOSIMA catalog for the two OBS. The histogram represents the proportion of labeled events in each clusters.

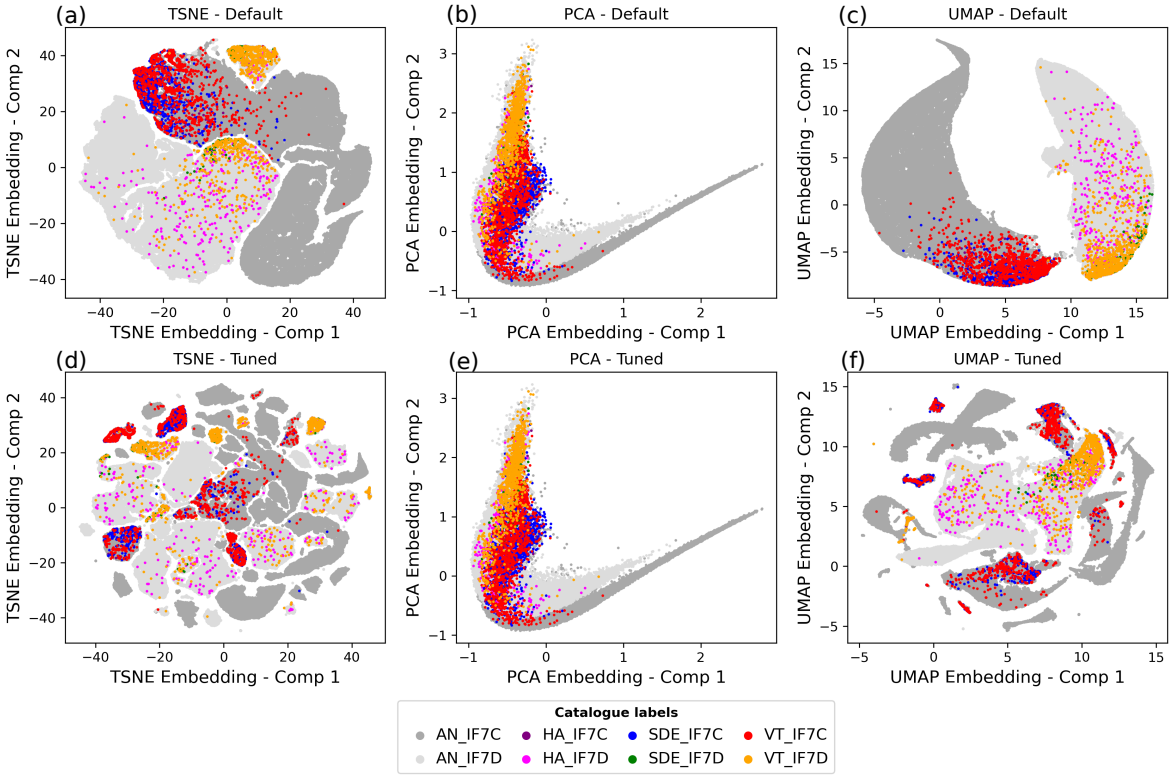

Figure 4. (a) and (d): 2d t-SNE embeddings with, respectively, default and chosen parameters. (b) and (e): 2d PCA (Principal Component Analysis) embeddings with, respectively, default and chosen parameters. (c) and (f): 2d UMAP (Uniform Manifold Approximation and Projection) embeddings with, respectively, default and chosen parameters. All the embeddings are colored based on the events present in the REVOSIMA catalog for the two OBS.

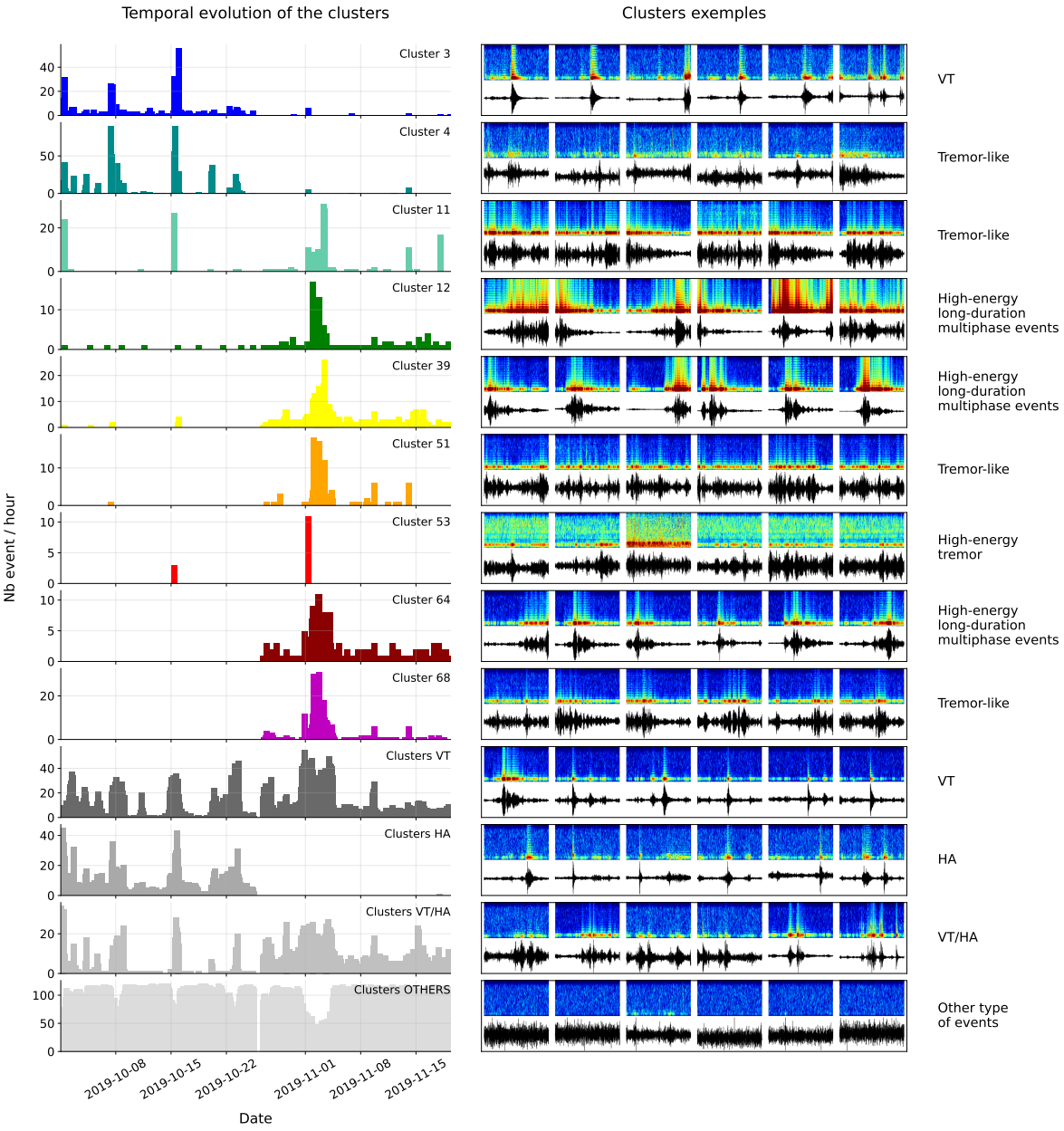

Figure 5. Number of events per hour for some specific clusters. See Fig. S2 (available online) for the entire representation. The right subplots give a representation of the clustered signals. VT, HA and VT/HA groups are defined based on our prior knowledge and are composed with the following clusters : VT (clusters 6, 55 and 60), HA (clusters 1 and 8) and VT/HA (clusters 2, 59 and 63). The OTHER group is composed by all the remaining clusters.

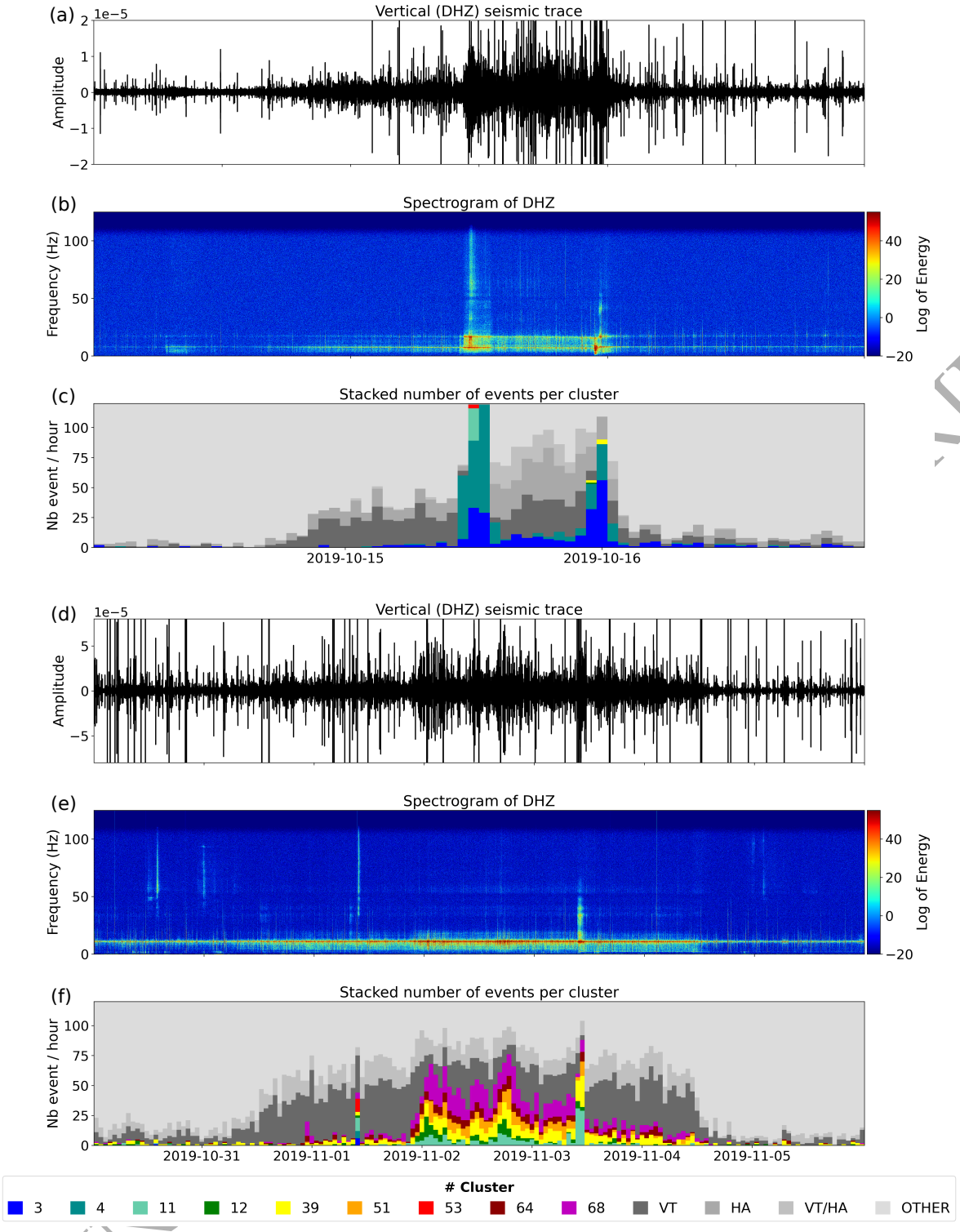

Figure 6. (a) and (d) Seismic traces after processing. (b) and (e) Spectrograms of the associated waveforms. (c) and (f) Number of events per hour for selected clusters and groups of seismic events. Selection and color of clusters are consistent with Fig.5.