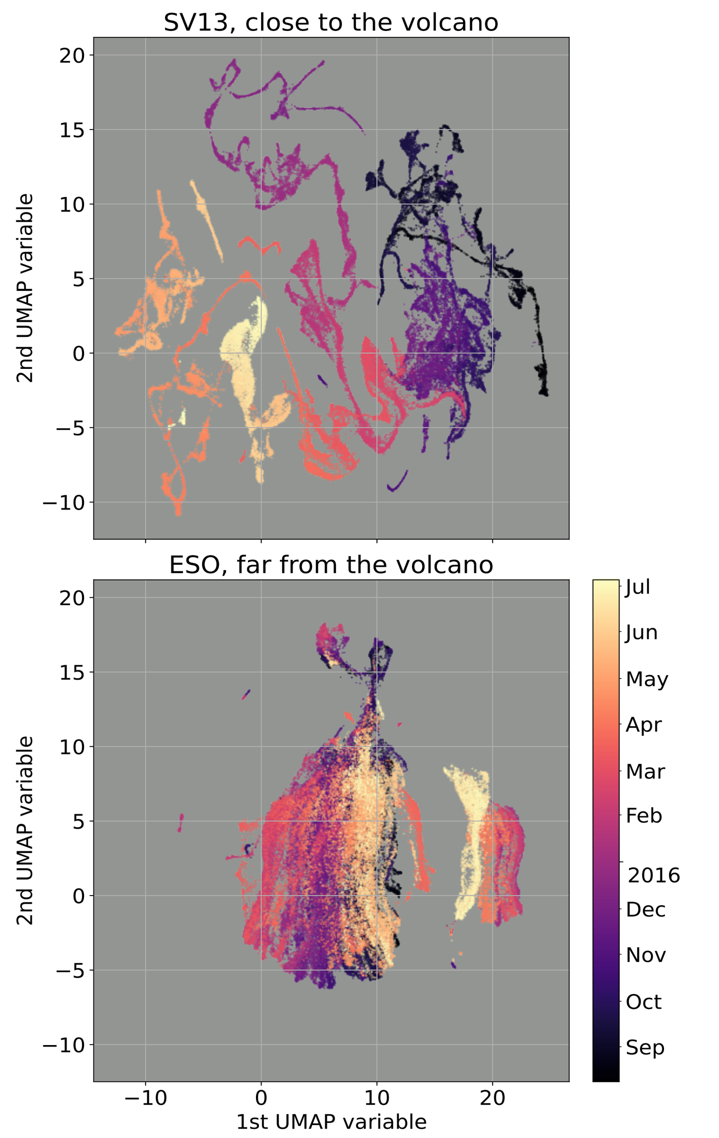
~~Non-stationarity~~ Evolution of volcanic tremor signals revealed by blind source separation and manifold learning

René Steinmann[1,2], Léonard Seydoux[3], Michel Campillo[1], Nikolai Shapiro[1], Cyril Journeau[1,4], Nataliya Galina[1]
[1]ISTerre, équipe Ondes et Structures, Université Grenoble-Alpes, France
[2]Helmholtz Centre Potsdam, German Research Centre for Geosciences GFZ, Germany
[3]Université Paris Cite, Institut de Physique du Globe de Paris, France
[4]Department of Earth Sciences, University of Oregon, 1255 E 13th Ave, Eugene, OR 97403, USA

In short:

The seismograms close to the volcano evolve continuously, changing permanent their signal characteristics (see **seismogram atlases** below).



How to read the atlases?
Each data point represents 20 minutes of three-component seismograms.
Close data points = **similar** seismograms.
Distant data points = **dissimilar** seismograms

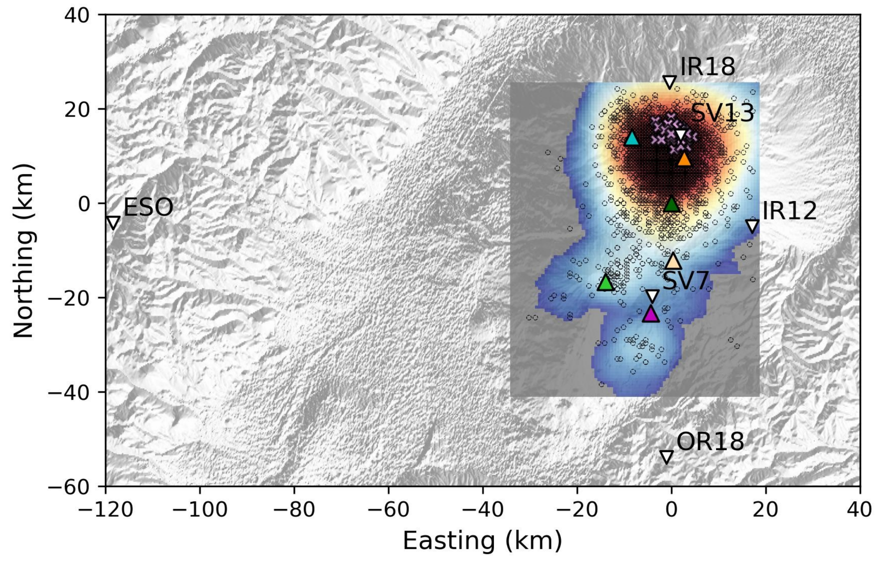
Main message:
seismograms contain interesting information about volcano beyond what is captured in catalogs and hand-designed features!

Outlook:
How to link these data-driven observation to physical processes?

1. Introducing the **idea** and the **data** of interest

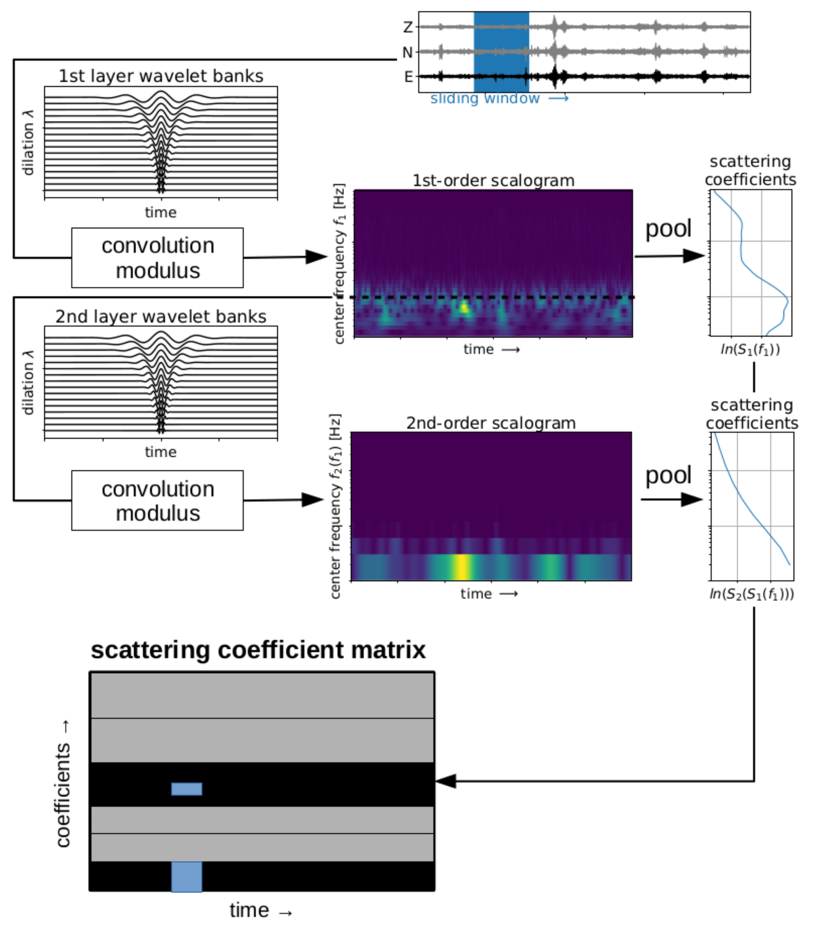
Motivation:
- Seismic monitoring of a volcano relies on discrete event catalog and hand-designed features
- Long-duration and emergent signals might contain information which is not captured by conventional methods

Our idea: we treat the continuous seismogram as a constant flux of information and retrieve data-driven features, which might capture new information from emergent signals



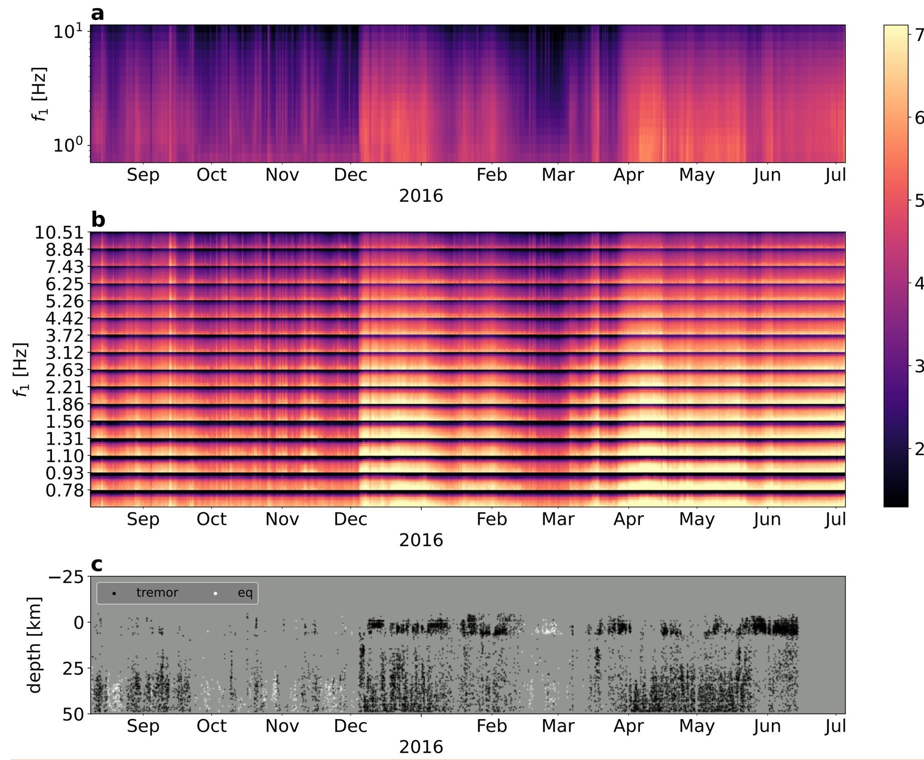
The data:
Continuous seismograms from station SV13 and other temprorary stations (August 2015 to July 2016) at the **Kamchatka** Volcano Group in Russia.

2. From **seismograms** to a data matrix with a **scattering network**

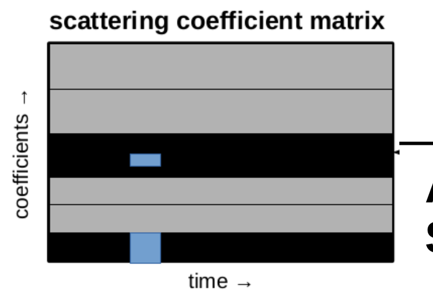


A **scattering network** is a **convolutional neural network** with fixed **wavelets** instead of learnable neurons. We apply the scattering network with a sliding window obtaining a set of **scattering coefficients** for each window.

3. The 1st and 2nd order **scattering coefficients** for the east component of SV13

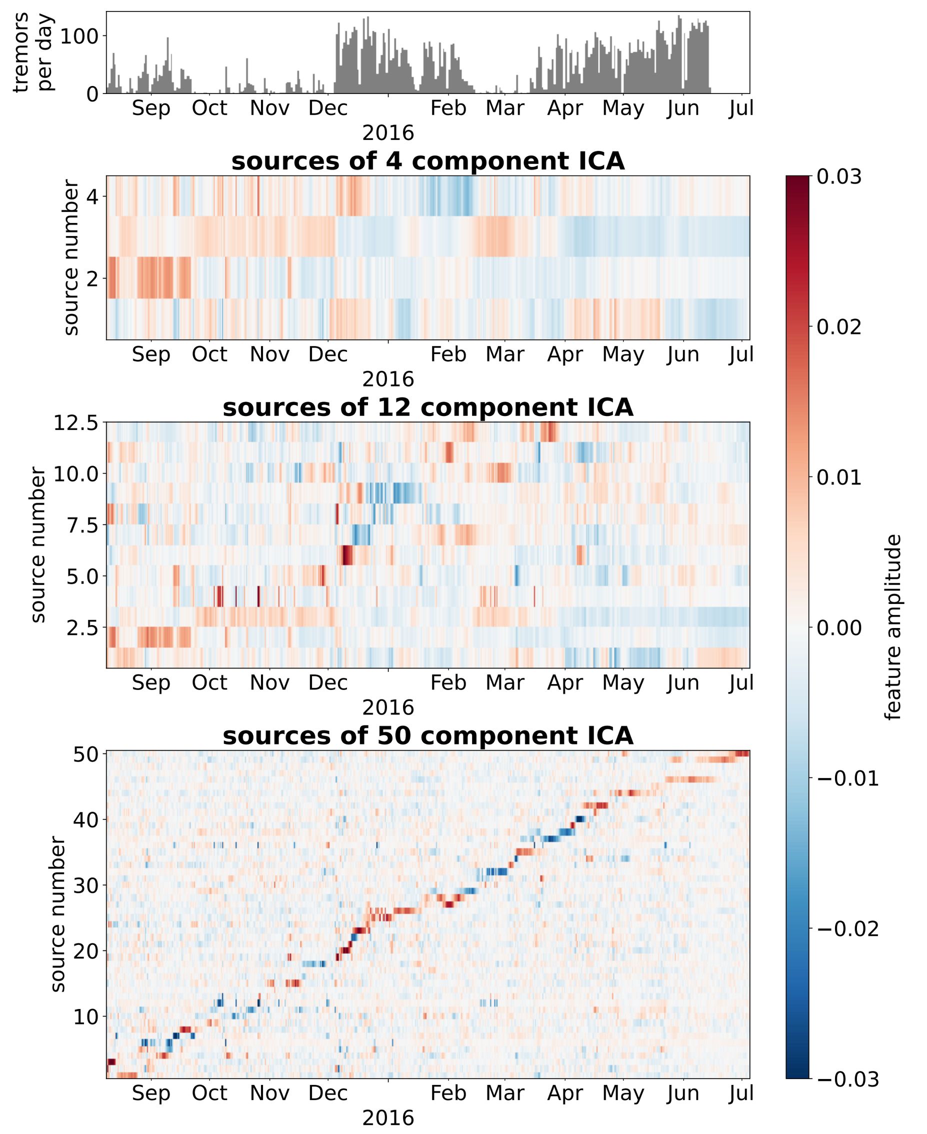


4. **Independent Component Analysis (ICA)** obtains continuous **features** for SV13

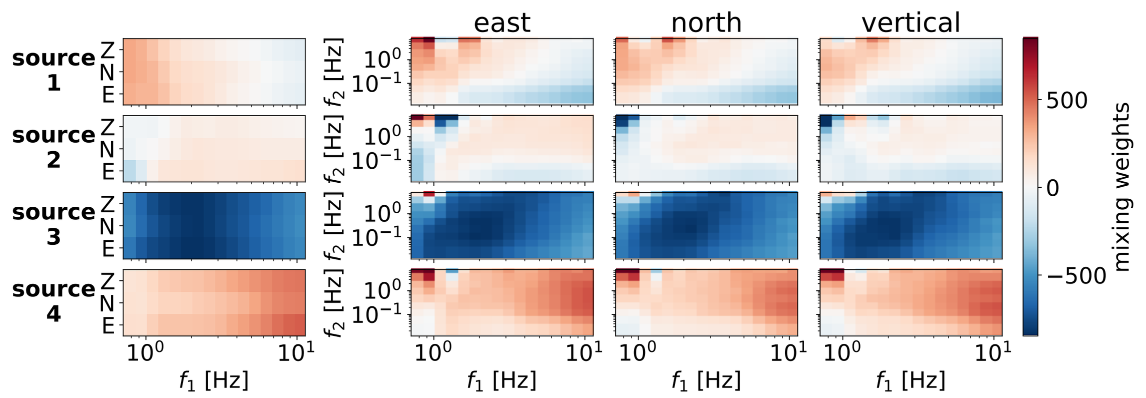


X = AS

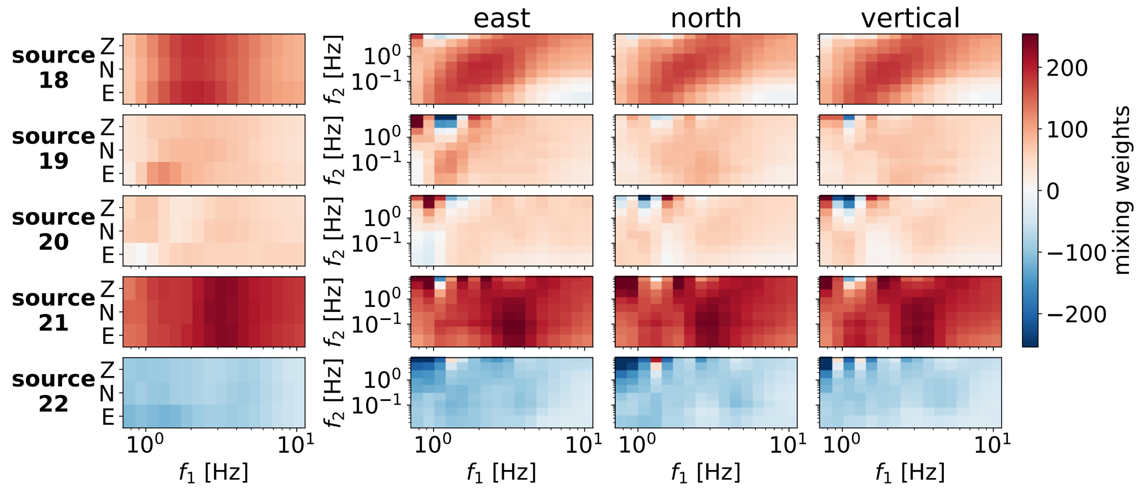
A = mixing matrix
S = source matrix (features)



5. The corresponding **mixing weights** for the **4 component model**

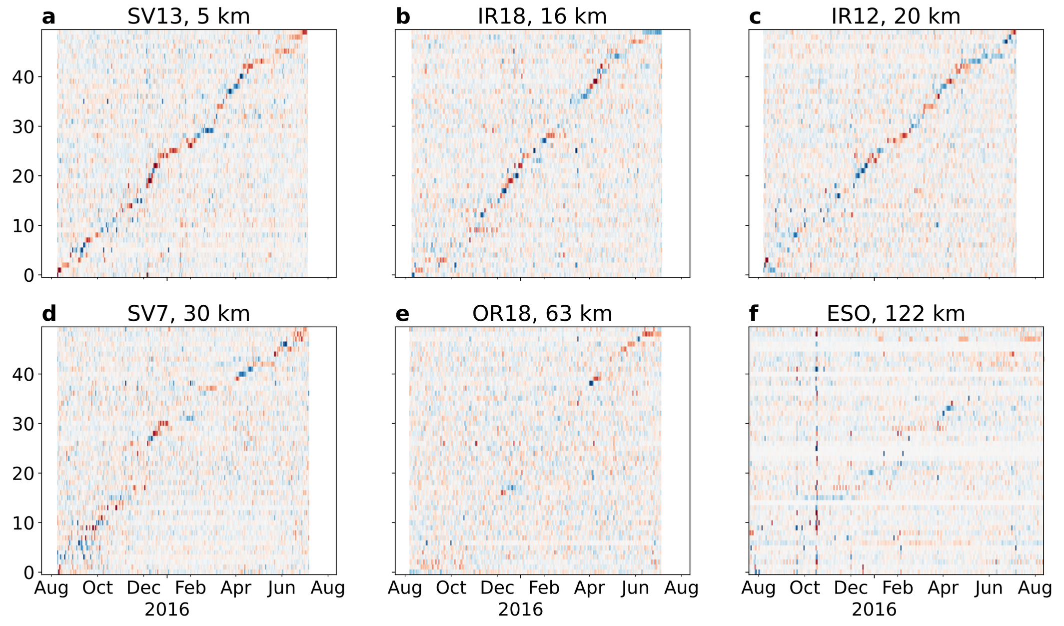


6. Some **mixing weights** for the **50 component model**



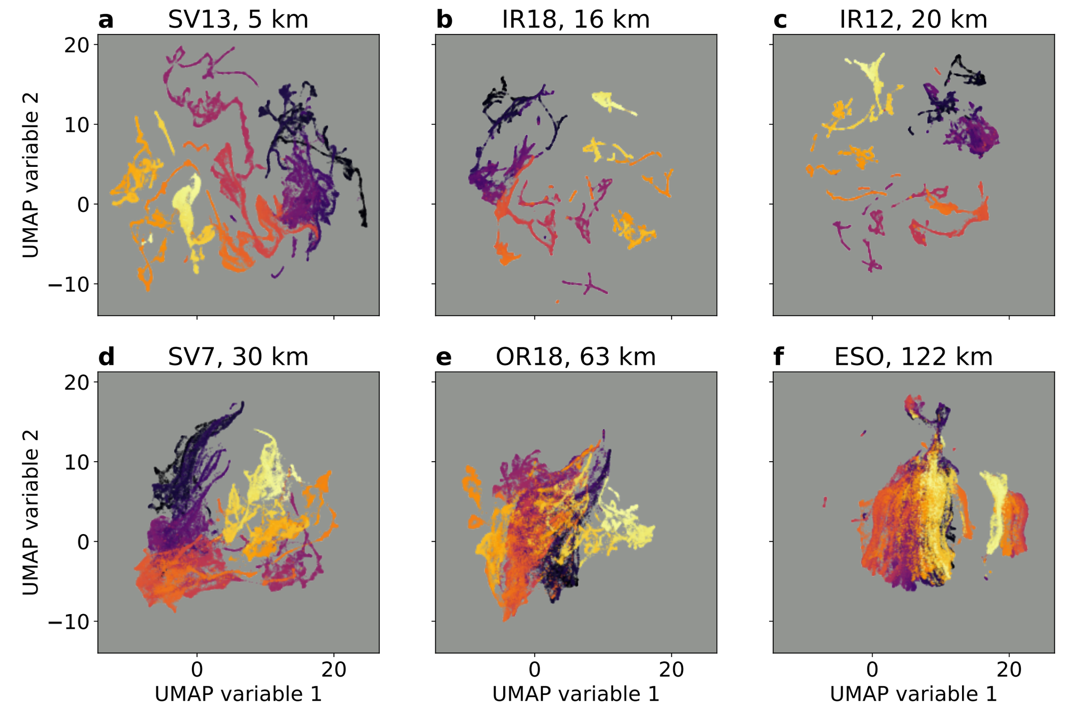
The mixing weights stored in matrix **A** help to understand the relation between the sources and the scattering coefficients. The weights of source 18-22 of the 50 component model show more variation between neighboring coefficients than the 4 component model, indicating pattern changes during the tremor onset in December.

5. How would a **50 component model** look like for other stations?



The stations closer to the volcano depict a **diagonal line** in the **time feature space**, indicating an interesting signal evolution due to the active volcano. With greater distance, the diagonal line degrades until it completely disappears.

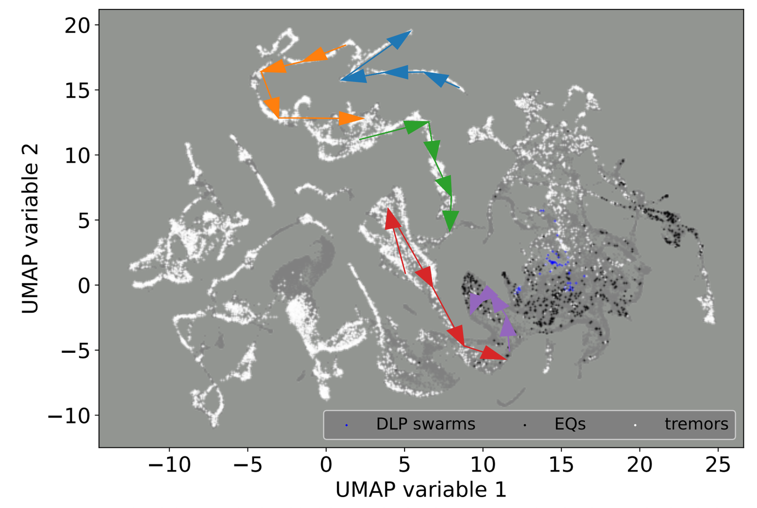
6. We complement the continuous features with **seismogram atlases** obtained with **manifold learning** (UMAP).



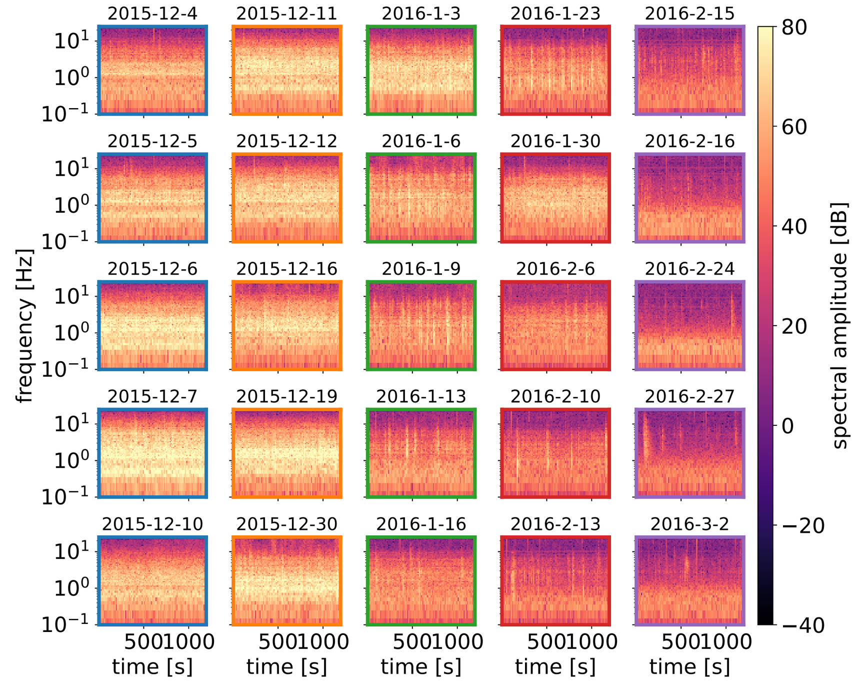
What is UMAP?
Uniform Manifold Approximation and Projection (UMAP) is a manifold learning technique, aiming at fitting a manifold to a high-dimensional dataset.

How is it different to ICA?
ICA reduces the dimension of a high-dimensional dataset by finding a set of independent sources through maximizing their independence. ICA performs well in preserving pair-wise distances while losing information about local neighborhood. UMAP performs better in preserving the local neighborhood while distorting the global properties such as cluster sizes or distances between clusters.

7. Exploring the **seismogram atlas** of **SV13**



The seismogram atlas depicts complex structures, especially during times of cataloged tremors. We see a continuous evolution of these signals, containing perhaps relevant information about underlying processes.



Preliminary conclusion:
- ICA and UMAP reveal interesting characteristics in the seismic time series close to the volcano: a complex evolution in time with continuous and sudden signal changes.
- It seems to be mainly related to tremors.
- It is difficult to relate these characteristics to physical processes (number of sources, moving source, change of source mechanism).
- Mapping the atlases between different stations might help to understand where and when these changes occur.