Classifying infrasound signals at Mount Etna using pattern recognition techniques

F. Eckel[1], H. Langer[2], M. Sciotto[2]

[1]Christian-Albrechts-Universität zu Kiel - Institut für Geowissenschaften [2]Istituto Nazionale di Geofisica e Vulcanologia - Osservatorio Etneo

Contact: felix.eckel@ifg.uni-kiel.de

Christian-Albrechts-Universität zu Kiel

Mathematisch-Naturwissenschaftliche Fakultät

ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA

Introduction

Active volcanoes in densely populated areas like Mt Etna (Italy) need constant monitoring to assess the risk of natural hazards.

Infrasound recordings are a sensitive tool in this context to detect volcanic activity that is coupled to the atmosphere, e.g. Strombolian explosions or degassing processes; even during night or in cloudy conditions.

Because volcanic infrasound signals are overlain by various (mostly weather induced) noise signals, an interpretation is often difficult to the untrained humand eye.

We therefore use a pattern classification method called *Self-Organizing maps* (SOMs) to automatically classify typical infrasound regimes at Mt Etna, e.g. wind induced noise, low amplitude background noise, or different levels of volcanic activity.

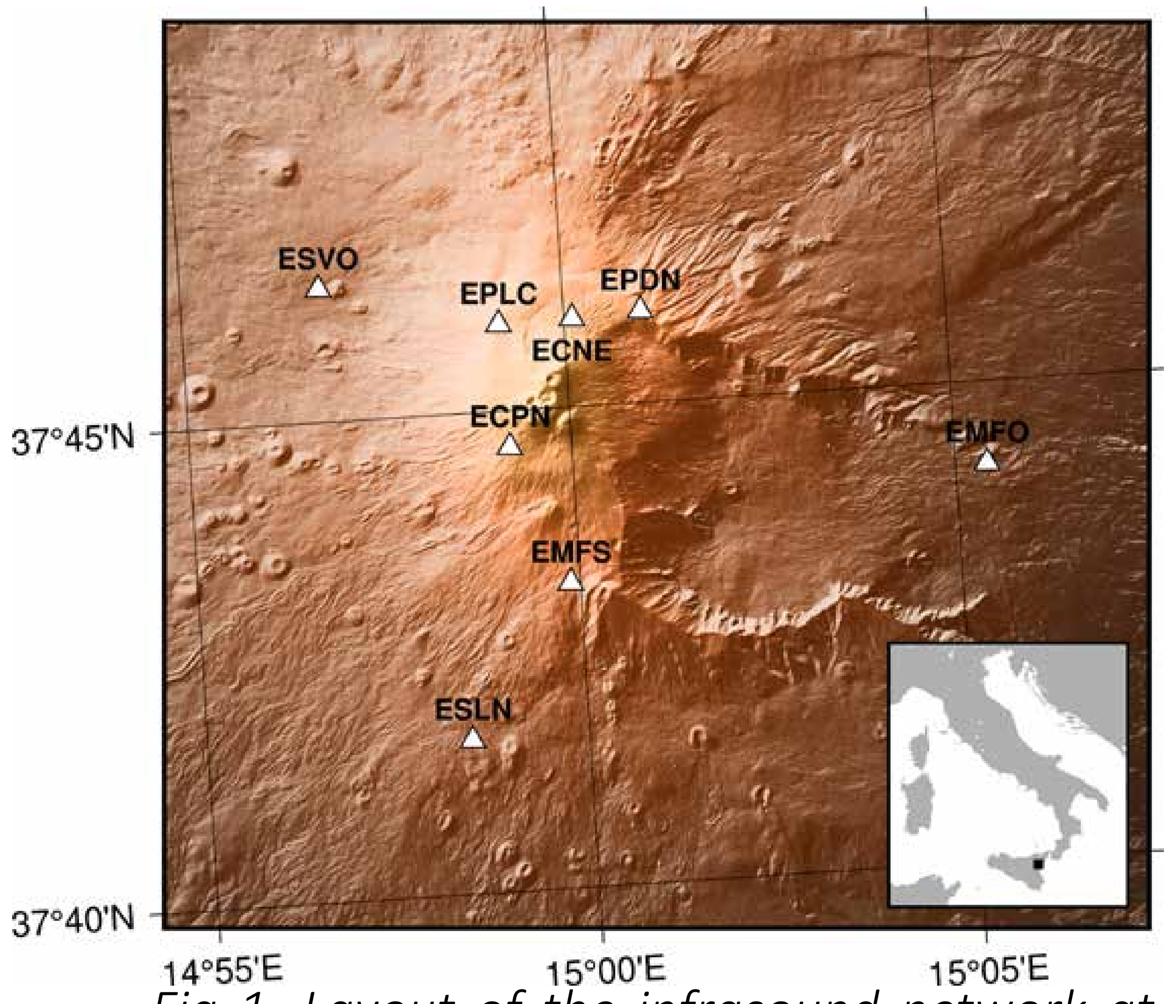

Fig 1. Layout of the infrasound network at Mt Etna. We use data from station ESLN at approx. 1,800 m a.s.l.

Infrasound Signals

We analyze data from station ESLN (see fig. 1) in the time frame from Dec 28, 2018 to Feb 29, 2020. We found this station to be less affected by wind noise or data gaps.

Typical infrasound signals (fig. 2) of volcanic origin are infrasound tremors (e.g. induced by degassing processes) or infrasound events which are identifiable by transients with durations of 1 - 30 s (induced by explosive activity).

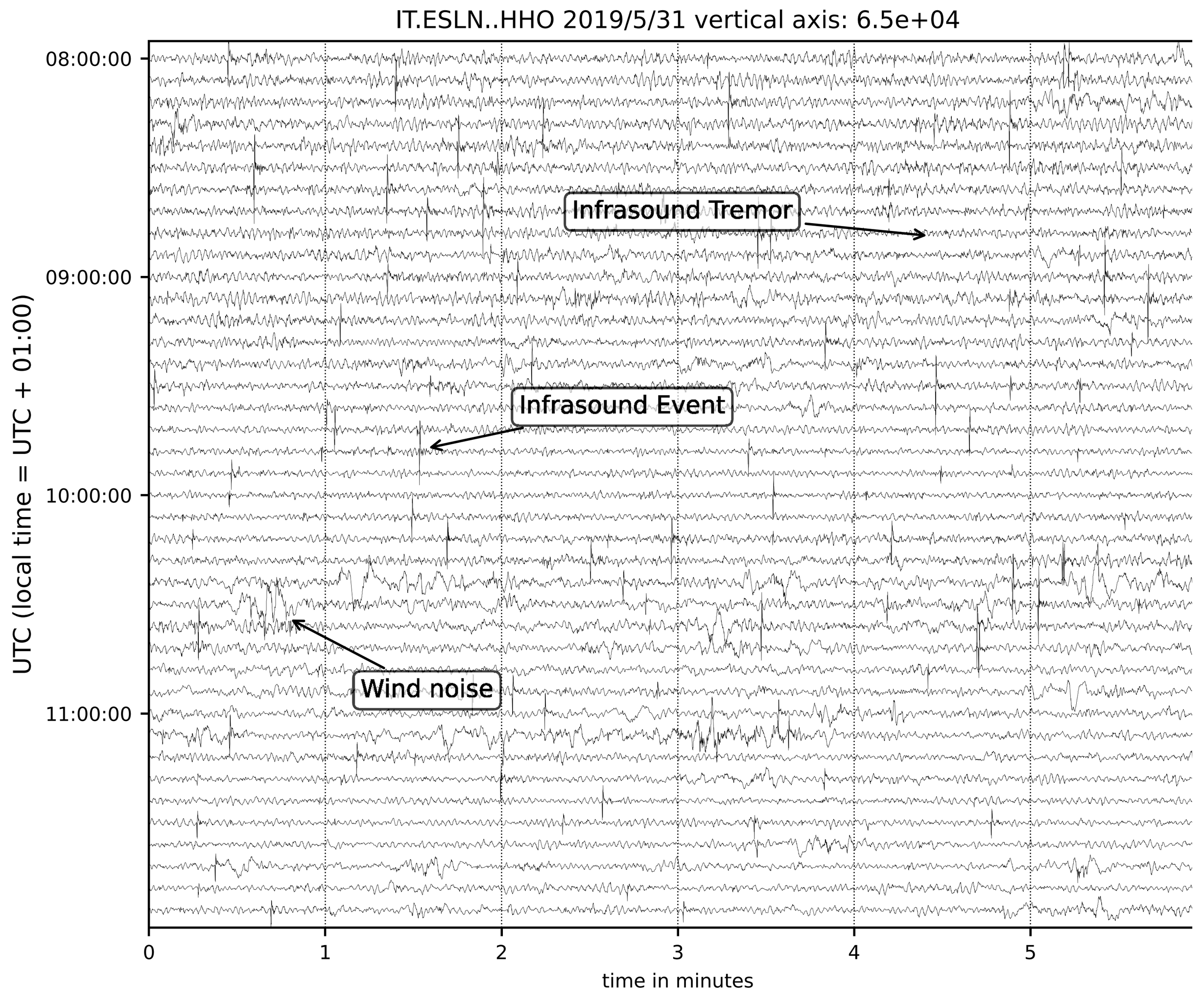

Fig 2. 4 h data example of the unfiltered infrasound waveform. Infrasound tremor is visible as the low amplitude background signal. It is overlain by infrasound events which manifest as high frequency transients. More chaotic periods can be described as wind noise.

Method

A pattern is described by as set of *features*. SOM recognizes pattern similarites by a distance metric in the feature space. The recognition is performed on a training data set. The trained model can then later be used to classify any newly acquired patterns.

We extract the features from the spectral content of the waveform in time windows of 250 s. In these windows we obtain a set of ten values describing the median amplitude in logarithmically spaced frequency bins (0.05 Hz to 25 Hz). A normalization is applied on each of these feature dimensions to ensure equal weighting for all frequency contents.

We manually select a set of 8000 patterns for the training process. This set should represent all expected and relevant scenarios at Mt Etna.

SOMs first apply a data reduction to the input feature vectors of the training data set by performing a micro-clustering approach. The cluster centroids, the *nodes*, function as prototype patterns for all feature vectors for which they are the closest node. During this process the summed distance between each node and its associated feature vectors is minimized. The nodes are connected in a lattice where each node has six neighboring nodes.

Then, the node lattice is projected onto a 2D plane using the first two components of the Principal Component Analysis. Ideally, the topological information is conserved in this process, i.e. two nodes that are close in the higher dimensional feature space are also close in the projected space.

Finally, a color is assigned to each node based on its position on the projection plane. This only supports the visualization of the results.

Classification Results

The SOM result is displayed as the projected lattice of nodes with their assigned colors (fig. 3). Similar colors refer to similar pattern representations. The color information of the nodes can be transfered to the individual patterns that are represented by the respective node. Combined with the time series of the infrasound waveform, pattern changes in the time series become easily identifiable (fig. 4).

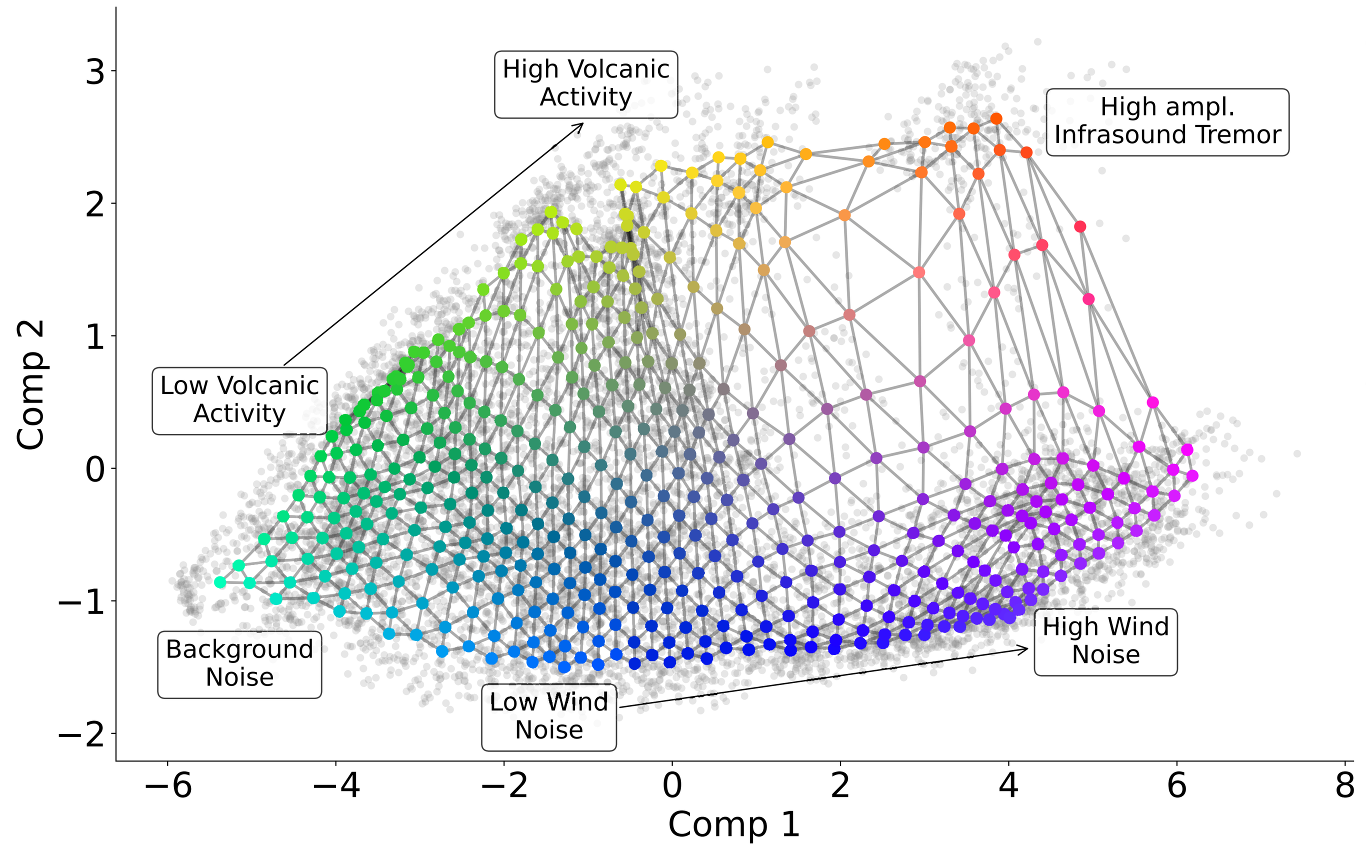

Fig 3. Projection view of the SOM node lattice (interlinked colors dots). Grey dots mark the projected feature vectors of the training data set. Nodes that are close/have similar colors represent similar pattern characteristics. The axes units are based on the Principal Component analysis and are not relevant here. Briefly annotated are the main regimes and related nodes and colors.

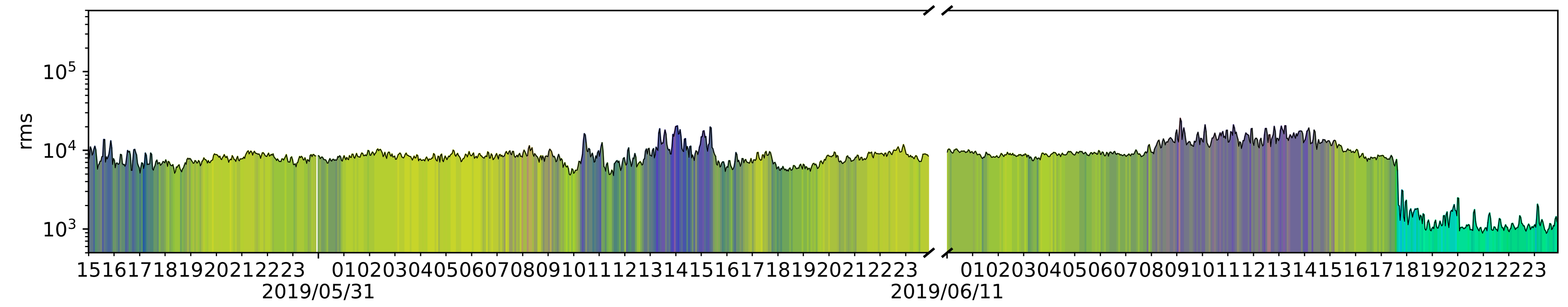

Fig 4. Root-Mean square of the infrasound waveform (black line) with the color of the closest node to the respective feature vector. Each x-tick is one hour. Greenish colors mark periods of infrasound tremor. Purplish/bluish colors indicate wind noise interferrence and mint colors represent periods of low amplitude background noise.

Conclusion

SOMs allow an effective visualization of multivariate pattern characteristics and their change over time. By using ten spectral features we were able to identify well defined infrasound regimes and their relation to the state of volcanic activity at Mt Etna.