



HAL
open science

Machine Learning for Volcano-Seismic Signals: Challenges and Perspectives

Marielle Malfante, Mauro Dalla Mura, Jean-Philippe Métaxian, Jerome I.
Mars, Orlando Macedo, Adolfo Inza

► **To cite this version:**

Marielle Malfante, Mauro Dalla Mura, Jean-Philippe Métaxian, Jerome I. Mars, Orlando Macedo, et al.. Machine Learning for Volcano-Seismic Signals: Challenges and Perspectives. IEEE Signal Processing Magazine, 2018, 35 (2), pp.20 - 30. 10.1109/MSP.2017.2779166 . hal-01742506

HAL Id: hal-01742506

<https://hal.science/hal-01742506>

Submitted on 7 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Machine Learning for Volcano-seismic Signals: Challenges and Perspectives

Marielle MALFANTE, Mauro DALLA MURA *Member IEEE*,
Jean-Philippe MÉTAXIAN, Jérôme I. MARS *Member IEEE*,
Orlando MACEDO, Adolfo INZA.

Abstract

Environmental monitoring is a topic of increasing interest, especially concerning the matter of natural hazards prediction. Regarding volcanic unrest, effective methodologies along with innovative and operational tools are needed to monitor, mitigate and prevent risks related to volcanic hazards. In general, the current approaches for volcanoes monitoring are mainly based on the manual analysis of various parameters, including gas leaps, deformations measurements and seismic signals analysis. However, due to the large amount of data acquired by in situ sensors for long term monitoring, manual inspection is no longer a viable option. As in many Big Data situations, classic Machine Learning approaches are now considered to automatize the analysis of years of recorded signals, thereby enabling monitoring at a larger scale. This paper focuses on integrated and operational tools dedicated to the automatic analysis of volcano-seismic signals. Namely we review (i) tools for the optimal representation of volcano-seismic signals (feature space) and the available methods for volcano-seismic events (ii) detection and (iii) classification. We then propose an architecture for the automatic classification of volcano-seismic events. Our prediction system is tested on 6 years of recordings containing 109434 volcano-seismic events acquired from Ubinas volcano (the most active volcano in Perú). Our new proposed model is build using supervised machine learning algorithms (Support Vector Machine) and reaches 92.2% of correct classification over six classes. This prediction model is then used to fully analyze the 6 years of recorded signals.

Index Terms

Volcano-seismic signals, Machine Learning, Data Mining, Automatic Statistical Classification, Signal Representation, Random Forest, Support Vector Machine, Feature Space.



1 INTRODUCTION

VOLCANIC unrest monitoring and the evaluation of associated risks is still an open and timely issue. Following the United Nations International Strategy for Disaster Reduction recommendations [1], it is clear that the development of effective methodologies along with innovative and operational tools to mitigate risks related to volcanic unrest (prevention, crisis management and recovery) are of key importance for the society. Volcanoes are usually mostly monitored through their seismic activities recorded by seismometers. By interpreting recorded seismic events, geoscientists obtain meaningful knowledge on the volcano internal activity. This is done by looking the seismic recordings for signals that can be associated to a specific kind of volcano activity. In terms of data analysis, this results in a pattern recognition (semantic classification) problem in which observed signals are assigned to semantic classes (each associated to a different volcano activity). Operational volcanoes monitoring can range from routine surveillance of volcanic activity to more focused inspections aimed at detecting specific signatures (e.g., associated to precursors of an eruption). This can then be carried out by analyzing on a time lapse the frequency of occurrence, rate, energy level and source location of the different classes. Volcanoes produce a great variety of seismic signals that are related to the transport of fluids (water, magma, gas) and their interaction with solid rock. The most recurrent events that can be associated to seismic signals are listed below:

- *Volcano-Tectonic* (VT) events are related to brittle failure generated by the response of the rock to stress changes induced by magma movement; VT events are characterized by high-amplitude P and S phases and broad spectra up to 15 Hz [2].
- *Long-Period* (LP) events are related to resonance of fluid-filled cracks or conduits induced by pressure transients in the fluid ([3], [4], [5]). They are different from VT events as they do not show a double couple mechanism [6], [7]. LP events involve volumetric modes of deformations [8].

• M. Malfante, M. Dalla Mura and J. I. Mars are with Univ. Grenoble Alpes, CNRS, Grenoble INP*, GISPA-Lab, 38000 Grenoble, France.

* Institute of Engineering Univ. Grenoble Alpes.

• J. P. Metaxian is with Univ. Savoie Mont Blanc, ISTerre, CNRS, IRD CNRS, 73376 Le Bourget du Lac and Institut de Physique du Globe de Paris, Université Paris-Sorbonne, CNRS, France.

• O. Macedo and A. Inza are with Instituto Geofísico del Perú, IGP Perú

- *Tremors* (TR) are denoted by a sustained amplitude lasting from minutes to days and occur over a frequency range from 1-9 Hz [9]. Observations show a large variety of tremor duration, amplitude and frequency contents. On some volcanoes, tremor is created by the superposition of increasingly frequent and regular individual LP events [10]. Eruptions also produce broadband tremor directly associated with sustained explosions [11] and magma degassing.
- *Explosions* (EXP) quakes are generated by sudden magma, ash and gas extrusion. Ubinas volcano [12] (whose data are used in Section 4) produces Vulcanian explosions which are related to fragmentation processes in the conduit as observed on many andesitic volcanoes [13], [14], [15], [16].

Other volcano-seismic signals that do not fit this nomenclature are not uncommon, but are not necessarily found on all volcanoes (see Section 4 for more details on classic and less frequent volcano-seismic signals).

With the increasing number of seismic observatories, the greater autonomy of recording stations and the availability of new sensors, huge databases of volcano-seismic signals are now accessible to the community. Consequently, the processing of volcano-seismic data (i.e. (i) detection of volcano-seismic events and (ii) their classification) registers in the Big Data context and manual inspection of those signals is no longer possible [17]. New techniques such as Machine Learning methods are taken in consideration but despite studies on the subject, a large majority of volcanic observatories still process their data manually, stressing the need for fully operational models. In particular, in most operational situations, the detection stage is semi-automatic, but the classification task remains manual and highly time consuming. Classification is particularly complex and difficult given that there is a great variety of volcano-seismic signatures (classes) at a given volcano and that intra-class properties and characteristics evolve through time (e.g., inactive or eruptive period, from an eruption to another). The number of events for each class can also increase drastically in active phases making it difficult to distinguish waveforms which can be overlapped. According to the experts and technicians (manually) classifying those signals in volcanic observatories, the classification task is particularly complex and some signals are extremely difficult to analyze. In particular, it is not uncommon that experts change their understanding on the recorded signals over time, making interpreted data evolve.

Machine learning algorithms have proven to be very efficient in many different application such as image and speech processing, medical imaging, finance, robotics or data analytics [18]. This article focuses on Machine learning algorithms and methods in a geophysical context. In particular, we propose to (i) explain the issue of signals representation in Section 2 and give a deep review of representations used for seismo-volcanic signals in the literature. We also compare it with classic representations used for more general transient signals. In this section, we also propose a representation for transient signals that goes beyond the state of the art proposed representations and that is adapted for seismo-volcanic signals.

(ii) In Section 3, we survey existing Machine Learning approaches devoted to the automatic analysis (detection and classification) of volcano-seismic signals. The issue has previously been addressed in the literature but is not resolved and we give a summary and analysis of the suggested methods.

(iii) We then propose to illustrate the automatic classification of volcano-seismic signals by proposing and testing an architecture based on Support Vector Machines (SVM). In particular, the method we propose is tested on a very large dataset (more than 105000 volcano-seismic events) and gives excellent results in term of classification and volcano physical understanding. The analysis of this 6 years long database is carried out in Section 4. Finally, we leverage the obtained results by reporting the challenging aspects of this analysis and some envisioned new perspectives in Section 5.

2 SIGNALS REPRESENTATION

The optimal representation of a signal is a lasting issue in signal processing and data analysis: many applications have proven more reliable or effective when using alternative representations of the data with respect to the original space (e.g., temporal domain). These representations are obtained by extracting some characteristics (features) from the data and by using these as bases of a new space of representation [18] (i.e., the feature space). Many representation (feature) spaces can be considered and their choice is related to the data and applications at hand. Typical characteristics of features spaces used for classification tasks are (i) a lower dimension compared to the one of the original space of representation and (ii) a greater separability of data belonging to different classes [18]. We here review conventional representations used for automatic classification applications, and we propose an adapted representation for volcano-seismic signals that will be tested in Section 4.

2.1 A review on the representation of transient signals

In this part, we review features that have been used in the literature for classification of transient signals. Specifically, we give an account on features that are used to represent volcano-seismic signals and compare them with features used for the representation of other transient signals such as bioacoustics [19], [20], [21], anthropic [22], speech and music [23], [24] signals.

The analysis of volcano-seismic recordings has been historically carried out on the raw time series [25], [26], [27], or on some transforms of the recordings including Fourier transform [25], the recordings envelope [25], the autocorrelation function [25], [28], [29] or the wavelet transform [30]. Results are variable in terms of the reported precisions and in general with a margin of improvement using such features. [31], [32] and [33] used classic features from speech processing called Mel Frequency Cepstral Coefficients (MFCC) and transposed them to the frequency range of volcano-seismic signals.

Indeed, even if in some aspects volcano-seismic recordings are similar to speech signals, they remain incompatible with speech dedicated features since their frequency and distribution range are incompatible (see Figure 2 for some examples of volcano-seismic signals). Volcano-seismic signals rarely exceed 15Hz and do not follow the mel scale while the human speech range can be as high as 20kHz. This feature extraction process leads to an ordinate representation of the signals in the cepstral domain (frequency of frequency, more detail in Section 2.2). However, results vary significantly from one study to another. [34] also used low level coefficients (ordinate representation): Linear Predictive Coding coefficients (LPC) and observed good classification results, however they worked on two classes only. LPC coefficients were also used in [35] and in [36] along with temporal features describing the amplitude range of the considered signals. One common point of those features is that they are low level and ordered (i.e. features cannot be permuted). Typical learning algorithms do not respond well to ordered representations with a potential impact to the results.

Higher level features have started to appear in the representation volcano-seismic signals, for example in [37] with six features describing temporal and spectral properties of the signals which lead to very good results. [38] also obtained good results by using six statistical features extracted from the time series, for a three classes classification task. [39] proposes to use more than 30 high level features resulting of trace or polarization analysis and of spectral parameters for example. The same features were also used in [40]. Finally, [41] uses 40 features extracted both from time and spectral domains.

It is also worth noticing that some studies consider learning the signals representation from the dataset instead of handcraft features. [42] for example used dissimilarity representation. However those representation are once again ordinate and therefore lead to limited results on the considered dataset. On a similar idea, [43] uses HMM models to represent the signals in a generative embedding space, but with limited results.

As a comparison, we also consider some works based on features used for the classification of transient signals from different domains. Generally speaking, features used for wider applications tend to be more numerous and are worth reviewing since they could be used in other applications, including volcano-seismic signal processing. For the automatic classification of transient sonar sounds, [22] consider more than 20 features including signals shape descriptors (e.g., rate of attack and decay, temporal occurrence of the main peak), statistical moments (e.g., mean, skewness and kurtosis) and signal power (e.g., peak power, average and power standard deviation). In bioacoustic, similar features can be used in: see for example [19], [20] to respectively distinguish boats from whales and to automatically identify bird species. In [21] descriptors based on entropy are used for frogs sounds classification. Similar features are also found to represent music for genres classification, for example in [24] in which entropy, centroid, centroid ratio, bandwidth and silence ratio are considered. [23] also perform a classification of orchestral instruments representing recordings in terms of their centroid, skewness, kurtosis or centroid velocity. Most of those features are included in the feature set we propose in Section 2.2.

Another issue related to signals representation is the feature selection process: how to select the best features among a large feature set for a given application? This question is of great interest and several approaches can be used. This topic however is out of the scope of this paper, and the reader is welcome to consult the literature. Some paper related to the feature selection issue for the automatic processing of volcano-seismic data have approached the subject, see for example [32], [39], [41], [44].

2.2 Proposed representation for volcano-seismic signals

As previously explained, the automatic classification of volcano-seismic signatures is still an open issue where reliable methods still need to be tested. Real-time operational systems are especially lacking. In this scenario, automatic classification methods take a fundamental role since they allow (i) to carry out a supervised automatic discrimination of signals when reference data (e.g., defined by the experts) is available and (ii) to analyze large datasets without human prior knowledge and interpretation over the observed phenomenons in an unsupervised approach. The role of features is critical in this task, and we propose here a feature set appropriate for transient signals representation, in particular for volcano-seismic signals. The main differences between the state of the art and the feature set we propose is (i) the large set of features we consider (102 features) which allow a thorough and complete description of the signals and (ii) the feature extraction process. In particular, we propose to compute the features from three different domains of representation of the observations. By doing so, we aim at better underlining complementary signals properties, and eventually give a more complete characterization of the observations. The three domains of representation are listed below:

- *Temporal domain.* The representation of the signal of interest in this domain ($x[t]$) is useful to get access to the waveform properties in the original domain of acquisition [45].
- *Frequency domain.* A Fourier transform ($\mathcal{F}\{\cdot\}$) of the discrete temporal signal $x[t]$ leading to $X[f] = \mathcal{F}\{x[t]\}$ which represents the spectral content of the observation.
- *Cepstral domain.* Cepstral domain was originally used in speech processing. By computing the Fourier transform twice (i.e., the Fourier transform is computed on the frequency representation of a temporal signal), cepstral domain highlights harmonic properties of a given signal: $\mathcal{X}[q] = \mathcal{F}\{X[f]\}$.

A benefit of computing handcraft features instead of learning them blindly from data is to keep a physical sense associated with data descriptors. Such physical interpretation of features is extremely significant for volcanologists, since some of these features can be directly related to specific characteristics of the physical phenomena in the volcano [47]. The set of features used in this work is listed in Table 1. We organized them in three groups: statistical features, entropy features and shape descriptors.

TABLE 1

Features computed on a signal $s[i]_{i=1}^n$ (in which i might correspond to a temporal, frequency or cepstral sample). $E = \sum_{i=1}^n s[i]^2$ and $E_i = s[i]^2$ respectively describe the signal energy and the energy at sample i .

STATISTIC FEATURES		
Feature	Definition	Ref.
Length	$n = \text{card}(s)$	[22]
Mean	$\mu_s = \frac{1}{n} \sum_i s[i]$	[22]
Standard deviation	$\sigma_s = \sqrt{\frac{1}{(n-1)} \sum_i (s[i] - \mu_s)^2}$	
Skewness	$\frac{1}{n} \cdot \sum_i \left(\frac{s[i] - \mu_s}{\sigma_s} \right)^3$	[46]
Kurtosis	$\frac{1}{n} \cdot \sum_i \left(\frac{s[i] - \mu_s}{\sigma_s} \right)^4$	[46]
\bar{i} of Central Energy	$\bar{i} = \frac{1}{E} \cdot \sum_i E_i \cdot i$	[22]
RMS bandwidth	$B_i = \sqrt{\frac{1}{E} \sum_i i^2 \cdot E_i - \bar{i}^2}$	[22]
Mean skewness	$\sqrt{\frac{\sum_i (i - \bar{i})^3 E_i}{E \cdot B_i^3}}$	[22]
Mean kurtosis	$\sqrt{\frac{\sum_i (i - \bar{i})^4 E_i}{E \cdot B_i^4}}$	[22]
ENTROPY FEATURES (with $p(s_j)$ the probability of amplitude level s_j)		
Feature	Definition	Ref.
Shannon entropy	$-\sum_j p(s_j) \log_2(p(s_j))$	[21], [24]
Rényi 'entropy'	$\frac{1}{1-\alpha} \cdot \log_2 \left(\sum_j p(s_j)^\alpha \right)$	[21]
SHAPE DESCRIPTOR FEATURES		
Feature	Definition	Ref.
Rate of attack	$\max_i \left(\frac{s[i] - s[i-1]}{n} \right)$	[22]
Rate of decay	$\min_i \left(\frac{s[i] - s[i+1]}{n} \right)$	[22]
Specific values	Ratios, min, max, mean, etc.	[22], [46]

- *Statistical features* are interesting for their immediate interpretation regarding the signals shapes. For instance standard deviation describes the spread of data around its mean, skewness describes the asymmetry of the signal as a distribution (compared to the Gaussian distribution) and kurtosis is related to the flatness of the distribution. Considering the feature 'i of central energy' in the temporal domain, it displays the time around which the signal energy is centered and can be referred as time centroid.. Computed on frequency domain, it is related to the fundamental frequency in the case of a periodic signal; and similarly, if the original observation is harmonic, the feature will describe the harmonic frequency if computed in the cepstral domain. Some of those features are used in [22], [46].
- *Entropy*: Shannon entropy describes the distribution of the amplitude levels of a given signal. For a periodic signal, amplitude levels would be equally likely and the entropy high. On the contrary, a signal containing a single impulse on a continuous (constant amplitude) signal would have a lower entropy given the fact that the distribution of amplitude levels would be very biased towards the mean amplitude level. Entropy features are used in [21], [24].
- *Shape Descriptors*: Some ratios have a very helpful physical interpretation, for example the ratio maximum value over mean value can describe the contrast of a signal: if the ratio is large in time, it means that the waveform is not constant. This can be related to the cause of the event, an explosion for example will lead to a strong peak in the seismic signature, followed by a fast decay in terms of amplitude. In the frequency domain, the maximum over the mean ratio describes the spectral richness of the signature: a white noise would have an unitary ratio, while the ration of a pure sinusoid would be infinite. Finally, the ratio also describes the harmonic content of an observation if computed in the cepstral domain: a harmonic signal has a periodic spectra and consequently, is represented by a peak in the cepstral domain, the maximum over mean ratio is then infinite. On the contrary, a non harmonic observation would have a low ratio. This ratio is thus particularly interesting to describe a signal shape: depending on the computation domain, various interpretation can be given, leading to various physical interpretations. Sensitivity to outliers is a drawback of features of this family. Some of those features are used in [22], [46].

3 STATE OF THE ART IN MACHINE LEARNING TECHNIQUES FOR VOLCANO-SEISMIC SIGNALS

In this section we give an account on the techniques that have been proposed in the literature for the automatic analysis of volcano-seismic events. Namely, we focus on the two following tasks: (i) detection of relevant volcano-seismic events and (ii) their classification into semantic classes. We also review hybrid methods which perform both tasks with the same architecture.

Regarding the detection it is still done manually in numerous studies [25], [28], [30], [35], [36], [39]. Among the automatic detection processes, STA/LTA method is, by far, the most popular method and was originally presented in [48], [49]. STA/LTA is widely used in operational context and in published works, including [37], [43]. Detection systems based on the signal kurtosis have also been considered, for example in [47]. Optimal filtering has also been used, for example in [26]. In all cases results are satisfactory or promising but need be improved. In particular, they have proven to be efficient for well-separated events, or for some specific volcano-seismic classes. However, to the authors knowledge, there is no established procedure to detect volcano-seismic events in continuous recordings when (i) numerous signals occur in a short period of time (hundreds per hour) which is the case during an eruption: in this case signals associated to different events (not necessarily of the same type) can occur overlapped in time and show very different amplitudes; and (ii) for emergent signals, i.e. signals whose amplitude increases and decreases very slowly (Tremors in particular). Those signals are difficult to detect because their starting and end points are not always clearly detectable, especially when the analysis is carried out on relatively short temporal windows. Another issue of volcano-seismic events detection is the high variability in events durations: they can last less than 10 seconds for some (Volcano-Tectonics for example) to several days (Tremors). Furthermore, methods such as STA/LTA need to be manually tuned at each application (setting thresholds, window lengths, etc). And many of those approaches are tested on relatively small datasets (few hundreds or less than a hundred samples in some cases) or on datasets including only a given class of signals.

Once extracted - manually or using an automatic detection algorithm - volcano-seismic events need to be classified into one of several classes related to a physical behavior of the volcano. This information is then used to analyze the volcano and predict eruptions. In many observatories, this classification task is still done manually but the literature offers some studies on the subject. Hidden Markov Models (HMM) have been used [30]. Neural Network are also popular, but with very various results [25], [28], [29], [34], [35], [36], [37], [38], [50]. Bayesian classifiers were also tested in [42], but with limited results. Support Vector Machines algorithm also have been used, for instance in [35], [43], [50] with excellent results and very similar method than studies using Random Forests as learning algorithm, see for example [41]. Some studies also tried using unsupervised models, such as [34] with Principal Component Analysis (PCA), [27], [39], [50] with Self Organizing Maps (SOM) or [50] Cluster Analysis (CA). Results however, are highly variable.

Finally, a few studies have proposed to develop architecture to process continuous volcano-seismic recording by performing detection and classification in one step. Such architecture are based on HMM, see for instance [31], [33], [40], [51]. Once again however, results are very fluctuating.

Methods proposed in the literature for the automatic classification of volcano-seismic data are engaging but are often limited by the results which are promising in many cases but need improving, or by the testing process which involve very few signals (few hundreds or less than a hundred samples in some cases) or only a small number of classes compared to the variety of signals produced by the volcano. The main issue of today's state of the art however, is the lack of operational systems: only very few studies mention an applicative context and very few have been deployed in volcanic observatories [31], [39], [41].

4 EXAMPLE OF AUTOMATIC CLASSIFICATION OF VOLCANO-SEISMIC SIGNALS

In this section, we describe the architecture we propose for the automatic classification of volcano-seismic events.

4.1 Proposed Architecture

The proposed architecture is composed of three steps and will be assessed on volcano-seismic data recorded at Ubinas volcano in Perú [12]. The dataset is detailed in Section 4.2.

- 1) *Preprocessing*. Given the amount of data to handle which is usually very large (6 years of continuous recordings in this example), a preprocessing step is typically necessary to standardize the signals. In this work, we standardize the sampling frequency to 100Hz: the recording period being very long, some sensors or their settings have changed and need to be made up for. A first step to standardize the data is therefore necessary and all signals have been resampled to the highest sampling frequency (100Hz) available in the dataset. This preprocessing stage also includes the semi-automatic detection and extraction of volcano-seismic events from the continuous recordings. STA/LTA methods along with manual analysis were used to build a labeled dataset of observations of volcano-seismic events: each event was manually associated to a class by the experts. In our example, $N = 109434$ observations $\{x_j[t]_{t=1}^{n_j}\}_{j=1}^N$ are extracted and labeled $\{l_j\}_{j=1}^N$, thereby forming a labeled dataset. This considerable work was done by the staff at Ubinas observatory. Each observation is then normalized in energy and limited in time. The idea is here to build a classifier based on the observations shapes rather than on their energies, as it is usually the case in the state of the art. Such models would give access to a more precise analysis and prediction of volcanic eruptions and are therefore needed.

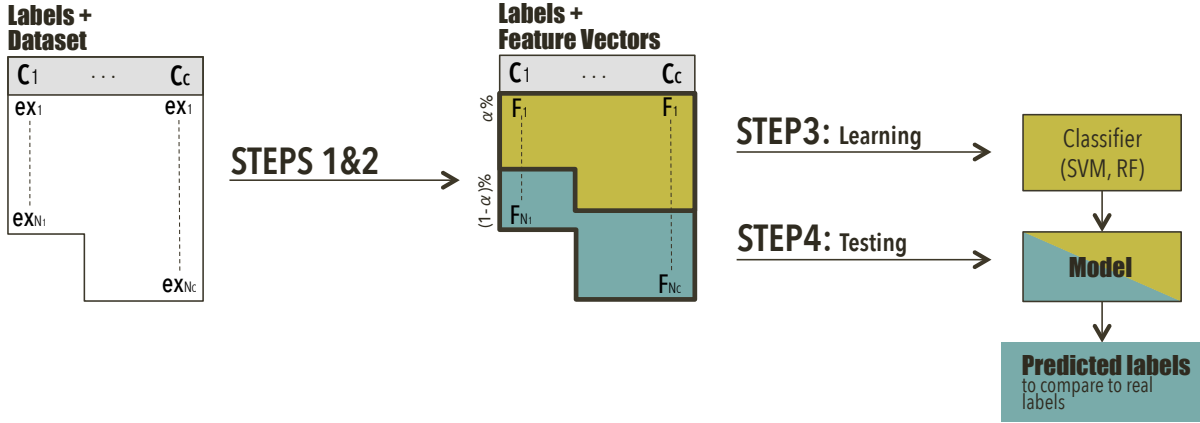


Fig. 1. Flowchart depicting the procedure for learning and testing a discriminative model.

- 2) *Features Extraction.* As reported in Section 2, the purpose of extracting features is to represent the data in a space where an automatic decision rule can be established. Features used for this study were presented in Section 2.2. Eventually, each observation is represented by a feature vector of dimension 102 that describes the signal in time, frequency and cepstral domains. For each domain, features are ordered as presented in Table 1.
- 3) *Classification.* Finally, a machine learning algorithm is used to train a classification model by learning an automatic decision rule. In this work, we propose to use Support Vector Machine (SVM) [52]. The leading idea of SVM is to build a hyperplan in the feature space, thereby separating the data into classes. The hyperplan is chosen to maximize the margin, which is the distance between the hyperplan and the *support vectors* (i.e., the samples closest to the hyperplane). Features that are used as input of the model are also displayed in a space of higher dimension (using Gaussian kernel), where the data might become linearly separable. The choice of learning algorithm should not have a significant impact on results if features have been correctly chosen [18]. The model can then be used to predict the class of unknown observations.

4.2 Results

The data on which our models are tested have been recorded at Ubinas, Peru. In this study, six years of recording are available (including 3 years of eruption). Six classes of interest have been identified by volcanologists. Among the six classes, some are quite generic for volcano-seismic signals: Long Period (LP), Tremors (TR), Explosions (EXP), Volcano-Tectonic (VT), but we also considered classes that are more specific to the Ubinas volcano, such as Hybrids (HYB) and Tornillo (TOR). Hybrids have both characteristics of VT and LP events with a high-frequency onset followed by low-frequencies. A hybrid is defined as an event with characteristics of both shear-failure and resonance [7]. Hybrid events are typical of andesitic magma and were observed at Redoubt volcano (USA) where they were first described [2]. They were also observed at Soufrière Hills Volcano, Montserrat (UK) [53] and Mount St. Helens Volcano (USA) [54]. Tornillos also called ‘screws’ events are related to resonating fluid-filled conduits or cavities. They can be considered as a specific type of LP event with a long duration coda composed of harmonic oscillations. They were observed at Galeras volcano (Colombia) before several eruptions in 1992 and 1993 [55].

On Ubinas, a total of 109434 seismic events have been manually labeled into one of those six classes, and truncated into chunks of a maximum of 5 minutes. We would like to underline the uniqueness of this dataset in terms of duration of the period of observation and diversity of the activity monitored. All results are presented using an SVM with Radial Basis Function kernel as learning algorithm, with parameters $C_{SVM} = 10$ and $\gamma = 0.01$ chosen to optimize the results on a subset of the data. In this paper we only present the results associated to SVM algorithm, but the same experiments were conducted with Random Forest with similar results.

We detail below three experimental scenarios we considered in the analysis.

4.2.1 First result: Performance Evaluation

In order to validate the methodology and to estimate results that can be expected by the proposed method, we perform cross validation (process illustrated in Figure 1). For each class, $N_{i,training} = \min(\alpha \cdot N_i, 800)$ with 70% ($\alpha = 0.7$ the learning rate) of the signals are randomly chosen to train the model. The remaining 30% of the labeled observations are used to test the model. The choice of $N_{i,training}$ is done in order to have enough data for representing each class, but also limited to 800 observations, since in some classes the number of samples is more limited. The process is repeated 50 times to obtain statistically stable results. The numerical results reported in the following correspond to the mean values over

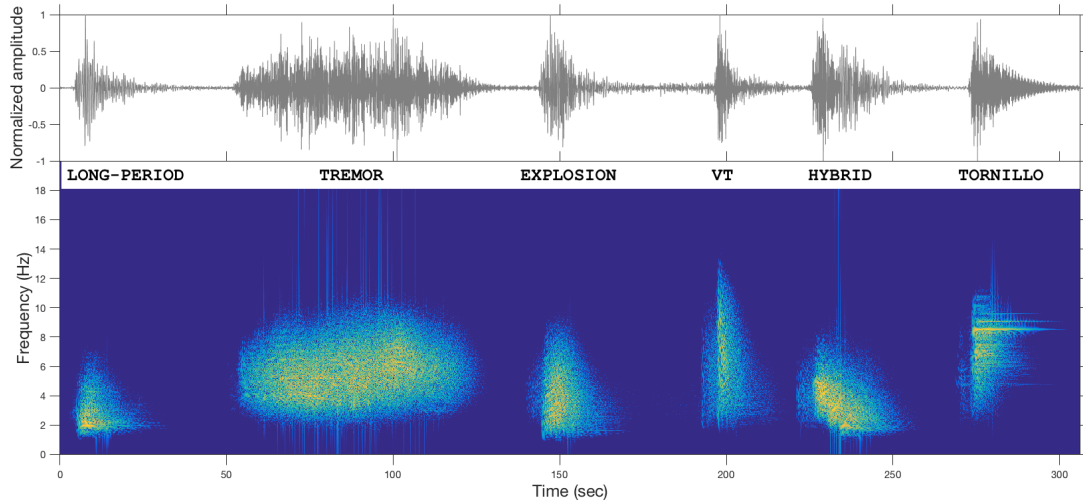


Fig. 2. Waveform and spectrogram (Gaussian window of 512 samples width) of a volcano-seismic signal recorded at Ubinas. Six observations are presented, respectively related to Long-Period (LP), Tremors (TR), Explosion (EXP), Volcano-Tectonic (VT), Hybrid (HYB) and Tornillo (TOR) events. Amplitude is linear and has been normalized.

TABLE 2
Confusion Matrix for $\alpha = 70\%$ and a model trained with SVM, 50-fold cross-validation. Overall accuracy: 92.2%.

		True Class (ground truth)					
		LP	TR	VT	EXP	HYB	TOR
Predicted Class	LP	57260	474	5	0	6	1
	TR	4261	4757	3	5	3	2
	VT	400	7	479	6	11	2
	EXP	47	6	6	36	0	0
	HYB	57	1	16	1	121	0
	TOR	4	0	3	0	0	28
Accuracy:		92.3%	90.7%	93.6%	75.3%	85.9%	84.4%

the 50 runs. To consider a coherent number of observations per class, cross validation is performed on one year of data for the most abundant classes (Long Period LP and tremors TR), and on the observations of the whole dataset for the less represented classes: in total, 70856 volcano-seismic events are considered for this first experiment.

Using this configuration, our model reaches 92.2% of overall accuracy. The confusion matrix is reported in Table 2, in which a column displays the number of correct and wrong predictions for a given class c , with $1 \leq c \leq C$. Two main sources of error are detected from this analysis: first, LP have an accuracy of 92.3%, but are mistaken with TR in most of the errors. This can be explained by the similarity in the frequency content of both events, but also by the original dataset in which some Volcano-Tectonics VT and Tremors TR are overlapped. The second main source of error comes from Hybrids HYB being mistaken for Long Periods LP and Volcano-Tectonics VT: this is explained by the physical nature of Hybrids signals, which is a class of signals with characteristics similar to both VT and LP signals. It is relevant to notice that mistakes made by the model have a physical explanation: the automatic analysis conducted by the model we propose matches the reasoning made by the experts while performing the manual labeling.

An additional analysis was carried out for evaluating the influence of the considered features. The results show that mean accuracy is systematically lower if using features extracted from only one of the three domains of representation: 86.1%, 83.0% and 79.4% for temporal, spectral or cepstral features, respectively. The choice of features describing the signal general shape in various domains is therefore well-founded.

4.2.2 Second result: Data Evolution

Depending on the volcanic activity and the volcano structure which could significantly evolve with time, observations within a same class can have a significant variability and change over time. To estimate this evolution, we propose to train a model on the 800 first observations of each class (less if 800 are not available), and to test it on the six years recordings where the mean accuracy reaches 59.7%. More specifically, we propose to study the evolution of the dominant class: Long Period LP, with 95094 events. We point out that the 800 first LP observations are recorded in few days only (June 2006). LP mean accuracy is 61.0% and its evolution with time is presented in Figure 3. This plot is interesting: we can see a clear change in the trend around May-June 2007, which can be related to a change in the observations compared to the samples

used for training. The accuracy drops from more than 95% in average to less than 10%. A drop in the number of LP can be also noticed after August 2007.

After communicating those results to the experts in charge of the Ubinas monitoring (from IGP, Arequipa), they analyzed the dataset in order to interpret the drop in accuracy. This additional work reveals significant changes in the dataset labels, mostly due to the a posteriori analysis of signals and the current better understanding of the volcano-seismic signatures by the experts with respect to 2007. First, the number of LP events in the new classification system is 10 times fewer than in the original dataset (from several hundreds per week to several dozen only). Secondly, many observations originally labeled as LP are now considered to be VT events after May 2007. This result is of capital importance, since it shows that the system we propose was able to detect dissimilarities between signals and incongruence in the manual labeling. Regarding the physical interpretation, LP come from superficial sources while VT are emitted by much deeper sources. The seismic activity therefore becomes deeper from the end of May and the increase in VT probably meaning a recharge of the system in depth which leads to a new volcanic phase of activity. Figure 3 is a proof of change starting on May 2007, with physical reason such as a volcanic phase leading to a change in Ubinas structure and/or a internal activity.

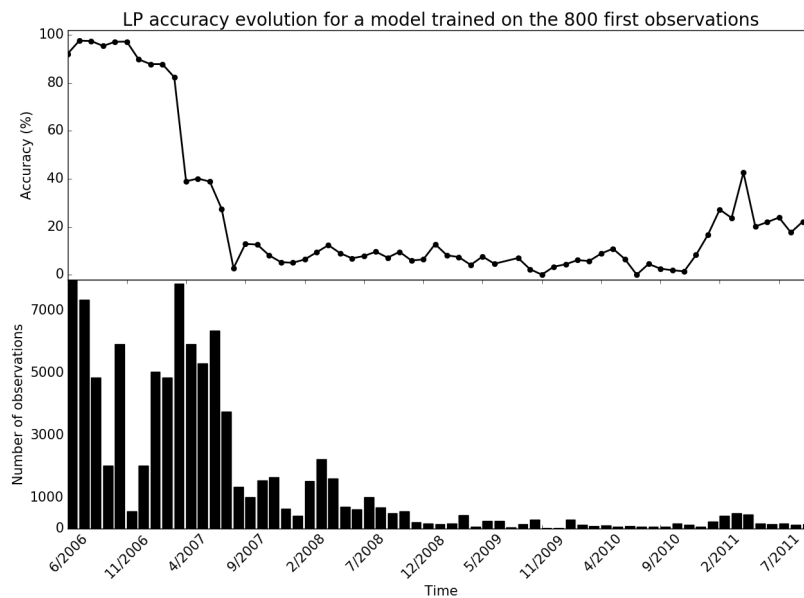


Fig. 3. Accuracy evolution of Long Period classification, from 2006 to 2011. The model is trained on the first 800 LP observations (all recorded in June 2006) and is then used to analyze the 6 years of recordings.

4.2.3 Third result: Continuous Analysis

In order to work with continuous analysis on the six years recording, we propose to build an evolving model where a new classifier is trained each month. This kind of specification is very important because it can be seen as a real-time analysis. To reach this analysis, we consider to train the models with a maximum of 800 past observations for each class. For very frequent classes such as Long Period LP, only a few days are necessary to gather 800 observations: new observations are therefore considered for each newly trained model. For less frequent classes such as Tremors TR, few months are needed to gather the 800 observations: the first models will be then trained on less than 800 examples, but once this limit is reached, only the 800 most recent observations will be used for training. Finally, for very sporadic classes such as Hybrids HYB, the whole dataset does not contain 800 observations, therefore each monthly model will be trained on the available observations. For the first months, the number of observations can be very low (≤ 10).

Global accuracy for this configuration reaches 80.3% of correct classification. The accuracy evolution on the six years of recordings can be seen in Figure 4. It is interesting to interpret the performance decay compared with cross validation results: firstly, and seen in section 4.2.2, observations within a same class do evolve with time. By using a newly trained model every month, we manage to follow and learn the classes evolutions. Global performances are indeed much higher in this configuration than in the previous one. However, the observations evolution is still visible with sudden accuracy drops in the plot: for example in September 2007, the accuracy drops from 95% to 38% and increases again to 98% in October. With the sudden change in signals shapes, observations from September are predicted with a model trained on data that are no longer representative of the class. As soon as the model has been updated however (in October), accuracy increases again. This phenomenon is observed on the six years and is notably visible between stable phases in the volcanic activity. Secondly, sporadic classes such as Hybrids HYB present a difficulty since very few observations are available (only six for the first model including this class): accuracy levels are then notably low (see Figure 5). With an increasing number

of observations used in training however (blue plot displaying cumulative number of observations), the accuracy level steadily increases to very good performances (100% in November 2011).

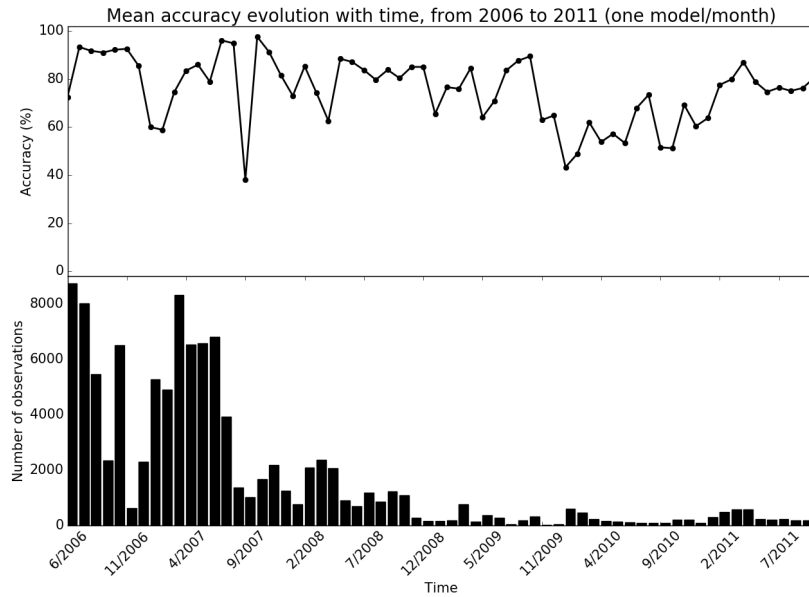


Fig. 4. Mean accuracy evolution, from 2006 to 2011. A new model is trained every month (SVM), on a maximum of 800 observations per class.

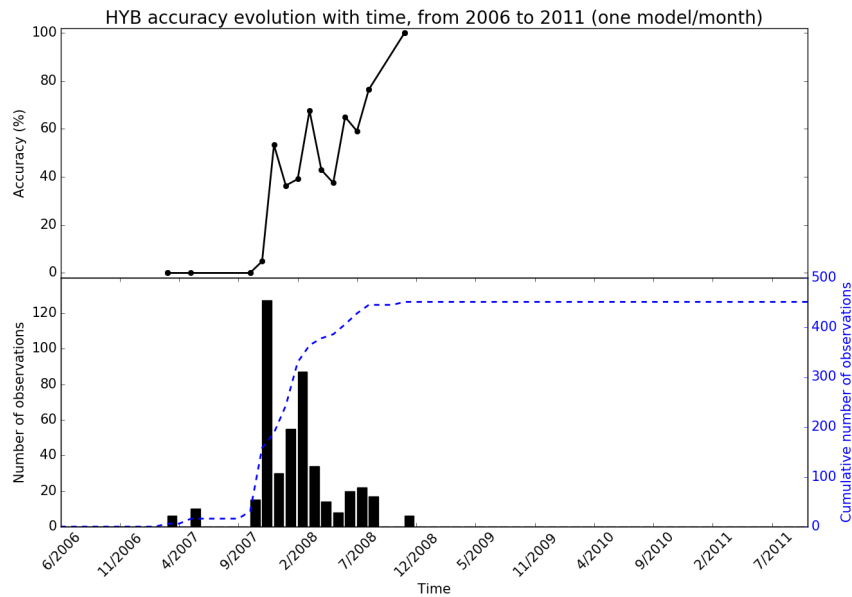


Fig. 5. Accuracy evolution of Hybrid classification, from 2006 to 2011. A new model is trained every month (SVM), on the past data (the maximum limit of 800 observations is never reached, the class is particularly sporadic).

5 CONCLUSION

As a conclusion to this paper, we insist on the issue of monitoring the volcanic hazard and the evaluation of associated risks. With the increasing number of volcano-seismic stations that record continuous signals, machine learning methods have become essential. Tools for the automatic analysis of volcano-seismic signals have started to appear, but are often limited and operational systems are still very scarce. In this paper, we reviewed the existing methods and proposed an efficient procedure for the automatic classification of volcano-seismic events. The model is tested using cross validation on 70856

observations, and reaches 92.2% of accuracy. Furthermore, we investigated the generalization capacities by building models in experimental conditions close to real-time analysis and displayed the need of more a couple hundreds of observation to represent the variety within a considered class. Finally, such tools can also be used to monitor the evolution of intra-class observations: we illustrated the process by analyzing six years of volcano-seismic recordings.

Prospects such tools are therefore more than promising, and machine learning architectures can be used to automatize the analysis of volcano-seismic data. Challenges for the future are the number of observations needed to classify a new class: in many applied cases, hundreds of signals are not available. The lowering of this constraint is a straightforward prospect in order to diffuse the use of such techniques in volcanic observatories. Semi-supervised learning methods can be considered and some study have used HMM in this purpose [40]. The use of several sensors and of hierarchical models could also be considered to improve the robustness of detection and classification results. Finally, alternative approaches to classify known data *without any a priori knowledge* on classes and detect new classes is a challenging prospect that would allow the monitoring of volcano previously non examined. The use of unsupervised learning methods can thus be considered.

6 ACKNOWLEDGMENTS

This work has been supported by a grant from Labex OSUG@2020 (Investissements d'avenir - ANR10 LABX56) and DGA/MRIS. GIPSA-Lab SIGMAPHY is part of Labex OSUG@2020 (ANR10 LABX56).

REFERENCES

- [1] UN-ISDR, "International Strategy for Disaster Reduction," *70th UN General Assembly*, no. 20c, p. 22 December 2015, 2015.
- [2] J. C. Lahr, B. A. Chouet, C. D. Stephens, J. A. Power, and R. A. Page, "Earthquake classification, location, and error analysis in a volcanic environment: implications for the magmatic system of the 1989-1990 eruptions at redoubt volcano, alaska," vol. 62, no. 1-4, pp. 137-151, 1994.
- [3] K. Aki, M. Fehler, and S. Das, "Source mechanism of volcanic tremor: fluid-driven crack models and their application to the 1963 kilauea eruption," *Journal of Volcanology and Geothermal Research*, vol. 2, no. 3, pp. 259 - 287, 1977.
- [4] B. Chouet, "Resonance of a fluid-driven crack: Radiation properties and implications for the source of long-period events and harmonic tremor," *Journal of Geophysical Research: Solid Earth*, vol. 93, no. B5, pp. 4375-4400, 1988.
- [5] J. Neuberg, R. Luckett, B. Baptie, and K. Olsen, "Models of tremor and low-frequency earthquake swarms on montserrat," *Journal of Volcanology and Geothermal Research*, vol. 101, no. 1, pp. 83 - 104, 2000.
- [6] E. Del Pezzo, F. Bianco, and I. Borgna, "Magnitude scale for lp events: a quantification scheme for volcanic quakes," *Geophysical Journal International*, vol. 194, no. 2, pp. 911-919, 2013.
- [7] B. A. Chouet and R. S. Matoza, "A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption," *Journal of Volcanology and Geothermal Research*, vol. 252, no. Supplement C, pp. 108 - 175, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0377027312003435>
- [8] B. A. Chouet, "Long-period volcano seismicity: its source and use in eruption forecasting," *Nature*, vol. 380, no. 6572, pp. 309-316, Mar 1996.
- [9] S. McNutt, "Volcanic tremor," *Encyclopedia of Earth System Science*, vol. 4, pp. 417-425, 1992.
- [10] A. J. Hotovec, S. G. Prejean, J. E. Vidale, and J. Gombert, "Strongly gliding harmonic tremor during the 2009 eruption of redoubt volcano," *Journal of Volcanology and Geothermal Research*, vol. 259, no. Supplement C, pp. 89 - 99, 2013, the 2009 Eruption of Redoubt Volcano, Alaska.
- [11] S. R. McNutt and T. Nishimura, "Volcanic tremor during eruptions: Temporal characteristics, scaling and constraints on conduit size and processes," *Journal of Volcanology and Geothermal Research*, vol. 178, no. 1, pp. 10 - 18, 2008, dynamics of Volcanic Explosions: Field Observations, Experimental Constraints and Integrated Modelling of Volcanic Explosions: Field Observations, Experimental Constraints and Integrated Modelling.
- [12] P. Traversa, O. Lengliné, O. Macedo, J.-P. Métaxian, J.-R. Grasso, A. Inza, and E. Taipe, "Short term forecasting of explosions at ubinas volcano, Perú," *Journal of Geophysical Research: Solid Earth*, vol. 116, no. B11, 2011.
- [13] M. Iguchi, H. Yakiwara, T. Tameguri, M. Hendrasto, and J. ichi Hirabayashi, "Mechanism of explosive eruption revealed by geophysical observations at the sakurajima, suwanosejima and semeru volcanoes," *Journal of Volcanology and Geothermal Research*, vol. 178, no. 1, pp. 1 - 9, 2008, dynamics of Volcanic Explosions: Field Observations, Experimental Constraints and Integrated Modelling of Volcanic Explosions: Field Observations, Experimental Constraints and Integrated Modelling.
- [14] T. H. Druitt, S. R. Young, B. Baptie, C. Bonadonna, E. S. Calder, A. B. Clarke, P. D. Cole, C. L. Harford, R. A. Herd, R. Luckett, G. Ryan, and B. Voight, "Episodes of cyclic vulcanian explosive activity with fountain collapse at soufrière hills volcano, montserrat," *Geological Society, London, Memoirs*, vol. 21, no. 1, pp. 281-306, 2002.
- [15] T. Ohminato, M. Takeo, H. Kumagai, T. Yamashina, J. Oikawa, E. Koyama, H. Tsuji, and T. Urabe, "Vulcanian eruptions with dominant single force components observed during the asama 2004 volcanic activity in japan," *Earth, Planets and Space*, vol. 58, no. 5, pp. 583-593, May 2006.
- [16] "Analysis of dynamics of vulcanian activity of ubinas volcano, using multicomponent seismic antennas," *Journal of Volcanology and Geothermal Research*, vol. 270, no. Supplement C, pp. 35 - 52, 2014.
- [17] S. Haykin, V. Tresp, and J. A. Benediktsson, "Big data: Practical applications [scanning the issue]," *Proceedings of the IEEE*, vol. 104, no. 11, pp. 2082-2084, Nov 2016.
- [18] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*. Springer series in statistics Springer, Berlin, 2001, vol. 1.
- [19] S. Zaugg, M. Van Der Schaar, L. Houégnigan, C. Gervaise, and M. André, "Real-time acoustic classification of sperm whale clicks and shipping impulses from deep-sea observatories," *Applied Acoustics*, vol. 71, no. 11, pp. 1011-1019, 2010.
- [20] S. Fagerlund, "Bird Species Recognition Using Support Vector Machines," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, 2007.
- [21] N. C. Han, S. V. Muniandy, and J. Dayou, "Acoustic classification of Australian anurans based on hybrid spectral-entropy approach," *Applied Acoustics*, vol. 72, no. 9, pp. 639-645, 2011.
- [22] S. Tucker and G. J. Brown, "Classification of transient sonar sounds using perceptually motivated features," *IEEE Journal of Oceanic Engineering*, vol. 30, no. 3, pp. 588-600, 2005.
- [23] I. Fujinaga and K. MacMillan, "Realtime recognition of orchestral instruments," in *Proceedings of the International Computer Music Conference (ICMC2000)*, vol. 141, 2000, pp. 141-143.
- [24] S. Esmaili, S. Krishnan, and K. Raahemifar, "Content based audio classification and retrieval using joint time-frequency analysis," in *ICASSP-2004, IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, 2004, pp. 9-12.
- [25] S. Falsaperla, S. Graziani, G. Nunnari, and S. Spampinato, "Automatic Classification Of Volcanic Earthquakes By Using Multi-Layered Neural Networks," *Natural Hazards*, vol. 13, pp. 205-228, 1996.

- [26] R. A. Lara-Cueva, A. S. Moreno, J. C. Larco, and D. S. Benitez, "Real-Time Seismic Event Detection Using Voice Activity Detection Techniques," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 12, pp. 5533–5542, 2016.
- [27] A. Esposito, F. Giudicepietro, L. DAuria, S. Scarpetta, M. Martini, M. Coltelli, and M. Marinaro, "Unsupervised neural analysis of very-long-period events at stromboli volcano using the self-organizing maps," *Bulletin of the Seismological Society of America*, vol. 98, no. 5, pp. 2449–2459, 2008.
- [28] H. Langer, S. Falsaperla, and G. Thompson, "Application of artificial neural networks for the classification of the seismic transients at soufriere hills volcano, montserrat," *Geophysical research letters*, vol. 30, no. 21, 2003.
- [29] H. Langer, S. Falsaperla, T. Powell, and G. Thompson, "Automatic classification and a-posteriori analysis of seismic event identification at Soufrière Hills volcano, Montserrat," *Journal of Volcanology and Geothermal Research*, vol. 153, no. 1-2 SPEC. ISS., pp. 1–10, 2006.
- [30] P. Alasonati, J. Wassermann, and M. Ohrnberger, "Signal classification by wavelet-based hidden markov models application to seismic signals of volcanic origin," 01 2006.
- [31] M. C. Benítez, J. Ramírez, J. C. Segura, J. M. Ibáñez, J. Almendros, A. García-Yeguas, and G. Cortés, "Continuous HMM-based seismic-event classification at deception Island, Antarctica," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 1, 2007, pp. 138–146.
- [32] I. Alvarez, L. Garcia, G. Cortes, C. Benitez, and A. De La Torre, "Discriminative feature selection for automatic classification of volcano-seismic signals," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 2, pp. 151–155, 2012.
- [33] J. M. Ibáñez, C. Benítez, L. A. Gutiérrez, G. Cortés, A. García-yeguas, and G. Alguacil, "The classification of seismo-volcanic signals using Hidden Markov Models as applied to the Stromboli and Etna volcanoes," *Journal of Volcanology and Geothermal Research*, vol. 187, no. 3-4, pp. 218–226, 2009.
- [34] E. Pezzo Del, A. Esposito, F. Giudicepietro, M. Marinaro, M. Martini, and S. Scarpetta, "Discrimination of Earthquakes and Underwater Explosions Using Neural Networks Using Neural Networks," *Bulletin of the Seismological Society of America*, vol. 93, no. February 2003, pp. 215–223, 2003.
- [35] F. Giacco, A. Esposito, S. Scarpetta, Giudicepietro F, and M. Marinaro, "Support Vector Machines and MLP for automatic classification of seismic signals at Stromboli volcano," in *Proceedings of the 19th Italian Workshop on Neural Nets*, 2009.
- [36] S. Scarpetta, F. Giudicepietro, E. C. Ezin, S. Petrosino, E. Del Pezzo, M. Martini, and M. Marinaro, "Automatic classification of seismic signals at Mt. Vesuvius volcano, Italy, using neural networks," *Bulletin of the Seismological Society of America*, vol. 95, no. 1, pp. 185–196, 2005.
- [37] M. Ibs-von Seht, "Detection and identification of seismic signals recorded at Krakatau volcano (Indonesia) using artificial neural networks," *Journal of Volcanology and Geothermal Research*, vol. 176, no. 4, pp. 448–456, 2008.
- [38] G. Curilem, J. Vergara, G. Fuentealba, G. Acuña, and M. Chacón, "Classification of seismic signals at Villarrica volcano (Chile) using neural networks and genetic algorithms," *Journal of Volcanology and Geothermal Research*, vol. 180, no. 1, pp. 1–8, 2009.
- [39] A. Köhler, M. Ohrnberger, and F. Scherbaum, "Unsupervised pattern recognition in continuous seismic wavefield records using Self-Organizing Maps," *Geophysical Journal International*, vol. 182, no. October, pp. 1619–1630, 2010.
- [40] C. Hammer, M. Beyreuther, and M. Ohrnberger, "A seismic-event spotting system for volcano fast-response systems," *Bulletin of the Seismological Society of America*, vol. 102, no. 3, pp. 948–960, 2012.
- [41] A. Maggi, V. Ferrazzini, C. Hibert, B. François, B. Patrice, and A. Amandine, "Implementation of a multi-station approach for automated event classification at Piton de la Fournaise volcano - in press," *Seismological Research Letters*.
- [42] M. Orozco, M. Garcia, R. Duin, and C. Castellanos, "Dissimilarity Based Classification of Seismic Signals at Nevado del Ruiz Volcano," *Earth Sciences Research Journal*, vol. 10, no. 2, pp. 57–65, 2006.
- [43] M. Bicego, C. Acosta-Munoz, and M. Orozco-Alzate, "Classification of seismic volcanic signals using hidden-markov-model-based generative embeddings," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 6, pp. 3400–3409, 2013.
- [44] G. Cortes, M. Carmen Benitez, L. Garcia, I. Alvarez, and J. M. Ibanez, "A Comparative Study of Dimensionality Reduction Algorithms Applied to Volcano-Seismic Signals," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 1, pp. 253–263, 2016.
- [45] Ö. Yilmaz, *Seismic data analysis: Processing, inversion, and interpretation of seismic data*. Society of exploration geophysicists, 2001.
- [46] N. Langet, "Détection et caractérisation massives de phénomènes sismologiques pour la surveillance d' événements traditionnels et la recherche systématique de," Ph.D. dissertation, 2014.
- [47] N. Langet, A. Maggi, A. Michelini, and F. Brenguier, "Continuous kurtosis-based migration for seismic event detection and location, with application to Piton de la Fournaise volcano, La Réunion," *Bulletin of the Seismological Society of America*, vol. 104, no. 1, pp. 229–246, 2014.
- [48] R. V. Allen, "Automatic earthquake recognition and timing from single traces," *Bulletin of the Seismological Society of America*, vol. 68, no. 5, pp. 1521–1532, 1978.
- [49] R. Allen, "Automatic phase pickers: Their present use and future prospects," *Bulletin of the Seismological Society of America*, vol. 72, no. 6, pp. S225–242, 1982.
- [50] H. Langer, S. Falsaperla, M. Masotti, R. Campanini, S. Spampinato, and A. Messina, "Synopsis of supervised and unsupervised pattern classification techniques applied to volcanic tremor data at Mt Etna, Italy," *Geophysical Journal International*, vol. 178, no. 2, pp. 1132–1144, 2009.
- [51] M. Ohrnberger, "Continuous Automatic Classification of Seismic Signals of Volcanic Origin at Mt . Merapi , Java , Indonesia Continuous Automatic Classification of Seismic Signals of Volcanic Origin at Mt . Merapi , Java , Indonesia," no. July 2001, 2003.
- [52] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [53] R. A. White, A. D. Miller, L. Lynch, and J. Power, "Observations of hybrid seismic events at soufriere hills volcano, montserrat: July 1995 to september 1996," *Geophysical Research Letters*, vol. 25, no. 19, pp. 3657–3660, 1998.
- [54] R. M. Harrington and E. E. Brodsky, "Volcanic hybrid earthquakes that are brittle-failure events," *Geophysical Research Letters*, vol. 34, no. 6, 2007.
- [55] L. Narváez, H. Cepeda, J. Stix *et al.*, "tornillo-type seismic signals at galeras volcano, colombia, 19921993," *Journal of Volcanology and Geothermal Research*, vol. 77, no. 1, pp. 159 – 171, 1997.