Geophysical Research Abstracts
Vol. 19, EGU2017-8842-2, 2017
EGU General Assembly 2017

Automatic classification of seismo-volcanic signatures

Marielle Malfante (1), Mauro Dalla Mura (1), Jérôme Mars (1), Orlando Macedo (2), Adolfo Inza (2), and Jean-Philippe Métaxian (3)

(1) Univ. Grenoble Alpes, CNRS, GISPA-Lab, 38000 Grenoble, France (marielle.malfante@gipsa-lab.fr)., (2) Instituto Geofísico del Perú, Lima, Peru., (3) ISTerre, IRD, CNRS, Université Savoie Mont Blanc, F-73376 Le Bourget du Lac, France.

The prediction of volcanic eruptions and the evaluation of their associated risks is still a timely and open issue. For this purpose, several types of signals are recorded in the proximity of volcanoes and then analysed by experts. Typically, seismic signals that are considered as precursor or indicator of an active volcanic phase are detected and manually classified. In this work, we propose an architecture for automatic classification of seismo-volcanic waves.

The system we propose is based on supervised machine learning. Specifically, a prediction model is built from a large dataset of labelled examples by the means of a learning algorithm (Support Vector Machine or Random Forest). Four main steps are involved: (i) preprocess the signals, (ii) from each signal, extract features that are useful for the classes discrimination, (iii) use an automatic learning algorithm to train a prediction model and (iv) classify (i.e. assign a semantic label) newly recorded and unlabelled examples. Our main contribution lies in the definition of the feature space used to represent the signals (i.e. in the choice of the features to extract from the data). Feature vectors describe the data in a space of lower dimension with respect to the original one. Ideally, signals are separable in the feature space depending on their classes. For this work, we consider a large set of features (79) gathered from an extensive state of the art in both acoustic and seismic fields. An analysis of this feature set shows that for the application of interest, 11 features are sufficient to discriminate the data.

The architecture is tested on 4725 seismic events recorded between June 2006 and September 2011 at Ubinas, the most active volcano of Peru. Six main classes of signals are considered: volcanic tremors (TR), long period (LP), volcano-tectonic (VT), explosion (EXP), hybrids (HIB) and tornillo (TOR). Our model reaches above 90% of accuracy, thereby validating the proposed architecture and the features used. We illustrate the limited influence of the choice of the classifier used (Random Forest or Support Vector Machine) by showing that accuracy results are stable regardless the classifier.

In order to build a model capable to analyse continuous recordings with few inputs from experts, we are currently considering semi-supervised classification techniques.