Geophysical Research Abstracts
Vol. 18, EGU2016-15705, 2016
EGU General Assembly 2016

Automatic classification of endogenous seismic sources within a landslide body using random forest algorithm

Floriane Provost (1), Clément Hibert (1), Jean-Philippe Malet (1), André Stumpf (1,2), and Cécile Doubre (1)

(1) Institut de Physique du Globe; UMR7516, Université de Strasbourg/EOST, CNRS, Strasbourg, France (f.provost@unistra.fr), (2) Laboratoire Image, Ville, Environnement, UMR 7362, Université de Strasbourg, Strasbourg, France (andre.stumpf@unistra.fr)

Different studies have shown the presence of microseismic activity in soft-rock landslides. The seismic signals exhibit significantly different features in the time and frequency domains which allow their classification and interpretation. Most of the classes could be associated with different mechanisms of deformation occurring within and at the surface (e.g. rockfall, slide-quake, fissure opening, fluid circulation). However, some signals remain not fully understood and some classes contain few examples that prevent any interpretation. To move toward a more complete interpretation of the links between the dynamics of soft-rock landslides and the physical processes controlling their behaviour, a complete catalog of the endogeneous seismicity is needed.

We propose a multi-class detection method based on the random forests algorithm to automatically classify the source of seismic signals. Random forests is a supervised machine learning technique that is based on the computation of a large number of decision trees. The multiple decision trees are constructed from training sets including each of the target classes. In the case of seismic signals, these attributes may encompass spectral features but also waveform characteristics, multi-stations observations and other relevant information. The Random Forest classifier is used because it provides state-of-the-art performance when compared with other machine learning techniques (e.g. SVM, Neural Networks) and requires no fine tuning. Furthermore it is relatively fast, robust, easy to parallelize, and inherently suitable for multi-class problems.

In this work, we present the first results of the classification method applied to the seismicity recorded at the Super-Sauze landslide between 2013 and 2015. We selected a dozen of seismic signal features that characterize precisely its spectral content (e.g. central frequency, spectrum width, energy in several frequency bands, spectrogram shape, spectrum local and global maxima) and its waveform (e.g. duration, ratio between the maximum and the mean/median of the envelope amplitude, envelope kurtosis and skewness, polarization). This preliminary study shows that the classification accuracy is high, and insensitive to sampling permutations of training/validation sets.