

EGU2020-7495

<https://doi.org/10.5194/egusphere-egu2020-7495>

EGU General Assembly 2020

© Author(s) 2020. This work is distributed under the Creative Commons Attribution 4.0 License.



## Integration of Machine Learning on Distributed Acoustic Sensing Surveys

**Camille Jestin**<sup>1</sup>, Clément Hibert<sup>2</sup>, Gaëtan Calbris<sup>1</sup>, and Vincent Lanticq<sup>1</sup>

<sup>1</sup>FEBUS OPTICS, Pau, France (camille.jestin@febus-optics.com)

<sup>2</sup>Institut de Physique du Globe de Strasbourg/CNRS, Strasbourg University, Strasbourg, France

Distributed Acoustic Sensing (DAS) is an innovative technique which has been recently employed for near-surface geophysics purposes. It involves the use of fibre-optic cable as a sensor. The fibre is analysed by sending a laser pulse from an interrogator unit. The phase of the backscattered signal contains the information on the strain on the cable, enabling the detection of a passing acoustic wave with enough energy for the cable excitation. Allowing the interrogation of long profiles and the generation of a dense spatial sampling, uneasy to obtain with classic geophysical techniques, DAS instrumentation then proved its relevance for seismic applications but also for infrastructure monitoring.

During DAS acquisition, and more precisely when closely looking at infrastructures integrity, it is necessary to clearly identify the source of the acoustic vibrations at the structure neighbourhood. Indeed, in the context of pipeline monitoring for example, it appears important to be able to classify events which generate seismic signals recorded by DAS systems and which can be related to a potential threat for the structure. In order to launch an alarm if necessary, the source identification must be fast, accurate and robust. Moreover, because DAS acquisition can generate traces every few meters along fibres of tens of kilometres, the used machine-learning algorithm must demonstrate its ability to handle a big amount of data.

In this study, we analyse the efficiency of the Random Forests (RF) machine-learning algorithm applied to data acquired with DAS system for the discrimination of event sources. RF algorithm has been selected because of its ability to handle large numbers of attributes related to signal characteristics and to enable a good reliability for the discrimination of sources. This algorithm has already proved its efficiency for automated classification of seismic waveforms (e.g. earthquakes, volcanic tremors, rock falls, avalanches, etc.).

We focus our study on tests lead along a gas pipeline instrumented with fibre-optic cable. Different third-party works have been conducted: excavation, saw sections, drill, jackhammer, etc. We work on the discrimination of six classes of seismic source. After running a detection phase based on a threshold on signal energy, we obtain several hundred of exploitable seismic traces to inject to the RF algorithm. We demonstrate the efficiency of the application of machine learning on DAS data to discriminate seismic waveforms from the correct class, with an overall precision on our test set of 99%.

