Automatic classification of endogenous landslide seismicity using the Random Forest supervised classifier

F. Provost, C. Hibert, J.-P. Malet

AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2016GL070709

Key Points:
- Landslide seismic sources are automatically classified using the Random Forest supervised classifier with 71 attributes
- The sensitivity of the classification is up to 93% in the case of four different classes of seismic events
- The large amount of attributes enables the method to be implemented easily and automatically in many contexts

Correspondence to:
F. Provost,
f.provost@unistra.fr

Citation:
Provost, F., C. Hibert, and J.-P. Malet (2017), Automatic classification of endogenous landslide seismicity using the Random Forest supervised classifier, *Geophys. Res. Lett.*, *44*, 113–120, doi:10.1002/2016GL070709.

Received 3 AUG 2016
Accepted 5 DEC 2016
Accepted article online 9 DEC 2016
Published online 5 JAN 2017

Automatic classification of endogenous landslide seismicity using the Random Forest supervised classifier

F. Provost[1], C. Hibert[1], and J.-P. Malet[1]

[1] Institut de Physique du Globe de Strasbourg, CNRS UMR 7516, University of Strasbourg/EOST, Strasbourg, France

Abstract The deformation of slow-moving landslides developed in clays induces endogenous seismicity of mostly low-magnitude events ($M_L < 1$). Long seismic records and complete catalogs are needed to identify the type of seismic sources and understand their mechanisms. Manual classification of long records is time-consuming and may be highly subjective. We propose an automatic classification method based on the computation of 71 seismic attributes and the use of a supervised classifier. No attribute was selected a priori in order to create a generic multi-class classification method applicable to many landslide contexts. The method can be applied directly on the results of a simple detector. We developed the approach on the seismic network of eight sensors of the Super-Sauze clay-rich landslide (South French Alps) for the detection of four types of seismic sources. The automatic algorithm retrieves 93% of sensitivity in comparison to a manually interpreted catalog considered as reference.

1. Introduction

Recent studies have demonstrated the presence of endogenous seismicity induced by the deformation of slow-moving clay-rich landslides, whereas aseismic creeping was previously assumed [*Gomberg et al.*, 1995; *Tonnellier et al.*, 2013; *Walter et al.*, 2013]. Slidequakes have been recorded on these unstable slopes proving the presence of material failures and shearing at the contact with the bedrock or directly within the moving mass. Locally, rockfalls can also be recorded on steep slopes [*Tonnellier et al.*, 2013]. Tremor-like signals have also been observed [*Gomberg et al.*, 2011] and may be linked to fluid transfer or transient slip. In the case of clay-rich landslides, analysis of the microseismicity is a challenging task because the signals are of low magnitude ($M_L < 1$), low amplitude (<10000 nm/s), and are generally highly attenuated at short distances (<200 m). Dense seismic arrays should therefore be installed over long observation periods to obtain numerous signals with high signal-to-noise ratio (SNR). Analysis of long seismological records (>2 years) has been realized [*Spillmann et al.*, 2007; *Helmstetter and Garambois*, 2010] resulting in the detection of several thousands of landslide seismic events. In these studies, the seismic events are classified manually after detection and rely on the personal experience of the human operator which can be subjective and time-consuming.

Automatic classification methods have been developed for detecting the sources in volcanic areas [*Langer et al.*, 2006; *Curilem et al.*, 2009] to differentiate earthquakes and blasts [*Fäh and Koch*, 2002; *Laasri et al.*, 2015] or for characterizing large rockslides [*Dammeier et al.*, 2016]. For multiclass problems, many classifiers were used such as hidden Markov models (HMMs), artificial neural networks, and support vector machines (SVMs), mainly on a reduced number of seismic attributes [*Curilem et al.*, 2009; *Hibert et al.*, 2014; *Ruano et al.*, 2014]. Recently, some studies [*Beyreuther and Wassermann*, 2008; *Ruano et al.*, 2014; *Quang et al.*, 2015] focused on the classification of continuous seismic records discriminating the background noise from the signal of interest. HMM was modified to detect one type of signal from few to one single example which is interesting for the detection of rare seismic sources [*Hammer et al.*, 2012, 2013; *Dammeier et al.*, 2016]. However, the use of a unique seismic signal as reference lacks to capture the influence of the travel path effects on the waveform and the frequency content of the recorded signal [*Hammer et al.*, 2013]. The authors hence suggested to use one example for different source-receiver distances. Finally, *Ruano et al.* [2014] applied SVM for the discrimination of earthquake and explosion from background noise, but the method requires a further interpretation of the detected events.

We propose a generic (applicable to various objects) and automatic (no fine-tuning required) method to classify the endogenous seismicity of slow-moving landslides.The Super-Sauze landslide data set is used as an

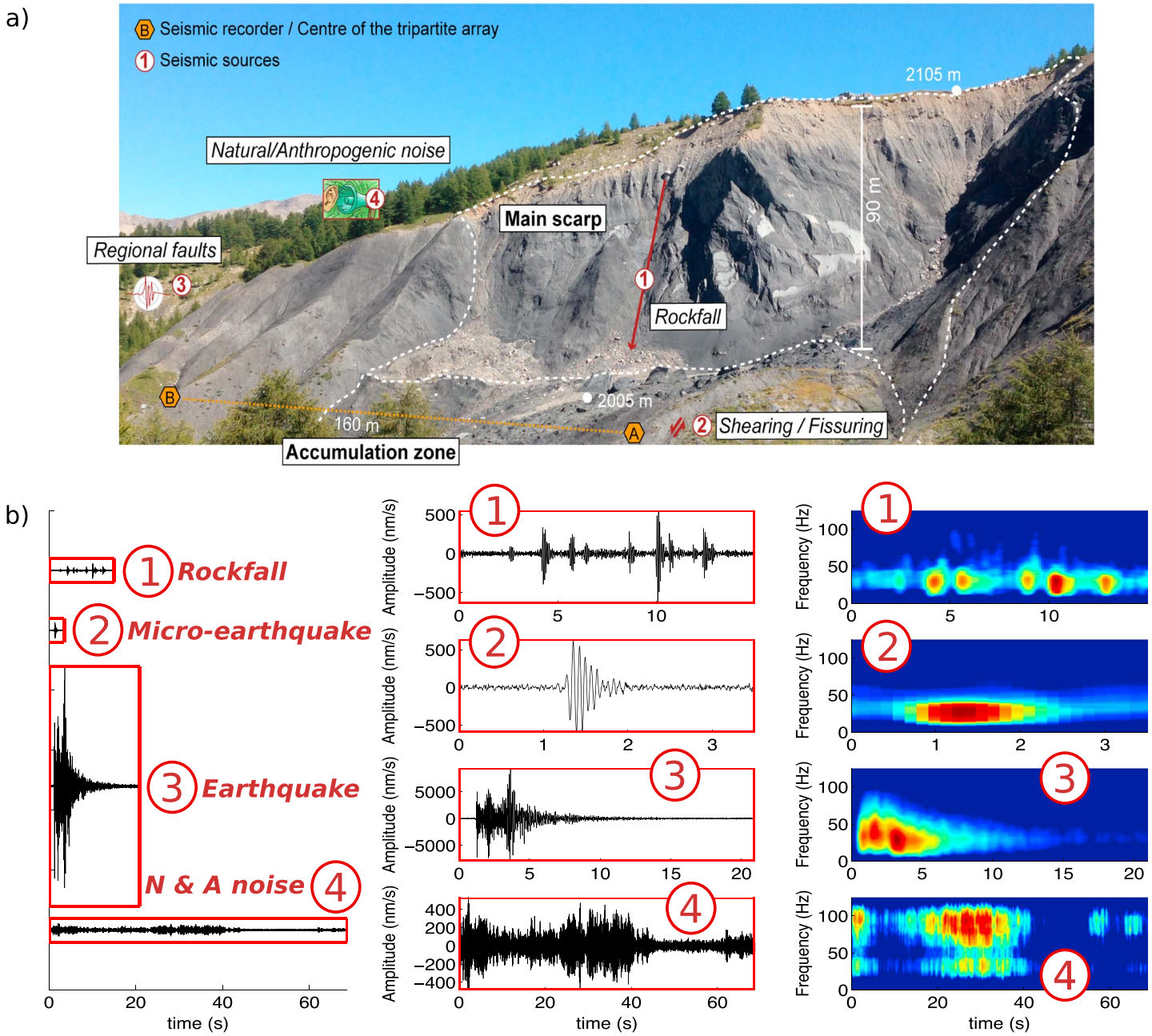

Figure 1. Seismic signals, sources, and mechanisms observed on clay-rich landslides based on analyst interpretation: (a) Location of the microseismic arrays (A and B) and of the sources at the Super-Sauze landslide (Southeast France); (b) examples of seismic signal for each class with the display of the waveform (trace with the higher SNR, first and second column) and (third column) stacked spectrogram.

example to test the method. We decided to use the Random Forest (RF) supervised classifier [*Breiman*, 2001] on a large training set. Four classes of seismic events are identified: two classes are related to events associated with the landslide deformation (e.g., quakes and rockfalls) and two classes of external sources (e.g., regional/global earthquakes and natural and anthropogenic noise). To train the model, 71 seismic attributes are computed (waveform, spectral content, spectrogram content, polarization, and attributes related to the seismic network geometry). We here focused on the results of the classification method on a sample data set and discuss possible improvement for its implementation as a near real-time classifier.

2. Data

The seismic records are acquired by two permanent arrays of the French Landslide Observatory OMIV (Observatoire Multi-disciplinaire des Instabilités de Versants) installed at the east and west sides of the Super-Sauze landslide (Southeast France) developed in weathered black marls [*Malet et al.*, 2005]. The seismic stations consist of short period seismometers (Noemax and Sercel L4C) with a flat response in the range 5–100 Hz; the signals are recorded with two broadband seismic recorders (RefTek 130S-01) at a 250 Hz sample frequency. The seismometers are arranged as tripartite array of 40 m layout around one three-component center site and three vertical one-component organized as equilateral triangle. The array thus forms a six-channel seismic recorder. The experiment data set consists of three acquisition periods from 11 October to 19 November 2013, from 10 to 30 November 2014, and from 9 June to 15 August 2015. The investigated data set consists of 418 "Rockfall" events, 239 "Quake" events, 407 "Earthquake" events (EQ), and 395 "Natural/Anthropogenic noise" events ("N and A" noise) (Figure 1).

1. The Rockfall events take place mostly in the landslide main scarp where rigid blocks fall from steep slopes (>100 m high). The block impacts are visible both in the signal waveform and in the stacked spectrogram for most of the events but can also present cigar shapes when finer material is falling.
2. The Quake events are likely to be triggered by material failures, surface fissure openings, and shear stress release at the landslide boundaries or at the contact with the bedrock. They are usually strongly attenuated and not recorded by all the seismometers and last less than 5 sec.
3. The Earthquake events cluster all the regional seismic events triggered in the region [*Jenatton et al.*, 2007] and the teleseisms. Their pseudospectrograms have typically a triangle shape with a decrease of high-frequency content with time.
4. The Natural/Anthropogenic noise events cluster all the anthropogenic (footsteps, car and helicopter motors, ski lifts, etc.) and environmental (wind, storm, water streams, etc.) noise in the vicinity of the landslide. Those events usually last several tens of seconds and illuminate either several frequencies or only specific ones in the spectrogram.

For more details on the endogenous signals observed at Super-Sauze landslide, the reader may refer to [*Helmstetter and Garambois*, 2010; *Walter et al.*, 2012; *Tonnellier et al.*, 2013]. It must be noted that it can be difficult to differentiate certain signals such as a succession of quakes from small-volume rockfalls or to distinguish footsteps and small-volume rockfalls.

3. Methods

The processing chain consists in successive stages with (1) the detection of signals of potential interest, (2) the computation of the seismic attributes, and (3) the classification of the signals. The method relies on a detection algorithm that must be carefully chosen depending depending on the signals of interest and the study area. We used a spectrogram analysis [*Helmstetter and Garambois*, 2010] to detect the signals. The method is equivalent to a STA/LTA algorithm applied in the frequency domain. An event is detected when the seismic signal spectrum is larger than 1.5 times the noise spectrum level. The spectrogram is computed in the frequency range 5–100 Hz for each vertical sensor. Second, we calculated 71 seismic attributes (Table 1) to describe the signals in terms of signal waveform, signal frequency content, pseudospectrogram, polarity, and some attributes related to the seismic network geometry. The signal waveform and the frequency content are computed on the trace with the highest SNR. The spectrogram attributes are computed on the summed spectrogram, and the polarity attributes are computed separately on the three-component seismometers. Some of these attributes are commonly used in signal classification and have been described in previous studies [*Bessason et al.*, 2007; *Curilem et al.*, 2009; *Hammer et al.*, 2012; *Hibert et al.*, 2014]. We explored here additional attributes related to the seismic network geometry. All the attributes are computed automatically on the raw signal window and do not require any human interpretation such as wave onset identification or location. Only the kurtosis attribute requires setting manually various frequency bands (here 5–10 Hz, 5–50 Hz, 10–70 Hz, 50–100 Hz, and 5–100 Hz).

The selected supervised classifier is the *Random Forest* algorithm (RF), which is an ensemble learning method based on decision trees [*Breiman*, 2001]. It is based on several (> 500) decision trees trained on the data set. The class is assigned by the majority of the decision tree votes. RF has proven to be one of the most efficient algorithm for the classification of 121 complex data sets among 17 families of classifiers [*Fernández-Delgado et al.*, 2014]. Moreover, it is simple to use and does not require any fine-tuning [*Stumpf and Kerle*, 2011]. It also

Table 1. Attributes Table

Number	Description	Formula
	Waveform Attributes:	
1	Duration	$t_f - t_i$, with t_i and t_f: beginning and end of the signal
2	Ratio of the mean over the maximum of the envelop signal	–
3	Ratio of the median over the maximum of the envelop signal	–
4	Ratio between ascending and descending time	$\frac{t_{max}-t_i}{t_f-t_{max}}$, with t_{max} : time of the largest amplitude
5	Kurtosis of the raw signal (peakness of the signal)	$\frac{m_4}{\sigma^4}$, with m_4: fourth moment, σ: standard deviation
6	Kurtosis of the envelop	see 5
7	Skewness of the raw signal	$\frac{m_3}{\sigma^3}$, with m_3: third moment
8	Skewness of the envelop	see 7
9	Number of peaks in the autocorrelation function	–
10	Energy in the first third part of the autocorrelation function	$\int_0^{\frac{T}{3}} C(\tau)d\tau$, with T: signal duration, C: autocorrelation function
11	Energy in the remaining part of the autocorrelation function	see 10
12	Ratio of 11 and 10	–
13–17	Energy of the signal filtered in 5–10 Hz, 10–50 Hz, 5–70 Hz, 50–100 Hz, and 5–100 Hz	$\int_0^T y_f(t)dt$, with y_f: filtered signal in the frequency range [f1-f2]
18–22	Kurtosis of the signal in 5–10 Hz, 10–50 Hz, 5–70 Hz, 50–100 Hz, and 5–100 Hz frequency range	see 5
23	RMS between the decreasing part of the signal and $l(t) = Y_{max} - \frac{Y_{max}}{t_f-t_{max}}t$	$\sqrt{\overline{Y(t)-l(t)}^2}$, with Y: envelop of the signal
	Spectral attributes:	
24	Mean of the DFT	DFT: discrete Fourier transform
25	Max of the DFT	–
26	Frequency at the maximum	–
27	Central frequency of the 1st quartile	–
28	Central frequency of the 2nd quartile	–
29	Median of the normalized DFT	–
30	Variance of the normalized DFT	–
31	Number of peaks ($> 0.75\ DFT_{max}$)	DFT_{max}: maximum of the DFT
32	Number of peaks in the autocorrelation function	–
33	Mean value for the peaks	–
34–37	Energy in $[0, \frac{1}{4}]Nyf, [\frac{1}{4}, \frac{1}{2}]Nyf, [\frac{1}{2}, \frac{3}{4}]Nyf, [\frac{3}{4}, 1]Nyf$	$\int_{f_1}^{f_2} \mathrm{DFT}(f)\mathrm{d}f$ with f_1, f_2: the considered frequency range
38	Spectral centroid	$\gamma_1 = \frac{m_2}{m_1}$, with m_1 and m_2 are the first and second moment
39	Gyration radius	$\gamma_2 = \sqrt{\frac{m_3}{m_2}}$, with m_3 is the third moment
40	Spectral centroid width	$\sqrt{\gamma_1^2 - \gamma_2^2}$
	Spectrogram[a] attributes:	
41	Kurtosis of the maximum of all discrete Fourier transforms (DFTs) as a function of time t	$\mathrm{Kurtosis}\left[\max_{t=0,..,T}(\mathrm{SPEC}(t,f))\right]$ with SPEC(t,f): the spectrogram
42	Kurtosis of the maximum of all DFTs as a function of time t	see 41
43	Mean ratio between the maximum and the mean of all DFTs	$\mathrm{mean}\left(\frac{\max(\mathrm{SPEC})}{\mathrm{mean}(\mathrm{SPEC})}\right)$
44	Mean ratio between the maximum and the median of all DFTs	see 43
45	Number of peaks in the curve showing the temporal evolution of the DFTs maximum	–
		–
46	Number of peaks in the curve showing the temporal evolution of the DFTs mean	–
47	Number of peaks in the curve showing the temporal evolution of the DFTs median	–
48	Ratio between 45 and 46	–
49	Ratio between 45 and 47	–

Table 1. (continued)

Number	Description	Formula
50	Number of peaks in the curve of the temporal evolution of the DFTs central frequency	–
51	Number of peaks in the curve of the temporal evolution of the DFTs maximum frequency	–
52	Ratio between 50 and 51	–
53	Mean distance between the curves of the temporal evolution of the DFTs maximum frequency and mean frequency	–
54	Mean distance between the curves of the temporal evolution of the DFTs maximum frequency and median frequency	–
55	Mean distance between the 1st quartile and the median of all DFTs as a function of time	–
56	Mean distance between the 3rd quartile and the median of all DFTs as a function of time	–
57	Mean distance between the 3rd quartile and the 1st quartile of all DFTs as a function of time	–
58	Number of gaps in the signal	–
	Network attributes:	
59	SNR maximum	–
60	Station with maximum SNR	–
61	Station with maximum amplitude	–
62	Station with minimum amplitude	–
63	Ratio between attributes 62 and 61	–
64	Mean correlation	–
65	Maximum correlation	–
66	Mean correlation lag in between station	–
67	Standard deviation correlation lag in between station	–
	Polarity attributes:	
68	Rectilinearity	$1-\frac{\lambda_{11}+\lambda_{22}}{2\lambda_{33}}$ with $\lambda_{33}>>\lambda_{22}>>\lambda_{11}$
69	Azimuth	$\arctan(\lambda_{23}/\lambda_{13})\times 180/\pi$
70	Dip	$\arctan(\lambda_{33}/\sqrt{\lambda_{23}^2+\lambda_{13}^2})\times 180/\pi$
71	Planarity	$1-\frac{2\lambda_{11}}{\lambda_{33}+\lambda_{22}}$

[a]The spectrogram is the collection of the DFTs computed for signal windows of 1s with an overlap of 90%. The spectrogram is represented as a two-dimensional matrix representing the evolution of the frequency content (rows) through time (columns).

enables to compute which attributes are the most discriminant for the classification. This is done through the estimation of the *Variable Importance* as defined by *Breiman* [2001] and consists in randomly swapping the values of one attribute over all the samples of the training data set. The *Variable Importance* is the variation in the out-of-bag error computed before and after the permutation. The larger the error variation the more important is the attribute. We worked with the TreeBagger version of RF implemented in the MATLAB Statistical Toolbox. For the tests, successively 10%, 30%, 50%, and 70% of each class were randomly selected as training set for the classifier. The RF model was evaluated for each test on 70 events randomly chosen in each class in the events not used in the training set. One hundred model runs were performed for each test, and the classification results are averaged.

Table 2. Mean Confusion Matrix for 100 Runs of Random Forest[a]

	Class. Rockfall	Class. Quake	Class. EQ	Class. N and A Noise
Ref. Rockfall	94%	1%	3%	2%
Ref. Quake	3%	93%	3%	1%
Ref. EQ	1%	2%	94%	3%
Ref. N and A Noise	2%	3%	3%	92%

[a]The classified (Class.) events are represented with respect to the events of the reference catalog (Ref.).

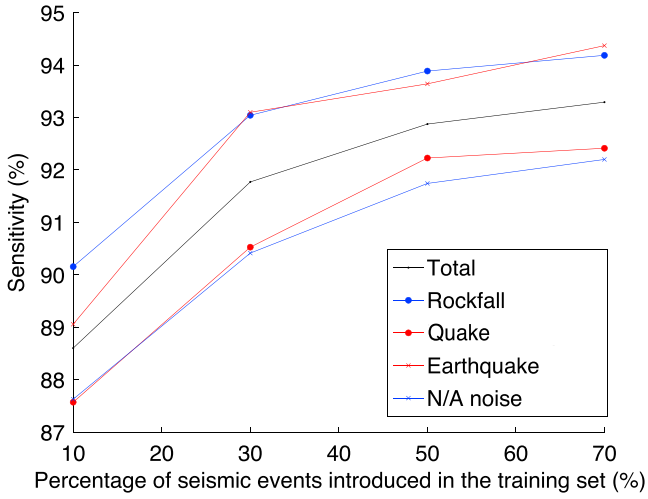

Figure 2. Sensitivity for different sizes of the training set: 10%, 30%, 50%, and 70% of each class. The training set is randomly selected 100 times, and the sensitivity is averaged for each run.

4. Results

The sensitivity reaches 93% ± 1.5% and the specificity 97% ± 0.5% in average for 100 runs of Random Forest. Table 2 presents the confusion matrix that represents the comparison between the reference interpretation and the automated classification of the data set. The Earthquake and Rockfall events are better classified with a sensitivity of 94%, whereas the Quake and the Natural/Anthropogenic noise events have a sensitivity of 93% and 92%, respectively.

The sensitivity increases with the number of examples in the training set (Figure 2). We observe that for above 50% of examples introduced in the training set for each class, the increase in sensitivity is not significative anymore (< 1%). However, it seems that the sensitivity of the classes Earthquake and Natural/Anthropogenic noise increases more than for the two others; thus, more examples of these two classes may improve the sensitivity.

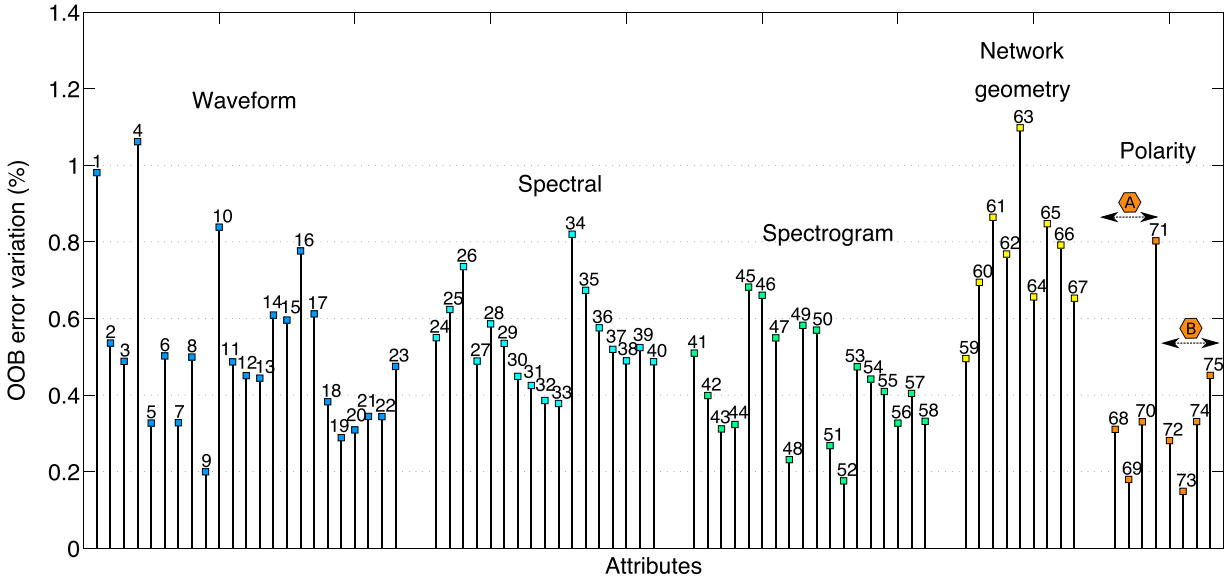

Figure 3. Mean variable importance for 100 runs of Random Forest. The attributes are presented with their respective numbers (Table 1) and by category: Waveform, Spectral (attributes computed on the FFT of the signal), Spectrogram (attributes computed on the pseudospectrogram of the signal), Network geometry (attributes taking into account the geometry of the network), and Polarity. The polarity attributes of the three-component seismometers are represented successively for microseismic array A and B.

The importance of the attributes in the classification are presented in Figure 3. In our case, we observe that the most discriminant attribute is the ratio between the maximum and minimum amplitude recorded in the network (A63). The other most discriminant attributes are, successively, the ratio between the ascending and descending duration of the signal (A4), the signal duration (A1), the station with the maximum amplitude (A61), the median correlation (A65), the energy in the first third of the autocorrelation function (A10), the energy in the signal spectrum between 0 and 31.25 Hz (A34), the planarity (A71), the mean correlation lag in between traces (A66), and the energy of the signal filtered in 50–100 Hz (A16). No attribute from the spectrogram analysis appears in the 10 most discriminant attributes. The less discriminant attributes are respectively the azimuth (A69, A73) and the ratio between the number of peaks in the pseudospectrogram function (A52).

5. Discussion and Conclusion

The more classes are discriminated by an attribute, the more important the attribute is. Therefore, all the seismic sources can be separated by attribute 63 ($A63_{EQ} >= A63_{N\ and\ Anoise} > A63_{ROC} > = A63_{Q}$) because the waves created by the local sources travel through the subsurface and may be highly attenuated at short distances (<100 m), while for the external sources, the waves travel in the consolidated bedrock and are thus less attenuated. In the same manner, the ratio between the ascending and descending phase can separate the highly nonsymmetrical sources (Rockfall and Earthquake) from the symmetrical signal generated by the quakes ($A4_{ROC} > = A4_{Q} > A4_{EQ}$). The duration finally discriminates mostly the "Natural and Anthropogenic noise" events ($A1_{N\ and\ Anoise} > 20s$) and the quakes ($A1_{Q} < 5s$). Finally, the next attributes in the order of importance mainly discriminate only one class: the "N and A noise" events are highly not correlated compared to the other signals (A65), and the rockfall events mainly occur in the scarp so the sensors with the maximum amplitude are the closer ones (A61).

The azimuth of the particle motion is, in theory, a useful information to discriminate the sources because it gives the source location direction. Here it seems that the station with the maximum amplitude (A61) gives more robust information on the source location by discriminating the rockfall events. The attributes implying the number of peaks are also not very discriminating here probably because it is rather difficult to set an efficient threshold above which a peak is considered relevant.

Considering only the 10 most discriminant attributes slightly deteriorates the classification with a sensitivity of 92% $\pm$ 1.5% while not reducing significantly the computing time (e.g., the attributes can currently be calculated from the raw signal in approximately 1 second). Therefore, no attribute should be removed from the model even the one with the lower discriminating rate. Moreover, the list of attributes enables the method to be easily applied to various contexts since they allow to fully describe each signal. The most discriminant attributes will probably be different at other landslides, but this will not decrease the accuracy of the method. The number of attributes also enables the method to be adaptative to possible temporal variations of the source mechanisms.

Sensitivity is a common measure used to evaluate a classifier, but this criterion requires a reference catalog constructed by one analyst. Because the human interpretation is subjective [*Langer et al.*, 2006; *Hibert et al.*, 2014; *Laasri et al.*, 2015], the sensitivity may not reflect the complexity of the data set. We, thus, tested the human subjectivity on a smaller data set of 60 events with 20 human analysts and compared the results. The total of the analysts' votes agrees fully with the reference interpretation, thus validating the reference catalog used to compute the sensitivity. The mean sensitivity for the human interpretation is 82% (Table 3). It must be noted that the information given to the human analysts was slightly different to the one introduced in the statistical model as the seismic network geometry and the polarity attributes were not presented to the analysts. We ran the RF model with the same information as given to the analysts (A1 to A58) and obtained a mean sensitivity of 90% (Table 3). In both cases, the automatic method obtains higher or similar maximal sensitivity than the analysts (Table 3). The automatic method is hence comparable to the human analysis, while a larger sensitivity of the automatic method would overfit the interpretation of one particular analyst.

We propose an automatic classifier based on RF and a large number of attributes to describe the seismic signals. The obtained sensitivity is 93% $\pm$ 1.5% for a complex multiclass problem (low-magnitude events and intraclass heterogeneity). The method requires at least 150 examples for the different event types to train the model; it further allows separating the highly heterogeneous class (e.g., "N and A noise") with the same success rate than the other classes. RF provides probability estimates on the classification that are useful to accept or reject a new classification and determine when the model needs to be retrained. The latter could also provide

Table 3. Sensitivity Results for the Set of 20 Human Analysts, for the RF Model With All Attributes Taken Into Account and for the RF Model With A1 to A58[a]

Mean Sensitivity	82%	93%	90%
Maximum Sensitivity	95%	96%	94%
Minimum Sensitivity	58%	90%	86%

[a]The RF model is tested 100 times using 70% of the data set as training set and testing the model on the 30% of the data set (not selected as training set).

new insights to detect changes of the seismicity with time. In summary, the method is easily applicable to classify the seismicity of various objects (volcanoes, geothermal fields, and earthquake detection) and can be used even for studies where only one single sensor is available. The implementation of the method for real-time applications is in progress.

Acknowledgments
This work was carried out with the support of the French National Research Agency (ANR) through the projects HYDROSLIDE "Hydrogeophysical Monitoring of Clayey Landslides" and SAMCO "Adaptation de la Société aux Risques Gravitaires en Montagne dans un Contexte de Changement Global" and of the Open Partial Agreement Major Hazards of Council of Europe through the project "Development of cost-effective ground-based and remote monitoring systems for detecting landslide initiation." Continuous seismic data used here were provided by the Observatoire Multi-disciplinaire des Instabilités de Versant (OMIV) [*RESIF/OMIV*, 2015]. The authors thank A. Stumpf (EOST) for constructive discussion on supervised classification algorithms, A. Helmstetter (ISTerre) for discussion on the classification of the event types, and the volunteers at EOST who participated in our classification experiment. The authors also acknowledge the anonymous reviewers for their helpful remarks.

References

Bessason, B., G. Eiríksson, O. Thórarinsson, A. Thórarinsson, and S. Einarsson (2007), Automatic detection of avalanches and debris flows by seismic methods, *J. Glaciol.*, *53*(182), 461–472, doi:10.3189/002214307783258468.

Beyreuther, M., and J. Wassermann (2008), Continuous earthquake detection and classification using discrete hidden Markov models, *Geophys. J. Int.*, *175*(3), 1055–1066, doi:10.1111/j.1365-246X.2008.03921.x.

Breiman, L. (2001), Random forests, *Mach. Learn.*, *45*(1), 5–32, doi:10.1023/A:1010933404324.

Curilem, G., J. Vergara, G. Fuentealba, G. Acuña, and M. Chacòn (2009), Classification of seismic signals at Villarrica volcano (Chile) using neural networks and genetic algorithms, *J. Volcanol. Geotherm. Res.*, *180*(1), 1–8, doi:10.1016/j.jvolgeores.2008.12.002.

Dammeier, F., J. R. Moore, C. Hammer, F. Haslinger, and S. Loew (2016), Automatic detection of alpine rockslides in continuous seismic data using Hidden Markov Models, *J. Geophys. Res. Earth Surf.*, *121*, 351–371, doi:10.1002/2015JF003647.

Fäh, D., and K. Koch (2002), Discrimination between earthquakes and chemical explosions by multivariate statistical analysis: A case study for Switzerland, *Bull. Seismol. Soc. Am.*, *92*(5), 1795–1805, doi:10.1785/0120010166.

Fernández-Delgado, M., E. Cernadas, S. Barro, and D. Amorim (2014), Do we need hundreds of classifiers to solve real world classification problems, *J. Mach. Learn. Res.*, *15*(1), 3133–3181.

Gomberg, J., P. Bodin, W. Savage, and M. E. Jackson (1995), Landslide faults and tectonic faults, analogs?: The Slumgullion earthflow, Colorado, *Geology*, *23*(1), 41–44, doi:10.1130/0091-7613(1995)023<0041:LFATFA>2.3.CO;2.

Gomberg, J., W. Schulz, P. Bodin, and J. Kean (2011), Seismic and geodetic signatures of fault slip at the Slumgullion Landslide Natural Laboratory, *J. Geophys. Res.*, *116*, B09404, doi:10.1029/2011JB008304.

Hammer, C., M. Beyreuther, and M. Ohrnberger (2012), A seismic-event spotting system for volcano fast-response systems, *Bull. Seismol. Soc. Am.*, *102*(3), 948–960, doi:10.1785/0120110167.

Hammer, C., M. Ohrnberger, and D. Fäh (2013), Classifying seismic waveforms from scratch: A case study in the alpine environment, *Geophys. J. Int.*, *192*(1), 425–439, doi:10.1093/gji/ggs036.

Helmstetter, A., and S. Garambois (2010), Seismic monitoring of Séchilienne rockslide (French Alps): Analysis of seismic signals and their correlation with rainfalls, *J. Geophys. Res.*, *115*, F03016, doi:10.1029/2009JF001532.

Hibert, C., et al. (2014), Automated identification, location, and volume estimation of rockfalls at Piton de la Fournaise volcano, *J. Geophys. Res. Earth Surf.*, *119*, 1082–1105, doi:10.1002/2013JF002970.

Jenatton, L., R. Guiguet, F. Thouvenot, and N. Daix (2007), The 16,000-event 2003–2004 earthquake swarm in Ubaye (French Alps), *J. Geophys. Res.*, *112*, B11304, doi:10.1029/2006JB004878.

Laasri, E. H. A., E.-S. Akhouayri, D. Agliz, D. Zonta, and A. Atmani (2015), A fuzzy expert system for automatic seismic signal classification, *Expert Syst. Appl.*, *42*(3), 1013–1027, doi:10.1016/j.eswa.2014.08.023.

Langer, H., S. Falsaperla, T. Powell, and G. Thompson (2006), Automatic classification and a-posteriori analysis of seismic event identification at Soufrière Hills volcano, Montserrat, *J. Volcanol. Geotherm. Res.*, *153*(1–2), 1–10, doi:10.1016/j.jvolgeores.2005.08.012.

Malet, J.-P., D. Laigle, A. Remaître, and O. Maquaire (2005), Triggering conditions and mobility of debris flows associated to complex earthflows, *Geomorphology*, *66*(1–4), 215–235, doi:10.1016/j.geomorph.2004.09.014.

Quang, P. B., P. Gaillard, Y. Cano, and M. Ulzibat (2015), Detection and classification of seismic events with progressive multi-channel correlation and hidden Markov models, *Comput. Geosci.*, *83*, 110–119, doi:10.1016/j.cageo.2015.07.002.

RESIF/OMIV (2015), RESIF/OMIV: French Multidisciplinary Observatory of Versant Instabilities,RESIF:Réseau Sismologique et Géodésique Français, doi:10.15778/RESIF.MT.

Ruano, A., G. Madureira, O. Barros, H. Khosravani, M. Ruano, and P. Ferreira (2014), Seismic detection using support vector machines, *Neurocomputing*, *135*, 273–283, doi:10.1016/j.neucom.2013.12.020.

Spillmann, T., H. Maurer, A. G. Green, B. Heincke, H. Willenberg, and S. Husen (2007), Microseismic investigation of an unstable mountain slope in the Swiss Alps, *J. Geophys. Res.*, *112*, B07301, doi:10.1029/2006JB004723.

Stumpf, A., and N. Kerle (2011), Object-oriented mapping of landslides using Random Forests, *Remote Sens. Environ.*, *115*(10), 2564–2577, doi:10.1016/j.rse.2011.05.013.

Tonnellier, A., A. Helmstetter, J.-P. Malet, J. Schmittbuhl, A. Corsini, and M. Joswig (2013), Seismic monitoring of soft-rock landslides: The Super-Sauze and Valoria case studies, *Geophys. J. Int.*, *193*(3), 1515–1536, doi:10.1093/gji/ggt039.

Walter, M., C. Arnhardt, and M. Joswig (2012), Seismic monitoring of rockfalls, slide quakes, and fissure development at the Super-Sauze mudslide, French Alps, *Eng. Geol.*, *128*, 12–22, doi:10.1016/j.enggeo.2011.11.002.

Walter, M., J. Gomberg, W. Schulz, P. Bodin, and M. Joswig (2013), Slidequake generation versus viscous creep at softrock-landslides: Synopsis of three different scenarios at Slumgullion landslide, Heumoes slope, and Super-Sauze mudslide, *J. Environ. Eng. Geophys.*, *18*(4), 269–280, doi:10.2113/JEEG18.4.269.