

A Neural-based Algorithm for Landslide Detection at Stromboli Volcano: Preliminary Results

Antonietta M. ESPOSITO^{a,1}, Flora GIUDICEPIETRO^a, Luca D'AURIA^a,
Rosario PELUSO^a, Marcello MARTINI^a

^a *Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Napoli Osservatorio Vesuviano, Napoli, Italy*

Abstract. This study presents a neural-based algorithm for the automatic detection of landslides on Stromboli volcano (Italy). It has been shown that landslides are an important short-term precursor of effusive eruptions of Stromboli. In particular, an increase in the occurrence rate of landslides was observed a few hours before the beginning of the February 2007 effusive eruption. Automating the process of detection of these signals can help analysts and represents a useful tool for the monitoring of the stability of the Sciara del Fuoco flank of Stromboli volcano. A multi-layer perceptron neural network is here applied to continuously discriminate landslides from other signals recorded at Stromboli (e.g., explosion quakes, tremor signals), and its output is used by an automatic system for the detection task. To correctly represent the seismic data, coefficients are extracted from both the frequency domain, using the linear predictive coding technique, and the time domain, using temporal waveform parameterization. The network training and testing was carried out using a dataset of 537 signals, from 267 landslides and 270 records that included explosion quakes and tremor signals. The classification results were 99.5% predictive for the best net performance, and 98.7% when the performance was averaged over different net configurations. Thus, this detection system was effective when tested on the 2007 effusive eruption period. However, continuing investigations into different time intervals are needed, to further define and optimize the algorithm.

Keywords. Landslides, detection, neural network, seismic signals, volcano seismicity.

Introduction

During the last effusive eruption of Stromboli volcano (February 27, 2007) it was noted that the opening of the effusive vents was preceded by a few hours by an increase in the landslide rate [10]. Detection of the landslides can be considered a significant short-term precursor of such effusive eruptions of Stromboli. Furthermore, automating this task will help analysts in the classification task of these signals, and will simultaneously allow monitoring of the stability of the northwest flank of the volcano, known as “Sciara del Fuoco”, which showed significant instabilities during the 2002-2003 eruption.

¹ Corresponding author: Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Napoli Osservatorio Vesuviano, Via Diocleziano, 328, 80124 Napoli, Italy. e-mail: antonella.esposito@ov.ingv.it.

In this study, we implement a system for automatic detection of landslides that is based on a multi-layer perceptron (MLP) neural network [1]. Neural networks have been successfully applied in the field of volcano seismology [2, 6, 7, 12, 13] for the detection and discrimination of different types of seismic signals [5, 11]. This has been possible due to their flexibility and the possibility of arranging their architecture according to specific applications.

Landslides on Stromboli volcano have a characteristic seismic waveform, with a higher spectral content than typical explosion quakes and volcanic tremor signals, and a cigar-shaped amplitude envelope. Usually, these two features are used by analysts to visually detect such signals. The neural algorithm also exploits these features to discriminate them. A first step of data manipulation is expected to reduce their dimension, extracting only the distinctive attributes that uniquely identify the seismic signals. In a second stage, the network performs the landslide classification and the system uses the net output to carry out the detection.

In the following, we first briefly describe the seismic activity of Stromboli volcano and its monitoring network, then we introduce the dataset and illustrate the pre-processing techniques that were applied for data parameterization. In section 4, we describe the MLP network and the automatic system. Section 5 shows the results, and the last section is devoted to conclusions.

1. The volcanic activity of Stromboli

The Strombolian activity is typically characterized by individual explosions that emit gasses and pyroclastic fragments, and which usually occur 6-7 times per hour. However, occasionally effusive phases occur. The December 2002 effusive phase was characterized by a large landslide on the “Sciara del Fuoco” flank of Stromboli, which generated a tsunami with a maximum wave height of about 10 m. After this episode, the northwest flank became unstable, and as many as 50 landslide signals were recorded per day. Furthermore, the last effusive episode of February 2007 was preceded by an increment in the landslide rate a few hours before.

During these past two effusive eruptions, Vulcanian explosions took place on April 5, 2003, and on March 15, 2007. The eruptive processes at Stromboli are such that the seismic signals generated span a wide range of frequencies [3]. At present, the typical signals recorded for Stromboli are: volcanic tremor, a continuous signal with a 1-3 Hz frequency range; explosion quakes, signals with no distinct seismic phases and with a 1-6 Hz frequency range; landslides, with an emergent onset and a 1-10 Hz frequency range; and other seismic signals, such as long-period earthquakes (LP), volcano-tectonic earthquakes (VT) and very long period events (VLP) that are associated with the explosions.

In January, 2003, the “Istituto Nazionale di Geofisica e Vulcanologia” (INGV) started to install a broadband seismic network for the monitoring of the volcanic activity of Stromboli [4]. The network is composed of 13 digital stations (as shown in Figure 1) that are equipped with three-component broadband Guralp CMG-40T seismometers, with frequency responses of 0.02-60s. The data are acquired with a sampling rate of 50 sps and are continuously transmitted via transmission control protocol/ internet protocol (TCP/IP) to the recording center in Naples at the Vesuvius Observatory (INGV) (www.ov.ingv.it/stromboli/).

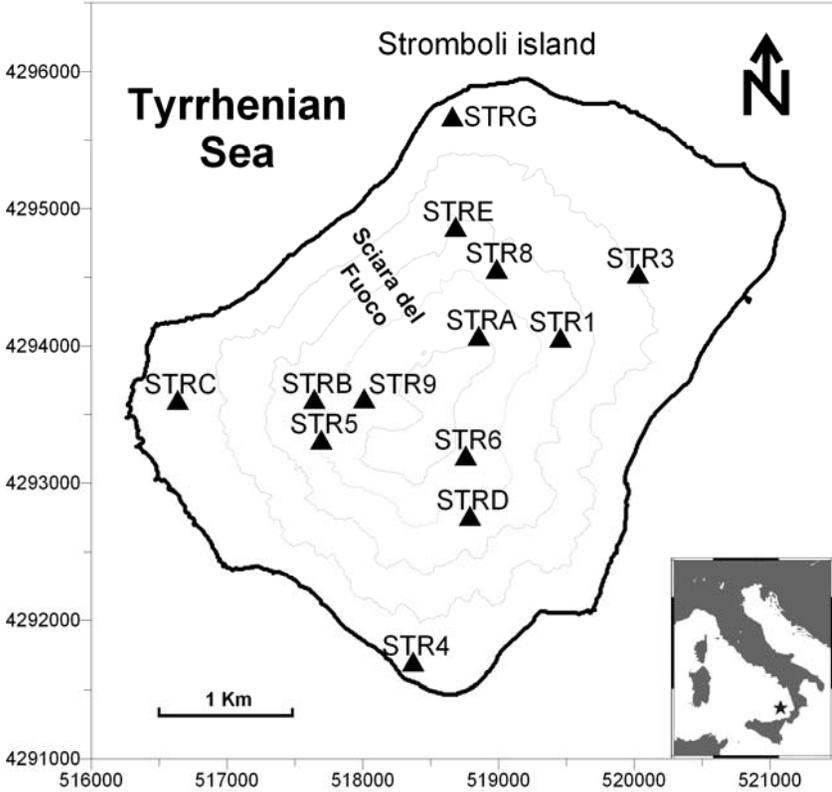


Figure 1. The Stromboli Island monitoring network, indicating the 13 seismic stations (black triangles). The scale is expressed as UTC coordinates. Inset: location of Stromboli Island with respect to the Italian peninsula.

2. The seismic data

The dataset used to train and test the MLP network consisted of 537 records from five of the seismic stations (STR1, STRA, STR8, STR5, STRB; Figure 1). In particular, it was composed of 267 landslides and 270 signals that included 130 explosion quakes and 140 tremor signals. For each event, a 24-s-long record was taken at a 50-Hz sampling frequency.

Figure 2 shows the waveforms and the associated spectrograms of three typical signals recorded at Stromboli: an explosion quake (Figure 2A), a landslide (Figure 2B) and a tremor window (Figure 2C). From Figure 2 it is possible to see that the landslide has a particular cigar-like shape, it is higher in frequency than the explosion quakes and tremor signals, and its typical waveform has an emergent onset.

The automatic system was tested for the analysis of the February, 2007, effusive eruption. Here, the time interval considered ranged from January 1, 2007 to June 1, 2007 to include the beginning (February 27) and the end (April 2) of the eruption.

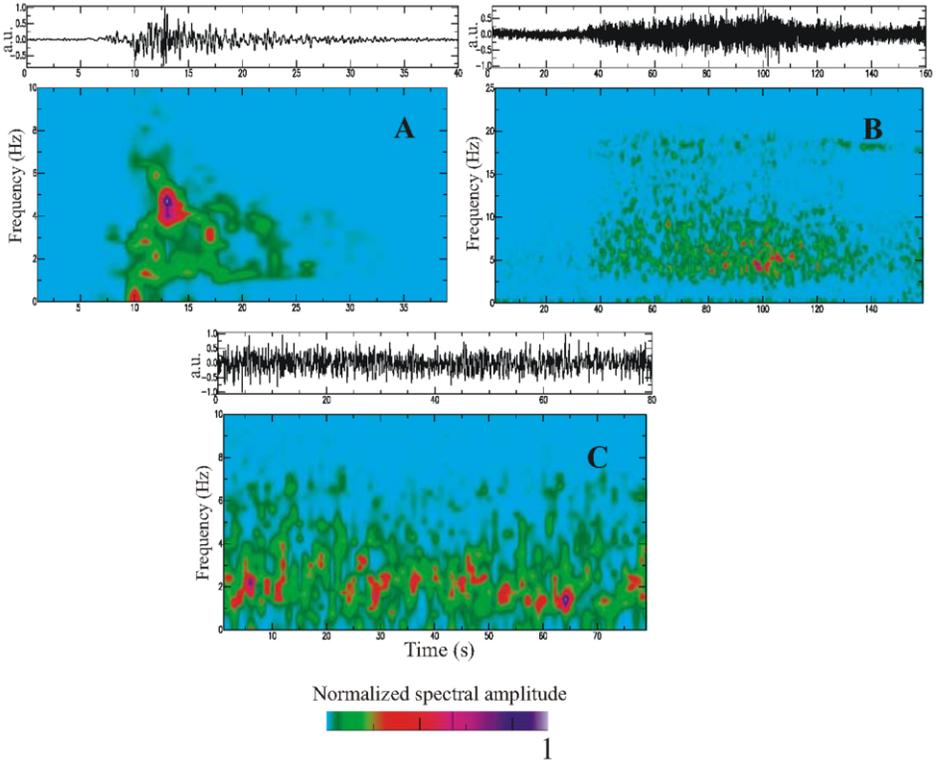


Figure 2. Seismograms and the associated spectrograms of (A) an explosion quake, (B) a landslide, and (C) a tremor window, recorded at Stromboli.

3. Data pre-processing

The MLP-based neural algorithm requires a preliminary phase of signal pre-processing to perform the landslide classification. Finding which and how many features uniquely identify and describe a seismic signal is a critical task. It is important to obtain a vector representation of the seismic data with significant, and not redundant, features, and with a reasonable dimension, in order to reduce the computational burden.

As indicated above, landslides have a well-known frequency content and a cigar-like waveform, so that analysts can visually distinguish them from explosion quakes and tremor signals on seismograms. Thus, to exploit previous results [8], the proposed algorithm uses the linear predictive coding (LPC) technique [9] to obtain the spectral features. LPC predicts a signal sample through a linear combination of its p previous samples:

$$s^*(n) \approx c_1 s(n-1) + c_2 s(n-2) + \dots + c_p s(n-p) \quad (1)$$

where $s(n)$ is the signal sample at time n , $s^*(n)$ is its prediction, and p is the model order, the value of which is problem-dependent. The prediction coefficients c_i , for $i = 1, \dots, p$, are estimated through an optimization procedure that reduces the error between

the real signal and its LPC estimate. In our case, we extracted $p = 6$ LPC coefficients from each of the eight 5.12-s-long Hanning windows into which we divided the signals, with an overlap of 2.56 s. In this way, each signal is described by a vector of 48 spectral features.

The time-domain information of the signal is also taken into account, and computed as the correctly normalized difference between the maximum and minimum signal amplitudes within a 1-s-long window. Thus for a 24-s-long signal, we obtain a vector of 24 temporal elements.

As a result, the pre-processing step is able to reduce the size of each seismic signal from 1200 samples to 72 coefficients ($6 \times 8 = 48$ frequency features + 24 time features). This compression provided a good compromise between robustness and compactness of the signal representation. Robustness refers to the fact that such encoding accurately describes the properties of the seismic signals, while compactness means that it does not contain redundant information allowing a faster computational processing.

4. MLP-based algorithm

The neural network used for this landslide classification was a supervised feed-forward MLP. The network architecture (shown in Figure 3) had an input layer of 72 units, a hidden layer with a varying number of nodes, and an output layer with only one unit. The activation function of the hidden units is a non-linear hyperbolic-tangent function, expressed as:

$$\tanh(a/2) \equiv \frac{1 - e^{-a}}{1 + e^{-a}} \quad (2)$$

while the output nodes use a logistic sigmoidal function, computed as:

$$\sigma(a) \equiv \frac{1}{1 + e^{-a}} \quad (3)$$

The network output y is calculated as:

$$y = \sigma\left(\sum_j w_{jk} \tanh\left(\sum_i W_{ij} x_i\right)\right) \quad (4)$$

where x_i is the i th input unit, W_{ij} and w_{jk} are the weights on the connections from the input to the hidden layer and from the hidden layer to the output layer, respectively.

The network undergoes supervised learning, meaning that the input is presented to the net together with the desired response or target vector. The difference between the target vector and the net output is calculated, and the weights on the connections between the units are adjusted according to a learning rule that minimizes the error in this approximation. The process is repeated until the net performance is acceptable. In the present study, a Quasi-Newton [1] learning algorithm and a cross-entropy error function [1] were used. For this binary problem of discrimination between landslides and the other seismic signals, this function is defined as:

$$E = - \sum_n \left\{ t^n \ln y^n + (1 - t^n) \ln(1 - y^n) \right\} \quad (5)$$

where n is the number of training samples, t is the target vector, and y is the net output.

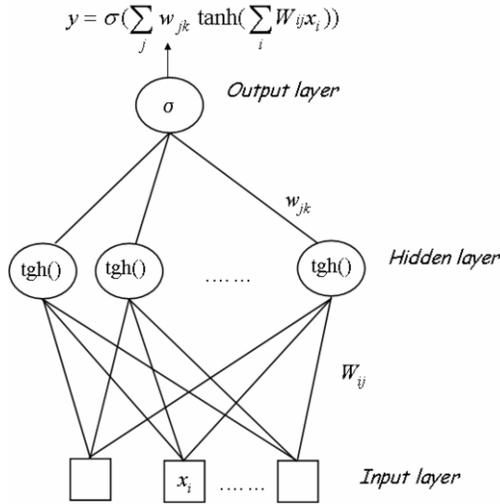


Figure 3. The architecture of a multi-layer perceptron network.

The network was trained using just over 60% (5/8) of the total dataset, while the remaining data were used for the testing. Furthermore, to assess the net robustness, the final network performance was computed by averaging the percentages of correct classification over six different network configurations that were obtained through six different data permutations and random weight initializations (see Table 1). Other network parameters, such as the number of hidden units and the training cycles, were established empirically through trial-and-error processes (see Table 2), as they depended on the specific task.

Table 1. The correct classification for each of the six net configurations, and the averaged network performance.

Classification Task	Performance (% correct classification)						
	Different network configurations						% Average
Landslides/ (explosion quakes, tremor signals)	99.50	99.50	97.01	97.51	99.50	99.00	98.67

Table 2. The number of the hidden and output units, and the training cycles of the MLP network used for the classification task.

Classification Task	Hidden units	Output units	Training cycles
Landslides/ (explosion quakes, tremor signals)	4	1	70

The automatic system performed the analysis by applying the neural algorithm over 10-min-long seismic traces. The algorithm then used 24-s-long sliding windows of

the signal that overlapped by 12 s, and outputs its probability of being a landslide. Figure 4 shows the net response for a typical landslide. Setting a user-defined threshold (e.g. 0.9), the system identifies the start and the end of each landslide signal. Figure 5 graphically illustrates the system output over a 48 h time interval.

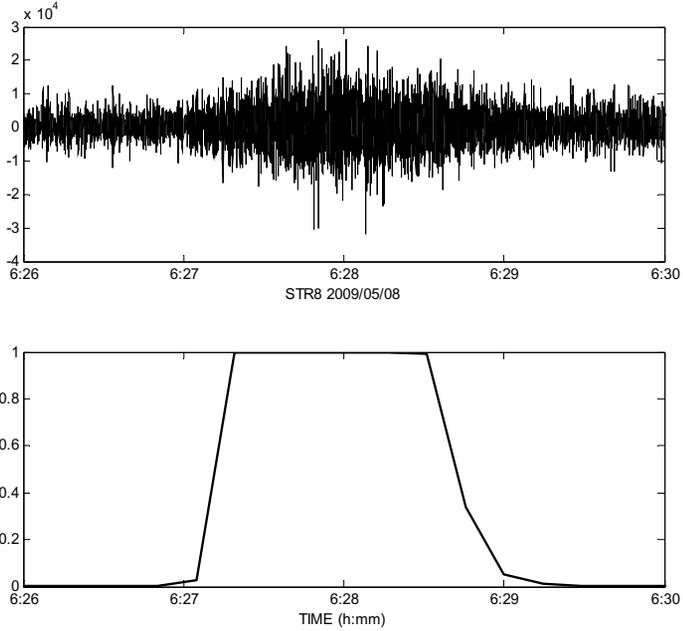


Figure 4. An example of the network response for a typical landslide.

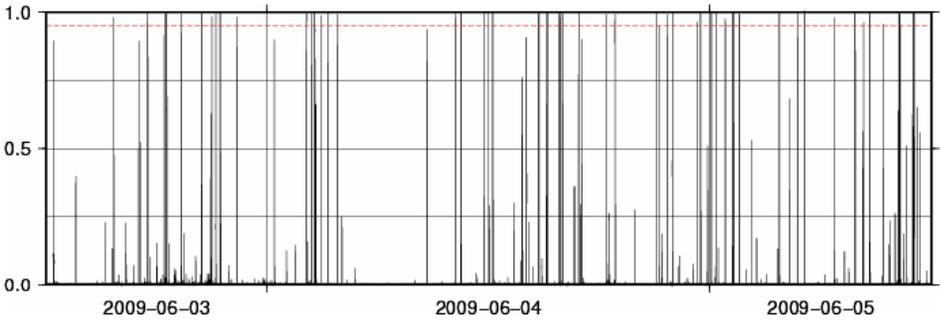


Figure 5. The output of the automatic system over a 48 h time interval. The dashed red line represents the user-defined threshold.

A set of tools was also developed for the use of the system output for different types of analysis, such as the percentage duration of the landslides, their daily number, and the duration of each landslide detected.

5. Results

The implemented system has been used for the analysis of the February 2007 effusive eruption. Figure 6 shows an example of the analysis performed over the period from 1/1/2007 to 1/6/2007. Figure 6A shows the percentage duration of the landslide signals over 24-h periods, where the red lines indicate the beginning and the end of the lava flow. Figure 6B shows the number of landslides per day, and Figure 6C shows the duration of each detected landslide (blue squares) and their daily average duration (red squares). An increase in March 2007 is evident, when the lava flow was active, and also during April 2007.

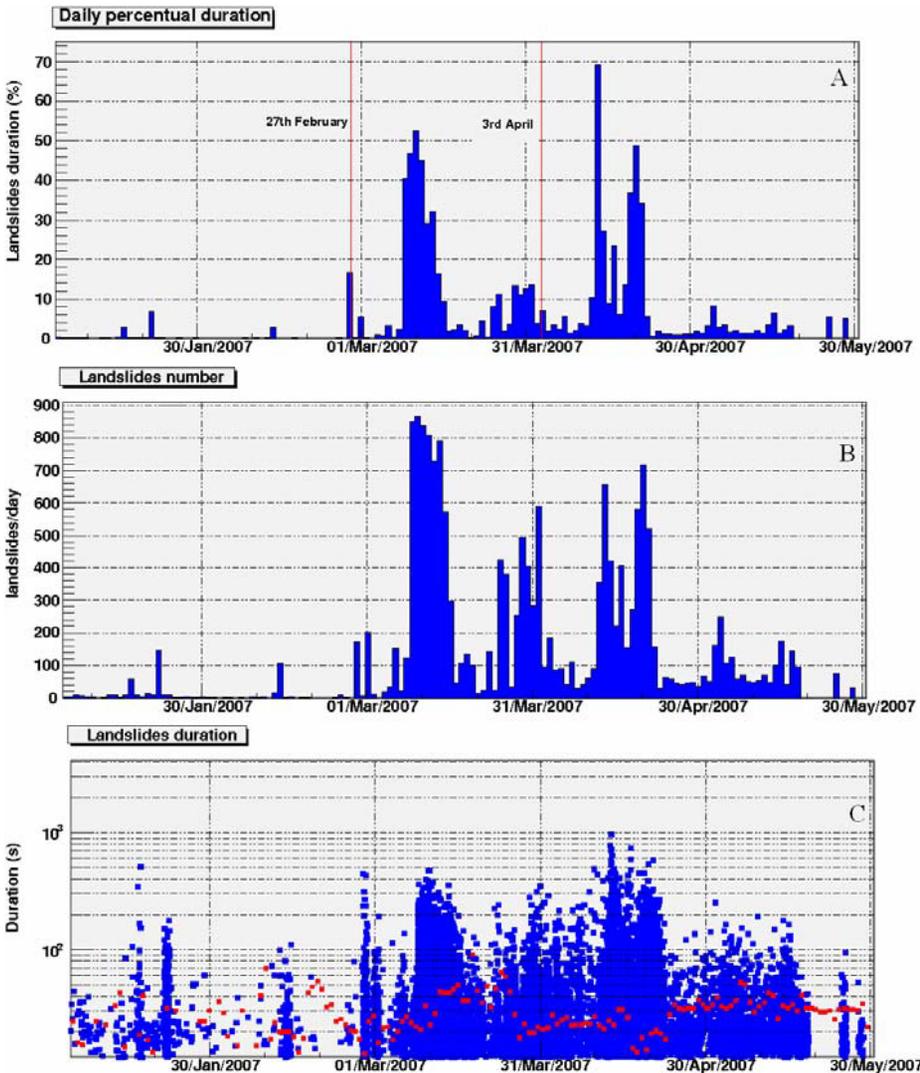


Figure 6. Landslide analysis over the period from 1/1/2007 to 1/6/2007. (A) Percentage duration of landslides over 24-h periods, with the beginning and the end of the lava flow indicated by the red lines. (B) Daily number of landslides. (C) Duration of each landslide detected (blue squares) and the daily average duration (red squares).

Figure 7A shows the seismogram of the February 27, and Figure 7B its associated percentage duration of landslide signals over a 10 min periods. Considering the occurrence of the first landslide and the beginning of the eruption (Figure 7A, first and second red arrows) and simultaneously the percentage duration of landslides (Figure 7B), it can be seen that there is an evident increase in the landslide signals rate from 09:00 GMT, about 4 h before the eruption started.

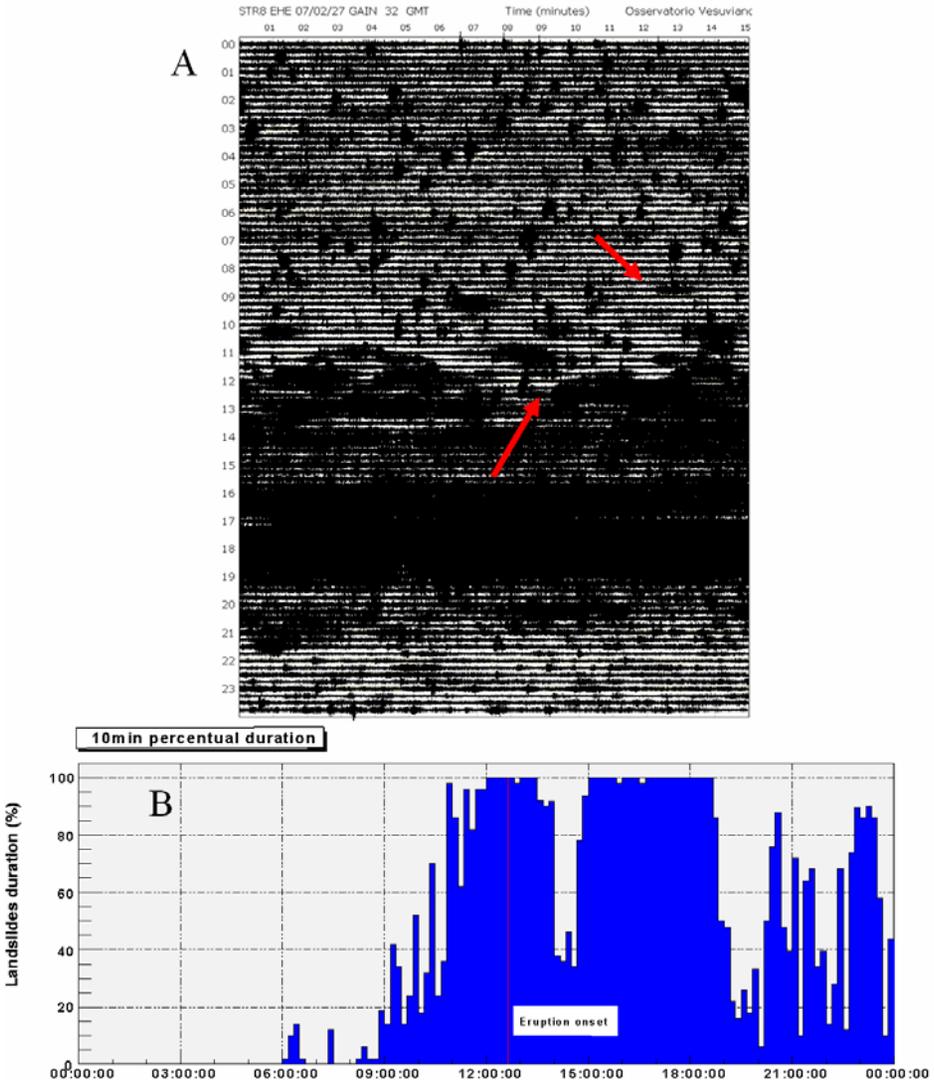


Figure 7. (A) Seismogram of February 27. (B) Percentage duration of the landslides over 10-min periods. The first red arrow in (A) indicates the occurrence of the first landslide, while the second one indicates the eruption onset.

6. Conclusions

In this study, we have presented a system for automatic landslide detection on Stromboli volcano. Usually this task is carried out visually by analysts, on the basis of their personal knowledge and using the spectral and temporal information of the signals analyzed. Automating this process would help analysts and allow objective event interpretation. Moreover, it will also be useful for monitoring of the “Sciara del Fuoco” flank.

Therefore, we have here realized an early landslide classification stage through a MLP-based algorithm. The proposed system then needs to apply a user-defined threshold to identify the start and end of each landslide signal, and to provide a graphical representation of the detection, as illustrated in Figure 5. Moreover, other procedures have been implemented that can perform several types of analysis on the system output, such as those illustrated in Figure 6. Finally, the results shown in Figure 7 suggest that the proposed system can be used as a short-term precursor detector.

However, this study is still in progress with the testing of the algorithm on different time intervals, so as to improve its robustness and to provide a system for landslide analysis in real time.

References

- [1] C. Bishop, *Neural Networks for pattern recognition*, Oxford University Press. 500 pp., 1995.
- [2] J.M. Cercone and J.R. Martin, An application of neural networks to seismic signal discrimination, Phillips Laboratory, report no. 3, PL-TR-94-2178, Hanscom, AFB, Massachusetts (1994).
- [3] B. Chouet, P. Dawson, T. Ohminato, M. Martini, G. Saccorotti, F. Giudicepietro, G. De Luca, G. Milana and R. Scarpa, Source mechanisms of explosions at Stromboli Volcano, Italy, determined from moment-tensor inversions of very-long-period data, *J. Geoph. Res.* **108** (B1) (2003).
- [4] W. De Cesare, M. Orazi, R. Peluso, G. Scarpato, A. Caputo, L. D’Auria, F. Giudicepietro, M. Martini, C. Buonocunto, M. Capello, A. M. Esposito, The broadband seismic network of Stromboli volcano, Italy, *Seismol. Res. Lett.* **80** (2009) 435-439; doi: 10.1785/gssrl.80.3.435.
- [5] E. Del Pezzo, A. Esposito, F. Giudicepietro, M. Marinaro, M. Martini and S. Scarpetta, Discrimination of earthquakes and underwater explosions using neural networks, *Bull. Seism. Soc. Am.* **93** (2003) 215-223.
- [6] F.U. Dowla, S.R. Taylor and R.W. Anderson, Seismic discrimination with artificial neural networks: preliminary results with regional spectral data, *Bull. Seism. Soc. Am.* **80** (1990), 1346-1373.
- [7] F.U. Dowla, Neural networks in seismic discrimination, in *Monitoring a Comprehensive Test Ban Treaty*, E.S. Husebye and A.M. Dainty (Editors), NATO ASI, Series E, 303, Kluwer, Dordrecht, The Netherlands, (1995), 777-789.
- [8] A. M. Esposito, F. Giudicepietro, S. Scarpetta, L. D’Auria, M. Marinaro, M. Martini, Automatic discrimination among landslide, explosion-quake and microtremor seismic signals at Stromboli volcano using neural networks, *Bull. Seismol. Soc. Am.* **96** (2006) 1230-1240; doi: 10.1785/0120050097.
- [9] J. Makhoul, Linear prediction: a tutorial review, *Proc. IEEE* **63** (1975) 561-580.
- [10] M. Martini, F. Giudicepietro, L. D’Auria, A.M. Esposito, T. Caputo, R. Curciotti, W. De Cesare, M. Orazi, G. Scarpato, A. Caputo, R. Peluso, P. Ricciolino, A. Linde, S. Sacks, Seismological monitoring of the Feb. 2007 effusive eruption of the Stromboli volcano, *Ann. Geophys.* **50** (2007) 775-788.
- [11] 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, *Bull. Seism. Soc. Am.* **95** (2005) 185-196.
- [12] T. Tiira, Detecting teleseismic events using artificial neural networks, *Comp. Geosci.* **25** (1999) 929-939.
- [13] J. Wang and T. Teng, Artificial neural network based seismic detector, *Bull. Seism. Soc. Am.* **85** (1995) 308-319.