Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Volcanology and Geothermal Research 153 (2006) 1–10

Journal of volcanology and geothermal research

www.elsevier.com/locate/jvolgeores

Automatic classification and a-posteriori analysis of seismic event identification at Soufrière Hills volcano, Montserrat

H. Langer [a], S. Falsaperla [a,*], T. Powell [b], G. Thompson [c,d]

[a] Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Catania, P.zza Roma 2, 95123 Catania, Italy
[b] Department of Earth Sciences, University of Leeds, Leeds, W Yorkshire, LS2 9JT, England, United Kingdom
[c] Montserrat Volcano Observatory, Montserrat, British West Indies, United Kingdom
[d] British Geological Survey, Keyworth, Notts, NG12 5GG, England, United Kingdom

Received 13 April 2004; received in revised form 21 September 2004; accepted 17 August 2005
Available online 18 January 2006

Abstract

Seismic energy radiation at Soufrière Hills volcano, Montserrat, is made up by various types of transient signals, which are distinguished by the Montserrat Volcano Observatory (MVO) in different classes with respect to their characteristics and/or origin. There are five fundamental classes, i.e., Volcano-Tectonic Events, Regional Events, Long-Period Events, Hybrid Events, and Rockfalls. Identification and classification of these transients, which have been hitherto carried out manually by various staff members, yield important information for the assessment of the state of the volcano system. In the frame of the MULTIMO project, we proposed the application of Artificial Neural Networks (ANN) for the classification of these kind of data in order to handle large data sets, and to achieve reproducible results, emulating the expert's analysis. Using the manual routine classification as a-priori information, we obtained a fair performance of such an automatic processing, with 70% of the automatic classifications being consistent with the original ones. From an analysis of the misclassified events, however, we found that for most of them the original a-priori classification was incorrect.

In this study, we first revised manually the original a-priori classification. Based on a data set of 6000 events, we carried out a reanalysis of the seismic traces recorded at different seismic stations. Then, using this new information, we trained and tested the ANN, obtaining a successful classification in ca. 80% of records. Particularly, the automatic classification was excellent in the identification of Rockfalls and Volcano-Tectonic Events. Among the misfits, we observe the erroneous attribution of Long-Period and Hybrid Events to Rockfalls. This may be partly explained by the fact that signals addressed to as Rockfalls contain frequently contributions of various sources. Overall, the failure in the classification between some types of transients highlights the problem of the concurrent activation and/or unclear separation of distinctive sources from which the signals stem. We conclude that the automatic classification with ANN is a powerful tool for handling large data masses as well as for the a-posterior analysis of the consistency of the classification problem.

Keywords: automatic classification; Montserrat; neural network; volcano seismicity

1. Introduction

At the beginning of the 20th century, volcanism in the Caribbean was famous for two spectacular and devastating eruptions at the Soufrière of St. Vincent

* Corresponding author. Tel.: +39 95 7165845; fax: +39 95 435801.
E-mail address: falsaperla@ct.ingv.it (S. Falsaperla).

0377-0273/$ - see front matter
doi:10.1016/j.jvolgeores.2005.08.012

and Mt. Pelée, Martinique, where in particular the phenomenon of *nuées ardentes* was reported for the first time. In 1995, the focus of scientific interest turned again to the Caribbean, where after 400 years of relative quiescence the Soufrière Hills volcano (Fig. 1) on Montserrat resumed its activity. Since then Montserrat has provided a large volume of data, boosting the study of volcanic processes in the framework of research projects such as MULTIMO, which has been financed by the EU.

Montserrat (Fig. 2) is located in the northern part of the Lesser Antilles. The island is made up by four volcanic centres whose age ranges from 2.5 Ma to the present. Among these the Soufrière Hills volcano is the youngest, its activity being started at least 170 ka ago (Sparks and Young, 2002). The volcano complex has been monitored regularly since 1992, when enhanced seismic activity was locally recorded. Immediately after the phreatic explosions on July 18, 1995, an operational centre was established in the island. This centre acted as a seed for the Montserrat Volcano Observatory (MVO), which has been provided with a sizable scientific staff to conduct round the clock monitoring of volcano activity (Aspinall et al., 2002). The monitoring at Soufrière Hills volcano encompasses visual observations, ground deformations, seismic radiation, and geochemical analyses. Continuous seismic monitoring has proven to be the key technique. On active volcanoes, a wide variety and amount of seismic signals may occur. The problem of handling the huge data masses accumulating during continuous monitoring was recently addressed in a paper by Langer and Falsaperla (2003), where specific aspects concerning both persistent signals – the so-called volcanic tremor – and transients were discussed. The authors highlighted the need and benefits of data reduction and parameter extraction. On volcanoes like Soufrière Hills with its andesitic and SiO_2-rich magma, seismic radiation is characterized by transient signals. Although tremor episodes do occur, they are often due to the superposition of repeated Hybrid or Long-Period Events (Baptie et al., 2002) and are not continuous background tremor, such as that observed on open-conduit volcanoes like Mt. Etna or Stromboli (Falsaperla et al., 1998; Schick and Riuscetti, 1973). There are typical signal classes with distinctive waveform, frequency content, and duration. Therefore, the goal of data compression can be achieved by event classification. In a general scheme, transient seismic signals on volcanoes are categorised as (see McNutt, 2000): (i) High Frequency Events, with clear P- and S-phases and dominant frequencies greater than 5 Hz, (ii) Long-Period Events, with no clear onset and dominant frequencies between 1 and 5 Hz, (iii) Hybrid Events, with sharp onset and coda similar to Long-Period Events, (iv) Explosion Quakes, accompanying eruptive activity, and often accompanied by air waves, (v) Very Long-Period Events, with dominant periods between 3 and 20 s, and (vi) Superficial Events, such as landslides and rockfalls. The MVO staff modified this scheme in order to match the specific situation on Montserrat. The relationships between the various event types and volcanic activity have been widely described and investigated in several volcanic areas (e.g., McNutt, 2000). In the following, we refer to the

Fig. 1. Snapshot of the Soufrière Hills volcano (taken from east) during a pyroclastic flow.

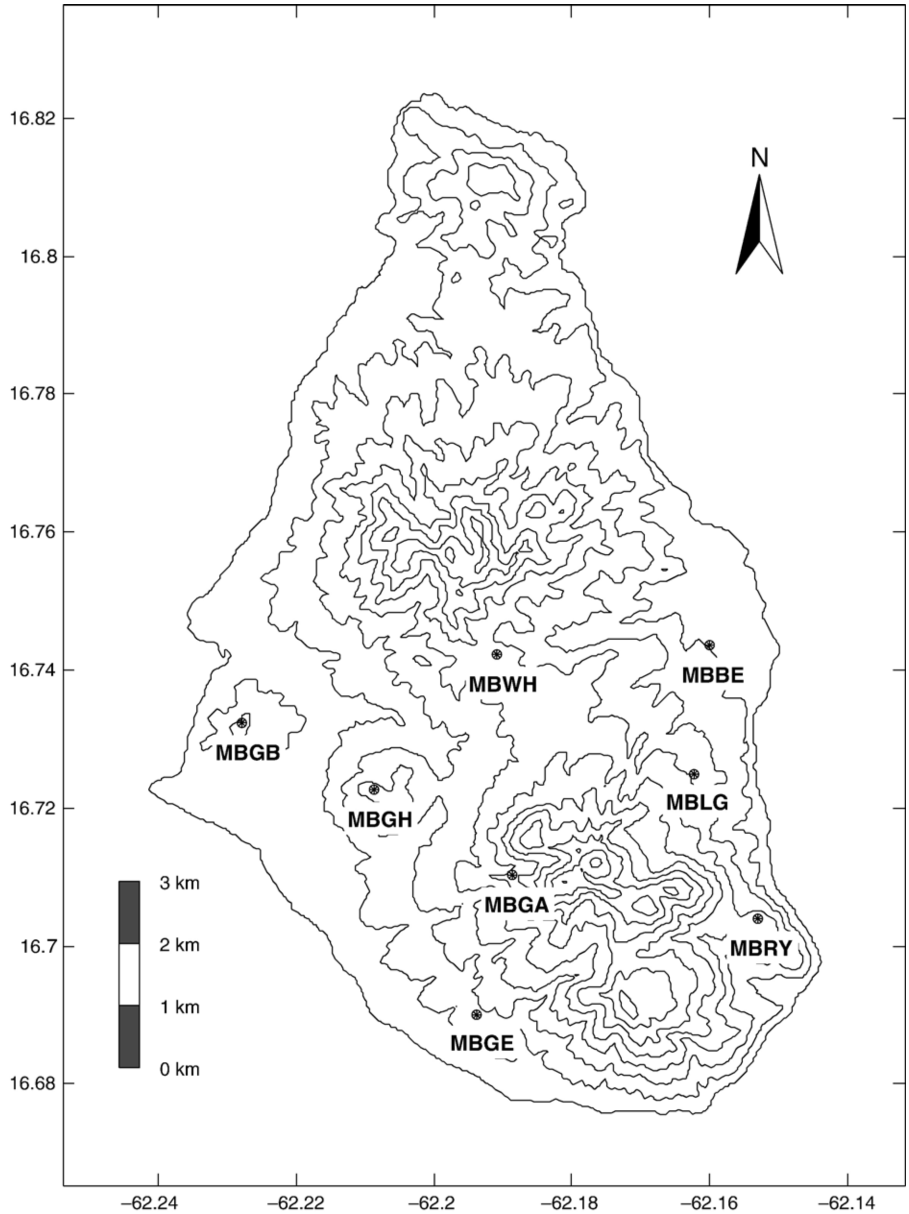

Fig. 2. Map of the island of Montserrat. Stars mark location of the stations of the digital seismic network. Some stations have been destroyed by volcanic activity, and relocated.

scheme adopted by the MVO staff, and identify five event classes (Fig. 3), i.e., Volcano-Tectonic Events (hereafter VTE), Regional Events (REG), Long-Period Events (LPE), Hybrid Events (HYB), and Rockfalls (ROC). Besides these five types, a further class was originally distinguished as Long-Period Events + Rockfalls (LPE + ROC). In this class, the MVO staff identified a mixed transient consisting of a rockfall shortly preceded by a low frequency signal.

Volcano-Tectonic Events (VTE) have sharp onsets, dominant high frequencies, and short durations. They are commonly due to brittle fracture in response to stress changes associated with magma dynamics. According to the MVO's classification scheme, VTE originate from within the volcanic edifice, and often occur before impending eruptions and in post-eruptive times during stages of magma withdrawal and stress relaxation.

Regional Events (REG) are tectonic earthquakes caused by brittle fracture. Unlike VTE, however, they have longer delays between P- and S-wave arrival times, and originate outside of the volcanic edifice.

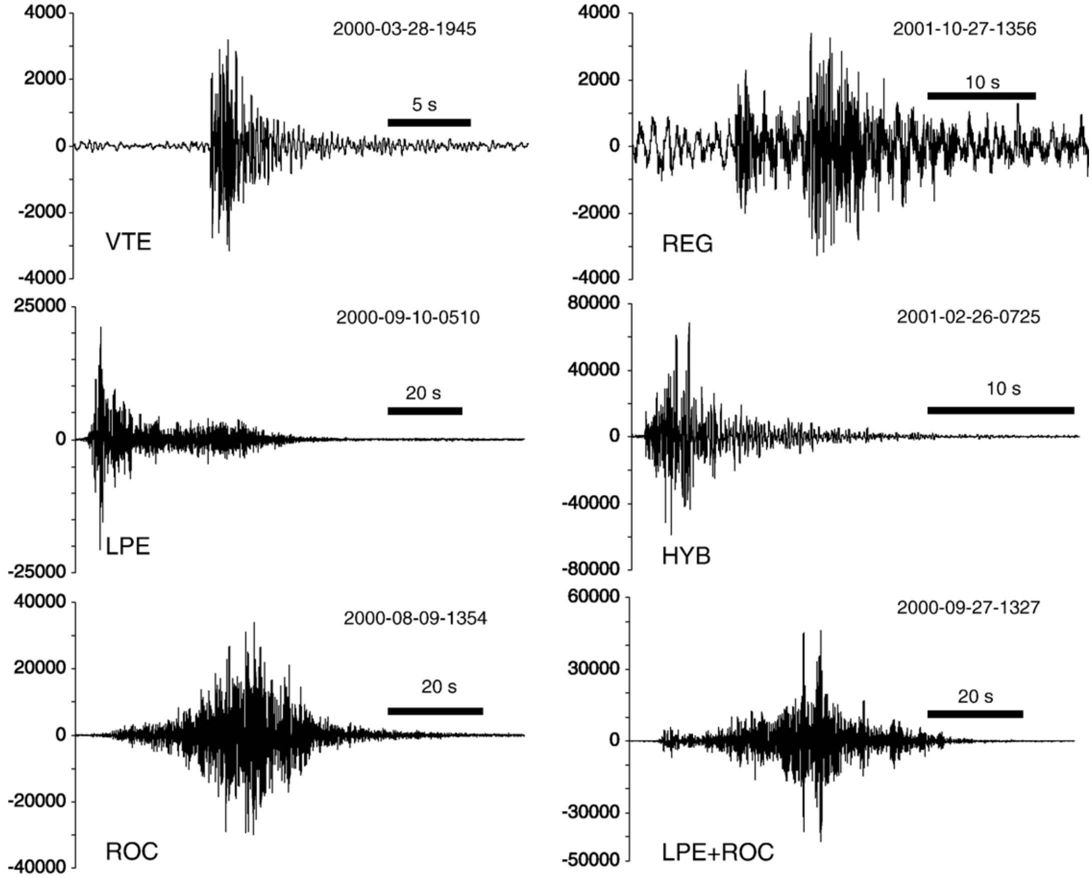

Fig. 3. Waveforms of the six classes of transients considered. Year, month, day, hour, and minute at the onset of each transient are indicated at the upper right of each trace.

Long-Period Events (LPE) at Soufrière Hills volcano have distinct spectral peaks in the frequency range between 0.2 and 5 Hz. They are interpreted as the result of a resonance phenomenon of a magma-filled conduit.

Unlike LPE, Hybrid Events (HYB) have an additional high frequency phase up to 10 Hz, which precedes the long-period phase. HYB are often recorded in swarms and sometimes merge to tremor accompanying explosive activity. Cyclic behaviour of the swarms has been linked to cyclic tilt behaviour, suggesting that the two phenomena are interconnected (Baptie et al., 2002). Neuberg et al. (2000) consider LPE and HYB as two end-members of a single event class, and claim that a continuum exists between them.

On Montserrat specific attention has been devoted to Rockfalls (ROC), as these are related to the growth and collapse of the lava dome (Calder et al., 2002). Rockfalls could be as small as a single block bouncing down the slope of the dome. In extreme cases, however, they may assume the dimension of pyroclastic flows, which run over distances of several kilometres. Whenever a rockfall is observed, a characteristic seismic signal is recorded at the nearby seismometers. The seismic record, on the other hand, allows the detection of ongoing rockfall activity when its visibility is concealed from meteorological conditions. Although the seismic signals associated with ROC are usually characterized by dominant frequencies between 2 Hz and 8 Hz, in many cases significant signal energy is present in the frequency range below 2 Hz as well (Luckett et al., 2002).

Commonly, hundreds of events are recorded each week by MVO. A visual inspection by human operators of all these events, forms a rather arduous task in terms of time – especially during phases of enhanced activity – and whose results may be severely affected by the subjectivity and motivation of the data analyst. We therefore investigated the application of an automatic classification tool, known as Artificial Neural Networks (ANN). We use the ANN in a classification scheme with a supervisor. In this scheme, the discrimination function among the various classes is estimated using the a-priori classification provided by an expert for a specific training set of data, and then it is applied to new, unknown data. The advantage of the ANN over classical classification techniques has been discussed in various studies (e.g., Lippmann, 1987; Raiche, 1991). A principal feature of ANN is that it is possible to solve classification problems of arbitrary complexity. Once the discrimination function is established, the classification becomes extremely fast and suitable for both real time applications as well as the processing of very large data sets piled up in the past.

Supervised learning essentially depends on the quality of the training data set, which has to be of sufficient size in order to represent the parent population of the events recorded on the volcano. There is no general rule to fix a proper size for the training set. As ANN are good at generalization, even a small data set can be sufficient to train a network successfully. A further request is the consistency of the a-priori information (here the a-priori classification carried out by the human operator for the training data set). The request of consistency of the a-priori information, however, is not easy to match. The staff in a remote observatory like MVO is subject to change (see Aspinall et al., 2002 for a thorough discussion of the development of MVO), and the reproducibility of expert opinions is widely questioned. Moreover, the staff who routinely categorise events are not trained seismologists. Thus, the scope of the present study is twofold. One aspect deals with the technical characteristics of the ANN application, the other aspect focuses on the re-analysis of the data used in the training set.

From a first application to a limited data set with 336 records, Langer et al. (2003a) concluded that automatic classification of seismic transients on Montserrat should be successful if a training set of sufficient size were available. In this paper, we analyze a considerably larger data set, consisting of several thousands of traces. These signals had been routinely classified by various members of the MVO staff. The original classification was used in a first step as-priori information for the application of the ANN.

2. Data set

The data set we analyze here covers a time span from 1996 to 1999. The seismic signals were recorded at 9 digital stations belonging to the permanent seismic network run by MVO (Fig. 2). The stations were equipped with Guralp CMG-40T 3C broadband sensors, having a bandwidth between 0.03 and 30 Hz, and dynamic range of 144 dB. Additionally, Integra LA-100, vertical component, 1 Hz seismometers were also used. The sampling rate was 75 Hz. Signals from each station were transmitted to MVO, where they were recorded on a PC-based digital acquisition system.

For our application, we used up to 6000 seismic records, considering the vertical component only. These records were originally divided into the six classes aforementioned by MVO staff. The time series have durations from ca. 1 min to several minutes. Following Langer et al. (2003a), we chose to use the time series recorded at the different stations for each single transient as separate examples, i.e., the classification is carried out considering each single channel. By classifying each trace on its own, we avoid complex reasoning about how to weigh the significance of the various stations. In fact, these weighting schemes might become quickly obsolete, as a seismic network on active volcanoes often undergoes technical changes, and especially during climactic stages, stations considered very significant may be missing from one day to the other.

3. Application of the ANN

The architecture of ANN we use for our application is the so-called Multi-layer Perceptron (Rumelhart et al., 1986, Fig. 4). This ANN has the advantage to define discrimination functions of arbitrary complexity in principle, and the procedures followed for the estimation of the discrimination function do not require any a-priori knowledge about its mathematical structure.

For the set-up of the ANN, we carry out the three following steps: (i) definition of a suitable form of data representation, (ii) identification of representative examples of the patterns (here the signals or their transforms, respectively) whose class membership is known, (iii) choice of an ANN topology. We chose a simple topology consisting of three layers of nodes: an input layer, where the input data vectors (patterns) are stored, a hidden layer, where a nonlinear weighting function is applied and which is necessary to guarantee the generality of the discrimination function, and the output layer representing the resulting vector of the ANN application. Particularly, the hidden layer is formed by the computing units of the network which work as feature detectors. The terms node and neuron are synonyms, and indicate the individual computation elements of the ANN. We consider topologies in the form U–NH–O, where U and NH are the number of neurons in the input and hidden layers, respectively, and O are the classes of patterns to identify. The output vectors concerning each single pattern contain

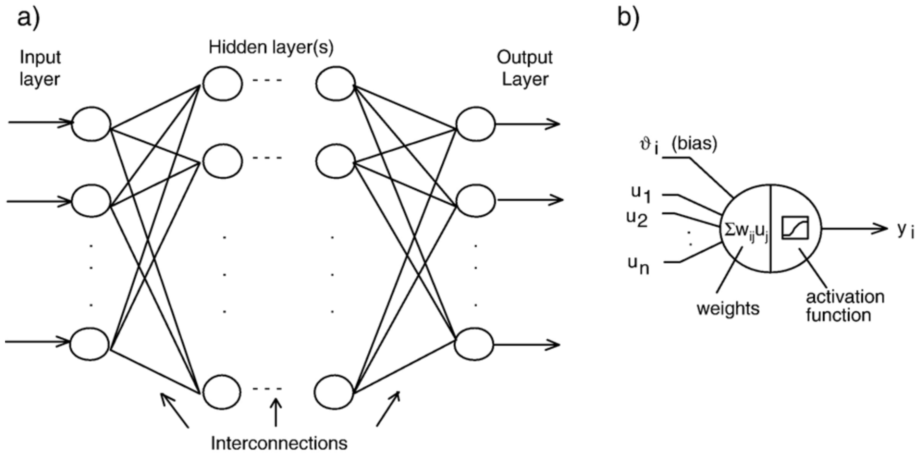

Fig. 4. (a) Topology of an ANN, and (b) scheme of the single neuron (modified from Langer and Falsaperla, 2003).

the class memberships of the signals calculated with the ANN, which are compared to the target values defined by the human operator. The target output has assigned "1" to the class which the signal belongs to, and "0" to the other classes.

Formally the classification is carried out by applying a mapping function of the input vector **U** (which represents our signals) to an output vector **Y** (here the class membership values). The mapping function is given by

$$\hat{y}_k(\mathbf{U}) = \sum_{j=1}^{\mathrm{NH}} c_j \sigma\left(\mathbf{w}_j^T \cdot \mathbf{U} + t_j\right) + c_0$$

where $\hat{y}_k$ is the k-th element of **Y** estimated by the network, **U** is the input vector, $\mathbf{w}_j$ are the vectors of the weights between input and hidden layer, c_j are the weights between hidden and output layer, t_j are biases, $\sigma(\cdot)$ is the sigmoid activation function $\sigma(z)=1/(1+\mathrm{e}^{-z})$, c_0 is a constant. In simple words, the quantities $\mathbf{w}_j$ and c_j are weights which express the strength of the connection between two individual nodes. The sigmoid activation function regulates the activation of the nodes according to thresholds fixed by its typical shape, and consequently determines the output value. An in-depth discussion of this function can be found in Grossberg (1982).

A classification is considered successful if the maximum of the $\hat{y}_k$ is obtained for the same class k for which the target output was set to "1". The ANN performance improves when the length of the input data vectors can be limited and phase alignment problems are avoided (Falsaperla et al., 1996; Langer and Falsaperla, 2003). Following Langer et al. (2003a), we find convenient to use an information code which provides a constant length of all the input data vectors, regardless of the original signal duration. Particularly, for the preparation of this information code, we use a combination of: (i) the autocorrelation function, (ii) statistical parameters such as the sums of the amplitude signal A, (iii) the amplitude ratio between filtered and unfiltered traces, (iv) the ratio of the maximum amplitude of the signal versus the RMS amplitude measured before and after the maximum. The autocorrelation functions (acf) represent the spectral content of the signal. The acf were obtained in the frequency domain using a window length of 16,384 points for the Fast Fourier Transform. In each input data vector, we first provide the acf taking the first n points that we find sufficient to represent the function. Being zero-phase functions, acf always have their maximum at the beginning of the trace. This avoids the problems of phase alignment which may occur if plain waveforms are used (Langer and Falsaperla, 2003). Then, we consider the amplitude values of the signal A, followed by the sums over A^2, A^3, A^4. These sums of the amplitude resemble statistical moments. Finally, we introduce the amplitude ratio between the bandpass filtered (between 0.7 and 1.5 Hz) and unfiltered traces, as well as the ratio between the maximum amplitude of the signal and the RMS amplitude measured in two time windows over 5 s before and 30 s after the maximum, respectively. The statistical moments and amplitude ratios should help to distinguish brief, peaked transients (like VTE) from signals with long duration (like ROC), but having similar frequency content.

Following the common practise in ANN applications, we divide our data set into two groups, by randomly selecting a training set, which was used for the estimation of the ANN coefficients, and a test set. The application of the ANN to a test set not being used during the training phase provides information about how the ANN performs when applied to new, unknown data. The random selection makes reasonable the assumption that training and test data sets belong to the same parent population, as requested by theory.

4. Previous tests with original a-priori classification

The years from 1996 to 1999 encompass a time span in which volcanic activity on Montserrat reached a climax with devastating eruptions in 1997. For our analyses, we selected about 6000 records corresponding to ca. 1000 events, which were representative of the activity in the years 1996–1999. We continued to use the a-priori classification assigned by the MVO staff during routine analyses, with the six classes of signals: VTE, REG, LPE, HYB, ROC, LPE+ROC. Table 1 highlights the distribution for event classes in the data set as inferred from the a-priori classification. There is a high number of ROC and VTE, whereas the number of events identified a-priori, in particular, as REG and LPE+ROC is small in comparison to the total number. During the pre-processing, we removed from our data set all the events with peak-to-peak amplitude less than 1000 counts, in order to limit the effects of noise. Next, 3500 events were randomly selected, which made up the training data set for the estimation of the ANN mapping function. The remaining 900 events were used as test set for the assessment of the performance of the ANN when applied to new data. After some trial-and-error experiments, we chose the topology 103–20–6, i.e., an input layer consisting of 96 neurons for the autocorrelation function plus 7 neurons for the other parameters, a hidden layer of 20 neurons, and an output layer with 6 neurons corresponding to the number of classes to recognize. The results obtained with this topology are provided in Table 2, and highlight a success rate of about 70% both in the test and training data set. This percentage is of the same order of the success rate reported by Langer et al. (2003a). These authors found that for their 336 signals the results of the classification for the training set improved augmenting the complexity of the ANN topology, whilst concurrently the mismatch for the test set increased. They interpreted this effect as a consequence of overfitting, and concluded that the mismatch rate of 30% for their data set was partly due to an insufficient number of records available for the training of the ANN, which made useless the application of complex topologies.

Conversely, the tests we carried out did not show overfitting problems even with topologies as large as 300–50–6 (293+7 neurons between acf and the other selected parameters, 50 neurons for the hidden layer, 6 neurons for output). In a further test, we also excluded events with peak-to-peak amplitude less than 5000 counts. The remaining data set consisted of ca. 2200 events, 1400 of which were used in the training set and about 800 in the test set. Nonetheless, no significant change in the mismatch rate was noticed. Overall, the tests we carried out led us to surmise that the success rate of the ANN applications was bounded for intrinsic reasons. These reasons might be the somehow unsuitable representation of the input data (here autocorrelation functions, statistical parameters, and amplitude ratios, as described above). Another possible reason might be related with the a-priori information, and the possibility that it had flaws. Visual inspections of some misclassified events – for which indeed the a-priori classification turned out questionable (see Fig. 5) – led us to carry out a careful re-analysis of the a-priori classification.

Table 1
Composition of the data set with peak-to-peak amplitude greater than 1000 counts

VTE	REG	LPE	HYB	ROC	LPE+ROC
895	226	504	589	2094	83

Table 2
Results of the automatic classification using original a-priori information

T/C[a]	VTE	REG	LPE	HYB	ROC	LPE+ROC
Training set						
VTE	608	51	11	43	63	0
REG	1	12	2	2	1	1
LPE	5	5	220	50	25	4
HYB	21	10	30	195	26	1
ROC	79	103	141	180	1548	62
LPE+ROC	0	0	0	0	0	0
Test set[a]						
VTE	148	17	1	10	22	0
REG	0	0	0	0	0	1
LPE	2	2	49	13	5	1
HYB	3	4	10	48	3	0
ROC	28	31	40	48	401	13
LPE+ROC	0	0	0	0	0	0

[a] T/C=Target (column) versus calculated (line).

5. Reclassification with revised a-priori information

A first, partial revision was carried out by one of the co-authors of the present paper, Tanya Powell, who was

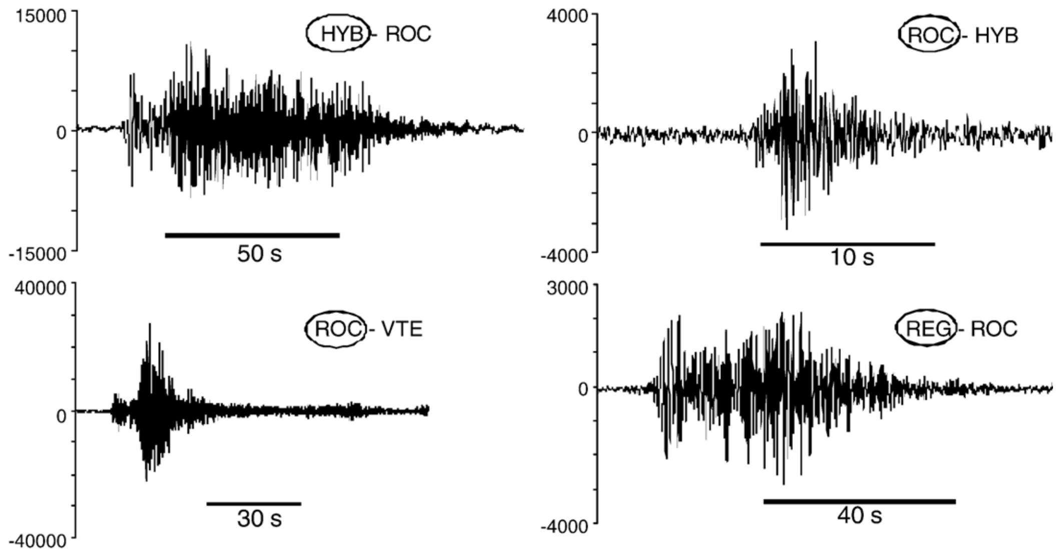

Fig. 5. Examples of doubtful or erroneous a-priori classifications. For each transient, the class encircled at the upper right indicates the a-priori classification, the following one is the class assigned by the ANN.

not involved in the previous analysis, and whose judgment could not be biased by the research hitherto developed. About 300 of the misclassified traces were re-classified and compared to both the original classification and the automatic classification provided by the ANN. The results shown in Table 3 clearly demonstrated that the flaws of the a-priori classification could be a serious problem. Overall, in more than 50% of the revised events, the classification of the ANN was preferred to the original one previously assigned by the human operators.

Moving on from these results, we decided to reanalyze ca. 2400 events with peak-to-peak amplitude greater than 5000 counts, and to reclassify them by our own, before proceeding with a new application of the ANN. About 200 events turned out to be not classifiable, for the signal quality was too poor. Among the remaining 2200 patterns, 1400 transients were randomly selected as training set patterns, whereas 800 traces were used to form the test set. Similar to our former applications of the ANN, we used the autocorrelation function, the sums over A, A^2, A^3, and A^4, and the amplitude ratios previously described, as input parameters. We assumed, however, a slightly simplified scheme with five classes, including the transients of the class LPE+ROC within that of the ROC. The automatic classification with ANN matched the target classification in about 80% of cases (81% in the training set and 78% in the test set, see Table 4). Most of the ROC were classified consistently, and the same holds for the VTE. Conversely, many HYB and LPE were confused with ROC, whereas a minor number of misclassifications was noticed among the two classes themselves. On the other hand, the identification of the REG mostly failed.

Table 3
Manual reclassification of 300 apparently misclassified events

T/C[a]	VTE	REG	LPE	HYB	ROC	LPE+ROC
VTE	8	15	0	0	9	0
REG	0	3	0	0	2	0
LPE	0	0	15	10	0	3
HYB	9	4	2	50	14	5
ROC	4	5	12	17	102	35
LPE+ROC	0	0	0	0	0	0

[a] T/C=New Target (column) versus calculated (line).

Table 4
Results of the automatic classification using revised a-priori information

T/C[a]	VTE	REG	LPE	HYB	ROC
Training set					
VTE	182	10	0	3	19
REG	3	21	0	0	5
LPE	2	1	113	23	15
HYB	12	5	17	76	5
ROC	26	26	57	31	748
Test set					
VTE	102	6	2	7	9
REG	5	3	0	1	2
LPE	2	1	47	16	8
HYB	3	4	10	31	4
ROC	14	22	40	30	458

[a] T/C=New Target (column) versus calculated (line).

6. Discussion and conclusions

Seismic monitoring has become a key tool for the surveillance of active volcanoes worldwide. The continuous data acquisition, however, brings along the problem of accumulating huge data masses – such as 100 MB per day per each single three-component station – which are cumbersome to handle. Data compression and parameter extraction are therefore necessary to exploit the huge amount of information for forecast and warning purposes. The method used to achieve this goal depends on the type of volcanic activity and related seismic signals. The Soufrière Hills volcano on Montserrat forms an andesitic complex with highly viscous magma (Sparks and Young, 2002). Seismic radiation on volcanoes of this type is characterized by transient signals, which prevail throughout long time spans over tremor. In this context, we tackle the task of data compression and parameter extraction performing the automatic classification of signals with ANN. Our applications follow a supervised classification scheme, in the sense that the target classification of the training set is defined by an expert. The advantage of the ANN compared to more traditional techniques is their generality. In other words, it is possible, in principle, to attain classification functions of arbitrary complexity, provided there is a sufficient number of available examples to achieve a stable estimation of the coefficients of the classification function. Once this function is obtained, its application is simple and fast, allowing the processing of large data sets off-line as well as the automatic classification on-line of continuously recorded signals.

An important aspect of automatic classification by ANN is that it is undoubtedly objective, for it is formally reproducible. In addition, its results can be exploited in a-posteriori analyses to verify the consistency of the a-priori classification. In the first tests carried out by Langer et al. (2003a), the performance of the ANN was good enough to be considered encouraging. The percentage of misfits, which reached 30%, compelled the authors to carry out further analyses. From their experiments with various topologies, the authors suspected that their data set containing 336 traces was probably insufficient for getting better results, as overfitting problems were encountered for large topologies. Moving on from these findings, we considered a large data set with up to 6000 traces, corresponding to ca. 1000 events. We carried out various experiments to choose the best topology and define a convenient amplitude threshold which excluded noisy signals with small amplitude. Although overfitting problems were definitely fixed, the output of the automatic classification in the following tests diverged from the desired output in about 30% of the records, both in the training set and test set (Langer et al., 2003b).

It has been observed (see, e.g., Langer et al., 1996) that persisting errors in both training and test set might be an indication for an ill-conditioned problem rather than an ANN failure. In our specific case, we suspected that the a-priori classification had flaws. This could be indeed verified by one of the co-authors of the present paper, who had not taken part in the analysis until then. A re-evaluation of 300 apparently misclassified records revealed that at least 50% of the mismatches were due to debatable or even erroneous a-priori classification. As a consequence, we redefined the target classifications for all the records with peak-to-peak amplitudes greater than 5000 counts, visualizing them one by one and assigning a class membership to each of them. With this new definition, we achieve a success rate of ca. 78% for the ca. 800 examples of the test set, whereas the score in the training set amount to 81%. ROC and VTE are recognized with a high success rate (close to 90%). We also note a proper separation between HYB and LPE, which comes to some degree as a surprise, as these two types might form two end-members of a single event class and a continuum has been claimed between them (Neuberg et al., 2000).

We believe that the remaining 20% of misclassified traces places an intrinsic limit of our classification problem. Firstly, similar to the original a-priori classification, our revised classification may also contain inconsistencies and errors. We also note an overall failure in the identification of the REG. This failure is probably due to the scarce (about 5%) number of records in the data set. Nevertheless, we cannot exclude that the coding of the input information (autocorrelation function, statistical moments, and amplitude ratios) is not appropriate for this event class. On the other hand, regional events may be easily recognized by simply carrying out an automatic location.

Apart from that, LPE and HYB are frequently confused with ROC. Besides the reasons earlier discussed, this failure may be explained as the result of the fact that ROC form a rather heterogeneous class. Luckett et al. (2002) identified three subclasses of rockfalls based on waveforms and frequency content: (i) rockfall signals with dominant frequencies above 2 Hz, (ii) rockfall signals with significant contributions of energy radiation between 1 and 2 Hz, and (iii) rockfall signals where the energy radiation at 1–2 Hz starts before the rest of the signal (the class Long-Period Events + Rockfall used in Langer et al., 2003a,b and then dropped here). Following Luckett et al. (2002), there is no simple relation between the frequency content of rockfall signals and the speed of dome growth, but there seems to be a link between dome growth and the percentage of energy of the rockfall signals below 2 Hz. Luckett et al. (2002) speculate that two separate sources contributing to the rockfall signals may exist. The high-frequency part above 2 Hz then may represent the radiation due to the material tumbling downhill, whereas the low-frequency signal would be generated by resonance phenomena similar to Long-Period or Hybrid events. In this context, one must expect that the results of the automatic classification are affected by these intrinsic difficulties for separating signal sources. In other words, automatic classification is valuable tool not only for its success in the identification of the majority of the transients, but also in the context of the a-posteriori analysis of failures, shedding light on the intrinsic difficulties in separating sources. We therefore recommend that automatic classification is applied to MVO's entire data set of transient events, as this will provide a much clearer picture of the seismicity associated with the eruptive activity of the Soufrière Hills Volcano, and improve the quality of the scientific advice MVO is able to give to the authorities on Montserrat.

Acknowledgments

We thank Cataldo Godano and Laura Sandri for their useful comments and suggestions. We also are grateful to Roberto Carniel, as guest editor of this issue. Finally, we gratefully acknowledge the contribution of MVO staff. Published by permission of the Director, British Geological Survey (NERC). This paper was financially supported by the EC project MULTIMO (Contract No. EVG1-CT-2000-00021).

References

Aspinall, W.P., Loughlin, S.C., Michael, F.V., Miller, A.D., Norton, G.E., Rowley, K.C., Sparks, R.S.J., Young, S.R., 2002. The Montserrat volcano observatory: its evolution, organization, role and activities. In: Druitt, T.H., Kookelaar, B.P. (Eds.), The Eruption of Soufrière Hills Volcano, Montserrat, from 1995–1999, Geological Society, vol. 21. Memoirs, London, pp. 71–91.

Baptie, B., Luckett, R., Neuberg, J., 2002. Observation of low-frequency earthquakes and tremor at Soufrière Hills Volcano, Montserrat. In: Druitt, T.H., Kookelaar, B.P. (Eds.), The Eruption of Soufrière Hills Volcano, Montserrat, from 1995–1999, Geological Society, vol. 21. Memoirs, London, pp. 611–620.

Calder, E.S., Luckett, R., Sparks, R.S.J., Voight, B., 2002. Mechanisms of lava dome instability and generation of rockfalls and pyroclastic flows at Soufrière Hills volcano, Montserrat. In: Druitt, T.H., Kokelaar, B.P. (Eds.), The Eruption of Soufrière Hills Volcano, Montserrat from 1995 to 1999, Geological Society, vol. 21. Memoirs, London, pp. 173–190.

Falsaperla, S., Graziani, S., Nunnari, G., Spampinato, S., 1996. Automatic classification of volcanic earthquakes by using multi-layered neural networks. Nat. Hazards 13, 205–228.

Falsaperla, S., Langer, H., Spampinato, S., 1998. Statistical analysis and characteristics of volcanic tremor at Stromboli volcano (Italy). Bull. Volcanol. 60, 75–88.

Grossberg, S. (Ed.), 1982. Studies of Mind and Brain, Boston Studies in the Philosophy of Science, vol. 70. D. Reidel Publishing Compagny, Boston.

Langer, H., Falsaperla, S., 2003. Seismic monitoring at Stromboli volcano (Italy): a case study for data reduction and parameter extraction. J. Volcanol. Geotherm. Res. 128, 233–245.

Langer, H., Nunnari, G., Occhipinti, L., 1996. Estimation of seismic waveform governing parameters with neural networks. J. Geophys. Res., 20109–20118.

Langer, H., Falsaperla, S., Thompson, G., 2003a. Application of artificial neural networks for the classification of the seismic transients at Soufrière Hills volcano, Montserrat. Geophys. Res. Lett. 30 (21), doi:10.1029/2003GL018082.

Langer, H., Falsaperla, S., Thompson, G., 2003b. New tests of automatic classification of seismic transients recorded at Montserrat using ANN. Oral Presentation at the 3rd Annual Meeting of the MULTIMO Project, Sao Miguel, Azores, May 21–24, 2003.

Lippmann, R.P., 1987. An introduction to computing with neural nets. IEEE ASSP Mag., 4–22.

Luckett, R., Baptie, B., Neuberg, J., 2002. The relationship between degassing and rockfall signals at Soufrière Hills volcano, Montserrat. In: Druitt, T.H., Kokelaar, B.P. (Eds.), The Eruption of Soufrière Hills Volcano, Montserrat from 1995 to 1999, Geological Society, vol. 21. Memoirs, London, pp. 595–602.

McNutt, S.R., 2000. Volcanic seismicity. In: Sigurdsson, H., Houghton, B., McNutt, S.R., Rymer, H., Stix, J. (Eds.), Encyclopedia of Volcanoes. Academic Press, pp. 1095–1119.

Neuberg, J., Luckett, R., Baptie, B., Olsen, K., 2000. Models of tremor and low-frequency earthquake swarms on Montserrat. J. Volcanol. Geotherm. Res. 101, 83–104.

Raiche, A., 1991. A pattern recognition approach to geophysical inversion using neural nets. Geophys. J. Int. 105, 629–648.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representation by error propagation. In: Rumelhart, D.E., Mc Clelland, J.L. (Eds.), Parallel Distributed Processing: Explorations in the Microstructures of Cognitions. MIT Press, Cambridge, Mass, pp. 318–362.

Schick, R., Riuscetti, M., 1973. An analysis of volcanic tremors at south Italian volcanoes. Zeits. Geophys. 39, 247–262.

Sparks, R.S.J., Young, S.R., 2002. The eruption of Soufrière Hills volcano, Montserrat (1995–1999): overview of scientific results. In: Druitt, T.H., Kokelaar, B.P. (Eds.), The Eruption of Soufrière Hills Volcano, Montserrat from 1995 to 1999, Geological Society, vol. 21. Memoirs, London, pp. 45–69.