# Discrimination of quarry blasts from tectonic microearthquakes in the Hyblean Plateau (Southeastern Sicily)

**Andrea Ursino** ([1]), **Horst Langer** ([1]), **Luciano Scarfì** ([1]), **Giuseppe Di Grazia** ([1]) **and Stefano Gresta** ([2])
([1]) *Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Catania, Italy*
([2]) *Dipartimento di Scienze Geologiche, Università di Catania, Italy*

**Abstract**
The seismic network set up in the Hyblean Plateau (Southeastern Sicily) in the framework of the POSEIDON project is aimed at the seismic surveillance of the zone, and in particular the identification of faults with enhanced activity. The seismic activity as inferred from the records of the years 1994-1998 showed an apparent concentration of events in the zone between Augusta and Syracuse where important petrochemical facilities are present, with a resulting elevated secondary seismic risk. However, the heterogeneity in the distribution of events with respect to the time of day made us suspect that these seismicity maps are severely biased by artificial events, such as quarry explosions. We distinguished between tectonic earthquakes and quarry blasts by the inspection of waveforms of certain key stations, and by spectral analysis. As a general rule we found that the local tectonic microearthquakes are richer in high frequencies than the quarry blasts. All events which were identified as quarry blasts occurred during the daytime between 08:00 a.m. and 03:00 p.m. GMT and on weekdays from Monday to Friday. The aforementioned concentration of seismicity between Augusta and Syracuse disappeared when filtering out these events. Automatic discrimination was carried out in a straightforward way using Artificial Neural Networks (ANN) in a supervised classification. The application of the ANN to various test data sets gave a success of about 95%. This showed that our results obtained with a 'visual' discrimination are mathematically reproducible and not arbitrary.

**Key words** *Hyblean Plateau – quarry blasts – tectonic microearthquakes – discrimination – artificial neural network*

## 1. Introduction

Southeastern Sicily, in particular the Hyblean plateau, belongs to the zones in Europe with the highest seismic potential. During the past millenium Southeastern Sicily has been struck by several strong earthquakes (fig. 1), which reached magnitudes of about $M \cong 7.0$ and caused some 10000 deaths each. The Ionian coast currently belongs to the most populated areas in Sicily. Between the cities of Augusta and Syracuse there are important industrial petrochemical facilities with an elevated secondary seismic risk.

In the light of the considerable seismic potential, a permanent seismic network in the Hyblean Plateau was set up originally in the framework of the POSEIDON project. This network, now operated by the Istituto Nazionale di Geofisica e Vulcanologia, consists of nine digital three-components stations (fig. 1, table I) and is aimed at the seismic surveillance of the zone, and in particular the identification of faults with

*Mailing address*: Dr. Stefano Gresta, Dipartimento di Scienze Geologiche, Università di Catania, Corso Italia 55, 95129 Catania, Italy; e-mail: gresta@mbox.unict.it

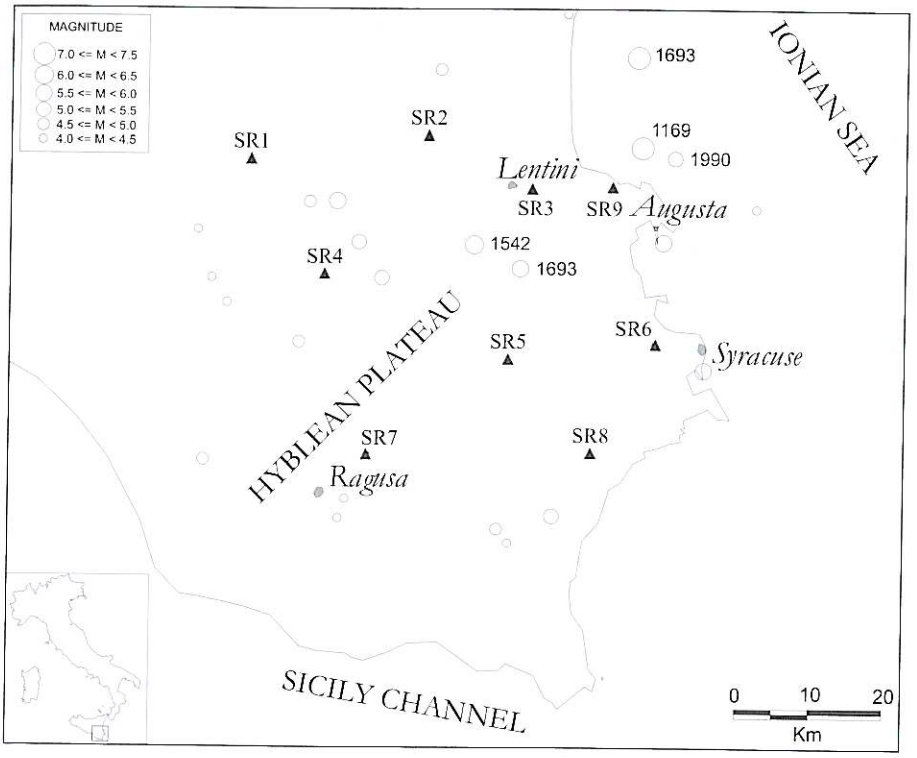


**Fig. 1.** Historical seismicity (open dots) in Southeastern Sicily. Epicenter coordinates were taken from Azzaro and Barbano (2000), coordinates of the event of 13 December 1990 from Amato *et al.* (1995). Black triangles represent stations of the seismic network.

**Table I.** Station parameters of the seismic network on the Hyblean Plateau.

| Station code | Lat. deg. N | Long. deg. E | Elev. (m) |
|---|---|---|---|
| SR1 | 37.318 | 14.600 | 530 |
| SR2 | 37.348 | 14.872 | 95 |
| SR3 | 37.283 | 15.032 | 190 |
| SR4 | 37.178 | 14.715 | 730 |
| SR5 | 37.075 | 14.998 | 600 |
| SR6 | 37.093 | 15.222 | 100 |
| SR7 | 36.958 | 14.782 | 540 |
| SR8 | 36.960 | 15.125 | 455 |
| SR9 | 37.285 | 15.155 | 126 |

Dynamic range 24 bit, Sampling Frequency 125 Hz

enhanced activity. The earthquakes recorded during the operation time of the network are of small magnitudes which have not yet exceeded $M$ = 4.0. Despite the modest seismic energy release, the distribution of hypocenters may give useful seismotectonic hints as to the existence of active faults, the orientation of seismic dislocations, the relation of seismicity to tectonic structures visible at the surface, and so on.

The geometrical distribution, the technical specifications of the network and the low noise level recorded at the stations permit a sufficient completeness of the catalog for events whose epicenters fall within the area shown in fig. 1 and whose magnitudes are greater than 1.0. The hypocentral coordinates have been calculated with the program HYPOELLIPSE (Lahr, 1989)

using the velocity model by Sharp *et al.* (1980). The errors, as inferred from the output of this routine, are in general of the order of 2 km (or less) for the epicenter, and 5 km (or less) for the focal depth (Sistema Poseidon, 2000).

## 2. Seismicity distribution in time and space

The seismicity map in the Hyblean Plateau for the events recorded in the years 1994-1998 (fig. 2) shows an apparent concentration of activity along the coastal zone between the towns of Augusta and Syracuse. In the distribution of earthquake occurrence with respect to the time of day a maximum between 09:00 a.m. and 10:00 a.m. GMT can be noted (fig. 3). This distribution makes us suspect that the data set is affected by the presence of artificial events, in particular quarry blasts. Quarry blasts occur more frequently during working hours of the day (*e.g.*, USGS, 1999). We therefore may obtain useful hints on possible biases introduced to seismicity maps by filtering the data set with respect to the time of day.

Between 03:00 p.m. and 08:00 a.m. GMT quarry explosions can be assumed to be substantially absent, and the seismicity pattern during these hours should not be seriously affected by these events. The activity in this time of the day shows a quite uniform geographical distribution over the area covered by the seismic network, whereas the aforementioned concentration of events between Augusta and Syracuse disappears (fig. 4a). *Vice versa*, in the hours from 08:00 a.m. to 03:00 p.m. GMT the majority of events occurs in the coastal range between Augusta and Syracuse, some more close to the

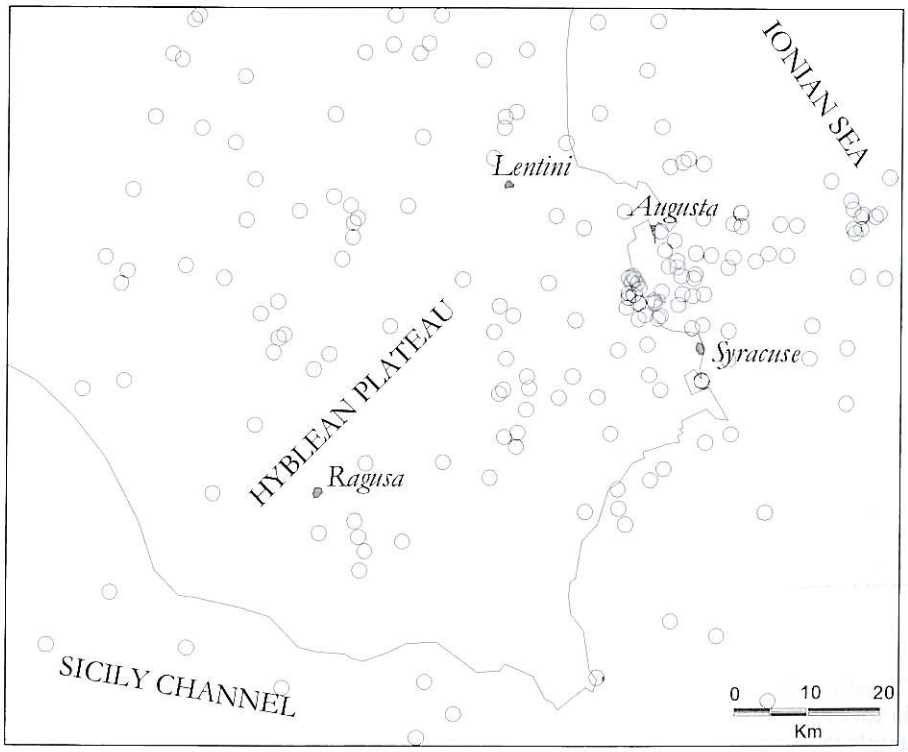


**Fig. 2.** Preliminary locations of the overall «seismicity» in the years 1994-1998 (after ISMES, 1999).

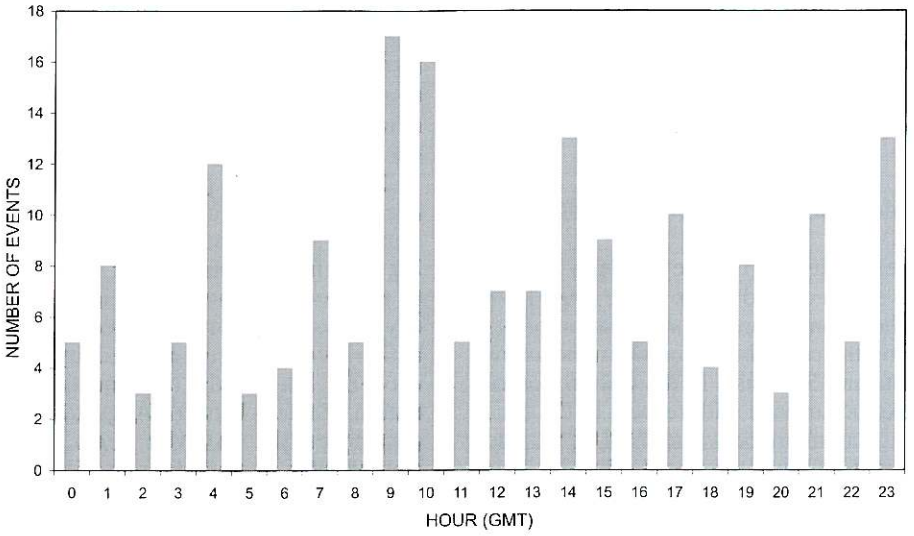


**Fig. 3.** Daytime distribution of overall «seismicity» in the years 1994-1998 (after ISMES, 1999).

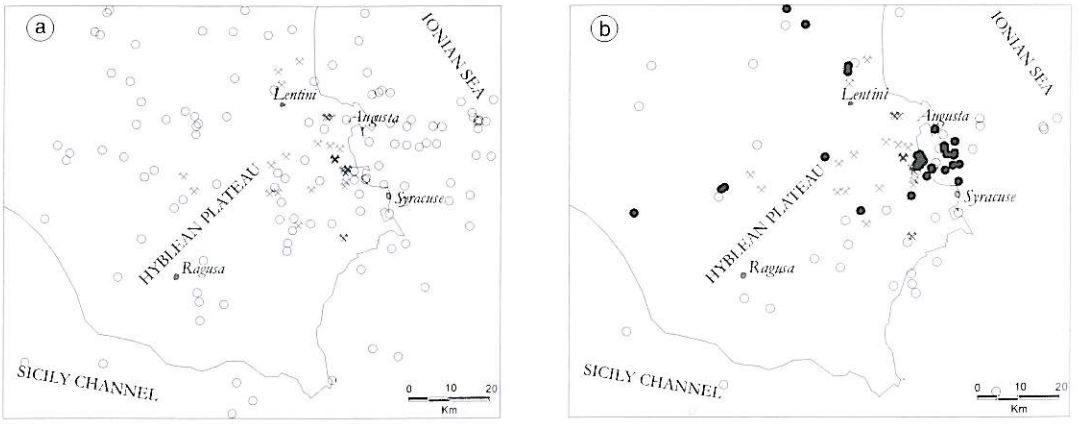


**Fig. 4a,b.** Distribution of the overall seismicity in the years 1994-1998: a) between 03:00 p.m. and 08:00 a.m.; b) between 08:00 a.m. and 03:00 p.m. where filled circles designate events which occurred between 09:00 a.m. and 10:00 a.m. Hours are in GMT. The hammer symbols give the location of quarries as reported by the *Distretto Minerario* of Catania.

town of Lentini (fig. 4b). The concentration of events in this zone is even more evident if we consider only the hours from 09:00 a.m. to 10:00 a.m. GMT, *i.e.* the hours where distribution of the activity with respect to the time of the day shows a maximum. It is not surprising that many of these events were located close to quarries, even though some of these locations are affected by high uncertainties. A good deal of these events have waveforms and spectra which we

found to be quite typical for quarry blasts, at least at certain stations (see below). In the data recorded during night-time we expect, on the other hand, that tectonic microearthquakes strongly prevail since quarries are unlikely to launch explosions during the night. As we show later, we were able to corroborate this hypothesis as we did not find signals during night time which showed the features recognized as typical for quarry blasts launched during the day.

## 3. Criteria for the separation of earthquakes from artificial events

A guideline for the discrimination of earthquakes from mine explosions has been given on the Web by USGS/NEIC (table II). This document points out that for many events not all types of evidence are available. Also in our specific situation the discrimination problem cannot be resolved by applying the criteria reported in the USGS/NEIC document in a schematic way. So the discrimination on the basis of the results of event location can be misleading, since the quarries fall into the same zone where earthquakes occur and, as we show below, the determination of the focal depth is ambiguous in certain cases. Contrary to what is stated by USGS/NEIC and despite their shallow depth, the quarry blasts are sometimes felt by the population in the vicinity (*i.e.* at distances of few kilometres) of the quarries, even though the computed magnitudes are low ($M \cong 1.0$). Occasionally, an acoustic signal is also reported. Note that most blasts are not quoted in the documents of the «Distretto Minerario». So data concerning the exact origin time of the blasts and quantity of used explosives are not available.

The interpretation of waveforms without *a priori* knowledge is tricky because the characteristics of the seismograms may vary significantly from station to station. The observation of similar waveforms – reported as one of the typical features for artificial events in the USGS/NEIC document – is not limited to explosion seismograms because multiplet earthquakes are common phenomena in the area. Also the similar amplitude criterion of USGS/NEIC has often failed since the explosions are not always carried out using the same quantity of explosive material.

**Table II.** Guideline for the discrimination of earthquakes from mine explosions (from USGS, 1999).

| USGS-NEIC discrimination criteria |
|---|
| 1. Computed locations – Locations of many of the provisionally identified explosions occur within or near well-known mining districts that have large surface mines, and provisionally identified explosions occur where other similar size events have regularly occurred at the same time of day. |
| 2. Time of day – Mine explosions tend to be set off during local daytime hours, even if the mines are operating 24 hours a day. |
| 3. Seismic waveforms – Seismograms at a given station for explosions at the same mine tend to be similar from event-to-event, both in the relative times and amplitudes of different seismic phases within each seismogram and in the absolute amplitudes of the seismic phases. Seismograms may have the general characteristics expected for mine explosions – emergent beginning of phases due to ripple-firing, no S, presence of Rg phase. |
| 4. Events not reported as felt – Calculated magnitudes of seismic events in some mining districts are large enough that, if the events were earthquakes, they would probably have been felt at nearby towns and reported to the USGS/NEIC or to regional seismographic network operators. |
| 5. Independent knowledge of operators of regional seismographic networks. |
| Seismic events that are in mining regions but that do not have strong characteristics of «normal» mining explosions have usually been cataloged as possible earthquakes by the USGS/NEIC if their magnitudes were of a size that the shocks might have been felt. Mine-associated rockbursts usually have been reported by the USGS/NEIC if their magnitudes were of a size that the rockbursts might have been felt. |

## 4. The problem of locations

In order to highlight the problem we examined the location of quarry blasts which were identified on the basis of waveform analysis and spectra (see below). The data set of quarry blasts we considered was recorded from September 1999 to February 2000. The small magnitudes, the often emergent *P*-wave onsets and unclear *S*-wave arrival times make impossible to precisely locate the explosions. Due to the small magnitudes, these events are recorded only at a few stations. As before, the locations were obtained with the HYPOELLIPSE program (Lahr, 1989)

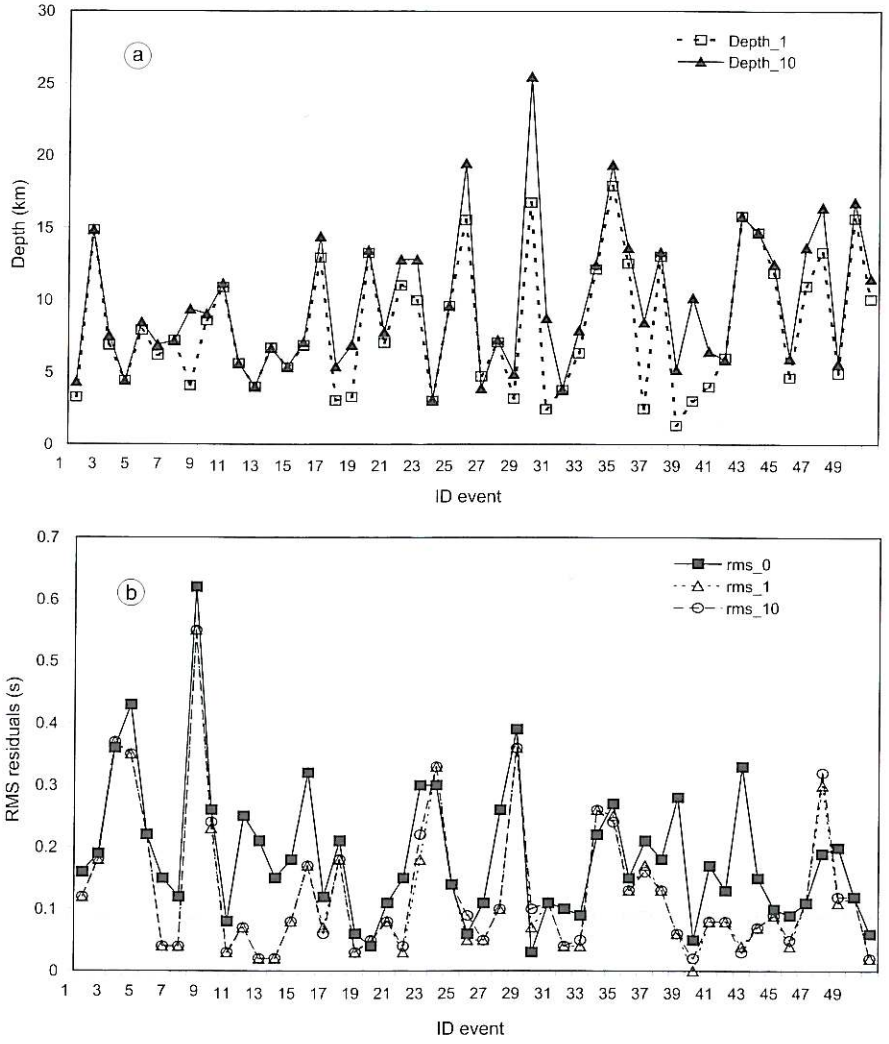


**Fig. 5a,b.** a) Focal depth of explosion events obtained with HYPOELLIPSE using a starting depth of 1 and 10 km; b) rms-travel time residuals for the solutions with focal depth fixed at 0 km, and for solutions obtained using depth of 1 and 10 km as starting values.

using arrival times of *P*-waves, and, when readable, *S*-waves (in the sense of «secondary» wave). The computed focal depths range from 1 to 25 km. The errors on hypocentral locations, as inferred from the output of the standard location, are often larger than for tectonic earthquakes, on average *ca.* 5 km for the horizontal coordinates and about 20 km for focal depth. In particular, the reported error on focal depth varies strongly from event to event. In some cases, the standard location fails to constrain the focal depth at all, *i.e.* yields an error of 99 km. In other cases, the calculated focal depths are around 15 km with errors in the order of only 2 km, which is an obvious contradiction of the nature of these events. Unfortunately, the problem of fixing the focal depth of the sources of these signals cannot be solved by simply choosing a shallow depth for the starting solution for the standard location. Even when we used 1 km as starting depth, the final computed focal depths are frequently greater than 15 km (fig. 5a). In a further step, we carried out the location using a fixed focal depth of 0 km, and compared the rms-residuals of travel times with the ones obtained before (fig. 5b). On the whole, the rms-residuals reported for the solutions with the depth fixed at 0 km are worse than the ones where the depth can vary. One possible explanation for this somewhat surprising result is an intrinsic ambiguity arising from the presence of a thick sedimentary layer stack in the Hyblean Plateau whose real thickness and wave propagation velocities, however, are not yet well known. Thus direct waves radiated by shallow sources, or refracted waves at some discontinuity within the layer stack, may have long travel times since they travel at low velocities within the shallower parts of the propagation medium. In this sense, the long travel times can be misinterpreted by standard location procedures as deriving from sources located at greater depth, particularly if the superficial structure in the crustal model of the standard location is not represented in a suitable way.

## 5. Waveforms and spectra

One might expect that the superficial position of explosion sources together with the firing techniques for quarry explosions produces specific signal characteristics, which could be exploited for discrimination purposes. According to USGS/NEIC in ripple firing, *P*-onsets are often emergent, clear *S*-waves are lacking, whereas dispersed Rayleigh-waves are observed. Unfortunately not all explosion seismograms show these features. In fig. 6a the waveforms and spectra of a microearthquake and of a quarry blast, both recorded at station SR9 are compared. The *P*-wave onset of the blast is even more impulsive on all three components than that of a nearby located tectonic earthquake, and we also note a rather clear *S*-wave onset for the blast. Amplitudes of the horizontal components, in particular the E-W component, are considerably higher than those of the vertical one. Dispersion is observed on all three components after the *S*-wave arrival for the blast but this may also happen in the case of a shallow earthquake. On the other hand, in the seismogram of the same quarry blast recorded at station SR6 (fig. 6b, right), we note an emergent *P*-onset, and a clear *S*-onset (in terms of a «secondary» phase) cannot be identified. A typical feature of explosion seismograms at this station is an envelope with a «fish-like» shape. Amplitudes of the vertical component are in the same order as those on the horizontal components.

The quarry blasts close to the city of Lentini can be distinguished rather clearly using the station SR2, where we note a very sharp *S*-wave onset on all three components of the earthquake seismogram (fig. 6c). Here peak amplitudes of the horizontal components are about three times higher than the amplitudes of the vertical component. For the explosion seismogram, the amplitudes of all three components are comparable. At station SR3 (fig. 6d), earthquake seismograms differ from explosions with respect to the envelope of the corresponding seismograms. However, earthquakes as well as explosions show larger peak amplitudes on the horizontal components than on the vertical one. We therefore focus on SR2, and SR6 as key stations, where a first visual distinction between earthquakes and explosions may be carried out more safely. Conversely, the station SR9 may lead to confusing results.

The discrimination by visual inspection of waveforms, using linguistically formulated cri-

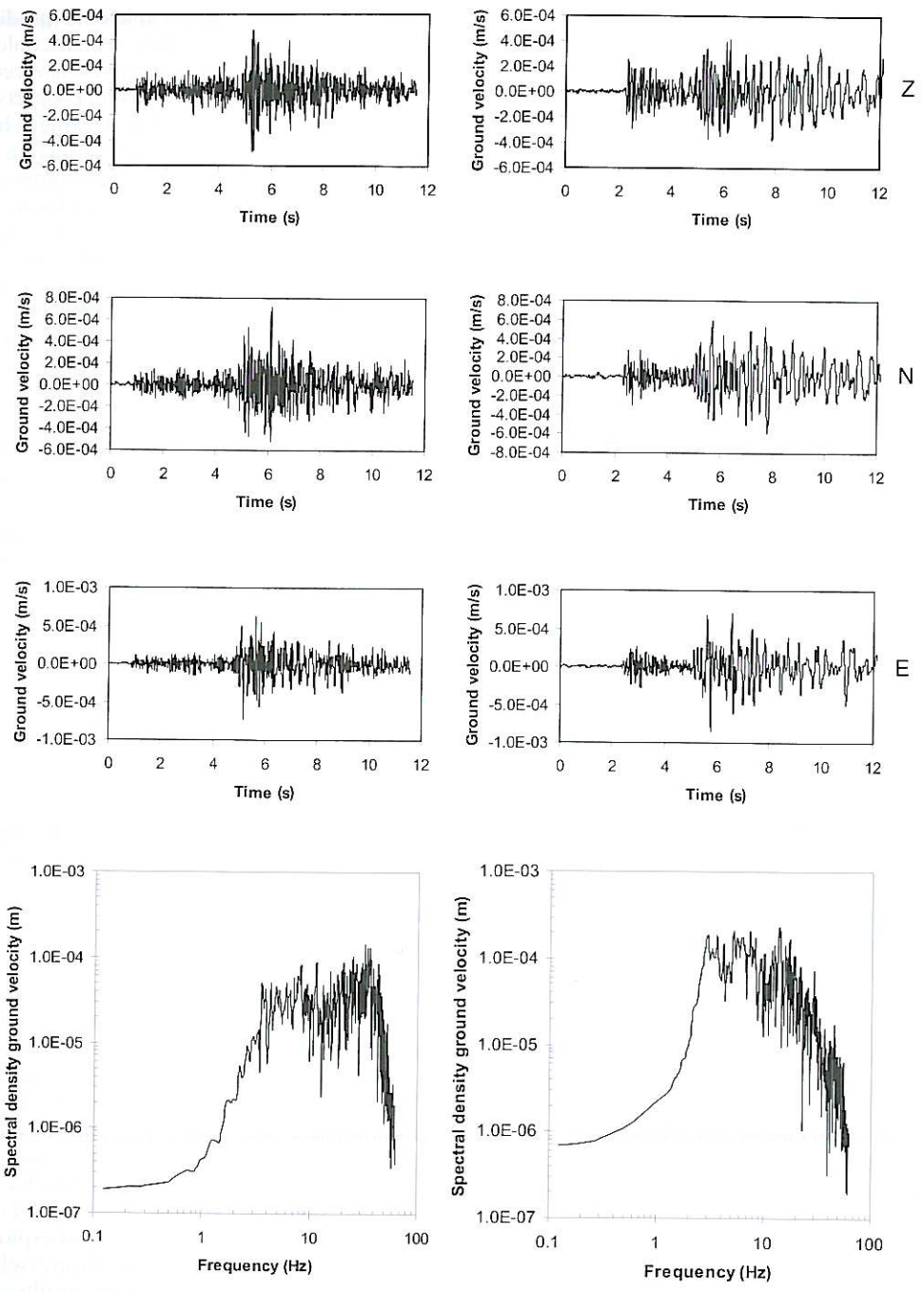


**Fig. 6a.** Example of seismograms and spectra at station SR9: earthquake (left); explosion (right). From top to bottom: vertical (Z), north-south (N) and east-west (E) components.

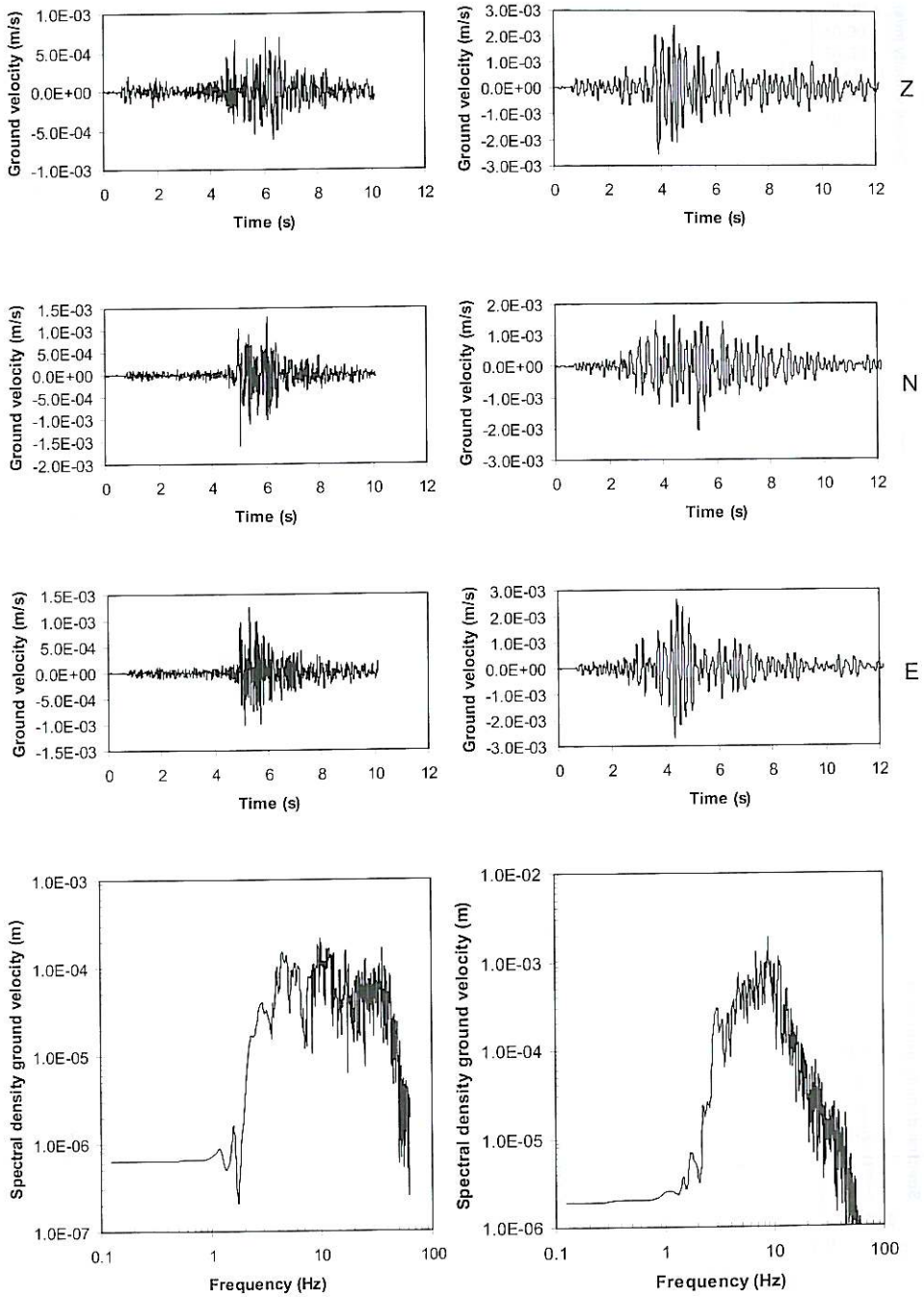


**Fig. 6b.** Example of seismograms and spectra at station SR6: earthquake (left); explosion (right). From top to bottom: vertical (Z), north-south (N) and east-west (E) components. The events are the same as in fig. 6a.

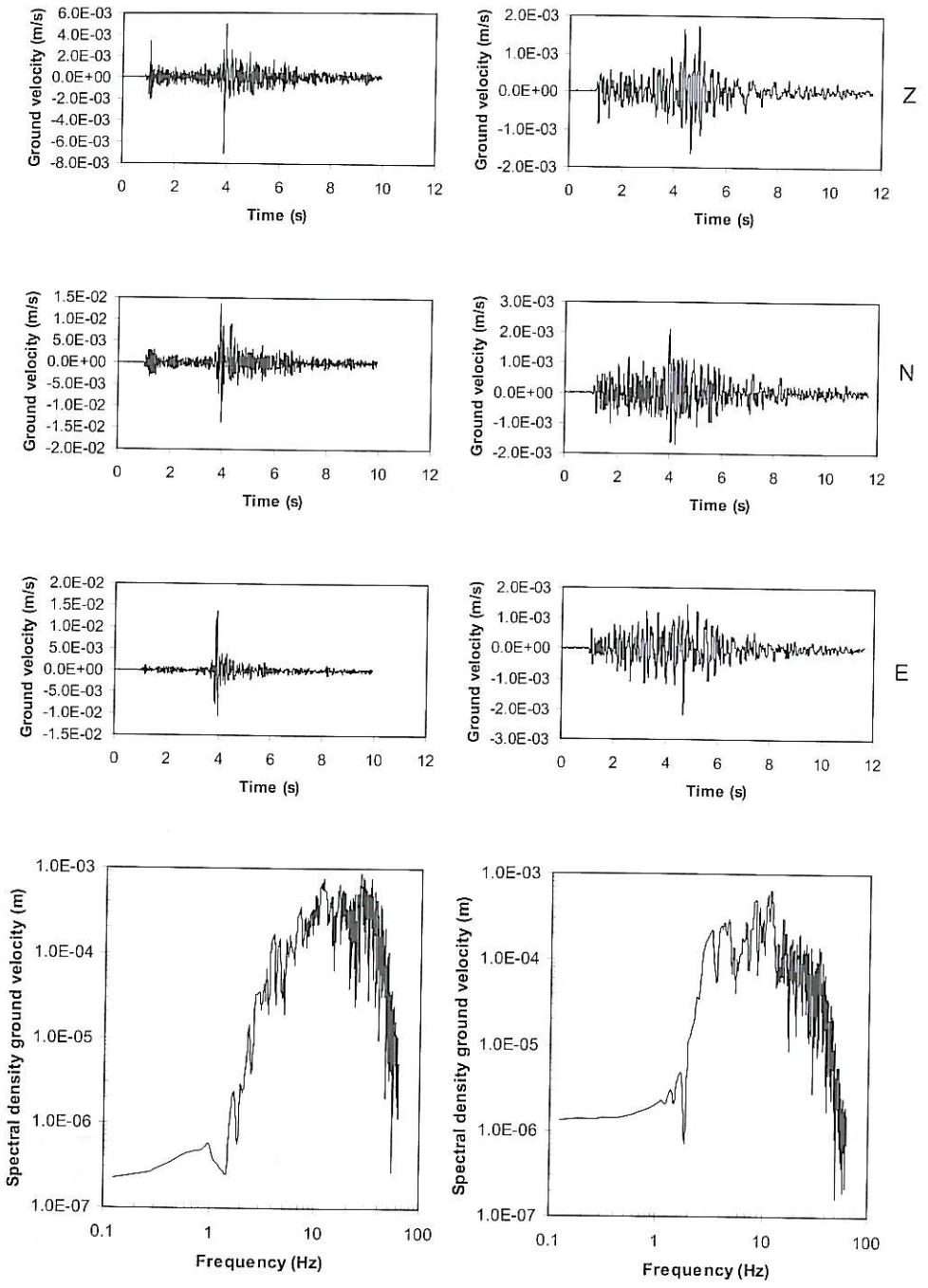


**Fig. 6c.** Example of seismograms and spectra at station SR2: earthquake (left); explosion (right). From top to bottom: vertical (Z), north-south (N) and east-west (E) components.

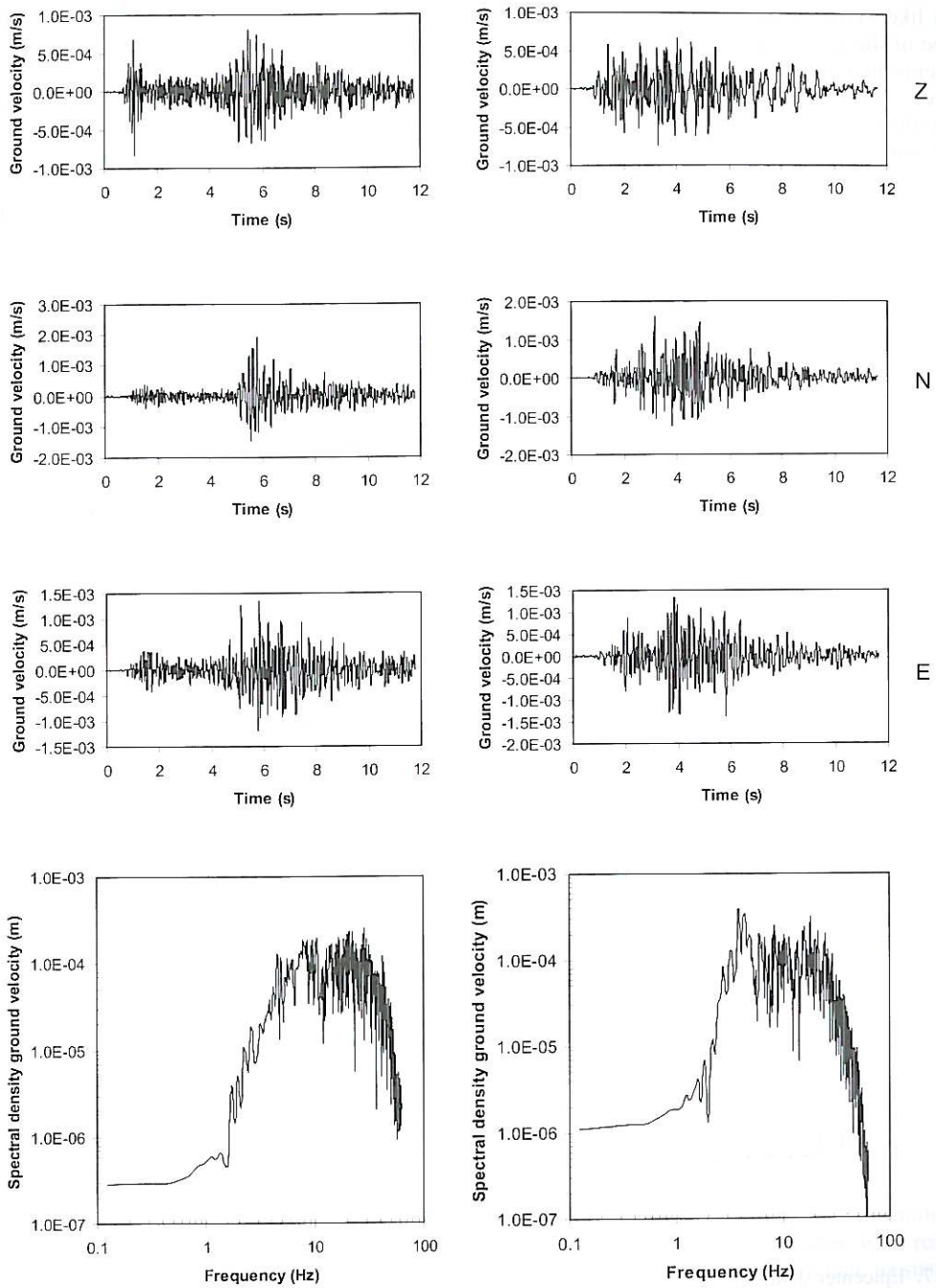


**Fig. 6d.** Example of seismograms and spectra at station SR3: earthquake (left); explosion (right). From top to bottom: vertical (Z), north-south (N) and east-west (E) components. The event are the same as in fig. 6c.

teria like «clear *S*-wave onsets» or «fish-like» shape of the seismogram envelope, is not readily reproducible and remains arbitrary to some degree. In order to find a way out of these difficulties, we analyzed the spectra of the vertical component seismograms to establish easier rules for the discrimination between quarry blasts and tectonic earthquakes. For the sake of simplicity, we reduced the preprocessing steps to a minimum. The spectra for which we show examples (figs. 6a-d) were obtained from the traces after applying a Butterworth band-pass filter with corner frequencies of 1.5 and 20 Hz and two filter sections. We considered a time window of length 8 s, starting at the *P*-wave arrival without considering the effects of mixing *P* and *S*-waves. Again, we considered data which were recorded in the time span September 1999 - February 2000.

The differences between the spectra of quarry explosions and earthquakes are clearly visible. Contrary to nuclear explosions, the seismograms of quarry blasts are poorer in high frequencies than seismograms of earthquakes with comparable magnitudes (Evernden, 1969; Elvers, 1974). At the key stations SR2 and SR6, the ground velocity spectra of the earthquakes studied here show dominating frequencies of up to 20 Hz, whereas the peaks in the spectra of the explosion seismograms fall into ranges between 3 and 10 Hz (see figs. 6b,c). Dominating frequencies are particularly low for the quarries

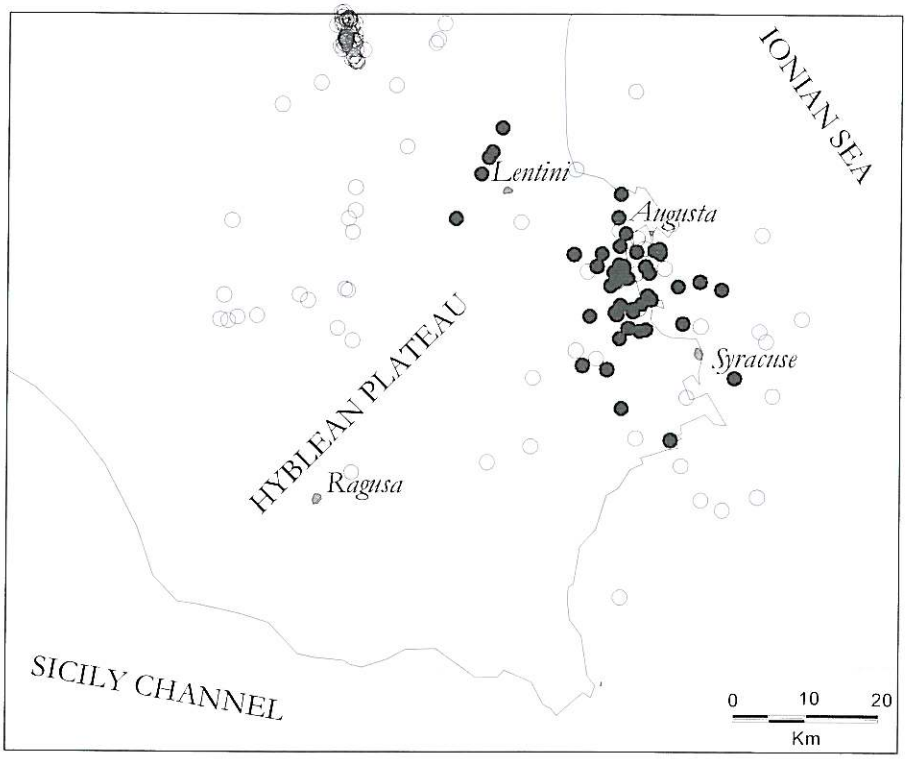


**Fig. 7.** Epicenter distribution of earthquakes (open circles) and quarry explosions (full dots) on the base of location with HYPOELLIPSE. The quarry explosions have been discriminated from earthquakes by visual inspection of the spectra. The time span considered is September 1999 to February 2000. Note that some of the quarry blasts were localized in the sea which is probably an effect of the small number of good onset readings available.

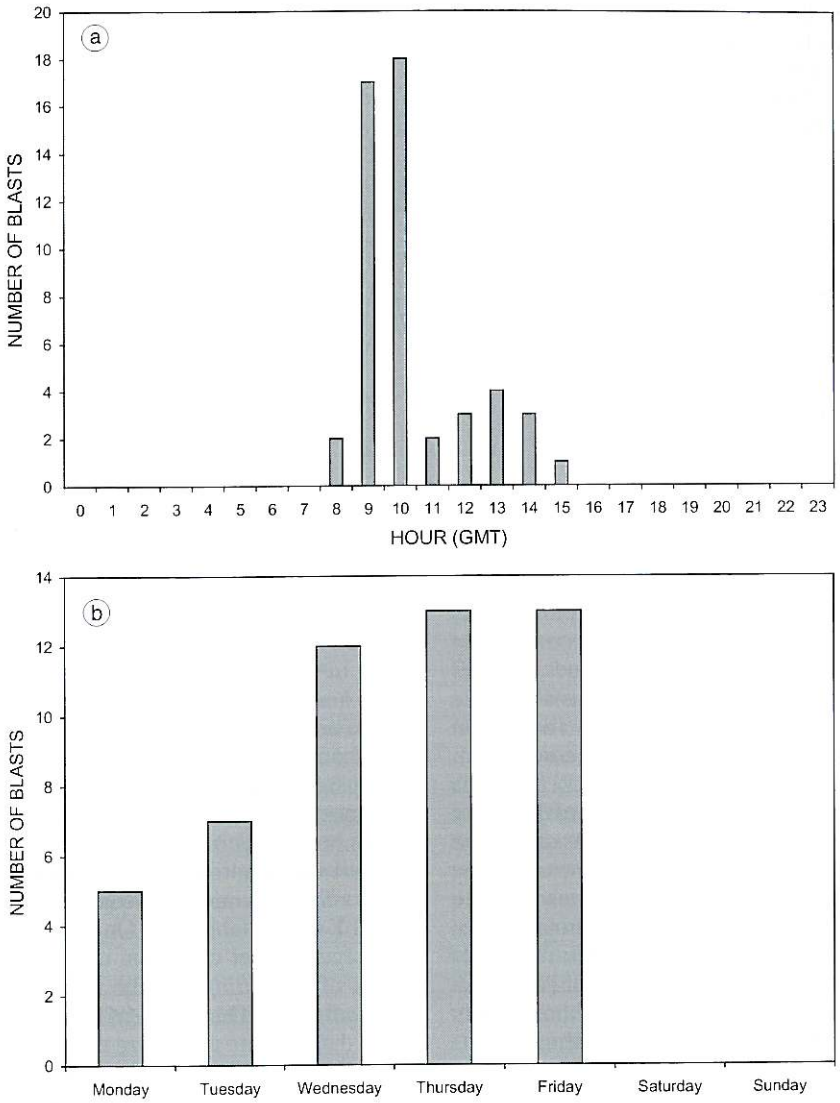


**Fig. 8a,b.** Daytime (a) and weekday (b) distribution of quarry blasts recorded in the time span September 1999 - February 2000.

between the towns of Augusta and Syracuse, where they hardly exceed 5 Hz. The spectra of the blasts close to Lentini show dominating frequencies between 5 and 10 Hz, in any case fairly below the dominating frequencies of earthquake spectra.

We now may re-examine the distribution of the quarry blasts and earthquakes with respect to space and time of the day. In fact, as suspected earlier, the «true» seismicity, *i.e.* comprising events identified as earthquakes on the basis of the spectral shapes, is distributed rather uni-

formly over the coastal zone and inland area (fig. 7). The cluster to the north is related to two earthquake swarms occurring in November 1999 and January 2000 (Scarfì *et al.*, 2001). On the other hand, quarry blasts are concentrated between Augusta and Syracuse, some more blasts were localized close to the town of Lentini. It turns out further, that most of the identified quarry explosions are launched between 09:00 a.m. and 10:00 a.m. GMT (fig. 8a) and are restricted to the working days Monday to Friday (fig. 8b).

## 6. Discrimination of quarry blasts form earthquakes with Artificial Neural Networks

In order to verify our 'visual' classification, either on waveforms or spectra, we applied so called Artificial Neural Networks (ANN) to our discrimination problem. This technique is straightforward for this task since, contrary to conventional classification techniques, it does not require any *a priori* knowledge about the mathematical structure of the discrimination function. A further important feature comes from the generality of ANN, *i.e.* the fact, that it is possible to approach any function of arbitrary degree of complexity. That way classes whose envelopes are of complicated shape can still be distinguished from each other, and it is possible to establish any kind of regression function between input and output vectors.

The aforementioned generality of ANNs requires at minimum a topology with three layers of nodes *i.e.* layers with input, hidden and output nodes (fig. 9a,b). The input layer serves for the storage of the data vector $\boldsymbol{U}$, with components normalized with respect to a maximum range between – 1 to 1. This input vector $\boldsymbol{U}$ is passed along the interconnections with the weights $w_{ij}$ to the next, the «hidden» layer. After applying an activation function to the values in the nodes of the hidden layer, the values are passed to the nodes of the output layer along the interconnections with the weights $w_{ik}$. In the output layer, the values obtained by the network are compared with the known ones $\boldsymbol{Y}$. During the «training phase» the input vectors $\boldsymbol{U}$ of the data set are passed to the ANN one by one. The weighting coefficients are adjusted in an iterative procedure minimizing the mismatch between target and calculated output vector $\boldsymbol{Y}$. Here we have been using the «Back Propagation Algorithm» (Werbos, 1974; Rummelhart *et al.*, 1986; see also Freeman and Skapura, 1992). Finally we obtain a function, which maps the input vector $\boldsymbol{U}$ to the output vector $\boldsymbol{Y}$

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

where:

$\hat{y}_k$ = $k$-th element of $\boldsymbol{Y}$ estimated by the network;
$\boldsymbol{U}$ = input vector;
$\boldsymbol{w}_j$ = vectors of weights between input and hidden layer;
$c_j$ = weights between hidden and output layer;
$t_j$ = biases;
$\sigma(\cdot)$ = sigmoid activation function $\sigma(z\cdot) = 1/(1 + e^{-z})$.

A second problem is the selection of a suitable network size: whereas the number of input and output components is given by the dimensions of $\boldsymbol{U}$ and $\boldsymbol{Y}$, there is no rule about the number of the neurons in the hidden layer. If the number of neurons in the hidden layer is too small the resulting regression function will be too smooth to match the true relation between $\boldsymbol{U}$ and $\boldsymbol{Y}$ in a suitable way. On the other hand, with a large number of hidden neurons one runs the risk of *overfitting*, *i.e.* the regression looses its significance. This problem is solved by dividing the data set into a *training set* and a *test set*. The training set is used for the determination of the weights. Applying the obtained weights to the data of the test set, which were not used during the training phase, one gains an idea about the general validity of the relation established with the neural network. It is useful to watch carefully the development of the global error encountered with both data sets. A commonly used measure for the global error is as the square sum of the differences between desired and calculated outputs of the data set, *i.e.*

$$GE = \Sigma_i \Sigma_k (\hat{y}_k - y_k)_i^2$$

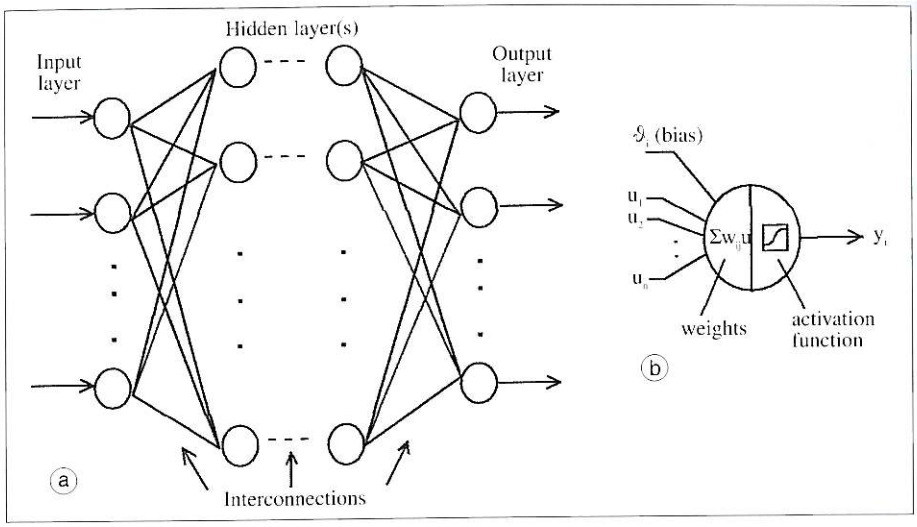


**Fig. 9a,b.** a) Topology of ANN; b) a single neuron. See text for details.

where $i$ runs from 1 to n (n = number of elements in the data set) and $k$ runs from 1 to the number of output parameters, here 2. During the iteration, the global errors encountered with the training and test data sets should decrease more or less monotonously. If the error of the test sets, however, inverts its trend and increases significantly after a number of iterations, one may suspect that the number of hidden nodes is too high and the mapping function is affected by overfitting.

## 7. Application of ANN to Hyblean Plateau data

A suitable transformation of the data facilitates the task of signal classification with ANN. As Falsaperla *et al.* (1996) and Langer *et al.* (1996) pointed out, the use of autocorrelation functions instead of plane waveforms, improves the results, probably because no problems with phase alignment occur. Besides this, autocorrelation functions represent the same information as amplitude spectra, which have proven to provide significant criteria during the 'visual' classification. Autocorrelation functions are normalized by definition, therefore, they can be directly used as input vectors in ANN applications. For the sake of simplicity of data preprocessing, we preferred autocorrelation functions to plane waveforms as well as spectra.

First we applied the ANN to the data recorded at the key stations SR2 and SR6 in the time interval from September 1999 to February 2000. Similar to the spectra discussed above, we obtained the autocorrelation functions (see examples shown in fig. 10) from the filtered vertical traces, taking a window of 8 s (1000 samples) starting from the $P$-wave onset. We divided the data set into two almost equal parts for the training and test data set, choosing the test set events randomly. 61 events made up the test set of station SR2 and 60 events at station SR6. We used a length of the input vector of 100, *i.e.* the first 100 points of the autocorrelation functions. The output vector is made up by two elements. Thus, the desired output vector $\boldsymbol{Y}$ for an earthquake reads as (1,0), and, conversely, (0,1) for a quarry blast. A classification of a test set event is considered as «failed» if the absolute difference between calculated and desired output of either of the two elements of $\boldsymbol{Y}$ is larger than 0.5. After some trial and error experiments we adopted an ANN with five hidden nodes (together with 100 nodes in the input layer and two nodes

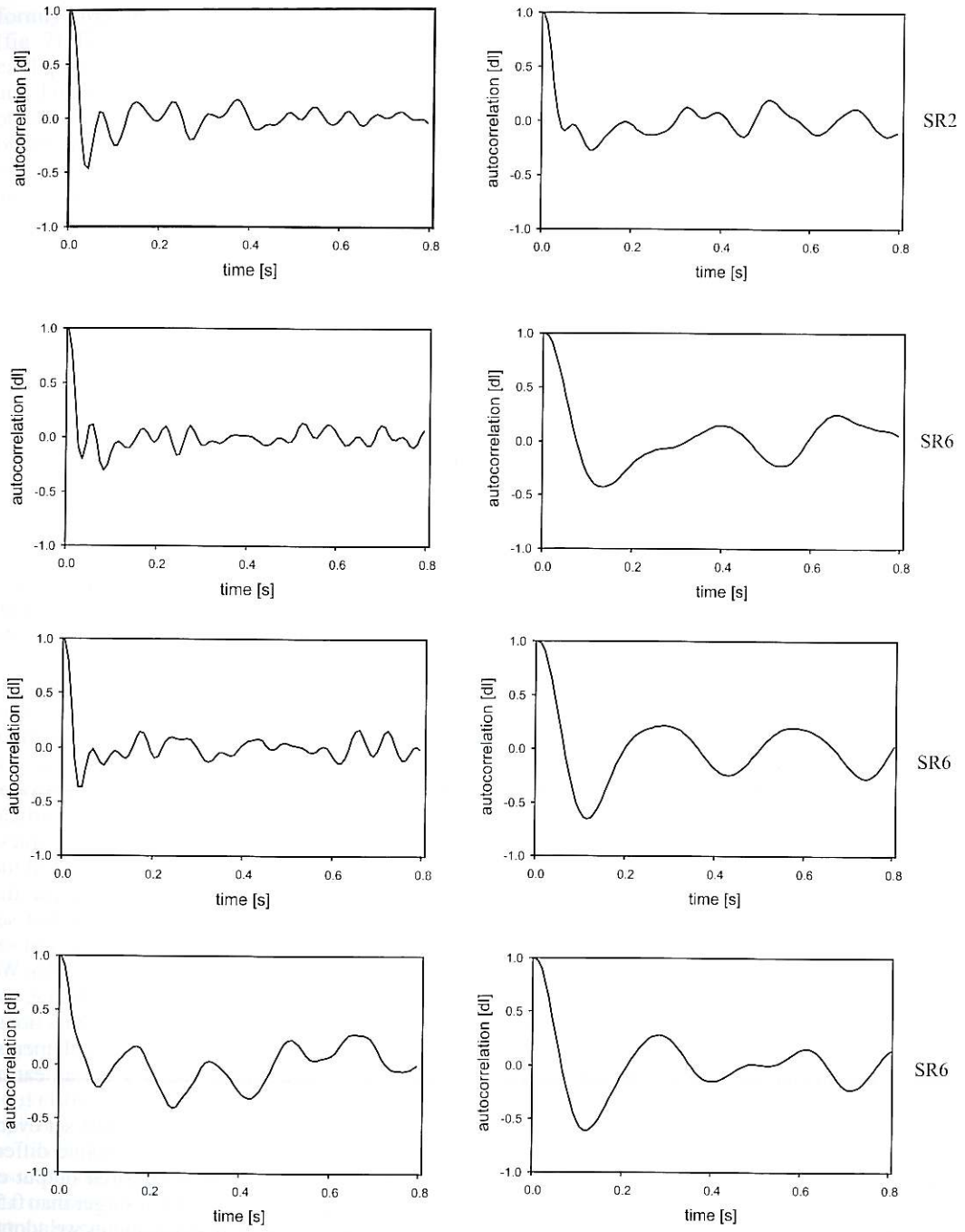


**Fig. 10.** Examples of autocorrelation functions at the two key stations (earthquakes on the left; quarry blasts on the right).

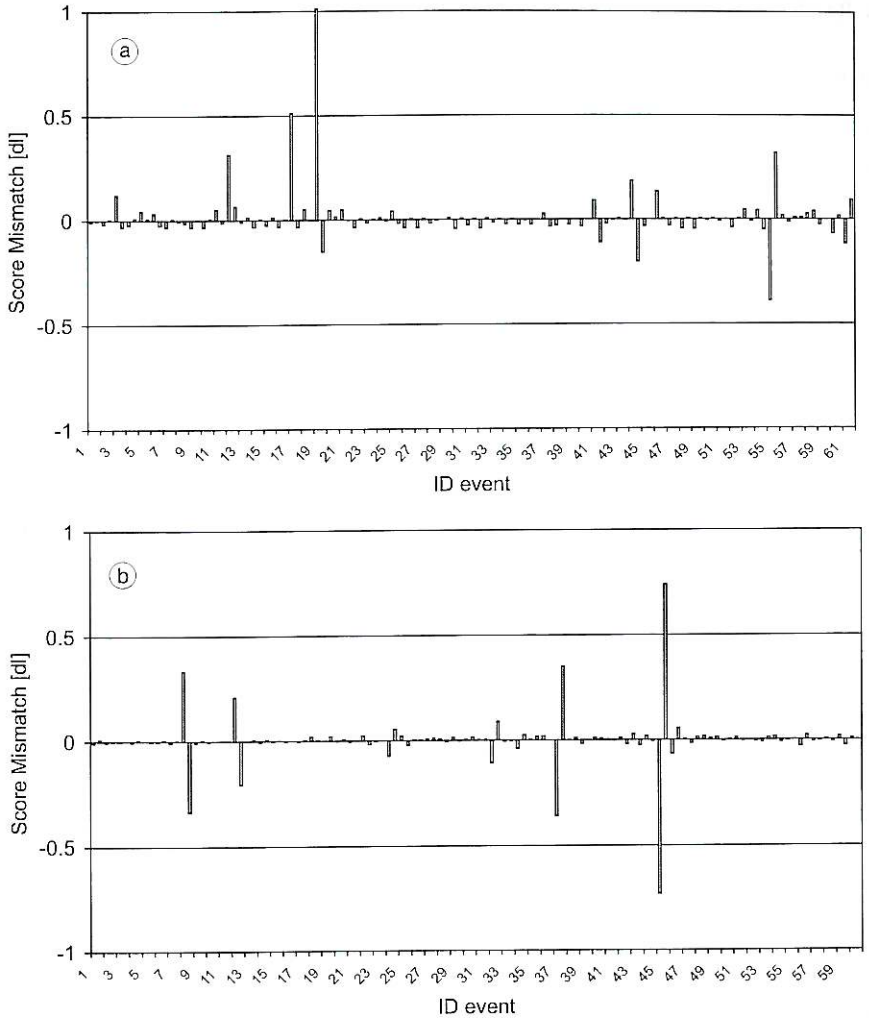


**Fig. 11a,b.** Classification mismatch of test set data at stations SR2 (a) and SR6 (b). The time span considered is September 1999 - February 2000. Note that each event is represented by a pair of score differences (see text). Thus a 'correctly' identified event has small differences with respect to both scores since the difference between desired and calculated output is small. An earthquake classified as quarry blast has a strong positive difference on the first score, and a strong negative one on the second score. A quarry blast classified as earthquake has a strong negative difference on first score and a strong positive one the second score.

in the output layer). The results of the classification are shown in fig. 11a,b where we note two misclassified signals (earthquakes) at station SR2, and one misclassified signal (quarry blast) at station SR6 (see also table III).

In the context of approximation, ANN can be considered as some kind of non-linear regression functional. That means their validity depends on whether the data set to which it is applied, *i.e.* the test set, belongs to the same

**Table III.** Results of ANN application to data recorded from September 1999 to February 2000 (see text for details).

| Station code | Number of events in training set | Number of events in test set | Misclassified earthquakes | Misclassified quarry blasts |
|---|---|---|---|---|
| SR2 | 57 | 61 | 2 | 0 |
| SR6 | 43 | 60 | 0 | 1 |

**Table IV.** Further results of ANN application: in this case the training set comprises data recorded from February 1994 to December 1998, whereas the test set comprises all data from September 1999 to February 2000.

| Station code | Number of events in training set | Number of events in test set | Misclassified earthquakes | Misclassified quarry blasts |
|---|---|---|---|---|
| SR2 | 73 | 118 | 1 | 2 |
| SR6 | 143 | 103 | 3 | 3 |

parent population as the training data which were used during the so called training phase, *i.e.* for the calculation of the weighting coefficients. In our first application set we guaranteed this condition by selecting the members of the test set randomly out of the whole data pool. In practice, one would like to use a once trained ANN to data recorded later. In this case, however, it is not guaranteed that the data used for ANN training belongs to the same parent population as the data to which the ANN is applied later. This is because the seismicity may change to some degree over the years: some faults become active, while others turn out to be quiescent, new quarries open and others may cease their activity. In order to emulate this situation, we separated test and training set by introducing a time cut, *i.e.* we considered the signals recorded from 1994-1998 as training set and all signals from September 1999 to February 2000 as test set. Using the same key stations (SR2 and SR6) and the same network topology we obtained the results summarized in fig. 12a,b and in table IV. Again we achieved a success rate of over 95% of correctly classified signals for both key stations.

## 8. Discussion and conclusions

The presence of quarries in an active seismic zone may cause a severe misrepresentation of the microseismicity. In our particular case, an uncritical representation of event location makes one believe that the coastal zone between Augusta and Syracuse has an enhanced level of activity compared to the inland areas of the Hyblean Plateau. However, the epicenters of events occurring during the night are uniformly distributed over the area covered by the seismic network. On the other hand, the events occurring at critical hours, *i.e.* between 09:00 a.m. and 10:00 a.m. GMT are concentrated in a dense cloud of epicenters between Augusta and Syracuse.

As it is difficult to distinguish between quarry blasts and tectonic earthquakes on the basis of hypocenter locations, we have been analyzing waveforms and spectra. As quarries are assumed to launch explosions only during working hours it is convenient to compare the signal characteristics of events occurring during night- and daytime. The visual distinction between waveforms of earthquakes and of quarry blasts, particularly when recorded on paper drum recorders, is not evident at all stations. Nevertheless, it was possible to identify two key stations (SR2 and SR6) where the differences between digital waveforms of earthquakes and of quarry blasts appear clear.

The differences between earthquakes and quarry blasts emerged most clearly from their spectra. Even reducing the signal preprocessing

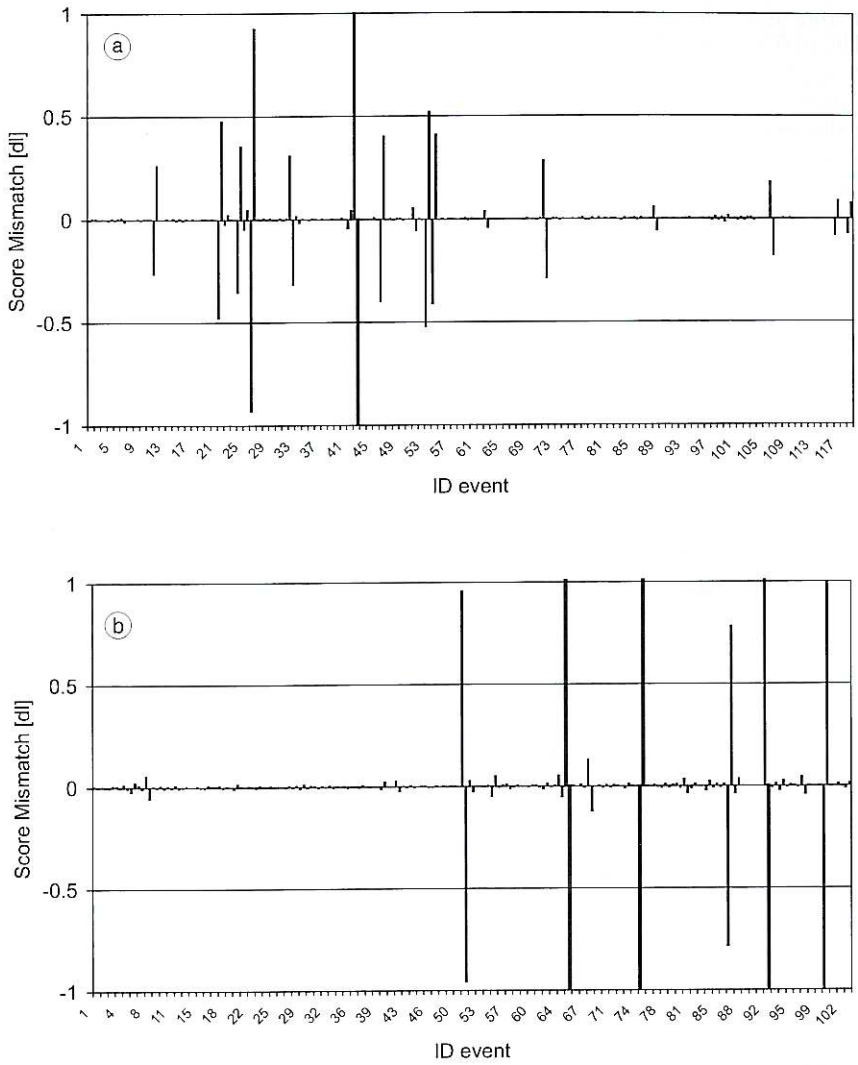


**Fig. 12a,b.** Classification mismatch of test set data at stations SR2 (a) and SR6 (b). The training set data were recorded from February 1994 to December 1998; the test set comprises all data from September 1999 to February 2000.

to a level which could be reached also by an automatic treatment, we find dominating frequencies of earthquakes being significantly higher than those of quarry blasts. This is also observed at stations where the visual inspection of waveforms led to some uncertainty of the correct event classification.

As a 'visual' discrimination can always be argued as arbitrary, we applied artificial neural networks in a supervised classification of signals as tectonic earthquakes or quarry explosions. This means that we defined training sets of events, whose origin (tectonic or artificial) was assumed to be known, and estimated a dis-

crimination function, which can be arbitrarily complex. The performance of this function was checked by applying it to a set of test events, again with known origin, but not used for establishing the discrimination function. Given the high rate of success, on average about 95%, we conclude that our 'visual' classification first carried out is mathematically reproducible and thus far from arbitrary. As the preprocessing steps are simple, it is possible to use ANN as a reliable tool for an automatic discrimination of quarry blasts from local tectonic microearthquakes.

Note that all events recognized as quarry blasts occurred during the daytime and in the working days of the week. This confirms our initial idea that quarry blasts, *i.e.* signals with characteristics identified as typical for those events, are essentially absent during the night times. If there were blasts during the night, then they should have characteristics significantly different from those launched during daytime. Otherwise, we should have picked them, at least when we carried out the automatic classification. Thus the quite uniformly distributed seismicity of the night-time events is closer to the true picture, whereas the apparent concentration of events in fig. 2 along the coastal zone between Augusta and Syracuse is certainly artificial.

Further, the second experiment, where we used the data from 1994-1998 as training set and the data from September 1999 - February 2000 as test set, still gives a success rate of about 95%. This shows the stability of the method as the two data sets – due to the possible non stationary behavior of seismic activity – are not fully representative for each other. Thus the success of event classification with ANN should not be not seriously affected by small variations of seismicity even though it may be useful to repeat ANN training with an updated data set from time to time.

## Acknowledgements

We are grateful to Dr. Giuseppe Montalbano (Distretto Minerario of Catania), who provided information on local quarries. We thank Dr. Rodolfo Console (Istituto Nazionale di Geofisica e Vulcanologia, Rome) and an anonymous reviewer for their constructive comments.

*REFERENCES*

AMATO, A., R. AZZARA, A. BASILI, C. CHIARABBA, M. COCCO, M. DI BONA and G. SELVAGGI (1995): Main shock and aftershocks of the December 13, 1990, Eastern Sicily earthquake, *Ann. Geofis.*, **38** (2), 255-266.

AZZARO, R. and M.S. BARBANO (2000): Analysis of the seismicity of Southeastern Sicily: a proposed tectonic interpretation, *Ann. Geofis.*, **43** (1), 171-188.

ELVERS, E. (1974): Seismic event identification by negative evidence, *Bull. Seismol. Soc. Am.*, **64**, 1671-1683.

EVERNDEN, J.F. (1969): Identification of earthquakes and explosions by use of teleseismic data, *J. Geophys. Res.*, **74**, 3823-3856.

FALSAPERLA, S., S. GRAZIANI, G. NUNNARI, and S. SPAMPINATO (1996): Automatic classification of volcanic earthquakes by using multi-layered neural networks, *Natural Hazards*, **13**, 205-228.

FREEMAN, J.A., D.M. SKAPURA (1992): *Neural Networks - Algorithms, Applications and Programming Techniques* (Addison-Wesley Publishing Company), pp. 401.

ISMES (1999): Progetto Poseidon: risultati preliminari della rete sismica digitale, *Technical Report*, pp. 48.

LAHR, J.C. (1989): HYPOELLIPSE/version 2.0: a computer program for determining local earthquake hypocentral parameters, magnitude, and first motion pattern, *U.S. Geol. Surv., Open File Rep., 89/116*, pp. 81.

LANGER, H., G. NUNNARI and L. OCCHIPINTI (1996): Estimation of seismic waveform governing parameters with neural networks, *J. Geophys. Res.*, **101**, 20109-20118.

RUMMELHART, D.E., G.E. HINTON and R.J. WILLIAMS (1986): Learning internal representation by error propagation, in *Parallel Distributed Processing*, vol. 1, 318-362 (MIT Press, Cambridge, MA, U.S.A.).

SCARFÌ, L., H. LANGER, G. DI GRAZIA, A. URSINO and S. GRESTA (2001): Analysis of two microearthquakes swarms in Southeastern Sicily: evidence for active faults?, *Ann. Geofis.*, **44** (4), 671-686 (this volume).

SHARP, A.D.L., P.M. DAVIS, F. GRAY, and G.C.P. KING (1980): Preliminary results from a seismic array study of Mount Etna, *United Kingdom Research on Mount Etna, 1977-1979*, The Royal Society, London, 15-18.

SISTEMA POSEIDON (2000): http://www.poseidon.nti.it/Report/ReportList.

USGS (1999): www.neic.cr.usgs.gov/neis.

WERBOS, P.J. (1974): Beyond regression: new tools for prediction and analysis in behavioral sciences, *Master Thesis*, Harvard University.

(received February 8, 2001;
accepted June 25, 2002)