

Signal enhancement with generalized ICA applied to Mt. Etna volcano, Italy

G. CABRAS¹, R. CARNIEL² and J. WASSERMAN³

¹ *Dipartimento di Ingegneria Elettrica Gestionale e Meccanica, Università di Udine, Italy*

² *Laboratorio di Misure e Trattamento dei Segnali, DIEA, Università di Udine, Italy*

³ *Dpt. of Earth and Envir. Sciences, Ludwig Maximilians Universität, München, Germany*

(Received: November 10, 2009; accepted: December 12, 2009)

ABSTRACT Independent Component Analysis (ICA) is an emerging new technique in the blind identification of signals recorded in a variety of different fields. ICA tries to find the most statistically independent sources from an observable random vector, with the only restriction that all sources, but at the most one, are non-Gaussian; no other a priori information on sources and mixing dynamic system are needed. The applications of this technique to the analysis of volcanic time series are until today relatively few. In this paper, we show that ICA is a suitable technique to separate a volcanic source component from ocean microseisms in a seismic data set recorded at the Mt. Etna volcano, Italy. The encouraging results obtained with this methodology in the presented case study support its wider applicability in the volcano seismology context. The separation and consequent elimination of noise components from the continuous seismic signal can in fact facilitate tasks such as the characterization of volcanic regimes, their relationship with tectonic activity and the identification of possible precursors of paroxysmal phases.

1. Introduction

The seismic wavefield near an active volcano consists of the mixture of signals in a wide range of frequency bands both from inside the volcano (our signal of interest) and from outside (i.e., noise components). Identifying and extracting a continuous volcanic component (e.g., volcanic tremor) embedded in background noise, possibly with a low signal-to-noise ratio, is challenging and becomes an important issue in broadband, multi-sensor signal processing. The problem can be restated as the decomposition of a multi-sensor data set into at least two complementary subspaces, namely the signal of interest and the background noise. In a volcanic data set, the background noise is often dominated by microseisms, i.e. seismic noise originating from ocean waves (Kadota and Labianca, 1981), and/or by a noise of anthropogenic origin (e.g., continuous industrial activities), which may show similar characteristics during the whole day or show, on the contrary, strong 24 hour cycles (Carniel *et al.*, 2008). Other noise components can have e.g., meteorological origin and can even modulate the volcanic signal itself (Neuberg, 2000). There are cases in which the traditional frequency band related to oceanic microseismic also contain non negligible volcanic source signals as demonstrated by De Lauro *et al.* (2005), relative to Stromboli volcano. The effect of wind on the volcano edifice is also reported by De Lauro *et al.* (2006). All these sources of noise are caused by independent processes that usually excite signals

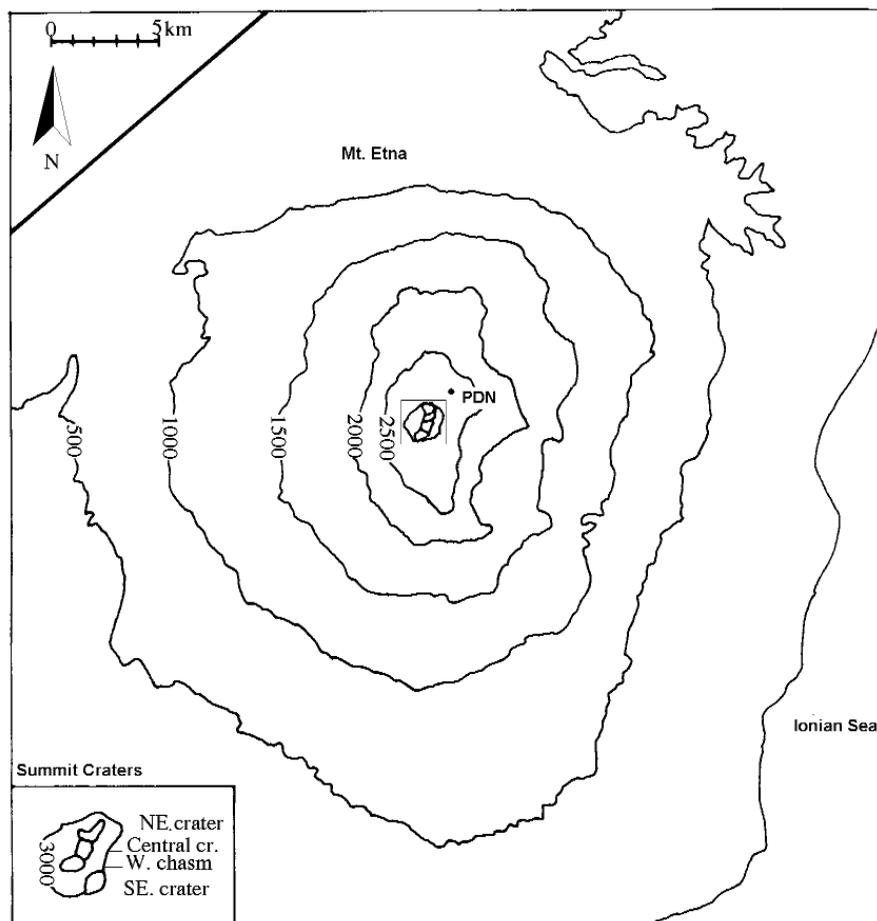


Fig. 1 - Sketch map of Mt. Etna with location of Pizzi Deneri (PDN) observatory and Summit Craters detail [modified from Wassermann (1997)].

in specific frequency bands (e.g., 0.1–0.3 Hz for microseisms) but they can sometimes contaminate the seismic signal over a wider frequency band, that is relevant to a volcanic tremor analysis, making a simple spectral filtering solution quite ineffective (Webb, 1998). Sometimes, even the volcanic signal itself can be seen as the superposition of different, simpler signals that share the same volcanic origin but may be e.g., located in different parts of the volcano [e.g., different vents or craters; Carniel and Iacop (1996)] and/or show distinctive characteristics due to the presence of multiple sources, which may share the same frequencies. An interesting example is the Erta Ale volcano in Ethiopia, where low and high regimes alternate and the question whether the high regime is characterized by a seismic signal that “sums up” over the low regime one, or on the contrary, by a seismic signal that “substitutes” the low regime one, is still under debate (Harris *et al.*, 2005; Jones *et al.*, 2006).

In this paper, we analyze continuous seismic recordings from Mt. Etna, in order to separate the volcanic component from a background noise mainly attributable to ocean microseisms. Ocean microseisms often share the same frequency band with low frequency tremor or volcanic

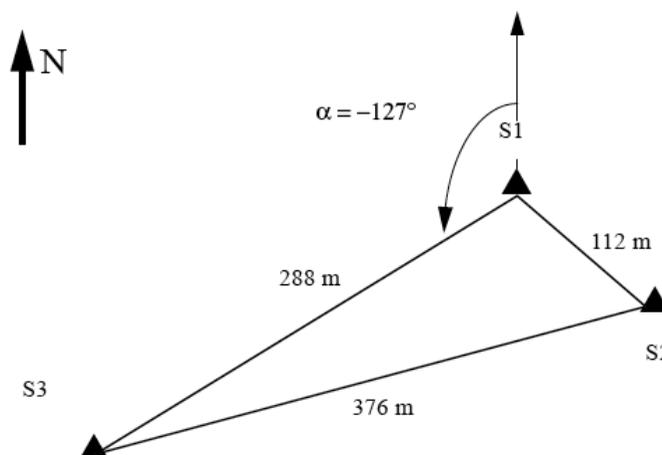


Fig. 2 - Sensor location: the reference sensor S1 was placed in the NW-end of the tiltmeter gallery of the observatory at Pizzi Deneri [adapted from Wassermann (1997)].

long period events thus preventing volcanic signals from further detailed analysis (e.g., Bartosch and Wassermann, 2004).

Mt. Etna (Italy) is an essentially basaltic volcano that is rather known for effusive eruptions, although it produced ignimbrites of more evolved composition about 15000 years ago and, more recently, peculiar endogenous lava domes (Behncke *et al.*, 2003). The persistent volcanic activity in the summit craters is characterized by phases of quiet degassing alternating with mild Strombolian activity, which occasionally evolves into fountaining and lava overflows. Its seismic activity includes a continuous volcanic tremor (Cannata *et al.*, 2008) and discrete events that can involve long and very long periods (Falsaperla *et al.*, 2002; Cannata *et al.*, 2009).

2. Data acquisition setup and observation

The deployment site is located at Pizzi Deneri (PDN) observatory, located NE of the NE crater of Mt. Etna (Fig. 1). The three Wielandt-Streckeisen STS-2 broadband seismometers were positioned as shown in Fig. 2. Vertical (Z) and horizontal (N and E) sensor components were recorded on a PDA data logger with a sampling rate of 20 Hz; continuous recorded waveforms were split into files containing 1 hour of data (i.e., 72000 samples). The observation period extended for several days between June 30, 1993 and July 23, 1993. In Fig. 3 the smoothed Welch Power Spectrum Density (PSD) of one hour of data of the reference station S1 are plotted in the broadband frequency range 0.04-8 Hz for the vertical (we denote it by Z1) and horizontal (N1 and E1) sensor components. The PSD show that the energy recorded from the horizontal components increases rapidly at very low frequencies (< 0.08 Hz), probably due to installation problems and/or wind pressure effects, whereas the vertical component has a narrower energy band-width; for this reason, we selected the vertical component of all 3 sensors for the blind

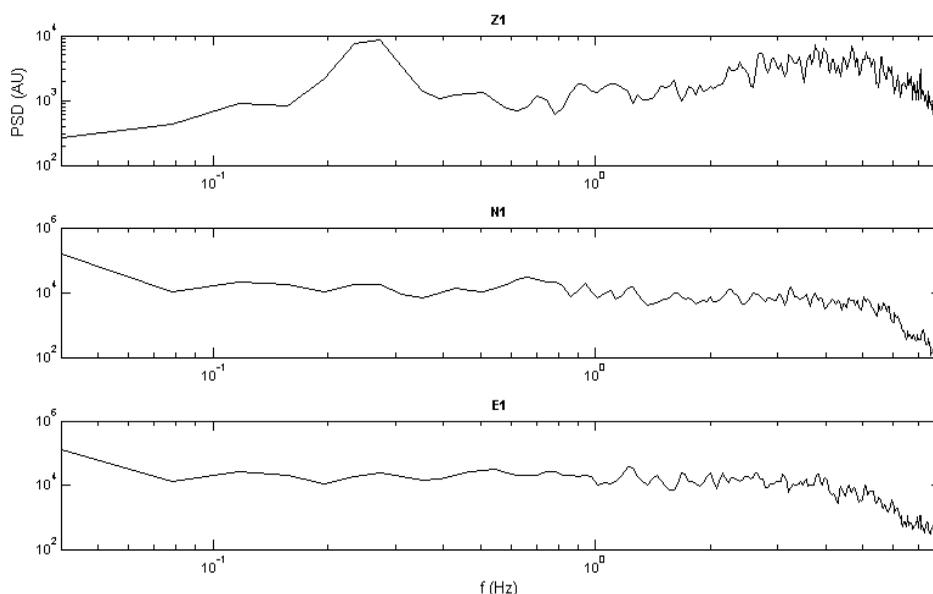


Fig. 3 - Broadband smoothed Welch PSD of vertical (Z1) and horizontal (N1 and E1) sensor S1 components, from July 23, 1993 11:00 to 12:00.

signal processing described below.

3. Blind Signal Processing

Blind Signal Processing (BSP) has received wide attention from many scientific and engineering disciplines concerned with understanding and extracting useful information from data as diverse as biology and bioinformatics (Ladisa *et al.*, 2005), neuroscience and neuroinformatics (Hasegawa, 2009), communications, complex stochastic dynamics (Capobianco, 2008), the world wide web, audio (Rufiner *et al.*, 2006), video, sensor signals, data mining and geophysical data processing. It is currently one of the most exciting areas of research in neural networks, statistical signal processing, unsupervised machine learning, information theory and exploratory data analysis. A general BSP problem can be formulated as follows (Cichocki and Amari, 2003): we observe records of m sensor signals $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ (where t is the time or sample index) from a multiple input/multiple output (MIMO) dynamical mixing system H ; the objective is to find an inverse system, termed a reconstruction system \mathbf{W} , if it exists and it is stable, in order to estimate the n primary source signals $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$. The estimation is performed on the basis of the n output signals $\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_n(t)]^T$ outcome from the reconstruction system, m sensor signals $\mathbf{x}(t)$ and a priori knowledge of the mixing system. Often, source signals are simultaneously linearly filtered and mixed: the aim is to process these observations in such a way that the original source signals are extracted by the reconstruction system. The problem of separating and estimating the original source waveforms from the sensor array, without knowing the transmission channel characteristics and the sources, can be expressed as a Blind Source Separation (BSS) process.

Our objective, in this simple but real case study, is to separate the volcanic origin component (our signal of interest) from the other, non-volcanic, source, dominated by ocean microseisms (background noise), and analyze the volcanic tremor in the low frequency range.

4. Generalized Independent Component Analysis technique

Independent Component Analysis (ICA) is a robust statistical technique in the field of BSS (Hyvärinen *et al.*, 2001). It separates a set of observed signals into the statistically most independent components by appealing to higher-order statistics (HOS), with the further a priori restriction that all, but at the most one, independent components are non-Gaussian.

In the simplest linear case, a number m of mixed signals $\mathbf{x}(t)$ are noise-corrupted linear combinations of n ($\leq m$) unknown mutually statistically independent, zero-mean source signals $\mathbf{s}(t)$:

$$x_i(t) = \sum_{j=1}^n h_{ij} s_j(t) + v_i(t), \quad (i=1, 2, \dots, m) \quad (1)$$

or in the matrix notation:

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{v}, \quad (2)$$

where \mathbf{v} is the vector of additive sensor noise and \mathbf{H} is an unknown full rank $m \times n$ mixing matrix. In other words, it is assumed that the signals received by an array of sensors are linear mixtures of primary sources. The separating (unmixing) matrix \mathbf{W} defined by

$$\mathbf{y} = \mathbf{W}\mathbf{x}, \quad (3)$$

combines the observations \mathbf{x} to generate estimates of the source signals:

$$\hat{s}_j(t) = y_j(t) = \sum_{i=1}^m w_{ji} x_i(t), \quad (j=1, 2, \dots, n) \quad (4)$$

The optimal weights correspond to the statistical independence of the output signals, $y_j(t)$, which are as independent as possible by evaluation of an information-theoretical cost function, e.g., as a minimum of Kullback-Leibler divergence (Cichocki and Amari, 2003).

Despite the success of using ICA to analyze synthetic or simple real data, some caution should be shown when using it to analyze real world problems, like in broadband seismic data analysis. In fact, the basic assumptions of ICA, which cannot estimate statistically dependent original sources, may not hold for some kinds of signals. Multiresolution Subband Decomposition ICA (MSD-ICA) is a generalization of ICA which relaxes the assumption regarding mutual independence of primary sources (Cichocki and Amari, 2003; Cichocki and Georgiev, 2003; Tanaka and Cichocki, 2004; Cichocki *et al.*, 2007). In this approach, the key assumption is that

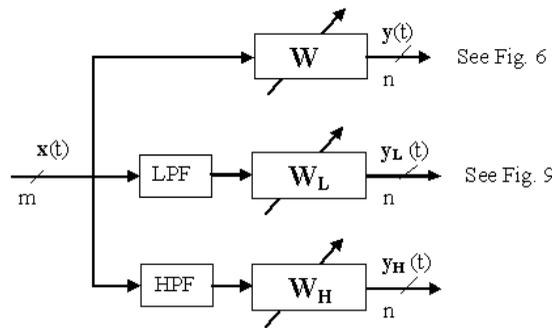


Fig. 4 - Block diagram illustrating the simplest case of MSD-ICA.

the unknown broadband sources can be dependent, but in some of the subbands, sources are independent. In other words, we assume that each unknown source can be modeled as a sum of the subband components:

$$s_j(t) = \sum_b s_{jb}(t), \quad (j=1, \dots, n). \tag{5}$$

In the simplest case, source signals can be modeled into their low- and high-frequency subband components:

$$s_j(t) = s_{jL}(t) + s_{jH}(t), \quad (j=1, \dots, n). \tag{6}$$

If, for instance, the low-frequency subband components $s_{jL}(t)$ can be identified to be mutually independent by some a priori knowledge, while the high-frequency subband components $s_{jH}(t)$ are weakly dependent, we can use a Low Pass Filter (LPF) to sensor observed signals and then apply any standard ICA algorithm to extract the low frequency subband components (Fig. 4).

In order to generalize this for more than two subband components, let us assume that only a certain set of subband components are independent. Provided that for some of the frequency subbands (at least one) all (multi)subband components, say $\{s_{jb}(t); j=1, \dots, n\}$, are mutually independent, we can easily estimate the mixing or separating system under the condition that these subbands are to be identified by some a priori knowledge. For this purpose, we simply apply any standard ICA algorithm, however not for all available raw sensor data but only for suitably (multi)subband sensor signals. After extracting the independent components from the mixture with ICA or MSD-ICA, we can examine the effects of discarding some of the components by reconstructing the sensor signals from the remaining components. The reconstruction allows us to remove undesirable components that are hidden in the mixture data. In other words, the reconstruction permits us to extract and remove one or more independent components from the mixture x . Back-projection of extracted interesting components allows us to remove noise, i.e. undesired components, and enhance information-

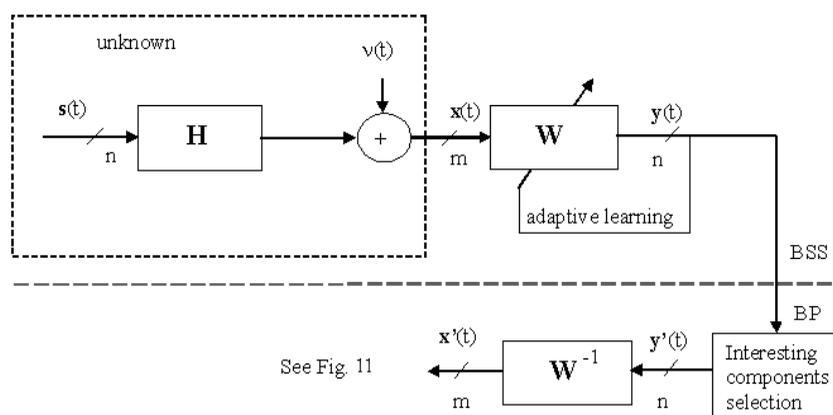


Fig. 5 - Block diagram illustrating standard BSS process followed by back-projection (BP) of selected components.

rich data.

A simple model to eliminate undesirable components from a multi-sensor data set is sketched in Fig. 5. In the first step, BSS is performed using a suitably chosen robust algorithm with respect to noise by a linear transformation of sensor data as $\mathbf{y} = \mathbf{W}\mathbf{x}$, where the vector $\mathbf{y}(t)$ represents the independent components. In the second step, the back-projection of selected components \mathbf{y}' onto the sensors is performed. The denoised sensor signals are obtained by a linear transformation $\mathbf{x}' = \mathbf{W}^{-1}\mathbf{y}'$, where \mathbf{W}^{-1} is the inverse (if $m=n$) or generalized pseudo-inverse (if $m>n$) of the estimated unmixing matrix \mathbf{W} and \mathbf{y}' is the vector obtained from the vector \mathbf{y} after removal of all the undesirable components, setting their values to zero. The entries of estimated attenuation (mixing) matrix \mathbf{W}^{-1} indicate how strongly each sensor picks up each individual component.

5. Application to volcano seismic signals: previous studies

As mentioned in the introduction, in recent years, ICA was adopted only sporadically to analyze multi-sensor volcano seismic data sets. Attention was mainly directed to the analysis of explosion quakes (Acernese *et al.*, 2000, 2004) and the very long period tremor (De Martino *et al.*, 2005; De Lauro *et al.*, 2006) of the Stromboli volcano, showing that HOS techniques are suitable for the isolation of components of the corresponding wavefield. ICA indicates, in particular, the presence of very-low-frequency content in the range of 0.1–0.5 Hz. Cabras *et al.* (2008) showed the feasibility of applying BSS techniques in order to separate volcanic and ocean microseism activity at Mt. Merapi volcano, Indonesia. Extensive applications of the ICA technique, complemented by other (in some sense more seismological) techniques, can be found in recent papers regarding the Erebus volcano (De Lauro *et al.*, 2009a, 2009b) as well as in the paper by Acernese *et al.* (2003) regarding Stromboli. The results from the Mt. Merapi volcano and the present application to the Mt. Etna volcano further support the importance of adopting nonlinear techniques in time domains in volcanology.

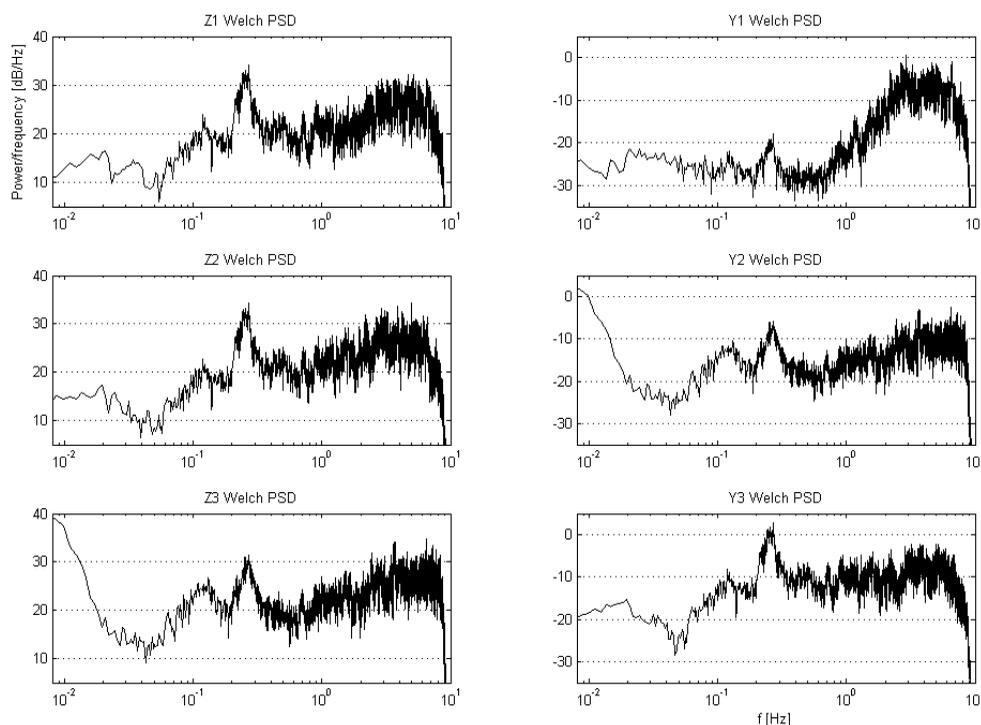


Fig. 6 - Welch PSD of three Z channel sensors (left) and respective reordered estimated sources Y (right) by ICA algorithm: Y1 is the estimated tremor source (0.28 peak suppressed), Y3 is the estimated microseism source (0.28 peak increased), Y2 is almost equivalent to Z3 (i.e. non-separated source).

6. Application of generalized ICA to the Mt. Etna volcano

The goal of the present study is the application of BSP to a seismic multi-sensor data set recorded at the Mt. Etna volcano. Volcanic tremor features can, in fact, provide precious information about the evolution of the eruptive activity (Cannata *et al.*, 2008). Its lower frequency band, in particular, was found to be, possibly, informative of deeper processes, not directly related to the observed eruptive activity (Alparone *et al.*, 2007). The frequency band 0.1-0.3 Hz during a low volcanic activity period with some evident discrete seismic observed events, is dominated by ocean microseism, a seismic noise originated from ocean waves (Webb, 1998), as we can see on the PSD of the STS-2 broadband seismometer vertical channel (Fig. 3, top), where the very sharp peak centered at 0.28 Hz dominates the low frequency spectrum. Our primary objective is to separate the component of volcanic origin (our signal of interest) from the other sources, dominated by ocean microseisms (background noise). To accomplish this, we apply ICA to the signal recorded by the three STS-2 vertical channel seismometers. As we can see in Fig. 6, the data set shows an evident background noise in the range 0.1-0.3 Hz, while the volcanic tremor shows most of its energy in the range 2-10 Hz. However, these different frequency ranges only indicate where each signal dominates, but do not exclude that there could still be superpositions between the frequency ranges of noise and signal, respectively, so that a classical e.g., high-pass

filter would not be an optimal solution. Moreover, in a volcanic scenario, like in many other real-world applications, recorded time series have complex features and properties: true latent sources are seldom all spatio-temporal decorrelated, or all statistically independent or all sparse. Thus, we need to apply several BSP techniques in order to estimate the desired latent sources in an optimal way.

Several BSS algorithms are available as powerful automatic procedures, which a user can apply to given data sets. Unfortunately, a simple direct application of these techniques to noisy volcanic data sets puts the need for a priori knowledge, expertise and proper pre-processing of signals and post-processing of models in a coherent and robust environment immediately in evidence. Otherwise, meaningless, inconsistent and even erroneous results can be obtained. Our approach was to accomplish this by applying and comparing several different Second Order Statistic (SOS) and Higher Order Statistic (HOS) algorithms in a stable and reliable software environment, the ICALAB (ICA in MATLABTM) Package for Signal Processing (Cichocki *et al.*, 2007) a toolbox developed and maintained by Laboratory for Advanced Brain Signal Processing, RIKEN, Japan.

We now briefly introduce the principal SOS and HOS BSS algorithms used in this analysis. More details can be found in the respective cited references.

AMUSE (Algorithm for Multiple Unknown Source Extraction) is a Second Order Statistic Spatio-Temporal Decorrelation (SOS-STD) BSS algorithm, which exploits the fact that the estimated components should be spatio-temporally decorrelated and less complex than any mixture of those sources (Tong *et al.*, 1993). AMUSE is very fast, parameter free and the output singular eigenvalues of the time-delayed covariance matrix are always uniquely ordered in decreasing order (i.e. in increasing complexity in the sense of best linear predictability). Its main drawback is the sensitivity to additive noise, particularly when the number of sensors is equal to the number of sources. This problem may be reduced using the SOBI (Second Order Blind Identification) algorithm (Belouchrani *et al.*, 1993), which tries to exploit the fact that the source has a stationary time structure, while the noise has not. This hypothesis, however, is not necessarily satisfied.

AMUSE and SOBI are able to estimate colored Gaussian distributed sources and estimate primary sources only if the sources have a temporal structure and different spectra. Moreover, these algorithms are not able to exploit statistical independence, since only second order statistical information is involved.

JADE (Joint Approximate Diagonalization of Eigenmatrices) is a HOS extension of SOBI to extract independent components (Cardoso and Souloumiac, 1993).

A widely used HOS algorithm is the Fixed-Point ICA (FPICA), also known as FastICA for its efficient convergence. FPICA is based on a fixed-point iteration scheme for finding a maximum of the non-Gaussianity of $\mathbf{y} = \mathbf{w}^T \mathbf{x}$. It combines the superior algorithmic properties resulting from the fixed-point iteration with the preferable statistical properties and extracts independent non-Gaussian distributed sources from the mixture in a sequential fashion, but not in a meaningful and determined order like AMUSE (Hyvärinen and Oja, 1997).

An improved version of FPICA is EFICA: an Efficient variant of algorithm FastICA. EFICA is, in fact, asymptotically statistically efficient, i.e. its accuracy given by the residual error variance attains the Cramér-Rao lower bound (CRB), which is the variance's theoretical minimum

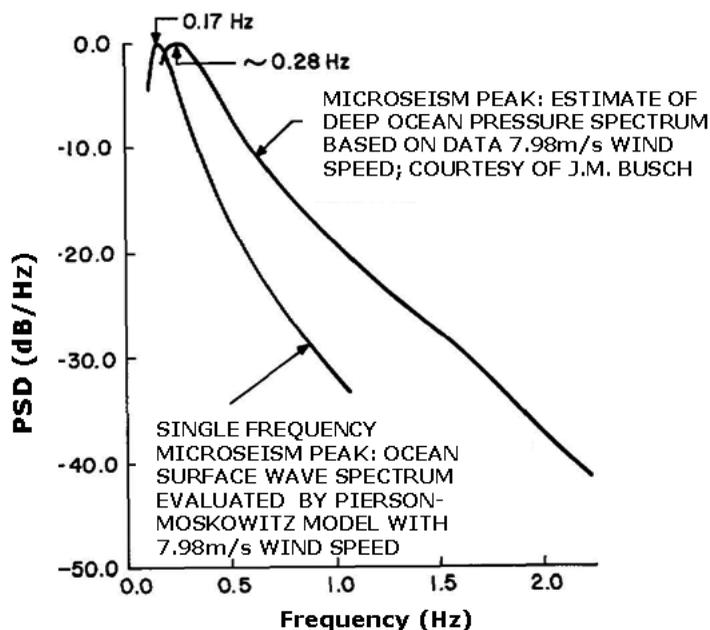


Fig. 7 - Ocean surface wave power spectrum density obtained by evaluating the Pierson-Moskowitz equation with the wind speed value of 7.98 m/s showing a peak at 0.17 Hz, and ocean bottom pressure power spectrum density estimated on a 24 hour average of measurements taken at a sensor on the ocean floor with surface wind speed value of 7.98 m/s showing a peak at ~ 0.28 Hz [modified from Kadota and Labianca (1981)].

thus the error is the smallest possible (Koldovský *et al.*, 2006; Tichavský *et al.*, 2006, 2008) .

7. Results and discussion

In order to understand the complexity of the tremor wavefield, embedded in background noise, we have processed several 1 hour records (72000 data samples) in the observation period of low Mt. Etna activity from June 30, 1993 to July 23, 1993, when data were collected from all three seismometers S1, S2 and S3 (Fig. 2). The most evident result, in correspondence of data recorded during windy days on the Ionian Sea (windspeed > 7 m/s recorded at the weather station in Catania), is the presence of a sharp peak centered at 0.28 Hz on all seismometers' vertical channel Welch PSD, as depicted on the left-hand side of Fig. 6. If we apply ICA to the three channels and look at the estimated three sources Y, it is noteworthy that all the used SOS and HOS BSS algorithms clearly separate the 0.28 Hz peak in the Y3 from the tremor spectral energy in the Y1 on the right-hand side of Fig. 6. The non-propagating, spatially incoherent energy below 0.02 Hz in sensor Z3, left unchanged in estimated source Y2 while not projected to estimated source Y1 and Y3 (i.e., is local to sensor Z3), could be generated by temperature-induced seismometer noise, in fact, seismometer S3 was located in a hole dug in a layer of lapilli and the hole was covered by an aluminium foil for temperature isolation.

To reinforce our confidence in attributing the energy in Y3 to the ocean microseism, we can

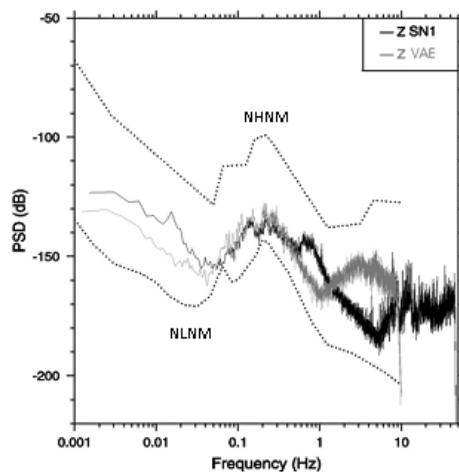


Fig. 8 - Welch PSD broadband Z channel seismometer of: SN1 (black line) installed at 2105 m deep floor station in open Ionian Sea, in proximity of Catania, and the VAE (grey line), a Qx80 STS-1 of Med-Net installed at ground in the middle of Sicily (37.469 lat 14.3533 lon, 735.1 m elev.), dataset collected during low Mt. Etna activity in 2002; dotted lines delimit the New High- and Low-Noise Models [NHNM and NLNM in Peterson (1993)]. From Monna *et al.* (2004).

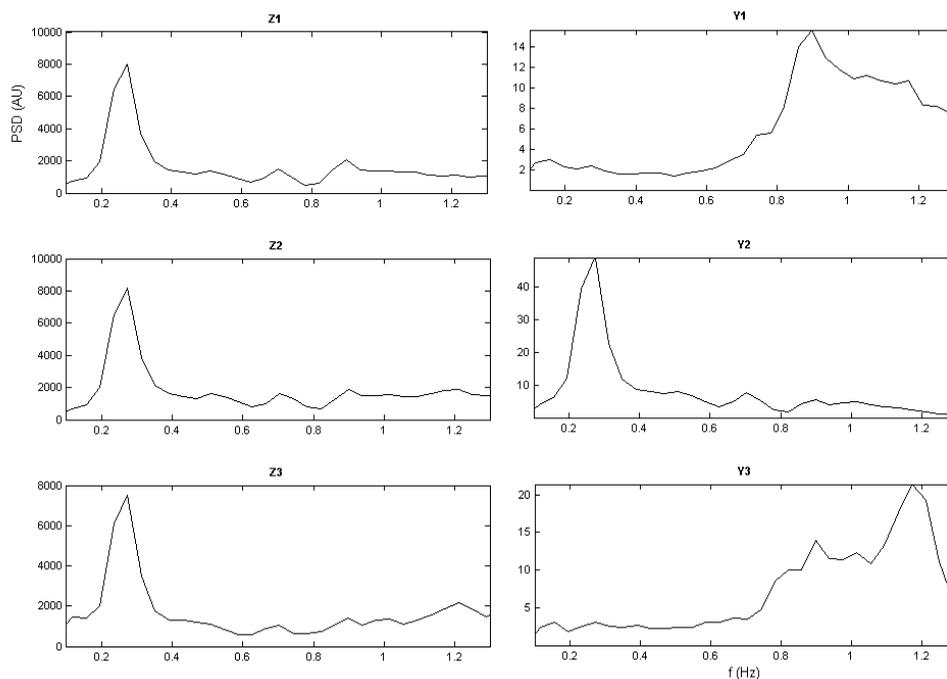


Fig. 9 - On the left, smoothed Welch PSD of three sensors vertical channel Z in the frequency range 0.1-1.3 Hz; on the right the respective estimated sources Y with ICA.

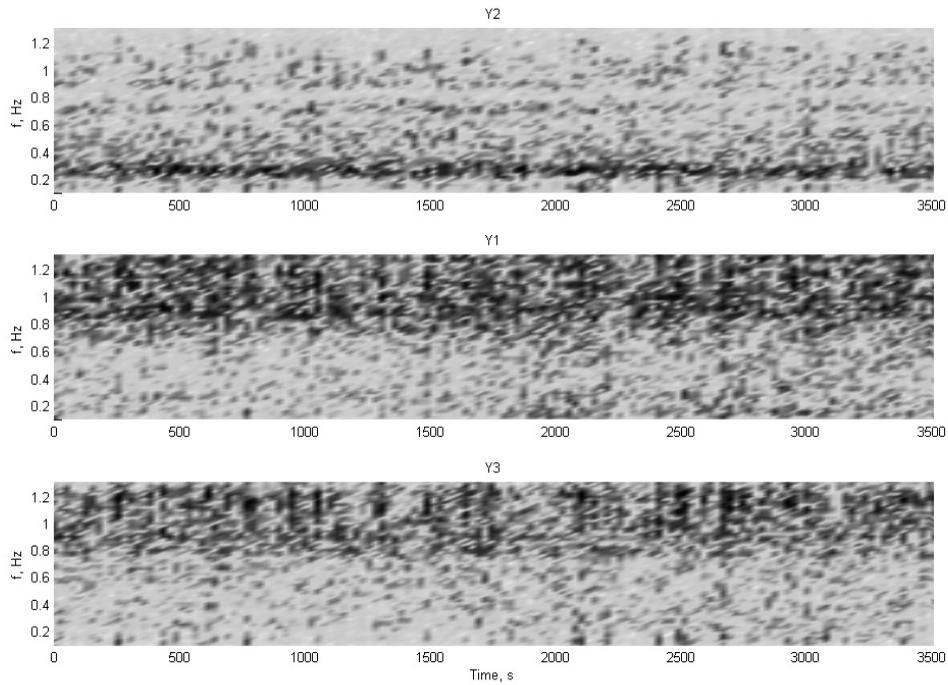


Fig. 10 . Spectrogram of estimated sources Y of Fig. 9, reordered by increasing frequency, black is coding higher spectrogram energy.

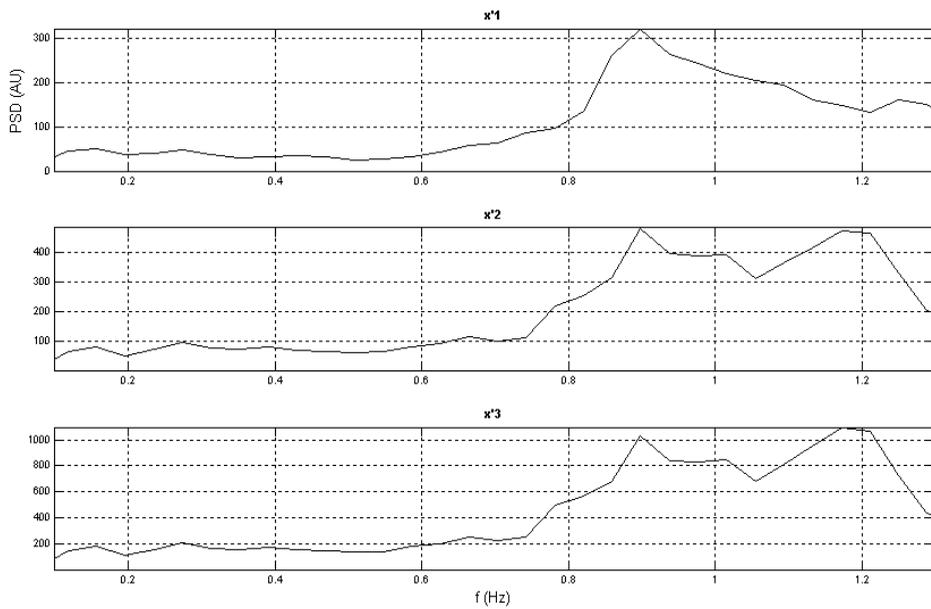


Fig. 11 - Reconstruction of denoised sensors x' by back-projection of estimated independent component Y1 and Y3, microseism component Y2 removed.

refer to a mathematical model of wind induced pressure fluctuations in the deep ocean (Kadota and Labianca, 1981), applied to experimental data from a local deep sea floor station SN1.

In the mathematical model, the ocean-surface-wave PSD and the ocean-bottom pressure PSD are obtained by evaluating the Pierson-Moskowitz equation with a windspeed of 7.98 m/s, resulting in an energy peak at about 0.28 Hz, as depicted in Fig. 7.

Experimental data were recorded at station SN1 (Seismic Network 1), installed in open Ionian Sea at 2105 m depth floor, in proximity to Catania; data were collected during low Mt. Etna activity in 2002 (Monna *et al.*, 2004). Again, as we can see in Fig. 8, the Z channel of SN1 has a peak at about 0.2 Hz. We can conclude that sustained wind at sea acts as a prominent ocean microseism source that can be recorded by land seismometers in the proximity of the volcano during a low activity phase.

Now, we focus our interest to the 0.1-1.3 Hz, range, on a data set where sea microseism origin dominates the wavefield. If we simply bandpass the Z channels of the 3 sensors, apparently, we

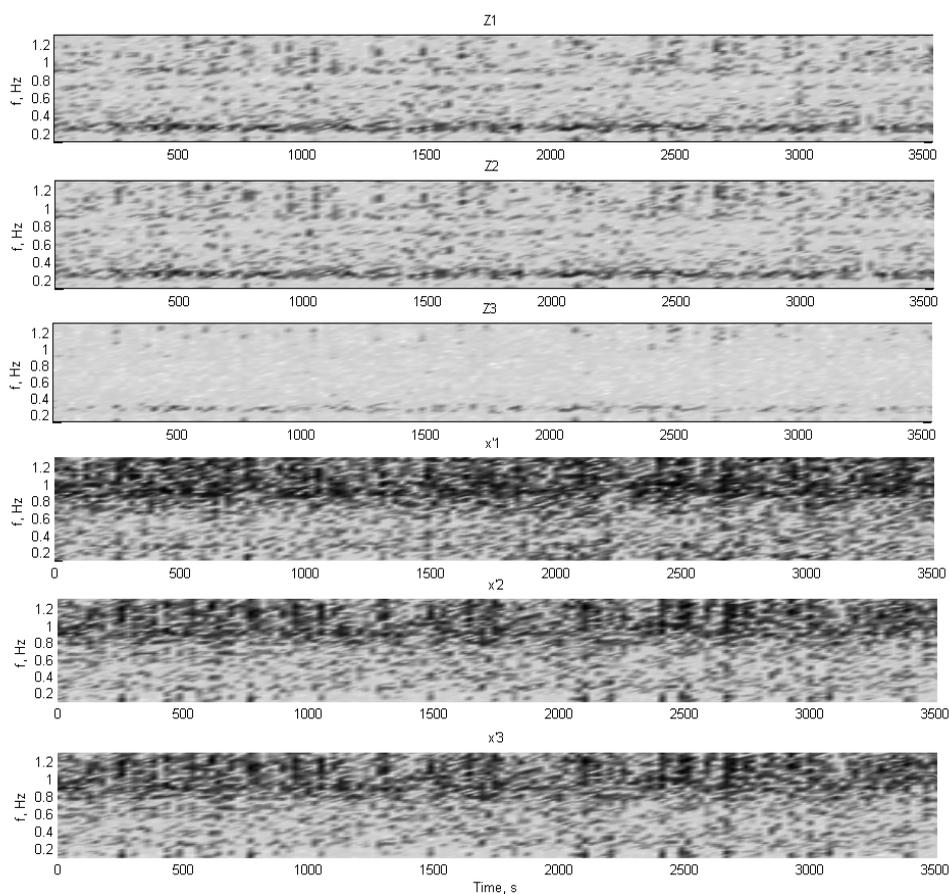


Fig. 12 - Signal enhancement (black is coding higher spectrogram energy). Top: spectrogram of original vertical channel seismometer signals Z1, Z2, Z3 in the frequency range 0.1-1.3 Hz, dominated by stationary microseism peak. Bottom: spectrogram of denoised sensors $x'1$, $x'2$, $x'3$ of Fig. 11, showing better time-frequency structures of volcanic origin without oceanic microseism contamination.

do not get more information from volcano sources (see Fig. 9 left side), because ocean microseism energy dominates the frequency range, but if we analyze the data with MSD-ICA, i.e. processing with ICA the band-passed sensor array, we can obtain enhanced information on the primary sources in that range (see Fig. 9 right-hand side). In this case, we continue to get the sea microseism source in Y2, but we also get two energy sources, Y1 and Y3, peaked at 0.9 and 1.2 Hz respectively, which are known Mt. Etna tremor spectral dominant peaks, observed during the pre-eruptive period (Lombardo *et al.*, 1996; Alparone *et al.*, 2007). In order to analyze the evolution of spectral features over time, a spectrogram of separated components is graphed in Fig. 10: on the top, we easily recognize the (relatively constant) ocean microseism, while Y1 and Y3 appear as more time-variable, and partially spectrally overlapping, signals. We do not attempt to identify Y1 and Y3 as individual volcanic meaningful sources but rather we recombine them, thus dividing the whole data sources into a volcanic subspace with similar features (Y1 and Y3), and a noise subspace, dominated by sea microseism (Y2). We can now simply enhance the volcanic sensor data by applying a back projection to the selected subspace given by the combination of Y1 and Y3, thus removing the noisy component Y2. Fig. 11 shows the smoothed PSD of the resulting band-passed and denoised sensor vector \mathbf{x}' , which could be further analyzed with standard techniques like time-frequency analysis as shown in the spectrogram of Fig. 12, in order e.g., to better investigate the time evolution of its spectral content. This can facilitate e.g., the identification of volcanic regimes (Jones *et al.*, 2006), the time scales of preparation of paroxysmal events (Carniel *et al.*, 2006) and/or the effect of external triggers such as tectonic events (Carniel and Tárrega, 2006).

8. Conclusions

In this paper, we proposed an application of BSS and ICA to the analysis of continuous seismic data recorded at Mt. Etna volcano, showing that these techniques are suitable to separate tremor components from ocean, microseism-dominated, background noise. The results confirm what was previously obtained at the Mt. Merapi volcano analyzing data recorded by one broadband and two short-period seismometers (Cabras *et al.*, 2008), although in that work only HOS methods separated tremor from ocean microseism successfully; this underlines the fact that in general, we must apply and compare several different BSS techniques to separate meaningful volcanic sources. In the present work, where the data recorded by three broadband seismometers were analyzed, we had the opportunity of exploiting the use of MSD-ICA as a generalization of ICA in order to analyze the complexity of the volcanic tremor wavefield in the frequency band 0.1-1.3 Hz. In this frequency band, dominated by the presence of microseismic noise, a blind source separation applied after a properly chosen prefiltering phase clearly highlighted the presence of significant components of volcanic origin, with spectral peaks at 0.9 and 1.2 Hz. After this identification we could enhance the signals of interest, and only those, recorded at each sensor, by back projecting only the selected estimated tremor components. The result is a significantly less noisy volcanic tremor signal, that can be used to perform any subsequent, more sophisticated analysis. This case study underlines that it is possible to filter broadband seismic data recorded by an array of seismometers even in the presence of a low signal-to-noise ratio and a non trivial intersection between the frequency bands of the signal of interest and noise, i.e. the

part of the signal we are not interested in. In the specific case, this was dominated by ocean microseism, but in other cases the source of noise could be obviously of completely different origin, including an anthropogenic one. The generalized ICA technique can then assume an important role in the efficient filtering of low signal-to-noise volcanic seismic data by extracting coherent volcanic components from a, possibly much more energetic, non-volcanic background noise.

Acknowledgements. The analyzed data set was collected within the framework of the DFG grant Schi 120/14-2&3. The authors would like to thank the Dipartimento di Fisica dell'Università degli Studi di Udine for logistic and technical support. Comments by Luciano Telesca and by an anonymous reviewer helped to improve the manuscript.

REFERENCES

- Acernese F., Ciaramella A., De Martino S., Falanga M. and Tagliaferri R.; 2000: *Neural networks for blind sources separation of Stromboli explosion quakes*. In: Pajunen P. and Karhunen J. (eds), Proceeding of ICA 2000, International workshop on independent component analysis and blind signal separation 19-22 June 2000, Helsinki, Finland, pp. 441-446.
- Acernese F., Ciaramella A., De Martino S., Falanga M. and Tagliaferri R.; 2003: *Neural Netw., for blind sources separation of Stromboli explosion quakes*. IEEE Trans. Neural Netws., **14**, 167-175.
- Acernese F., Ciaramella A., De Martino S., Falanga M., Godano C. and Tagliaferri R.; 2004: *Polarisation analysis of the independent components of low frequency events at Stromboli volcano (Aeolian Islands, Italy)*. J. Volcanol. Geotherm. Res., **137**, 153-168.
- Alparone S., Cannata A. and Gresta S.; 2007: *Time variation of spectral and wavefield features of volcanic tremor at Mt. Etna (January-June 1999)*. J. Volcanol. Geotherm. Res., **161**, 318-332.
- Bartosch T. and Wassermann J.; 2004: *Wavelet-coherence analysis of broadband array data at Mt. Stromboli volcano*. Bull. Seim. Soc. Am., **94**, 44 -52.
- Belouchrani A., Abed-Meraim K., Cardoso J.F. and Moulines E.; 1993: *Second-order blind separation of temporally correlated sources*. In: Proc. Int. Conf. on Digital Sig. Proc., Limassol, Cyprus, pp. 346-351.
- Behncke B., Neri M. and Carniel R.; 2003: *An exceptional case of endogenous lava dome growth spawning pyroclastic avalanches: the 1999 Bocca Nuova eruption of Mt. Etna (Italy)*. J. Volcanol. Geotherm. Res., **124**, 115-128.
- Cabras G., Carniel R. and J. Wassermann J.; 2008: *Blind source separation: an application to the Mt. Merapi volcano, Indonesia*. Fluctuation and Noise Letters, **8**, 3-4.
- Cannata A., Catania A., Alparone S. and Gresta S.; 2008: *Volcanic tremor at Mt. Etna: inferences on magma dynamics during effusive and explosive activity*. J. Volcanol. Geotherm. Res., **178**, 19-31.
- Cannata A., Hellweg M., Di Grazia G., Ford S., Alparone S., Gresta S., Montalto P. and Patané D.; 2009: *Long period and very long period events at Mt. Etna volcano: characteristics, variability and causality, and implications for their sources*. J. Volcanol. Geotherm. Res., **187**, 227-249.
- Capobianco E.; 2008: *Kernel methods and flexible inference for complex stochastic dynamics*. Physica A: Statistical Mechanics and its Applications, **387**, 4077-4098.
- Cardoso J.F. and Souloumiac A.; 1993: *Blind beam-forming for non Gaussian signals*. Radar and signal processing, IEEE Proceedings-F, **140**, 362-370.
- Carniel R. and Iacop F.; 1996: *On the persistency of crater assignment criteria for Stromboli explosion-quakes*. Annali di Geofisica, **39**, 347-359.

- Carniel R. and Tárrega M.; 2006: *Can tectonic events change volcanic tremor at Stromboli?* Geophys. Res. Lett., **33**, L20321, doi:10.1029/2006GL027690.
- Carniel R., Del Pin E. and Tárrega M.; 2008: *Event recognition by detrended fluctuation analysis: an application to Teide-Pico Viejo volcanic complex, Tenerife, Spain.* Chaos, Solitons and Fractals, **36**, 1173–1180.
- Carniel R., Ortiz R. and Di Cecca M.; 2006: *Spectral and dynamical hints on the time scale of preparation of the 5 April 2003 explosion at Stromboli volcano.* Canadian Journal of Earth Sciences, **43**, 41-55.
- Cichocki A. and Amari S.; 2003: *Adaptive Blind Signal and Image Processing.* John Wiley, Chichester, UK, 552 pp.
- Cichocki A. and Georgiev P.; 2003: *Blind source separation algorithms with matrix constraints.* IEICE Trans. Fund. Electron. Commun. Comput. Sci, E86-A, 513–522.
- Cichocki A., Amari S., Siwek K., Tanaka T. et al.; 2007: ICALAB toolboxes, <http://www.bsp.brain.riken.jp/ICALAB>.
- De Lauro E., De Martino S., Falanga M., Palo M. and Scarpa R.; 2005: *Evidence of VLP volcanic tremor in the band [0.2–0.5] Hz at Stromboli volcano, Italy.* Geophys. Res. Lett., **32**, L17303, doi:10.1029/2005GL023466.
- De Lauro E., De Martino S., Falanga M. and Palo M.; 2006: *Statistical analysis of Stromboli VLP tremor in the band [0.1–0.5] Hz: some consequences for vibrating structures.* Nonlin. Processes Geophys., **13**, 393-400.
- De Lauro E., De Martino S., Falanga M. and Palo M.; 2009a: *Modelling the macroscopic behavior of Strombolian explosions at Erebus volcano.* Physics of the Earth and Planetary Interiors, **176**, 174-186.
- De Lauro E., De Martino S., Falanga M. and Palo M.; 2009b: *Decomposition of high-frequency seismic wavefield of the Strombolian-like explosions at Erebus volcano by independent component analysis.* Geophys. J. Int., **177**, 1399-1406.
- De Martino S., Falanga M., Scarpa R. and Godano C.; 2005: *Very long period volcanic tremor at Stromboli, Italy.* Bull. Seism. Soc. Am., **95**, 1186–1192.
- Falsaperla S., Privitera E., Chouet B. and Dawson P.; 2002: *Analysis of long-period events recorded at Mount Etna (Italy) in 1992, and their relationship to eruptive activity.* J. Volcanol. Geotherm. Res., **114**, 419-440.
- Harris A.J.L., Carniel R. and Jones J.; 2005: *Identification of variable convective regimes at Erta Ale Lava Lake.* J. Volcanol. Geotherm. Res., **142**, 207-223.
- Hasegawa H.; 2009: *Population rate codes carried by mean, fluctuation and synchrony of neuronal firings.* Physica A: Statistical Mechanics and its Applications, **388**, 499-513.
- Hyvärinen A. and Oja E.; 1997: *A Fast Fixed-Point Algorithm for Independent Component Analysis.* Neural Computation, **9**, 1483-1492.
- Hyvärinen A., Karhunen J. and Oja E.; 2001: *Independent Component Analysis.* John Wiley & Sons Inc., New York, 481 pp.
- Jones J., Carniel R., Harris A.J.L. and Malone S.; 2006: *Seismic characteristics of variable convection at Erta 'Ale lava lake, Ethiopia.* J. Volcanol. Geotherm. Res., **153**, 64-79.
- Kadota T. and Labianca F.; 1981: *Gravity-wave induced pressure fluctuations in the deep ocean.* IEEE J. Oceanic, OE-6 (2), 50-58.
- Koldovský Z., Tichavský P. and Oja E.; 2006: *Efficient variant of algorithm FastICA for Independent Component Analysis attaining the Cramér-Rao Lower Bound.* IEEE Trans. Neural Netw., **17**, 1265- 1277.
- Ladisa M., Lamura A., Nico G. and Siliqi D.; 2005: *The independent component analysis as a new tool to determine solvent content in protein crystals.* Physica A: Statistical Mechanics and its Applications, **349**, 571-581.
- Lombardo G., Coco G., Corrao M. and Gresta S.; 1996: *Features of seismic events and volcanic tremor during the preliminary stages of the 1991-1993 eruption of Mt. Etna.* Annali di Geofisica, **2**, 403-410.
- Monna S., Frugoni F., Montuori C., Beranzoli L. and Favali P.; 2004: *Valutazione della qualità del segnale sismico registrato dall'osservatorio geofisico marino SN-1.* In: Slejko D. and Rebez A. (eds), GNGTS 2004 – 23° Convegno Nazionale, Riassunti Estesi delle Comunicazioni, Tipografia Masetti, Trieste, pp. 21-23.
- Neuberg J.; 2000: *External modulation of volcanic activity.* Geophys. J. Int., **142**, 232-240.
- Peterson J.R.; 1993: *Observations and modelling of background seismic noise.* Open-file report 93-322, U.S. Geological Survey, Albuquerque, New Mexico, 94 pp.
- Rufiner H.L., Goddard J., Rocha L.F. and Torres M.E.; 2006: *Statistical method for sparse coding of speech including a linear predictive model.* Physica A: Statistical Mechanics and its Applications, **367**, 231-251.

- Tanaka T. and Cichocki A.; 2004: *Subband decomposition independent component analysis and new performance criteria*. In: Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'04), Montreal, Canada, pp. 541–544.
- Tichavský P., Koldovský Z. and Oja E.; 2006: *Performance analysis of the FastICA algorithm and Cramér-Rao Bounds for Linear Independent Component Analysis*. IEEE Trans. on Signal Processing, **54**, 1189-1203.
- Tichavský P., Koldovský Z. and Oja E.; 2008: *Corrections to 'Performance analysis of the FastICA algorithm and Cramer-Rao Bounds for Linear Independent Component Analysis' TSP 04/06*. IEEE Trans. on Signal Processing, **56**, 1715-1716.
- Tong L., Inouye Y. and Liu R.; 1993: *Waveform-preserving blind estimation of multiple independent sources*. IEEE Trans. on Signal Processing, **41**, 2461-2470.
- Wassermann J.; 1997: *Untersuchung seismischer Signale vulkanischen Ursprungs anhand von Breitband - Arrayregistrierungen an den Vulkanen Etna und Stromboli*. Berichte des Institutes fuer Geophysik der Universitaet Stuttgart, Ph.D. Thesis, 169 pp. in German.
- Webb S.C.; 1998: *Broadband seismology and noise under the ocean* Rev. Geophys., **36**, 105–142.

Corresponding author: Giuseppe Cabras
University of Udine, Department of Physics
Via delle Scienze 208, 33100 Udine, Italy
phone: +39 0432 558212; fax: +39 0432 558222; e-mail: giuseppe.cabras@uniud.it