

Acoustic Emission spectra classification from rock samples of Etna basalt in deformation-decompression laboratory experiments

De Rubeis V.(1), Vinciguerra S.(1), Tosi P.(1), Sbarra P.1) and Benson P.M.(2)(3)

(1) Istituto Nazionale di Geofisica e Vulcanologia, Roma, Italy.

(2) Lassonde Institute, University of Toronto, Toronto, Canada.

(3) Department of Earth Sciences, University College London, London, UK

Abstract

Recent laboratory experiments on Etna basalt have permitted the generation of an extensive catalogue of acoustic emissions (AE) during two key experimental phases. Firstly, AE have been generated during triaxial compressional tests and formation of a complex fracture/damage zone. Secondly, rapid fluid decompression through the damage/shear zone after failure. We report new results from an advanced analysis method using AE spectrograms, allowing us to qualitatively identify high and low frequency events; essentially comparable to seismicity in volcanic areas. Our analysis, for the first time, quantitatively classifies ‘families’ of AE events belonging to the same experimental stage without prior knowledge. We then test the method using the AE catalogue for verification, which is not possible with field data.

FFT spectra, obtained from AE, are subdivided into equal log intervals for which a local slope is calculated. Factor analysis has been then applied, in which we use a data matrix of columns representing the variables considered (frequency data averaged in bins) vs. rows indicating each AE data set. Factor analysis shows that the method is very effective and suitable for reducing data complexity, allowing distinct factors to be obtained.

We conclude that most of the data variance (information content) can be well represented by three factors only, each one representing a well defined frequency range. Through the factor scores it is possible to represent data in a lower dimension factor space. Classification is then possible by identifying clusters of AE belonging to the same experimental stage. This allows us to propose a deformation/decompression interpretation based solely on the AE frequency analysis and to identify a third type of AE related to fluid movements in the deformation stage.

Introduction

An acoustic emission (AE) is defined as a transient elastic wave generated by the rapid release of energy within a material. In the Earth Sciences, studies of AE and seismology show a significant overlap. Both approaches deal with the radiation of elastic waves, although at different scales and frequencies. Generally, AE recorded in the laboratory are generated by flaws at the grain size scale with source dimensions between micron and millimeter scale and frequency ranges between 100 kHz – 2 MHz (Lockner, 1993).

In the brittle regime, the damage evolution in a rock under load involves the growth of microcracks from stress concentrators such as voids, inclusions and grain contacts, resulting in both inelastic strain and acoustic emissions. The acoustic signals that are spontaneously generated from the microcracking provide information about the size, location and deformation mechanisms of the events as well as properties of the medium through which the acoustic wave travel (e.g. velocity, attenuation and scattering). Importantly, laboratory rock fracture and earthquake rupture are processes obeying similar statistics for source dimensions over more than eight orders of magnitude (e.g. Hanks, 1992; Zang et al., 1998).

The main goals of AE studies so far aimed to:

- (a) Characterize individual AE events in terms of their frequency content, amplitudes and durations so that they can be related to the micro-mechanisms that produce them. A temporal correlation between the onset of AE and dilation in samples under loading was found, showing that dilation was caused by pervasive microcracking, primarily oriented parallel to the maximum compressive principal stress (Brace, 1966; Scholz, 1968). AE amplitudes and frequency have been observed to increase before failure (Zang et al., 1998 and references therein)
- (b) Locate the source of AE events in 3D in order to image the localization of failure processes. Insights have been provided into the nucleation phase of fracture, using AE source locations to map the temporal and spatial evolution of fracture. Two different processes operate, a process zone in front of the fault tip and a damage zone following the process zone (Lockner et al., 1992; Lei et al., 2000).
- (c) Analyze the statistics of recorded events to gain insights into the deformation processes and their rates. The Gutenberg-Richter relationship between frequency and magnitude of earthquakes also applies to experimental rock failure (Meredith et al., 1990; Sammonds et al., 1992; Ponomarev et al., 1996; Lei et al., 2003)
- (d) Study AE in the presence of fluid. The influence of fluid pressure on deformation has been investigated in drained conditions at constant pore pressure. A weakening effect of water in the brittle faulting and cataclastic flow regime has been found, because of reduction of both specific surface energy and friction coefficient (Baud and Meredith, 1997; Baud et al., 2000).

New directions are focussing on the fundamental similarity of the physical processes involved in generating different frequency content seismic signals, such as seismic tremor under the subduction zones in tectonic areas and volcanic tremor and long-period events in volcanic areas. To date, comparisons have been made to field scale seismicity by comparing the spectral character and frequency/size scaling of acoustic emissions (e.g. Burlini et al., 2007, 2008; Benson et al., 2008). Frequency scaling offers the strongest argument to assess the equivalence of the physical processes between laboratory experiments and natural volcanic seismic signals. For example, experimental low frequency events and tremor have frequency of about 100 kHz for intrusion lengths ranging between few tens to 100 mm. In natural earthquakes, dominant frequencies around 1-2 Hz are associated with fracture lengths of some hundreds meters to 1 km. Considering that dominant frequencies of earthquakes scale inversely with source dimension (Aki and Richards, 1980), one may write $d_1 \times f_1 = d_2 \times f_2$, where d_1 , d_2 and f_1 , f_2 are the dimension and frequency of laboratory (1) and nature (2), respectively. Comparing laboratory data with typical frequency (1-2 Hz) and size (1 km) of low frequency earthquakes, we obtain $d_1/d_2 = 5 \times 10^6$ and $f_2/f_1 = 2.5-5 \times 10^6$, which indicates excellent agreement between laboratory information and natural cases (Burlini et al., 2007;2009; Benson et al., 2008). Here, we present a new and alternative analysis, aiming to quantitatively classify ‘families’ of AE through the analysis of spectral ratios and associated factors, without a-priori knowledge of the spatio-temporal evolution.

The data set

Samples of Etna Basalt, Italy were deformed in using a triaxial cell installed at the University of Toronto (Fig. 1). Two experimental stages are performed, designed to separate the signals generated from deformation (stage 1) from those specific to fluid movement (stage 2), as described in Benson *et al.*, (2008). Irrespective of the stage of experiment (deformation or decompression), events are located on or in close proximity to the fault (Benson et al., 2008). During the formation of the fault, waveforms show a distinctive, high frequency character (600-900 kHz), exhibiting a rapid onset (i.e. rapid acceleration) and relatively short duration (400 us). These features are well known and understood from Rock Physics studies on brittle processes in rock fracture. During the decompression stage, however, waveforms exhibit a much lower frequency component (typically

80-100 kHz). The onset of these waveforms are highly emergent, and with longer duration, of the order of ms. In the following analysis we refer to these characteristic signals as high frequency (HF) and low frequency (LF) respectively.

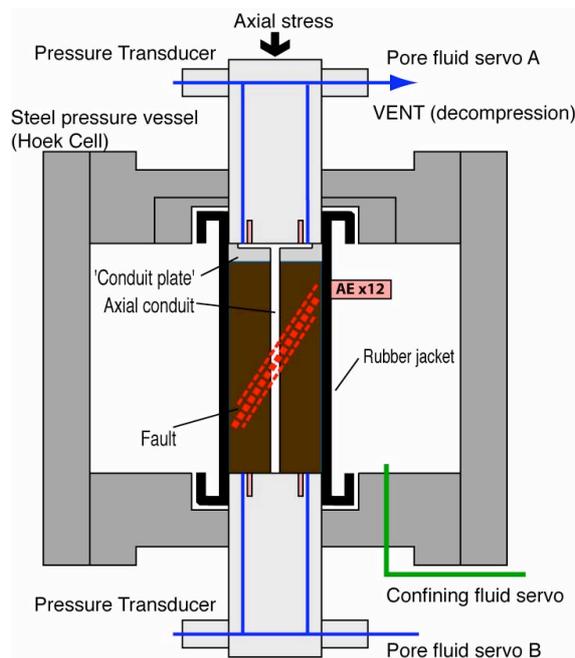


Figure 1. Scheme of the experimental setup (after Benson et al., 2008). Note that, for clarity, the precise locations of the 16 AE sensors (12 lateral and 4 axial) have been omitted.

Method

Input data consist of AE FFT spectra. The aim of the analysis is to statistically apply a classification method that is able to recognize different AE and their relation within each of the dynamic parts of the experiment. A key point is to classify the set of AE solely through spectral analysis, and then to test such classifications with other experimental information as occurrence time, load history, AE spatial location, etc. We look for an AE classification showing a consistent interpretation. If the interpretation is in agreement with the other experiment variables not involved, then we interpret this to have yielded two methodological useful results: we receive an independent confirmation of the reliability of classification method, and can elucidate the dynamical process.

Our classification is based on the application of factor analysis (Davis, 2002 and reference therein). It consists of a statistical multivariate approach useful to highlight relations among variables. Let us define the dataset as a matrix $[X]$, in which rows are the samples and columns are variables ($n \times m$). The statistical assumption of the factor analysis is that the m variables, of the original dataset, derives from a smaller number of p independent factors ($p < m$). The principal component approach to factor analysis consists of the following steps:

- the original data matrix $[X]$ is standardized in respect of the variables simply subtracting the mean and dividing each data by the standard deviation of the corresponding variable, thus obtaining $[X_{st}]$;
- a variance-covariance (correlation) matrix $[s^2]$ is extracted from $[X_{st}]$ (note that being $[s^2]$ extracted from standardized data, it is a correlation matrix). It is symmetric and all diagonal elements are equal to unit);

- eigenvectors and eigenvalues are extracted from $[s^2]$; the eigenvectors define the matrix $[U]$, the square root of eigenvalues gives a diagonal matrix of singular values $[\lambda]$; the sum of the eigenvalues represents the total variance (information content) of data;
- eigenvectors are converted into factors using this relation: $[A^R] = [U] \cdot [\lambda]$, where $[A^R]$ is an $(m \times m)$ -sized matrix, each column representing a factor whose elements are referred to as factor loadings.

Due to the precise manner in which the factors are calculated, each factor loading is weighted proportionally to the square root of the amount of variance contributed by that variable to the factor. Looking at the proportion of variance accounted for by each factor, it is possible to select a smaller number of factors (p) accounting for most of the total variance in the dataset. The independence of all factors is guaranteed by orthogonal decomposition of the correlation matrix, given by the Eckart–Young theorem, through the extraction of eigenvalues and eigenvectors. Therefore, we can reduce the size of $[A^R]$ to $(m \times p)$.

Factor analysis results in factor scores $[\hat{S}^R]$, which represent the values of any single factor for every observation, as follows: $[\hat{S}^R] = [X_{st}] \cdot [A^R] \cdot ([A^R]^T \cdot [A^R])^{-1}$.

In this way it is possible to associate with each observation the value of a factor score related to each one of the p factors. Summarizing the benefits of factor analysis approach, we have a statistical tool able to reduce the complexity of a multi-variable data set into few factors, which they retain the biggest amount of variance. The key to perform such process is inside the idea that many original variables are redundant, having a certain degree of mutual correlation. Moreover by the factor model application we can consider each original variable X_j as constitute by the composition of p factors f , following a specific loading a , plus a random variation ε . If we assume that p is less than the number of original variables we have reduced the complexity of the problems and separate a small amount of system information, which we consider random noise:

$$X_j = \sum_{r=1}^p a_{jr} f_r + \varepsilon_j.$$

The specific application of factor analysis to AE spectra consists in the conversion of the data into a $[X]$ matrix: the logarithmic amplitude of each spectrum is standardized in order to remove any effect due to AE magnitude. The frequency range is then subdivided into equal log intervals and mean amplitude for each interval is calculated. The columns of $[X]$ matrix thus represent the binned frequency variables. The rows are each AE as recorded during the loading experiment. The assumption is that near frequencies may have the same behaviour, quantified by their mutual correlation. The unknown is the extent of the similar-correlated behaviour of contiguous frequencies. The method is to create frequencies variables of short intervals in order to remove noise. Then, by the application of factor analysis, the frequencies variables are grouped together in function of their mutual correlation and true independent groups are extracted. They represent the true dimension of the process and give the basis for a useful data classification.

Results and discussion

The data matrix is constituted by 280 FFT spectra calculated on AE waveforms in the frequency range 29 KHz - 493 KHz. The logarithmic amplitude of each spectrum is standardized in order to remove any effect due to AE magnitude. The range is then subdivided into 10 equal log intervals and mean amplitude for each interval is calculated. Boundaries (in KHz) of each of the ten variables defined for the analysis are shown in table 1.

variable	Lower freq. limit (KHz)	Upper freq. limit (KHz)
1	29.52	39.13
2	39.13	51.85
3	51.85	68.72
4	68.72	91.07
5	91.07	120.69
6	120.69	159.95
7	159.95	211.97
8	211.97	280.91
9	280.91	372.27
10	372.27	493.35

table 1

Figure 2 shows the variance percentage expressed by each factor, the first 3 factors account of about 74% of total data variance: this value is satisfactory, ensuring a proper representation of the whole data set.

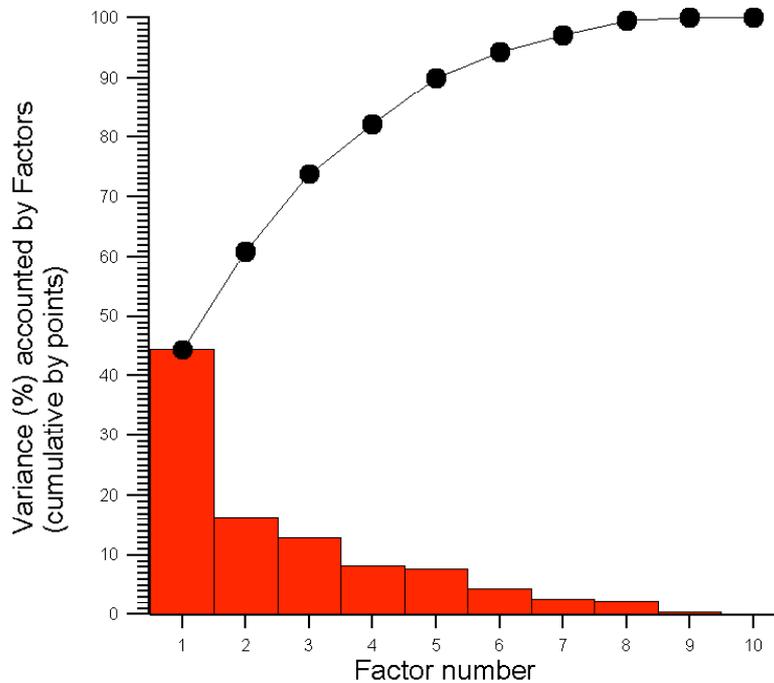


Figure 2. Individual and cumulative variance of original data accounted by factors.

Figure 3 shows the factor loadings values for the three factors (respectively first factor, second, and third). Factor loadings express the correlation of each original variable – in this study the averaged frequency range – to each factor, which represents the new variable.

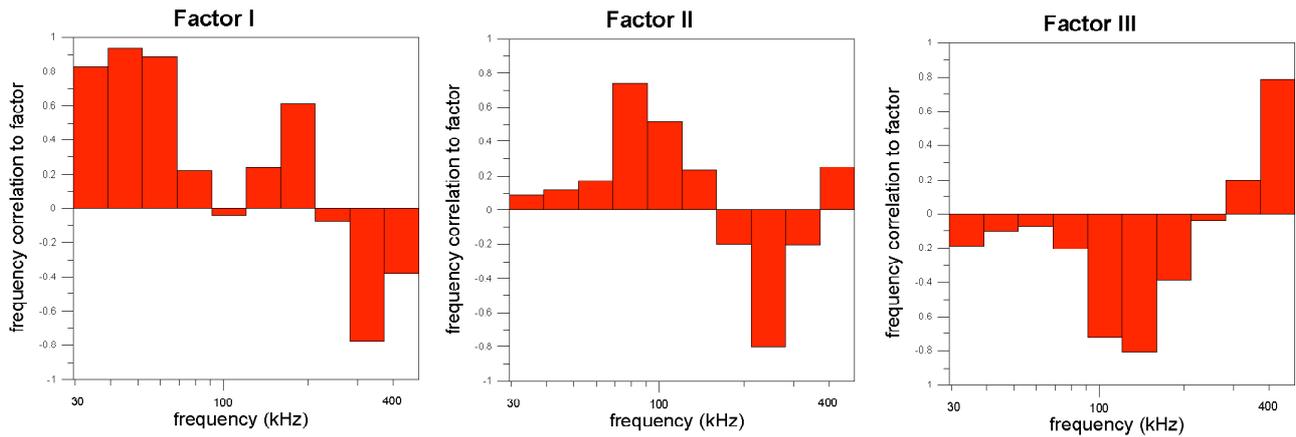


Figure 3. Factor loadings values for the first three factors as extracted by factor analysis. See text for comments on ability of factors to represent original AE frequency bins.

The first factor shows a positive correlation of the low frequency range up to 70 kHz, while a negative correlation is found related to the highest frequency range (280-480 KHz). A moderate positive correlation occurs in the frequency interval 150-200 KHz (for the precise frequency interval see table 1). The second factor shows also a positive correlation in the frequency range 60-120 KHz interval, while a negative correlation is found in the 200-370 KHz interval.

Third factor is negatively correlated in the 90-200 KHz interval and positively in the highest frequencies interval 370-500 KHz.

These factors represent different groups of interrelated frequency ranges. Positive and negative correlations allow interpreting the relationship among the different frequencies. The first factor indicates that, when we have an AE where lower frequencies (up to 70 KHz) prevail, higher frequencies (280-480 KHz) tend to reduce, and vice versa. Second factor shows similar trends, but for the frequencies range of 60-120KHz and 200-370 KHz respectively. Finally, the third factor mirrors the second, but correlations are shifted towards higher frequency ranges (90-200 KHz and 370-500 KHz).

Through the factor scores, it is hence possible to represent data in a lower dimension factor space and to see the similar behaviour of each AE associated in specific factor space portion or clustered. Since the first factor is the most informative (figure 2), we show two scatter diagrams representing original AE spectra as defined respectively by I-II factors and I-III factors (Figures 4a, 4b).

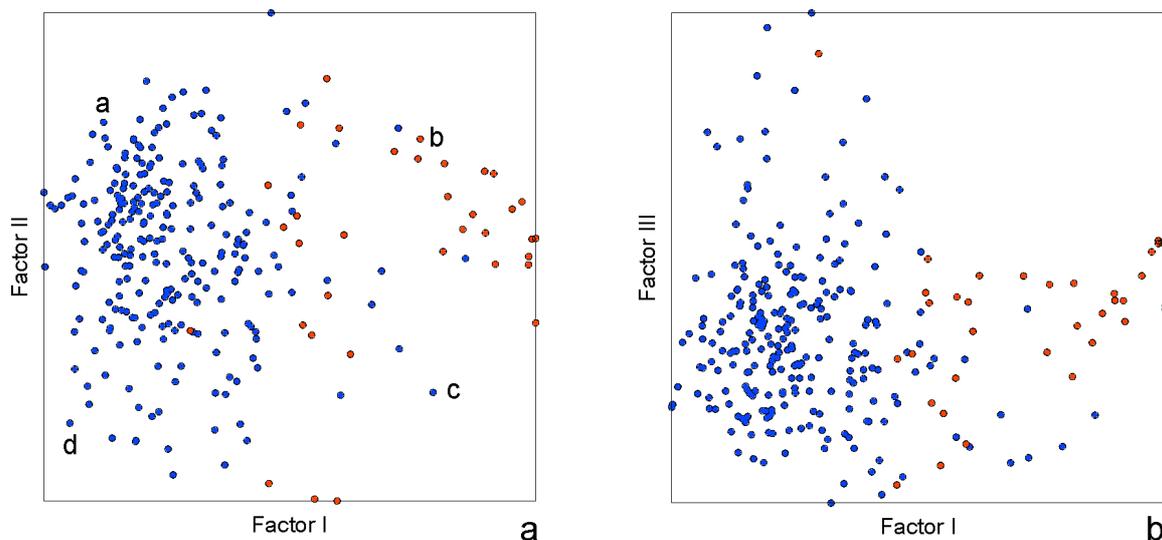


Figure 4. All 280 FFT spectra represented on the factor space following factor I with factor II scores (4a) and factor I with factor III scores (4b). In blue colour AE spectra recorded during deformation phase, in red those recorded during rapid decompression of the pore fluid. Four spectra (a,b,c,d) are marked (4a) in order to interpret their factor space position under their spectral content. Compare with Figure 5 for complete interpretation.

In Figure 4(a) and 4(b), all 280 AE samples are represented in the factor space. Each factor holds the frequency spectral content discussed above (Figure 3). By construction factors are orthogonal, which it means they have no correlation: in fact points are distributed with zero linear correlation between the two factors. In order to achieve a full understanding of the inter-factor relationship, we selected (a, b, c, d in figure 4a) four binned spectra, which represent the end terms in the factor space, i.e. the frequency ranges analyzed.

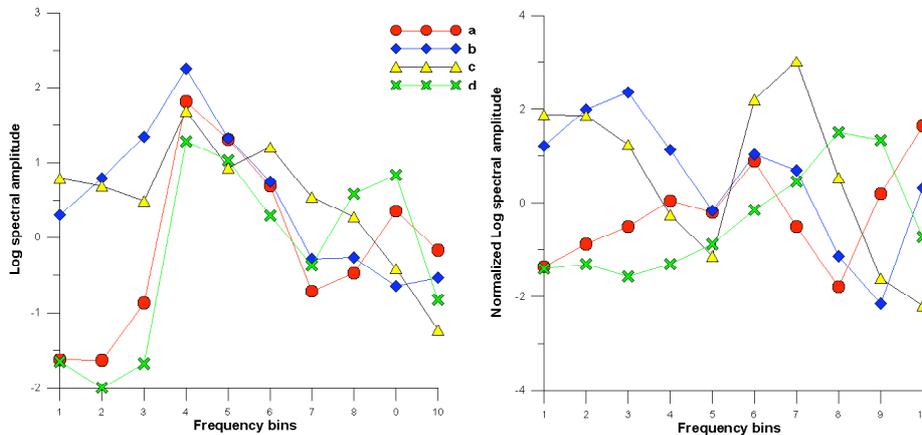


Figure 5. Four binned log amplitude spectra from AE sample data (*a, b, c, d* samples as referred into Figure 4a) as original values (left) and normalized values (right). Normalized spectra are represented to follow variables standardization applied to AE data matrix prior the extraction of eigenvalues and eigenvectors, as operated by factor analysis. See text for further details.

In detail, AE spectra *a-d* and *c-b* have similar factor I values while spectra *a-b* and *c-d* have similar factor II values. This similarity can be investigated by plotting the log of spectral amplitude vs. the frequency bins (Figure 5). AE spectra *a-d* and AE spectra *c-b* show correspondent amplitudes (bins 1,2,3,7,9) in spectral frequencies represented by factor I. AE spectra *a-d* hold low amplitudes in frequency bins 1,2,3,7 (positive correlation) and high amplitudes in bin 9 (negative correlation). AE spectra *c-b* show an opposite behaviour. Analogue considerations apply for the factor II.

Factor analysis provides a quantitative tool to systematically discriminate AE events generated during deformation phase of the experiment (blue dots in Figure 4) from those recorded during rapid decompression of the pore fluid (red dots in Figure 4). High (HF) and low frequency (LF) events fill specific sectors in the factors space. In detail the LF events reflect the pure origin of fluid driven events, due to the pore fluid flashing onto the damage/shear zone. No reactivation of cracks in terms of shearing/propagation of existing microfractures is triggered from the fluid propagation onto the fault plane, as no HF components are found from the factor analysis. HF shows a wider range of amplitudes involved. This can be explained from the fact that the energy and size of HF events increases over time as micron to mm scale cracks are involved in the early stage of deformation, followed by cm scale faults formation at the failure.

The most intriguing observation is given from the position of several events generated during deformation phase (where HF events only were believed to occur) in the factor space domain of the LF events (blue dots in Figure 4). This suggests that these events were generated from analogous mechanisms, which led to the formation of the LF, i.e. fluid migration onto the faults. We interpret this behaviour due to the fact that during loading cracks open and shear. At the same time pressurized fluids rapidly transit into the crack surfaces and generate same LF signals, similarly to those generated from pure fluid decompression in stage 2 (red dots in Figure 4).

Taken together this emphasizes as factor analysis applied to a given AE data set can allow to highlight bulk AE patterns, that otherwise can not be easily seen by a pure visualization of single events.

Conclusions

We have statistically analysed AE spectra generated during deformation and failure of an Etna basalt sample (stage 1) and the subsequent rapid fluid decompression of the pore pressure (phase 2). A factor analysis has been applied on 280 FFT spectra calculated on AE waveforms in the frequency range 29 KHz - 493 KHz. The first 3 factors account of about 74% of total data variance. Through the factor values we reported data in a lower dimension factor space and we analysed the similar behaviour of each AE associated in specific factor space portion or clustered.

Factor analysis provides a quantitative tool to systematically discriminate between HF and LF events generated during the two experimental stages (deformation and fluid movement), which essentially acts as to produce a higher proportion of HF and LF AE respectively. Importantly, several events generated during the deformation phase occur in the factor space domain of the LF events. This suggests the existence of LF events forming during phase 1, due to the fluid migration into forming and propagating faults.

References

- Aki, K., and Richards, P.G., 1980, *Quantitative Seismology: Theory and Methods*, Freeman & co., San Francisco (USA). 932 p.
- Baud, P., and P. B. Meredith (1997), Damage accumulation during triaxial creep of Darley Dale sandstone from pore volumetry and acoustic emission, *Int. J. Rock Mech. Min. Sci.*, 34, Pap. 024.
- Baud, P., W. L. Zhu, and T.-F. Wong (2000), Failure mode and weakening effect of water on sandstone, *J. Geophys. Res.*, 105, 16,371– 16,389.
- Benson, P. M., B. D. Thompson, P. G. Meredith, S. Vinciguerra, and R. P. Young (2007), Imaging slow failure in triaxially deformed Etna basalt using 3D acoustic-emission location and X-ray computed tomography, *Geophys. Res. Lett.*, 34, L03303, doi:10.1029/2006GL028721.
- Benson P. M., Vinciguerra, S. Meredith, P. G. and Young, R. P. (2008) Laboratory simulation of volcano seismicity, *Science*, 322, 249, doi: 10.1126/science.1161927.
- Brace, W.F., Paulding, B.W., and Scholz, C. (1966), Dilatancy in the fracture of crystalline rocks, *J. Geophys. Res.* 71, 3939–3953.
- Burlini L., Vinciguerra S., Di Toro G., De Natale G., Burg J.P., Seismicity preceding volcanic eruption: new experimental insight, *Geology*, 35,2,183-186; doi: 10.1130/G23195A, 2007
- Burlini L., Di Toro G., Meredith P.G., Seismic tremor under the subduction zones: the rock-physics interpretation. *GRL*, VOL. 36, L08305, doi:10.1029/2009GL037735, 2009.
- Davis, J.C. (2002). *Statistics and Data Analysis in Geology*, third edition, Wiley and Sons, Inc., New York.
- Hanks, T.C., 1992, Small earthquakes, tectonic forces: *Science*, v. 256, p. 1430-1432.

- Lei, X.-L., Kusunose, K., Rao, M.V., Nishizawa, O., and Satoh, T. (2000), Quasi-static fault growth and cracking in homogenous brittle rock under triaxial compression using acoustic emission monitoring, *J. Geophys. Res.* 105, 6127–6139.
- Lei, X.-L., Kusunose, K., Satoh, T., and Nishizawa, O. (2003), The hierarchical rupture process of a fault: An experimental study, *Phys. Earth Planet. Inter.* 137, 213–228.
- Lockner, D. A., Byerlee, J. D., Kuksenko, V., Ponomarev, A., and Sidorin, A., *Fault mechanics and transport properties of rocks* (eds. Evans, B., and Wong, T-F) (Academic, London 1992).
- Lockner D., The role of acoustic emission in the study of rock fracture, *IJRM*, 30, 7, 883-899, 1993.
- Meredith, P.G., Main, I.G., and Jones, C., 1990, Temporal variations in seismicity during quasi-static and dynamic rock failure: *Tectonophysics*, v. 175, p. 249-268.
- Ponomarev, A.V., Zavyalov, A.D., Smirnov, V.B., and Lockner, D.A., 1997, Physical modelling of the formation and evolution of seismically fault zones: *Tectonophysics* v. 277, p. 57-81.
- Sammonds, P.R., Meredith, P.G., and Main I.G., 1992, Role of pore fluids in the generation of seismic precursors to shear fracture: *Nature*, v. 359, p. 228-230.
- Scholz, C. H. (1968a), Microfracturing and the inelastic deformation of rock in compression, *J. Geophys. Res.* 73, 1417–1432.
- Zang, A., C. F. Wagner, and G. Dresen (1996), Acoustic emission, microstructure, and damage model of dry and wet sandstone stressed to failure, *J. Geophys. Res.*, 101, 17,507– 17,521.