An Automatic S-phase Picking Algorithm for Volcano-tectonic Earthquakes based on Spectral Dissimilarity Analysis

Luz García, Gerardo Alguacil, Manuel Titos, Angel De la Torre, and Carmen Benítez

***Abstract*—We present an S-phase picking algorithm for volcano-tectonic earthquakes, based on analyzing the changes of frequency and amplitude expected in the plane transverse to the ray direction at S-phase arrival. A measure of these changes, called *spectral dissimilarity*, is proposed. Picking is performed in a particular waveform transformation that underlines such variations foreseen: horizontal instant power. Then, the algorithm provides a measure of its reliability, grounded on the low or high fluctuations of the picking instant obtained when applied to other horizontal components of the sismogram. Experiments are performed to test the algorithm with a challenging database of volcano seismic earthquakes from Mt. Etna, carefully picked and labeled by a human expert. The technique is compared to two other S-phase pickers existing in the literature: one based on the damped predominant period analysis, and a second one based on polarization and kurtosis rate analysis. The algorithm improves these techniques for the particular scenario of volcano-tectonic earthquakes, providing interesting results and possibilities of application.**

***Index Terms*—Automatic S-phase picking, spectral dissimilarity, instant power, picking quality, volcano-tectonic earthquakes.**

I. Introduction

Picking a phase arrival consists in detecting the time instant when its energy first arrives to a seismic station. The arrival produces a change from the background noise in amplitude, frequency contents and signal polarization. Identification of P- and S-phase arrival times in earthquake-originated waves is a powerful tool in observational seismology. It is used for identification and location of seismic events, analysis of source mechanisms, determination of their origin time, calculation of their magnitude or determination of subsurface velocity structures for example. In the case of volcano-tectonic earthquakes, phase arrival picking provides useful insights to the inside of the volcano. As the volume of seismic data has increased remarkably, many approaches to automatic phase picking have been explored in the last decades. Signal processing algorithms have been proposed to reduce noise and transform signals to domains where onset times are more easily detected. Besides offering effective processing tools, the automatization permits to sustain more easily consistent picking criteria along time for big amounts of data, alliviating the tedious task usually performed by humam experts. Early warning systems based on automatic remote sensing of seismic activity in areas of seismic risk, are also facilitated through automatic phase picking.

The first automatic phase pickers proposed [1] [2] [3] were based on detection of changes in energy related to the signal's amplitude. Allen's P-phase picker [4], based on variations of short-/long- term energy ratio (*STA/LTA*), was the pioneer of a set of techniques focused on energy changes like [5], analysis of kurtosis and skewness like [6] or [7], or rate of change of a dissipated *damping energy* [8] among others. Time domain processing presents limitations for emergent onsets and low SNR signals. Other authors have used frequency domain analysis like the multiband adaptive frequency processing of [9] [10], extraction of the damped predominant period [11] or wavelet processing like [12] [13] [7]. With a higher computational cost and good performance, Akaike Information Criterium (AIC) [14] on auto-regressive (AR) models has also been largely used for phase-picking in works like [15] [16] for P-phase picking, or [17] and [18] for P- and S-phase picking.

S-phase picking is more challenging than P-phase picking. Even if amplitudes are larger for S-waves, their onset is often weak and emergent, overlapped with the preceding P-wave coda and mixed with P to S conversions. Picking methods for S-phases use often 3 components and polarization analysis, firstly proposed by [19] and widely used [20] [21] [22] [23] or [24]. Data-driven pattern recognition like Artificial Neural Networks [25] [26] or K-Nearest Neighbours [27] has also been explored. Many authors present combinations of approaches [28] [29] [30] using kurtosis and polarization analysis for P and S-phase arrivals respectively, AIC together with STA/LTA plus polarization analysis [31], AIC plus wavelet domain processing [32], or AR modelling, polarization analysis and STA/LTA used as features for pattern recognition [25] [33] for example. As a common feature, methods are often dependent on the specific database they are applied to, and they work well with seismograms recorded with good SNR and clear phases. Few of them are applied to volcano-tectonic earthquakes [30], or microearthquakes [7]. When the earthquake source mechanism occurs in a volcano and is registered by a seismic station of the area, the shallowness, proximity of the source, lower magnitude and complex propagation path among other factors, make the S-phase detection more difficult than in the case of tectonic earthquakes.

This work proposes a novel S-phase picking technique for volcano-tectonic earthquakes focused on two items: (i) S-

Luz García, Manuel Titos, Carmen Benítez and Ángel De la Torre are with the Department of Signal Theory, Telematics and Communications of the University of Granada, Spain

Gerardo Alguacil is with the Andalusian Institute of Geophysics and Prevention of Seismic Disasters and the Department of Theoretical and Cosmos Physics of the University of Granada, Spain

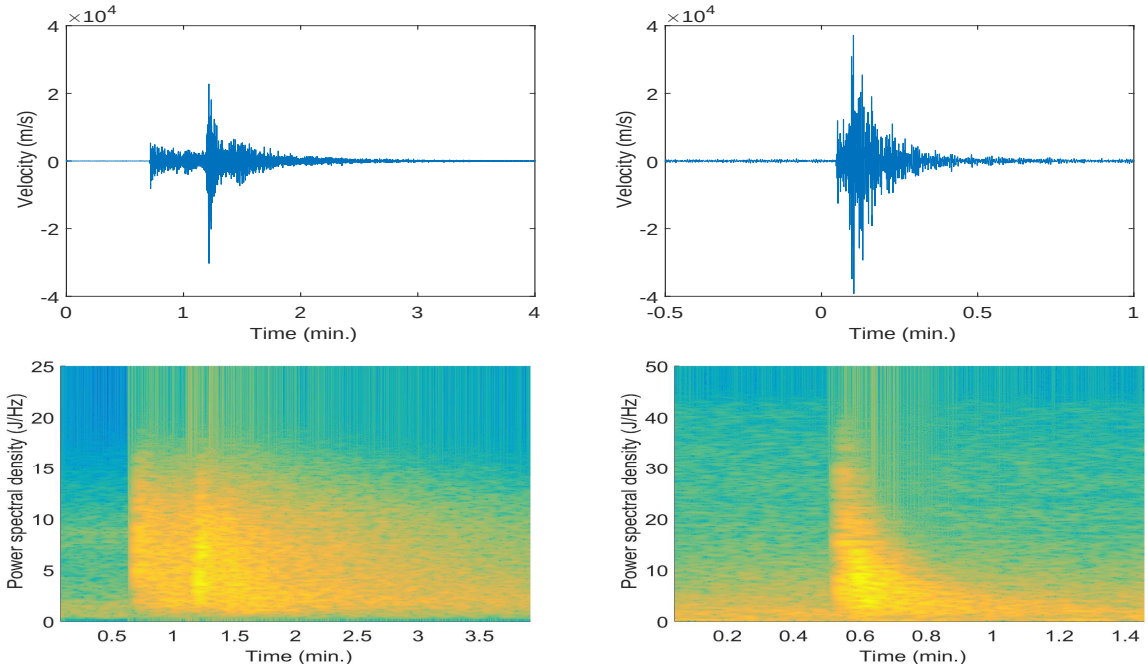

Figure 1. Waveform and spectrogram for regional (upper and lower left-side subfigures) and volcano-tectonic earthquakes (upper and lower right-side subfigures).

phase arrival will be detected searching for instant changes of energy in the frequency domain around the time when the S arrival is expected; (ii) the signal used to perform the picking is not the original ground velocity registered by the seismometer, but its specific instant power [34] used to underline these changes in frequency and energy compared to the original signal. The rest of this manuscript is organized as follows. Section II describes the S-phase and its peculiarities when originated by volcano-tectonic earthquakes. Section III presents the automatic S-phase picking algorithm proposed in this work. Section IV presents a set of experiments on a database of volcano-tectonic earthquakes registered at Mt. Etna during 2010 to test the quality of the technique proposed. Section V provides conclusions on the work.

II. Volcano-tectonic earthquakes

Earthquakes are originated when the brittle part of the Earth's crust is subjected to stress that exceeds its breaking strength [35][36]. A fracture occurs then generating a seismic elastic wave that propagates along the earth with two independent phases: a longitudinal (*compressional-dilatational*) P wave and a transverse (*shear*) S wave. P waves, involving volume change, travel faster than S waves (pure shearing) and therefore arrive ahead in the seismogram. Fluids (liquids or gases) cannot propagate S waves as they have no shear strength. Amplitudes of S waves are about five times larger than those of P waves. Both P and S waves have linear polarization in the longitudinal and transverse plane respectively. In presence of anisotropy, P linear polarization is more cleanly maintained as S waves split into fast and slow components that propagate with different velocities superimposing an elliptical polarization. Period of S waves are longer than those of P waves, at least by a factor of $\sqrt{3}$ due to the difference in wave propagation velocity and the related differences in the corner frequency of the P- and S-wave source spectrum. Additionally S waves are more strongly attenuated than P waves, filtering out higher frequencies. Picking S-phase has added difficulties. Its onset is superposed by the P-wave coda. Also S-wave splitting and the presence of Sp-converted precursors might happen.

Volcano-tectonic earthquakes are earthquakes induced by changes in pressure produced by injection or withdrawal of magma that generate fractures in the rocks surrounding the magmatic chamber of a volcano. This source mechanism also generates a seismic wave with P and S phases, that present several differences compared to those generated by tectonic earthquakes [36] [37]:

- Other simultaneous seismic events related to liquid and/or gas-solid processes take place in the volcanic scenario. Tremors, long-period events, or surface effects such as rockfalls, landslides or pyroclastic density flows might happen simultaneously, overlapping and therefore *contaminating* the seismic register.
- Volcanic regions present often a rough topography with changing propagation and site properties.
- Sismo-volcanic sources are often shallower compared to tectonic ones. That implies that instead of clean impulsive onsets, their P and S phases are often emergent and their frequencies shift to a lower bandwidth. The shallower hypocenter location implies a larger amount of scattering during wave propagation, especially of higher frequencies. In some cases, S phase arrival is not detectable.
- Near-source field effects and near surface propagation effects complicate the polarization analysis of the particle

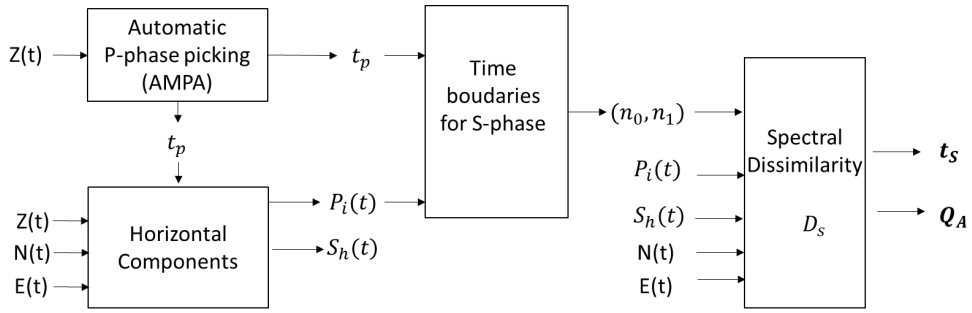

Figure 3. Flowchart of the algorithm proposed and described along Section III.

motion.

Figure 1 shows the waveform and spectrogram for two example regional and volcano-tectonic earthquakes on which these differences can be seen. Far distances from hypocenters for regional earthquakes, produce separations between P and S phase arrivals long enough to ease their picking. High frequencies are attenuated by the long distance trip of the waveform, providing therefore regional earthquakes with a better polarization analysis if needed.

III. Proposed Method: Simultaneous Frequency and Energy Variations

As Section II describes, once P-phase has arrived, S-phase arrival will imply a change in the wave frequency and energy, specially in the plane perpendicular to the ray direction. S-phase frequency is lower than that of the previous P, while its energy is higher due to the shear mechanism of the earthquake source. This change in frequency and energy can be seen in Figure 2. The figure compares the average power spectral density (PSD) envelope for a window of signal of half second starting: (i) at the instant of arrival of P-phase (t_p) in the vertical component Z (blue line) and, (ii) at the instant of arrival of S-phase (t_s) in the horizontal component of the transverse plane S_h (green line). Signals used to average the PSDs depicted belong to the database of Mt. Etna Volcano (fully described in Section IV-A) provided by the Instituto Nationale di Geofisica e Vulcanologia, Section of Catania, Italy [41]. Arrival times for P- and S- phases used in the figure were manually picked by an expert geophysicist.

Based on such change of frequency and energy produced with the S- phase arrivals, we propose a *spectral dissimilarity* measure (D_s) to compare consecutive frames of signal in the time interval when S-phase is expected, setting the arrival time when this D_s reaches a maximum. The method follows four steps depicted in Figure 3:

s1. Horizontal components on which S-phase is searched for are identified.
s2. Based on the automatic detection of the P-phase arrival, a time interval for the potential arrival of the S-phase is defined.
s3. The instant corresponding to the maximum spectral dissimilarity (D_s) in that time interval, is chosen as S-phase picking time t_s.

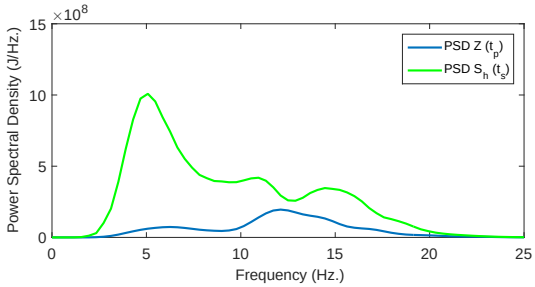

Figure 2. Average Power Spectral Density for a window of 0.5 seconds around the P- (blue) and S- (green) phase arrivals in vertical and horizontal seismic registers. Values averaged for the set of 300 volcano-tectonic earthquakes from Mt. Etna used in this work.

s4. An automatic measure of the quality of the S-phase picking is provided.

A. Horizontal components

According to linear theory of wave propagation in homogeneous media, P and S waves are independent solutions of the equation of motion which result in linear particle motion parallel to the direction of wave propagation for P-phase , and transverse to the direction of wave propagation for S-phase. The following horizontal components will be used in the algorithm presented:

i. Velocity horizontal components of the seismogram: N(t) and E(t).
ii. Based on the definition of *damage index* provided by Nakamura [34] [38] [39], we can calculate the *instant power per unit of mass* ($P_i(t)$) of the horizontal components N(t) and E(t). Given the definition of power (P) as work (W) per unit of time:

$$P = \frac{dW}{dt} = \frac{\overrightarrow{F} \cdot \overrightarrow{dr}}{dt} = \overrightarrow{F} \cdot \overrightarrow{v} \tag{1}$$

The power per unit of mass (m) is related to acceleration ($\overrightarrow{a}$) and velocity ($\overrightarrow{v}$) as:

$$\frac{P}{m} = \overrightarrow{a} \cdot \overrightarrow{v}. \tag{2}$$

Using the horizontal velocity components N(t) and E(t), $P_i(t)$ can be obtained as:

$$P_i(t) = \frac{P}{m} = N(t) \cdot \frac{dN(t)}{dt} + E(t) \cdot \frac{dE(t)}{dt} \tag{3}$$

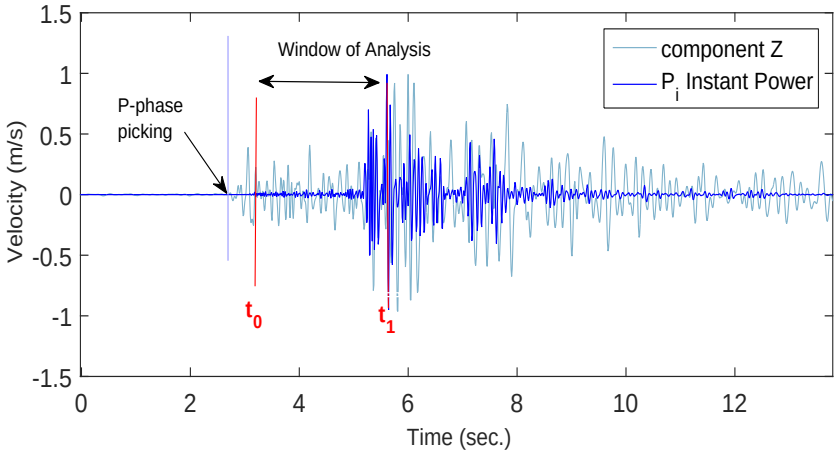

Figure 4. Boundaries delimitation for the window of analysis to search for the S-phase arrival. t_0 is set 0.5 seconds after picking the P-phse. t_1 is set in the time instant when instant power P_i reaches its maximum. Z and P_i amplitudes are normalized to one for descriptive purposes in the figure.

$P_i(t)$, used to measure earthquake damages produced mainly by its S-phase, expresses the distribution in time of the seismic intensity and the propagation of the earthquake motion. It emphasizes the changes of frequency and energy that occur with the S-phase arrival. It will be used as main horizontal component to detect the S-phase arrival using the spectral dissimilarity measure proposed in this work.

iii P-phase is automatically picked in the vertical trace Z using the algorithm AMPA [9] based on multiband adaptive processing and noise removal. Using the P-phase arrival and the classical polarization analysis [19], the direction of wave propagation S_r, backazimut and incidence angles are obtained. Original components Z, N, and E are rotated into components L, Q and T, corresponding respectively to the ray propagation direction S_r, and transverse plane components S_v and S_h. S_h will be used for *quality-check* phase picking, avoiding Sp precursors present in S_v that could be misinterpreted as too early onset of S-phase [40].

With the aim of searching for the biggest change in energy and frequency as symptom of S-phase arrival, the onset time will be picked measuring spectral dissimilarity in $P_i(t)$. Spectral dissimilarity in $S_h(t)$, N(t), and E(t) will serve as *quality-check* to validate the picking in original and rotated components of the transverse plane. As a result of the validation, a measure of the reliability of the picking will be given.

B. Boundaries of the window of analysis

S-phase arrival takes place in the time interval (t_0, t_1), starting after the P-phase picking time t_p (e.g. $t_0 = t_p$ + 0.5 seconds) and ending before $P_i(t)$ reaches its maximum value, instant noted as t_1, when S-phase has certainly already arrived. Figure 4 depicts Z and P_i components normalized to their maximum amplitudes, showing an example of delimitation of the window of analysis. P-phase arrival (t_p) is automatically obtained using the automatic P-phase picking algorithm AMPA as described in the former subsection.

C. Spectral Dissimilarity

Because we are searching for the change of frequency and energy expected when S-phase arrives (see Figure 3), we propose a spectral dissimilarity measure (D_s) between overlapped consecutive frames of samples within the window of analysis (t_0,t_1):

i. Each sample n from the interval of analysis (t_0,t_1) is framed into a window of L samples, (x_{orig}), containing the $\frac{L}{2}-1$ samples prior to it (x_{orig-l}), and the $\frac{L}{2}-1$ samples posterior to it (x_{orig-r}). Then the frame of signal x_{orig} thus created, is weighted by a triangular window $w_t(n)$ to underline the analysis on its central sample, maximizing the left/right side dissimilarities around it and minimizing the contribution of the more distant ones.

$$w_t(n) = \begin{cases} \frac{n}{W_0} & \text{if } n \leq \frac{L}{2} \\ \\ \frac{L-n}{W_0} & \text{if } \frac{L}{2}+1 \leq n \leq L \end{cases} \quad (4)$$

Left and right sub-windows $x_l(n)$ and $x_r(n)$ are calculated as:

$$x_l(n) = x_{orig-l}(n) \cdot w_t(n) \quad (5)$$

$$x_r(n) = x_{orig-r}(n) \cdot w_t(n) \quad (6)$$

Figure 5.(a) depicts an example of original window of analysis, with left and right sub-windows in blue and green respectively, overlapped with the triangular window $w_t(n)$. Subfigure 5.(b) shows the frame of signal after being multiplied by the triangular window.

ii. Spectral densities of left and right sub-windows are calculated. If we note $X_L(w)$ and $X_R(w)$ the Fast Fourier Transforms of sub-windows $x_l(n)$ and $x_r(n)$ respectively, and we define a Hamming window of length L_0 ($w_H(L_0)$) as:

$$w_H(l) = 0.54 - 0.46 \cdot cos(\frac{2\pi \cdot l}{L_0 - 1}) \quad l \leq L_0 \quad (7)$$

The convolution of the modules of $X_L(w)$ and $X_R(w)$ with the Hamming window, will produce the smoothed

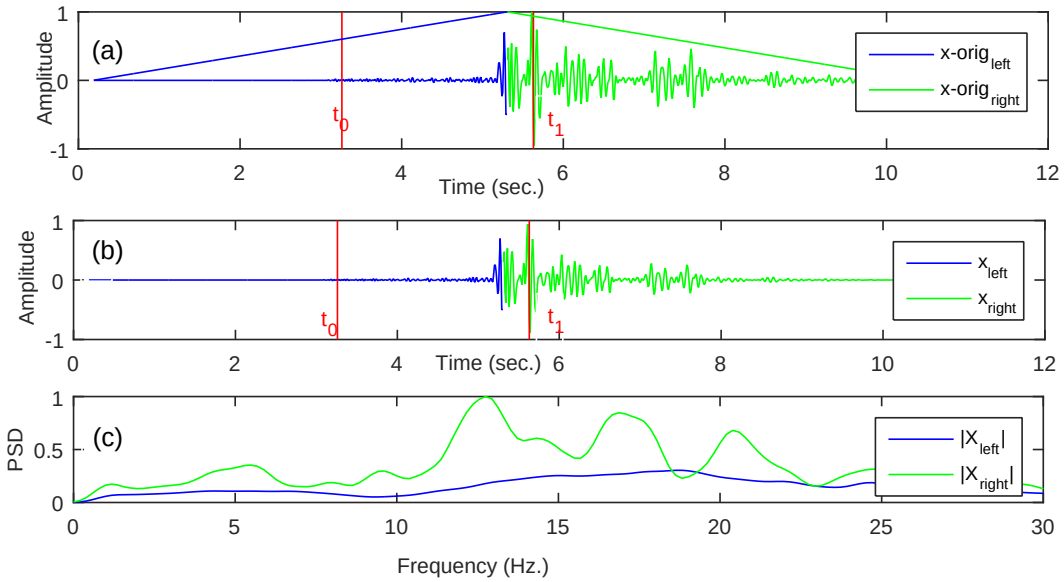

Figure 5. Frame based processing to calculate spectral dissimilarity. Signal is first split into subwindows of 1024 samples (a). Each window is multiplied by a triangular window that emphasizes the central samples (b). Spectral density of the left (blue) and right (green) halves of the window are calculated (c)

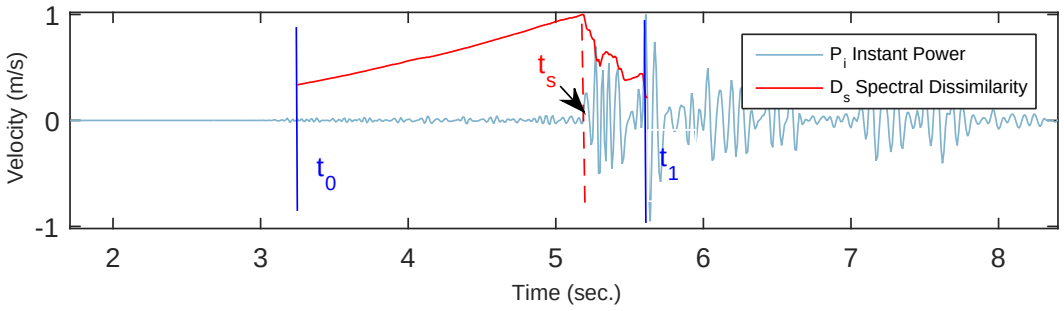

Figure 6. Spectral dissimilarity D_s (red) in the window of analysis calculated over horizontal instant power $P_i(t)$. S-phase arrival takes place when D_s reaches its maximum value

spectral densities of left and right sub-windows depicted in blue and green respectively in Figure 5.(c):

$$X_{LS}(w) = |X_L(w)| * w_H(l) \quad (8)$$

$$X_{RS}(w) = |X_R(w)| * w_H(l) \quad (9)$$

iii. *Spectral dissimilarity* between left and right sub-windows of central sample n under analysis, is then defined as:

$$D_s(n) = \sum_{i=1}^{\frac{L}{4}} (X_{LS}(w_i) - X_{RS}(w_i))^2 \quad (10)$$

The sample on which $D_s(n)$ obtains the highest value in the time interval (t_0,t_1) , is selected as arrival time for the S-phase t_S. Figure 6 shows an example of the time function *spectral dissimilarity (D_s)* plotted in red in the window of analysis. It can be seen that D_s has a maximum in the time sample when frequency and energy of the horizontal component $S_h(t)$ simultaneously change.

Very few parameters have to be set in the algorithm presented:

- The window of analysis L should be long enough to provide the frequency resolution needed, but short enough to capture spectral changes with the convenient temporal resolution. In our case, for a sampling frequency $f_s = 100$ Hz and an average $ts - tp$ interval of 2.37 sec., we tested the performance of the algorithm with L being 2048, 1024, 512 and 256 samples. Table II shows the analysis done, obtaining best results for $L = 1024$ samples.
- The maximum amplitude of the triangular window $w_t(n)$, W_0 in equation 4, should be set to keep the range of signal amplitudes after multiplying by it. With such criterion we have set it equal to the amplitude of each central sample $x(n)$ in the window-based analysis.
- Finally, a low pass filter is used to smooth the spectral density of the windows of analysis eliminating short-term variations of frequency not related to the change of phase. A Hamming window of length $L_0 = 10$ frequency bins (1.95 Hz) has been set after testing a range of values between 5 and 20 frequency bins (0.98 and 3.9 Hz respectively) without variations.

Spectral dissimilarity is measured in $P_i(t)$, and simultaneously it is evaluated in the rest of horizontal components $S_h(t)$, $N(t)$ and $E(t)$, thus providing an estimation of the reliability of the

picking in noisy signals.

D. Quality of the Automatic Picking

Impulsive and energetic S-phase arrivals, separated enough in time from P-phase arrivals, will produce a simultaneous clear maximum of Ds in the four horizontal components $P_i(t)$, $S_h(t)$, $N(t)$ and $E(t)$. P-phase coda, scattering, proximity between phases arrival, rough topographies and other near-source effects generate noisy unclear emergent arrivals not easy to pick. Although a clear consistent picking is desirable, when the quality of the phase does not permit it, it is at least necessary to know its degree of reliability. A quality measure based on the component's SNR around t_s could mislead results if a big dissimilarity occurs in a component for spurious reasons. Averaging t_s obtained in all horizontal components would also degrade performance of potential good estimations in one/some of them. As an alternative, the algorithm proposes a conservative and reliable measure of the quality of the automatic picking (Q_A) based on a criterion of *'consensus'*. Picking is performed on horizontal component $P_i(t)$ after empirical verification of its best functioning detailed in latter Section 4.B. In addition, arrival times ts obtained in the four components are compared in pairs, leading to 6 comparisons used to measure the quality of the picking: $|ts_{Pi} - ts_{Sh}|$, $|ts_{Pi} - ts_N|$, $|ts_{Pi} - ts_E|$, $|ts_{Sh} - ts_N|$, $|ts_{Sh} - ts_E|$ and $|ts_N - ts_E|$. Automatic quality assignments criteria are enumerated in table I, on which '0', '1' and '2' represent *high* , *medium* and *low* picking qualities respectively.

more than 3 comparisons of ts differ in less than 0.1 sec.	$Q_A = 0$
3 comparisons of ts differ in less than 0.1 sec.	$Q_A = 1$
2 or less comparisons of ts differ in less than 0.1 sec.	$Q_A = 2$

Table I
AUTOMATIC ASSIGNMENT OF QUALITY OF THE PICKING BASED ON DIFFERENCES BETWEEN ts DETECTED IN THE DIFFERENT HORIZONTAL COMPONENTS.

Figure 7 shows two examples of automatic S-phase picking using the algorithm proposed. Maximum spectral divergence is located in the window of analyis for the 4 horizontal components, and set as time arrival for S-phase t_S. The subfigure on the left shows a case on which pickings in all components differ among each other in less than 0.1 seconds. Automatic quality of the picking is set to 0. Subfigure on the right presents a case on which 5 out of the 6 comparisons of t_S in the 4 components differ more than 0.1 seconds. The quality is set then to 2.

IV. EXPERIMENTAL WORK

A. Database description

The algorithm proposed has been tested using 90 volcano-tectonic earthquakes occurred in the surroundings of Mt. Etna during 2010. The events were registered in a network of 29 permanent stations provided by the Instituto Nazionale di Geofisica e Vulcanologia di Catania, originating a total of 300 registers of VT phase arrivals with a sampling frequency of 100 Hz. Mt. Etna, located in eastern Sicily, is one of the most active volcanoes in the world, with a great geological and structural complexity that produces a high rate of shallow earthquakes of low and moderate energy [41]. Figure 8 shows the histogram of depths of hypocenters and magnitude of the events used. Earthquakes under analysis have low magnitudes and in most of the cases were generated at shallow depths in the surroundings of the volcano. A few of them even have their hypocenter above sea level.

To evaluate the quality of the algorithm proposed in this work, all the S-phase arrivals have been manually picked by an expert geophysicist, that also rated the reliability of the manual picking according to the characteristics of the VT. Based on its signal-to-noise ratio, its energy, depth, etc., VTs have been divided into *'clear S-phase arrival'* (labeled as having $Q_E = 0$), *'S-phase can be detected'* ($Q_E = 1$), *'unclear and difficult S-phase detection'* ($Q_E = 2$) and *'very unclear S-phase arrival'* ($Q_E = 3$). Figure 10 shows the distribution of the four sets of VTs in the database of 300 events used in the experiment.

B. Experiments: S-Phase Picking using Spectral Dissimilarity

Four sets of experiments have been performed to test the algorithm. S-Phase picking assessed by the human expert has been taken as reference, defining the picking error (e) as the unsigned difference between human and automatic picking in units of seconds.

Firstly, table II analyzes the performance of the technique for different sizes of the window of analysis L, (x_{orig} in section III.C), used to calculate the spectral dissimilarity along consecutive frames of signal. The size of the window will search for a compromise between spectral and temporal resolution in the frequency domain analysis. Results show that for an average $t_s - t_p$ of 2.37 sec. best performance is achieved tuning the window size to around 10 sec. (1024 samples at a sampling frequency f_s=100 Hz).

Table III compares the percentage of pickings with error lower than 0.2 sec., 0.5 sec. and 1 sec., when applying the maximum spectral divergence criterium to the four horizontal components used. Results show that instant power of horizontal components, $P_i(t)$, produces the best results closely followed by the horizontal component of the transverse plane $S_h(t)$. The usage of only one original horizontal component from the seismometer, N(t) or E(t), degrades slightly the performance. This behavior is coherent with the fact that none of them is capturing all the information of the transverse plane, but only its projection in the particular axis N or E. The slightly better results of $P_i(t)$ compared to $S_h(t)$, added to its independence of a correct P-Phase picking to perform rotation of axis (like $S_h(t)$), make it the best horizontal component for t_S estimation.

Table IV analyzes in detail the automatic picking results on $P_i(t)$, linked to the automatic quality assessment given by the algorithm. From the 300 signals to be picked, the algorithm assesses a high quality picking ($Q_A = 0$) to 120 pickings. The 97.05 % of these high quality pickings differ less than 0.2 sec. from the picking performed by the human expert. The goodness of the method decreases progressively

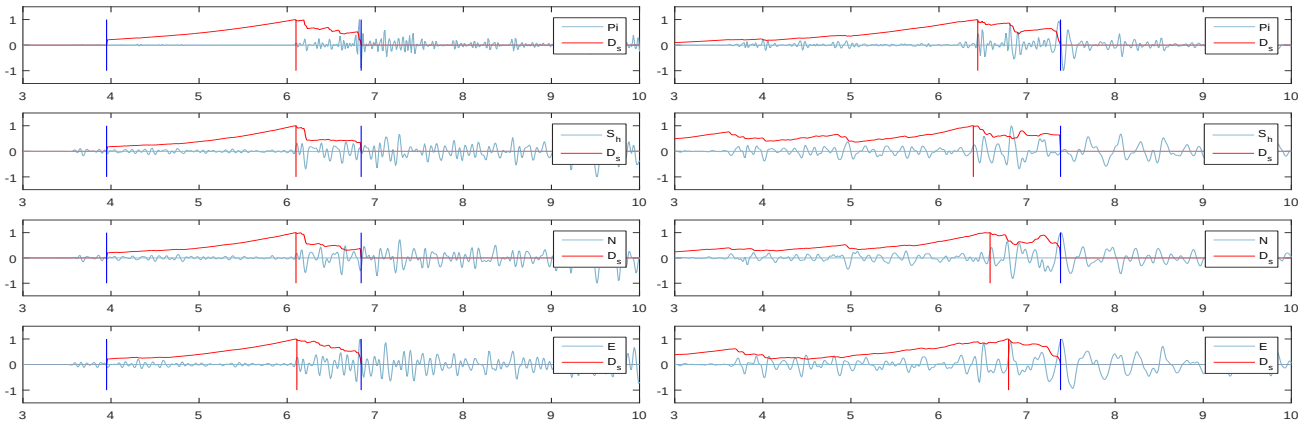

Figure 7. Subfigure on the left shows an example of high quality picking using D_S. The algorithm assignes $Q_A = 0$ as t_S time in all components differs less than 0.1 seconds. Right side subfigure depicts an example of low quality picking using D_S. Given the difference in t_S for all components, the automatic quality assigned is $Q_A = 2$.

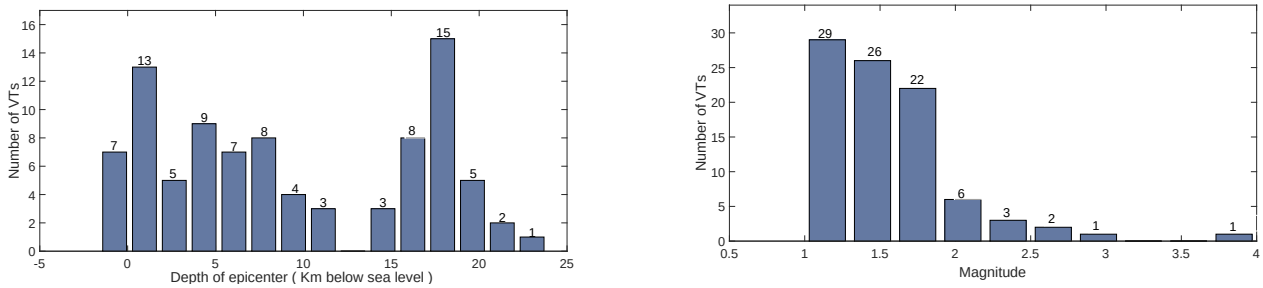

Figure 8. Histograms of depths in Km. (left subfigure) and magnitude (right subfigure) for the database of 90 volcano-tectonic earthquakes from Etna volcano used in the experiments

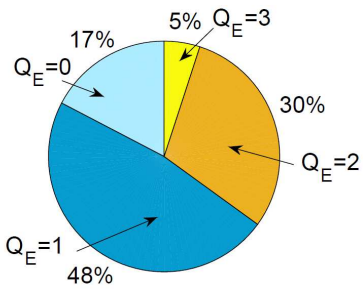

Figure 9. Percentage of VT registers (over a total of 300) manually labeled by the human expert with qualities 0, 1, 2 and 3, in the database used in the experiment

	L = 2048	L = 1024	L = 512	L = 256
e $<$ 0.2 s	77.00%	77.66%	70.66%	42.66%
e $<$ 0.5 s	86.33%	88.00%	79.66%	60.33%
e $<$ 1 s	91.66%	94.00%	85.66%	75.33%

Table II

S-PHASE PICKING RESULTS COMPARED FOR 4 DIFFERENT SIZES (IN NUMBER OF SAMPLES) OF WINDOW L (x_{orig} IN SEC. III.C) FOR CALCULATION OF SPECTRAL DISSIMILARITY. RESULTS OVER A TOTAL OF 300 REGISTERS.

	$P_i(t)$	$S_h(t)$	N(t)	E(t)
e $<$ 0.2 s	77.66%	77.00%	74.33%	73.33%
e $<$ 0.5 s	88.00%	87.66%	85.33%	83.33%
e $<$ 1 s	94%	94%	91.67%	90.33%

Table III

COMPARATIVE RESULTS ON S-PHASE PICKING APPLYING SPECTRAL DISSIMILARITY D_s ON ALL HORIZONTAL COMPONENTS. PERCENTAGE OF PICKINGS WITH AN ERROR (e) LESS OR EQUAL THAN 0.2 , 0.5 AND 1 SECONDS OVER A TOTAL OF 300 REGISTERS

	$P_i(t)$		
	$Q_A = 0$ (120 VTs)	$Q_A = 1$ (62 VTs)	$Q_A = 2$ (118 VTs)
e $<$ 0.2 s	97.05%	79.03%	56.78%
e $<$ 0.5 s	99.17%	91.93%	74.58%
e $<$ 1 s	100%	95.16%	87.29%

Table IV

S-PHASE PICKING APPLYING SPECTRAL DISSIMILARITY D_s ON INSTANT POWER $P_i(t)$. PERCENTAGE OF PICKINGS WITH AN ERROR (e) LESS OR EQUAL THAN 0.2 , 0.5 AND 1 SECONDS.

as expected for the 118 VTS with assessed with low quality $Q_A = 2$. These results show that quality assessment given by the method proposed is conservative, being a useful filtering criterium to warranty the reliability of the picking needed, depending on its usage or specific needs of accuracy.

Finally Table V analyzes the correlation between the original quality assessment performed by the human expert (Q_E) and the automatic quality assessment (Q_A). It shows that 80.77% of VTs labeled with $Q_E = 0$ by the human expert are

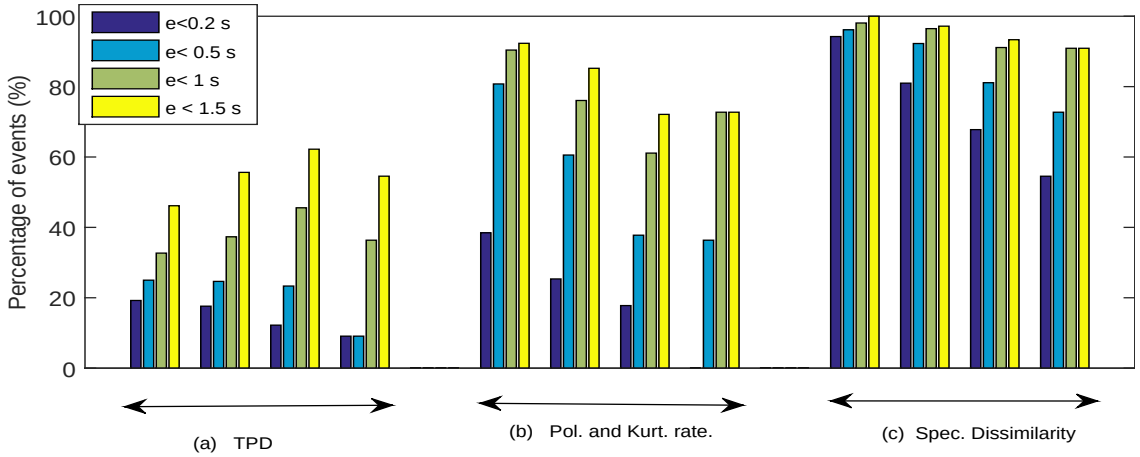

Figure 10. Picking error percentages for the 4 sets of VTs with qualities 0, 1 2 and 3. Comparison of results for: (a) algorithm based on the damped predominant period *Tpd* [11] on the left side of the figure, (b) algorithm based on polarization analysis and kurtosis rate proposed by [29] om the center of the figure, and (c) technique proposed in this work on the right side of the figure

			e<0.2 s	**e<0.5 s**	**e<1 s**
Q_E**=0** (52 VTs)	Q_A**=0**	80.77%	97.6%	100%	100%
	Q_A**=1**	9.61%	100%	100%	100%
	Q_A**=2**	9.61%	60%	60%	80%
Q_E**=1** (142 VTs)	Q_A**=0**	42.25%	100%	100%	100%
	Q_A**=1**	23.94%	76.47%	94.12%	94.12%
	Q_A**=2**	33.80%	58.33%	81.25%	93.75%
Q_E**=2** (90 VTs)	Q_A**=0**	16.67%	93.33%	100%	100%
	Q_A**=1**	24.44%	77.27%	90.90%	100%
	Q_A**=2**	58.89%	56.60%	71.70%	84.90%
Q_E**=3** (16 VTs)	Q_A**=0**	18.76%	66.67%	100%	100%
	Q_A**=1**	6.25%	0%	0%	0%
	Q_A**=2**	75.00%	50.00%	66.67%	75.00%

Table V
PERCENTAGES OF DIFFERENT QUALITIES OF PICKINGS PROVIDED BY THE ALGORITHM, (Q_A), FOR THE 4 SUBSETS OF VTS WITH QUALITIES 0 TO 3 TO PROVIDED BY THE HUMAN EXPERT (Q_E)

provided with a high quality automatic picking ($Q_A = 0$) by the algorithm. The coherence of qualities assignment remains for rest of values of Q_E. 12 out of the 16 VTs labeled with $Q_E = 3$, (75%), are assessed with a low reliance $Q_A = 2$.

C. Comparison with other techniques

Shallowness and low magnitude of the events of the database under analysis, together with the near source-effects, degrade the performance of picking techniques based on polarization analysis. The proximity of P- and S- phases hinders the usage of autoregressive methods based on AIC criterion, as they need a certain length to model correctly the signal before and after S-Phase arrival.

To place the technique proposed in the framework of existing S-Phase picking algorithms, we have compared it to other two S-phase picking methods. The first one, is a the technique based on the estimation of the *damped predominant period function* T^{pd} proposed by [11]. Based on the fact that P- and S- phases will typically have different frequency distributions, the damped predominant period of the ground velocity will change with the arrival of S- phase, and therefore can be used as automatic picker. Parameters of the T^{pd} picker have been set according to [11] with the following values: the time window τ_ω has been set to 5 sec., and the upper limit for T^{pd} function prior to phase arrival has been set to $\tau_{mx} = 0.1$. The second S-phase algorithm compared was proposed by [29]. It defines a polarization filter to separate P- and S- phases. Then it uses an STA/LTA detector to search for the time range when S-phase is expected, and picks it when the kurstosis rate of the S- filtered component is maximum on that range. Windows of 4 sec. and 1 sec. have been used for polarization analysis, and kurtosis calculation respectively. STA and LTA windows were set to 1 and 10 sec. Figure 10 compares the results of applying these two well known techniques and the algorithm proposed in this work to our database of 300 volcano-tectonic earthquake registers. Percentages of automatic pickings with an error minor to 0.2 sec., 0.5 sec., 1 sec. and 1.5 sec. are compared for all techniques. Results show the interesting better performance of our method for all the range of VTs, specially increasing the percentage of picking errors lower than 0.2 sec. Spectral dissimilarity improves significantly results for the specific challenging conditions of our database.

V. CONCLUSIONS

This work presents an automatic S-phase picking algorithm specially devoted to volcano-tectonic earthquakes. It is based on a windowed search of maximum simultaneous changes of frequency and energy in several horizontal components of the seismic signal. A measure of these changes, called spectral dissimilarity D_s, is evaluated in the horizontal instant power of the signal. S-phase arrival is set as the instant when D_s reaches its maximum within a time interval on which it is expected. In addition, the algorithm provides a measure of the quality of the picking based on its consistency when applied to the rest of horizontal components. This quality evaluation is a useful tool to differentiate picking results depending on their application, or to give them different weights when used

as input for seismic tomographies. Two main results of the experiments performed are remarkable. Firstly, the algorithm works satisfactorily for volcano-tectonic events in spite of the difficulties that these low magnitude and shallow events present. Secondly, the automatic picking-quality provided (coherent with quality evaluation performed by human experts) permits to distinguish reliable from non-reliable pickings, and therefore select a quality threshold depending on the needs of the user.

VI. Acknowledgemtens

This research was funded by Research Project TEC2015-814 68752 (MINECO/FEDER) and by research grant José Castillejo CAS17/00411.

References

[1] R. Allen, *Automatic phase pickers: Their present use and future prospects* Bull. Seismol. Soc. Amer., vol. 72, no. 6B, pp. S225–S242. (1982)

[2] L. Küperkoch, T. Meier, T. *Automated Event and Phase Identification* In: Bormann, P. (Ed.), New Manual of Seismological Observatory Practice (NMSOP-2), Potsdam: German Research Centre of Geosciences, GFZ, 1-52, (2011)

[3] J. Akram, D.W. Eaton *A review and appraisal of arrival-time picking methods for downhole microseismic data* Geophysics, 81(2), KS67-KS87, (2016)

[4] R.V. Allen *Automatic earthquake recognition and timing from single traces* Bull. Seism. Soc. Am., 72, 225-242 (1978)

[5] M. Baer, U. Kradolfer *An automatic phase picker for local and teleseismic events* Bull. Seism. Soc., 77, 1437-1445, (1987)

[6] C. Saragiotis, L.J. Hadjileontiadis, S.M. Panas *PAI-S/K: A Robust Automatic Seismic P Phase Arrival Identification Scheme* IEEE Transactions on Geoscience and Remote Sensing, 40(6), 1395-1404, (2002)

[7] X. Li, X. Shang, Z. Wang, L. Dong, L. Weng *Identifying P-phase arrivals with noise: An improved Kurtosis method based on DWT and STA/LTA* Elsevier Journal of Applied Geophysics, 133, 50-61,(2016)

[8] E. Kalkan *An automatic P-Phase Arrival-Time Picker* Bull. Seism. Soc., 106(3), 1-15, (2016)

[9] I. Álvarez, L. García, S. Mota, G. Cortés, C. Benítez, A. D. la Torre., *An automatic p-phase picking algorithm based on adaptive multiband processing* IEEE Geoscience and Remote Sensing Letters, 10(6), 1488-1492, (2013)

[10] L. García, I. Álvarez, C. Benítez, M. Titos, A. Bueno, A. De la Torre, A. Díaz-Moreno, J. Prudencio, A. García-Yeguas, G. Alguacil, J.M. Ibánez, *Advances on the automatic estimation of the P-wave onset time* Annals of Geophysics, 59(4), S0434, (2016)

[11] M.W. Hildyard, S.E.J. Nipress, A. Rietbrock *Event Detection and Phase Picking Using Time-Domian Estimate of Predominate Period* T^{pd} Bull. of Seism. Soc., 98(6), 3025-3032, (2008)

[12] K.S. Anant, F.U. Dowla, *Wavelet transform methods for phase identification in three-component seismograms* Bull. of Seismol. Soc. of America 87, 1598–1612, (1997)

[13] J.J. Galiania-Merino, J.L. Rosa-Herranz, S. Parolai *Seismic P-Phase Picking Using a Kurtosis-Based Criterion in the Stationary Wavelet Domain* IEEE Trans. on Geoscience and Remote Sensing, 46(11), 3815-3826 (2008)

[14] H. Akaike *Information theory and an extension of the maximum likelihood principle, in Proc. of the 2nd International Symposium on Information Theory*, 267-281 (1973)

[15] T. Takanami, G. Kitagawa *A new efficient procedure for the estimation of onset times of seismic waves* J. Phys. Earth, 36, 267-290, (1988)

[16] T. Takanami, G. Kitagawa *Estimation of the arrival times of seismic waves by multivariate time series models* Ann. of the Institute of Statistical Mathematics, 43, 407-433, (1981)

[17] M. Leonard, B.L.N. Kennet *Multi-component autoregressive techniques for the analysis of seismograms* Physics of the Earth and Planetary Interiors

[18] T. Akazawa *A technique for automatic detection of onset time of P- and S-phases in strong motion records* 13th World Conference on Earthquake Engineering, Vancouver, B.C. Canada, No. 786, (2004)

[19] A. Cichowicz, *An automatic S-phase picker* Bull. of Seis. Soc., 83(1), 180-189, (1993)

[20] D. Patanè, F. Ferrari, *ASDP: a PC-based program using a multi-algorithm approach for automatic detection and location of local earthquakes* Physics of the Earth and Planetary Interiors 113, 57–74, (1999)

[21] A.M. Reading, W. Mao, D. Gubbins, *Polarization filtering for automatic picking of seismic data and improved converted phase detection* Geophys. J. Int., 147, 227–234, (2001)

[22] D. Patanè, F. Ferrari, E. Giampiccolo, S. Gresta *A PC-based computer package for automatic detection and location of earthquakes: application to a seismic network in eastern Sicily (Italy)*

[23] C.R. Pinnegar, *Polarization analysis and polarization filtering of three-component signals with the time–frequency S transform* Geophys. J. Int. 165, 596–606,(2006)

[24] L. D'Auria, F. Giudicepietro, M. Martini, M. Orazi, R. Peluso, G. Scarpato, *Polarization Analysis in the Discrete Wavelet Domain: An Application to Volcano Seismology* Bull. of Seismol. Soc. of America, 100(2), 670-683, (2010)

[25] J. Wang, T. Teng, *Identification and Picking of S Phase using Artificial Neural Networks* Bulletin of the Seismological Society of America, Vol. 87, No. 5, pp. 1140-1149, October (1997)

[26] S. Gentili, A. Michelini, *Automatic picking of P and S phases using a neural tree* Jrnl. of Seismology, Springer Verlag, 10, 9-63, (2006)

[27] C. Rawles, C. Thurber, *A non-parametric method for automatic determination of P-wave and S-wave arrival times: application to local micro-earthquakes* Geophys. J. Int. 202, 1164–1179, (2015)

[28] Z.E. Ross, Y. Ben-Zion, *Automatic picking of direct P, S seismic phases and fault zone head waves* Geophys. J. Int. 199, 368-381, (2014)

[29] Z.E. Ross, M.C. White, F.L. Vernon, Y. Ben-Zion, *An improved Algorithm for Real-Time S-Wave Picking with application to the (Augmented) ANZA Network in Southern California* Bull. of Seism. Soc., 106(5),2013-2022, (2016)

[30] C. Baillard, W.C. Crawford, V. Ballu, C. Hibert, A. Mangeney *An Automatic Kurtosis-Based P- and S-Phase Picker Designed for Local Seismic Networks* Bull. of Seism. Soc., 104(1), 1-16 (2014)

[31] T. Diehl, N. Deichmann, E. Kissling, S. Husen, *Automatic S-Wave Picker for Local Earthquake Tomography* Bull. of the Seism. Soc., 99(3), 1906–1920, (2009)

[32] H. Zhang, C. Thurber, C.Thurber, C. Rowe, *Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings* Bull. Seism. Soc. 93(5), 1904-1912, (2003)

[33] C. Bai, B.L.N. Kennet, *Automatic Phase-detection and identification by full use of a single three-component broadband seismogram* Bull. of Seism. Soc. of America, 90(1), 187-198, (2000)

[34] Y. Nakamura, *A New Concept for the Earthquake Vulnerability Estimation and its Application to the Early Warning System* J. Zschau et al. (eds.), Early Warning Systems for Natural Disaster Reduction, Springer-Verlag Berlin Heidelberg, 693-699, (2003)

[35] P. Bormann, B. Engdahl, R. Kind *Seismic wave propagation and earth models- In Bormann, P. (Ed.) new Manual of seismological observatory practices (NMSOP)* Postdam: German Research Centre of Geosciences, GFZ, 649-690, (2009)

[36] J. Wasserman *Volcano Seismology- In Bormann, P. (Ed.) new Manual of seismological observatory practices (NMSOP)* Postdam: German Research Centre of Geosciences, GFZ,1-77 (2009)

[37] J. Neuberg, T. Pointer, *Effects of volcano topography on seismic broadband waveforms* Geophys. J. Int. 143, 239–249, (2000)

[38] Y. Nakamura, *Earthquake motion of the 2011 off the Pacific Cost of Tohoku earthquake and the situation of issuing various earthquake alarms* Joint. Conf. Proc. 9th Int. Conf. on Urban Earthquake Eng. Tokyo, Japan, (2012)

[39] G. Alguacil, F. Vidal, *Medidas instrumentales de la intensidad del movimiento del suelo. Aplicación a terremotos europeos.* In Homenaje A La Profesora María Dolores Romacho Romero, pp 15-32. University of Almería, Spain, (2012)

[40] T. Diehl, E. Kissling, P. Bormann *Tutorial for consistent phase picking at local to regional distances* In: Bormann, P. (Ed.), New Manual of Seismological Observatory Practice 2 (NMSOP-2), Potsdam : Deutsches GeoForschungsZentrum GFZ, pp. 1—21, (2011)

[41] S. Alparone, V. Maiolino, A. Mostaccio, A. Scaltrito, A. Ursino, G. Barberi, S. D'Amico, G. Di Grazia, E. Giampiccolo, C. Musumeci, L. Scarfi, L. Zuccarello, *Instrumental Seismic catalogue of Mt. Etna earthquakes (Sicily, Italy): ten years (2000-2010) of instrumental recordings.* Annals of Geophysics, 58(4), (2015)