AUTOMATIC S-PHASE ARRIVAL IDENTIFICATION FOR LOCAL EARTHQUAKES

IZIDOR TASIČ and FRANC RUNOVC

About the authors

Izidor Tasič
Environmental Agency of the Republic of Slovenia,
Vojkova cesta 1b, 1000 Ljubljana, Slovenia
E-mail: izidor.tasic@gov.si

corresponding author

Franc Runovc
University of Ljubljana,
Faculty of Natural Sciences and Engineering
Aškerčeva cesta 12, 1000 Ljubljana, Slovenia
E-mail: franc.runovc@ntf.uni-lj.si

Abstract

In the case of a strong local earthquake, a quick report about the earthquake's location is expected. Such reports are usually performed automatically, where the identification of the seismic-phase arrival of the various seismic waves on the seismogram is the most important task. For this purpose, numerous detecting methods for the first P-wave arrival identification and determination are used. But in some cases, where the number of seismic stations in a local seismic network is very small, an automatic reading of the S-wave arrival is required. An algorithm for the automatic picking of the S wave arrival from three-component seismic data has been developed. Three parameters of the signal are calculated from these data and the S phase arrival is declared when the product of the three parameters increases above a reference level. Such a so-called S-phase picker is used to automatically analyze the data from local earthquakes in Slovenia's seismic network.

Keywords

seismic waves, P and S waves, analysis of seismogram data, algorithm of the automatic picking of the S-wave arrival, S-phase picker, seismic station, seismic network

1 INTRODUCTION

The seismic-phase arrival identification of the various seismic waves on the seismogram is the most important task for a seismologist; it is the basic input data for additional geophysical and seismological studies, where earthquakes are the source of information. The complexities of different earthquake source mechanisms, the effects of propagation through inhomogeneous media, scattering, reflection and refraction on different boundaries of the earth's interior and the vast amount of seismic data mean that this task is very complicated and is performed by human experts who have to recognize the arrival of different types of seismic waves and their phases, which are frequently covered by seismic noise or the seismic waves themselves. In addition, the further the earthquake is from a seismic station, the more complex is the seismic record of such an earthquake.

For local earthquakes there are two important seismic wave identifications. A seismic station first detects the arrival of the P (primary) waves, which are compressional elastic waves; then it detects the arrival of the S (secondary) waves or transverse waves. The experts, who analyze the records of a local earthquake, usually denote just the entry of the first P seismic waves on the vertical record of the earth's movements, while the S seismic waves are marked on the horizontal record. The entry of the S seismic wave appears within the scattered P seismic waves; therefore, it is less conspicuous. To distinguish one from another, an expert must pay attention to the amplitude, the frequencies and the motion change, as well as to the position of the extreme of the amplitude. Consequently, it is not always easy to fix the entry of the phases in time. When both entries are well defined, the time difference between the entries is used to calculate the distances of the seismic station form the earthquake's hypocenter, and when this distance is known, the magnitude of the earthquake can also be estimated.

The very first and most basic purpose of a local seismic network is an accurate determination of the earthquakes' locations. This is also the goal of governments, which often finance the local seismic network. In the present era of a rapid exchange of information, many organizations, such as civil defense, expect a quick report about the earthquake's location. The civil defense uses this information to mitigate the social and economic consequences in the case of a strong earthquake. Such reports are usually performed automatically and depend on the quality of the local network of seismic stations, as well as on the information infrastructure, and on the algorithms, which perform the automatic phase-arrival identifications and consequently automatically estimate the basic parameters of an earthquake, such as its location and its magnitude. A modern seismic station consists of an acquisition unit, which digitizes the analogue signals from the seismometer and sends seismic data in real time to the supervision center, and a broadband seismometer, where three sensors detect the ground motion in the vertical and two horizontal directions (usually the east-west and north-south directions). When the seismic signal is received at the supervision center, a seismic-wave (or seismic-phase) arrival identification is automatically performed, and based on this information the basic parameters of the earthquake are calculated.

In recent years, numerous detection methods for the first P-wave arrival identification and determination have been developed and brought into use, from a very simple threshold algorithm to more sophisticated, adaptive methods, including neural networks and pattern matching [1],[2],[3],[4],[5],[6],[7],[8],[9],[10],[11]. Although these methods still cannot completely replace expert knowledge, they play a significant role in any automated location procedure, which provides the first information about an earthquake's parameters for an emergency centre. The number of seismic stations that provide earthquake-waveform data (or information about trigger times) and pass them to the center in real-time play an important role in the automatic procedure. For sufficient information, the number of seismic stations should be as high as possible. In some cases, however, the number of seismic stations in a local seismic network, which contribute their information to supervision center, can be very low. For this reason, an automatic reading of the secondary-waveform arrival is recommended. But the automatic S-wave arrival-identification methods are very few [12],[13],[14],[15], and they tend to be relatively complex. The reason why there are so many first P-wave arrival detectors compared to just a few first S-wave arrival detectors lies in fact that the seismic signal is, especially for a strong earthquake, significantly higher than the seismic noise, while the S seismic waves travel with a lower velocity than the P seismic waves, and the first S-wave arrival can be hidden in the P seismic waves.

For this reason we have developed our own S phase detector for the local earthquake waveforms. The algorithm is fairly simple, efficient and also provides good information in spite of the higher seismic noise. Since the end of 2002 this algorithm has been working automatically in the Slovenian seismic network, and in the case of the strong earthquake in 2004, contributed to the rapid automatic location of this earthquake.

2 ALGORITHM AND METHOD

A three-component seismic station monitors the ground motions, usually the velocities along the east-west (E-W), north-south (N-S) and vertical (V) directions. With the help of an automatic first P-wave arrival detector the measuring system prepares a set of data containing a record of the local earthquake. Let us define the vectors for all three components as

$$\mathbf{Y}_e = (e_1, e_2, ..., e_k) \qquad (1)$$

$$\mathbf{Y}_n = (n_1, n_2, ..., n_k) \qquad (2)$$

$$\mathbf{Y}_v = (v_1, v_2, ..., v_k) \qquad (3)$$

The subscripts e, n and v indicate the east, north and vertical velocity components, respectively, and k is the number of samples. Let the first P seismic wave arrive at the sample i_p and the first S seismic wave arrive at the sample i_s, where $1 < i_p < i_s < i_k$. The following vector

$$\mathbf{O} = (o_1, o_2, ..., o_k) \qquad (4)$$

with the components

$$o_i = e_i^2 + n_i^2 + v_i^2 \ \ \text{for} \ \ i = 1, ..., k - l$$

can be used to estimate the relative total energy of an earthquake [16], [17]. This vector can also be expressed as the square of an envelope function, which is thought of as a positive outline of the seismogram [18]. In a similar way, we define the envelope function for earthquake records in horizontal planes, in an non-rotated data system, as the following vectors

$$\mathbf{E} = (|e_1|, |e_2|, ..., |e_k|) \qquad (5)$$

$$\mathbf{N} = (|n_1|, |n_2|, ..., |n_k|) \qquad (6)$$

The S-wave picker algorithm was developed on the basis of an energy analysis, where the idea of minimalism is followed: "the less the filtering, the better the algorithm" [19]. The energy analysis was based on the ratio of the short-term average (STA) energy to the long-term average (LTA) energy level, derived from the same seismogram [17]. These types of algorithms are referred to as STA/LTA detectors. Today, the 'short-time-average through long-time-average trigger' (STA/LTA) is the most broadly used algorithm for a P phase picker because of its simplicity and efficiency. Several different algorithms have been developed, but the most popular is the procedure where the average values of the absolute amplitude of a seismic signal are continuously calculated in two consecutive moving-time windows [24]. The short-time average is sensitive to seismic events, while the long-time average provides information about the temporal amplitude of the seismic noise at the measurement site. When the ratio of both exceeds a pre-set value, an event is 'declared' and data starts being recorded in a computer data file. Because the first S-wave arrival is hidden in the P seismic waves, this algorithm is not useful for S-wave detectors. For this reason, we define a transformation $f(x)$ for the vector x, where $f(x_i)$ is given by

$$f(x_i) = \frac{\frac{1}{l}\sum_{j=i}^{i+l-1}|x_j|}{\frac{1}{k-i+1}\sum_{j=i}^{k-l}|x_j|} \quad \text{for} \quad i = 1, ..., k-l \qquad (7)$$

and l is the number of samples in an STA window. While the nominator still follows the idea of STA and is 'sensitive' to local changes in the amplitudes, the denominator provides information about the amplitude variation from the sample i to the end of the signal. Because, generally, the amplitudes of the S seismic waves for a local earthquake are about five times larger than those of P seismic waves, and also S seismic waves are much more strongly attenuated than P seismic waves [24], the value of the denominator is expected to be higher in the cases where i is before the first S-wave arrival, $i < i_s$. Using equation (7), the transformation of the total energy presented by (4) reduces to

$$\mathbf{f_O} = (f(o_1), f(o_2), ..., f(o_{k-l})) \qquad (8)$$

Similarly, the transformations of the envelopes of the horizontal components, presented with (5) and (6), are then given by

$$\mathbf{f_E} = (f(e_1), f(e_2), ..., f(e_{k-l})) \qquad (9)$$

$$\mathbf{f_N} = (f(n_1), f(n_2), ..., f(n_{k-l})) \qquad (10)$$

The relative position of the seismic station with reference to the source mechanism of an earthquake means that it is not possible to predict on which horizontal component the S-wave's arrival will be more accurately detected. Because of this uncertainty, we create the following, so-called characteristic, vector [12]

$$\mathbf{C} = (c_1, c_2, ..., c_k) \qquad (11)$$

with the components

$$c_i = f(e_i)f(n_i)f(o_i) \quad \text{for} \quad i = 1, ..., k-l \qquad (12)$$

The S-wave arrival is declared when the characteristic function increases above the threshold. It is chosen at a sample c_s, where the threshold is for the first time lower than the values in the characteristic function

$$c_s \ni: (c_1, ..., c_{i-1} \leq threshold) \wedge (c_i > threshold)$$

$$\text{for } i = 1, ..., k-l \qquad (13)$$

To simplify the search procedure, the S-wave arrival can be searched in the interval between the previous automatically defined P-wave arrival and the maximum value of the characteristic function. It is unlikely that the first S-wave arrival will fall outside this time window. For each record the value of the threshold is defined with regard to the characteristic vector. For example, in his picking procedure Cichowicz [12] uses the average value and the variation of its characteristic function. We simplify the procedure. The threshold is related to the maximum value of our characteristic function:

$$threshold = constant(\max(\mathbf{C})); \ 0 < constant < 1 \qquad (14)$$

The constant is defined from the set of earthquake records with known P and S wave arrival times.

2.1 CALCULATION OF THE WINDOW WIDTH L AND THE THRESHOLD CONSTANT

Two unknown parameters have to be evaluated, the length of the STA window 'l' and the values of the constant in the threshold definition. The STA window's length and the constant are parameters that can be chosen arbitrarily after gaining some experience from the real data. These parameters were prepared from the set of 109 earthquake records from the years 1997 and 1998, with the local magnitudes ranging from 1.3 to 5.6, and with the epicenter distances ranging from 16 km to 160 km. The distribution of the epicenter distances with
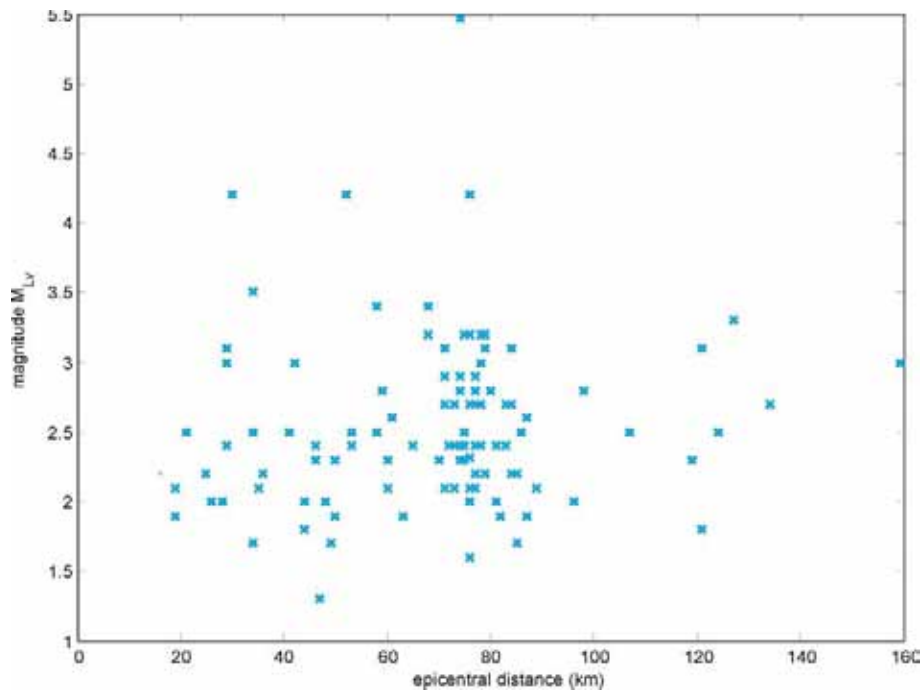
Figure 1. The distribution of epicentral distances with regard to the local magnitude of 109 earthquake records, recorded at the seismic station LJU, which were used to estimate the parameters used in the algorithm.
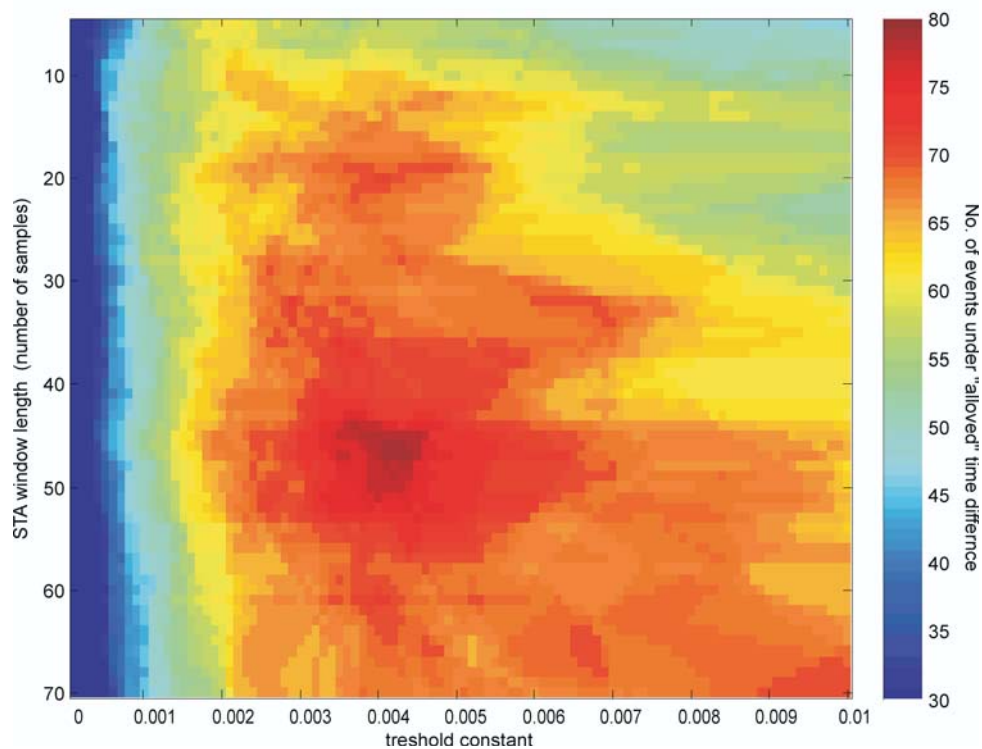

Figure 2. The number of events (colored), where the time difference between the "correct" first S seismic waves and the automatically estimated ones is less than 0.3 seconds. The sampling rate is 200 sps.

regard to the local magnitude is shown in Figure 1. All the events were detected and recorded automatically at the seismic station LJU, which is located near Ljubljana (Slovenia) and is equipped with a Guralp CMG40T seismometer and a Nanometrics RD3 digitizer. The sampling rate used at this station is 200 sps, and the pre-trigger time for automatic P seismic wave arrival detector is set to 10 seconds. Some of these earthquake records were very noisy. All these events were then manually checked by an expert (a seismologist on duty), who manually declared the time of the first P and S wave arrivals. Let us define that this is our reference time, which might not always be correct [17]. The fact is that a simple automatic procedure cannot replace expert knowledge, and for an automatic procedure it is difficult to estimate the allowed time difference between the "correct" manual picking of the first S seismic wave and the automatically estimated one. We will "define" the allowed time difference on the basis of using a single station for an earthquake's location [18]. The time difference between the first S and P entries is proportional to the distance between the seismic station and the earthquake's hypocenter [19]. If the error between the "correct" and the estimated time difference is about 0.3 s, then the error in the estimated epicenter distance is less than 3 km, and in the case of using a single station for an earthquake's location, this error can be taken into account. If more than two seismic stations contributed in the automatic evaluation and these stations were distributed around the hypocenter, the error with respect to the earthquake's hypocenter location could be smaller, or even neglected [24]. We will use the value of 0.3 s as an allowed time difference in the procedure of the window length 'l' and the threshold constant definition. From Figures 2 and 3, the following parameters were evaluated: the window length (l) is 50 samples and the relative threshold level is 0.004.
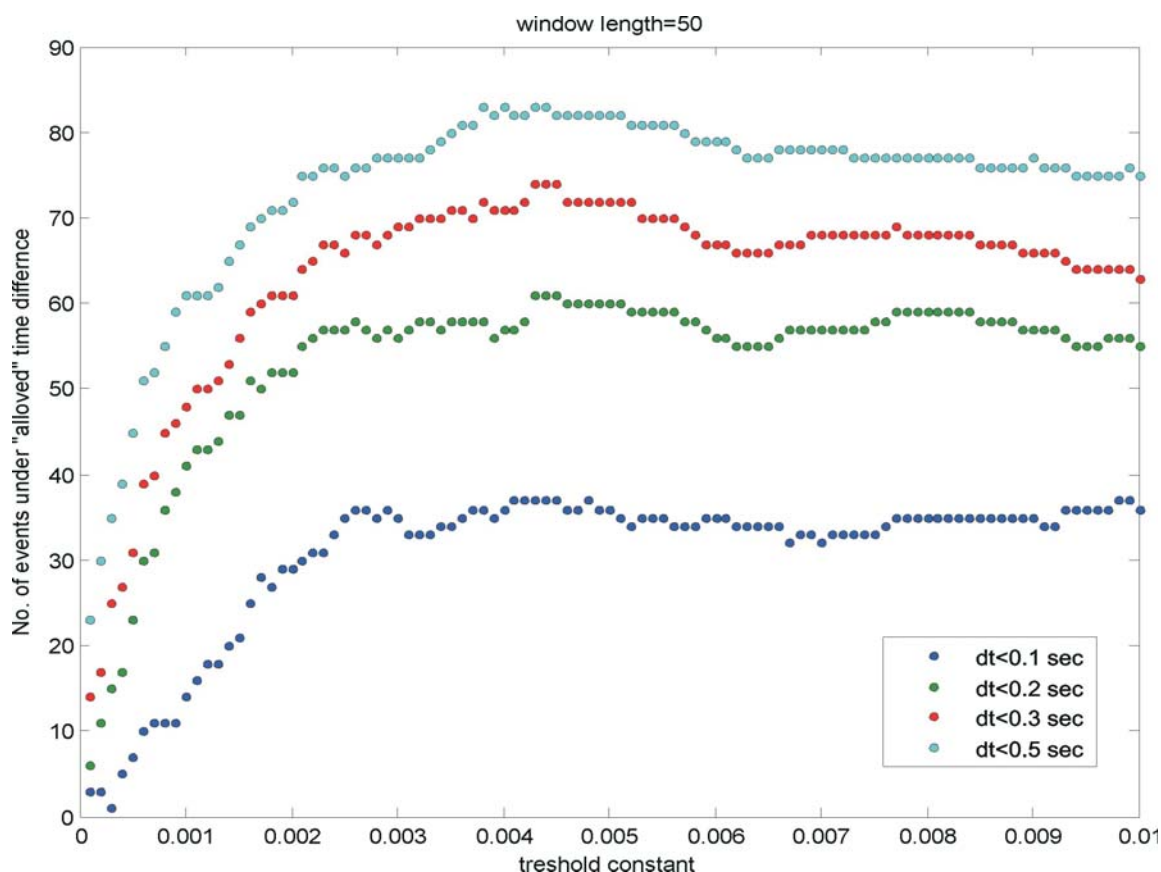

Figure 3. The number of events for different time differences between the "correct" first S seismic waves and the automatically estimated ones, for the STA window with 50 samples (the sampling rate is 200 sps), with regard to the constant value of the threshold level.

3 TEST AND RESULTS

Figures 4 and 5 depict two different earthquake records (their E-W, N-S and Z components), their characteristic signal and the evaluated S-wave arrival. The local magnitude of both events M_{LV} is 2.5, they differ with respect to the epicentral distances. In both cases the error between the "correct" and the estimated time difference is less than 0.1 s. For this time difference no manual retiming is needed [13]. However, for a deviation larger than 0.1 s, a retiming has to be performed, but under the assumption that more seismic stations will contribute their data in the procedure of the automatic location definition of the earthquake's parameters, a time interval, which was less than 0.3 s with regard to the "correct" time, will still give well-defined earthquake parameters. From our test set of 109 earthquake records this happens in 70% of cases. In reality, the automatic procedure is important in the case of a strong earthquake, where the earthquake's signal in a record is not affected by the seismic noise, and the probability that the location of an earthquake is defined correctly, is higher than from our test sets of data. Also, the time difference, which is larger than 0.3 s, can still present a satisfactory automatically calculated location of the earthquake. This statement can be presented in the next example of an earthquake (Figures 6 and 7), which occurred on 23 January 2009 at 04:28 local time. Its

local magnitude was 3.0 and it was felt in Ljubljana and Škofja Loka (Slovenia). The epicenter of this earthquake was estimated by an expert (using data from 29 seismic stations) as 14.41º longitude and 46.09º latitude. Figure 5 shows three seismograms of the vertical components recorded at seismic stations in the Slovenian seismic network (DOBS Dobrina, CESS Cesta, LJU Ljubljana). These stations are equipped with an NMX acquisition system [20] where our system of automatic location works, and for this earthquake only these stations contributed their data to the software for the automatic epicenter location. On each recording the arrivals of the P (red line) and S (green line) waves were automatically denoted and the earthquake's parameters, such as the coordinates of the epicenter (14.44º lon, 46.07º lat) and its magnitude (3.0), were calculated. For a comparison, the automatic and 'expert' times of the S wave's arrival at the station LJU are given in Table 1. The time difference between the manual and the automatic evaluation was 0.85 s, but the automatically defined location is still close to the real one (Figure 6).

Table 1. The automatic and the 'expert' times of the S-wave arrival at the station LJU for the earthquake in Figure 5.

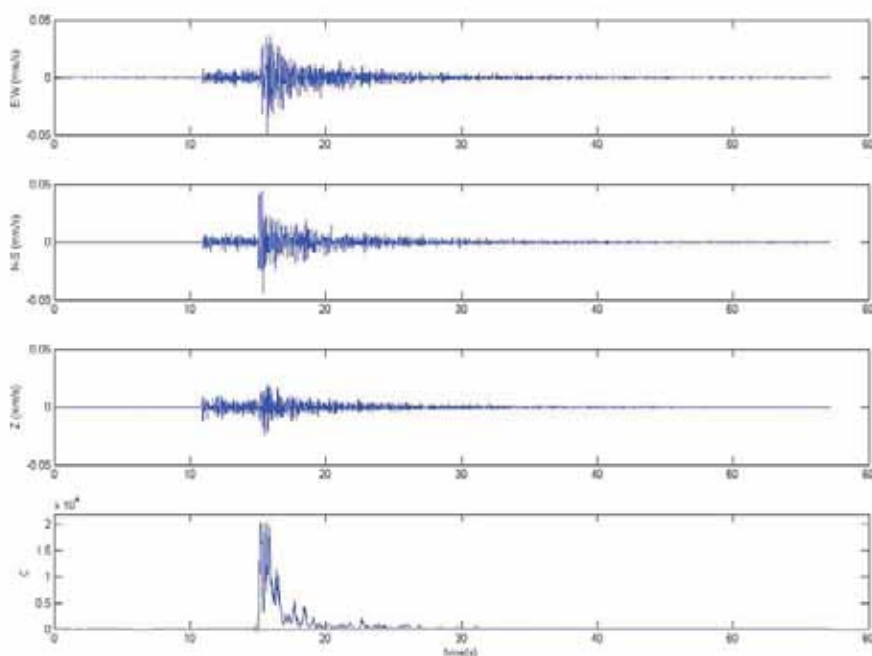|	Time (UTC)
Manual (expert)	03h 28min 05.98sec
Automatic	03h 28min 06.82sec

Figure 4. An event with magnitude M_{LV} 2.5 recorded at the three-component seismic station LJU at an epicentral distance of 29 km (E-W, N-S, Z traces). The system output is represented by a green circle, which represents the S arrival (C trace). The error between the "correct" and the estimated time difference is 0.07 s.
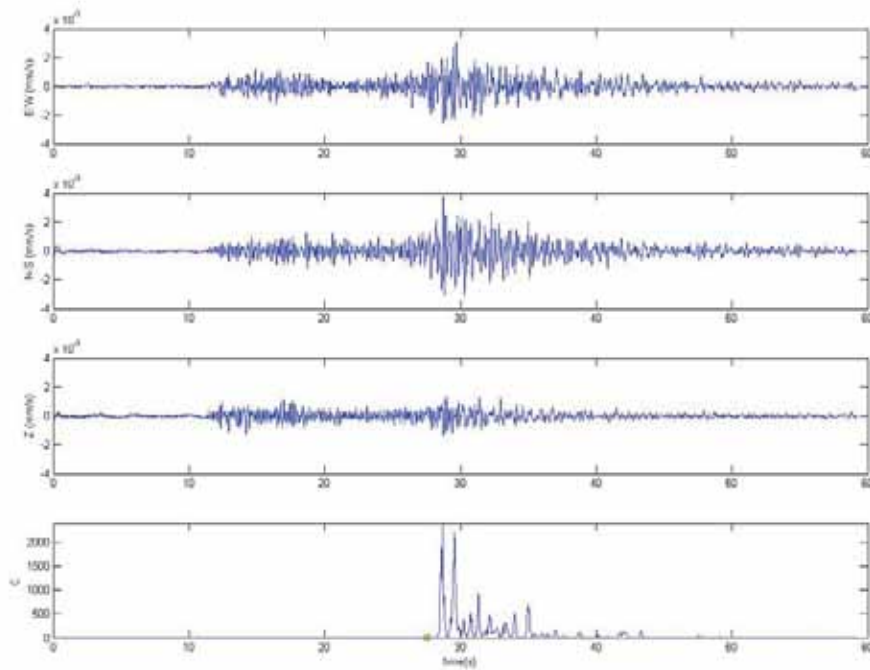
Figure 5. An event with magnitude M_{LV} 2.5 recorded at the three-component seismic station LJU at an epicentral distance of 124 km (E-W, N-S, Z traces). The system output is represented by a green circle, which represents the S arrival (C trace). The error between the "correct" and the estimated time difference is 0.09 s.
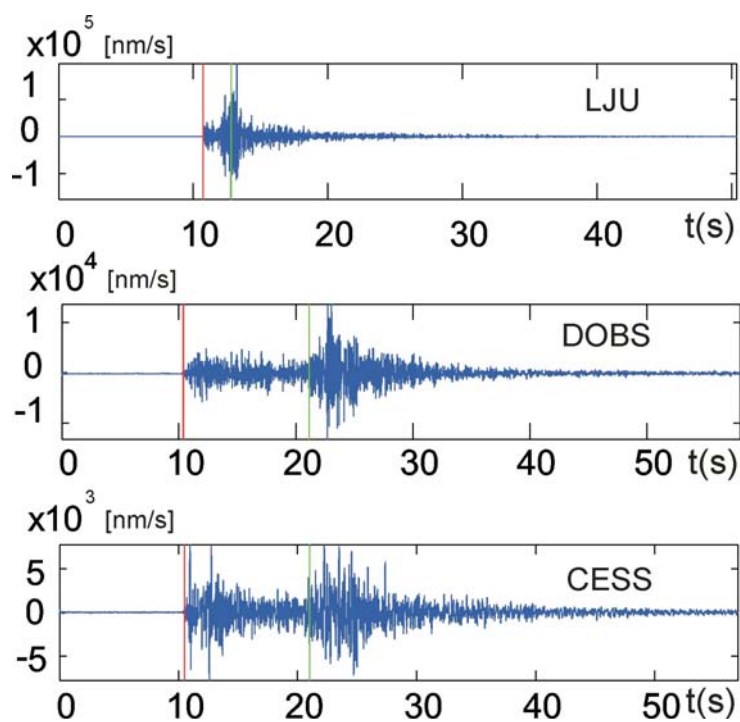

Figure 6. The record of an earthquake that occurred on 23 January 2009 at 04:28 local time. The seismograms of the vertical components were recorded at the seismic stations DOBS (Dobrina), CESS (Cesta) and LJU (Ljubljana). The arrival of the P (red line) and S (green line) waves was automatically denoted.
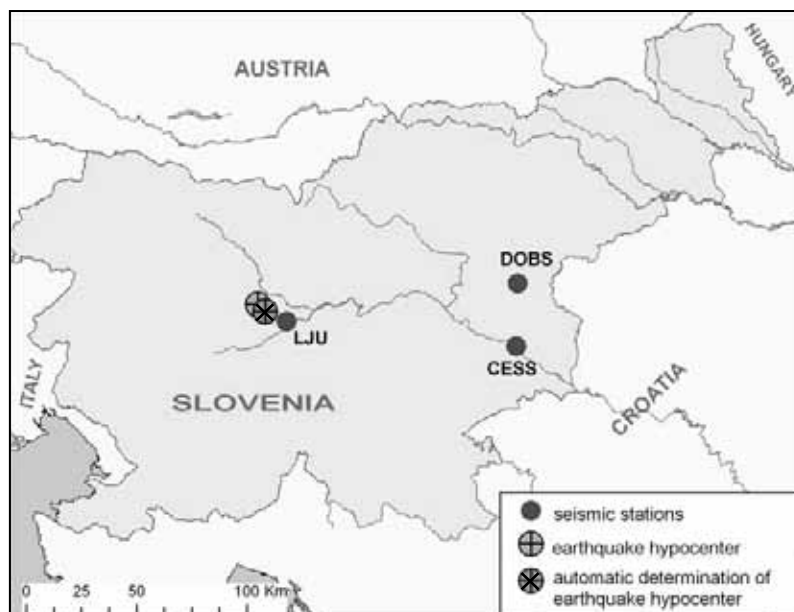
Figure 7. The map showing the locations of the seismic stations that contributed their data to the automatic determination of the earthquake's parameters, the hypocenter of the earthquake and the automatic determination of the hypocenter's location. Even with this small number of seismic stations, the automatic location is still close to the real one.

4 CONCLUSIONS

An automatic procedure for the P and S phase picking could not completely replace the human experts and their knowledge, but are very important for the primary and rapid estimation of the earthquake's parameters. The clearest phase to pick is usually the P wave, which is also the case in the automatic procedure. When a small number of seismic stations contribute their data to the automatic procedure, S-phase picking is desired. Using phases with different velocities enables a better estimation of the epicenter's location. Our procedure to estimate the S-wave arrival time has been in practical daily use in the seismic network of Slovenia since 2002. The energy analysis used in our procedure is primarily suitable for local earthquakes. A procedure, also taking the frequency content into account, should be used for earthquakes with larger epicentral distances.

REFERENCES

[1] Freiberg, W.F. (1963). An approximation method in signal detection. *Quart. J. App. Math*. 20, 373-378.

[2] Joswing, M. (1990). Pattern recognition for earthquake detection. *Bulletin of the Seismological Society of America*, 80, 170-186.

[3] Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America*, 68, 1521-1532.

[4] Houliston, D.J., Waugh, G., Lazughlin. J. (1984). Automatic Real-Time Event Detector for Seismic Networks. *Computers and Geosciences*, 10, 431-436.

[5] Bear, M., Kradofler, U. (1987). An automatic phase picker for local and teleseismic event. *Bulletin of the Seismological Society of America*, 77, 1437-1445.

[6] Pisarenko, V. F., Kushnir, A. F., Savin, I.V. (1987). Statistical adaptive algorithms for estimation of onset moments of seismic phases. *Physics of Earth and Planetary Interiors*, 47,4-10.

[7] Roberts, R. G., Christofsen, A., Casidy, F. (1989). Real time event detection, phase identification and source location estimation using single station three-component seismic data. *Geophysical Journal International*, 97, 471-480.

[8] Takanami, T., Kitagawa, G. (1993). Multivariate time-series model to estimate the arrival times of S waves. *Computers and Geosciences*, 19, 295-301.

[9] Charutti, C., Salemi, G. (1993). Artificial intelligence techniques in the analysis of digital seismograms. *Computers and Geosciences*, 19, 149-156.

[10] Joswing, M. (1995). Automated classification of local earthquake data in the BUG small array. *Geophysical Journal International*, 120, 262-286.

[11] Dai, H., MacBeth, C. (1995). Automatic picking of seismic arrivals in local earthquake data using an artificial neural network. *Geophysical Journal International*, 120, 758-774.

[12] Cichowitcz, A. (1993). An automatic S-phase picker, *Bulletin of the Seismological Society of America*, 83, 180-189.

[13] Wang, J., Teng, T. (1997). Identification and picking S phase using an artificial neural network. *Bulletin of the seismological Society of America*, 87, 1140-1149.

[14] Tasič, I. (2000). Characterization of seismic waves arrival with simulated neural networks. Faculty of Mechanical Engineering, Ljubljana 2000.

[15] Saragiotis, C.D., Hadjileontiadis, L.J., Savvaidis, A.S., Papazachos, C.B. Panas, S.M. (2000). Automatic S-phase arrival determination of seismic signals using nonlinear filtering and higher-order statistics, Geoscience and Remote Sensing Symposium, 2000. *Proceedings. IGARSS 2000. IEEE 2000 International*, Vol 1, 292-294., 24-28 July 2000.

[16] Tong, C. and Kennett, B.L.N. (1996). Automatic seismic event recognition and later phase identification for broadband seismograms. *Bull. Seism. Soc. Am.* 86, 1896–1909.

[17] Bai, C.Y., Kennett, B.L.N. (2001). Phase identification and attribute analysis of broadband seismograms at far-regional distances. *Journal of Seismology* 5, 217–231.

[18] Earle, P. (1997). The Picking Procedure http://mahi.ucsd.edu/research_projects/Autopicker/autopicker/node3.html.

[19] Douglas, A. (1997). Bandpass filtering to reduce noise on seismograms: Is there a better way? *Bull. Seism. Soc. Am*. 87, 770–777.

[20] Tasič, I. (2001). Nanometrics's Digital Seismic Network, Potresi v letu 1999 (ed. R. Vidrih), ARSO, Urad za seizmologijo, 68-72.

[21] Tasič, I. (2001). Classification Of Arrival Times Of Local Earthquake Seismic Waves, Potresi v letu 1999 (ed. R. Vidrih), ARSO, Urad za seizmologijo, 83-93.

[22] Tasič, I, (2006). Automatic Quantification Of Earthquake Parameters At A Single Seismic Station, Potresi v letu 2004 (ed. R. Vidrih), ARSO, Urad za seizmologijo, 83-93.

[23] Lay, T., Wallace, T.C. (1995). *Modern Global Seismology*. Academic Press, S. Diego, USA.

[24] Bormann P. Ed. (2002). *New Manual of Seismological Observatory Practice*. GFZ Potsdam, Germany.