Seismological Research Letters

Fast discrimination of local earthquakes using a neural approach

--Manuscript Draft--

Manuscript Number:	SRL-D-16-00222R2
Full Title:	Fast discrimination of local earthquakes using a neural approach
Article Type:	Article - Regular Section
Corresponding Author:	Flora Giudicepietro Istituto Nazionale di Geofisica e Vulcanologia Napoli, Napoli ITALY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Istituto Nazionale di Geofisica e Vulcanologia
Corresponding Author's Secondary Institution:	
First Author:	Flora Giudicepietro
First Author Secondary Information:	
Order of Authors:	Flora Giudicepietro
	Antonietta Maria Esposito
	Patrizia Ricciolino
Order of Authors Secondary Information:	
Manuscript Region of Origin:	ITALY
Suggested Reviewers:	
Opposed Reviewers:	

Fast discrimination of local earthquakes using a neural approach

F. Giudicepietro (1), A.M. Esposito (1), P. Ricciolino (1)

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

Corresponding author:

Flora Giudicepietro, via Diocleziano 328, 80124, Napoli, Italy.

Email: flora.giudicepietro@ingv.it

Abstract

In this paper, we describe a Neural Network method for the fast discrimination between local earthquakes and regional and teleseismic earthquakes by using seismic records from a single station. Neural networks are data-driven nonlinear classifiers that learn from experience and can model real world complex relationships. For the discrimination task we implement a two-layer feed-forward Multi Layer Perceptron (MLP). MLP is a supervised technique that accomplishes the learning process by using a pre-classified dataset for the training phase. The dataset includes 70 teleseisms, 79 regional earthquakes and 103 local earthquakes. The seismic events are recorded at a single station, equipped with a short period sensor. We parameterize the seismograms in the frequency domain, using the Linear Predictive Coding (LPC). This technique is mostly used in audio signal processing for efficiently encoding frequency features of digital signals in a compressed form. The obtained spectral features, or LPC coefficients, are the input to the neural model. We carry out several tests by shortening from 4 to 1 second the time window duration used for the LPC analysis. The proposed algorithm achieves a correct classification of 98.5% and 97.7% in discriminating local versus regional and local versus teleseismic earthquakes, respectively, on 1-second time window. These results indicate that our discrimination algorithm can be profitably exploited in automatic analysis of seismic data, which require fast responses, such as seismological monitoring systems and earthquake early warning systems.

Introduction

The problem of fast discrimination among different seismic events is crucial for timely communications to Civil Protection Authorities and to people exposed to natural hazards (earthquakes, tsunamis, volcanic eruptions and floods). Current technology allows to perform advanced automatic analysis to pick seismic phases (Allen 1978; Rowe et al., 2002; Lomax et al., 2012), to locate earthquakes (Johnson et al., 1995; Olivieri and Clinton, 2012; www.earthwormcentral.org/; www.seiscomp3.org), to estimate the shaking intensity in the area where the earthquake occurred (Wald et al., 1999; Earle et al., 2009), to locate volcanic seismic events (Wassermann and Ohrnberger 2001; Giudicepietro et al. 2009; De Cesare et al. 2009; Olivieri and Clinton, 2012), and VLP events (Auger et al., 2006) and to classify seismic events (Langer et al., 2006; Esposito et al., 2006; Ochoa et al., 2014; Mousavi et al., 2016; Kortström et al., 2016; Vargas et al., 2016). A critical aspect of all of these advanced automatic analyses is a fast and robust recognition of seismic event types.

In the last decades, the scientific community has made considerable efforts to develop Earthquake Early Warning (EEW) systems, aimed to warn people, allowing them to take protective actions several seconds before the arrival of potentially damaging seismic waves (Basher 2006; Böse et al., 2008; Alcik et al., 2009; Allen et al. 2009, Zollo et al. 2009; Given et al. 2014; Zollo et al. 2014). For these advanced seismic systems too, the problem of a fast and robust seismic event discrimination is critical, in order to avoid false alarms. For this reason, researchers have focused their attention on improving the robustness and reliability of fast automatic seismic analyses.

In this study we deal with discrimination among teleseisms, regional earthquakes and

53 local earthquakes, according to the traditional seismogram classification based on the
54 epicentral distance (Lay and Wallace, 1995). This discrimination is important because it
55 allows to recognize events that occurred outside the network of seismic stations, such as
56 teleseismic and distant regional earthquakes, that are among the most common causes of
57 failure of real-time earthquake systems (Lee and Stewart, 1981; Nakamura et al. 2009; Li
58 et al., 2013; Zschau and Küppers, 2013, Zollo et al. 2016; Errata for Latest Earthquakes
59 (https://www.usgs.gov/)).

60 To this aim, we propose a neural classification of seismograms recorded at a single
61 station, which can be used along with other methods based on the information produced by
62 a seismic network of many stations (e.g. Earthworm, (Johnson et al., 1995)) to attain a
63 more robust and fast earthquakes' detection and classification. In the last years, artificial
64 neural networks have been successfully used to approach seismic signal automatic
65 classification (Wang and Teng 1995; Falsaperla et al., 1996; Ezin et al., 2002; Avossa et
66 al., 2003; Scarpetta et al., 2005; Langer et al., 2006; Esposito et al., 2006a,b; Esposito et
67 al., 2008; Esposito et al., 2013a; Horstmann et al.,2013), and detect events of particular
68 interest (Del Pezzo et al., 2003; Esposito et al., 2013a,b; Wiszniowski et al., 2014). These
69 techniques require a data preprocessing, aimed at feature extraction from seismograms,
70 which typically exploits several seconds of the seismic signal. The proposed neural
71 analysis uses just a few seconds of the beginning of a seismogram, to allow a fast
72 discrimination among local, regional and teleseismic earthquakes.

73 **Dataset**

74 Our dataset includes seismograms of local earthquakes, defined as earthquakes with
75 epicentral distance <= 100 km, regional earthquakes with epicentral distance of 100 to

1400 km, and teleseisms, with large epicentral distance ($\Delta > = 30°$) (see Lay and Wallace, 1995). We use data from SGG seismic station, operated by the Osservatorio Vesuviano (INGV) that is located near San Gregorio Matese, a village situated in a seismically active area in the Southern Apennines, Italy (Fig.1). This station, equipped with a short period sensor (Geotech S-13), is operating since 1997. It has an analog system for data transmission. The data are then acquired by a 16-bit A/D converter system. Our dataset includes 70 teleseisms, 79 regional earthquakes, recorded between 1999 and 2015, and 103 local earthquakes, most of which were recorded between 29 December 2013 and 30 January 2014, when a seismic swarm occurred in that area. Data are courtesy of the Osservatorio Vesuviano (INGV) seismic Lab. In our analysis we use the vertical component. In order to apply a uniform automatic criterion for picking the onset of the seismic events, we exploit the ObsPy module for automatic P phase picking (Withers et al., 1998; Trnkoczy 2012; Krischer et al., 2015). In general, seismic automatic systems include specific modules for avoiding false triggers through the analysis of sets of automatic picking (e.g. "binder" module in the Earthworm system (Johnson et al., 1995); Zollo, 2016). For this reason, our data does not include false triggers, being their detection beyond the scope of our analysis.

Figure 2 and Figure 3 show examples of the three types of seismic events considered for the analysis, with the markers of the automatic picking. The teleseismic earthquakes recorded at SGG station are characterized by a frequency range < 3 Hz (Fig 4). The regional earthquakes show higher frequency than the teleseismic ones (Fig 4). The local earthquakes have typical waveforms, characterized by an impulsive onset, and are higher in frequency than teleseisms and regional earthquakes (Fig 2, 3 and 4). SGG station sometimes records earthquakes with magnitude greater than 3.5, occurred at a small

100 distance (<100 km). These events lead to the saturation of the seismic records and are not
101 considered in this paper.

102 **Analysis**

103 For each event, we cut an initial data window of 20 seconds. The data sampling rate
104 is 100 Hz. The starting point for the windows is the P phase picking (Fig. 3). Next, we
105 select smaller signal windows, with a variable length of 400, 200 and 100 samples, i.e.
106 from 4 to 1 second. We divide the time window of 400 samples (4 seconds) in two
107 windows of 200 samples (2 seconds). Then, we apply the Linear Predictive Coding (LPC)
108 technique (Makhoul, 1975). This technique provides a compact data encoding that is
109 suitable for reducing the input dimension to the neural network. In particular, the LPC
110 models the signal spectrum in the frequency domain with an all-pole filter (Del Pezzo et
111 al., 2003), which compresses spectral information. In seismology, the spectrogram, that is
112 based on the Fourier transform, is typically used to characterize the signals, but it does not
113 allow to obtain a compact representation as input to the neural network. The linear
114 prediction is a very efficient technique for eliminating correlation and redundancy from a
115 signal. It is typically used in audio signal processing and speech analysis for representing
116 the spectral envelope of a speech signal in a compact way. The LPC algorithm models each
117 signal window s_n as a linear combination of its p previous samples as follows (Equation 1):

118 $$s_n^* = \sum_{k=1}^{p} c_k s_{n-k} + G \qquad (1)$$

119 where s_n is the signal at time n, s^*_n is the LPC estimated or predicted signal, p is the
120 model order that is problem dependent, c_k, $k = 1,.., p$, are the prediction coefficients and G
121 is the gain. The c_k estimation is obtained through a procedure that minimizes the error

between the true signal and its LPC estimate. This procedure computes the following misfit function (Equation 2):

$$E(c)=\sum_{n}(s(n)-s*(n))^2 \qquad (2)$$

where c is the vector of the c_k prediction coefficients. The c_k estimation is not time-consuming (Vaidyanathan, P. 2007) therefore it can be performed in real time.

In our experiments, we extract p = 14 coefficients for each signal window of 1 and 2 seconds (100 and 200 samples), in order to maximize the data compression and minimize the corresponding error. So, for each of these time windows we obtain a vector of 14 spectral features (LPC coefficients). Furthermore, the time window of 4 seconds is divided in two sub-windows of 2 seconds, and for each of them 14 LPC coefficients are extracted. So, the 4 second window is encoded by a vector of 14+14=28 spectral features. Figures 5, 6 and 7 show the feature vectors obtained for signal windows of 1, 2 and 4 seconds, respectively.

Method

To discriminate among the different typologies of signals we use a Multi-Layer Perceptron (MLP) network (Bishop, 1995; Haykin, S., 1999). MLP networks are neural supervised techniques, meaning that the learning process is realized by training the net on a pre-labelled subset of data. The MLP architecture (Fig. 8) presents an input layer, one or more hidden layers, and an output layer. In our case, the input layer is the vector of the LPC coefficients computed for each signal, while the output layer provides the network response for each of the three examined classifications, i.e. Regional/Local, Teleseisms/Local and Teleseisms/Regional. There are neither intra-state not feedback

connections in the network so the input signal (x_1, ..., x_n), propagates from the input layer to the output one in a forward direction (i.e. feed-forward). MLP networks have been widely used for function approximation, pattern classification and recognition due to their structural simplicity and fast learning abilities.

The dataset is divided into two subsets, one for training and the other one for testing the neural network. In particular, we use 5/8 of the available dataset for the training phase and the remaining 3/8 for the testing one. In this way, the testing set is large enough to evaluate the network performance. Table 1 shows the adopted data distribution for training and testing set in each classification task.

The setting of the net parameters (e.g. the initial weights on the connections, the number of hidden nodes, the activation functions, the learning algorithm, the number of learning cycles) is tuned on the basis of previous works (Esposito et al., 2006b; Giudicepietro et al., 2008; Esposito et al., 2016). Moreover, we evaluate the Final Prediction Error (FPE) function (Equation 3) (Akaike, 1970; Akaike, 1974):

$$FPE = s^2(p)\frac{N+p+1}{N-p-1} \qquad (3)$$

where N is the number of samples and $s^2(p)$ is the prediction error. $(N+p+1)/(N-p-1)$ increases with p and represents the inaccuracies in estimating the prediction parameters (Esposito et al. 2006b). We use a trial and error procedure to choose the appropriate number of LPC coefficients, in order to represent the envelope of the signal spectrum in a compressed form (Fig. 9). Therefore, we use five hidden nodes for the MLP architecture and, as node activation functions, the hyperbolic-tangent for the hidden units and the logistic sigmoidal for the output node.

166 The Quasi-Newton algorithm (Bishop, 1995; Dennis et al., 1983) is chosen for the
167 weight optimization and the Cross-Entropy Error Function (Bishop, 1995) is used as error
168 function for the output. The combined use of this algorithm with this function allows a
169 probabilistic interpretation of the net response. We carry out 54 classification experiments,
170 with a permutation of the training and testing subsets, to verify the stability of the network
171 output, as described in the next section.

172 The network training task takes less than one second on a PC with standard
173 configuration. Once the network is trained, the recognition of a single event takes
174 insignificant time (less than a hundredth of a second).

175 **Results**

176 Table 2 summarizes the results of our analyses, showing the net performances for the
177 three pairs of signals, namely Regional/Local, Teleseisms/Local and Teleseisms/Regional,
178 for a total of 54 classification experiments. For each classification task (Reg/Loc, Tel/Loc
179 and Tel/Reg), we consider signal windows of the onset of the earthquakes with decreasing
180 lengths (4 seconds, 2 seconds and 1 second). Each row in Table 2 reports the net
181 classification accuracy (in percentage) obtained on six different permutations of the
182 training and testing sets, and six random initial configuration of the net weights. Finally,
183 the last column shows the average performance computed on each row. In Table 2, we can
184 observe that good performance are obtained for Regional/Local and Teleseism/Local pair.
185 In particular, an average of 98.53% of correct classification is achieved in the
186 Regional/Local pair experiment, using only one second of the signal. For the
187 Teleseism/Local pair an average of 97.69% correct classifications is obtained using one
188 second of the signal and an average of 99.49% by using the first four seconds of signal.

189 The worst performances are those of the Teleseismic/Regional pair, showing only 61% of
190 correct classification on a one second signal window. This performance increases up to
191 78.27% of correct classifications when the first four seconds of the signal are considered,
192 but it is still significantly lower than that of the other two classification tasks (Reg/Loc and
193 Tel/Loc). In our application, the short-period station introduces a high-pass filter on the
194 signal, due to the instrumental response. For this reason the regional earthquakes and the
195 teleseisms are not sharply separated.

196 **Conclusions**

197 The goal of this work is to provide a reliable neural network algorithm, which works
198 on signal windows as short as possible, to allow an early identification of the different
199 types of earthquakes and detect events outside the seismic network. In our case, the events
200 outside the seismic network are regional and teleseismic earthquakes. The results of the
201 proposed automatic neural classification are encouraging, as good performance are
202 preserved on Regional/Local and Teleseism/Local classification by shortening the signal
203 window duration from 4 seconds up to 1 second. The method can be used with the signals
204 of any short-period station, however, to achieve optimal performance, the neural network
205 should be trained with examples of earthquakes recorded at that station, in order to realize
206 a specialized detector (Scarpetta et al. 2005). For broadband sensors, which ensure better
207 quality of the teleseismic and regional seismograms, the spectral encoding does not allow
208 very small time window (e.g. window length < T, where T is the period of the dominant
209 components of the first P-wave arrival). However, we can not exclude that, by using other
210 parameterization techniques (D'Auria et al. 2006; Colombelli and Zollo, 2016), it is
211 possible to overcome this problem.

In our experiments, we are able to discriminate between local and regional earthquakes, and local earthquakes and teleseisms, using only the first second of the seismograms, with a percentage of correct classifications of 98.5% and 97.7% respectively. The majority of local earthquakes belong to a seismic sequence occurred between December 2013 and January 2014. Nonetheless, the neural network has correctly classified also the local earthquakes that do not belong to this seismic sequence. The seismic signal preprocessing is based on a speech signal processing technique, which allow effective data compression while maintaining the spectral information. Once network is trained, the computation time for the classification is negligible. Therefore, we can obtain a robust discrimination on very short windows of the earthquake onset signal, in real time. The analysts typically recognize different types of seismic events through visual analysis of the waveforms. However, even the most experienced analysts can hardly distinguish in real time among the three examined types of events, just considering the first second of the seismograms.

This type of application of neural networks for the classification of seismic signals allows a profitable exploitation of old analog seismic stations, which have the advantage to provide a large number of regional earthquake and teleseism recordings. This neural method, in conjunction with real time earthquake analyses performed by advanced dense seismic networks, can improve the reliability of seismological monitoring systems. In particular, this technique can detect earthquakes that are located outside the seismic network. Furthermore, due to its very rapid response, the neural algorithm, with an appropriate tuning, can be also suitable for EEW systems, to reduce the incidence of false alarms and the resulting erroneous communications to the public.

Data & Resources

Data are courtesy of Osservatorio Vesuviano, Istituto Nazionale di Geofisica e Vulcanologia, Italy (www.ov.ingv.it).

Acknowledgments

We thank the technical and technological staff that takes care of the seismic network and the data analysis Lab. at Osservatorio Vesuviano, Istituto Nazionale di Geofisica e Vulcanologia (INGV). We thank Anna Esposito and Novella Tedesco for their precious help with the English language improvements.

245

246 **References**

247 Akaike, H. (1970). Statistical predictor identification, Ann. Inst. Statist. Math., 22,
248 202-217.

249 Akaike, H. (1974). A new look at the statistical model identification, IEEE Trans.
250 Automat. Control, 19, 716-723.

251 Alcik, H., Ozel, O., Apaydin, N., and Erdik, M. (2009). A study on warning
252 algorithms for Istanbul earthquake early warning system. Geophysical Research Letters,
253 36(5). DOI: 10.1029/2008GL036659.

254 Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces,
255 Bull. Seismol. Soc. Am. 68, 1521–1532.

256 Allen, R. M., Gasparini, P., Kamigaichi, O. (2009). New methods and applications of
257 earthquake early warning. Geophysical Research Letters, 36, L00B05. doi:
258 10.1785/gssrl.80.5.682.

259 Auger E., D'Auria L., Martini M., Chouet B., Dawson P. (2006). Real-time
260 monitoring and massive inversion of source parameters of very-long-period seismic
261 signals: an application to Stromboli Volcano, Italy Geophysical Research Letters, 33
262 (2006), p. L04301 http://dx.doi.org/10.1029/2005GL024703.

263 Avossa, C., Giudicepietro, F., Marinaro, M., & Scarpetta, S. (2003). Supervised and
264 unsupervised analysis applied to Strombolian EQ. In Neural Nets (pp. 173-178). Springer
265 Berlin Heidelberg. Doi:10.1007/978-3-540-45216-4_19.

266 Basher, R. (2006). Global early warning systems for natural hazards: systematic and
267 people-centred. Philosophical Transactions of the Royal Society of London A:
268 Mathematical, Physical and Engineering Sciences, 364(1845), 2167-2182. DOI:
269 10.1098/rsta.2006.1819.

270 Bishop, C. (1995). Neural Networks for Pattern Recognition, p. 500. Oxford
271 University Press, New York

272 Böse, M., Wenzel, F., & Erdik, M. (2008). PreSEIS: A neural network-based
273 approach to earthquake early warning for finite faults. Bulletin of the Seismological
274 Society of America, 98(1), 366-382. doi: 10.1785/0120070002.

275 Colombelli, S., & Zollo, A. (2016). Rapid and reliable seismic source
276 characterization in earthquake early warning systems: current methodologies, results, and
277 new perspectives. Journal of Seismology, 20(4), 1171-1186.

278 De Cesare, W., Orazi, M., Peluso, R., Scarpato, G., Caputo, A., D'Auria, L.,
279 Giudicepietro, F., Martini, M., Buonocunto, C., Capello, M., and Esposito, A. M. (2009).
280 The broadband seismic network of Stromboli volcano, Italy. Seismological Research
281 Letters, 80(3), 435-439. doi: 10.1785/gssrl.80.3.435.

282 D'Auria, L., Giudicepietro, F., Martini, M., & Peluso, R. (2006). Seismological
283 insight into the kinematics of the 5 April 2003 vulcanian explosion at Stromboli volcano
284 (southern Italy). Geophysical research letters, 33(8).

285 Del Pezzo, E., Esposito, A., Giudicepietro, F., Marinaro, M., Martini, M., &
286 Scarpetta, S. (2003). Discrimination of earthquakes and underwater explosions using
287 neural networks. Bulletin of the Seismological Society of America, 93(1), 215-223. doi:

10.1785/0120020005.

Dennis, J.E., Schnabel, R.B. (1983), Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Earle, P. S., D. J. Wald, K. S. Jaiswal, T. I. Allen, K. D. Marano, A. J. Hotovec, M. G. Hearne, and J. M. Fee (2009). Prompt Assessment of Global Earthquakes for Response (PAGER): A system for rapidly determining the impact of global earthquakes worldwide, U.S. Geol. Surv. Open-File Report 2009-1131.

Esposito, A. M., D'Auria, L., Giudicepietro, F., Caputo, T., & Martini, M. (2013a). Neural analysis of seismic data: applications to the monitoring of Mt. Vesuvius. Annals of Geophysics. doi:10.4401/ag-6452.

Esposito, A. M., D'Auria, L., Giudicepietro, F., Peluso, R., & Martini, M. (2013b). Automatic recognition of landslides based on neural network analysis of seismic signals: an application to the monitoring of Stromboli volcano (Southern Italy). Pure and Applied Geophysics, 170(11), 1821-1832. Doi:10.1007/s00024-012-0614-1

Esposito, A. M., Giudicepietro, F., D'Auria, L., Scarpetta, S., Martini, M. G., Coltelli, M., & Marinaro, M. (2008). Unsupervised neural analysis of very-long-period events at Stromboli volcano using the self-organizing maps. Bulletin of the Seismological Society of America, 98(5), 2449-2459. doi: 10.1785/0120070110.

Esposito, A. M., Giudicepietro, F., Scarpetta, S., D'Auria, L., Marinaro, M., & Martini, M. (2006a). Automatic discrimination among landslide, explosion-quake, and microtremor seismic signals at Stromboli Volcano using neural networks. Bulletin of the Seismological Society of America, 96(4A), 1230-1240. doi: 10.1785/0120050097.

310 Esposito, A. M., Giudicepietro, F., Scarpetta, S. and Khilnani, S. (2016). A Neural
311 Approach for Hybrid Events Discrimination at Stromboli Volcano, to be published in
312 “Multidisciplinary approaches to neural computing”, Springer Series Smart Innovation,
313 Systems and Technologies, 2016.

314 Esposito, A. M., Scarpetta, S., Giudicepietro, F., Masiello, S., Pugliese, L., &
315 Esposito, A. (2006b). Nonlinear exploratory data analysis applied to seismic signals. In
316 Neural Nets (pp. 70-77). Springer Berlin Heidelberg. Doi:10.1007/11731177_11.

317 Ezin, E. C., Giudicepietro, F., Petrosino, S., Scarpetta, S., & Vanacore, A. (2002).
318 Automatic discrimination of earthquakes and false events in seismological recording for
319 volcanic monitoring. In Neural Nets (pp. 140-145). Springer Berlin Heidelberg.
320 Doi:10.1007/3-540-45808-5_15.

321 Falsaperla, S., Graziani, S., Nunnari, G., & Spampinato, S. (1996). Automatic
322 classification of volcanic earthquakes by using multi-layered neural networks. Natural
323 Hazards, 13(3), 205-228.

324 Giudicepietro, F., D'Auria, L., Martini, M., Caputo, T., Peluso, R., De Cesare, W.,
325 Orazi, M., Scarpato, G. (2009). Changes in the VLP seismic source during the 2007
326 Stromboli eruption. Journal of Volcanology and Geothermal Research, 182(3), 162-171.

327 Giudicepietro, F., Esposito, A., D’Auria, L., Martini, M., Scarpetta, S. (2008).
328 Automatic Analysis of Seismic Data by using Neural Networks: Applications to Italian
329 Volcanoes. In: Warner Marzocchi and Aldo Zollo (Eds.): “Conception, verification, and
330 application of innovative techniques to study active volcanoes”, Copyright © 2008 Istituto
331 Nazionale di Geofisica e Vulcanologia, pp. 399-415.

Given D. D., Cochran E. S., Heaton T., Hauksson E., Allen R., Hellweg P., Vidale J. and Bodin P. (2014). Technical implementation plan for the ShakeAlert production system: an Earthquake Early Warning system for the West Coast of the United States. U.S. Geological Survey, Open-File Report 2014-1097 DOI: 10.3133/ofr20141097.

Haykin, S. (1999) – Neural Networks: A Comprehensive Foundation, 2nd ed. (Englewood Cliffs, NJ: Prentice-Hall).

Horstmann, T., Harrington, R. M., & Cochran, E. S. (2013). Semiautomated tremor detection using a combined cross- correlation and neural network approach. Journal of Geophysical Research: Solid Earth, 118(9), 4827-4846.

Johnson, C. E., A. Bittenbinder, B. Bogaert, L. Dietz, and W. Kohler (1995). Earthworm: a flexible approach to seismic network processing, IRIS Newsl. 14, no. 2, 1–4.

Kortström, J., Uski, M., & Tiira, T. (2016). Automatic classification of seismic events within a regional seismograph network. Computers & Geosciences, 87, 22-30.

Krischer L., Megies T., Barsch R., Beyreuther M., Lecocq T., Caudron C. and Wassermann J., (2015). ObsPy: a bridge for seismology into the scientific Python ecosystem. Computational Science & Discovery, 8, 014003, doi:10.1088/1749-4699/8/1/014003

Langer, H., Falsaperla, S., Powell, T., and Thompson, G. (2006). Automatic classification and a-posteriori analysis of seismic event identification at Soufriere Hills volcano, Montserrat. Journal of volcanology and geothermal research, 153(1), 1-10. doi:10.1016/j.jvolgeores.2005.08.012.

353 Lay, T., & Wallace, T. C. (1995). Modern global seismology (Vol. 58). Academic
354 press.

355 Lee, W. H. K., & Stewart, S. W. (1981). Principles and applications of
356 microearthquake networks (Vol. 2). Academic press.

357 Li, J., Jin, X., Zhang, H., & Wei, Y. (2013). Comparison of two earthquake early
358 warning location methods. Earthquake Science, 26(1), 15-22.

359 Lomax, A., Satriano, C., & Vassallo, M. (2012). Automatic picker developments and
360 optimization: FilterPicker—a robust, broadband picker for real-time seismic monitoring
361 and earthquake early warning. Seismological Research Letters, 83(3), 531-540. doi:
362 10.1785/gssrl.83.3.531.

363 Makhoul, T. (1975) Linear Prediction: a Tutorial Review. Proceeding of IEEE, 561-
364 580.

365 Mousavi, S. M., Stephen, P. H., Langston, C. A., & Samei, B. (2016). Seismic
366 features and automatic discrimination of deep and shallow induced-microearthquakes
367 using neural network and logistic regression. Geophysical Journal International, ggw258.

368 Nakamura, H., Horiuchi, S., Wu, C., Yamamoto, S., & Rydelek, P. A. (2009).
369 Evaluation of the real- time earthquake information system in Japan. Geophysical
370 Research Letters, 36(5).

371 Ochoa, L. H., Niño, L. F., & Vargas, C. A. (2014). Severity Classification of a
372 Seismic Event based on the Magnitude-Distance Ratio Using Only One Seismological
373 Station. Earth Sciences Research Journal, 18(2), 115-122.

374 Olivieri, M., & Clinton, J. (2012). An almost fair comparison between Earthworm
375 and SeisComp3. Seismological Research Letters, 83(4), 720-727.

376 Rowe, C. A., Aster, R. C., Borchers, B., & Young, C. J. (2002). An automatic,
377 adaptive algorithm for refining phase picks in large seismic data sets. Bulletin of the
378 Seismological Society of America, 92(5), 1660-1674. doi:10.1785/0120010224.

379 Scarpetta, S., Giudicepietro, F., Ezin, E. C., Petrosino, S., Del Pezzo, E., Martini, M.,
380 & Marinaro, M. (2005). Automatic classification of seismic signals at Mt. Vesuvius
381 volcano, Italy, using neural networks. Bulletin of the Seismological Society of America,
382 95(1), 185-196. doi: 10.1785/0120030075.

383 Trnkoczy, A. (2012), Understanding and parameter setting of STA/LTA trigger
384 algorithm, in New Manual of Seismological Observatory Practice 2 (NMSOP-2), IS 8.1, 20
385 pp.

386 Vaidyanathan, P. P. (2007). The theory of linear prediction. Synthesis lectures on
387 signal processing, 2(1), 1-184.

388 Vargas, M. G., Rueda, J., Blanco, R. M. G., & Mezcua, J. (2016). A Real-Time
389 Discrimination System of Earthquakes and Explosions for the Mainland Spanish Seismic
390 Network. Pure and Applied Geophysics, 1-16.

391 Wald, D. J., V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, and C. B.
392 Worden (1999). TriNet ShakeMaps: Rapid generation of peak ground motion and intensity
393 maps for earthquakes.

394 Wang, J.&Teng, T.L., 1995. Artifcial neural network-based seismic detector, Bull.

seism. Soc. Am., 85(1), 308–319.

Wassermann, J., and Ohrnberger, M. (2001). Automatic hypocenter determination of volcano induced seismic transients based on wavefield coherence—an application to the 1998 eruption of Mt. Merapi, Indonesia. Journal of volcanology and geothermal research, 110(1), 57-77.doi:10.1016/S0377-0273(01)00200-1.

Wiszniowski, J., Plesiewicz, B., & Trojanowski, J. (2014). Application of real time recurrent neural network for detection of small natural earthquakes in Poland. Acta Geophysica, 62(3), 469-485. Doi: 10.2478/s11600-013-0140-2.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., & Trujillo, J. (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection. Bulletin of the Seismological Society of America, 88(1), 95-106.

Zollo, A. (2016, April). A new tool for rapid and automatic estimation of earthquake source parameters and generation of seismic bulletins. In EGU General Assembly Conference Abstracts (Vol. 18, p. 15638).

Zollo, A., Colombelli, S., Elia, L., Emolo, A., Festa, G., Iannaccone, G., Martino, C. & Gasparini, P. (2014). An integrated regional and on-site Earthquake Early Warning system for southern Italy: concepts, methodologies and performances. In Early Warning for Geological Disasters (pp. 117-137). Springer Berlin Heidelberg. Advanced Technologies in Earth Sciences, DOI: 10.1007/978-3-642-12233-0_7.

Zollo A., Iannaccone G., Lancieri M., Cantore L., Convertito V., Emolo A., Festa G., Gallovicˇ F., Vassallo M., Martino C., Satriano C., and Gasparini P. (2009) Earthquake early warning system in southern Italy: Methodologies and performance evaluation. GRL

417 36, L00B07, doi:10.1029/2008GL036689.

418 Zschau, J., & Küppers, A. N. (Eds.). (2013). Early warning systems for natural

419 disaster reduction. Springer Science & Business Media.

420

421 **Tables**

422 Table 1. Training and testing data sets used for each type of classification task.

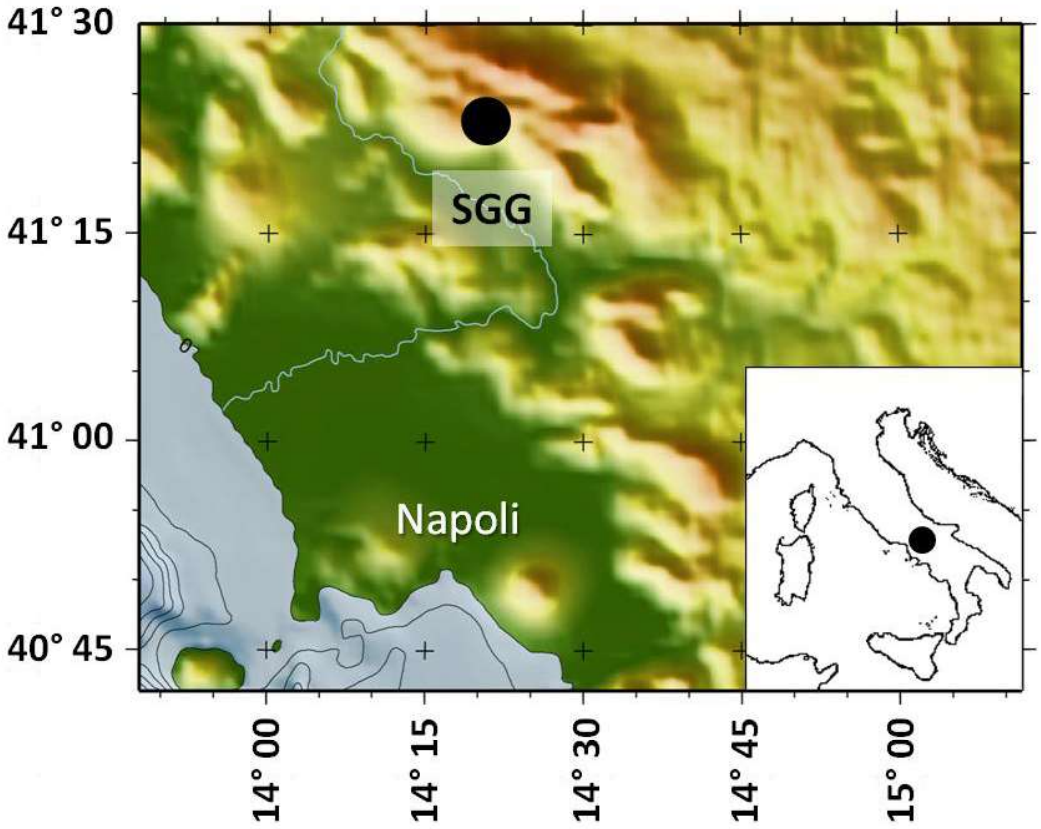
TASKs	Training (5/8)	Testing (3/8)	Total
Regional-local	114	68	182
Teleseismic-Local	108	65	173
Regional-Teleseismic	93	56	149

423

Table 2. MLP Classification performances. For each classification task, the first column reports the signal window duration (in seconds), the corresponding number of samples and the number of the neural network input nodes. We performed six different classification experiments for each window length, varying the set of events used for the network training and testing phase (6 permutations of the training and testing datasets, e.g. T1#1 … T1#6 for Regional/ Local classification task). The last column reports the average performance value obtained on each signal length under examination.

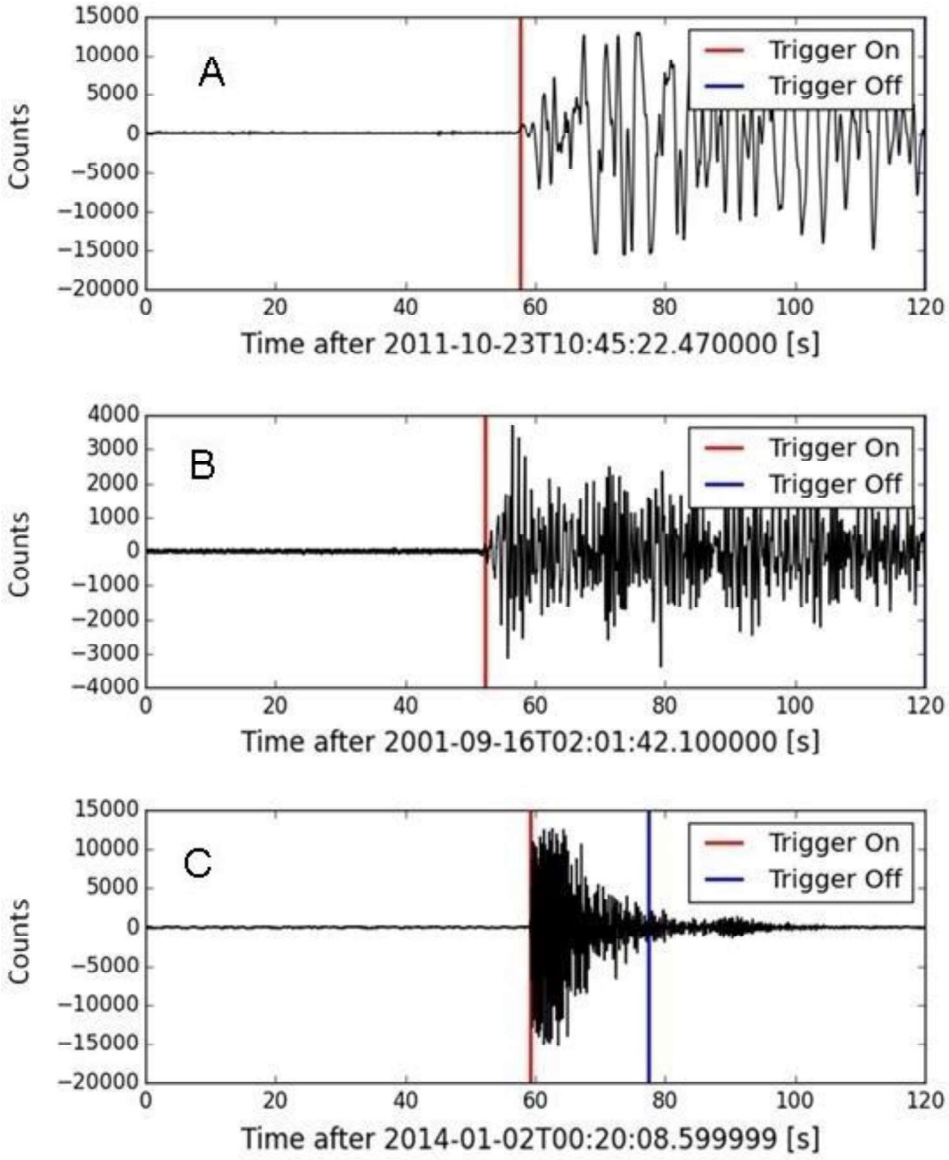
Regional – Local	**%Performance**						
Sec; Npt; input node	**T1#1**	**T1#2**	**T1#3**	**T1#4**	**T1#5**	**T1#6**	**Average**
4s; 400 samples; 28 input	98.52	100.00	100.00	100.00	97.05	98.52	99.02
2s; 200 samples; 14 input	97.05	100.00	100.00	100.00	92.64	98.52	98.04
1s; 100 samples; 14 input	98.52	97.05	100.00	100.00	95.58	100.00	98.53
Teleseism – Local	**%Performance**						
Sec; Npt; input node	**T2#1**	**T2#2**	**T2#3**	**T2#4**	**T2#5**	**T2#6**	**Average**
4s; 400 samples; 28 input	100.00	98.46	100.00	100.00	98.46	100.00	99.49
2s; 200 samples; 14 input	100.00	98.46	96.92	96.92	98.46	96.92	97.95
1s; 100 samples; 14 input	100.00	96.92	95.38	95.38	98.46	100.00	97.69
Regional – Teleseism	**%Performance**						
Sec; Npt; input node	**T3#1**	**T3#2**	**T3#3**	**T3#4**	**T3#5**	**T3#6**	**Average**
4s; 400 samples; 28 input	78.57	78.57	76.78	82.14	71.42	82.14	78.27
2s; 200 samples; 14 input	73.21	75.00	71.42	71.42	71.42	67.85	71.72
1s; 100 samples; 14 input	56.35	60.71	58.92	58.92	66.07	66.07	61.17

435 **Figures**



436

437 Fig. 1 Location of SGG Station.



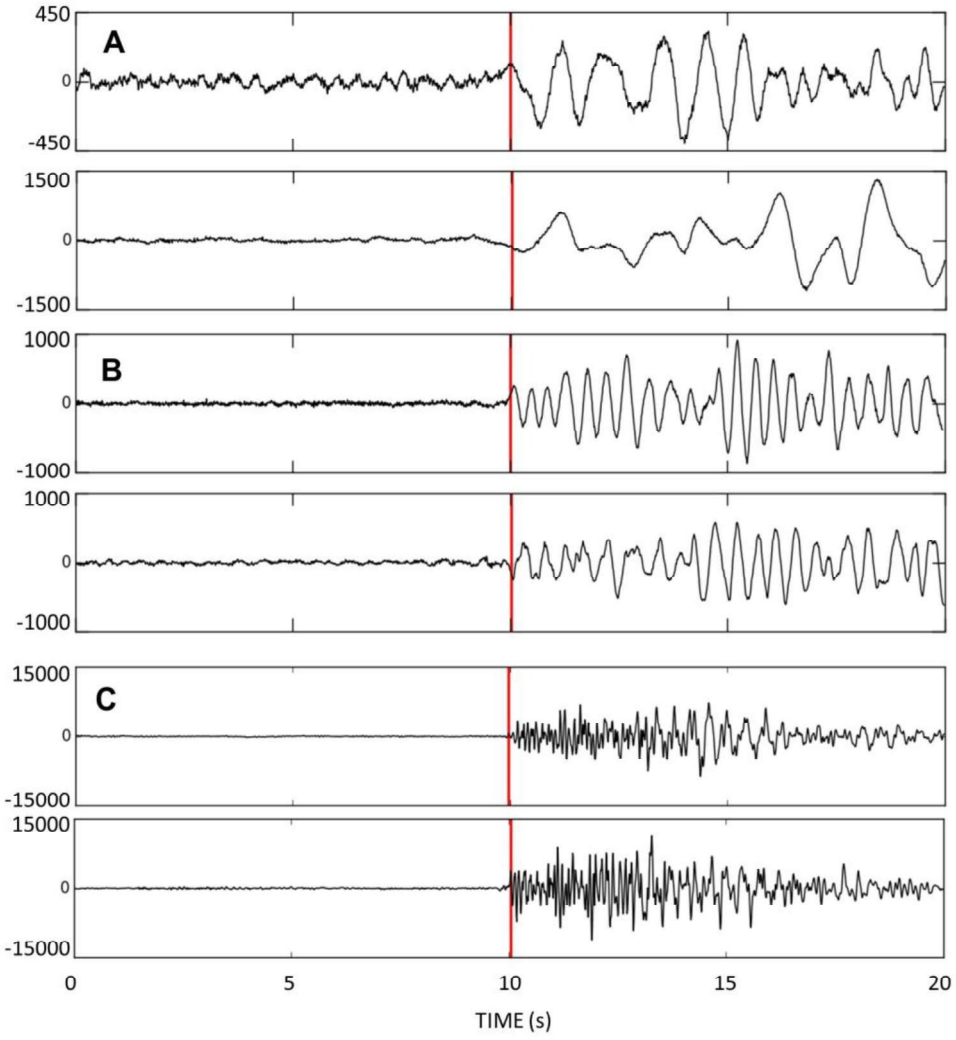
438

439 Fig. 2 Waveforms of a teleseism (panel A), a regional earthquake (panel B) and a

440 local earthquake (panel C). The red markers indicate the P wave onset obtained by using

441 ObsPy automatic picker. In blue the "trigger off" marker.

442



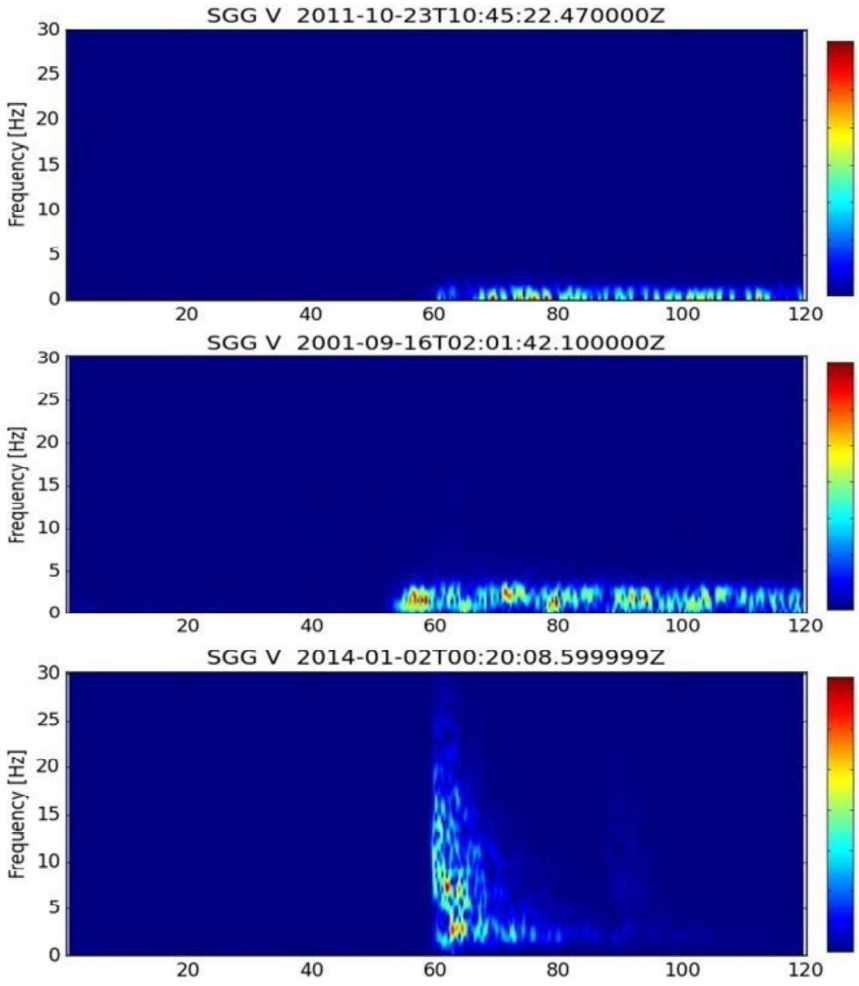
443

444 Fig. 3 Onset examples of two teleseisms (panel A), two regional earthquakes (panel

445 B) and two local earthquakes (panel C). The red markers indicate the P wave onset

446 obtained by using ObsPy automatic picker.

447



448

449 Fig. 4 Spectrograms of a teleseism (panel A), a regional earthquake (panel B) and a

450 local earthquake (panel C).

451

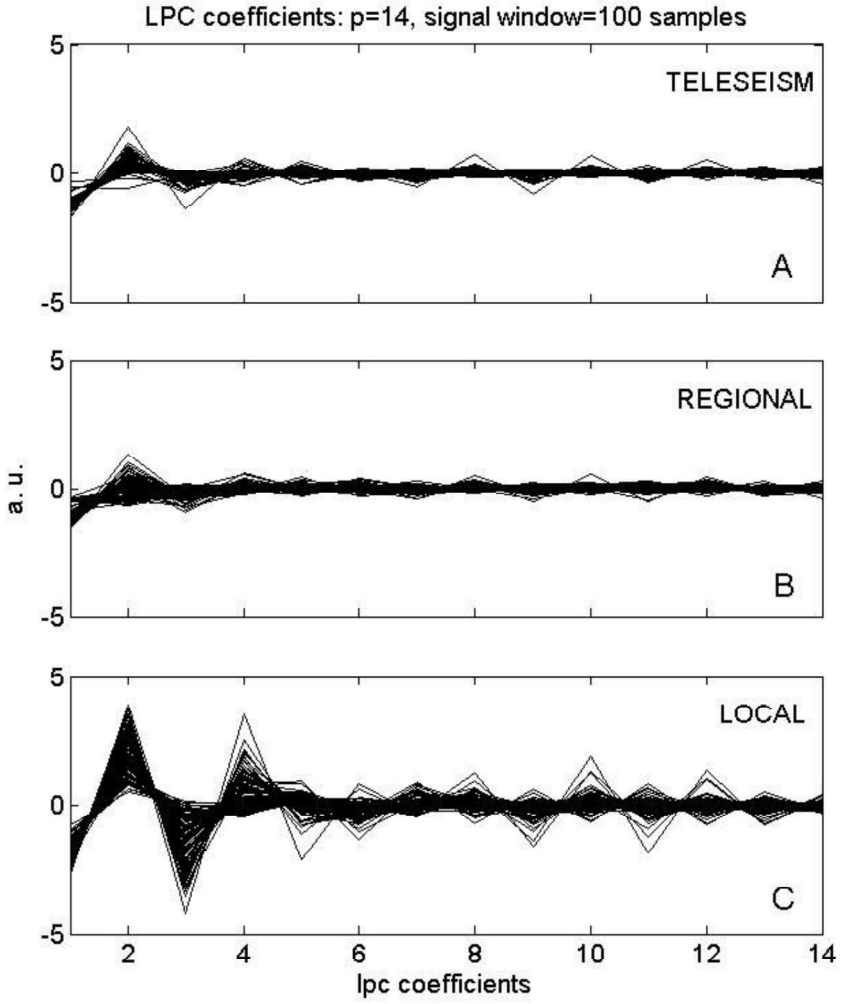


Fig. 5 Parameterization of one second of the seismic event onset. Teleseisms (panel A), regional earthquakes (panel B) and local earthquakes (panel C).

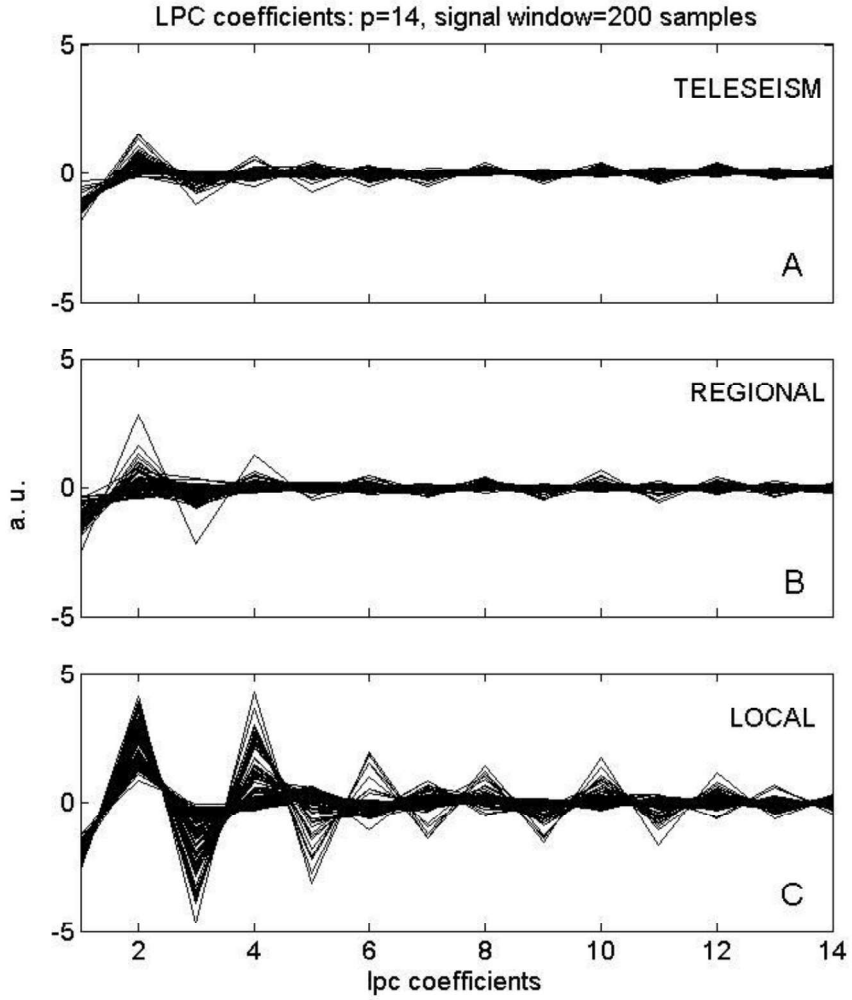
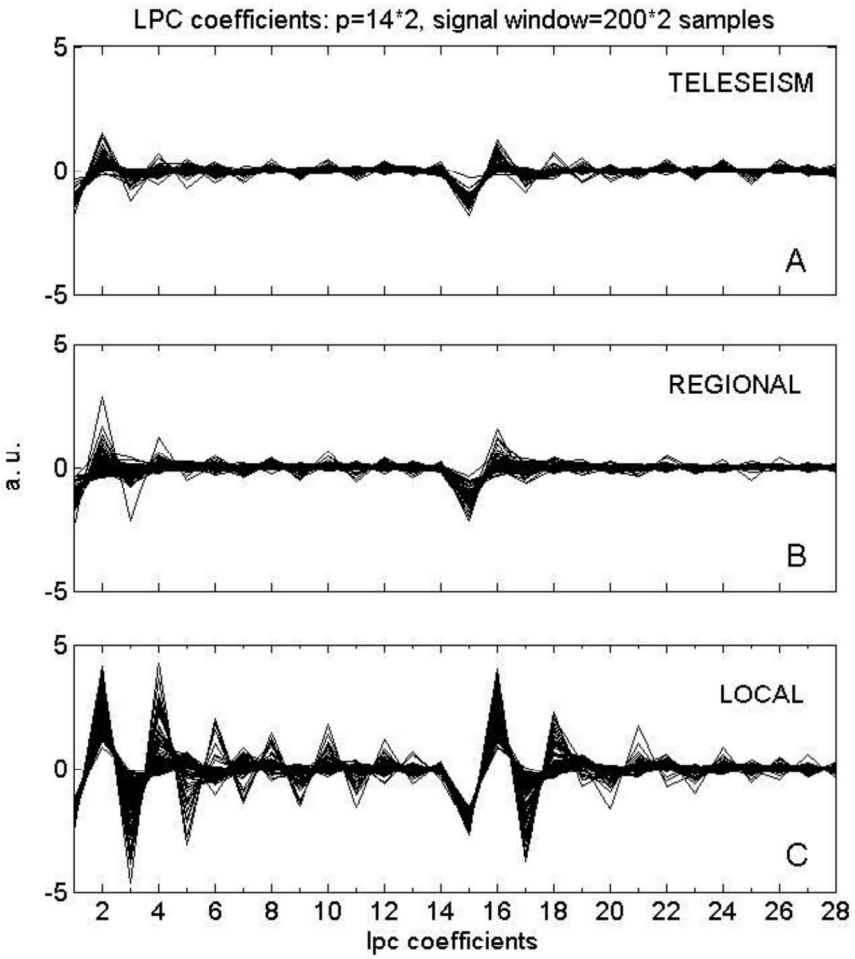


Fig. 6 Parameterization of 2 seconds of the seismic event onset. Teleseisms (panel A), regional earthquakes (panel B) and local earthquakes (panel C).



461
462

463 Fig. 7 Parameterization of 4 seconds of the seismic event onset. Teleseisms (panel

464 A), regional earthquakes (panel B) and local earthquakes (panel C).

465

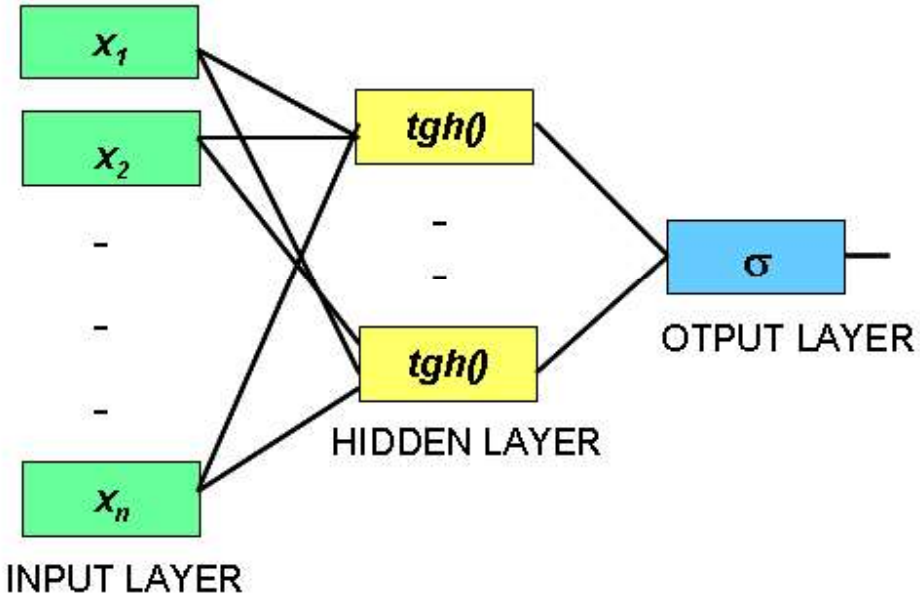


Fig. 8 An example of two layer MLP architecture: the input vector (x1, ..., xn), i.e. the LPC coefficients, moves forward from the hidden layer to the output one, that is one three examined classification tasks, i.e. Regional/Local, Teleseisms/Local and Teleseisms/Regional. There are not cycles and cross-connections between the layers.

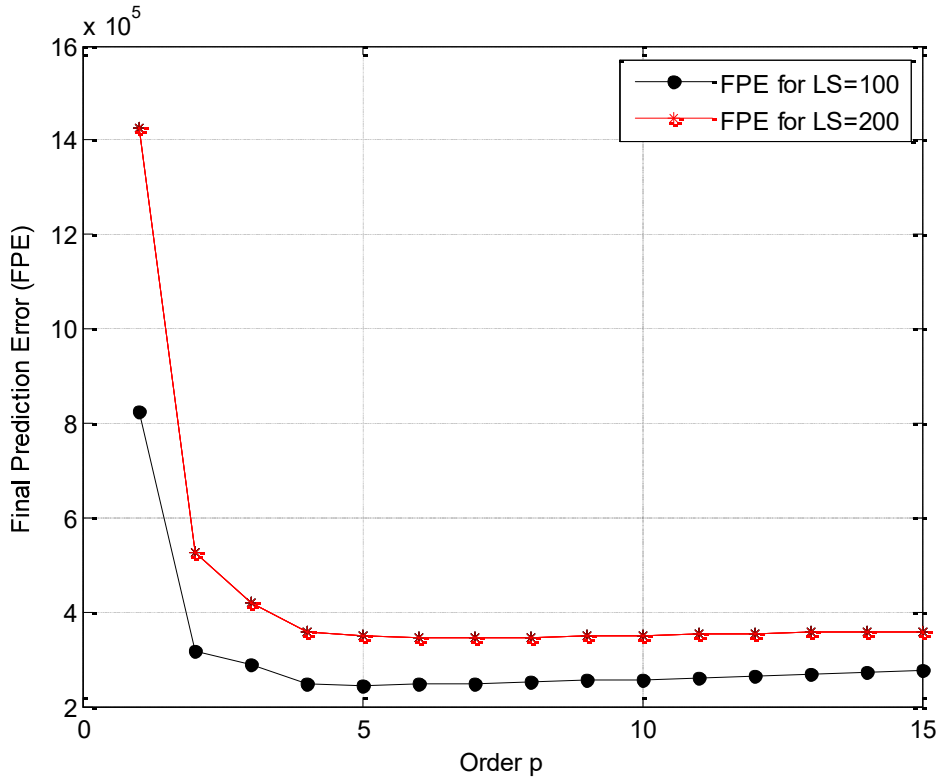


Fig. 9 Final Prediction Error of the linear prediction coding evaluated for signal windows of 100 (black line) and 200 (red line) samples (1 and 2 seconds respectively).

Caption of tables

Table 1. Training and testing data sets used for each type of classification task.

Table 2. MLP Classification performances. For each classification task, the first column reports the signal window duration (in seconds), the corresponding number of samples and the number of the neural network input nodes. We performed six different classification experiments for each window length, varying the set of events used for the network training and testing phase (6 permutations of the training and testing datasets, e.g. T1#1 … T1#6 for Regional/ Local classification task). The last column reports the average performance value obtained on each signal length under examination.

491 **Caption of figures**

492 Fig. 1 Location of SGG Station.

493 Fig. 2 Waveforms of a teleseism (panel A), a regional earthquake (panel B) and a
494 local earthquake (panel C). The red markers indicate the P wave onset obtained by using
495 ObsPy automatic picker. In blue the "trigger off" marker.

496 Fig. 3 Onset examples of two teleseisms (panel A), two regional earthquakes (panel
497 B) and two local earthquakes (panel C). The red markers indicate the P wave onset
498 obtained by using ObsPy automatic picker.

499 Fig. 4 Spectrograms of a teleseism (panel A), a regional earthquake (panel B) and a
500 local earthquake (panel C).

501 Fig. 5 Parameterization of one second of the seismic event onset. Teleseisms (panel
502 A), regional earthquakes (panel B) and local earthquakes (panel C).

503 Fig. 6 Parameterization of 2 seconds of the seismic event onset. Teleseisms (panel
504 A), regional earthquakes (panel B) and local earthquakes (panel C).

505 Fig. 7 Parameterization of 4 seconds of the seismic event onset. Teleseisms (panel
506 A), regional earthquakes (panel B) and local earthquakes (panel C).

507 Fig. 8 An example of two layer MLP architecture: the input vector (x1, ..., xn), i.e. the LPC
508 coefficients, moves forward from the hidden layer to the output one, that is one three examined
509 classification tasks, i.e. Regional/Local, Teleseisms/Local and Teleseisms/Regional. There are
510 not cycles and cross-connections between the layers.

511 Fig. 9 Final Prediction Error of the linear prediction coding evaluated for signal

512 windows of 100 (black line) and 200 (red line) samples (1 and 2 seconds respectively).