ITACA (ITalian ACcelerometric Archive): a Web Portal for the Dissemination of the Italian Strong Motion Data

Lucia Luzi, Salomon Hailemikael, Dino Bindi, Francesca Pacor, Francesco Mele
Istituto Nazionale di Geofisica e Vulcanologia, Italy

Fabio Sabetta
Dipartimento della Protezione Civile, Italy

INTRODUCTION

The Italian strong-motion database contains 2182 three component waveforms generated by 1004 earthquakes with a maximum magnitude of 6.9 (1980 Irpinia earthquake) covering the period range from 1972 to 2004. The database is devoted to serving the seismological and engineering communities.

The database can be accessed on-line at the site http://itaca.mi.ingv.it, where a wide range of search tools enables the user to interactively search events, recording stations and retrieve waveforms with particular characteristics, whose parameters can be specified, as needed, through a graphical user interface. A range of display options allows users to view data in different contexts, extract and download time series and spectral data.

The database was created during a joint project between Istituto Nazionale di Geofisica e Vulcanologia (INGV, *Italian institute for geophysics and vulcanology*) and Dipartimento della Protezione Civile (DPC, *Italian civil protection*). The aim of the project was the collection, homogenization and distribution of data acquired over the time period 1972-2004 in Italy by different Italian institutions, namely Ente Nazionale per l'Energia Elettrica (ENEL, *Italian electricity company*), Ente per le Nuove tecnologie, l'Energia e l'Ambiente (ENEA, *Italian energy and environment organization*) and DPC. The project had multiple purposes, such as permanent strong motion monitoring and temporary monitoring during seismic sequences or before permanent installation.

This database brings up to date the ENEA strong motion data collection which ended in 1993 and the European database where the most recent Italian data pertain to the Umbria-Marche sequence of 1997-1998 (Ambraseys at al., 2002). In addition effort was spent reviewing strong motion metadata and data processing to increase data quality and reliability.

This article describes the steps that led to the completion of the project and provides an overview of the search capabilities available at the database interface website.

THE DATABASE CONSTRUCTION

The database was implemented by development of three main activities:

1 – Definition of the database schema

2 – Collection and processing of waveforms

3 – Review of seismic event, station and instrument metadata

1. Definition of the database schema

The database schema and selection of a commercial software for the Relational Database Management System (RDBMS) was the first issue addressed. The database is handled through two different RDBMSs: Microsoft Access® 2003 for data input and DVD release, and Sun Microsystems MySQL for the web distribution. The selection of the former product is driven by the simplicity of the software, the worldwide distribution and the possibility of being linked to software for the management of spatial data, such as ESRI® products or software for scientific implementations such as Matlab®.

Forty-eight tables were created in order to store the information concerning the seismic events, the recording stations, the installed instruments and the strong-motion parameters, which are connected through a relational structure in order to avoid data redundancy.

A standard has been defined for the waveform format: the file name is composed of 33 characters that include event date and time, recording network, recording site, component, correction flag and time series type (acceleration, velocity, displacement or acceleration spectrum). This naming convention was chosen to facilitate file organization and management with simple operating system (OS) commands. The waveform file is in ASCII (American Standard Code for Information Interchange) format: it contains a metadata header of 43 rows that describes the event, the recording site and the instrument metadata. Processing information is also included for correct use of the data. Each waveform record has a self-consistent structure. The waveforms are also available in SAC (Seismic Analysis Code, Goldstein et al., 2003), a format used by the seismological community. They are distributed both in processed and raw format, so that expert users can re-process the data according to their needs.

2. Collection and processing of waveforms

The strong-motion dataset is composed by waveforms recorded by a wide variety of instruments. Before 1997 the accelerometers were almost entirely analog, while, after 1997, they were progressively substituted with digital ones. The heterogeneity of the records convinced us of the necessity of an individual waveform processing to be preferred to an automatic processing. The strong motion dataset was therefore processed in a homogeneous way, in spite of their different sources, in order to obtain reliable estimates of acceleration and velocity time-series and acceleration response spectra.

The linear trend of each raw analog record has been removed and the obtained signal has been convolved with the instrument response. Then, the time series has been band-pass filtered, selecting the high pass frequency by visual inspection of the Fourier spectrum, in order to detect the deviation from an omega-square shape. The low-pass frequency has been generally selected close to the instrument corner frequency, usually centred on 20-25 Hz.

The convolution of the digital records with the instrument response has not been performed, as in these cases, the instrument response is generally flat beyond 50 Hz. As few records had a usable pre-event, we removed the linear trend fitting the entire record. A band pass filter has been applied selecting the high pass frequency similarly to the analog records, while the low-pass frequency was generally applied in the range 25-30 Hz.

The filter type was selected in order to avoid phase shifts in the signal, which can alter the calculation of velocity and displacement time-histories and the shape of the elastic response spectra at frequencies higher than the applied low-cut (Boore and Bommer, 2005). A raised cosine filter was used for the analog records, often triggered on the S-phase, and an acausal 4th order Butterworth was used for the digital signals, after applying a cosine taper at the beginning and at the end of the record in order to avoid the filter transients. Preserving the low frequency content of these records was of primary importance during the processing phase.

3. Review of seismic event, station and instrument metadata

Reviewing of seismic event, station and instrument metadata, required the largest effort because the data were extremely inhomogeneous and sparse. In many cases new field surveys were undertaken for those stations that recorded the mainshock of strong earthquakes, the aftershocks of important seismic sequences, or when evident site effects were detectable from the records.

The seismic event metadata have an accuracy which reflects the network and instrument evolution during the 30 year time-span covered by the strong-motion data-base. The Italian seismic network managed by INGV (ING until 1999), increased from 12 instruments in 1972 to about 180 in 2004. As an example, the Italian Seismic Bulletin of 1972 included only 6 seismometers phases for the magnitude 5 event recorded offshore Ancona in 1972, while an event with the same magnitude that occurred in 2004 was recorded by over 70 instruments of the National Seismic Network.

Different catalogues were used to retrieve the hypocentral parameters and magnitudes for different periods:

- ING Catalogue (internal database of INGV) for the events in the period range 1972 – 1982;
- Catalogue of Italian Seismicity - CSI, version 1.1. (Castello *et al.*, 2006) and version 2.0 (R. Di Stefano, personal communication) for the events subsequent to 1982;

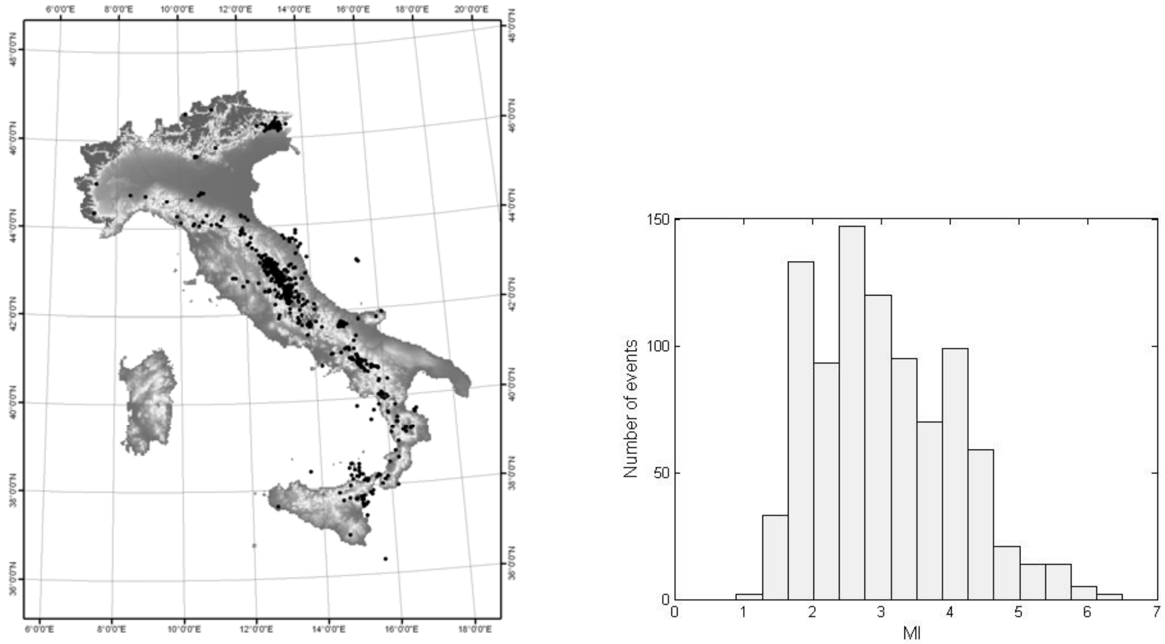
- Bollettino Sismico italiano (Istituto Nazionale di Geofisica e Vulcanologia – CNT, 2007) for some events subsequent to 1982;
- Catalogo Parametrico dei Terremoti Italiani CPTI04 (Gruppo di lavoro CPTI , 2004) used only for some surface wave magnitude values and as a reference for alternative hypocenters.

The CSI catalogues were preferred because the hypocentral parameters are instrumentally determined by integrating the Italian seismic network with regional and non-Italian networks. Events localised offshore or showing large horizontal errors in these catalogues, were relocated using the IPOP procedure (Mele *et al.* 2002).

Each event is associated with one or more magnitude estimates (local magnitude, Ml; surface wave magnitude, Ms; moment magnitude Mw; body wave magnitude, Mb). The Mw is evaluated from the solution of the parameters of the Centroid Moment Tensor (CMT), or from the Regional Centroid Moment Tensor (RCMT, Pondrelli *et al.*, 2006) and Earthquake Mechanisms of the Mediterranean Area (EMMA version 2, Vannucci and Gasperini, 2004). The Mb and Ms are attributed on the base of the International Seismologic Centre (ISC) Bulletin or the National Earthquake Information Centre (NEIC) catalogue. For earthquakes with low magnitude values (< 4), the reference is the local magnitude, Ml, obtained from INGV instrumental catalogues.

For the focal mechanisms the classification of Zoback (1992) was adopted, which discriminates among 5 main types: normal faulting, predominately normal with strike-slip component, strike-slip faulting (with eventual minor normal or thrust component), thrust faulting and predominately thrust faulting with strike-slip component. The fault geometry, strike, dip and rake are reported for the major events. They are obtained from the DISS catalogue, version 3.0.2 (DISS Working Group, 2006).

The geographic distribution of the seismic events included in the database is shown in Figure 1A. The events are mainly located in the Apennines and in the eastern Alps. The magnitude distribution is shown in Figure 1B: only 38 events have local magnitude equal to or higher than 5, representing a small percentage of the total (about 4%). The distribution of peak ground acceleration values reflects the event distribution. Only 15% of the waveforms (about 640, if all the components are examined) have Peak Ground Acceleration (PGA) values greater than 50gal (Figure 2).



(A) (B)

Figure 1. (A) Map of Italy showing the geographical distribution of seismic events and (B) local magnitude distribution.

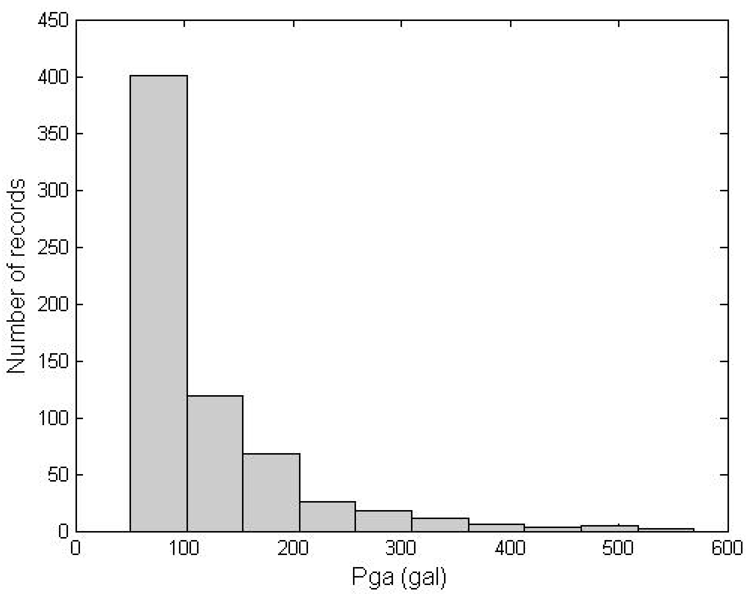


Figure 2. Distribution of peak ground acceleration after applying a threshold of 50gal.

The station metadata were included in the database after collection of pre-existing data and field investigations performed during the project. In particular the metadata concerning 620 strong motion stations have been collected: 330 among them are inoperative since they were part of temporary networks or equipped with old analog instruments, which have since been removed.

Station metadata are composed of different levels. The descriptive level includes all the synthetic information regarding the site, such as name, code, address, coordinates, topographic map location, geotechnical class, type of installation, etc. The map level includes the station location on a topographic map or an aerial photograph and a geological map. The table level concerns different useful parameters for the geotechnical site characterization, including stratigraphic logs, Standard Penetration Test (SPT) logs, Vs/Vp profiles, dispersion curves, fundamental frequencies, site response functions, etc. Furthermore the station metadata have been stored in specific metadata reports (about 400 pdf files have been compiled), which can be downloaded through the web interface to the database.

Field investigations were undertaken during the project in order to characterize the sites that recorded the strongest Italian events (i.e. 1980 Irpinia, M_w 6.9; 1984 Lazio-Abruzzo, M_w 5.9; 1997 Umbria-Marche, M_w 6.0). Different geophysical techniques were applied, depending on the nature and importance of the site: downhole, cross-hole, seismic refraction, seismic reflection, SASW, noise measurements (single station or array). In addition, in a few sites detailed geological surveys and geo-mechanical surveys were performed.

THE ONLINE DATABASE

Strong motion data dissemination is achieved through the database web portal at

http://itaca.mi.ingv.it.

The data acquired by different institutions are stored in a RDBMS on the web server. Intuitive graphical user interfaces allow queries selecting station, seismic event and waveform parameters.

The strong-motion recordings selected with a search query can be downloaded by web clients in raw and processed format, together with the velocity and displacement time-series and acceleration response spectra. A schematic figure of the user interface is illustrated in Figure 3.

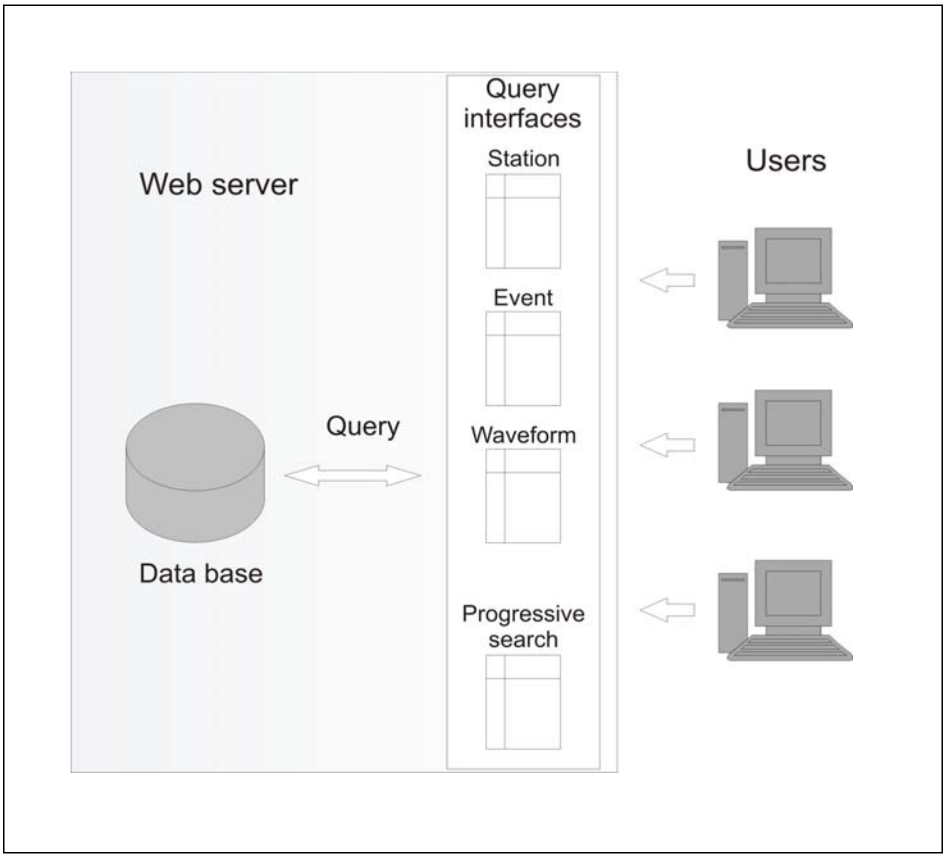


Figure 3. Schematic figure of the on-line user interface.

The database can be explored through searchable key fields: 10 for the stations, 8 for the seismic events and 9 for the waveforms. The design philosophy of the database is that separate queries can be performed through three distinct database interfaces: stations, events and waveforms. Alternatively a progressive search can be done starting with the selection of seismic event parameters and progressively constraining the results with station and waveform parameters, keeping a record of the choices made that are displayed at the top of each webpage.

Each query (event, station, waveform or progressive search) returns a list of matching results and the single outcome can be explored in detail. Figure 4 shows the user interface for the selection of the recording stations. From the list of outcomes the recordings and the details relative to a single station can be retrieved. Both recording stations and events are plotted using the Google Maps© interface that allows the user to display the points either on a satellite image or a basic map (or both), as shown in Figure 5.

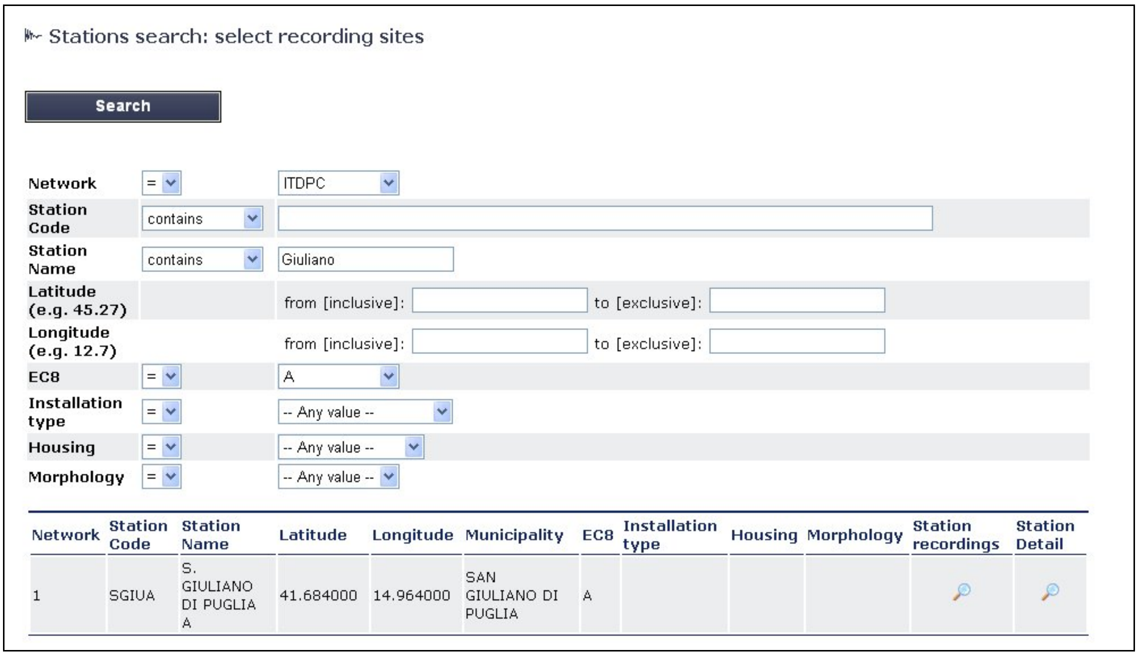


Figure 4. Screenshot of the graphical user interface for the station search (the search criteria are: network = *ITDPC*, station name contains *Giuliano* and EC8 site class = *A*).

Station Name	COLFIORITO				
Lat	43.035898	**Long**	12.920538	**Projection**	GEOWGS84
Elev [m.a.s.l.]	753	**EC8 Code**	C	**Estimate**	DH

Install. Date	1991-05-16 00:00:00	**Removal date**			
Address					
Municipality	SERRAVALLE DI CHIENTI				
Proximity	Far from buildings				
Permanent	Permanent	**Housing**			
Installation	Pillar				
IGM sheet	123	**Sector**	II	**Orientation**	SE
Morphology	Alluvia plain				

GEOTECHNICAL LOGS

Code	Reference	Latitude	Longitude	Nspt profile	Vs/Vp profile	Cu profile
1	Di Giulio et al. (2006)	43.020944	11.932786			

Figure 5. Screenshot of an example of station mapping using the Google Maps interface©.

The geotechnical information (stratigraphy, NSPT or Vs/Vp profile, etc.), when available, is displayed in a table format, as shown in Figure 6.

Vs / Vp profile

Network	ITDPC				
Station code	CLF				
Log code	1				
Lat **Proj**	43.020944 Geographic coordinates referred to WGS84	**Long**	11.932786	**Elev [m.a.s.l.]**	
Reference	Di Giulio et al. (2006)				

VS PROFILE

Code	Top [m from g.l.]	Bottom [m from g.l.]	Vs [m/s]	Vp [m/s]
1	0.0	4.0	126.0	
2	4.0	4.5	126.4	
3	4.5	5.2	129.5	
4	5.2	6.2	126.4	
5	6.2	7.0	126.4	
6	7.0	7.7	123.3	
7	7.7	8.5	129.5	
8	8.5	9.3	126.4	
9	9.3	10.1	129.5	
10	10.1	10.9	126.4	
11	10.9	11.5	126.4	
12	11.5	12.2	126.4	
13	12.2	13.0	126.4	
14	13.0	13.8	126.4	
15	13.8	14.4	126.4	
16	14.4	14.9	126.4	
17	14.9	15.6	126.4	
18	15.6	16.1	129.5	
19	16.1	16.3	154.2	
20	16.3	16.8	154.2	

Figure 6. Screenshot of the Vs profile for station Colfiorito (CLF).

For a single event the epicentre is shown together with the triggered stations. Each station can be selected in order to display the recorded waveform (Figure 7).

When a progressive search is made, the user can start with the event characteristics and gradually refine the search by specifying the station and the waveform parameters (Figure 8). In this case only the selected waveforms can be downloaded in a zipped file that contains unprocessed and processed acceleration, velocity, displacement and acceleration response spectra.

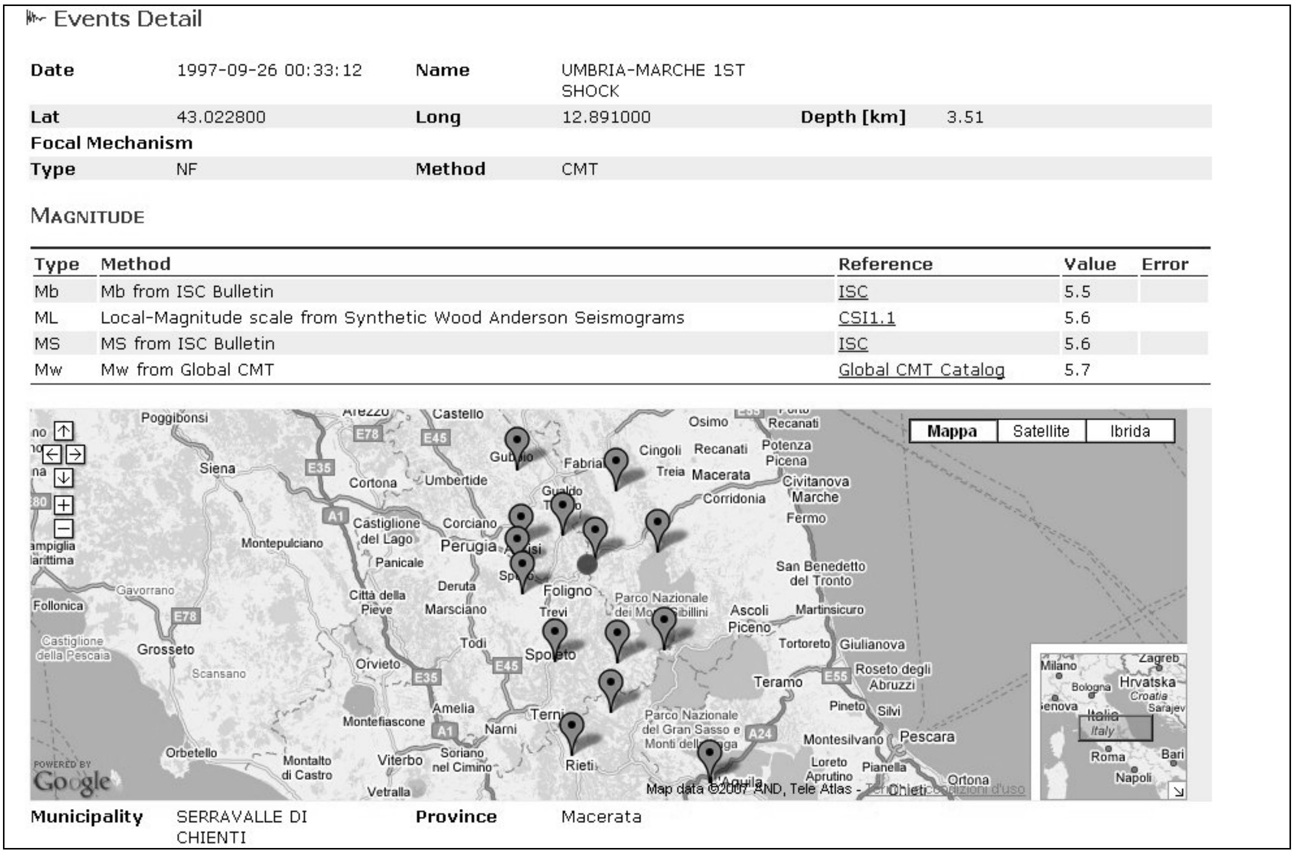


Figure 7. Screenshot of an example of epicentre location and triggered stations.

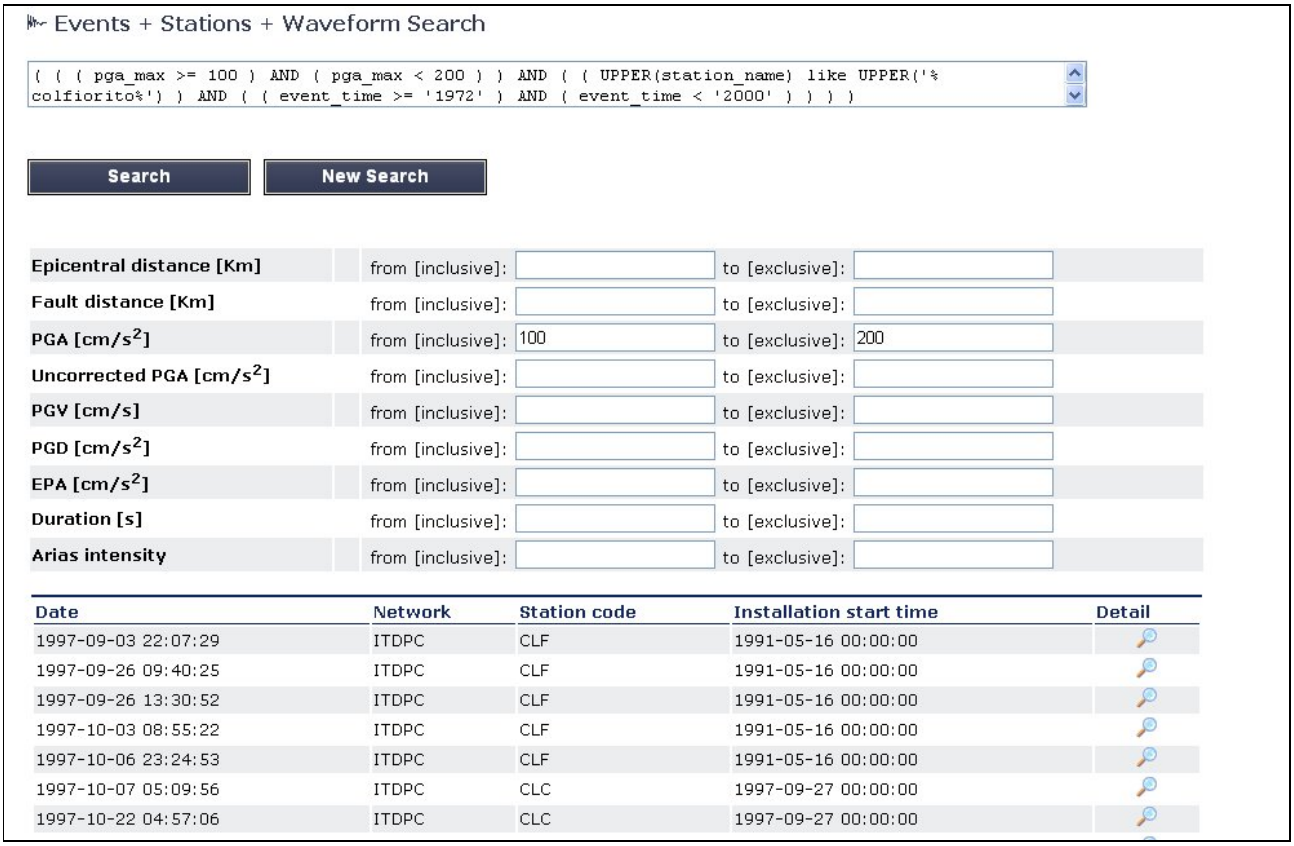


Figure 8. Screenshot of an example of data query and returned results with the following constraints: earthquake date from *1972* to *2000*, station name contains *colfiorito* and absolute PGA between *100* and *200* gals.

The waveforms that satisfy the required conditions can be displayed with the aid of a Sun Microsystems Java © applet (Figure 9) that allows the user to perform simple operations like zoom in/out, change plot options (axis labels, axis limits, background and foreground colour, etc.) and save or print the plot.

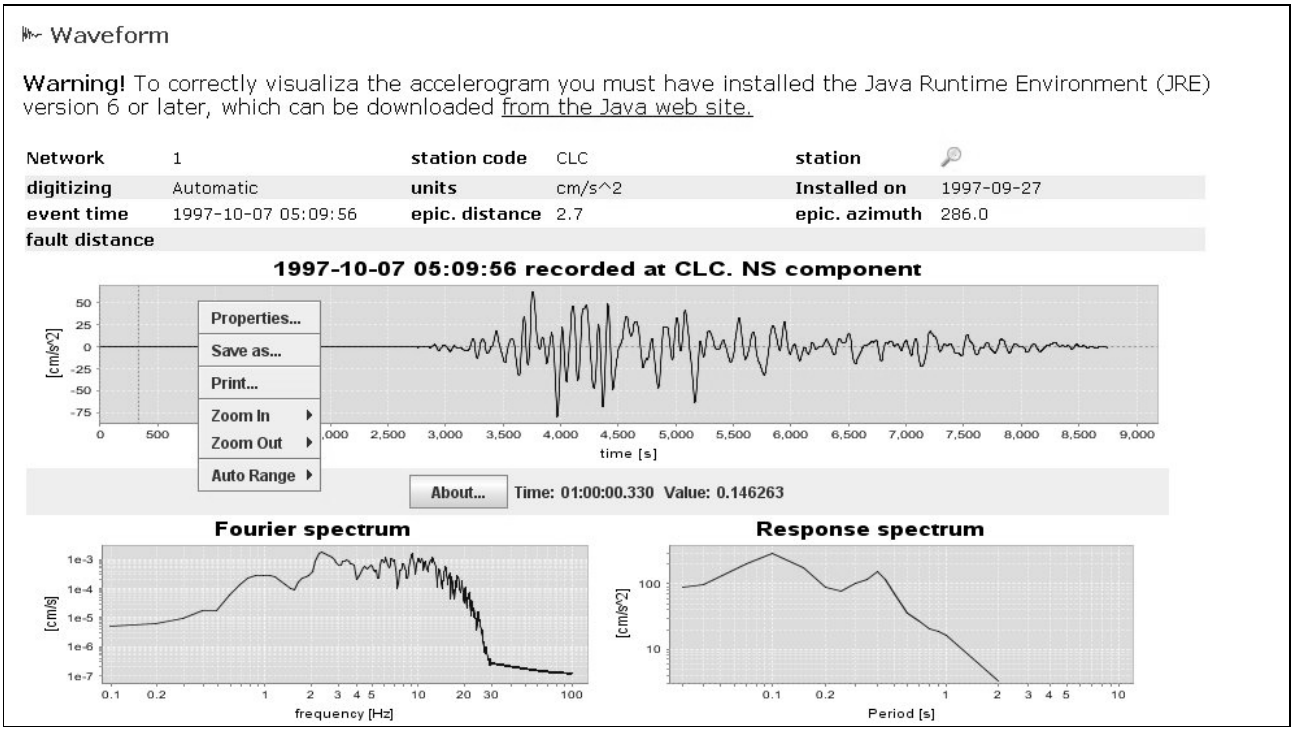


Figure 9. Screenshot of web-based waveform visualization.

CONCLUSION AND FUTURE DEVELOPMENTS

The main features of the web interface and the relational database management system for the dissemination of the archived Italian strong motion dataset have been presented. The database has been designed to be a very useful tool for scientific research in the seismological and engineering fields and in particular for data analysis focused on seismic risk assessment.

Several decisions need to be made in order to keep the database constantly updated in the future and improve the amount of information regarding recording sites, waveforms and seismic events.

One of the most important future challenges will be to perform a correct evaluation of the transfer functions of most of the recording sites, in order to improve the knowledge on the site response.

This goal will be reached through the analysis of the recorded waveforms and direct geotechnical site characterization.

In the project the importance of this task has been acknowledged, and many activities have been conducted for this purpose for sites of relevant interest. Moreover, cooperation with the European project Network of Research Infrastructures for European Seismology (NERIES, 2006) – Joint Research Geotechnical Site Characterization (JRA4) has been established in order to plan and perform geophysical field investigations concerning 10 Italian recording sites. Even with the amount of work that has already been done with respect to site characterization and site response, there is still much to do. In particular, the nonlinear behaviour of soils will require continued monitoring to record a greater number of ground motions in the 100 gal or larger range, and continued analysis of these records.

ACKOWLEDGEMENTS

This work was supported by grants from the understanding between Istituto Nazionale di Geofisica e Vulcanologia and Dipartimento della Protezione Civile (Italy) in the time span 2004-2006.

The authors are very grateful to the anonymous referee who contributed in improving the paper.

REFERENCES

Ambraseys, N., Smit, P., Sigbjornsson, R., Suhadolc, P. and Margaris, B. (2002). Internet-Site for European Strong-Motion Data. European Commission, Research-Directorate General, Environment and Climate Programme.

Boore D. M., and Bommer J.J. (2005). Processing of strong motion accelerograms: needs, options and consequences. *Soil Dyn. Earth. Eng.* **25**, 93-115.

Castello B., Selvaggi G., Chiarabba C., and Amato A. (2006). CSI Catalogo della sismicità italiana 1981-2002, version 1.1. INGV-CNT, Roma. http://www.ingv.it/CSI/.

DISS Working Group (2006). Database of Individual Seismogenic Sources (DISS), Version 3.0.2: A compilation of potential sources for earthquakes larger than M 5.5 in Italy and surrounding areas. http://www.ingv.it/DISS/, © INGV 2005, 2006 - Istituto Nazionale di Geofisica e Vulcanologia - All rights reserved.

Esri - Environmental Systems Research Institute, Inc. Copyright © 1995-2008 ESRI. All rights reserved.

Global Centroid Moment Tensor Project. http://www.globalcmt.org.

Google Maps. Copyright © 2008 Google Map data – 2008 Tele Atlas.

Goldstein, P., D. Dodge, M. Firpo, Lee Minner (2003). SAC2000: Signal processing and analysis tools for seismologists and engineers, Invited contribution to The IASPEI International Handbook of Earthquake and Engineering Seismology, Edited by WHK Lee, H. Kanamori, P.C. Jennings, and C. Kisslinger, Academic Press, London.

Gruppo di lavoro CPTI (2004). Catalogo Parametrico dei Terremoti Italiani, versione 2004 (CPTI04). INGV, Bologna. http://emidius.mi.ingv.it/CPTI04/.

International Seismological Centre. Bulletin of the International Seismological Centre. http://www.isc.ac.uk.

Istituto Nazionale di Geofisica e Vulcanologia – CNT (2007). Bollettino della sismicità strumentale. Centro Nazionale Terremoti, INGV, Roma. http://www.ingv.it/~roma/reti/rms/bollettino.

Istituto Nazionale di Geofisica e Vulcanologia MedNet - Mediterranean Very Broadband Seismographic Network. http://mednet.rm.ingv.it.

Istituto Nazionale di Geofisica. ING Catalogue (1450 B.C. - 1990). Roma.

Java Technology. Copyright © 2005 Sun Microsystems, Inc. All rights reserved.

Matlab - The Language of Technical Computing, Release 14. Copyrigth ® 1984-2005 The Mathworks Inc.

Mele F., Badiali L., Marcocci C. and Piscini A. (2002). Locator: i codici. Un sistema di localizzazione in tempo quasi reale. Technical notes of Centro Dati e Informazione sui Terremoti, n. 6 / 2002.

Microsoft Access 2003. Copyright © 1992-2003 Microsoft Corporation. All rights reserved.

MySQL 2008. Copyright © 1995-2008 Sun Microsystems. All rights reserved

National Earthquake Information Center – NEIC. Earthquake catalogue.
http:// earthquake.usgs.gov/regional/neic

Network of Research Infrastructures for European Seismology - NERIES.(2006). Sixth Framework Program (FP6) of the European Commission, contract no. RII3-CT-2006-026130. http://www.neries-eu.org/

Pondrelli, S., Salimbeni S., Ekström G., Morelli A., Gasperini P., and Vannucci G. (2006). The Italian CMT dataset from 1977 to the present. *Phys. Earth Planet. Int.*, doi:10.1016/j.pepi.2006.07.008, 159/3-4, 286-303. www.ingv.it/seismoglo/RCMT.

Vannucci G., and Gasperini P. (2004). The new release of the database of Earthquake Mechanisms of the Mediterranean Area (EMMA version 2). *Annals of Geophysics*, supplement to vol. **47**, N. 1.

Zoback M.L. (1992). First- and Second-Order Patterns of Stress in the Lithosphere: the World Stress Map Project. *J. Geoph. Res.* **97**, n.B8, 11703-11728.