The Italian strong motion database: design, data input and web distribution

L. Luzi[1], F. Sabetta[2], F. Mele[3], S. Hailemikael[2], D. Bindi[1], F. Pacor[1], M. Massa[1], S. Lovati[1], A. Gorini[3] and S6 project working group

[1] *Istituto Nazionale di Geofisica e Vulcanologia, sezione Milano-Pavia, Milano, Italy*

[2] *Dipartimento della Protezione Civile, Ufficio Valutazione Prevenzione e Mitigazione del Rischio Sismico ed Attività ed Opere Post-Emergenza, Roma, Italy*

[3] *Istituto Nazionale di Geofisica e Vulcanologia, Centro Nazionale Terremoti, Roma, Italy*

(Received: February 21, 2008; accepted: July 1, 2008)

ABSTRACT A new Italian strong-motion database was created during a joint project between the Istituto Nazionale di Geofisica e Vulcanologia (INGV, Italian Institute for Geophysics and Vulcanology) and the Dipartimento della Protezione Civile (DPC, Italian Civil Protection). The aim of the project was to collect, homogenize and distribute strong motion data acquired by different 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 the DPC with different purposes, such as permanent strong motion monitoring and temporary monitoring during seismic sequences or before permanent installation in Italy in the time span 1972-2004. The database contains 2182 three-component waveforms generated by 1004 earthquakes with a maximum moment magnitude of 6.9 (1980 Irpinia earthquake) and can be accessed on-line at the site http://itaca.mi.ingv.it. Here a wide range of search tools enables the user to interactively retrieve events, recording stations and waveforms with particular characteristics, whose parameters can be specified, as needed, through user friendly interfaces. A range of display options allows users to view data in different contexts, extract and download time series and spectral data. This article describes the database structure and the working steps which led to the completion of the project.

1. Introduction

The effort put into building an Italian strong motion database is motivated by the increasing demand from the scientific community for strong motion data and by the lack of a updated national databases. Scientific research in the seismological and engineering fields requires strong motion data for several purposes, such as evaluation of GMPE (Ground Motion Prediction Equations), verification of shaking scenarios and probabilistic hazard maps and the formulation of seismic codes. The correct use of the accelerometric data implies a strong background on the origin of the data, such as details on the instruments that recorded the time series, the recording site and the seismic event, for which the correct attribution of earthquake parameters, such as location and magnitude, should be made. Therefore, the database construction not only implies a mere collection of waveforms, but also a careful attribution of event, station and instrument metadata.

For this purpose a joint project was established between the Istituto Nazionale di Geofisica e Vulcanologia (INGV, Italian Institute for Geophysics and Vulcanology) and the Dipartimento della Protezione Civile (DPC, Italian Civil Protection) in the time span 2004-2006.

The aim of the project was to bring up to date an earlier Italian strong motion database, updated to 1993 and to make a considerable effort in improving the metadata associated to seismic events and recording stations in order to facilitate data accessibility for the scientific community. These characteristics should make this database different from the European strong motion database by Ambraseys *at al.* (2002, 2004), whose most recent Italian data pertain to the Umbria-Marche sequence of 1997-1998.

In order to fulfill these requirements four main activities were developed:

1 - definition of the database structure;

2 - collection and processing of waveforms;

3 - review of seismic event, station and instrument metadata;

4 - dissemination of data through a web portal.

This paper describes the working steps and the decisions made to fulfill the project requirements and the capabilities of the web database.

2. Definition of the database structure

The database contains 48 tables, created to store the information about the seismic events, the recording stations, the installed instrumentation and the strong-motion parameters, that are connected through a relational structure in order to avoid data redundancy. Fig. 1 exemplifies the relations existing among tables for the database block relevant to recording stations. The complete database structure is described at http://esse6.mi.ingv.it/, where the fields relevant to all tables are described in detail, together with the technical requirements for data input.

The database is handled through two different relational database management systems: Ms Access® 2003 for data input and DVD release, and MySQL for the web distribution. The choice of the former product is driven by the simplicity of the software, the worldwide diffusion and the interface capability with software for the management of spatial data, such as ESRI® products, or software for scientific implementations such as Matlab®. The two databases communicate through a protocol that allows us to transfer the database tables, stored locally in the MS Access format, to the web database, in order to update the information as soon as it becomes available.

Only authorized personnel has access to the database and to the web server for the modification of the data.

3. Collection and processing of waveforms

Raw recordings obtained by different institutions were collected and homogenized in a unique format. A standard was established for the strong motion file name composed of 33 characters which include: event date and origin time, recording network, recording site, component, correction flag and time series type (acceleration, velocity, displacement or spectral acceleration). This standard was chosen in order to exemplify the file organization and management with simple operative system commands. The waveform files, in ASCII format, contain a metadata

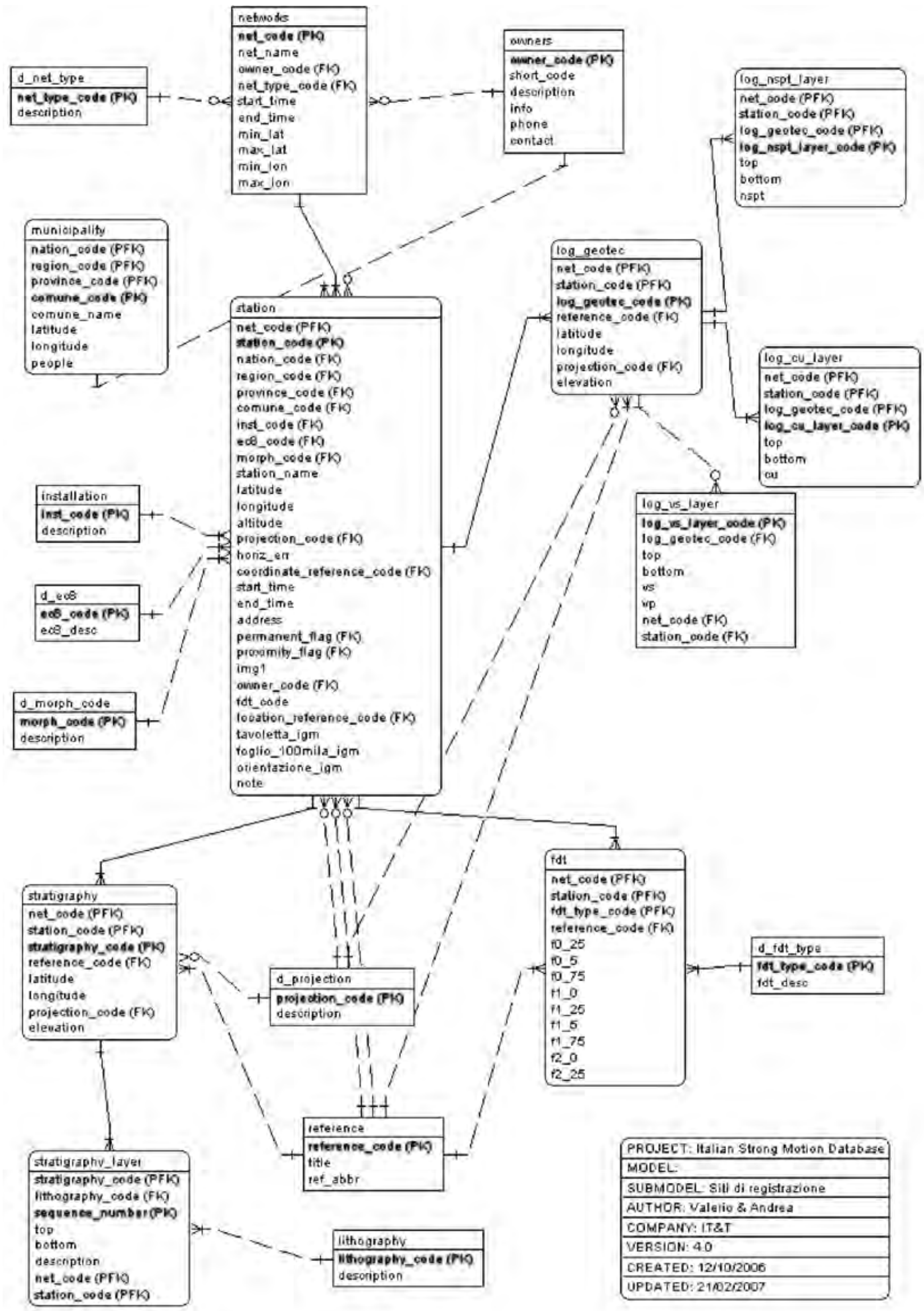


Fig. 1- Example of the relations existing among database tables for the stations.

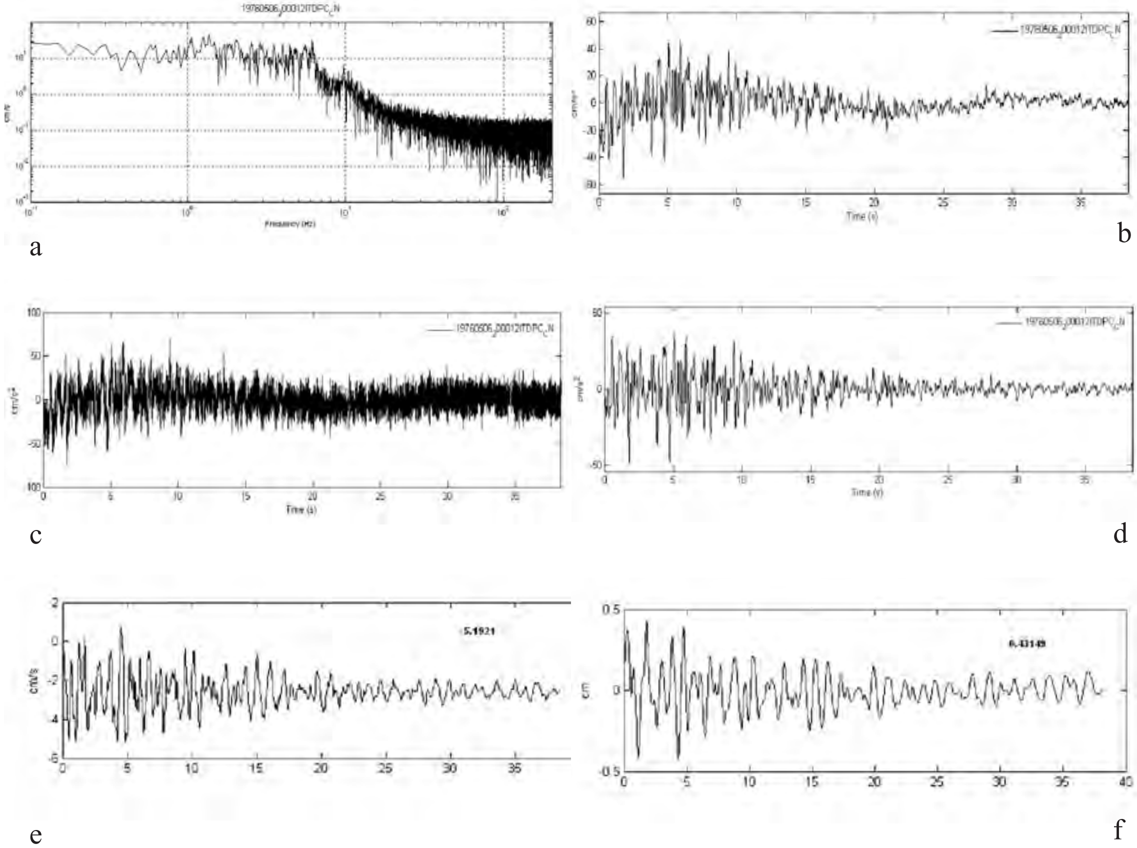
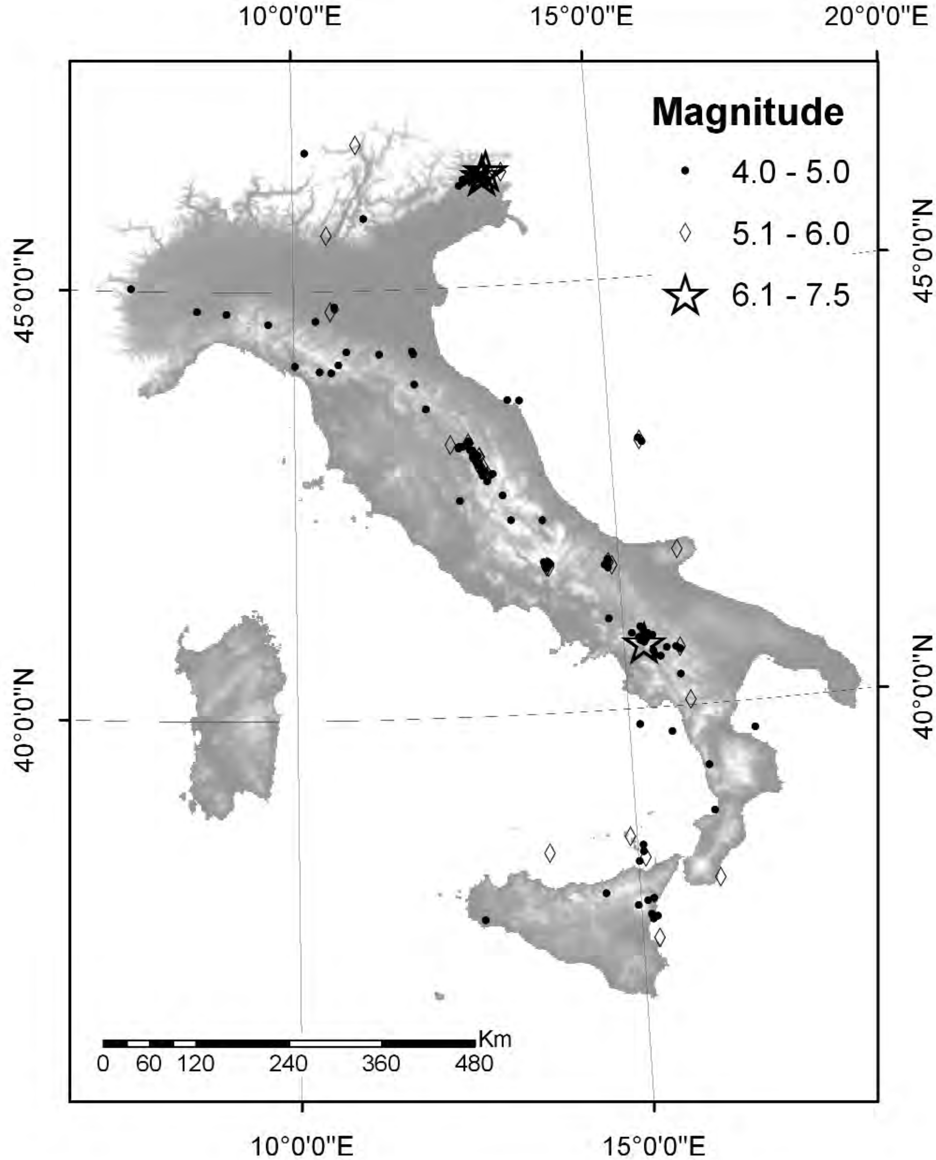


Fig. 2 - Data processing steps applied to the 1976-05-06 20:00:12 event recorded at Conegliano Veneto (N-S component): a) linear baseline correction, b) Fourier spectrum of the uncorrected data for high-pass frequency selection, c) convolution with the instrument response, d) processed acceleration after band pass filtering, e) velocity time series and f) displacement time series.

header of 43 rows which describes event, recording site and instrument metadata, together with processing information for a correct use of the data. The waveforms are also available in SAC format, mainly used by the seismological community. They are distributed both in processed and raw format, so that expert users can re-process data according to their needs.

In order to obtain reliable estimates of acceleration and velocity time-series and acceleration response spectra, the strong motion data were processed in a rather homogeneous way, although they were recorded by different instrument types. Before 1997 most of the accelerometers were analogue, while, after 1997, they were progressively substituted with digital ones. The heterogeneity of the records pushed us to use a one-by-one waveform processing instead of an automatic one.

The linear trend of each raw analogue record was removed and the signal obtained was corrected for the instrument response. Then, the time series was band-pass filtered, selecting the high-pass frequency by visual inspection of the record Fourier spectrum. The low-pass frequency was generally selected close to the instrument upper cut-off frequency, usually centred on 20-25 Hz.



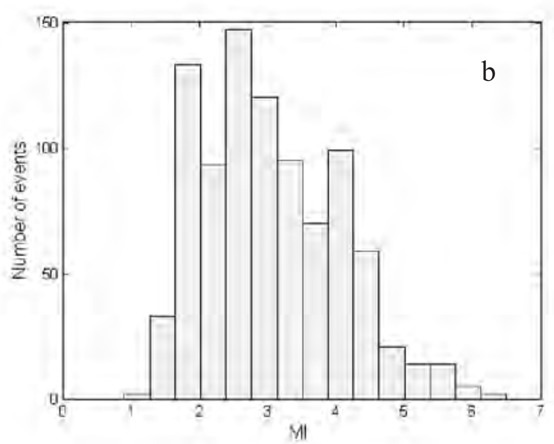


Fig. 3 - Geographic distribution of the events (a) and local magnitude distribution (b).

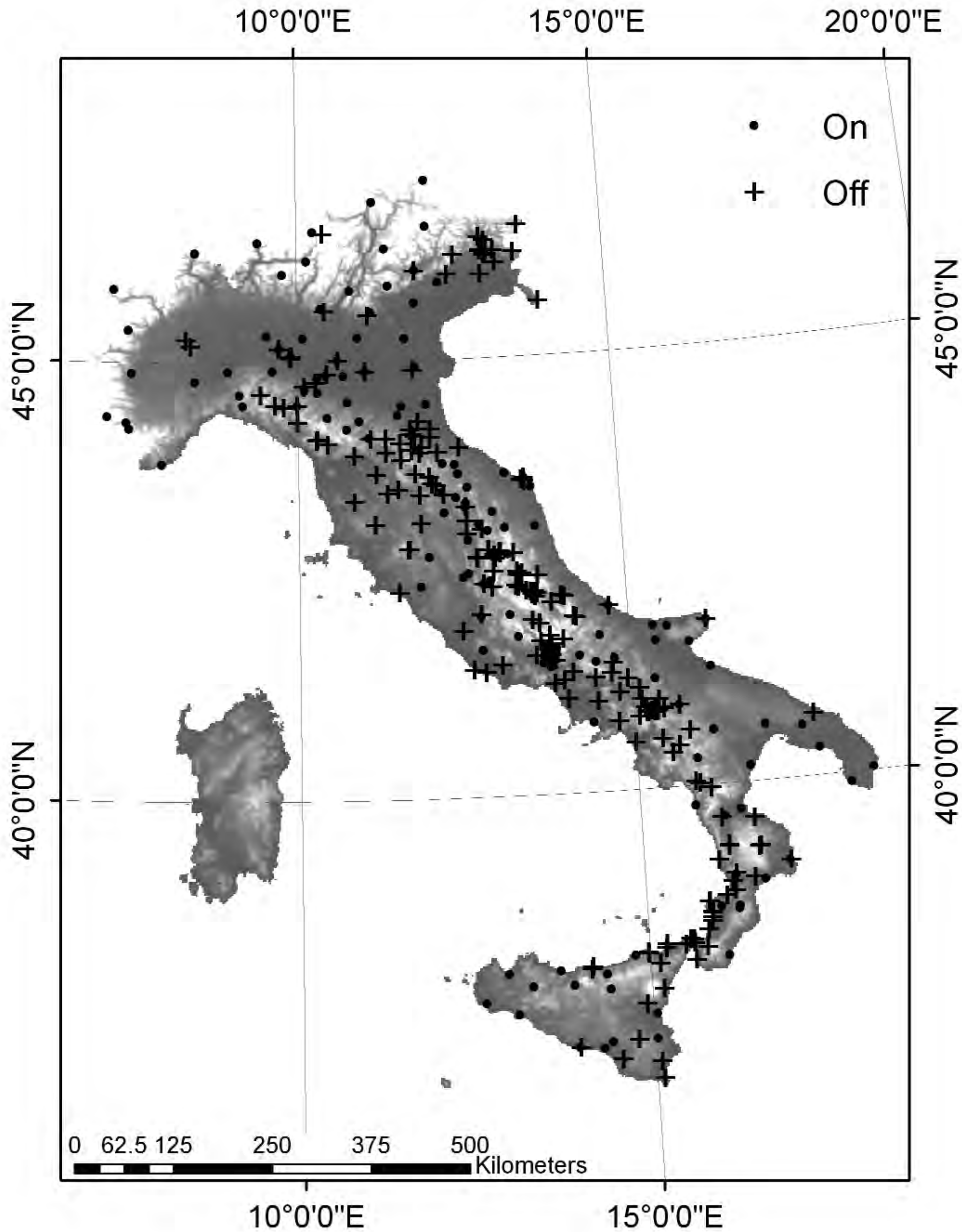


Fig. 4 - Distribution of the recording stations
(On = operational on 2007; Off = inoperative on 2007)

Convolution of digital records with the instrument response was not performed, as the instrument response is generally flat from 0 to frequencies higher than 50 Hz. As few records have a usable pre-event, we removed the linear trend fitting the entire record. A band-pass filter was applied selecting the high-pass frequency, similarly to the analogue records, while the low-

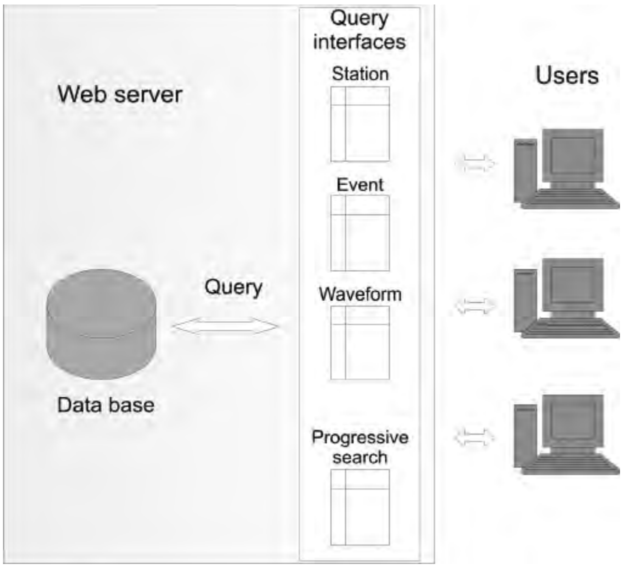


Fig. 5 - Structure of the web database.

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 analogue records, often triggered on the S-phase, and an acausal 4th order Butterworth filter was used for the digital signals, after applying a cosine taper at either side of the record in order to avoid the filter transients. The processing was devoted to preserve the low frequency content. Fig. 2 shows an example of the data processing steps for the time series recorded at Conegliano Veneto relevant to the 1976, Mw 6.4, Friuli earthquake.

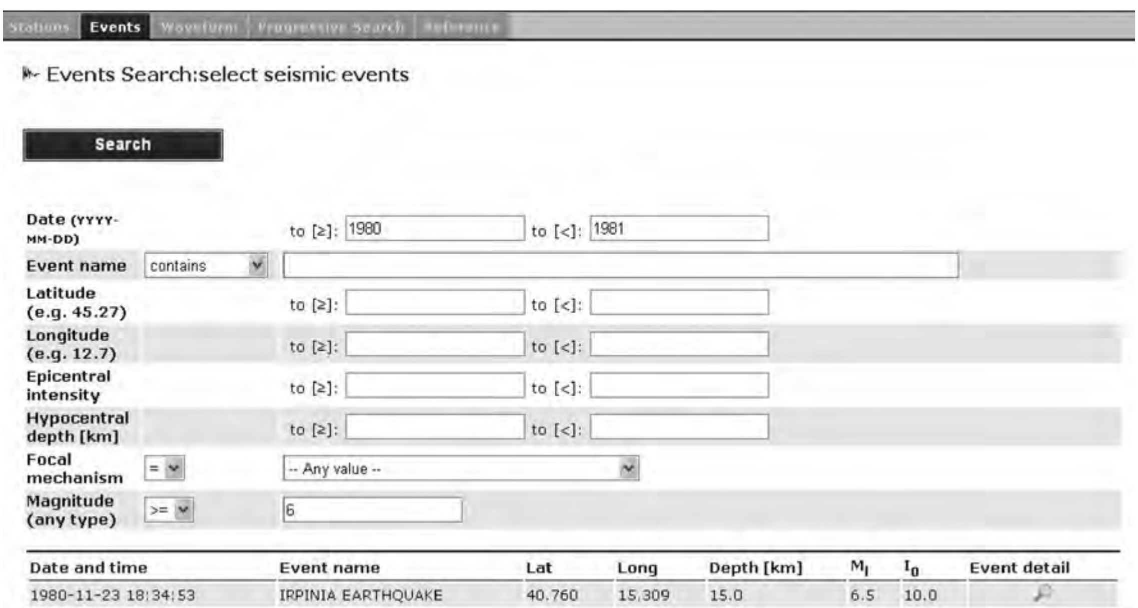


Fig. 6 - User interface for the event search (selection criteria are: Date ≥ 1980 and Date < 1981, Magnitude ≥ 6).

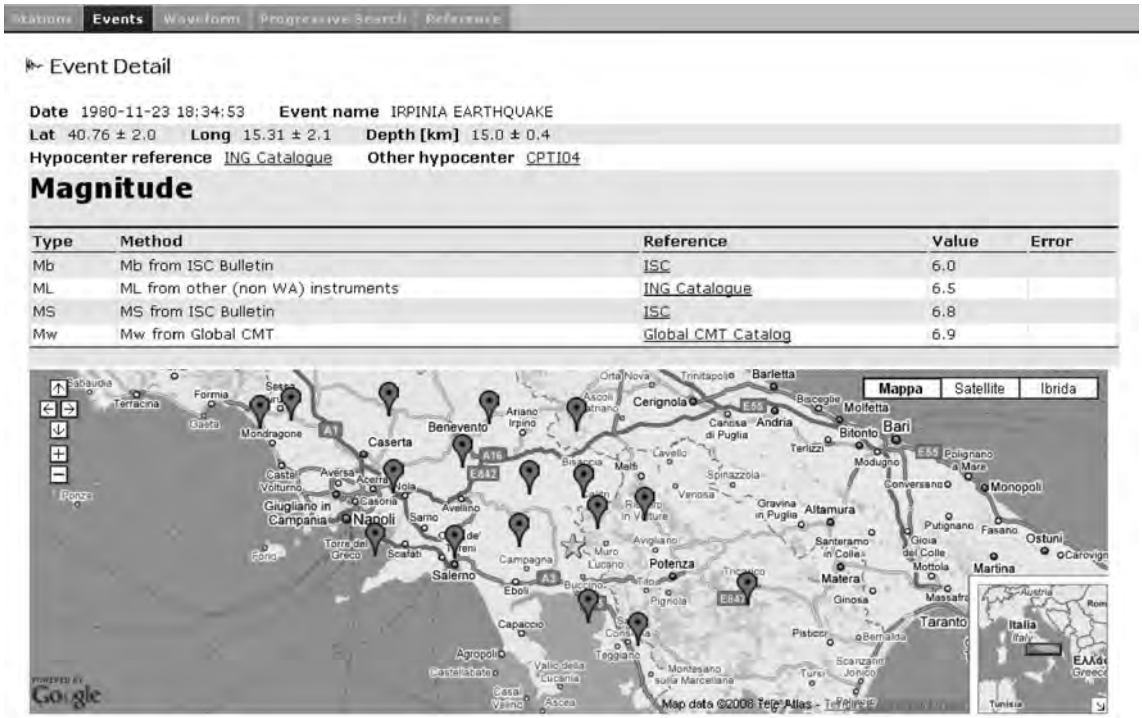


Fig. 7 - Example of event mapping through Google–Map data©; the epicentre is shown together with triggered stations. Each station can be selected in order to display the recorded waveform.

4. Review of seismic event, station and instrument metadata

The review of the seismic event, station and instrument metadata, was the biggest part of the project, as the data were extremely unhomogeneous.

The accuracy at the earthquake parameters reflects the network and instrument evolution covered by the strong-motion database during the 30-year time-span.

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

ING Catalogue (internal database of INGV) for 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 seismometric network with regional and non-Italian networks. Complex events, localised offshore or showing large horizontal errors in these catalogues, were relocated using the IPOP procedure (Mele *et al.*, 2002).

One or more magnitude estimates were attributed to each earthquake. The moment magnitude

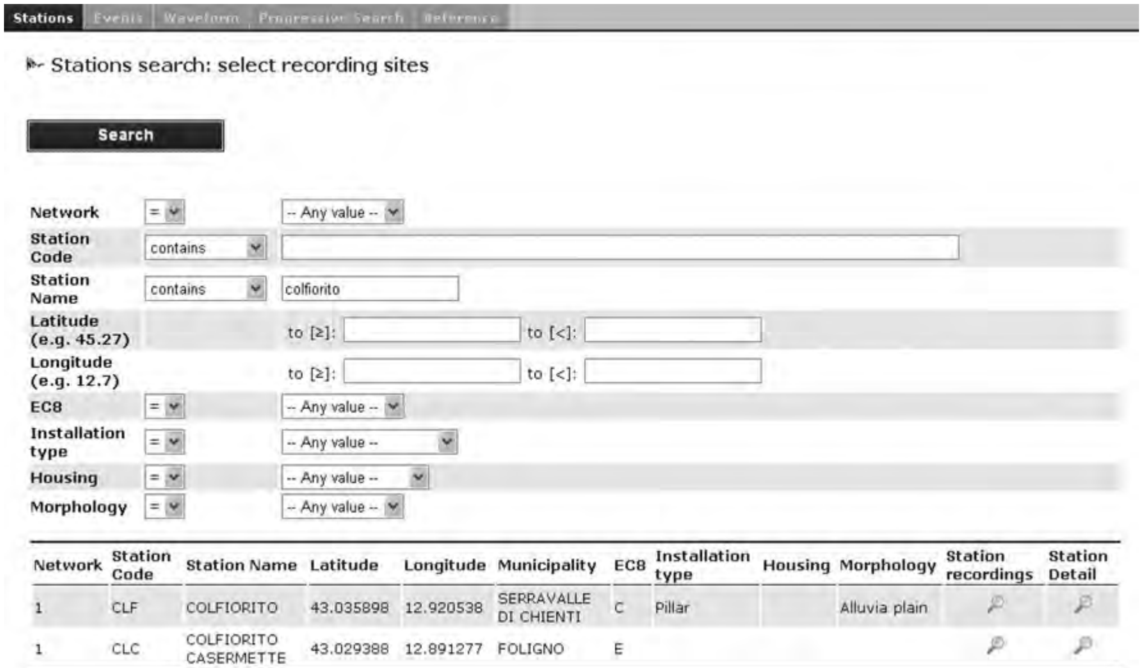


Fig. 8 - User interface for the station search (selection criteria are: network = ITDPC, station name contains Colfiorito).

is evaluated from the solution of the parameters of the Global Centroid Moment Tensor (www.globalcmt.org) or from the Regional Centroid Moment Tensor, (Pondrelli *et al*., 2006) and Earthquake Mechanisms of the Mediterranean Area [EMMA version 2, Vannucci and Gasperini (2004)], while the M_B and M_S were attributed on the basis of the International Seismological Centre ISC Bulletin or the National Earthquake Information Centre NEIC catalogue. For earthquakes with low magnitude values (< 4), the reference is the local magnitude, M_L, obtained from INGV instrumental catalogues.

The geographic distribution of the seismic events indicates their location along the Apennines and the eastern Alps (Fig. 3a). Only 38 events have a local magnitude equal to or higher than 5, which represents a small percentage of the total (about 4%), as shown in Fig. 3b.

In a few cases, it was possible to assign a focal mechanism estimation. The classification of Zoback (1992) was adopted, which discriminates 5 types, namely: 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 were reported for relevant events and were obtained from the DISS catalogue, version 3.0.2 (DISS Working Group, 2006).

The metadata regarding 620 strong motion stations are relevant to both pre-existing data and new field investigations. Station metadata are organized in different formats. The geographic distribution of the recording stations reflects the distribution of major earthquakes since 1972: they are mainly located along the Apennines, eastern Alps and Sicily (Fig. 4). The descriptive level includes the synthetic information regarding the site, such as name, code, address, coordinates, topographic map

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				

a

GEOTECHNICAL LOGS

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

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	

b

Fig. 9 - Colfiorito recording station: a) example of station detail and mapping through Google–Map data© and b) detail on the Vs profile

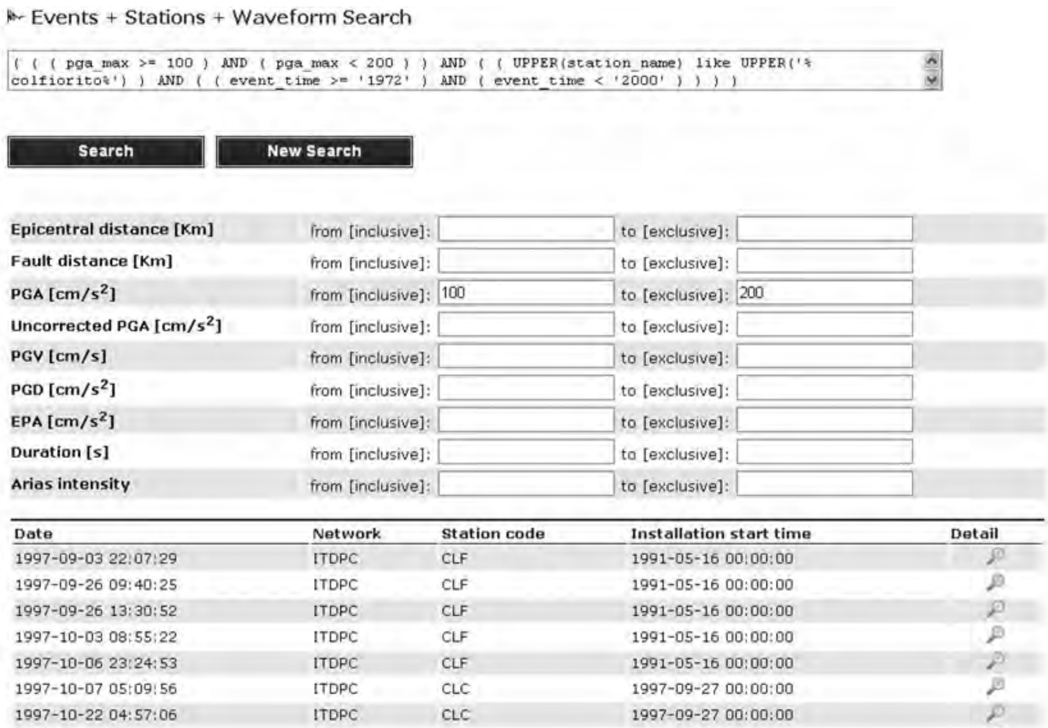


Fig. 10 - Example of data query with the following constraints: earthquake date from 1972 to 2000, station name contains Colfiorito and absolute PGA between 100 and 200 Gal.

location, ground type according to the EC8 classification, type of installation, etc. The map format includes the station location on a topographic map or an aerial photograph and/or a geological map, while the table format is used for the geothecnical parameters, such as: stratigraphy logs, NSPT logs, Vs/Vp profiles, dispersion curves, fundamental frequencies, site response functions, etc. As an alternative to the on-line database, station metadata have been stored in specific documents, that can be downloaded at http://itaca.mi.ingv.it.

Several field investigations were promoted during the project in order to characterize the sites that recorded the strongest Italian events (i.e. Irpinia 1980, Lazio-Abruzzo 1984, Umbria-Marche 1997). Borehole logs were obtained for a few selected sites, from drillings purposely performed during the project. Different geophysical techniques were applied, depending on the subsoil nature, characteristics of data recorded, and logistic considerations: downhole, cross-hole, seismic refraction, seismic reflection, SASW, noise measurements (single station or array).

5. Dissemination of data through a web portal

The dissemination of the Italian strong motion database is performed through the web portal http://itaca.mi.ingv.it.

Here, a fully relational database is stored, that can be accessed through user friendly interfaces which allow the user to perform queries to select station, seismic event and waveform parameters, in

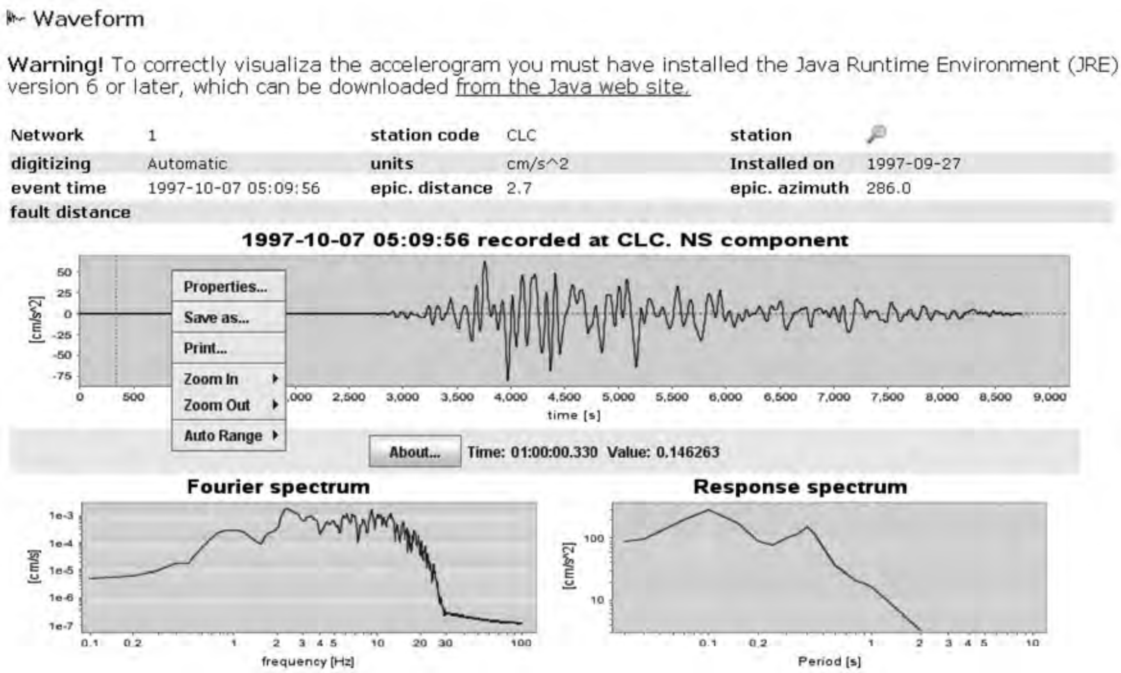


Fig. 11 - Waveform visualization.

order to download the strong motion data, as exemplified in Fig. 5.

The database can be explored through searchable key fields: 10 for the stations, 8 for the seismic events and 9 for the waveforms. The basic idea is that separate queries can be performed for three distinct data base blocks (stations, events and waveforms) and, as an alternative, a progressive search can be performed starting with the selection of seismic event parameters and progressively constraining the results with station and waveform parameters, keeping a memory of the choices made. Each query (event, station, waveform or progressive search) returns a list and the single outcome can be explored in detail. Fig. 6 shows the user interface for the selection of a seismic event. The details relative to a single event can be retrieved from the resulting list. Both recording stations and events are mapped through the Google-Map data©, which allows to display point data alternatively on a satellite image or a basic map (or both), as shown in Fig. 7.

In the same way, a single station can be searched, as shown in Fig. 8, and the geotechnical information can be retrieved from the station detail (stratigraphy, NSPT or Vs/Vp profile, etc.), as shown in Fig. 9.

When a progressive search is made, the user can start with the event characteristics and gradually constrain the search specifying station and the waveform parameters (Fig. 10). In this case, only the selected waveforms can be downloaded in a compressed file format which contains unprocessed and processed acceleration, velocity, displacement time-series and acceleration response spectra.

The waveforms that satisfy the required conditions can be displayed with the aid of a Java applet (Fig. 11), which allows the user to perform simple operations like zoom in/out, modify display options (axis labels, axis limits, background and foreground colour, etc.) and plot saving or printing. The strong-motion recordings selected through a query can be downloaded by the web clients in raw and processed format, together with the velocity and displacement time-series and acceleration response spectra.

6. Conclusions and future developments

The Italian strong motion database has been designed to be a useful tool for scientific research in the seismological and engineering fields and in particular for data analysis focused on seismic risk assessment. The main features of the relational database and the web portal for data dissemination have been illustrated. The web database can be accessed at http://itaca.mi.ingv.it, where a wide range of search tools enables the user to interactively retrieve events, recording stations and waveforms with particular characteristics, whose parameters can be specified, as needed, through user friendly interfaces. A range of display options allows users to view data in different contexts, extract and download time series and spectral data.

Several decisions have 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, to improve the knowledge on the site response. This goal can be reached through the analysis of the recorded waveforms and detailed geotechnical site characterization.

Ackowledgements. This work was supported by grants from the agreement between the Istituto Nazionale di Geofisica e Vulcanologia and the Dipartimento della Protezione Civile (Italy) in the time span 2004-2006, project S6.

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.

Ambraseys N.N., Smit P., Douglas J., Margaris B., Sigbjörnsson R., Ólafsson S., Suhadolc P. and Costa G.; 2004: *Internet site for European strong-motion data.* Boll. Geof. Teor. Appl., **45**, 113-129.

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.

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. ING Catalogue (1450 B.C. - 1990). Roma.

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.

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

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., **159**, 286-303. doi:10.1016/j.pepi.2006.07.008.

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**, 11703-11728.

Corresponding author: Lucia Luzi
Istituto Nazionale di Geofisica e Vulcanologia
via Bassini 15, 20133 Milano (Italy)
phone: +39 02 23699285; fax: +39 02 23699458; e-mail: luzi@mi.ingv.it