BOLLETTINO DI GEOFISICA TEORICA ED APPLICATA VOL. 45, N. 3, PP. 113-129; SEPTEMBER 2004

Internet site for European strong-motion data

N.N. AMBRASEYS[1], P. SMIT[1,*], J. DOUGLAS[1], B. MARGARIS[2], R. SIGBJÖRNSSON[3], S. ÓLAFSSON[3], P. SUHADOLC[4] and G. COSTA[4]

[1] *Department of Civil and Environmental Engineering, Imperial College of Science, Technology and Medicine, London, United Kingdom*
[2] *Ministry of Environment and Public Works, Institute of Engineering Seismology and Earthquake Engineering, Thessaloniki, Finikas, Greece*
[3] *University of Iceland, Earthquake Engineering Research Centre, Selfoss, Iceland*
[4] *Dipartimento di Scienze della Terra, Università degli Studi di Trieste, Italy*
[*] *Now at: National Emergency Operations Centre, Zürich, Switzerland*

(Received December 28, 2002; accepted June 11, 2003)

Abstract - The Internet Site for European Strong-Motion Data (ISESD) provides unlimited free access to over 2,000 strong-motion records of earthquakes from Europe, the Mediterranean and the Middle East (EMME). Four mirror sites of ISESD have been operating since 26th March 2002. The URLs of these sites are: www.isesd.cv.ic.ac.uk, smbase.itsak.gr, seismo.univ.trieste.it and www.isesd.hi.is. ISESD provides a basis for improved dissemination of strong-motion data in EMME. There are a number of future improvements to ISESD which would improve its usefulness to seismologists, earthquake engineers and insurance specialists.

1. Introduction

Strong-motion seismology is a rapidly growing research field of great practical value, providing data and models needed in earthquake engineering design. The number of strong-motion accelerometric stations and networks in EMME has been growing rapidly during the last two decades resulting in voluminous strong-motion data, which has stimulated both applied modelling and theoretical studies. This data collection has not been coordinated across state boundaries and within many countries there is more than one organisation involved, in most cases both governmental institutions and private industrial companies. This lack of formal

Corresponding author: N.N. Ambraseys, Department of Civil and Environmental Engineering, Imperial College London, South Kensington Compas, London, SW7 7AZ, U.K. Phone: +44 02075946059, Fax: +44 02072252716, e-mail: n.ambraseys@imperial.ac.uk

organisational structure has resulted in certain difficulties regarding availability of the data.

The strong-motion databank at Imperial College grew out of these environments as an attempt to enhance the availability of high-quality strong-motion data, the ultimate goal of which was to collect data required to predict earthquake-induced action needed for structural design as well as hazard and risk studies. The databank contains data from Europe, the Mediterranean and the Middle East (EMME) and has been disseminated on a CD ROM released in 2000 (Ambraseys et al., 2000).

In 2002 the access to the databank was made possible through the Internet Site for European Strong-Motion Data (ISESD), a project supported by the European Commission, Research-Directorate General XII, Environment and Climate Programme. Through this Internet site it is possible to access 2,213 strong-motion records obtained from 856 earthquakes recorded at 691 different stations. This includes the raw records, corrected and derivative records, linear earthquake response spectra as well as seismological information on the earthquakes and basic data on the sensors. The Internet site, which is based on a relational database concept, provides the users with a dynamic environment that gives direct, unlimited and free access to the most up-to-date information available in the databank. In the following sections the ISESD is described, including notes on the historic aspects as well as future perspective, emphasising the need for harmonisation and standardisation.

In this paper we define a strong-motion record as all three components (two mutually perpendicular horizontal components and a vertical component) of the time-histories recorded by accelerographs, which are seismographs which can record the ground motions that actually cause damage to structures and record ground motions that are almost equivalent to ground accelerations.

2. Past

The first accelerograms were recorded during the Long Beach (11/03/1933) earthquake in California. Accelerographs were not installed in EMME until later. The establishment of permanent strong-motion networks in EMME was slower than in the USA and Japan. The first accelerogram recorded in Europe was of an earthquake on 02/12/1967 recorded at Debar (Macedonia) on a Teledyne AR-240.

The routine collection of strong-motion data at Imperial College started in 1971. Research personnel involved with the collection and processing of such data includes: N.N. Ambraseys, J.J. Bommer, J. Crowther, J. Douglas, G. Eleftheriades, M. Free, J. Menu, S.K. Sarma, K. Simpson, P. Smit and M. Srbulov. Imperial College installed Kinemetrics SMA-1s in Ghir (Iran) in 1972 and in Peshawar (Pakistan) in 1975. The collection, processing and dissemination of strong-motion data by Imperial College has greatly benefitted from collaboration with A. Moinfar (Planning Organisation, Tehran, Iran), B. Mohammadioun (Commissariat à l'Energie Atomique, France), M. Basili, D. Rinaldis and R. Berardi (Ente Nazionale per l'Energia Elettrica and Ente per le Nuove Tecnologie, l'Energia e l'Ambiente, Italy) as well as the authors of this paper.

In 1992 it was estimated that about 2,400 permanent and temporary strong-motion instruments were then in operation in the Eurasian region run by about 200 different organisations (Bommer and Ambraseys, 1992).

Bommer and Ambraseys (1992) provides details of the VAX-based database and databank. In 1998, this database and databank was converted to Microsoft Access system running on a PC. As part of this process a new file format and slightly different database concept was adopted (see below).

In 2000, a CD ROM containing 1,068 mainly triaxial strong-motion records was released (Ambraseys et al., 2000). This CD ROM contains most of the strong-motion data from EMME that was then archived in the databanks of the partners of this project (Imperial College of Science, Technology and Medicine; Ente Nazionale per l'Energia Elettrica; Ente per le Nuove Tecnologie, l'Energia e l'Ambiente; Institut de Protection et de Sûreté Nucléaire).

The CD ROM provided a snapshot of strong-motion data then available in EMME. However, the amount of strong-motion data is continually increasing and so, although the CD ROM remains a useful resource, it no longer fully represents the current state of strong-motion data in EMME. Also, some of the associated parameters of the time-histories on the CD ROM have been re-evaluated because new information has become available and hence the CD ROM does not provide the most up-to-date information.

3. Present

An Internet site, known as the Internet Site for European Strong-Motion Data (ISESD), without the above mentioned short commings of the CD ROM was established from which users can freely download strong-motion data and associated parameters. Due to the nature of the Internet and of the relational database underlying the site, new data can be easily added to the databank and changes to the associated parameters can be made. Four mirror Internet sites were established in each of the partners' institutions so that there is redundancy in the system. This is important because of the, still, fragile nature of Internet sites.

Fig. 1 shows the increase in strong-motion records from EMME over the past 33 years. It shows that for the first decade of strong-motion data collection in EMME little data from earthquakes of engineering significance was collected because of the sparsity of instruments. However, since the early 1980s a considerable quantity of strong-motion data has been collected; in total 1,697 records are included in the Imperial College strong-motion database from earthquakes with $M_s \geq 4$. In the past five years temporary stations were installed in the areas of two aftershock sequences (the Umbria-Marche 1997–1998 sequence and the Kocaeli-Düzce 1999 sequence) leading to a steepening in the rate of increase of data. The graph also shows that ISESD provides significantly more data from earthquakes of engineering significance (686 more records) than does the CD ROM of Ambraseys et al. (2000), whose total stops at 1999 and excludes a significant number of records that have become available since the CD ROM was produced. In addition, the graph shows that there is a significant number of records (298 records) which are currently not available for download on ISESD because either

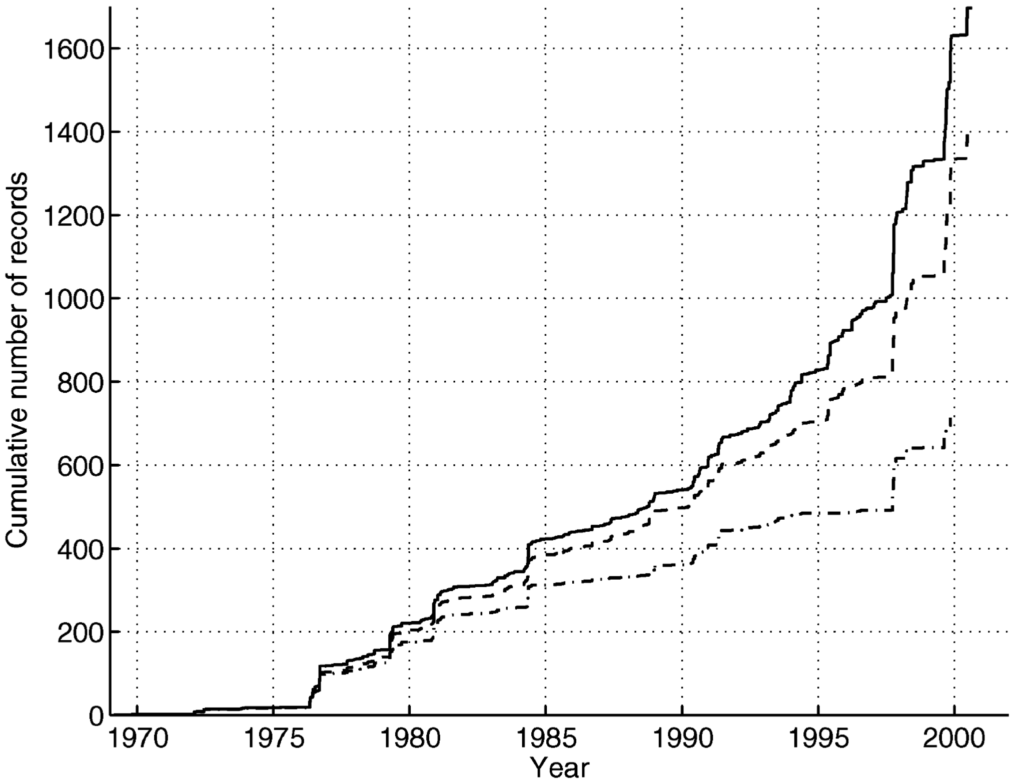

Fig. 1 - Cumulative number of strong-motion records from earthquakes within the complete strong-motion database of Imperial College (solid line), the ISESD database (dashed line) and on the CD ROM of Ambraseys et al. (2000; dash-dotted line).

the owners of the data do not allow the data to be disseminated or because the strong-motion records are thought not to be of high enough quality to be placed on ISESD. Records were rejected due to: insufficient bit range of the A/D convertor used, late triggering of the instrument, insufficient sampling rate, spikes, poor digitization and early termination during the coda.

The basic procedure for downloading data from ISESD is given below:

1. If a user has not already registered then the user completes an online form giving their name, organisation, email address and a choice of password. The three reasons why registration is included in the procedure are: a) to enable an automatic email to be sent to the user containing the associated parameters (e.g. magnitude and epicentral distance) of the strong-motion records they download, because the strong-motion files do not contain these parameters; b) to keep a list of records which have been downloaded so that if an organisation that provided data wants to know how many times their records have been downloaded this information could be provided; c) to justify to ISESD's funding agencies that it is a successful project;
2. once registered, the user specifies a selection criteria in terms of earthquake (e.g. magnitude and focal mechanism), station (e.g. local site category) and waveform (e.g. epicentral distance) parameters;

3. the strong-motion records fulfilling the selection criteria are then listed. The acceleration time-histories and response spectra for each record can then be viewed to further refine the selection;
4. the user selects which strong-motion records they want to download;
5. an email is sent to the user containing details of the project and also a comma-delimited file containing the associated parameters of the records selected;
6. the strong-motion records (uncorrected and corrected) and associated linear elastic response spectra files can be downloaded either via the URL given in the email (the files remain on the ISESD servers for two weeks so that users can choose when they download the records) or directly via the ISESD webpage.

At present no firm selection criteria has been established for what data is included in the ISESD databank. The seismicity of EMME varies considerably from high seismicity in the Mediterranean (e.g. Greece, Italy and Turkey) to low seismicity in north-western Europe (e.g. the United Kingdom and Norway). Therefore a single selection criteria based on magnitude would lead to the inclusion of many records from the high-seismicity regions and very few records from the low-seismicity regions. However, although most data that has been recorded in the low-seismicity areas is of low amplitude it is important for engineers and seismologists working in those areas and hence its inclusion in ISESD is important. Hence the current soft selection criterion is earthquakes with $M \geq 4$ (any magnitude scale) for high-seismicity regions and no selection criterion in low-seismicity regions where there is much less data.

A selection criterion based on the amplitude of the ground motions (for example peak ground acceleration greater than 0.05 g) is not used nor is a selection criterion based on distance (for example epicentral distance less than 200 km) because such criteria can introduce bias into studies based on data from ISESD.

Each strong-motion record is visually inspected to assess its quality. If the record is judged to be of low quality then it is included within the database and the associated parameters are shown on ISESD but the waveform files cannot be downloaded. Currently, whether a record is of too low a quality to be available for download is assessed simply by eye, however, a more rigorous selection procedure is to be introduced (see below).

The associated parameters (e.g. hypocentral location, magnitude estimates) are obtained from a number of different sources. If a special study has been made of the earthquake then the parameters given in the article are used. If such a study has not been undertaken then the parameters are usually taken from the Bulletin of the International Seismological Centre. Local magnitude estimates are taken from bulletins published by local or national networks. Locations given by local or national networks are used if they are thought to be more reliable than those in the Bulletin of the International Seismological Centre. If there is more than one local or national network operating in an area, for example in Greece, the location and local magnitude believed to be the most reliable is included in the database. The sources of the associated parameters are contained in database and are returned to the user.

The main source of local site information are the operators of the strong-motion station. However, articles about strong-motion data recorded at a particular station often have valuable information on the local site conditions. Unfortunately, for most strong-motion stations in the database detailed information on the site conditions is lacking or is of limited detail.

3.1. Databank and database structure

The relational database on which ISESD is based has five main tables: earthquake, station, waveform, address and reference. In addition there are seven other tables: Flinn-Engdahl region, intensity value, waveform quality, country, topography, building type and geology. These secondary tables simply contain the different categories that can be specified for the named associated parameters (e.g. country holds the names of EMME countries).

In line with the concept of relational databases no information is stored in more than one place so that if information is changed then this change is made throughout the database for all associated data. For example, if the local site classification of a station is altered because new information is available then when the database is queried for waveforms recorded at that site the new site classification is returned.

The database is based on five identity numbers, one for each entry in the main tables: earthquake ID, station ID, waveform ID, reference ID and address ID. These numbers are used to link together the information contained in each of the main tables.

The earthquake table of the ISESD database contains the associated parameters that characterise the source, for example the date and origin time, hypocentral location and magnitudes (M_w, M_s, m_b, M_L).

The station table contains parameters that characterise the stations which recorded strong-motion data in the databank, for example latitude, longitude and altitude and local site conditions. Currently, stations are classified using the Boore et al. (1993) scheme of four site categories: very soft soil ($V_{s,30}$ < 180 m/s, soft soil (180 ≤ $V_{s,30}$ < 360 m/s), stiff soil (360 ≤ $V_{s,30}$ < 750 m/s) and rock ($V_{s,30}$ ≥ 750 m/s), where $V_{s,30}$ is the average shear-wave velocity in the top 30 m. This classification is made using shear-wave velocity profiles if available (the actual $V_{s,30}$ value is then also stored in the database) or other information if these profiles are not available. For stations without shear-wave velocity profiles, descriptions of the local site conditions were used to assess the site classification. The source of the information used is given in the database and is returned to the user. Out of the current ISESD databank of 2,213 records, 32% have a known $V_{s,30}$ estimate, 44% have a site classification assessed from a description of the site and 24% have an unknown site classification.

The waveform table of ISESD contains information that categorises each strong-motion record, for example epicentral distance and distance to the surface projection of the rupture. It provides the link between the earthquake and station tables by including the earthquake ID of the earthquake recorded and station ID of the station where the waveform was recorded. The epicentral and hypocentral distances given in the waveform table are calculated using the information given in the earthquake and station tables (these are updated if the location of the earthquake is reassessed). Strong-motion parameters are also listed in the waveform table so that searches based on these parameters can be easily performed.

The reference table holds details of the references for each reference number. Each associated parameter given in the ISESD database is given a reference number. This is important for such a large database with data and associated parameters collected from many different sources. When strong-motion data is downloaded from ISESD the references for the original source strong-motion data, the main earthquake and station parameters and the local site conditions are given.

The address table contains the contact details of strong-motion network operators in EMME. The name of the strong-motion network from which the station that recorded the downloaded strong-motion record is returned by ISESD so that users can correctly cite the strong-motion data they use. Also the information in the address table is used for the “Network” pages of ISESD, which provide information on the strong-motion networks operating in EMME.

3.1.1. File format

For the CD ROM released in 2000 (Ambraseys et al., 2000) a new strong-motion data format was created. At present there are many other strong-motion data formats in use, for example: COSMOS (COSMOS Strong Motion Programs Board, 2001), CSMIP (Shakal and Huang, 1985), K-NET (Kyoshin Net, 2002), SAC (Goldstein, 2002), SMC (National Strong-Motion Program, 2002) and numerous unnamed ASCII formats. It was felt, however, that none of these existing formats was ideal for use with a relational database structure. Hence, a new ASCII format was developed (see the “CD ROM” pages of ISESD for details of the format and see the “Source code” pages of ISESD for subroutines in FORTRAN and MATLAB® for input and output of files in the ISESD format). Although a binary file format would result in smaller file sizes and quicker file input and file output times than an ASCII format, an ASCII format was selected so that the files are easy to view and manipulate. In addition the files are compressed before download so the slightly larger file size compared with a binary system is not a problem.

The ISESD format is specially designed for a relational database structure. This means that information is only stored once within the databank and database and so no inconsistencies between information given in two different places can occur. In other strong-motion file formats, such as SMC (National strong-motion Program, 2002), there are fields in the header of the record for magnitude and distance values. Such information is already stored in the ISESD database and so duplicating it in the header of each file would mean that if the magnitudes or distances were reassessed then the information would need to be changed in two places otherwise confusion could occur. Therefore the header of waveform files in ISESD format only includes information that is unlikely to change and also are required for appropriate processing of the accelerogram, for example: instrument type and natural frequency and critical damping ratio of the transducer. Also specified in the header of every file is the instrument operator so that the correct source of the data can be cited; this is particularly important for ISESD which includes data from over fifty different organisations. There is also space in the header for as many lines of comments as required for providing details of the digitisation and processing of the record. The three required identification numbers (earthquake ID, station ID and waveform ID) are given in the header of each file so that the user can link the record to the correct earthquake and station.

Each waveform in the databank is given an individual six digit number (the waveform ID) and the three components are labelled *x*, *y* (for horizontal directions) and *z* (for vertical directions) and the filename extension gives the type of information contained in the file (*raw* for uncorrected accelerations, *cor* for corrected accelerations and computed velocities and *spc* for response spectra). This file naming system is transparent but allows additional records to be incorporated into the databank.

Data samples in the uncorrected and corrected time-history files are given in six columns of exponential number format with five significant figure accuracy so that any amplitude of ground motions can be included without loss of accuracy (a loss of accuracy for small amplitude ground motions is a problem with some file formats because they do not use exponential number format). The six column format means that the data is readable but it leads to small file sizes, which is important for accessing large quantities of data on the Internet. Most strong-motion data is digitised at a constant sampling rate (usually 100 or 200 samples per second) therefore time ordinates do not need to be provided for each sample in the time-history files only the sampling rate is required. However, some old accelerograms were digitised using irregular time steps. By specifying a time step of -1 in the header of the ISESD file the waveform is identified as having irregular time-steps; the time abscissae are then given alternately with the acceleration ordinates. Seekins et al. (1992), overcame the problem of irregular sampling by linearly interpolating the irregular time-histories they provided to 200 samples per second. Linear interpolation to a higher sampling rate, than the original accelerogram was digitised at, does not retain information on the irregular sampling rate used for digitisation and so it is impossible to assess the highest frequency to which information can be extracted from the time-histories. SI base and derived units are used in all files in accordance with European standards.

3.2. Benefits of an EMME strong-motion Internet site

There are a number of benefits of an Internet site that enables users to download strong-motion data from stations across EMME.

Firstly, strong earthquakes are sometimes recorded on strong-motion instruments located in a number of different countries because they occur in border regions. Therefore for a complete set of strong-motion data from a particular earthquake it is important that data from each of the countries where the earthquake was recorded is available. An example of such an area is the southern Alps, where accelerograms of the same strong earthquake can be recorded in Austria, Italy, Slovenia and Switzerland. For example, important strong-motion records were obtained in both Italy and Slovenia of the Friuli earthquake sequence of 1976 and the Bovec (12/04/1998) earthquake.

Within some European countries there are a number of separate accelerographic networks operating. Instruments operated by different institutions are likely to record the same strong earthquakes occurring in the country. Therefore it is valuable to include data from each of the networks. An example of a country where a number of different accelerographic networks are currently operating is Greece, where there are at least three different networks (National Observatory of Athens; Institute of Engineering Seismology & Earthquake Engineering; and Broad-Band and Strong-Motion Stations in Western Corinth Gulf operated by Charles University and University of Patras). Another example where a number of different networks were operating in the same area was north-west Turkey after the Kocaeli (17/08/1999) earthquake, where at least eight strong-motion networks (Çelebi et al., 2001) were operating (Kandilli Observatory, Bogazici University; Earthquake Research Department, General Directorate of Disaster Affairs; Faculty of Civil Engineering, Istanbul Technical University; LGIT, Universite Joseph Fourier; Earthquake Research Institute, University of Tokyo; Earthquake Hazards Team, U.S. Geological Survey; Geologic Hazards Team, U.S. Geological Survey; Lamont Doherty Earth Observatory of Columbia University).

When data from the same earthquake is available from a large number of different networks it is important that the associated parameters (e.g. magnitude and source-to-site distances) are uniformly assessed for each strong-motion record. For example, if data is available from two different networks, which have used two different hypocentral locations for the calculation of epicentral and hypocentral distances then the inconsistency in source-to-site distances could lead to erroneous results when, e.g. the attenuation of ground motions is analysed. In addition, the data is easier to use if it is available in a uniform file format so that a single computer program can be used for all analysis. These two basic requirements for consistency and ease of use are both satisfied by ISESD, for which the associated earthquake parameters are consistently assessed for all strong-motion data and the strong-motion data is provided in a uniform file format (see above).

3.3. Current status of ISESD

As of 26/11/2002, there are 2,213 mainly triaxial strong-motion records available for download from 856 earthquakes recorded at 691 different stations. In addition there are 1,304 records that were recorded in EMME but are currently not available for download due to the above mentioned reasons.

So far 56 different organisations have contributed data to ISESD. These organisations are acknowledged on the ISESD 'Acknowledgements' page. We are grateful to all of these organisations for their invaluable contributions.

The strong-motion data available on ISESD is of mixed quality reflecting the high number of different sources of data, from different types of instruments used and the wide variety of different digitisation techniques employed. Table 1 shows the distribution of instrument types of strong-motion records currently on ISESD. This table shows that although digital instruments are being installed throughout EMME, a significant proportion of the strong-motion data currently available comes from analogue instruments, in particular SMA-1s. It is causing difficulties that

Table 1 - Distribution of records with respect to accelerograph type in the ISESD databank.

Type	Number of records	% of records	Type	Number of records	% of records
SMA-1	1008	45.5	TERRA TEK	77	3.5
Unknown	394	17.8	K2	73	3.3
SMACH SM2	106	4.8	SSA-2	70	3.2
SSA-1	104	4.7	GSR-16	61	2.8
SMACH SM1	81	3.7	Other types	239	10.8

the type of instrument is not known for a significant proportion of the records currently in the ISESD databank (17.8%). The instrument type is required to perform an appropriate correction procedure. The sampling periods of regularly digitised uncorrected waveforms archived on ISESD varies from 0.00119 to 0.0556 s, and 288 records have been irregularly digitised with sample periods varying from 0.0002 to 0.7896 s. Therefore some records on ISESD cannot be used for studies requiring high-frequency ground motions (> 25 Hz).

As of 26/11/2002, 334 users from over 170 different organisations have registered with one of the four ISESD mirror sites and have downloaded strong-motion data. Table 2 shows the distribution of organisations who have used ISESD with country of origin. It shows that there is wide interest in the data provided on ISESD. People from non-governmental agencies, research institutes, universities, civil engineering companies and insurance companies have registered and downloaded records from ISESD.

Table 2 - Distribution of organisations who have used ISESD with country of origin (only countries with more than two organisations are listed).

Country	Number of different organisations	Country	Number of different organisations
USA	29	Switzerland	5
United Kingdom	15	Austria	4
Germany	14	Romania	4
Italy	12	Australia	3
Turkey	11	Colombia	3
Greece	7	France	3
Iceland	5	Iran	3
Spain	5	Macedonia	3

So far 8,581 records have been downloaded by users of ISESD. The ten most commonly downloaded records are given in Table 3. This table shows that, as expected, the most commonly downloaded records are those from large magnitude earthquakes ($M_w \geq 6.0$) recorded at close distances ($d_f \leq 20$), i.e. those records of ground motions which are of high engineering significance. Another important point shown by Table 3 is that the most commonly downloaded time-histories are those that were recorded on rock and hence can be used as the input to site-specific studies. This finding demonstrates that it is important for station operators to locate some strong-motion instruments on rock sites and to provide evidence of the local site conditions so that the users of the records have confidence in using accelerograms classified as ‘rock’ records.

4. Future

The ISESD project was funded by the European Commission and the funding ended on 31/03/2002. Therefore currently ISESD is unfunded. The partners of the project have agreed to continue operating the four mirror sites and to update the sites when new data becomes available or updates are required. It is important, however, that a continuing source of funding is found if the Internet sites are to continue operating.

Table 3 - The ten most commonly downloaded strong-motion records from ISESD in descending number of downloads, where M_w is moment magnitude, M_s is surface-wave magnitude, d_e is epicentral distance and d_f is distance to surface project of rupture.

Date	M_w	d_f	Station	Site category	Horizontal PGA (ms^{-2})
06/05/1976	6.5	6	Tolmezzo-Diga Ambiesta	rock	3.60
13/03/1992	6.6	1	Erzincan-Meteorologij Mudurlugu	stiff soil	4.92
16/09/1978	7.4	3	Tabas	stiff soil	10.22
04/11/1973	5.8 (M_s)	11	Lefkada-OTE Building	soft soil	5.27
26/09/1997	6.0	4	Nocera Umbra	rock	5.51
18/05/1978	3.6 (M_s)	8 (d_e)	Patra-OTE Building	soft soil	0.28
23/11/1980	6.9	6	Bagnoli-Irpino	rock	1.87
06/04/1977	6.0	4	Naghan	rock	10.13
16/09/1978	7.4	11	Dayhook	rock	3.79
17/05/1976	6.7	4	Gazli	very soft soil	7.04

There are a few Internet sites providing strong-motion data on an almost real-time basis, for example Kyoshin Net (K-NET) in Japan, SeisWeb operated by the University of Bergen in Norway and Reseau Accelerometrique Permanent (RAP) in France. The development of ISESD to a near real-time system is considered as a long term goal although it would require considerable funding and effort.

Currently ISESD only includes ground response strong-motion data which has not been significantly affected by the response of the building in which the instrument was located. Strong-motion data from instruments that were located in the upper floors of structure or from dams are useful data for engineers who analyse structural response to earthquakes. These data would be a valuable addition to ISESD but would require that detailed information on the structure's geometry and material was given to the user of the data. This is not a current priority for ISESD.

One important area where the information of ISESD should be improved is in local site characterisation. The local site information currently included in the database and reported to the user of ISESD is not detailed enough for some purposes. For many stations (295 stations) no site information is available. The strong-motion records from these station (926 records) are therefore of limited use for most purposes where information about the local site conditions is vital. Some of these stations are no longer operating and so it is unlikely that the local site conditions at these stations will be investigated in the future. Even for stations that are still operating detailed investigation of the local site conditions is expensive and time consuming and so it may not be undertaken unless a large earthquake occurs in the region. Therefore, it is planned to attempt to classify the strong-motion stations without site categories using a variety of methods that do not rely on visiting the station. These techniques include: geological maps, sediment depth maps (e.g. Seber et al., 2001), horizontal to vertical ratio methods using strong-motion records (e.g. Zare et al., 1999), peak ground velocity to peak ground acceleration ratio methods (e.g. Decanini et al., 2000) and reference site techniques (e.g. Harmsen, 1997). However, detailed local site investigations, preferably including measurements of the near-surface wave velocities are preferable to these approximate methods and should be undertaken.

One important way that ISESD could be improved is for each time-history to be individually corrected to remove the short- and long-period noise. At present almost all the corrected strong-motion records accessible on ISESD have been bandpass filtered using an elliptical filter with cut-off frequencies of 0.25 and 25 Hz. The transducer response has not be removed because of a lack of detailed information on instrument type, natural frequency and critical damping ratio for many of the waveforms. This uniform correction procedure means that care must be taken when using the filtered strong-motion data downloaded from ISESD in the short- and long-period ranges. The uncorrected time-histories of almost all of the records are also available so that users can apply their own correction procedure. An individual correction of each time-history based on estimates of the noise present in the record and a removal of the instrument response would significantly improve the strong-motion data available on ISESD but this requires a significant effort. Providing individually corrected records would reduce the onus on the user to process the uncorrected records themselves and would increase their confidence in using the data in the short- and long-period ranges.

As mentioned above the current selection criterion for the exclusion of low quality time-histories from the downloaded databank is based on a visual inspection of the uncorrected acceleration time-histories. Douglas (2002, 2003) has recently suggested a more rigorous procedure for this selection based on the ability of poor-quality strong-motion records to yield good estimates of response spectral ordinates within the period range of most engineering interest. This more rigorous selection method will be used in future for the selection of strong-motion records for ISESD.

One additional table that could be added to the ISESD database is an instrument table containing instrument parameters such as instrument type and installation date. This would allow the instrument history of a single station to be known, for example when the instrument in a station changed from analogue to digital. Also it would allow more than one record for a single earthquake to be archived if two instruments were operating at the same station, an example of this is the Nocera Umbra station during the Umbria-Marche earthquake sequence of 1997 and 1998 when an SMA-1 instrument and a Kinemetrics Etna instrument were operating at the same station (Servizio Sismico Nazionale - Monitoring System Group, 2002). The reason an instrument table is not used is so that the database is less complex to maintain. In addition co-located instruments are not common in EMME. The header of each waveform contains information on instrument type (see below). In the current ISESD database a new station entry is created if two instruments were co-located and so the lack of an instrument table is not a large problem.

4.1. Estimates of number of new strong-motion records likely to be produced by European strong-motion networks

It is important to estimate how many accelerograms from earthquakes of engineering significance are likely to be recorded every year in EMME so that the resources required to keep the ISESD database and databank up-to-date can be assessed. These estimates are made herein assuming that there are no new permanent stations established and that no temporary instruments are installed to record aftershocks following mainshocks. Therefore such estimates are most likely lower bounds on the number of time-histories likely to be recorded in the future.

The methodology adopted to estimate the number of records likely to be produced is based on simulating an earthquake catalogue and then using an appropriate attenuation equation for peak ground accelerations to estimate which accelerographs will trigger, assuming a constant trigger threshold for simplification. The number of accelerographs that trigger for each earthquake are then summed over the entire catalogue to give the total number of records produced.

Source zones were categorised by their location, frequency-magnitude parameters (a and b in $\log N = a - bM$, where N is the number of earthquakes larger than magnitude M) and maximum magnitude expected in each zone (M_{max}). An artificial earthquake catalogue was generated using these parameters for 100 years assuming the earthquakes were uniformly distributed within each source zone. No time dependence was introduced into the catalogue so that the obtained estimated numbers of strong-motion records should be interpreted as the average totals for a given period of time. Estimated peak ground accelerations were computed for each of the earthquakes in the catalogue at all of the stations within the ISESD database using published equations for estimating strong ground motions, including the associated scatter of the equations (truncated to 1.5 times the standard deviation). When the estimated peak ground accelerations at a station was above the trigger level of the instrument then it was assumed to have recorded. This procedure generates an artificial strong-motion catalogue for the period of simulation and can be used for a number of purposes. For example, it can be used to assess the ability of strong-motion networks to adequately record the earthquakes that are likely to occur within the region. Furthermore, it can be used to test proposed network configurations before implementation. In this article only the estimated number of accelerograms recorded in the future will be considered.

Equations for the prediction of vertical peak ground accelerations were used because accelerographs such as the SMA-1 trigger on vertical accelerations above a threshold acceleration. Different threshold acceleration levels were used: 0.005 and 0.01 g (which are standard threshold trigger levels) and 0.05 g, which corresponds to a lower limit peak ground acceleration for recording small magnitude earthquakes and aftershocks (McLaughlin, 1991).

Since in this paper the interest is in predicting whether an instrument triggers, previously published equations for the prediction of peak ground acceleration are not completely suitable. This is because in regression analysis to find such equations only peak ground accelerations from instruments that triggered are used and not those from untriggered instruments. This leads to an overprediction of triggering of distant stations (McLaughlin, 1991).

For illustration, estimates of the number of strong-motion records from earthquakes with $M_s \geq 4.0$ are included for two countries, Italy and Iceland.

For Italy the source zones provided by Gruppo Nazionale per la Difesa dei Terremoti (1996) were used for the calculation and the catalogue provided by Boschi et al. (1999) was used to estimate a values for each of these source zones. b values for each of the source zones were taken from Molchan et al. (1997). The maximum magnitude used for all computations was M_s = 7.0. The equation by Ambraseys and Simpson (1996) for the prediction of vertical peak ground acceleration was used. The standard deviation of the equation was included in the estimation of the ground motions; the scatter was truncated at 1.5 times the standard deviation. The estimated number of strong-motion records from earthquakes with $M_s \geq 4$ is about 7,000 over the 100 year simulation (assuming a trigger level of 0.01 g), corresponding to about 70 per year. Currently the ISESD database has 508 strong-motion records from earthquakes that were recorded in Italy from 1972 to 1998 (27 years). Therefore the estimated number of records for the future 100 years may seem quite high. However, the number of accelerographs operating in Italy has grown significantly since 1972, especially in the early 1980s and with the establishment of the Friuli-Venezia Giulia Accelerometric Network (RAF; Costa et al., 1998) in 1993-1995, and so the number of earthquakes recorded also has increased with time. Also the simulation assumes strong-motion instruments are 100% reliable and that all records are digitised and retained even if the ground motion amplitudes are low.

For Iceland, nine source zones were defined based on the areas of highest seismicity (South Iceland seismic zone and the Tjornes Fracture zone) and their a, b and M_{max} values were assessed based on the earthquake catalogue for Iceland compiled by Ambraseys and Sigbjörnsson (2000). The seismicity associated with the volcanic zones and central volcanoes was not included in the simulation because there are no strong-motion instruments nearby. The equation by Sigbjörnsson and Baldvinsson (1992) was used for the prediction of horizontal PGA (most digital instruments in Iceland are triggered based on a combination of all three components and so an equation for vertical PGA was not used) and again the scatter was truncated at 1.5 times the standard deviation. The estimated number of strong-motion records from earthquakes with $M_s \geq 4$ is about 1,500 over the 100 year simulation (assuming a trigger level of 0.005 g), corresponding to about 15 per year. Currently the ISESD has 470 strong-motion records from earthquakes that were recorded in Iceland from 1986 to 2000 (15 years). Therefore the estimated number of records for the future 100 years seems slightly low. However, a large proportion (330 or 70%) of the records currently in the databank from Iceland either have no M_s given or are from earthquakes with $M_s < 4$. Therefore the future estimate of about 15 records from earthquakes with $M_s \geq 4$ seems reasonable.

5. Conclusions

In the past much money has been spent on the installation and maintenance of accelerographic networks and the collection and processing of strong-motion records. Hudson (1979) estimated that each important strong-motion record costs about $10,000 in terms of instrument installation and maintenance and processing. Due to inflation it is likely that the current cost of each important strong-motion record is three or four times higher than estimated in 1979. Therefore, so that most use can be made of the important data already recorded a freely-accessible source of strong-motion records and associated parameters is valuable to both the scientific, engineering and insurance sectors.

The ISESD, which seeks to provide such a resource, has been successfully operating since March 2002. A system is in place to keep ISESD expanding, therefore it is hoped that the databank can continue to grow.

At the ESC General Assembly in September 2002 it was pointed out that the main areas where ISESD should improve are in site characterisation and in the collection of data from proprietary sources. It was also mentioned that it is important that a standard strong-motion data format is established and standardised for easy data exchange. We believe that the ISESD file format provides such a data standard.

Many of the problems encountered during the creation of the ISESD and many of its short comings (e.g. poor site characterisation and file format standardisation) have been mentioned many times before including in a panel discussion session held ten years ago at the Tenth World Conference on Earthquake Engineering (Brady, 1992). Therefore it is important that these problems are finally addressed.

The estimated number of strong-motion records from earthquakes of engineering significance occurring within EMME and recorded on existing instruments is expected to equal a few hundred each year. This is based on 70 records per year from Italy and 15 records per year from Iceland, which are two of the most seismically active areas of the region covered by ISESD. To import new records into the database and databank of ISESD does not take much time therefore continual funding for one or two part-time personnel to continue updating and improving ISESD would be all that is required to maintain this resource.

Acknowledgments. We thank the European Commission, 4th and 5th Framework Programmes for funding the CD ROM (contract ENV4-CT97-0397) and ISESD (contract EVR1-CT-1999-40008) projects and the Engineering and Physical Sciences Research Council for previous grants which helped establish the Imperial College strong-motion database. We also thank all the individuals and organisations that have contributed data and time to these projects (see the ‘Acknowledgements’ page of ISESD for details), without their unselfish support these projects would have been impossible. Finally, we thank Ina Cecic, Mariano Garcia-Fernandez and Dario Slejko for inviting us to contribute this paper to Bollettino di Geofisica Teorica ed Applicata and to the two anonymous reviewers whose thoughtful reviews improved this paper.

References

Ambraseys N.N. and Sigbjörnsson R.; 2000: *Re-appraisal of the seismicity of Iceland.* Polytechnica: Engineering seismology, Earthquake Engineering Research Centre, University of Iceland, Selfoss, Iceland. 196 pp.

Ambraseys N.N. and Simpson K.A.; 1996: *Prediction of vertical response spectra in Europe.* Earth. Eng. Struct. Dyn., **25**, 401-412.

Ambraseys N.N., Smit P., Berardi R., Rinaldis D., Cotton F. and Berge C.; 2000: *Dissemination of European Strong-Motion Data.* CD-ROM collection. European Commission, Directorate-General XII, Environmental and Climate Programme, ENV4-CT97-0397, Brussels, Belgium, CD-ROM.

Bommer J.J. and Ambraseys N.N.; 1992: *An earthquake strong-motion databank and database.* In: Proc. of Tenth World Conference on Earthquake Engineering, A.A. Balkema. Rotterdam, The Netherlands. vol. 1, pp. 207-210,

Boore D.M., Joyner W.B. and Fumal T.E.; 1993: *Estimation of response spectra and peak accelerations from western North American earthquakes: an interim report.* Open-File Report 93-509, U.S. Geological Survey, 70 pp.

Boschi E., Gasperini P., Valensise G., Camassi R., Castelli V., Stucchi M., Rebez A., Monachesi G., Barbano M.S., Albini P., Guidoboni G., Ferrari E., Mariotti D., Comastri A. and Molin D.; 1999: *Catalogo parametrico dei terremoti italiani*. On Internet at: http: //emidius.mi.ingv.it/CPTI/.

Brady A.G.; 1992: *Panel discussion summary.* In: Proc. of Tenth World Conference on Earthquake Engineering, Balkema A.A. Rotterdam, The Netherlands. vol. 11, pp. 6951-6953.

Çelebi M., Akkar S., Gülerce Ü, Sanli A., Bundock H. and Salkin A.; 2001: *Main shock and aftershock records of the 1999 Izmit and Düzce, Turkey earthquakes.* Open-File Report 01-163, U.S. Geological Survey. CD-ROM.

COSMOS Strong Motion Programs Board; 2001: *COSMOS strong motion data format.* Technical report, Consortium of Organizations for Strong-Motion Observation Systems. On Internet at: http: //www.cosmos-eq.org/cosmos_format_1_20.pdf. 17 pp.

Costa G., Suhadolc P. and Panza G.F.; 1998: *The Friuli (NE Italy) Accelerometric Network: analysis of low-magnitude high-quality digital accelerometric data for seismological and engineering applications.* In: Proc. of the Sixth U.S. National Conference on Earthquake Engineering, Oakland, USA. Earthquake Engineering Research Institute. Seattle, USA. 31 May-4 June. CD-ROM.

Decanini L., Mollaioli F., Panza G.F., Romanelli F. and Vaccari F.; 2000: *Pericolosità sismica della Sicilia sud orientale. Terremoti di scenario per Augusta, Siracusa e Noto.* In: Decanini L. and Panza G.F. (eds), Scenari di pericolosità sismica ad Augusta, Siracusa e Noto. CNR-Gruppo Nazionale per la Difesa dai Terremoti, Rome, Italy, pp. 80-151, in Italian.

Douglas J.; 2002: *What is a poor quality strong-motion record?* In: Proc. of the XXVIII General Assembly of the European Seismological Commission (ESC). Genoa, Italy. 1-6 September. CD-ROM.

Douglas J.; 2003: *What is a poor quality strong-motion record?* Bull. Earth. Eng., **1**. In press.

Goldstein P.; 2002: *SAC data file format.* On Internet at: http: //www.llnl.gov/sac/.

Gruppo Nazionale per la Difesa dai Terremoti; 1996: *Zonazione sismogenetica del territorio nazionale ed aree limitrofe*. On Internet at: http: //emidius.mi.ingv.it/GNDT/ZONE/zone_sismo.html.

Harmsen S.C.; 1997: *Determination of site amplification in the Los Angeles urban area from inversion of strong-motion records.* Bull. Seism. Soc. Am., **87**, 866-887.

Hudson D.E.; 1979: *Reading and interpreting strong motion accelerograms.* Earthquake Engineering Research Institute, Berkeley, USA. 112 pp.

International Seismological Centre; 2002: *On-line Bulletin.* http: //www.isc.ac.uk/Bull. International Seismological Centre, Thatcham, United Kingdom.

Kyoshin Net; 2002: *About K-NET data format.* On Internet at: http: //knetwww.k-net.bosai.go.jp/k-net/man/knetform_en.html.

McLaughlin K.L.; 1991: *Maximum likelihood estimation of strong-motion attenuation relationship*. Earthquake Spectra, **7**, 267-279.

Molchan G., Kronrod T. and Panza G.F.; 1997: *Multi-scale seismicity model for seismic risk.* Bull. Seism. Soc. Am., **87**, 1220-1229.

National Strong-Motion Program; 2002: *SMC-format data files.* On Internet at: http: //nsmp.wr.usgs.gov/smcfmt.html.

Seber D., Sandvol E., Sandvol C., Brindisi C. and Barazangi M.; 2001: *Crustal model for the Middle East and North Africa region: implications for the isostatic compensation mechanism.* Geophys. J. Int., **147**, 630-638.

Seekins L.C., Brady A.G., Carpenter C. and Brown N.; 1992: *Digital Data Series DDS-7. 'Digitized strong-motion accelerograms of North and Central American earthquakes 1933-1986'.* CD-ROM.

Servizio Sismico Nazionale - Monitoring System Group; 2002: *The Umbria-Marche strong motion data set (September 1997-June 1998).* CD-ROM.

Shakal A.F. and Huang M.J.; 1985: *Standard tape format for CSMIP strong-motion data tapes.* Technical Report OSMS 85-03, California Department of Conservation Division of Mines and Geology Office of Strong Motion Studies. On Internet at: ftp: //ftp.consrv.ca.gov/pub/dmg/csmip/Formats/DMGformat85.pdf. 29 pp.

Sigbjörnsson R. and Baldvinsson G.I.; 1992: *Seismic hazard and recordings of strong ground motion in Iceland.* In: Proc. of Tenth World Conference on Earthquake Engineering, Balkema A.A., Rotterdam, The Netherlands. vol. 1, pp. 419-424.

Zare M., Bard P.Y. and Ghafory-Ashtiany M.; 1999: *Site characterizations for Iranian strong-motion network.* Soil Dyn. Earth. Eng., **18**, 101-123.