6ICEGE 6th International Conference on Earthquake Geotechnical Engineering
1-4 November 2015
Christchurch, New Zealand

Archival and near-real-time access to instrumented sites and structures through the European Integrated waveform Data Archive (EIDA), Rapid Raw Strong-Motion Database (RRSM) and Station Book.

J. Clinton[1], C. Cauzzi[2], R. Sleeman[3]
and the EIDA Management Board

ABSTRACT

We present the European Integrated Data Archive (EIDA), which can provide data archival and near-real-time dissemination of waveforms collected by European data centers from instrumented sites and structures. Current EIDA nodes are European institutions collecting and archiving data from seismic networks deploying broadband sensors, short-period sensors and strong-motion accelerometers. EIDA presently delivers seismic data from ~ 5000 stations from ~140 networks. All EIDA data is used to generate the European Rapid Raw Strong-Motion Database (RRSM) which provides earthquake information and engineering parameters within minutes of any event with magnitude > 3.5 in the greater European region. EIDA information is also included in a station book that comprises critical site characterisation parameters, including basic description of instrumented sites. As services in the European Plate Observing System (EPOS) EIDA, the RRSM and the station book are expected to be the long-term solutions for European dissemination of data from instrumented sites and structures.

Introduction and motivation

Seismology, engineering seismology, earthquake and geotechnical engineering today benefit from publicly available databanks that offer researchers and practitioners large amounts of waveform data and associated event and station metadata, often including recordings of geotechnical and structural arrays. These databanks are typically maintained by national institutions in charge of seismic monitoring, and / or governmental agencies with a research or civil protection mandate. Data are usually available through web interfaces often requiring user registration. Regional and global datasets have also been consolidated and homogenised in an effort to provide the community with a single source of high-quality data. Worldwide well-known examples include: i) the European Strong-Motion Database ESMD (Ambraseys et al., 2004); ii) the Pacific Earthquake Engineering Research Institute (PEER) ground-motion database (Ancheta et al., 2014); iii) the Center for Engineering Strong-Motion Data (CESMD). A recent European effort is that of the European integrated waveform data archive (EIDA, Clinton et al., 2014), that provides real-time (or close to real-time) seismic waveform data for the European Rapid Raw Strong-Motion Database (RRSM; Sleeman et al., 2014; Cauzzi et al., 2015), with

[1]Dr, Swiss Seismological Service (SED), ETH Zurich, Zurich, Switzerland, john.clinton@sed.ethz.ch
[2]Dr, Swiss Seismological Service (SED), ETH Zurich, Zurich, Switzerland, carlo.cauzzi@sed.ethz.ch
[3]Dr, ORFEUS Data Centre, Royal Netherlands Meteorological Institute (KNMI) - Seismology Division, De Bilt, Netherlands, reinoud.sleeman@knmi.nl

basic station metadata and station characterisation information being collected in a Station Book. It is the goal of this conference contribution to describe how these existing European infrastructures could offer a mature and sustainable framework for the dissemination of data and metadata of seismically instrumented sites and structures.

Overview of the European Integrated waveform Data Archive (EIDA)

In 2013 ORFEUS (Observatories and Research Facilities for European Seismology) established EIDA (Figure 1) a coordinated distributed data archive system, in order to expand and strengthen the availability of seismic data to the seismological community. This innovative waveform data infrastructure provides open, unified access to continuous waveform data from more than 5000 stations from 83 permanent and 61 temporary networks (July 2nd 2015). EIDA nodes are datacentres that collect and archive data from seismic networks deploying broadband and short-period velocity stations and broadband accelerometer stations. Current EIDA nodes are: ODC (Europe), GFZ (Germany), RESIF (France), INGV (Italy), ETH (Switzerland), BGR (Germany), IPGP (France), LMU (Germany), NIEP (Romania) and KOERI (Turkey). A common, uniform web interface (WebDC3) provides transparent access to EIDA (while each node may additionally provide unique, restricted data and additional services. Most permanent networks in Europe currently archive at least a portion of their data within EIDA. The networks contributing to EIDA are listed at http://www.orfeus-eu.org/eida/eida_network_lists.html. The ORFEUS Data Center (ODC) is a primary EIDA node; this means that all data from the Virtual European Broadband Seismograph Network) VEBSN are available through EIDA. The underlying technical architecture of EIDA is based on *ArcLink*, developed by GFZ (German Center for GeoSciences, Potsdam). The *ArcLink* software and protocol is based on TCP and uses simple ASCII coding that allows requesting waveform data based on time windows. An *ArcLink* request is associated with a request_ID that can be used by the client to get the status of the request, to download the data and to delete the request. The *ArcLink* server does not access the data archive directly, but delegates this job to a "request handler". Thus, it is possible to use *ArcLink* for accessing different data archives by using different request handlers. In addition to waveforms and metadata, it is also possible to request routing information from an *ArcLink* server. The routing information tells which *ArcLink* server provides the data of a given station. The routing database itself is supposed to be synchronized between all *ArcLink* servers. In this way a client can connect to any public *ArcLink* server, requests routing information and splits the request accordingly. Each EIDA node provides the resources for management and technical support that ensures the sustainability of this distributed archive system.

Under the mandate of the ORFEUS Board of Directors and Executive Committee, the EIDA nodes are responsible for steering and maintaining the technical developments and organization of EIDA and the integration within multidisciplinary frameworks. Embedding the EIDA infrastructure in the mature foundation of ORFEUS ensures a sustainable system. EIDA has become a key element in EPOS (European Plate Observing System), the European integrated research infrastructure for solid-Earth sciences. Current developments are directed towards ensuring coordination and compatibility with future EPOS thematic and integrated services; the integration of strong-motion data; improving data quality for research; standardising and expanding data collection; and developing quality parameter standards. EIDA also intends to integrate seismic data from mobile experiments by European organisations, including OBS, and

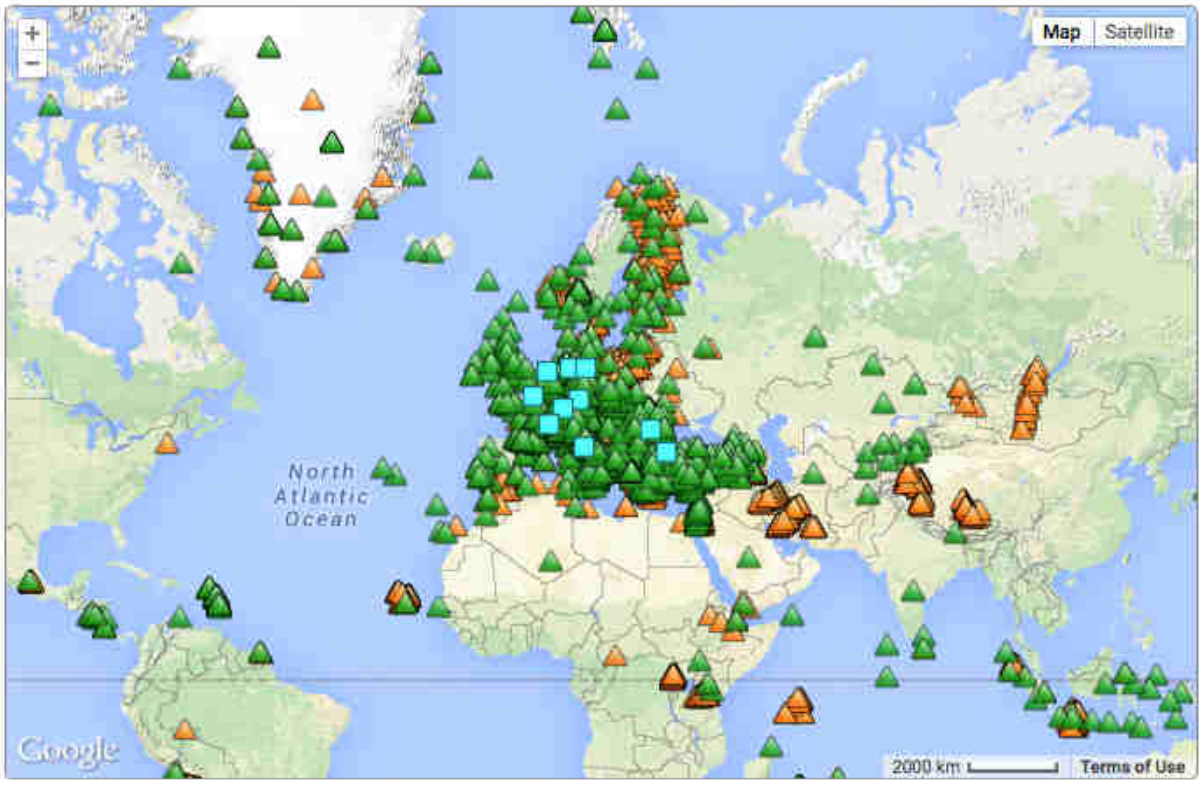

Figure 1. Screenshot of the EIDA homepage at ORFEUS Data Centre. The map shows all the stations presently available in the distributed archive (last accessed 02.07.2015). Permanent stations are depicted as green triangles, while orange triangles are temporary installations. The light blue squares are the EIDA nodes.

other types of data (e.g. infrasound, sequences from induced earthquakes, ground motion synthetics) in its archives.

The EIDA nodes also provide access to their data through web services that are compatible with international standards defined at the International Federation of Digital Seismograph Networks (FDSN). The FDSN web service implemented at ODC (http://www.orfeus-eu.org/man/fdsnws_station.html) provides access to metadata from all open EIDA stations. With this distributed infrastructure ORFEUS expands its archives with open broadband (BB) data from European networks and research infrastructures and creates a robust long-term archiving system for high quality data with modern data dissemination services (interactive web interfaces and webservices, http://www.orfeus-eu.org/eida/eida_odc.html).

The European Rapid Raw Strong-Motion Database (RRSM) and Station Book

The RRSM relies on the raw waveform data and station metadata available in EIDA and basic earthquake information (location and magnitude) automatically provided by the Euro-Mediterranean Seismological Center (EMSC) within 5-20 minutes of an earthquake origin time (OT). The RRSM database is populated by a waveform-processing module, namely *scwfparam*

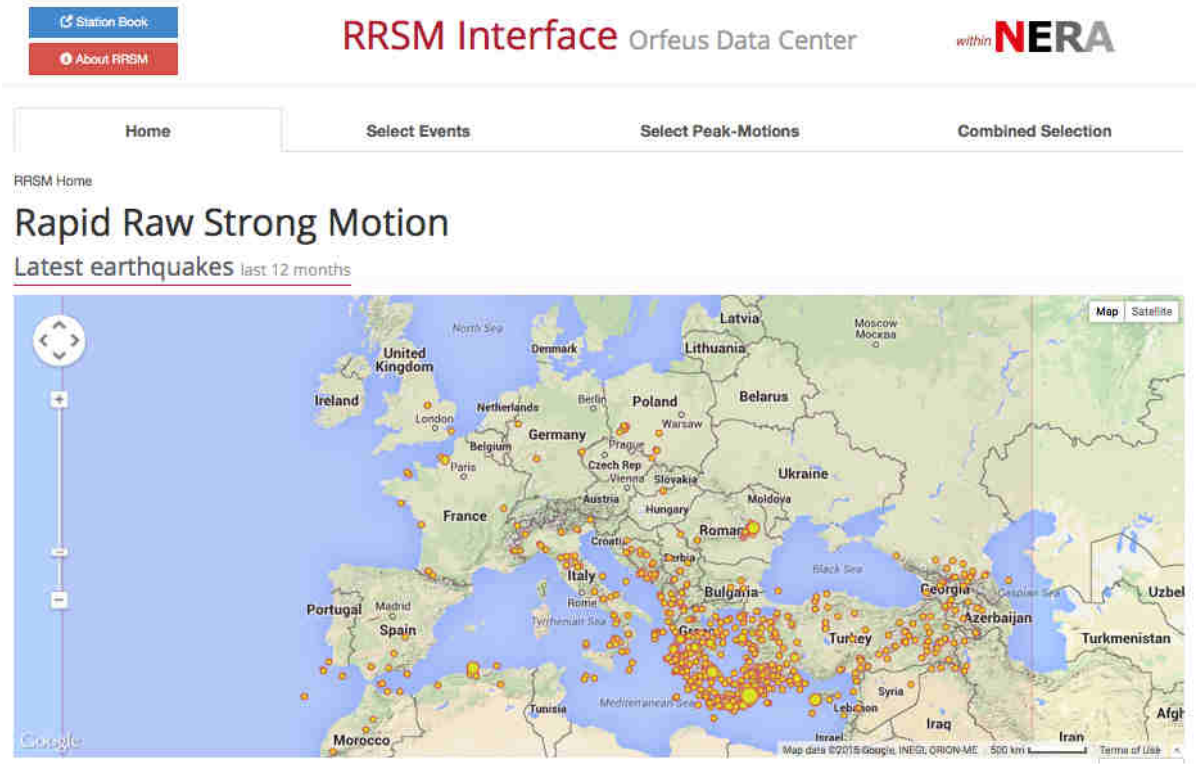

Figure 2. Screenshot of the RRSM web interface at ORFEUS. Shown on the map are the earthquakes with $M > 3.5$ occurred and processed in the last 12 months (last accessed 02.07.2015). Symbols are proportional to earthquake magnitude. Note the link to the Station Book in the upper left corner of the picture.

(Cauzzi et al., 2013), which is integrated in the earthquake monitoring software SeisComP3 (SC3; Hanka et al., 2010), that is free and openly available to interested users. *scwfparam* and the extension of the SC3 datamodel created to accommodate peak-motion data and information of engineering interest were designed by the Swiss Seismological Service (SED) at ETH Zurich within the framework of the EC-funded project NERA (Network of European Research Infrastructures for Earthquake Risk Assessment and Mitigation). The RRSM datamodel is designed to allow the inclusion of rupture characteristics and finite fault information if these are available (see https://quake.ethz.ch/quakeml/QuakeML2.0/StrongMotion). *scwfparam* computes peak ground acceleration (*PGA*), peak ground velocity (*PGV*), relative displacement elastic response spectra (*DRS*) and pseudo absolute acceleration elastic response spectra (*PSA*) in real-time or offline. It includes a process scheduler and handles reprocessing of data in a smart way. It supports ShakeMap XML and populates the strong-motion component of the SC3 database. The RRSM waveform processing is triggered by an EMSC earthquake alert, i.e., an automatic formatted message with basic information about the location and magnitude of the event. Earthquake data are processed if the earthquake magnitude *M* (any scale) exceeds 3.5 and the EMSC location is within the greater European region ($27° \leq$ Latitude $\leq 81°$, $-32° \leq$ Longitude $\leq 51°$). Waveforms are then requested from EIDA, only minutes after real-time, from all seismic stations installed within a given distance from the epicenter. This data request to EIDA is based on the EIDA core module *ArcLink* mentioned in the previous section.

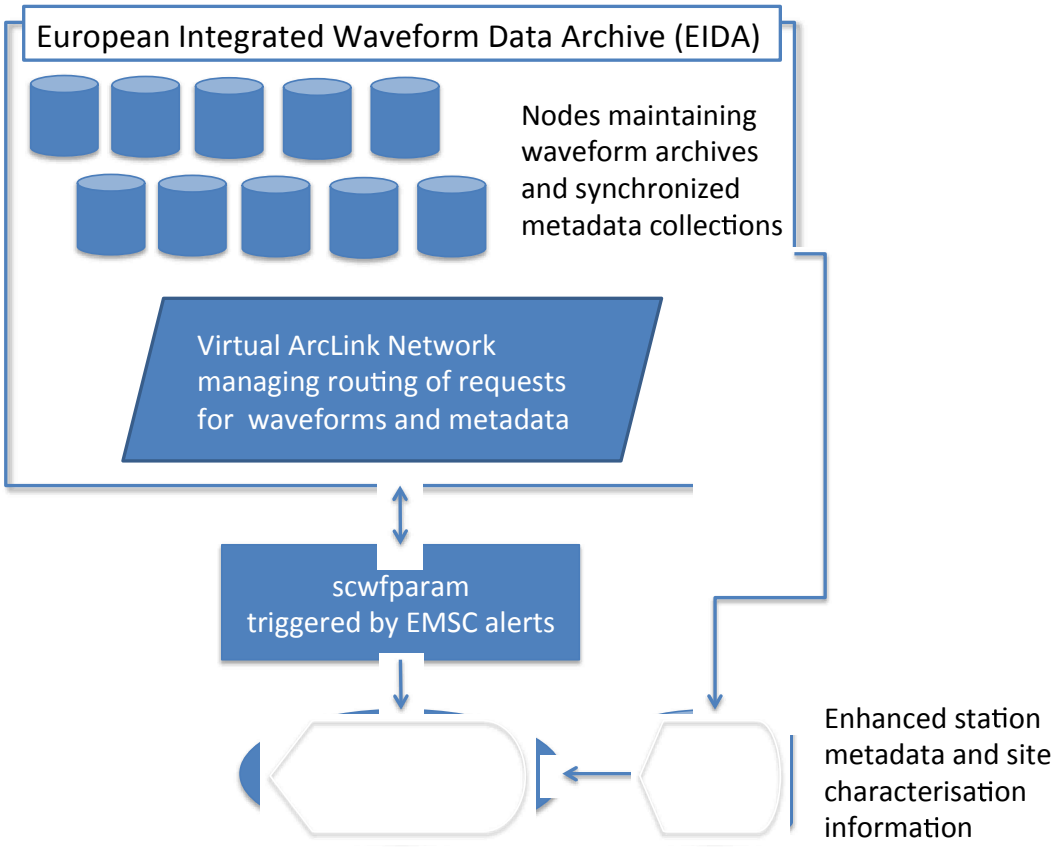

Figure 3. Simplified schematic of the data archival and dissemination infrastructure described in this paper.

The RRSM portal is openly available at http://www.orfeus-eu.org/rrsm. This interactive web interface (Figure 2) is based on OpenCMS, and users can download raw waveform data (in digital counts), peak ground-motion values and response spectra computed from *scwfparam*. The RRSM interface supports 3 different request types: a) Select Events allows querying the RRSM earthquake catalogue by event time, magnitude and location; b) Select Peak-Motions allows searching recorded waveforms based on peak-motion criteria; c) Combined Selection performs earthquake search by peak-motion criteria, station location, distance from and magnitude of the causative earthquake. In all cases the output of the query is events, streams and peak-motions fulfilling the search criteria. The interface provides plots of peak amplitude versus distance for a selected earthquake, and response spectra for selected stations. In addition, the user can select a set of events / stations / streams to download raw waveforms or metadata (SEED standard) from the EIDA web interface. Peak ground motion parameters and response spectra, earthquake and station metadata are available for download in various human-readable formats.

The RRSM is significantly different from previous or current strong-motion earthquake databanks in Europe, which usually provide processed data following manual review, often with a large delay. The RRSM concept is targeted to data users who require rapid access to strong-motion parameters, raw waveform data and metadata and do not wish to wait for delayed

manually processed waveforms. The RRSM targets users interested in having an overview of the openly available data available immediately following an earthquake, and providing easy and rapid access to the event waveform data. Key users are (a) seismologists and strong-motion data analysts; (b) earthquake and geotechnical engineers; (c) international earthquake response agencies; (d) the educated general public.

The RRSM is complimented by the Engineering Strong Motion (ESM) database, also developed during the NERA project, which provides manually reviewed processing and access to strong motion data from larger European earthquakes. The dataset from this services includes EIDA data. The RRSM and ESM are expected to be combined in order to provide a single service for strong motion data access during the EPOS project.

The basic station information available across EIDA is automatically included in the Station Book. The Station Book inventory can be manually modified by network operators, who can add new stations or update information. The aim of the Station Book is to set up a unique and reliable international registry of broadband and strong motion stations. Beyond the very basic information available at the ISC station book (which is dominated by broadband stations), such a registry does not exist so far and has an important role to serve the accelerometric data community, seismologists and engineers. The information can range from basic details such as location and names to acquisition chain description, morphology of the installation site or velocity profiles. Amongst the information already envisaged for site characterisation are: pictures, description of geological unit(s), site morphology, Eurocode 8 ground type (CEN, 2004), $V_{S,30}$, fundamental frequency at the site f_0 and related amplification $Amp(f_0)$, basin flags, groundwater depth, bedrock depth, borehole depth, borehole stratigraphy, housing details, building details. The Station Book interface was designed and implemented by ORFEUS Data Center (ODC) and ETH Zurich. The schema of the SeisComP3's database extension for Station Book was designed by ETH Zurich. The UML datamodel of this extension can be downloaded here and here. Available elements and attributes resulted from a careful scrutiny of the most common options adopted in strong-motion databanks worldwide.

Conclusions

We presented here the key features of the European Integrated waveform Data Archive (EIDA), the European Rapid Raw Strong-Motion Database (RRSM) and Station Book. We believe that the existing infrastructure (Figure 3) could with minor adaptations provide a consolidated and sustainable framework for archival and dissemination of structural and geotechnical arrays (seismically instrumented sites and structures) in Europe and worldwide.

Based on the SEED international standard for data format - nowadays also becoming familiar to the engineering community - EIDA can provide a mature and solid infrastructure for archival and dissemination of seismic and non-seismic data from structural and geotechnical arrays. Particularly attractive in this sense is the flexibility offered by the use of different band, stream, and location codes to rationally organise and archive data from sites equipped with multiple different sensors and sensor types, including dense structural and geotechnical arrays. Additionally, EIDA has developed services for distribution of both continuous and event-based waveforms, thus offering a unified access to scientists and practitioners interested in site /

structural response studies under low-strain or high-strain seismic loading. The RRSM, tailored to rapid computation and dissemination of parameters of engineering interest, can provide the interested users – both scientists and practitioners – a nearly immediate overview of the severity of shaking at all available seismic stations, and allow subsequent waveform data download from EIDA through either a web interface or webservices. The key feature of the RRSM is that it is automatically populated in near real-time; its growth is ensured any time an earthquake with magnitude larger than 3.5 occurs in the greater European region. The Station Book can be the infrastructure for collecting basic and advanced station metadata critical for the scientific and technical interpretation of the recorded waveforms.

Acknowledgments

The RRSM interface and the Station Book were designed and implemented by ORFEUS Data Center (ODC) and ETH Zurich, using funding from EC-funded FP7 project NERA (Network of European Research Infrastructures for Earthquake Risk Assessment and Mitigation). We are thankful to the NA3 working group (coordinated by Sinan Akkar) for supporting the development of the RRSM. We profited while developing these tools from the gentle and wise coordination and advice of Torild Van Eck, who passed away on 30 November 2014.

References

Ambraseys NN, Smit P, Douglas J, Margaris B, Sigbjornsson R, Olafsson S, Suhadolc P and Costa G. Internet site for European strong-motion data. *Bollettino di Geofisica Teorica ed Applicata* 2004; **45**: 113–129.

Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BSJ, Wooddell KE, Graves RW, Kottke AR, Boore DM, Kishida T and Donahue JL. NGA-West 2 Database. *Earthq Spectra* 2014; 140514111412006. doi: 10.1193/070913EQS197M

Cauzzi C, Sleeman R, Clinton J, Domingo Ballesta J, Kästli P, Galanis O. Kicking-off the European Rapid Raw Strong-Motion Database. *Seism Research Letters*, submitted.

Cauzzi, C, Clinton J, Becker J, Kästli P. Scwfparam: a tool for rapid parameterisation of ground motions and input to ShakeMap in SeisComP3. *Seismol Res Lett* 2013; **84**(2): 356. doi:10.1785/0220130011.

CEN, European Committee for Standardization. *Eurocode 8: design of structures for earthquake resistance—part 1: general rules, seismic actions and rules for buildings.* Bruxelles, 2004.

Clinton J, Hanka W, Mazza S, Pederson H, Sleeman R, Stammler K, Strollo A, Van Eck T. EIDA: The European distributed data archive for seismology. *Proceedings of 2ECEES* Istanbul 2014; paper no. 3322.

Hanka W, Saul J, Weber B, Becker J, Harjadi P, Fauzi and GITEWS Seismology Group. Real-time earthquake monitoring for tsunami warning in the Indian Ocean and beyond. *Nat Hazards Earth Syst Sci* 2010; **10**: 2611–2622. doi: 10.5194/nhess-10-2611-2010

Sleeman R, Cauzzi C, Domingo Ballesta J, Clinton J, Kästli P, Galanis O, Ballarin Dolfin E. The Rapid Raw Strong-Motion (RRSM) Database. *Proceedings of 2ECEES* Istanbul 2014; paper no. 2791.