Bollettino di Geofisica Teorica ed Applicata Vol. 61, n. 3, pp. 343-356; September 2020
DOI 10.4430/bgta0293

The pan-European Engineering Strong Motion (ESM) flatfile: comparison with NGA-West2 database

G. Lanzano, S. Sgobba, L. Luzi, F. Pacor, R. Puglia, C. Felicetta and M. D'Amico

Istituto Nazionale di Geofisica e Vulcanologia, Milan, Italy

(Received: 19 March 2019; accepted: 3 July 2019)

ABSTRACT In the first months of 2018, a parametric table (flatfile) related to the Engineering Strong Motion (ESM) database was released and disseminated through a website (http://esm.mi.ingv.it/flatfile-2018). The flatfile contains intensity measures of engineering interest and associated metadata of three-components manually processed waveforms. The uniform collection of strong motion data and the compiling of quality-checked metadata allow the users among practitioners and seismologists, to test and calibrate Ground Motion Models (GMMs) for hazard assessment purposes or for the analysis of the seismic structural response. In 2013, a database used for similar purposes was constructed in the framework of the NGA-West2 project, whose main objective is to update and improve the Next Generation Attenuation (NGA) models for active tectonic regions, such as California. In this framework, a flatfile containing several parameters, such as peak parameters and ordinates of the pseudo-acceleration elastic response spectra, along with metadata of events and stations was released. The scope of this paper is to highlight the main differences between the two tables in terms of structure, data statistics and qualification of metadata.

Key words: strong motion records, flatfile, metadata, engineering strong motion database, NGA-West2.

1. Introduction

The Engineering Strong-Motion (ESM) flatfile is a parametric table which contains verified and reliable metadata as well as intensity measures of manually processed waveforms included in the ESM database. The latest release of the flatfile was published at the beginning of 2018 and is freely available, after registration, through a website (http://esm.mi.ingv.it/flatfile-2018). This tool, developed in collaboration with GeoForschung Zentrum (GFZ), is one of the research products disseminated in the framework of the Seismology Thematic Core Service of the European project EPOS-IP (European Plate Observing System - Implementation Phase).

ESM flatfile supersedes the previous pan-European data sets for ground motion models (GMMs) calibration, such as ISESD [Internet-Site for European Strong motion Data: Ambraseys *et al*. (2004)] and RESORCE [Reference Database for Seismic Ground Motion in Europe: Akkar *et al*. (2014)], both for the number of records and for the type and quality of the metadata. Several details on the structure and organisation of the flatfile are discussed in Lanzano *et al*. (2019b). A consistency check of the data included in the flatfile was carried out by Bindi *et al*. (2019), performing a residual analysis using an *ad hoc* GMM. The flatfile was also recently used to update the GMMs for shallow active crustal regions in Italy by Lanzano *et al*. (2019a) and to implement the ground motion logic tree within the next generation European-Mediterranean Seismic Hazard Model (ESHM18) in the framework of EU SERA project (Weatherill *et al*., 2018).

The data set for flatfile compilation is extracted from ESM database [https://esm.mi.ingv.it; Luzi *et al*. (2016)], which is a data centre and provider of the European Integrated Data Archive (EIDA) waveforms for events with magnitude larger than 4.0. Data in ESM are also gathered from offline archives made available from other European and Mediterranean authority providers [see Luzi *et al*. (2016) for further details]. ESM was born in 2009 in the framework of the European Project NERA (Network of European Research Infrastructures for Earthquake Risk Assessment and Mitigation), along with the fruitful experience of ITACA [ITalian ACcelerometric Archive: http://itaca.mi.ingv.it; Pacor *et al*. (2011)], which contributes significantly to the number of ESM records.

Since 2003, in the western United States, the Pacific Earthquake Engineering Research (PEER) Center initiated a large research program to develop next generation attenuation relationships, called NGA (Bozorgnia *et al*., 2014). In 2013, the new worldwide database for the calibration of NGA models for shallow crustal earthquakes in active tectonic regions (named NGA-West2) was prepared and disseminated through several parametric tables (Ancheta *et al*., 2014). The database was used for NGA-West1 models updating (Abrahamson *et al*., 2008), which were introduced in the GMM branches of probabilistic hazard logic tree of western U.S. More recently a similar research initiative named NESS1 (http://ness.mi.ingv.it) was launched in Italy, with the aim to characterise the ground motion features in near source conditions (Pacor *et al*., 2018). Since part of the authors of NESS1 are also part of the ESM working group, the organisation of NESS1 table is very similar to that of the ESM flatfile and, in the majority of the cases, the metadata are the same.

In the following, a comparison between the ESM and NGA-West2 tables is carried out with a specific focus on the flatfile organisation and dissemination of events and station metadata. The statistical distribution of the most relevant explanatory variables for GMMs calibration are reported and discussed. A more direct comparison is also provided for a selected event in common to the two data sets, i.e. the main-shock (M_W 6.1) of the 2009 L'Aquila (Italy) seismic sequence, recorded by more than 50 strong-motion stations.

2. Processing scheme and intensity measures

2.1. Records processing

Both ESM and PEER NGA-West2 data sets contain waveforms that are uniformly processed: in detail, the selection of the corner frequencies of the Butterworth filter is guided by the examination of the Fourier amplitude spectrum on a component-by-component basis at both low and high frequencies (Boore and Bommer, 2005), but following different procedures for the two data sets. In ESM, the waveforms are processed according to the procedure described in Paolucci *et al*. (2011). The latter applies a second-order acausal time-domain Butterworth filter (Douglas and Boore, 2011) to the zero-padded acceleration time series and zero-pad removal to make acceleration and displacement consistent after double integration. The typical band-pass frequency range is between 0.1 and 40.0 Hz for digital records, whereas it is narrower for analogue ones (on average 0.3-25.0 Hz). The ESM processing is accessible through a web-interface available on http://esm.mi.ingv.it/processing that allows the users to perform their own manual processing and to download the processed data (Puglia *et al*., 2018).

The records of NGA-West2 data set are processed according to a procedure developed by the PEER team (Douglas and Boore, 2011; Ancheta *et al*., 2013). According to Boore *et al*. (2012), the ESM (or ITACA) and PEER processing schemes are quite similar and the results are coincident in several cases. The main difference is that ESM processing applies linear de-trending of the velocity and displacements obtained from the pad-stripped data, whereas PEER fits a sixth-order polynomial to the displacements time-series.

Fig. 1 shows the number of recordings within the usable frequency band for (pseudo-) acceleration response spectra.

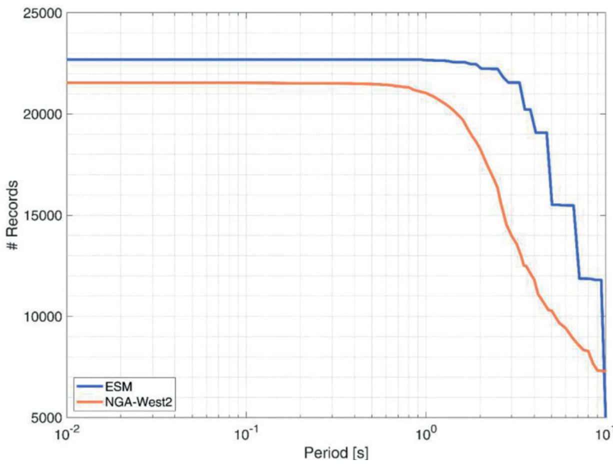

Fig.1 - Number of records with highest useable period versus period.

The spectral acceleration ordinates in both the data sets are considered valid until 100 Hz (T = 0.01 s), according to Douglas and Boore (2011), while all the processed records can be used up to about 0.6-0.7 s. At longer periods, the number of useable periods decreases, since several analogue records are processed using high pass corner frequencies larger than 0.3 Hz. At a period of 10.0 s the usable records in ESM and NGA-West2 reduces to about 20% and 30% of the entire data set, respectively.

2.2. Structure and dissemination of the flatfile

Fig. 2 schematically illustrates how the ESM and NGA-West2 parametric tables are organised and disseminated.

The ESM flatfile is arranged into three '.csv' files for 36 ordinates of the 5% damping acceleration and displacement elastic response spectra and 103 amplitudes of the Fourier spectrum. The table fields are related to event and station metadata, including several distance metrics, peak and integral intensity measures as well as some duration parameters. The waveforms are not disseminated but they are all accessible through the ESM website and detectable using the event id, the network and the station code, provided in the flatfile.

NGA-West2 provides 11 tables ('.xls' files) including the station and event metadata, the distances, the peak parameters and the ordinates of pseudo-spectral acceleration ranging from 0.01 to 20.0 s at 11 different damping levels in the range 0.5-30.0%. The raw waveforms of NGA-West2 database are made available by several worldwide agencies whereas the processed

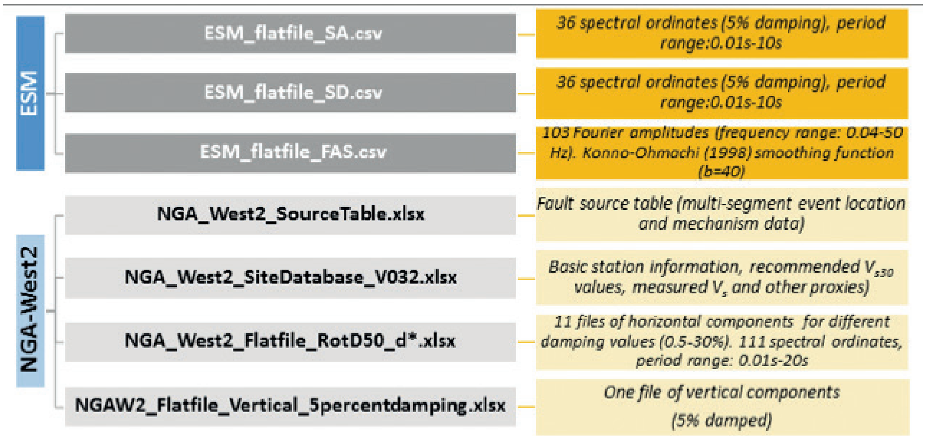

Fig. 2 - Schematic overview of the ESM and NGA-West2 flatfile format.

data are distributed for registered users through the PEER strong motion database portal (https://ngawest2.berkeley.edu).

The parameters for the horizontal components of strong motion are computed according to rotation independent measures, RotD50 (Boore, 2010) in both data sets. In NGA-West2, RotD100, i.e. the maximum value of the distribution of the intensity measure across all azimuths, is estimated through the empirical relation by Shahi and Baker (2014). ESM also contains the values of the parameters of each of the three components. The NGA-West2 strong motion parameters of the vertical component, the finite fault info table and the site database are instead distributed through distinct tables, as supplemental material (NGA_West2_supporting_data_for_flatfile).

Table 1 - Intensity measures included in ESM and NGA-West2 data sets.

Intensity measures	ESM	NGA-West2
Peak Ground Acceleration (*PGA*)	X	X
Peak Ground Velocity (*PGV*)	X	X
Peak Ground Displacement (*PGD*)	X	X
Arias Intensity, T90 duration (Trifunac and Brady, 1975)	X	X^
Housner Intensity	X	
Cumulative Absolute Velocity (*CAV*)	X	
(Pseudo-) Acceleration response spectra (*SA* or *PSA*)	X (36 periods with 5% damping ratio)	X (111 periods; 11 damping ratios from 0.5 to 30.0%)
Horizontal components combination RotDnn (Boore, 2010)	RotD00, RotD50, RotD100	RotD50, RotD100*
Displacement response spectra (*SD*)	X (36 periods with 5% damping ratio)	X (111 periods; 11 damping ratios from 0.5 to 30.0%)*
Fourier amplitude spectra	X (103 periods)	X^

^ available upon request to Y. Bozorgnia (yousef.bozorgnia@ucla.edu)
* not directly provided but easily computable from other parameters.

Table 1 reports a brief scheme of the intensity measures included in the parametric tables. Non-standard intensity measures, such as Arias and Housner intensities, significant duration, and the Cumulative Absolute Velocity (*CAV*) are included in ESM flatfile as well. For the NGA-West2 data, Arias intensity and significant duration are available upon request.

3. Event, station and path metadata

3.1. Event location and magnitude

One of the basic ideas of the ESM flatfile is to provide a reference field for each parameter included in the table. As a matter of fact, the event location and magnitude are assigned according to a specific hierarchy that firstly prefers earthquake-specific studies, secondarily the Moment Tensor solution (Ekström *et al*., 2012) and, finally, the regional or international bulletins, such as the International Seismological Centre (ISC). In addition, in the ESM flatfile the revised moment magnitude from EMEC [Euro-Mediterranean Earthquake Catalogue: Grünthal and Wahlström (2012)] is provided, when available. The focal mechanism is derived by the rake of the moment tensor solution (Aki and Richards, 2002), with the modification of Boore *et al*. (1997) for strike-slip events.

The main idea beneath the event metadata attribution in NGA-West2 is related to the availability of finite fault inversion solutions, even for low magnitudes events. In particular, the events in NGA-West2 database are divided into moderate-to-large worldwide earthquakes and small-to-moderate magnitude California earthquakes (events with moment magnitude less than about 5.0). A further distinction is also made on Class1 (mainshock) and Class 2 (aftershock) events to allow a more consistent data selection in GMM calibration.

The event parameters (i.e. epicentre location, focal depth, local magnitude, moment magnitude, and style of faulting) of small-to-moderate earthquakes for the California data are obtained from regional catalogues (i.e. earthquake catalogues of Northern and Southern California Seismic Network), while for moderate-to-large from literature studies or international catalogues.

3.2. Fault geometry

In the ESM flatfile, the fault geometry is provided only for events with $M > 5.5$. For the strongest events, the fault models from published studies are generally adopted as a primary reference. When no specific study is available, the regional and international databases are consulted, such as the Database of Individual Seismogenic Sources [DISS: http://diss.rm.ingv.it/diss; DISS Working Group (2018)] for Italy, the Greek Database of Seismogenic Sources [GreDaSS: http://gredass.unife.it; Caputo and Pavlides (2013)] and the Finite-Source Rupture Model Database [SRCMOD: Mai and Thingbaijam (2014)] for worldwide earthquakes.

For moderate-to-large events in NGA-West2 database, the finite fault geometry is typically obtained, in order of preference, from i) field observations of primary surface rupture, ii) co-seismic slip distribution from inversions of waveform and geodetic data, and iii) aftershock distributions. When available, slip inversion models are also used to extract information about rise time, rupture velocity, and other data related to the spatial characteristics of (co-seismic) fault slip, such as the existence of a shallow asperity producing significant (>20% of the total) moment release in the top 5 km of crust.

In both data sets, when there is more than one rupture model for an earthquake, the preferred model was carefully selected and referenced. On the contrary, when a published fault model is not available, a simulated finite fault is developed both in ESM and NGA-West2.

The ESM strategy for the rupture simulation or the calculation of some missing parameters is reported by Pacor *et al*. (2018) and it consists in a modification of the procedure by Kaklamanos *et al*. (2011), originally developed to convert the different distance metrics implemented in the GMMs. The procedure is applied only to events with magnitude larger than 5.5 lacking of a fault geometry.

The NGA-West2 methodology for simulating the fault plane is different and consists of randomly sampling from probabilistic distributions of fault rupture area, aspect ratio of ruptured area, and hypocentre position on the fault plane (Chiou *et al*., 2008). Differently to ESM, the simulated fault is obtained for all the events that miss the rupture geometry, thus including small earthquakes.

3.3. Distance metrics

A comparison between the distance metrics and path-related metadata provided by the two data sets is shown in Table 2. A detailed definition of the distance metrics can be found in Ancheta *et al*. (2014).

Table 2 - Distance metrics and path-related metadata included in ESM and NGA-West2 flatfile.

Parameter	ESM	NGA-West2
Epicentral distance: distance from epicentre R_{EPI}	X	X
Event-to-station azimuth	X	X
Hypocentral distance: distance from hypocentre R_{HYP}	X*	X
Joyner-Boore distance: distance computed from the surface projection of the fault R_{JB}	X	X
Rupture distance: shorter distance to the rupture plane R_{RUP}	X	X
Horizontal distance measured perpendicular to the fault strike, from the top edge of rupture plane R_X	X	X
Horizontal distance off the surface projection of rupture plane, measured parallel to the fault strike R_Y	X	X
Hanging-wall indicator	X*	X
Radiation pattern coefficients		X
Directivity parameters		X

* not directly provided but easily computable from other parameters.

In ESM, the distance from the fault rupture plane, R_{RUP}, the closest distance from the surface projection of the fault, R_{JB}, and the hanging/footwall distances, R_X and R_Y, are calculated only when the finite fault is available. For smaller magnitude events ($M_W < 5.5$), we consider point-like sources because the differences between the epicentral distance and R_{JB}, as well as between the hypocentral distance and R_{RUP}, can be neglected. Hypocentral distance and hanging wall indicator are not explicitly provided in the ESM flatfile, but they can be derived by other parameters, e.g. the sign of R_X is an indicator of hanging/footwall conditions.

In NGA-West2, R_{RUP}, R_{JB}, R_X and R_Y are calculated form the finite fault geometry for all the events in the database (Ancheta *et al.*, 2013). NGA-West2 also provides radiation pattern coefficients and some directivity parameters, useful for GMMs calibration.

3.4. Site characterisation parameters

In the ESM flatfile, the site response is characterised only by the time-averaged shear wave velocity in the uppermost 30 metres ($V_{S,30}$), following the Eurocode 8 (EC8, CEN 2004) classification scheme. Two different estimates of $V_{S,30}$ are included in the flatfile: i) *in-situ* geophysical measurements (preferred); ii) empirical correlation with the topographic slope according to the correlation provided by Wald and Allen (2007). The EC8 classification field is, instead, filled primarily with site categories derived by $V_{S,30}$ measurements, otherwise to those inferred by surface geology (Di Capua *et al.*, 2011), which are marked by an asterisk (e.g. A*).

In addition to measured $V_{S,30}$, the site database of NGA-West2 contains several estimates related to alternative proxies (Seyhan *et al.*, 2014), i.e. i) topographic slope (Wald and Allen, 2007), ii) geology or geology-slope hybrids, iii) terrain/geomorphology (Yong *et al.*, 2012), and iv) geotechnical tests. The table with waveform intensity measures solely includes the preferred $V_{S,30}$ and its associated uncertainty. The hierarchy to assign the preferred $V_{S,30}$ is reported in Seyan *et al.* (2014), but the first choice is always the measured value, if available. The basin depths of stations located in alluvial basin are provided on the basis on 3D velocity models and/or shear wave velocity profiles.

Fig. 3 shows the distribution of strong motion recording stations as a function of preferred $V_{S,30}$ in ESM and NGA-West2. The total number of stations in ESM (2,080) is about the half of the stations included in the NGA-West2 database (4,149). The amount of station with $V_{S,30}$ based on *in situ* measurements (with profile depths greater than 30 m) is 24% (474 out of 2,080) and 13% (552 out of 4,149) for ESM and NGA-West2, respectively.

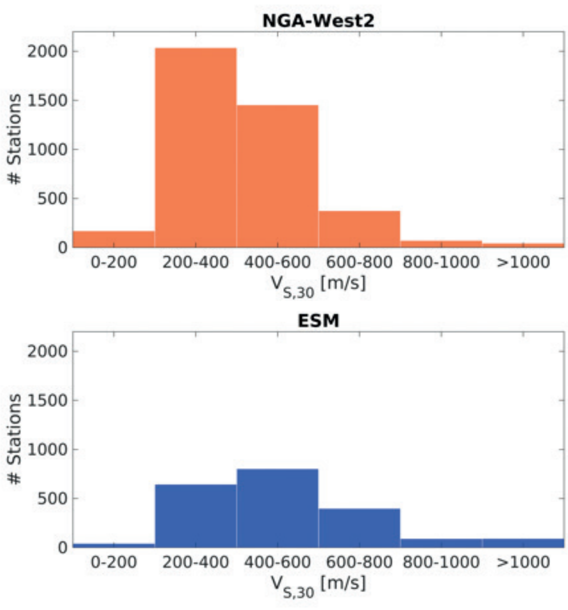

Fig. 3 - Distribution of strong motion recording stations as a function of preferred $V_{S,30}$ in NGA-West2 (on the top) and ESM (on the bottom).

4. Data statistics

The data set for the ESM flatfile includes 23,014 three-components recordings from 2,179 earthquakes and 2,080 stations from Europe and Middle-East in the time span 1969-2016. The events are characterised by magnitudes in the range 3.5-8.0 and refer to different tectonic regimes, such as shallow active crustal and subduction zones. The data set for NGA-West2 is relative to shallow active crustal events only and comprises 21,336 records of 600 worldwide events, recorded by 4,149 stations in the time interval 1998-2011.

The magnitude-distance distribution of the two data set is reported in Fig. 4, where in ESM the local magnitude M_L is used when M_W is not available. The minimum magnitude of ESM is 3.5, while the data collection of NGA-West2 includes events down to 3.0, mainly composed by weak events occurred in California. However, the weak events (lower than $M_W < 4.5$) dominate both the data set, especially for the ESM data set, that corresponds to 70% of the data. The magnitude-distance distribution of the two data sets is similar and the amount of near source data ($R_{epi} < 10$ km) is not negligible in both data sets, corresponding to 4% in ESM and 3% in NGA-West2, respectively (Fig. 4).

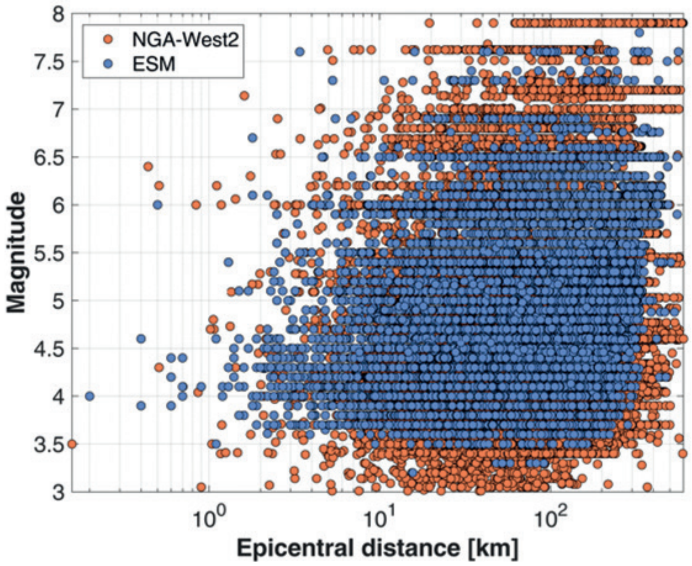

Fig. 4 - Magnitude-distance distribution of ESM flatfile and NGA-West2 data set.

Fig. 5 shows the amount of records in ESM and NGA-West2 data sets by earthquake-contributing countries. In ESM, most of the data (59%) are related to events occurred in Italy, but other relevant contributions come from earthquakes in Greece (16%) and Turkey (7%). In NGA-West2, the events of western U.S. and Japan contributes with the 38% and 30%, respectively. The contribution of the Mediterranean earthquakes, which are in common with ESM data, is only the 6%.

Fig. 6 shows the stacked histograms of moment magnitude, epicentral distance, focal depth and style of faulting of the two data sets. Magnitude histogram (Fig. 6a) is different for the two data sets, being characterised by a nearly uniform distribution over the whole magnitude range in NGA-West2 compared to ESM, which is dominated by small-to-moderate events in the

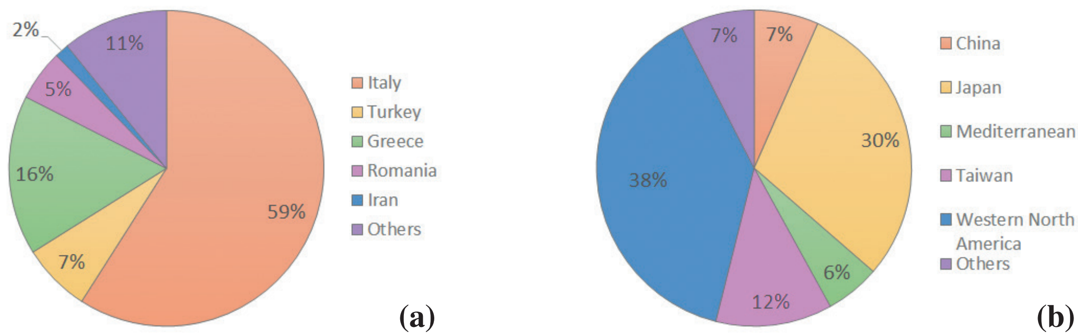

Fig. 5 - Pie charts of record numbers per country or area in ESM (a) and NGA-West2 (b).

magnitude range 4.0-5.0. A significant number of events with $M > 6.0$ is contained in NGA-West2 (14%) with respect to ESM (3%). Epicentral distance distribution, shown in Fig. 6b, indicates a good sampling of records in both the data sets near and far from the source without significant differences. The focal depth distribution in Fig. 6c shows a noticeable preponderance of shallow depth events both in ESM data set (70% at depth < 20 km); and in NGA-West2 (96% at depth < 20 km).

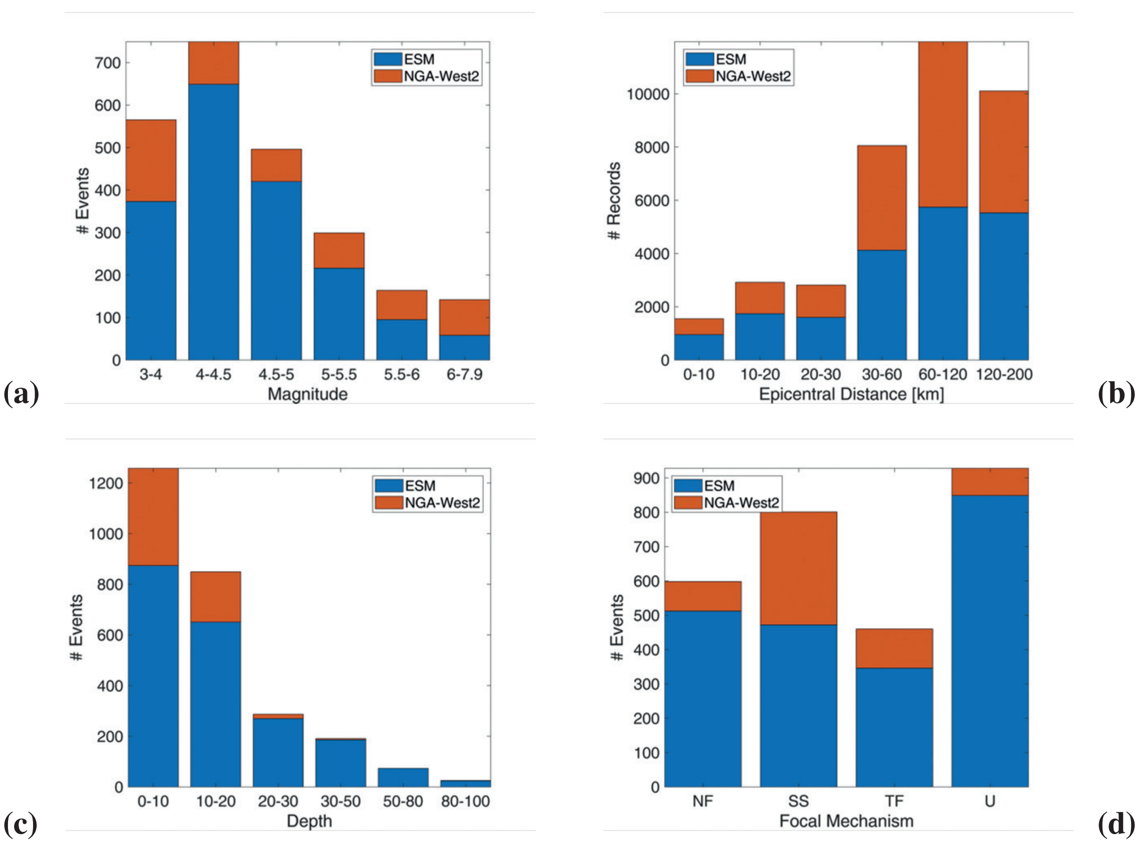

Fig. 6 - Stacked histograms of: a) moment magnitude, b) epicentral distance, c) focal depth and d) style of faulting in ESM flatfile and NGA-West2 database.

In Fig. 6d the comparison in terms of focal mechanisms shows a relevant amount of normal fault earthquakes in ESM data set (24%) due to the prevalence of Italian data. Conversely, the NGA-West2 has 54% of strike-slip style of faulting, thus reflecting the tectonic differences between Euro-Mediterranean earthquakes and the other parts of the world.

5. L'Aquila 2009 earthquake case study

The L'Aquila earthquake occurred on 9 April 2009 (01:32 UTC) in central Italy with a moment magnitude M_W equal to 6.1. It represents one of the first strong earthquake in Europe recorded by a consistent amount of stations, some of them in near source conditions. As a matter of fact, this event improved the knowledge of seismic motion in Italy and allowed the calibration of the GMM by Bindi *et al*. (2011), which is the reference model for regional hazard assessment. The L'Aquila earthquake was included in both data sets and the main characteristics are reported in Table 3.

The number of records in ESM is higher than NGA-West2, since in ESM some additional data recorded by temporary networks, collected in ITACA, were made available. Although different studies are considered as authoritative source, the parameters are quite similar. The preferred moment magnitude (M_w 63) of ESM (EMEC) and that provided by NGA-West2 are identical even if other authoritarian sources provide lower values [M_w 6.1: ISIDe Working Group (2016)].

The site parameters, the source-to-site distances, and the peak amplitudes (RotD50) of four L'Aquila earthquake waveforms, recorded by the stations of i) Avezzano, ii) Assisi, iii) L'Aquila-Valle dell'Aterno, and iv) Leonessa, are compared in Table 4.

Since ITACA database is the authoritative source for geophysical measurements in Italy both for ESM and NGA-West2, the preferred $V_{S,30}$ based on measured velocity profile is almost identical for Avezzano (AVZ) and L'Aquila (AQA) as derive from *in situ* tests performed before

Table 3 - Event metadata of 2009 L'Aquila earthquake in ESM and NGA-West2.

Parameters	ESM	NGA-West2
Event Name	L'Aquila	L'Aquila, Italy
Event time	2009-04-06 01:32:40	
Event ID	IT-2009-0009	0274
Number of records	64	48
Epicentre coordinates (degrees) (Reference)	42.342 (lat) - 13.380 (lon) (ISIDe Working Group, 2016)	42.342 (lat) - 13.380 (lon)
Hypocentre depth (km) (Reference)	8.3 (ISIDe Working Group, 2016)	9.27
Local magnitude M_L	5.9	-
Moment magnitude M_w (Reference)	6.1 (ISIDe Working Group, 2016) 6.3 (EMEC; Grünthal and Wahlström, 2012)	6.3
Strike, dip, rake (degrees) (Reference)	140, 50, -90 (Ameri *et al.*, 2009)	139, 48, -98 (Scognamiglio *et al.*, 2010)
Depth of the fault top (km)	0.5	0.8
Fault length (km), fault width (km) (Reference)	20, 15 (Ameri *et al.*, 2009)	20, 16 (Scognamiglio *et al.*, 2010)

Table 4 - Parameters of four recordings of 2009 L'Aquila earthquake.

Station code (ESM)	Station name (NGA-W2)	Parameters	ESM	NGA-West2
AVZ	Avezzano	Preferred $V_{s,30}$ [m/s]	199.000	199.000
		R_{JB} [m]	22.410	23.670
		R_{RUP} [m]	25.440	26.860
		PGV [cm/s]	10.920	10.910
		PGD [cm]	3.520	3.630
ASS	Assisi	Inferred from slope proxy $V_{s,30}$ [m/s]	1053.000	462.000
		R_{JB} [m]	97.510	96.120
		R_{RUP} [m]	97.570	96.140
		PGV [cm/s]	0.420	0.420
		PGD [cm]	0.160	0.190
AQA	L'Aquila	Preferred $V_{s,30}$ [m/s]	549.000	552.000
		R_{JB} [m]	0.000	0.000
		R_{RUP} [m]	5.740	6.550
		PGV [cm/s]	29.800	29.950
		PGD [cm]	4.410	4.440
LSS	Leonessa	Preferred $V_{s,30}$ [m/s]	1091.000	595.000
		R_{JB} [m]	32.250	33.980
		R_{RUP} [m]	37.070	35.950
		PGV [cm/s]	0.739	0.752
		PGD [cm]	0.228	0.253

2009 (DPC-INGV S4 project: http://esse4.mi.ingv.it). The most relevant difference of $V_{S,30}$ (from measured velocity profile) for Leonessa (LSS) station, is related to the update of the geophysical surveys (after 2009) included in ESM and not yet considered for the NGA-West2 database. The reported values of $V_{S,30}$ for Assisi (ASS) are inferred instead by topographic slope proxy both for ESM and NGA-West2. As the relationship adopted for the estimate of $V_{S,30}$ is the same in both the database (Wald and Allen, 2007), the observed gap (1053 vs. 462 m/s) could be related to different DEMs (Digital Elevation Models) used to calculate the slope. The differences in the rupture distances are small, in all the cases within about 2 km.

The acceleration spectral amplitudes (at 5% damping ratio) of the four records are also compared in Fig. 7 as a function of period. The peak parameters in Table 4 and the ordinates of acceleration response spectra in Fig. 7 are very similar, showing that the processing scheme and the calculation of RotD50 are consistent in the two data sets.

6. Conclusions

The main objective of this paper is to compare the recently generated (2018) ESM flatfile with that constructed in the 2013 for NGA-West2 project (Ancheta *et al.*, 2013). The latter aims

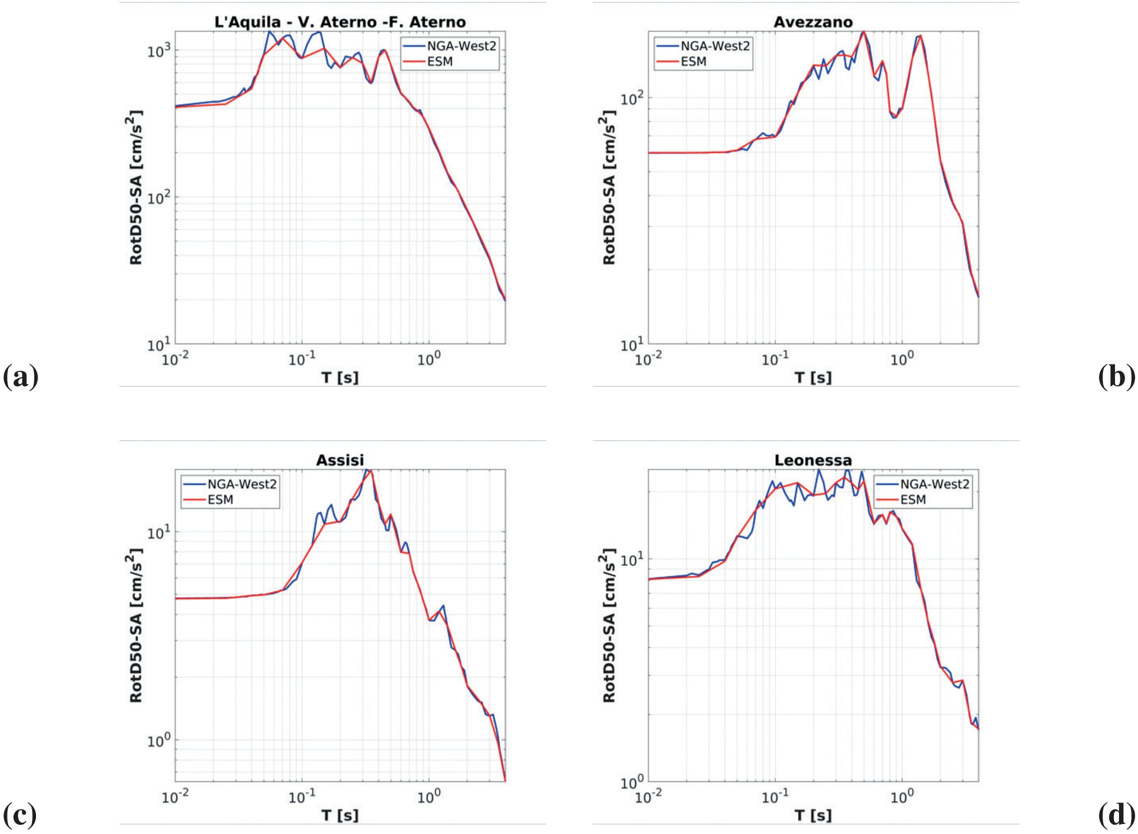

Fig. 7 - Comparison of RotD50 acceleration response spectra from: a) L'Aquila-Valle dell'Aterno (AQA), b) Avezzano (AVZ), c) Assisi (ASS), and d) Leonessa (LSS) recording stations.

at calibrating global ground motion models for shallow active crustal regions. ESM flatfile has similar scope for Europe, but also contains events derived by other different tectonic environments, such as volcanoes or subduction plates, whereas the NGA-West2 database only includes shallow crustal earthquakes from active tectonic regions.

The general idea of ESM flatfile is to include all the public data of ESM database as well as to allow the users to create their own table (a command-line application is provided on the ESM website to filter the table). The events of NGA-West2 database are instead tailored to the scope of calibrating GMMs and carefully selected among the well-sampled worldwide earthquakes with best quality of event and station metadata. In this way, the database is useful for the calibration of models with input parameters additional to the basic explanatory variables of standard GMMs (magnitude, distance, style of faulting, and site parameter). Indeed, the NGA-West2 table contains more fields than ESM to characterise the events and the stations, such as the stress drop of the event, directivity or radiation pattern parameters or depth of alluvial deposits for station located in basins.

However, the organisation of the two tables is quite similar and several fields are in common. Moreover, the processing scheme and RotD50 calculation, although different, produce similar results in terms of peak parameters and spectral amplitudes.

Our experience in the flatfile compilation led us to say that the most important effort is related to the definition of procedures to fill the table which must be, as much as possible, referenced and validated by the scientific community. Another important aspect is the traceability of the data and metadata, using a proper reference field for all the parameters provided in the table.

Acknowledgements. The European project EPOS (European Plate Observing System Implementation Phase) and its coordinator Massimo Cocco (Istituto Nazionale di Geofisica e Vulcanologia) are acknowledged by the authors for supporting and funding the preparation of the Engineering Strong Motion (ESM) flatfile. This study has been also funded by the H2020 project SERA (Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe, Grant Agreement Number 730900). The ESM flatfile is the result of the collaboration of INGV team (the authors) with German Research Centre for Geosciences (GFZ) team, coordinated by Fabrice Cotton and composed by Dino Bindi, Sreeram Reddy Kotha, and Graeme Weatherill.

REFERENCES

Abrahamson N., Atkinson G., Boore D., Bozorgnia Y., Campbell K., Chiou B., Idriss I.M., Silva W. and Youngs R.; 2008: *Comparisons of the NGA ground-motion relations*. Earthquake Spectra, **24**, 45-66.

Aki K. and Richards P.G.; 2002: *Quantitative seismology, 2nd ed*. University Science Books, Mill Valley, CA, USA, 704 pp.

Akkar S., Sandikkaya M.A., Şenyurt M., Sisi A.A., Ay B.Ö., Traversa P., Douglas J., Cotton F., Luzi L., Hernandez B. and Godey S.; 2014: *Reference database for seismic ground-motion in Europe (RESORCE)*. Bull. Earthquake Eng., **12**, 311-339.

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

Ameri G., Massa M., Bindi D., D'Alema E., Gorini A., Luzi L., Marzorati S., Pacor F., Paolucci R., Puglia R. and Smerzini C.; 2009: *The 6 april 2009, Mw 6.3, L'Aquila (central Italy) earthquake: strong-motion observations*. Seismol. Res. Lett., **80**, 951-966.

Ancheta T.D., Darragh R.B., Stewart J.P., Seyhan E., Silva W.J., Chiou B.S.J., Wooddell K.E., Graves R.W., Kottke A.R., Boore D.M., Kishida T. and Donahue J.L.; 2013: *PEER NGA-West2 database*. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA, USA, Report n. 2013/03, 134 pp.

Ancheta T.D., Darragh R.B., Stewart J.P., Seyhan E., Silva W.J., Chiou B.S.J., Wooddell K.E., Graves R.W., Kottke A.R., Boore D.M. and Kishida T.; 2014: *NGA-West2 database*. Earthquake Spectra, **30**, 989-1005.

Bindi D., Pacor F., Luzi L., Puglia R., Massa M., Ameri G. and Paolucci R.; 2011: *Ground motion prediction equations derived from the Italian strong-motion database*. Bull. Earthquake Eng., **9**, 1899-1920.

Bindi D., Kotha S-R., Weatherill G., Lanzano G., Luzi L. and Cotton F.; 2019: *The pan-European Engineering Strong Motion (ESM) flatfile: consistency check via residual analysis*. Bull. Earthquake Eng., **17**, 583-602.

Boore D.M.; 2010: *Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion*. Bull. Seismol. Soc. Am., **100**, 1830-1835.

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

Boore D.M., Joyner W.B. and Fumal T.E.; 1997: *Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: a summary of recent work (with 2005 erratum)*. Seismol. Res. Lett., **68**, 128-153.

Boore D.M., Azari Sisi A. and Akkar S.; 2012: *Using pad-stripped acausally filtered strong-motion data*. Bull. Seismol. Soc. Am., **102**, 751-760.

Bozorgnia Y., Abrahamson N.A., Al Atik L., Ancheta T.D., Atkinson G.M., Baker J.W., Baltay A., Boore D.M., Campbell K.W., Chiou B.S.-J., Darragh R., Day S., Donahue J., Graves R.W., Gregor N., Hanks T., Idriss I.M., Kamai R., Kishida T., Kottke A., Mahin S.A., Rezaeian S., Rowshandel B., Seyhan E., Shahi S., Shantz T., Silva W., Spudich P., Stewart J.P., Watson-Lamprey J., Wooddell K. and Youngs R.; 2014: *NGA-West2 research project*. Earthquake Spectra, **30**, 973-987.

Caputo R. and Pavlides S.; 2013: *The Greek Database of Seismogenic Sources (GreDaSS), version 2.0.0: a compilation of potential seismogenic sources (Mw>5.5) in the Aegean Region*. Doi: 10.15160/unife/gredass/0200, <gredass.unife.it>.

CEN; 2004: *Eurocode 8: design of structures for earthquake resistance - part 1: general rules, seismic actions and rules for buildings*. European Committee for Standardization, Brussels, Belgium, Directive 98/34/EC, Directive 2004/18/EC, BS EN 1998-1 2004, 231 pp.

Chiou B., Darragh R., Gregor N. and Silva W.; 2008: *NGA project strong-motion database*. Earthquake Spectra, **24**, 23-44.

Di Capua G., Lanzo G., Pessina V., Peppoloni S. and Scasserra G.; 2011: *The recording stations of the Italian strong motion network: geological information and site classification*. Bull. Earthquake Eng., **9**, 1779-1796.

DISS Working Group; 2018: *Database of Individual Seismogenic Sources (DISS), version 3.2.1: a compilation of potential sources for earthquakes larger than M 5.5 in Italy and surrounding areas*. Istituto Nazionale di Geofisica e Vulcanologia, Roma, Italy, doi: 10.6092/INGV.IT-DISS3.2.1, <diss.rm.ingv.it/diss>.

Douglas J. and Boore D.M.; 2011: *High-frequency filtering of strong-motion records*. Bull. Earthquake Eng., **9**, 395-409.

Ekström G., Nettles M. and Dziewonski A.M.; 2012: *The global CMT project 2004-2010: centroid-moment tensors for 13,017 earthquakes*. Phys. Earth Planet. Inter., **200-201**, 1-9.

Grünthal G. and Wahlström R.; 2012: *The European-Mediterranean Earthquake Catalogue (EMEC) for the last millennium*. J. Seismolog., **16**, 535-570.

ISIDe Working Group; 2016: *Italian Seismological Instrumental and Parametric Data-base, version 1.0*. Istituto Nazionale di Geofisica e Vulcanologia, Roma, Italy, doi: 10.13127/ISIDe.

Kaklamanos J., Baise L.G. and Boore D.M.; 2011: *Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice*. Earthquake Spectra, **27**, 1219-1235.

Lanzano G., Luzi L., Pacor F., Felicetta C., Puglia R., Sgobba S. and D'Amico M.; 2019a: *A revised ground motion prediction model for shallow crustal earthquakes in Italy*. Bull. Seismol. Soc. Am., **109**, 525-540, doi: 10.1785/0120180210.

Lanzano G., Sgobba S., Luzi L., Puglia R., Pacor F., Felicetta C., D'Amico M., Cotton F. and Bindi D.; 2019b: *The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and data statistics*. Bull. Earthquake Eng., **17**, 561-582.

Luzi L., Puglia R., Russo E., D'Amico M., Felicetta C., Pacor F., Lanzano G., Çeken U., Clinton J., Costa G., Duni L., Farzanegan E., Gueguen P., Ionescu C., Kalogeras I., Özener H., Pesaresi D., Sleeman R., Strollo A. and Zare M.; 2016: *The engineering strong motion database: a platform to access Pan-European accelerometric data*. Seismol. Res. Lett., **87**, 987-997.

Mai P.M. and Thingbaijam K.K.S.; 2014: *SRCMOD: an online database of finite-fault rupture models*. Seismol. Res. Lett., **85**, 1348-1357.

Pacor F., Paolucci R., Luzi L., Sabetta F., Spinelli A., Gorini A., Nicoletti M., Marcucci S., Filippi L. and Dolce M.; 2011: *Overview of the Italian strong-motion database ITACA 1.0*. Bull. Earthquake Eng., **9**, 1723-1739.

Pacor F., Felicetta C., Lanzano G., Sgobba S., Puglia R., D'Amico M., Russo E., Baltzopoulos G. and Iervolino I.; 2018: *NESS v1.0: a worldwide collection of strong-motion data to investigate near source effects*. Seismol. Res. Lett., **89**, 2299-2313.

Paolucci R., Pacor F., Puglia R., Ameri G., Cauzzi C. and Massa M.; 2011: *Record processing in ITACA, the new Italian strong-motion database*. In: Akkar S., Gulkan P. and van Eck T. (eds), Earthquake Data Eng. Seismol., Geotech., Geol., and Earthquake Eng. Series, **14**, 99-113.

Puglia R., Russo E., Luzi L., D'Amico M., Felicetta C., Pacor F. and Lanzano G.; 2018: *Strong-Motion processing service: a tool to access and analyse earthquakes strong-motion waveforms*. Bull. Earthquake Eng., **16**, 2641-2651.

Scognamiglio L., Tinti E., Michelini A., Dreger D.S., Cirella A., Cocco M., Mazza S. and Piatanesi A.; 2010: *Fast determination of moment tensors and rupture history: what has been learned from the 6 April 2009 L'Aquila earthquake sequence*. Seismol. Res. Lett., **81**, 892-906.

Seyhan E., Stewart J.P., Ancheta T.D., Darragh R.B. and Graves R.W.; 2014: *NGA-West2 site database*. Earthquake Spectra, **30**, 1007-1024.

Shahi S.K. and Baker J.W.; 2014: *NGA-West2 models for ground motion directionality*. Earthquake Spectra, **30**, 1285-1300.

Trifunac M.D. and Brady A.G.; 1975: *A study on the duration of strong earthquake ground motion*. Bull. Seismol. Soc. Am., **65**, 581-626.

Wald D.J. and Allen T.I.; 2007: *Topographic slope as a proxy for seismic site conditions and amplification*. Bull. Seismol. Soc. Am., **97**, 1379-1395.

Weatherill G., Bindi D., Cotton F., Danciu L. and Luzi L.; 2018: *Building a new ground motion logic tree for Europe: needs, challenges and new opportunities from European seismological data*. In: Proc. 16th European Conference on Earthquake Engineering, Thessaloniki, Greece, <papers.16ecee.org/files/10729%20final%20corr.pdf>.

Yong A., Hough S.E., Iwahashi J. and Braverman A.; 2012: *A terrain-based site-conditions map of California with implications for the contiguous United States*. Bull. Seismol. Soc. Am., **102**, 114-128.

Corresponding author: Giovanni Lanzano
Istituto Nazionale di Geofisica e Vulcanologia
Via Corti 12, 20133 Milan, Italy
Phone: +39 02 23699259; e-mail: giovanni.lanzano@ingv.it