Statistical analysis of European accelerogram parameters: a contribution to observational hazard in Europe

C.S. Oliveira
Instituto Superior Técnico, Lisboa, Portugal

G. Gassol, T. Susagna & X. Goula
Institut Geològic de Catalunya, Barcelona, Spain

C. Papaioannou
ITSAK, Thessaloniki, Greece

P. Gueguen
Laboratoire de Gépysique Interne, Grenoble, France

J. Clinton
ETHZ, Zurich, Switzerland

C. Zulfikar
KOERI, Istanbul, Turkey

S. Godey
EMSC-CSEM, Bruyères-le-Chatel, France

ABSTRACT:
A new accelerogram distributed database has been created recently (Roca et al., 2010), in the framework of NERIES Project (2006). The seven following European Networks, IGC, IST Azores, IST S. Portugal, LGIT/RAP, KOERI, ITSAK and ETHZ, have contributed with preliminary data to the database. Data are open to the scientific and engineering community through the European Earthquake Data Portal (www.seismicportal.eu). Several parameters with engineering importance were computed in a homogeneous way, for a total of circa 8350 three-components records.

In this work we use the collection of PGA and PGV parameters from accelerometric records at several regions to generate an estimate of the "observed hazard". Using the "ergodic assumption" we were able to estimate return periods (RP) with 30% uncertainty for "average sites" representative of the region covered by each network. We obtained results for RP=10 yrs in the 7 networks and for RP=100 yrs in 3 of them.

Keywords: European accelerograms; NERIES Project; statistic analysis; observational hazard.

1. INTRODUCTION

In the absence of a general strong motion archive for recent European data, the NERIES project (2006) included the initiative to create a distributed database for accelerometric waveforms in order to provide them to the scientific and engineering communities.

To achieve this goal several tasks were developed by the agencies (IGC, IST, LGIT, KOERI, ITSAK, ETHZ and EMSC-CSEM) participating in the Project: a detailed characterization of recording instruments and sites of the accelerometric stations; the development of a computer software to determine in a homogenized way a collection of parameter of engineering interest; and the development of a web-portal to manage the access of users to retrieve the parameter values and waveform data (Roca et al., 2010).

Earthquake Engineering usually, besides of accelerograms and response spectra, uses a set of significant parameters which are important for a better characterisation of ground motion and may be important to analyse structural behaviour including damage assessment for risk mitigation.

In this paper we have proceeded with a first analysis of the ground motion parameters from a large set of digital accelerograms assembled in the data base as October 2009, recorded since 1995, belonging to the seven different European Networks. The following parameters with engineering importance, computed in a homogeneous way with a standard procedure, were selected for the analysis: PGA (cm/s^2); PGV (cm/s); AI (cm/s); TD (s); CAV (cm/s); and HI (cm), together with PSV(f) for 28 frequencies. A total of circa 8350 three-components records from 424 stations associated to 1379 events characterized by magnitude (M1-7) and hypocentral location (0-863 km) were assembled in this work.

Certainly not complete, this set comprises a number of different situations related to event sizes, hypocenters, recording stations and distances, and they can be considered a representative sample of the Euro-Mediterranean Region.

A statistical analysis of each one of the above referred parameters as well as the correlation structure among different pairs is under progress. In this work we performed a statistical analysis of PGA and PGV for different geographical zones covered by each Network, in order to compare their values through an ”observed hazard indicator”.

2. DATA SETS CONTRIBUTING TO THE PRESENT STUDY

The seven following European Networks, participating in above mentioned NERIES project: IGC, IST Azores, IST S. Portugal, LGIT/RAP, KOERI, ITSAK and ETHZ, have contributed with preliminary data to the new distributed European database. Figure 2.1a shows a map with the location of the stations providing data and Figure 2.1b presents the events recorded by the stations.

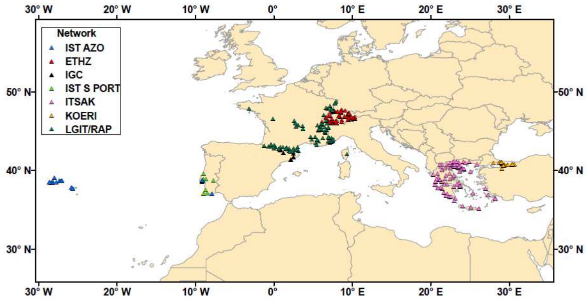

Figure 2.1a. European station distribution providing preliminary data.

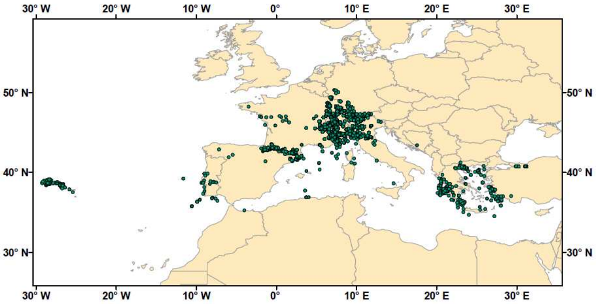

Figure 2.1b. Epicenters of recorded events.

Table 2.1 shows the number of records (components) processed by each Network, with reference to the time interval of events and range of magnitudes. A total of 25040 records were processed and their parameters are analysed after a detailed evaluation of data quality.

Table 2.1. - Number of accelerometric records (components) assembled in database of NERIES project (as December 2009) by Network, Stations, Dates, Magnitude range and Epicentral distance (from Roca et al., 2010).

NETWORK	# Stations	Dates	# Events	Magnitude	Epic. Distance (km)	# records
IST	46	1996-2006	238	2.1 - 5.9	1 - 490	1158
IGC	11	1996-2008	71	1.0 - 5.2	3 - 240	345
LGIT/RAP	103	1995-2007	378	3.0 - 6.8	1 - 863	5253
KOERI	38	Izmit-Eq.1999	7	5.2 - 7.4	13 - 273	369
ETHZ	113	2003-2009	286	2.5 - 5.5	0 - 495	15536
ITSAK	123	2003-2008	399	2.8 - 6.9	2 - 697	2379
Total	424	1995-2009	1379	1.0 - 7.4	0 - 863	25040

Data analyzed in this work, certainly not complete, comprises a number of different situations related to event sizes, hypocenters, recording stations and distances, though they can be considered a representative sample of the Euro-Mediterranean Region. A first observation can be made on the rather different number of stations per Network, allowing a clear separation of networks into two groups: LGIT/RAP, ETHZ and ITSAK are the networks with the larger number of stations and IST (divided into Azores and South Portugal), IGC and KOERI having a much less number (10-30%).

The selection process of data was made in two sequential steps: (i) by visual analysis of plotting different parameters as a function of magnitude and distances and also from analysis of correlations between pairs of parameters, it was possible to identify outliers from the general trends. This applies to distances from wrong event locations, records with poor signal/noise ratios, and records with erroneous units (eg. cm/s^2 instead of tenth of g); (ii) we retained events with M>3 for statistical analysis, and we have separated Azores Islands from South Portugal in the IST records.

Table 2.2 shows the number of events and records, respectively, by classes of magnitude for each Network and Figure 2.2 presents the distribution of records per Network corresponding to different Magnitudes and Distances, after the selection process.

Table 2.2. Number of events and records, respectively, by classes of magnitude for each Network after the selection process.

	3 < M < 4		4 < M < 5		5 < M < 6		M > 6		TOTAL	
NETWORK	# events	# records	# events	# records	# events	# records	# events	# records	# events	# records
IST Azores	140	621	31	168	5	39	0	0	176	828
IST S. Portugal	11	51	8	72	4	57	0	0	23	180
IGC	18	111	3	12	2	24	0	0	23	147
LGIT	325	4071	51	840	14	300	1	21	391	5232
KOERI	0	0	0	0	4	264	2	111	6	375
ETHZ	71	1648	7	227	2	27	0	0	80	1902
ITSAK	150	693	195	807	46	258	7	113	398	1871
All networks	564	6523	256	1886	68	873	10	245	1097	10535

This selection process have reduced the total number of records (components) to less than half, essentially due to the elimination of events with M<3.

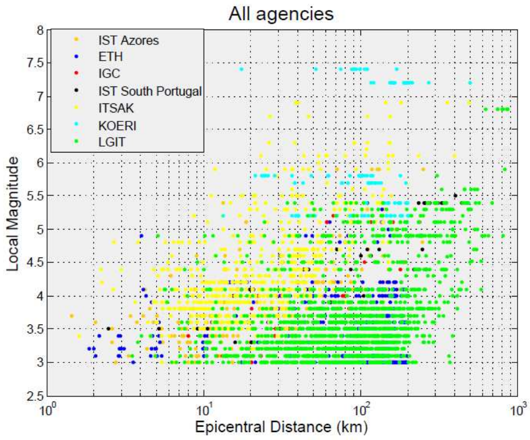

Figure 2.2. Distribution of records (components) corresponding to different Magnitudes and Distances.

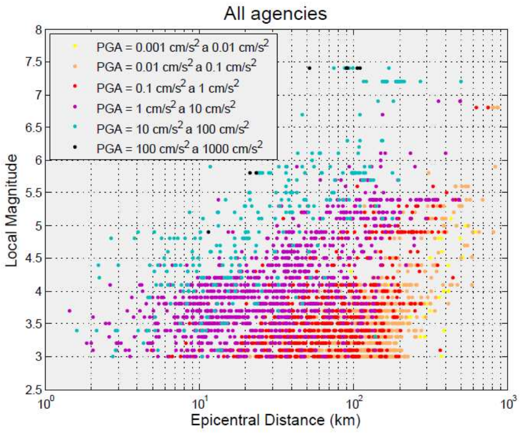

Figure 2.3. Distribution of PGA values of the recordings on a magnitude - distance plot.

3. DATA ANALYSIS

In addition to the acceleration time series, earthquake engineers usually use various simplified waveform parameterisations. These can allow simple characterisation of the complex ground motion for use in analysing expected structural response to the ground motion. Therefore, the distributed database provides not only the raw, complete accelerograms, but also the response spectra and several engineering parameters computed in a homogeneous way for each record (component). For more details on the standard processing procedures, see Roca et al., 2010.

Figure 2.3 shows the distribution of PGA values of records on a magnitude-distance plot. Colours of dots are different for different bins of PGA values. It is very clear that the data are not homogeneously distributed with larger amount referring to smaller magnitude events exhibiting a trend enforced by the attenuation. Similar trend is observed in plots of the other computed parameters.

3.1 Histograms of observed parameters (horizontal components)

In this section we first analyze PGA and PGV in statistical way, Network by Network. We have taken as measure of each parameter the average of the two horizontal components. Note in this analysis we did not include KOERI because the records correspond exclusively to few events of the Izmit 1999 earthquake sequence.

Figure 3.1 shows, as an example, the histograms of PGA and PGV by bins, in a logarithmic base (two per decade). The different histograms show that the networks have very different sensitivities. In fact, ITSAK and IST recorded essentially strong ground motions, i.e. ground motions with PGA<1 cm/s^2 are not well recorded. In the other extreme, ETHZ and LGIT/RAP recorded a lot of very weak motions.

If we concentrate only in the records with PGA higher than 10 cm/s^2 (0.01g which approximates to felt motions, the curves show a common pattern which can be seen as a “hazard” indicator. In fact, this pattern follows a trend similar to the typical hazard curve for a site (Oliveira and Campos-Costa, 2006).

The value of 10 cm/s^2 (0.01g) has been reached less than 10 times by IGC and IST S. Portugal; more than 20 times by ETHZ; near 50 times by LGIT/RAP and IST Azores and more than 100 times by ITSAK stations. The value of 100 cm/s^2 (0.1g) has been only reached 2 times by LGIT/RAP and 20 times by ITSAK.

PGV distribution (Figure 3.1b) shows for values higher than 0.03 cm/s a common pattern which can be seen as a “hazard” indicator, as it was referred for PGA, but in this case with higher resolution, suggesting that this parameter should be a better indicator of “hazard” than PGA.

The value of 0.1 cm/s was reached 10 times by IGC, more than 20 times by ETHZ and IST S. Portugal; more than 70 times by LGIT and IST Azores and more than 200 times by ITSAK network. The value of 1 cm/s was reached only 3 times by ETHZ, more than 10 times by LGIT and IST Azores and more than 60 times by ITSAK. The value of 10 cm/s was only reached by ITSAK, more than 10 times.

In Figure 3.2a) and b) we plot the number of records observed by all a networks organized by PGA and PGV bins, respectively, for different magnitudes. In both Figures we can see that smaller magnitudes decrease faster towards larger values of PGA and PGV.

PGA values less than 6 cm/s^2 (0.006g) are more represented by the lowest magnitudes (M3-4); between 6 cm/s^2 and 50 cm/s^2 (0.05g) earthquakes with M4-5 are contributing with a large number of records; between 50 cm/s^2 and 180 cm/s^2, M5-6 are the most represented; for higher values of PGA (>180 cm/s^2) only earthquakes with M>6 have contributed.

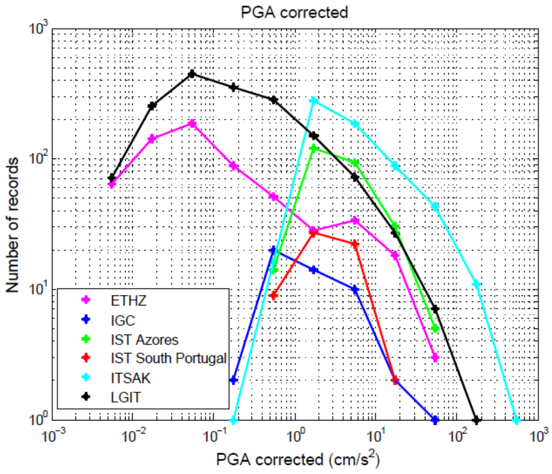

Figure 3.1a. Histogram of PGA values per networks.

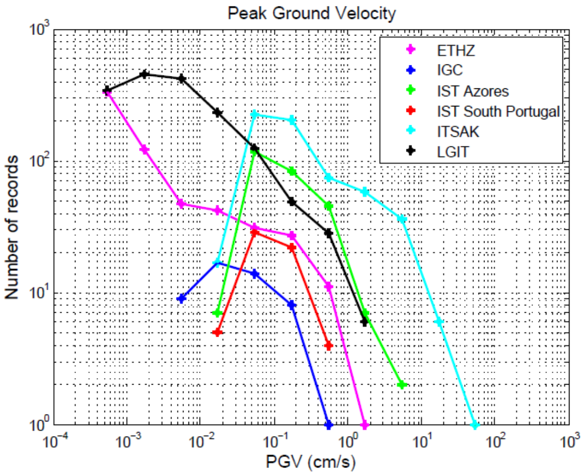

Figure 3.1b. Histogram of PGV values per networks.

In relation to PGV the tendency it is more regular. We see that as we can expect the decrease of the number of records is organized by magnitude values: for values lesser than 0.1 cm/s the larger number of records are obtained from earthquakes with M3-4; for values between 0.1 cm/s and 1 cm/s, the dominating number of records come from earthquakes with M4-5; from 1 cm/s to 20 cm/s records come mainly from M5-6; greater values, up to 50 cm/s, have only been observed once for an earthquake with M>6.

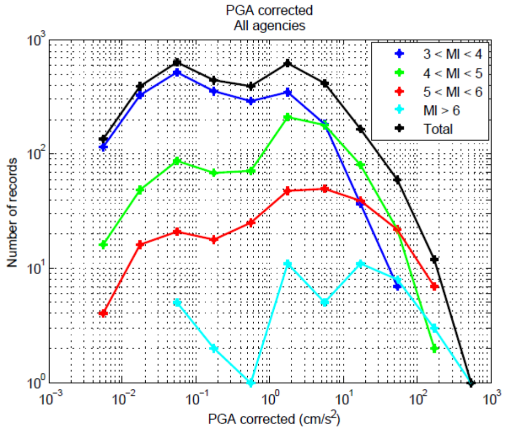

Figure 3.2a. Number of records observed by horizontal PGA bins, for different magnitudes.

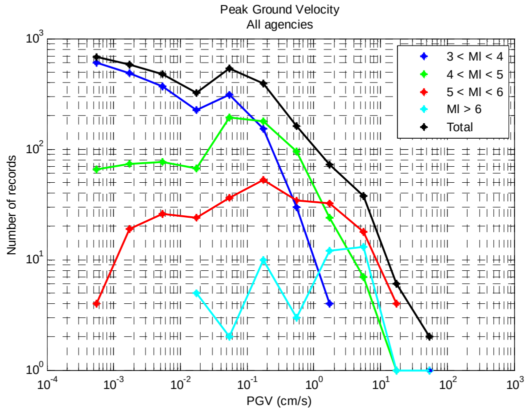

Figure 3.2b. Number of records observed by horizontal PGV bins, for different magnitudes.

3.2 Cumulated mean frequency of records as a “observed hazard” indicator

We propose to use the PGA and PGV observations quantified in the previous section in a way that can be used as an observed hazard indicator for the regions represented by the networks. In fact, discrepancies often encountered on the results of Probabilistic Seismic Hazard Assessments (PSHA) (Ordaz and Reyes, 1999, Ward, 1995 and Beauval et al. 2007) have highlighted the necessity to try to use available observations to constrain hazard estimates.

Considering that the number of stations operated by each network and the periods of operation are different, to perform a homogenised comparison among rates estimated from data recorded by different networks it is necessary to normalize the number of records of each bin by a factor which considers the number of stations and the period of operation of each station for which data was gathered.

In this way we can obtain mean frequency of records for each above mentioned bin per year, for an "average" site of the area covered by the stations of each network.

The above mentioned normalizing factor can be in fact computed for each station of each network. An average value of this factor, named "Equivalent number of years of observation" (ENYO) for each network is shown in Table 3.1, together with the number of stations and a mean number of years of operation.

Table 3.1. Total number of stations, mean value of years of operation and "equivalent number of years of observation (ENYO)" per each network.

NETWORK	Total # of stations	Mean # of Years of Operation	Equivalent # of years of operation (ENYO)
IST Azores	26	9	229
IST South Portugal	23	8	179
IGC	11	10	107
LGIT	103	7	756
ETHZ	113	7	744
ITSAK	113	5	613

This mean frequency for each bin can be cumulated in order to obtain an equivalent annual exceedence probability of different values of the chosen parameter for an "average site" among the recording stations. We are using the named "ergodic assumption" as it was proposed for example by (Ward, 1995), that permits to make an equivalence between counting a large series of records in a site and counting a shorter series of records in different sites, with homogeneous level of hazard. Of course this average site is a better representation of "hazard" if the zone covered by the network is smaller and more homogeneous.

In Figure 3.3 we present the same data shown in Figure 3.1, corrected now by the factor ENYO obtaining the mean frequency for each bin, which is cumulated in order to obtain an equivalent annual exceedence probability of different values of the chosen parameter. Thus, these values correspond to the lower value of the bin.

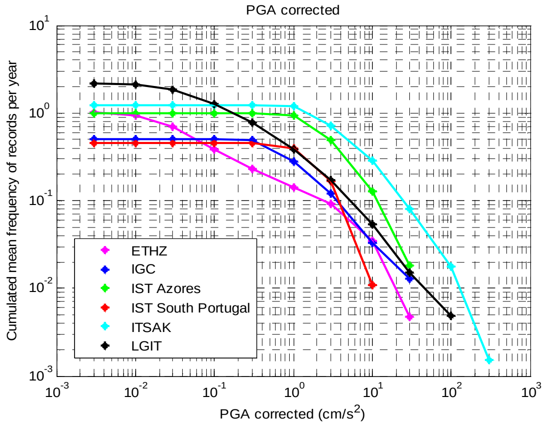

Figure 3.3a. Cumulated mean frequency of observed horizontal PGA values per year

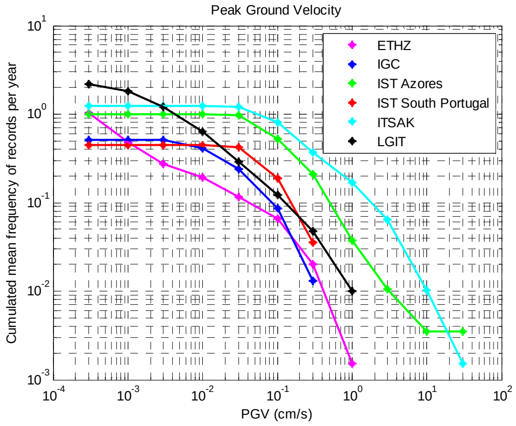

Figure 3.3b. Cumulated mean frequency of observed horizontal PGV values per year

We summarize all the results obtained for the different parameters in Tables 3.2 and 3.3, in terms of Return Periods (RP) in years, which can be defined as the inverse of cumulated mean frequency of observed horizontal values per year.

In a recent study based on the properties of Poisson processes, Beauval et al., 2007 have evaluated the minimum time windows insuring reliable rate estimates at a site. For example, for a ground-motion with a 10-years RP at a site 100 years observation time window is required; or for a value with 100-years RP a minimum of 1000 years observation time window is required for estimating the rate with 30% uncertainty (Figure 3 in Beauval et al., 2007).

In the Tables 3.2 and 3.3 we show in shaded cells, the RP calculated with a minimum time window leading to 30% uncertainty. Blank cells have higher uncertainties. Due to the larger uncertainties of these last values, they will not be considered for discussion.

Table 3.2. Observed PGA (cm/s^2) associated to different Return Period (RP) per network. Shaded cells correspond to estimations having 30% uncertainty. Blank cells are associated to higher uncertainties.

	NETWORK	IST South Portugal	ETHZ	IGC	LGIT/RAP	IST Azores	ITSAK
Return Period (RP - years)	10	4	2	3.5	5	11	24
	50	8	13	18	23	29	90
	100	10	20		45		130
	200		30		100		180
	500						260

Table 3.3. Observed PGV (cm/s) associated to different Return Period (RP) per agency. Shaded cells correspond to estimations having 30% uncertainty. Blank cells are associated to higher uncertainties.

	NETWORK	IST South Portugal	ETHZ	IGC	LGIT/RAP	IST Azores	ITSAK
Return Period (RP - years)	10	0.15	0.04	0.09	0.12	0.5	2
	50		0.3	0.22	0.30	1.8	6.5
	100		0.4		1	3	9
	200		0.6			6.5	17
	500		0.9				28

Only RP of 10 years have been estimated with a level of 30% of uncertainty for all the networks. The results show the lower value of PGA of 2 cm/s^2 for ETHZ data; a factor of near 2 appear for IST S. Portugal, LGIT/RAP and IGC; a factor of 5 for IST Azores and a factor of 12 for ITSAK values. For a RP of 100 years - the longest reliable estimations of RP - the lowest value is obtained for ETHZ records: 20 cm/s^2 (0.02 g); a factor of 2 is observed for LGIT/RAP data and a factor of more than 6, for ITSAK. An estimation for a RP of 500 years is obtained only for ITSAK data (260 cm/s^2 or 0.26 g), but this value is associated with large uncertainty.

Analogous results are obtained for PGV, but with larger differences between networks. For a RP of 10 years, ETHZ shows the lowest values: 0.04 cm/s; IGC a factor of 2; LGIT/RAP a factor of 3; IST S. Portugal: a factor of 4; IST Azores a factor 12 and ITSAK a factor of 50. For a RP of 100 years: ETHZ show the lowest value: 0.04 cm/s; LGIT/RAP a factor of 3 and ITSAK a factor of more than 20.

4. CONCLUSIONS

In this work we use the collection of PGA and PGV parameters from accelerometric records at regions of the seven European Networks, IGC, IST (Azores and South Portugal), LGIT/RAP, KOERI, ITSAK and ETHZ, contributing with preliminary data to the European-Mediterranean distributed accelerometric database (Roca et al., 2010), to generate an estimate of the “observed hazard”.

Using the “ergodic assumption” we were able to estimate the cumulated mean frequency of PGA and PGV values organized by bins, per year for “average sites” representative of the region covered by each network. From them we estimated return periods (RP) and uncertainties. In fact, following a recent study based on the properties of Poisson processes, Beauval et al. (2007) have evaluated the minimum time windows insuring reliable rate estimates at a site. For example, for a ground-motion with a 10-years RP at a site 100 years observation time window is required; or for a value with 100-years RP a minimum of 1000 years observation time window is required for estimating the rate with 30% uncertainty.

With these considerations we obtained estimations of PGA and PGV for RP=10 yrs in the 7 networks: PGA varies from 2 cm/s^2 (ETHZ) to 24 cm/s^2 (ITSAK); PGV varies from 0.04 cm/s (ETHZ) to 2.0 cm/s (ITSAK). For RP=100 yrs only estimations for ETHZ, LGIT/RAP and ITSAK are reliable with 30% uncertainty: PGA varies from 20 cm/s^2 (ETHZ) to 130 cm/s^2 (ITSAK); PGV varies from 0.4 cm/s (ETHZ) to 9.0 cm/s (ITSAK).

This is a first tentative to use strong motion data on a regional basis (European-Mediterranean Area) to contribute to PSHA studies. The approach presented here should be extended to more regions and covering longer RP of engineering interest. It requires the availability of more data which come from the installation of more instrumentation enlarging the time window of observations.

ACKNOWLEDGEMENTS

This paper has been partially supported by EC Project NERIES, Sixth Framework Programme, Contract number: RII3-CT-2006-026130. Special thanks are due to all author´s Institution for financial contributions. The authors would like to acknowledge all the networks that have contributed with their data, including the French RAP (Reseau Accelerometric Permanent), represented by LGIT. Special thanks are due to Antoni Roca for Project coordination, all support, and technical discussions and to Ramon Secanell for enlightening the statistical aspects of the paper.

REFERENCES

Beauval, C.; Bard, P.-Y.; Hainzl, S. ; Guéguen, P. (2008). Can Strong-Motion Observations be Used to Constrain Probabilistic Seismic-Hazard Estimates? , *BSSA*, **98, No. 2**, pp. 509–520.

NERIES Project (2006) Network of Research Infrastructures for European Seismology. (http://www.neries-eu.org/). Accessed 27 January 2010.

Oliveira C. S. and Campos-Costa. (2006). Overview on earthquake hazard assessment–Methods and new trends. In: C. S. Oliveira, A. Roca and X. Goula, (Editors), Assessing and Managing Earthquake Risk, 15 - 46, Springer.

Ordaz, M.; Reyes, C. (1999). Earthquake hazard in Mexico City: observations versus computations, *BSSA*, **5**, 1379-1383.

Roca, A., Guéguen, P., Godey, S., Goula, X., Susagna, T., Péquegnat C., Oliveira, C.S., Clinton, J., Pappaioanou, C., Zulfikar, C. (2010). The Euro-Mediterranean distributed accelerometric data-base. Geotechnical, Geological and Earthquake Engineering Book Series. Springer (Accepted for Publication).

Ward, S. (1994) A multidisciplinary approach to seismic hazard in Southern California, *BSSA*, **84**, 1293-1309.