16th World Conference on Earthquake, 16WCEE 2017

Santiago Chile, January 9th to 13th 2017

Paper N° 4117

Registration Code: S-M1461936734

A METHOD FOR COMPARING PSHA RESULTS TO HISTORICAL OBSERVATIONS USING FRAGILITY CURVES

A. Rosti[(1)], M. Rota[(2)], A. Penna[(3)], G. Magenes[(4)]

[(1)] *PhD student, ROSE Programme, UME School, Institute for Advanced Studies (IUSS), Pavia, Italy, via Ferrata 1, 27100 Pavia, e-mail address: annalisa.rosti@iusspavia.it*

[(2)] *Researcher, European Centre for Training and Research in Earthquake Engineering (EUCENTRE), Pavia, Italy, via Ferrata 1, 27100 Pavia, e-mail address: maria.rota@eucentre.it*

[(3)] *Associate Professor, Department of Civil Engineering and Architecture, University of Pavia, Italy, via Ferrata 3, 27100 Pavia and EUCENTRE, Pavia, Italy, via Ferrata 1, 27100, Pavia, e-mail address: andrea.penna@unipv.it*

[(4)] *Professor, Department of Civil Engineering and Architecture, University of Pavia, Italy, via Ferrata 3, 27100 Pavia and EUCENTRE, Pavia, Italy, via Ferrata 1, 27100, Pavia, e-mail address: guido.magenes@unipv.it*

Abstract

A method for comparing probabilistic seismic hazard analysis (PSHA) results with historical macroseismic observations is proposed, with specific application to the South-East French territory. As the mean damage, i.e. the average annual damage expected in the historical building stock, is the selected measure of comparison, procedures for respectively converting macroseismic intensities and PGA levels, for which PSH estimates are provided, into mean damage values are first presented. The originality of the method lies in the implementation of a logic tree approach, handling the different sources of uncertainty, and in the use of empirical fragility functions, representative of the vulnerability of the old building stock. Different approaches for comparing PSH results with historical observations at different scales are then outlined. Site-specific comparisons are first presented. To overcome the lack of macroseismic data at single sites, a method for aggregating the information available at a set of sites is then proposed. The comparison is then extended to the regional scale. To this aim, spatially-correlated PGA random fields, constrained to the available macroseismic observations, are generated.

Keywords: historical macroseismic observations, probabilistic seismic hazard analyses, empirical fragility curves, mean damage

1. Introduction

The number of PSH studies has been significantly increasing in the last decades. Attention has been addressed towards the definition of more precise and reliable methods and tools for better quantifying the seismic hazard and the associated uncertainties. In the case of low-to-moderate seismicity zones, the uncertainties due to the lack and inhomogeneous quality of the available data are significant. This is for instance the case of France, for which three different maps were recently established, leading to significant variations in the hazard assessment of the metropolitan territory [1]. Despite the absence of a general consensus on the criteria to be used for evaluating the performance of hazard maps (e.g. [2], [3], [4]), the need of testing and/or validating PSH results is evident. In order to obtain pertinent comparisons, independent observations, that is data not directly included in the PSHA, should be considered [5]. Different types of observations can be used, i.e. accelerometric data (e.g. [6], [7], [8], [9]), historical macroseismic observations (e.g. [1], [10]) and information from fragile geological structures (e.g. [11], [12]). The comparison of PSH results with observations can be directly carried out at single sites. Nonetheless, different works (e.g. [7], [8], [9], [13]) assembled several sites to make up for the lack of observations and/or the limited working period of accelerometric stations, if using accelerometric data. The scale of the comparison can be then extended, up to the regional level (e.g. [1]).

This paper presents different methods, developed in the framework of the research project Sigma[1], for comparing PSH results and historical macroseismic observations, with specific application to the South-East French territory. To this aim, PSH results derived by Carbon et al. [14] are considered. The mean damage annually expected in the old building stock is the selected measure of the comparison. Procedures for making PSH results and observations comparable are first outlined. PSH results and information from historical observations are then compared at different scales, namely at single sites, by aggregating a set of sites and at the regional scale.

2. Conversion of macroseismic intensities into mean damage values

The proposed methodology starts with the identification of sites, located within the area of interest, for which information on both historical macroseismic intensities and old building stock are available. To obtain pertinent comparisons with PSH estimates, macroseismic intensities observed at a given site need to be sufficient in terms of number, entity and distribution in time. Also, it is necessary to collect information on the building stock at the time of the historical observations and its subdivision into different structural typologies. Four building typologies were identified for the study area (i.e. South-East quadrant of France). They all consist of stone masonry buildings with flexible floors and differ one from the other for the number of stories and for the presence or absence of tie-rods and tie-beams (Table 1). Based on the environmental context of a given site (i.e. city, village on the Alps or small village), different weights were attributed to each building typology (Table 1) to account for the subdivision of the building stock in the different structural typologies.

Table 1 – Indentified building typologies and associated weights depending on the environmental context of a site

Typology	Description	Cities	Villages Alps	Small villages
Typ.1	with tie-rods and/or tie-beams -1-2 stories	0.05	0.10	0.10
Typ.2	w/o tie-rods and tie-beams - 1-2 stories	0.50	0.60	0.70
Typ.3	with tie-rods and/or tie-beams - >2 stories	0.15	0.10	0.05
Typ.4	w/o tie-rods and tie-beams - >1-2 stories	0.30	0.20	0.15

For each reported observed macroseismic intensity, the French macroseismic database, Sisfrance [15], provides a different code (i.e. A, B or C), depending on the reliability of the observation itself. Code A indicates

[1] Sigma was a research project supported by EDF, Areva, ENEL and CEA

that the value of the observed intensity is certain. Code B denotes a fairly certain intensity value, whilst code C refers to an uncertain intensity value. This type of uncertainty was accounted for by introducing weighted discrete intensity distributions, whose dispersion depends on the reliability of the observation. In the case of sure intensity value (i.e. code A), only the value of intensity reported in the catalogue, *I*, was considered. In the case of fairly certain intensity value (i.e. code B), the intensity value reported in the catalogue, *I*, and the intensity levels, *I*±0.5 were considered. In the case of code C (i.e. uncertain intensity value), the reported intensity value, *I*, together with the levels *I*±0.5 and *I*±1 were defined. In the case of reliability codes B and C, normal distributions centered on the reported intensity value and with standard deviations of 0.25 and 0.50, respectively, were defined to derive weights for each intensity level. Weights (Table 2) were obtained by integrating the area below the normal distribution bounded by midway percentiles (Fig.1).

Table 2 – Weights associated to the different intensity levels based on the reliability of the observation

Intensity Level	Code A	Code B	Code C
I-1	0	0	0.09
I-0.5	0	0.26	0.24
I	1	0.48	0.34
I+0.5	0	0.26	0.24
I+1	0	0	0.09

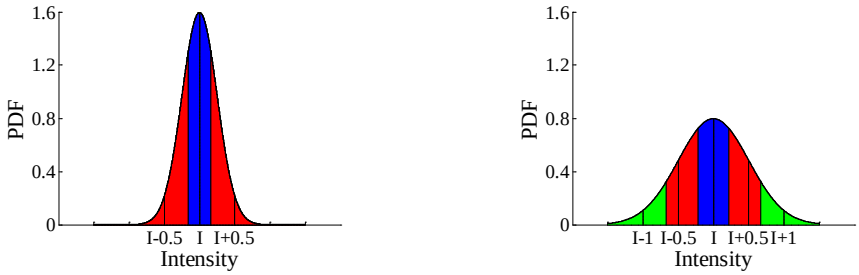

Fig. 1 – Normal distributions considered for the definition of the weights, centered on the reported intensity level: codes B (left) and C (right)

Based on this procedure, each observed macroseismic intensity, characterizing the seismic history of a given site (Fig.2, left), was converted into a weighted discrete distribution of intensity values (Fig.2, right). In the figure, the different colors correspond to the color of the weights associated to each intensity level, in accordance with Fig.1.

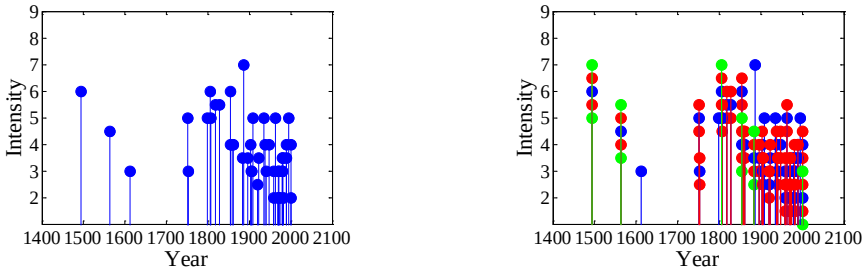

Fig. 2 – Seismic history of a generic site (left) and modified seismic history considering the uncertainty in the macroseismic intensities (right)

The macroseismic method [16] was then applied to convert macroseismic intensities into mean damage values. To this aim, the following closed-form analytical formula, correlating the macroseismic intensity, *I*, and the mean damage, μ_D, as a function of the vulnerability, was used:

$$\mu_D = 2.5 + 2.5 \tanh \frac{I + 6.25V_i - 13.1}{2.3} \quad (1)$$

where V_i are the vulnerability indexes, accounting for uncertainty in the attribution of the different structural typologies to the EMS-98 vulnerability classes [17]. For each building typology, five values of vulnerability index were defined, corresponding to the percentiles of the membership function [18] (Fig.3, left). For each building typology and intensity level, five mean damage values were hence obtained (Fig.3, centre). Similarly to the procedure for defining weights for the different intensity levels (Fig.1), different weights were assigned to vulnerability indexes (Table 3). It is observed that Sisfrance expresses macroseismic intensities according to the MSK-64 scale [19], whereas the macroseismic method refers to the EMS-98 scale [17]. In accordance with previous works [20], the equivalence of the mean values of each intensity class of the two macroseismic scales was assumed. The uncertainty on the intensity values was explicitly considered.

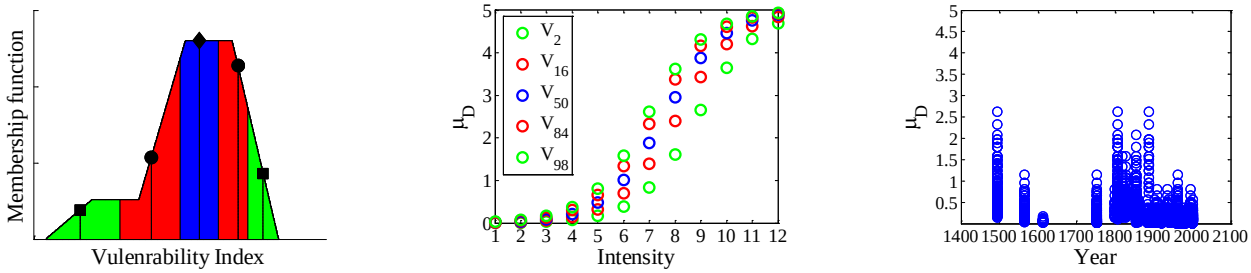

Fig. 3 – Membership function of typology 4 (left), mean damage values for different intensity levels (centre) and equivalent seismic history of a generic site in terms of mean damage (right)

The different uncertainties involved in the proposed methodology (i.e. the reliability of the macroseismic intensity values reported in Sisfrance, the uncertainty in the subdivision of the building stock in different building typologies and the uncertainty in the attribution of the different structural typologies to the EMS-98 vulnerability classes) were handled by a logic tree approach, whose outcome is an equivalent mean damage seismic history (Fig.3, right). It is noted that a weighted discrete mean damage distribution is associated with each year where a macroseismic observation was available. Single μ_D values were then extracted from the corresponding distributions through a Monte Carlo approach. At each sampling, the observation period and the best estimate of the empirically-derived annual rates of exceedance (i.e. number of exceedances over the corresponding observation period) were computed for each preselected μ_D level. In addition, the 90% confidence bounds on the best estimate were derived. By sampling many times, statistics of the best estimate of the empirically-derived annual rates of exceedance and the corresponding confidence limits were obtained.

Table 3 – Vulnerability index of each building typology and associated weight

Vulnerability Index	Typ.1	Typ.2 and 3	Typ.4	Weight
V_2	0.650	0.711	0.679	0.09
V_{16}	0.686	0.773	0.801	0.24
V_{50}	0.737	0.833	0.884	0.34
V_{84}	0.794	0.870	0.950	0.24
V_{98}	0.821	0.897	0.994	0.09

3. Conversion of PGAs into mean damage values

PSHA generally provides the annual rate of exceedance of PGA thresholds (Fig.4, left). A methodology for deriving mean damage values generated by different PGA levels through fragility functions was hence developed. Given the similarity between the South-East French and the Italian old building stock, empirical fragility curves derived from post-earthquake damage data collected after the main Italian events of the period

1980-2002 [21] and integrated by L'Aquila (2009) damage data [22], were selected. The updated damage database includes more than 140'000 buildings, where 58'408 belong to the four building typologies identified as relevant for the study area (Table 1). For each building typology, a mean damage curve as a function of PGA was derived by combining the corresponding fragility curves and assuming the different damage levels are binomially distributed (e.g. [16], [23]):

$$\mu_D = \sum_{k=0}^{5} p_k k \tag{2}$$

where p_k is the probability of having damage level D_k (k = 0÷5).

For each building typology and PGA threshold, a value of mean damage was derived from the corresponding PGA-μ_D curve. Then, for each PGA level, a single μ_D value was computed as the weighted average of the μ_D values corresponding to the different typologies (Fig.4, centre). Based on this procedure, mean damage values were associated to PSH rates of exceedance corresponding to different PGA thresholds (Fig.4, right).

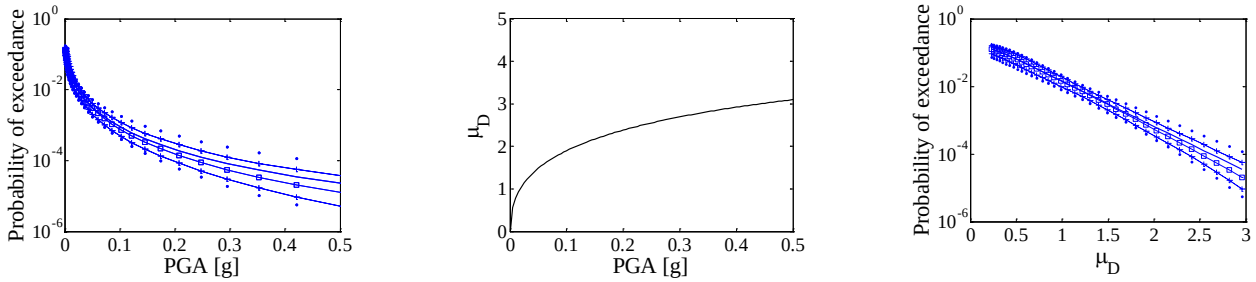

Fig. 4 – PSH curves for different percentiles (left), mean damage curve as a function of PGA (centre) and annual rates of exceedance for different mean damage thresholds

4. Proposed approaches for comparing PSH results and historical observations at different scales

Three different approaches are proposed for comparing PSH results with information from historical observations at three different scales: comparison at single sites, at a set of sites and at the regional scale.

4.1 Site-specific comparison

Comparisons of PSH results and historical macroseismic observations were first carried out at single sites, in terms of annual rates of exceedance of preselected μ_D thresholds. As an example, the comparison is applied to the site of Annecy, whose seismic history (Fig.5, left) shows a fairly good number of macroseismic observations exceeding intensity level 5 (i.e. the intensity level corresponding to slight cracks in the plasterwork, according to the MSK-64 scale). To account for the uncertainty in the macroseismic intensities reported in Sisfrance, each observation was transformed into a weighted discrete distribution of intensity values (Fig.5, centre), according to Section 2. Considering the environmental context of Annecy (i.e. city), weights equal to 0.05, 0.50, 0.15 and 0.30 were respectively attributed to the four identified building typologies (Table 1). The application of the macroseismic method to each observed macroseismic intensity led to an equivalent seismic history in terms of mean damage (Fig.5, right). To make the comparison of PSH results and observations feasible, μ_D values were associated to PGA thresholds, for which PSH estimates are available, by using fragility curves, according to Section 3. Fig.6 compares empirically-derived annual rates of exceedance and PSH results for different mean damage thresholds. In the figure, red corresponds to the best estimate of the empirically-derived annual rate of exceedance. Black and green correspond to the upper and lower bounds of the 90% confidence interval of the annual rates of exceedance. Circles correspond to the median, diamonds to the mean, whereas the error bars take into account the variability in the different Monte Carlo runs. It is noted that, starting from a μ_D threshold of 1, PSH results are in agreement with historical observations, as the best estimate of the empirically-derived rates of exceedance falls within the percentiles of the PSH estimates. Differently, PSH estimates tend to overestimate

results from historical observations at lower mean damage levels. This overestimation could be explained by the fact that some low-entity macroseismic observations may not be reported in the seismic catalogue. The uncertainty on the best estimate of the empirically-derived annual rates is smaller at lower μ_D levels, whereas it increases at higher mean damage thresholds. This can be due to the significant number of low μ_D values, allowing to obtain more confidence in the empirical results at lower mean damage levels.

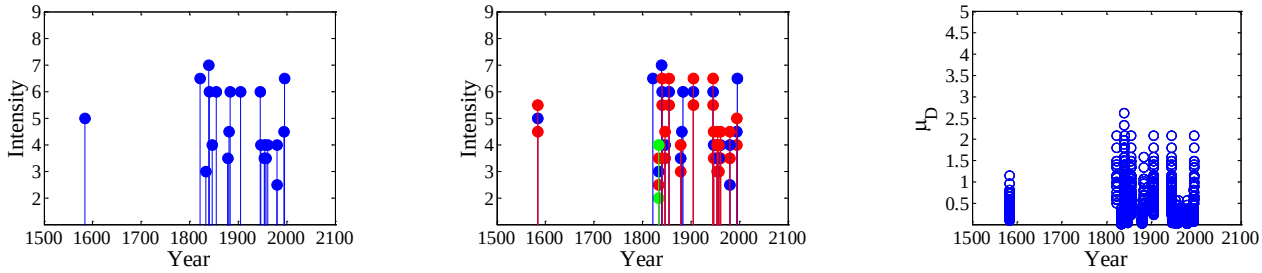

Fig. 5 – Seismic history (left), modified seismic history (centre) and equivalent mean damage seismic history of Annecy (right)

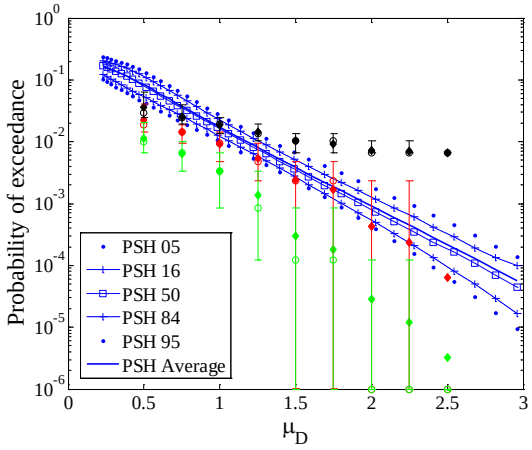

Fig. 6 – Comparison of PSH results for different percentiles with statistics of the best estimate (red) and the upper (black) and lower (green) bounds of the 90% confidence interval of the empirically-derived annual rates of exceedance of preselected mean damage levels (Annecy). Diamonds: mean; circles: median; error bars: 5th and 95th percentiles

4.2 Comparison for aggregated sites

Being direct and immediate, site-specific comparisons could be worthwhile for sites of particular interest, such as nuclear installations sites. Nonetheless, the seismic history of the selected site unavoidably impacts on the results of the comparison. In this sense, sparse macroseismic data of engineering interest may prevent pertinent comparisons with PSH results. This issue can be however faced by sampling in space. Under the assumption of ergodicity of the process of earthquakes' occurrence, time and space can be swapped. Since the number of observations at single sites may not be sufficient, several sites are assembled and treated as a single one. To this aim, observations need to be due to independent seismic events. However, when sampling in space, an issue is given by the stochastic dependency of observations generated by the same seismic event at different sites, that may lead to erroneous conclusions on the adequacy of the PSH model to be tested, if overlooked [24]. However, bearing in mind that sites are not independent, treating them like that, i.e. overlooking the stochastic dependence, does not affect the mean but only the variance of a given distribution of results [24]. Based on these considerations, a procedure for aggregating sites was developed. The comparison is performed in terms of mean annual rates of exceedance of preselected mean damage levels in at least one of the selected sites, so that no assumptions on sites independence is required. Seven sites located within the study area were selected for the

application of the methodology, i.e. Annecy, Albertville, Beaumont de Pertuis, Digne, Draguignan, La Mure, and L'Argentières La Bessée. These sites were afterwards classified according to the three considered environmental categories. In particular, Annecy and Draguignan were attributed to the environmental class of cities, Albertville and La Mure of villages on the Alps, whereas Beaumont de Pertuis, Digne and L'Argentières La Bessée of small villages. Based on the procedure described in Section 2, for each site an equivalent mean damage history was generated starting from its seismic history. Mean damage values were then sampled from the corresponding distributions through a Monte Carlo approach and annual rates of exceedance were computed for preselected μ_D thresholds. For each site and mean damage level, a distribution of annual rates of exceedance was hence obtained and the mean annual rate of exceedance was computed. For a given μ_D level, the mean annual rate of exceedance in at least one of the selected sites was given by the mean annual rates of the different sites divided by the number of sites to be aggregated.

On the other side, PSH rates of exceedance of PGA thresholds were associated to mean damage values, according to Section 3. The epistemic uncertainty in the hazard estimates was accounted for by fitting lognormal distributions through the different percentiles of the available PSH predictions, for each preselected μ_D threshold [25]. For each site and for each mean damage threshold, annual rates of exceedance were then sampled from the corresponding approximating lognormal distributions and the mean of the sampled annual rates of exceedance was computed. The expected annual rate of exceedance of mean damage thresholds in at least one of the selected sites was then derived by summing the mean rates of exceedance of each site and dividing the result by the number of the selected sites. Results obtained are shown in Fig.7, comparing the empirically-derived (red stars) and the PSH (black circles) mean annual rates of exceedance of different μ_D thresholds in at least one of the sites. Starting from a mean damage level of 1.5, PSH results are in agreement with historical observations. For lower mean damage values, PSH estimates tend to overestimate empirical results. Similarly to Section 4.1, these results could be explained by some low entity macroseismic observations not reported in the catalogue. The comparison seems however to be reasonable for μ_D thresholds larger than 1, because lower mean damage values correspond to observations of intensity level lower than or equal to 6 and, consequently, less reliable. Similarly, mean damage values lower than 1 correspond to very low PGAs. In the interval of interest (i.e. μ_D thresholds larger than 1), PSH estimates are consistent with historical observations.

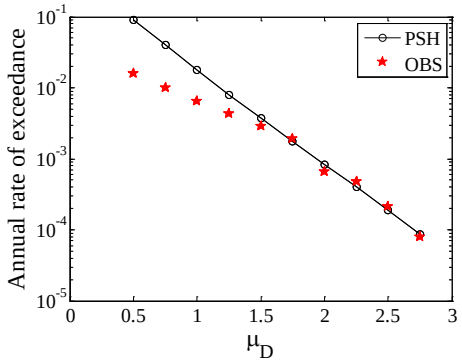

Fig. 7 – Comparison of the empirically-derived (red stars) and PSH-derived (black circles) annual rates of exceedance of mean damage levels in at least one of the selected sites

4.3 Comparison at the regional scale

Analyses of the French historical catalogue pointed out the lack of macroseismic observations of engineering interest for most of the sites located in the South-East French territory. Collecting and gathering together information available at different sites have the clear potentiality of significantly enlarging the size of the available dataset. Hence, a methodology for comparing PSH results with observations at the regional scale was developed. The proposed procedure starts by generating a set of spatially correlated random fields of PGA, constrained to the available macroseismic intensity observations. The advantage of this approach is the possibility of modeling past earthquake scenarios consistently with the available seismological data and observations. Calculations can be carried out on a regular grid of points or on selected sites of interest and the resulting probabilistic ground shaking scenarios can be used to supplement and integrate the available seismic

history at the site. The random fields approach starts from considering that intensity measures from a single event are spatially correlated from site to site, due to the characteristics of the generating earthquake (e.g. stress drop, rupture velocity), path effects and near fault effects (e.g. proximity to fault asperities for sites close to the fault plane). The general methodology is described in detail in Park et al. [26]. One possible method to introduce macroseismic intensities in the modeling of random fields is the Bayesian estimation approach [27].

In this work, random fields were used to derive PGA lognormal distributions at a grid of selected points, compatible with the available historical observations, and representing the ground motion that should have been experienced at the considered sites, due to the occurrence of selected seismic events. After selecting sites and seismic events, the comparison was then performed in terms of the mean annual rate of exceedance of preselected PGA thresholds in at least one of the selected sites, similarly to Section 4.2. Eleven French departments (i.e. Alpes Maritimes, Hautes-Alpes, Haute-Savoie, Vaucluse, Savoie, Isère, Rhône, Drôme, Alpes de-Haute-Provence, Bouches du Rhône, Var) of the South-East French quadrant were selected for this application. In particular, 580 grid points, approximately distributed at 10 km intervals and for which PSH estimates were available [14], were considered (Fig.8, left).

All the seismic events producing macroseismic intensity observations at least equal to 4 within the study area were considered for the generation of synthetic observations at the selected locations. For each event, epicentral location and intensity were retrieved from the Sisfrance online database [15]. Moment magnitude values were instead collected from the SHEEC catalogue [28], not being reported in Sisfrance. For all the seismic events not reported in the SHEEC catalogue, local magnitude values were computed by EDF-DIN-CEIDRE TEGG-Service Geologie Geotecnique. Dependent shocks and events with important information missing, such as the epicentral intensity and magnitude, were discarded. Random fields were generated using the Akkar et al. [29] ground motion prediction equation, together with the spatial correlation model of Jayaram and Baker [30] (long range version). As the Akkar et al. [29] GMPE is considered applicable to the magnitude range from 4 to 8 and for distances up to 200 km, only the seismic events with magnitude at least equal to 4 and located within 200 km of the hazard grid were considered. Furthermore, only seismic events with epicentral intensity at least equal to 5 were taken into account to model earthquakes that produced some level of damage on buildings. One additional event was discarded, as occurred in 463 a.C. This choice was based on considerations on the completeness in time of the catalogue of seismic events which, apart from this very old event, only includes earthquakes in the period 1397-2005. The final dataset hence included 196 seismic events of magnitude ranging from 4 to 6.62. Since the Akkar et al. [29] GMPE was developed for moment magnitude, in the case of earthquakes for which the moment magnitude was not available, the equivalence of local and moment magnitudes was assumed. Indeed, the magnitude range of these events is between 4 and 5.1 and, in this range, the two scales can be considered equivalent [31]. Fig.8 (centre) shows the epicentral intensity of each seismic event versus time. The color and size of the markers depend on the magnitude of the same event (i.e. green: 4≤M<5; orange: 5≤M<6; red: M≥6). Fig.8 (right) shows the epicentral location of each earthquake. Still, the different color and size of the markers correspond to different ranges of magnitude.

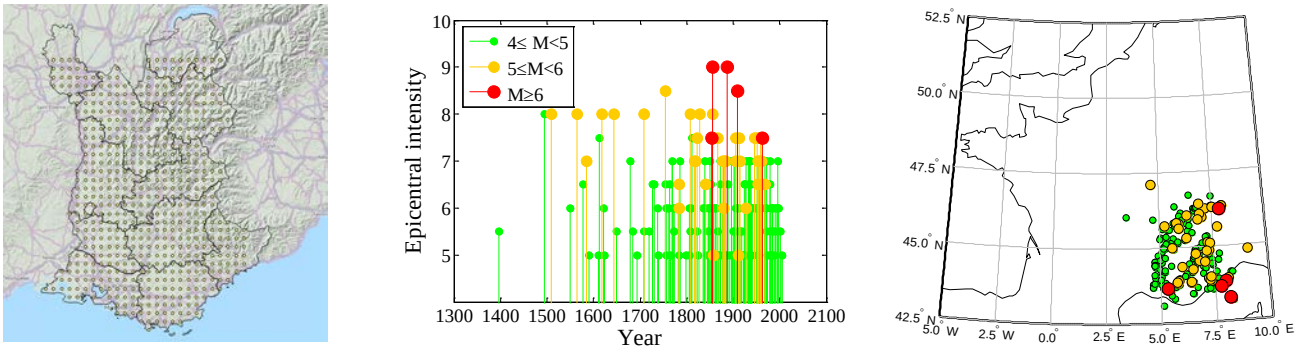

Fig. 8 – Selected sites (left), time history (centre) and epicenter location of the selected earthquakes (right)

Random fields were developed for PGA on rock sites, consistently with the hazard study to be tested. Modeling was carried out in terms of normalized residuals of PGA calculated with respect to the selected GMPE

model. To calculate the normalized residuals of PGA for PGA derived from the macroseismic intensity values, the knowledge of the soil type or shear wave velocity at the intensity locations is necessary. For this purpose, site conditions were evaluated based on the $V_{s,30}$ map produced by USGS [32]. No finite fault models or fault type mechanism were available for the calculations. The Akkar et al. [29] model, developed for epicentral source-to-site distances and a default fault mechanism (strike-slip), was hence adopted in the calculations. For all the selected earthquakes, the intensity points located at a distance smaller than 100 km from the grid points, with intensity at least equal to 4, were used as a constraint in the random fields modeling. The consideration of any intensity point located further than 100 km from the sites for which the simulation was carried out would have no effect on the results of the simulation since, for the model used, the effects of spatial correlation of PGA from site to site decrease quickly with increasing inter-site distance, becoming almost irrelevant already beyond 30 km. For each considered seismic event, the simulated PGA random fields allowed to obtain a lognormal distribution of PGAs at the considered sites, compatible with the characteristics of the event and conditioned on the available macroseismic observations.

The comparison at the regional scale level was carried out in terms of annual rates of exceedance of preselected PGA thresholds, in at least one of the selected sites, similarly to Section 4.2. With respect to Section 4.2, there are however some differences in the procedure for calculating the empirically-derived rates of exceedance. First, for each site, an equivalent seismic history in terms of synthetic PGA values was obtained, based on the results of the random fields approach. In particular, a lognormal distribution of PGA values was obtained for each event hitting the site and a single PGA value was sampled from this distribution. For each site, the observation period was calculated, as the difference between the year 2007 (i.e. the last year for which the French historical catalogue provides information) and the year of the first event hitting the site of interest. Finally, for each site, the mean annual rate of exceedance was computed for twenty PGA levels for which PSH estimates are available [14] and ranging from 0.001g to 1.02g. Fig.9 (left) compares the empirically-derived annual rates of exceedance of PGA thresholds in at least one of the selected sites with the expected ones. In the figure, red stars correspond to observations, whilst black circles correspond to PSH estimates. Also thanks to the logarithmic scale adopted in the plot, the comparison seems to provide very close results in the entire PGA range. To better explore the consistency of the obtained rates of exceedance, the ratio of the PSH over the empirically-derived rates of exceedance is plotted (Fig.9, right).

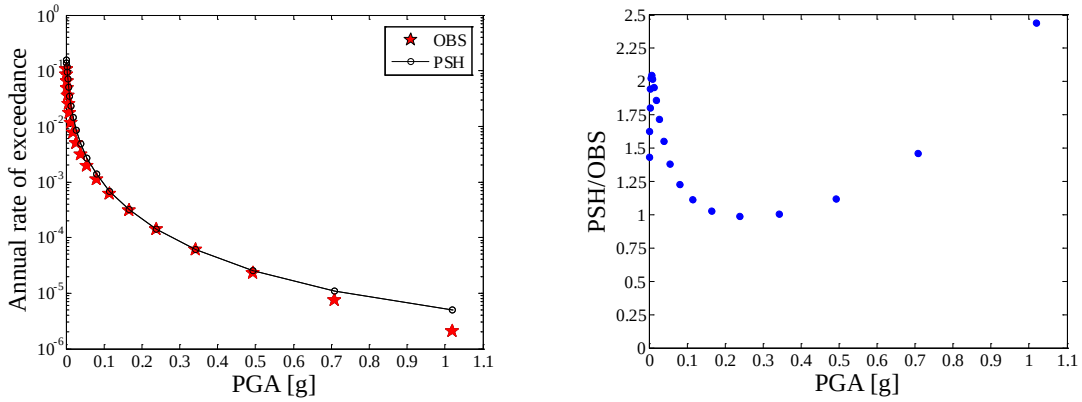

Fig. 9 – Comparison of the empirically-derived and PSH-derived annual rates of exceedance in at least one of the selected sites (left) and ratio of PSH-derived over empirically-derived annual rates of exceedance in at least one of the selected sites (right).

In almost all cases the ratio is higher than 1, suggesting the tendency of PSH estimates of overestimating empirical results. In the 0.1-0.5g PGA range, the ratio is approximately equal to 1, indicating a good agreement between predictions and observations. The overestimation at lower PGAs may be explained by the exclusion of seismic events with lower epicentral intensity and magnitude values from the list of events considered for the generation of random fields. It is however believed that these events would not significantly affect the results. Nevertheless, it can be noted that, selecting a value of the probability of exceedance (i.e. following the typical approach used for design), the overestimation in the corresponding acceleration is not so significant. Although a

comparison at the regional scale has the clear advantage of increasing the size of the available dataset, by exploiting all the information available within a study area, such a comparison can only test the average consistency of PSH predictions with observations.

5. Conclusions

This work presented different approaches for comparing PSH results with historical macroseismic observations. The expected mean damage on the building stock was selected as measure of the comparisons. Procedures for converting macroseismic observations and PGA thresholds, for which PSH estimates are available, in mean damage values were first outlined, to allow comparisons of PSH results and historical observations at different scales. Despite their immediacy, site-specific comparisons are affected by the seismic history of sites. In this sense, sites with sparse macroseismic data and/or low entity observations do not to allow pertinent comparisons. In these cases, the aggregation of multiple sites seems to be more appropriate as it allows to compensate for scanty macroseismic data at single sites and to enlarge the available macroseismic dataset. Results obtained from the application of the methodology to a set of sites suggested the development of a procedure for the comparison at the regional scale. To this aim, spatially correlated random fields of PGA, constrained to the available macroseismic observations, were generated. Results showed the tendency of PSH estimates to overestimate empirical results at low PGA values, whereas a good agreement between PSH outputs and observations was observed in the PGA range 0.1-0.5g. Despite exploiting all the information available within a given area, a regional comparison can only test the average consistency of PSH predictions with observations and does not allow to provide direct information on site-specific hazard. For the purpose of the comparison it is hence suggested to consider areas of rather homogeneous seismicity.

6. Acknowledgements

This work, carried out in the framework of the Sigma project, was financially supported by Areva. The authors would like to acknowledge Mr. J.M. Thiry (Areva), Dr. G. Senfaute (EDF), Dr. Ch. Martin (Geoter) and Prof. P. Bazzurro (IUSS-Pavia) for useful discussions. Special thanks are also due to Prof. E. Faccioli, Dr. A. Gurpinar, Dr. G. Woo and Dr. J. Savy for useful revisions of previous versions of this work. The authors would also like to acknowledge Dr. E. Fiorini (EUCENTRE) for her contribution in the generation of PGA random fields and Prof. I. Iervolino for valuable suggestions on a specific part of this work.

7. References

[1] Labbé PB, (2010): PSHA outputs versus historical seismicity. Example of France. *Proceedings of the 14th European Conference on Earthquake Engineering*, Ohrid, Macedonia.

[2] Stein S, Geller R, Liu M (2011): Bad assumptions or bad luck: why earthquake hazard maps need objective testing. *Seismological Research Letters*, **82**(5), 623-626.

[3] Stirling MW (2012): Earthquake hazard maps and objective testing: the hazard mapper's point of view. *Seismological Research Letters*, **83**(2), 231-232.

[4] Iervolino I (2013): Probabilities and fallacies: why hazard maps cannot be validated by individual earthquakes. *Earthquake Spectra*, **29**(3), 1125-1136.

[5] Beauval C (2011): On the use of observations for constraining probabilistic seismic estimates – brief review of existing methods. *International Conference on Earthquake Engineering*, Zurich, Switzerland.

[6] Ordaz M, Reyes C (1999): Earthquake hazard in Mexico City: observations versus computations. *Bulletin of the Seismological Society of America*, **89**(5), 1379-1383.

[7] Albarello D, D'Amico V (2008): Testing probabilistic seismic hazard estimates by comparison with observations: an example in Italy. *Geophysical Journal International*, **175**, 1088-1094.

[8] Stirling M, Gerstenberger M (2010): Ground motion-based testing of seismic hazard models in New Zealand. *Bulletin of the Seismological Society of America*, **100**(4), 1407-1414.

[9] Tasan H, Beauval C, Helmstetter A, Sandikkaya A, Guéguen, P (2014): Testing probabilistic seismic hazard estimates against accelerometric data in two countries: France and Turkey. *Geophysical Journal International*, **198**(3), 1554-1571.

[10] Stirling M, Petersen M (2006): Comparison of the historical record of earthquake hazard with seismic-hazard models for New Zealand and the Continental United States. *Bulletin of the Seismological Society of America*, **96**(6), 1978-1994.

[11] Purvance M.D, Brune JN, Abrahamson NA, Anderson JG (2008): Consistency of precariously balanced rocks with probabilistic seismic hazard estimates in Southern California. *Bulletin of the Seismological Society of America*, **98**(6), 2629-2640.

[12] Baker JW, Abrahamson NA, Whitney JW, Board MP, Hanks TC (2013): Use of fragile geologic structures as indicators of unexceeded ground motions and direct constraints on probabilistic seismic hazard analysis. *Bulletin of the Seismological Society of America*, **103**(3), 1898-1911.

[13] Ward SN (1995): Area-based tests of long-term seismic hazard predictions. *Bulletin of the Seismological Society of America*, **85**(5), 1285-1298.

[14] Carbon D, Drouet S, Gomes C, Leon A, Martin C, Secanell R (2012): Probabilistic analysis for France's southeast ¼ to produce a "classical" hazard map. *Deliverable SIGMA-2012-D4-24, Final Report*.

[15] Sisfrance – Catalogue des séismes français métropolitains, BRGM, EDF, IRSN, www.sisfrance.net.

[16] Lagomarsino S, Giovinazzi S (2006): Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings. *Bulletin of Earthquake Engineering*, **4**, 415-443.

[17] Grünthal G (1998): European Macroseismic Scale 1998 (EMS 1998). *Council of Europe, Cahiers du Centre Européen de Géodynamique et de Sismologie*, **15**.

[18] Dubois D, Prade H (1980): *Fuzzy Sets and Systems*. Academic Press, New York.

[19] Medvedev S, Sponheuer W, Karnik V (1964): Neue seismische Skala Intensity scale of earthquakes, 7. Tagung der Europäischen Seismologischen Kommission vom 24.9. bis 30.9.1962. In: Jena, Veröff. Institut für Bodendynamik und Erdbebenforschung. *Deutsche Akademie der Wissenschaften*, **77**, 69-76.

[20] Musson RMW, Grünthal G, Stucchi M (2010): The comparison of macroseismic intensity scales. *Journal of Seismology*, **14**(2), 413-428.

[21] Rota M, Penna A, Strobbia CL (2008): Processing Italian damage data to derive typological fragility curves. *Soil Dynamics and Earthquake Engineering*, **28**(10-11), 933-947.

[22] Rosti A, Rota M, Penna A, Magenes G (2015): Analisi statistica dei dati di danno raccolti in seguito ai principali terremoti italiani (1980-2009). *XVI Convegno ANIDIS*, L'Aquila, Italy. (In Italian).

[23] Braga F, Dolce M, Liberatore D (1982): A statistical study on damaged buildings and an ensuing review of the MSK76 scale. *Proceedings of the 7th European Conference on Earthquake Engineering*, Athens, Greece.

[24] Iervolino I, Giorgio M (2015): The effect of dependence of observations on hazard validation studies. *CSNI Workshop on Testing PSHA Results and Benefit of Bayesian Techniques for Seismic Hazard Assessment*, Pavia, Italy.

[25] Rota M, Penna A, Magenes G (2014): A framework for the seismic assessment of existing masonry buildings accounting for different sources of uncertainty. *Earthquake Engineering and Structural Dynamics*, **43**(7), 1045-1066.

[26] Park J, Bazzurro P, Baker JW (2007): Modeling spatial correlation of ground motion intensity measures for regional seismic hazard and portfolio loss estimation. *Proceeding of the 10th International Conference on Application of Statistic and Probability in Civil Engineering*, Tokyo, Japan.

[27] Ebel JE, Wald DJ (2003): Bayesian estimations of Peak Ground Acceleration and 5% Damped Spectral Acceleration from Modified Mercalli Intensity Data. *Earthquake Spectra*, **19**(3), 511-529.

[28] Stucchi M, Rovida A, Gomez Caprera AA, Alexandre P, Camelbeeck T, Demircioglu MB, Gasperini P, Kouskouna V, Musson RMW, Radulian M, Sesetyan K, Vilanova S, Beaumont D, Bungum H, Fäh D, Lenhardt W, Makropoulos K, Martinez Solares JM, Scotti O, Živčić M, Albini P, Batllo J, Papaioannou C, Tatevossian R, Locati M, Meletti C, Viganò D, Giardini D (2013): The SHARE European Earthquake Catalogue (SHEEC) 1000 - 1899. *Journal of Seismology*, **17**(2), 523-544.

[29] Akkar S, Sandikkaya MA, Bommer JJ (2014): Empirical ground-motion models for point- and extended- source crustal earthquake scenarios in Europe and the Middle East. *Bulletin of Earthquake Engineering*, **12**(1), 359-387.

[30] Jayaram N, Baker JW (2009): Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering and Structural Dynamics*, **38**(15), 1687-1708.

[31] Kramer SL (1996): *Geotechnical Earthquake Engineering*. Prentice Hall.

[32] Wald DJ, Allen TI (2007): Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, **97**(5), 1379-1395.