# On the use of observations for constraining probabilistic seismic hazard estimates - brief review of existing methods

C. Beauval
*ISTerre/LGIT, IRD-UJF-CNRS, Grenoble, France*

ABSTRACT: As probabilistic methods are now the standard for estimating long-term seismic hazard, it is tempting and legitimate to use as much as possible past observations to validate (or cast doubt) on PSHA estimations. Dealing with the output of a Probabilistic Seismic Hazard (PSH) calculation, these observations can be of at least three types: strong-motion recordings, intensity assignments and fragile geological structures. Probabilistic seismic hazard estimates by definition correspond to rare phenomena. For example, the ground-motion with return period 475 years, currently still in use for conventional building, has a 10% probability of being exceeded at least once in a 50 yrs-time window. In other words, this ground-motion has a mean occurrence rate of 1 every 475 years. In nuclear safety, the design ground-motions of reference correspond to much longer return periods. Here, the aim is to encourage the use of ground motions, intensities, or other observations, to derive methods to compare observed occurrences with probabilistic estimations. We discuss advantages and shortcomings of existing testing techniques, and encourage future authors to perform such studies with full transparency on the hypotheses underlying any of the testing methods. Otherwise, results might be over-interpreted, and in the case of validating or rejecting probabilistic hazard maps to be used in building regulation, consequences can be tremendous.

## 1 INTRODUCTION

Probabilistic seismic hazard is aimed at estimating accelerations corresponding to rather long return periods (500 to 10,000 years), and the estimation relies on a suite of models for earthquake and ground-motion occurrence. One legitimate question is whether the observations available in the region under study can be used to infer some constraints on the hazard curves obtained through PSHA studies. If many models are required in the course of calculating probabilistic hazard, it is because the time windows available are not long enough to use only recorded data to calculate hazard curves. Thus it is expected that this wish to constrain hazard curves with independent regional data might be difficult to achieve. However, some strong reasons sustain the idea:

- if using independent data from the data used in the process of developing hazard curves, the comparison is worthwhile. By independent data, we mean data that has not been used directly in the PSHA process. Such data might be varied: accelerometric records, macroseismic information, or precariously balanced landforms (e.g., Brune et al. 2004).

- statistical testing methods can be introduced to compensate short time windows by sampling in space: several methods mentioned in this paper are relying on this assumption.

Before more developments, one must bear in mind that there will be no unique method for “validating” or “invalidating” hazard curves, as time windows available, whatever the observable considered, are too short with respect to the return periods of interest in engineering seismology. However, we believe that if several methods based on very different hypotheses lead to the same conclusions, some consensus can be obtained.

The purpose of the present paper is to perform a brief overview of methods proposed up to now to take advantage of observations to try to constrain probabilistic hazard curves. Our aim is not to discuss the results of these comparisons and to conclude on an overall over- or under-estimation of the probabilistic hazard methodology, but rather to briefly discuss the methods themselves, their advantages and shortcomings. We plan to apply them and propose further developments for the French territory. We would like to stress that the aim here is the testing of the final output of a probabilistic hazard study, ie. hazard curves. We won’t address the methods implemented for testing strong-motion data against published GMPE (Ground-Motion Prediction Equations). We won’t address the short-term earthquake testing performed in the framework of the RELM initiative (e.g. Schorlemmer et al. 2007), but the statistical techniques developed for testing the methods aimed at predicting earthquake occurrence can for sure be useful and adapted for testing ground-motion occurrences at given return periods (see for ex. Albarello and D’Amico 2008).

## 2 PROBABILISTIC SEISMIC HAZARD AND POISSON PROCESSES

In a Probabilistic seismic hazard study, the aim is to determine probabilities of exceedance of given ground-motion levels over time windows of interest (equivalent to mean building life expectancies). Due to the Poisson hypothesis, estimating these probabilities of exceedance is equivalent to estimating the return periods (or annual rates) of these ground-motion levels. The main output of a PSHA study at a site is a hazard curve, a graph displaying probabilities of exceedance (or alternatively rates) versus ground-motion levels.

Obtaining a hazard curve through the classical PSHA method (Cornell 1968, Esteva 1968, Bommer and Abrahamson 2006) requires accepting the validity of various models. If willing to build a hazard curve directly from observations, one would need an accelerometric station installed for hundreds of years providing the complete history of recording during its lifetime. Then, different methods may lead to the empirical hazard curve. Let's assume here that we are interested in the probabilities of exceedance over 50 years. Assuming the Poisson hypothesis, one can count the ground-motion occurrences and derive exceedance rates. If one doesn't make the Poisson hypothesis, then another way is to obtain the distribution of the maximum acceleration observed in 50 years; the probability of non-exceedance of a given acceleration in 50 years is simply the associated percentile in this probability distribution (Beauval et al. 2006). In both methods, the high annual rates or the high percentiles will be better constrained than the low rates or low percentiles, and these constraints will depend directly on the observation windows available. Beauval et al. (2008) showed that for a Poisson process characterized by a 475-year return period at a site, a time window of 12,000 years was required to estimate the annual occurrence rate with 20% uncertainty. The methods addressed below making use of intensities or accelerations are variants of these simple counting techniques. One remark more, as observation time windows are always too short, if it is not trivial to test the hazard curve under the Poisson hypothesis, it would be even more difficult to test more realistic time-dependent occurrence models.

## 3 TESTING TECHNIQUES BASED ON ACCELERATION RECORDS

Most PSHA studies are for amplitudes of ground-motions (spectral accelerations). The most direct way to compare hazard curves with observations is to compare them to "empirical" hazard curves built from simple statistics on recordings. Accelerations recorded at sites in the region under study might have been used in the course of the PSHA analysis, if, for selecting ground-motion prediction equations, their fit with available data have been quantified. However we can still consider the acceleration recorded at a site as approximately independent from the PSHA analysis at this site. To our knowledge, the first authors to superimpose a PSHA hazard curve with an empirical hazard curve are Ordaz and Reyes (1999) in Mexico City. They take advantage of a seismological station that has been installed in the UNAM University since 1962, recording both crustal and subduction earthquakes. The hazard curve resulting of a probabilistic calculation is fitting quite well the observed annual rate calculated over ~40 years. As low annual rates calculated over short time periods can fluctuate largely (see e.g. Beauval et al. 2008), it is reasonable to state that the test at this station resulted positive partly by chance. The ground-motion prediction equations used were site-specific and had been derived from the same acceleration data, however this fact does not explain the quite good fit. Beauval et al. (2008) looked for other stations in the world with a long time life; they showed that if willing to constrain the hazard curve at a site with strictly only the data recorded at this site, it was impossible to say anything about return periods of interest for engineering seismology. They concluded that statistical methods should be proposed taking advantage of multi-sampling in space.

In Japan, accelerometric stations are densely covering the country over a grid of 25 km x 25 km on average. Intensities are estimated from seismic-intensity meters (JMA intensities). Fujiwara et al. (2009) proposed to compare observed occurrences of accelerations over 10 years (1997-2007) with the predictions of the national PSHA maps (time-dependent, based on earthquake data until 1997). For 3 levels of intensity, they determined the ratio of the stations that had recorded at least once an acceleration higher than the target level in the 10 years-period divided by the total number of stations (1028). The ratio they obtain is reasonably close to the mean probability of having an exceedance of the same intensity level in the next 10 years, averaged over the grid sites of the national PSHA map located in the same urban zones as the accelerometric stations. They do not provide uncertainty bars, which would be useful to quantify this fit. Moreover, they do not highlight the hypothesis behind such comparison: the occurrence of ground motions larger than a given threshold, at each of the accelerometric sites, during the 10 years time period is considered as realizations of the same stochastic variable. Theoretically acceleration occurrence should be independent at each site, which obviously is not the case as stations can be 25 km apart. However, it is the first time PSHA is compared with acceleration records in Japan, this effort must be encouraged and more methods must be developed to compare in a more rigorous way forecasts and observations.

One technique to compensate for short observation time windows is thus to use multiple sites and assume that the occurrence of accelerations at the different sites is a unique stochastic model (equivalent to the ergodic assumption performed while applying a ground-motion prediction equation at a site). An accelerometric network is therefore required. In most parts of the world, such networks have been developed since the beginning of the 1990s. For example, in France the accelerometric network started in the mid-1990s (Péquegnat et al. 2008). Albarello and D'Amico (2008) have proposed to statistically test PSHA predictions over a 30-year period against the observations of acceleration exceedance during a 30-year time window at 68 sites of the Italian accelerometric network (stiff rock stations). More exceedances were observed than forecast by the PSH estimations, which would mean that the PSHA estimates are underestimating the hazard. As stressed by the authors, the occurrences of ground-motions at the 68 sites are supposed to belong to the same stochastic process, implying that these sites should be independent. Acceleration occurrences might not be fully independent, as some seismic sources are impacting several sites at the same time. However such tests provide real insights into the match or mismatch between predictions and observations. Another example of testing hazard curves based on strong-motion recordings can be found in Stirling and Gerstenberger (2010), who applied in New Zealand a method comparable to Albarello and D'Amico (2008). We believe that these testing techniques deserve further developments that we plan to lead on the French dataset (e.g. understanding through synthetic catalogs what's the minimum number of sites required for a meaningful test).

## 4 TESTING TECHNIQUES BASED ON INTENSITIES

There are much more attempts on intensities than on accelerations to derive empirical hazard curve and compare them to the classical PSH methods. Here we will mention three of the latest papers on the topic. Probabilistic hazard is estimated mostly in terms of ground-motions, as estimating hazard in terms of intensity is impeded by the lack of constrained uncertainty (sigma) in intensity predictive equations. Therefore, a correlation between accelerations and intensities is compulsory. Such correlations are rather well constrained for few regions in the world (e.g. Atkinson and Kaka, 2007), but the choice of one equation will represent a high epistemic uncertainty for low-to-moderate regions of the world (e.g. France).

Intensities represent the unique observable of past earthquake effects at locations which are not instrumented. Furthermore, intensities are available since the existence of written documents, and thus can be extremely useful to extend the observation time windows. Intensities are reflecting the strength of the earthquake; the intensity degree is determined according to available descriptions of the effect of the shaking on humans, objects, nature, and buildings. Obviously such quantities bear large uncertainties, and some subjectivity. Uncertainties on the exact locations of earthquakes that occurred some centuries ago can be large, however they extend observation time windows well before instrumental periods. Depending on the region of the world, the time window described by macroseismic intensities varies largely: from 500 years in South America, to 3000 years in China. Note that intensity data is not strictly independent from the PSHA estimation process, as intensities have been used for estimating magnitude and location of pre-instrumental earthquakes. However, the intensity history at one site can still reasonably be considered independent from the probabilistic hazard evaluated at that site.

Stirling and Petersen (2006) compare the predicted annual rates of exceedance for ground-motion levels from the national PSH model with the historically based annual rates of ground-motions derived from intensity data, at sites located both in New Zealand and in United States. Note that no earthquake less than the minimum magnitudes considered in the PSH models have been included in the determination of the historically based hazard curves. The discrepancies between PSH model and historical data are assessed statistically. Highest mismatch appear at intraplate sites (low-to-moderate earthquake rates). The difficulty in interpreting the results is that there isn't a unique explanation for match or mismatch, and this is very clearly discussed in Stirling and Petersen (2006) paper. Mismatch can be explained by using non-Poissonian earthquake occurrences for the fault models, or by the fact that in the PSHA study earthquakes with large return periods are taken into account, which did not occur in the historical time window. Other explanations can be that local site conditions have been badly taken into account in the PSH calculations (whereas they are naturally included in the intensity records), or that the intensity-acceleration correlations are not adequate for the region under study.

As time windows might be too short for estimating reliable recurrence rates for ground motions with long return period, some authors propose to model the intensity recurrence at a site in a similar way as the modeling of the Gutenberg-Richter recurrence curve for magnitudes (with a previous completeness analysis and the assumption of a recurrence model for intensities). Bozkurt et al. (2007) establish a regional frequency-intensity curve for the Tokyo region, using a 400 years intensity catalog and assuming an exponential decay of frequencies. Then they assume that the regional decay slope can be applied to individual cell (5 km x 5 km) covering the region, and based on these recurrence curves they propose probabilities of exceedance of given intensity level over any exposure time of interest. Results are displayed only for reliable cells, the reliability for each cell being assessed combining the goodness of fit of the model, the number of intensity observations and the time period covered. The strong hypothesis underlying this method is the assumed similarity between past and future frequency of shaking, and the acceptance that the 400 years available are representative of the seismicity of the region. The authors' aim was to include as fewer models as possible in the derivation of probabilistic hazard estimates ("we view this data-rich method as an alternative that can be compared to assumption-rich approaches.", p. 544, Bozkurt et al. 2007). To compare intensity-based probabilistic estimates with the national PSHA values, they used a correlation intensity-acceleration derived from JMA intensities. Few countries in the world have such data available, and will face high uncertainties at the step of interpreting intensities in terms of accelerations. For this reason, Mucciarelli et al. (2008) preferred to compare ranking of hazard for sites in Italy rather than absolute values. For modeling the recurrence of intensities, they used a more elaborated method developed by D'Amico and Albarello (2008).

D'Amico and Albarello (2008) propose a scheme for developing a hazard curve in terms of intensities based 1) on an analysis of the intensity catalog at the site under study, 2) on an epicentral intensity catalog, and 3) on the other individual intensities assigned to neighboring sites. Uncertainties on intensities are integrated in the probability calculation, as well as completeness issues. This approach is not free of modeling assumptions, it builds on a model for the attenuation of intensity with distance and assumes isotropic behavior of intensity attenuation, thus producing an intensity hazard curve that must be more complete than hazard curves relying on simple counting techniques.

At last, in the framework of macroseismic intensity studies, one can be interested in mapping the maximum observed intensity over the observation time window. This maximum intensity cannot be compared easily to the hazard value associated with any given return period and is of little help for validating the PSHA estimates at a given site (Beauval et al. 2010). However, if considering different sites over a given time window (e.g. 500 years), and assuming a unique stochastic process, the probability distribution for the maximum acceleration over 500 years can be built, retrieving the acceleration at e.g. 10% over 500 years (4745 yrs return period) or 50% over 500 years (721 yrs return period). This value can be compared with the corresponding mean PSH value averaged over a grid site covering the region under study.

## 5 TESTING TECHNIQUES BASED ON PRECARIOUSLY BALANCED ROCKS OR OTHER PRECARIOUS NATURAL FEATURES

The "precarious rocks" are rock structures in nature which have been in equilibrium for thousands of years and thus subjected to ground motion from numerous historical earthquakes. Based on 3D finite-difference methods taking into account the shape of the rock and the rocking points, the acceleration able to topple the structure can be evaluated (Anooshehpoor et al. 2004). This acceleration is the maximum acceleration "observed" at this site over the time life of the precarious rock. Age dating techniques and numerical models for simulation of the acceleration able to topple the structure obviously bear some uncertainties. However, these precarious rocks represent one of the very few ways of enlarging the history of ground-motion occurrences at a site and providing maximum bounds. These rocks can provide important constraints on ground motion from events for which no ground-motion recording instruments were available. Fragility curves can be estimated for the precarious rocks, and one can determine the probability that the ground-motion occurrence (results of PSHA) overturns the rock during its life span, or the time period required to overturn with high probability (Purvance et al. 2009). These comparisons are not easy to interpret as again, the sources of discrepancies might be varied; however they can demonstrate inconsistencies and encourage improvement of the PSHA calculations (improvement of input data or improvement of modeling techniques). Precarious rocks have been studied in different tectonic environments; all located in relatively high seismicity regions (near-fault sites in southwestern US, California and Nevada). It would be interesting to study the potential of these methodologies in regions characterized by low-to-moderate seismic activity. Precarious rocks can be found in dry-climate regions. Elsewhere, other precarious natural features might be worth studying, such as speleothems (Kagan et al. 2005).

## 6 CONCLUSIONS

Different methods have been proposed in the last years for comparing probabilistic predictions of ground-motion frequencies with available observations (or in other words methods for testing these predictions against observations). All of these testing methods rely on some strong hypothesis, and the origin of match or mismatch can generally have more than one explanation, making it difficult if not impossible to conclude from the results of only one testing method if the probabilistic hazard has been under-estimated or over-estimated. The use of observations to test probabilistic predictions must be encouraged, but it must be stressed that results will depend highly on the testing method itself, and that a conclusion on the "validity" of PSHA estimates can be obtained only through the application of different testing techniques, using different observables. We would like to stress that all hypotheses underlying comparisons should be clearly and transparently explained; otherwise these studies will bring even more confusion instead of providing a means to discriminate between different hazard maps.

## 7 REFERENCES

Albarello D. and V. D'Amico (2008). Testing probabilistic seismic hazard estimates by comparison with observations: an example in Italy, Geophys. J. Int., 175, 1088-1094.

Anooshehpoor A., J. Brune and Y. Zeng (2004). Methodology for Obtaining Constraints on Ground Motion from Precariously Balanced Rocks, Bulletin of the Seismological Society of America, Vol. 94, No. 1, pp. 285–303.

Atkinson G.M., and S.I. Kaka (2007). Relationships between Felt Intensity and Instrumental Ground Motion in the Central United States and California, Bulletin of the Seismological Society of America, Vol. 97, No. 2, pp. 497–510.

Beauval C., S. Hainzl and F. Scherbaum (2006). Probabilistic seismic hazard estimation in low-seismicity regions considering non-Poissonian occurrence, Geophys. J. Int., 164, 543-550.

Beauval C., PY. Bard, S. Hainzl and Ph. Guéguen (2008). Can strong motion observations be used to constrain probabilistic seismic hazard estimates?, Bull. Seism. Soc. Am., 98, 509-520.

Beauval C., P-Y. Bard and J. Douglas (2010). Comment on "Test of seismic hazard map from 500 years of recorded intensity data in Japan" by Masatoshi Miyazawa and Jim Mori, Bull. Seismol. Soc. Am., 100, 3329-3331.

Bommer J.J., Abrahamson N.A. (2006). Why do modern probabilistic seismic hazrad analyses often lead to increased hazard estimates? Bulletin of the Seismological Society of America, 96(6): 1967-1977.

Bozkurt, S.B., R.S. Stein and S. Toda (2007). Forecasting Probabilistic Seismic Shaking for Greater Tokyo from 400 Years of Intensity Observations, Earthquake Spectra, 23(3), 525-546

Brune, J.N., A. Anooshehpoor, B. Shi, and Y. Zeng (2004). Precarious Rock and Overturned Transformer Evidence for Ground Shaking in the Ms 7.7 Kern County Earthquake: An Analog for Disastrous Shaking from a Major Thrust Fault in the Los Angeles Basin, Bulletin of the Seismological Society of America , 94, 6, 1993–2003.

Cornell, C. A. (1968). Engineering seismic risk analysis, Bull. Seism. Soc. Am. 58, no. 1, 1583–1606.

D'Amico V. and D. Albarello (2008). SASHA: A Computer Program to Assess Seismic Hazard from Intensity Data, Seismological Research Letters, 79; 663-671.

Esteva, L. (1968). Regionalizacion sısmica de Mexico para fines de Ingenieŕıa, PhD thesis. School of Engineering, National Autonomous University of Mexico (UNAM).

Fujiwara H., N. Morikawa, Y. Ishikawa, T. Okumura, J. Miyakoshi, N. Nojima and Y. Fukushima (2009). Statistical comparison of national probabilistic seismic hazard maps and frequency of recorded JMA Seismic intensities from the K-NET Strong-motion observation network in Japan during 1997-2006, Seism. Research Letters, 80, 458-464.

Kagan E.J., A. Agnon, M. Bar-Matthews, and A. Ayalon (2005). Dating large infrequent earthquakes by damaged cave deposits, Geology, 33, 4, 261-264.

Mucciarelli M., D. Albarello and V. D'Amico (2008). Comparison of probabilistic seismic hazard estimates in Italy, Bull. seism. Soc. Am., 98, 2652-2664

Ordaz, M., and C. Reyes (1999). Earthquake hazard in Mexico City: observations versus computations, Bull. Seismol. Soc. Am. 89, no. 5, 1379– 1383.

Péquegnat, C., P. Guéguen, D. Hatzfeld, and M. Langlais (2008). The French Accelerometric Network (RAP) and National Data Center (RAPNDC), Seism. Res. Lett. 79, no. 1, 79–89.

Purvance M.D., J.N. Brune, N.A. Abrahamson, and J.G. Anderson (2008). Consistency of Precariously Balanced Rocks with Probabilistic Seismic Hazard Estimates in Southern California, Bulletin of the Seismological Society of America, Vol. 98, No. 6, pp. 2629–2640.

Schorlemmer, D., M.C. Gerstenberger, S. Wiemer, D.D. Jackson, and D.A. Rhoades (2007). Earthquake Likelihood Model Testing, Seismological Research Letters, 78, 1, 7-29.

Stirling, M., and M. Petersen (2006). Comparison of the historical record of earthquake hazard with seismic hazard models for New Zealand and continental Unites States, Bull. Seismol. Soc. Am. 96, 1978–1994.

Stirling M. and M. Gerstenberger (2010). Ground motion-based testing of Seismic hazard models in New Zealand, Bull. Seismol. Soc. Am., 100, 1407-1414.