On the Use of a Bayesian Updating Technique to Get Probabilistic Seismic Hazard Assessment More Rugged

E. Viallet[1], N. Humbert[1], C. Martin[2], R. Secanell[2]

[1] *EDF SEPTEN, Dept. of Dynamics & Earthquake Engineering, Lyon, France*
[2] *Geoter International, Roquevaire, France*
Email: emmanuel.viallet@edf.fr

ABSTRACT :

Since the basic work of Cornell, many studies have been conducted in order to evaluate the probabilistic seismic hazard (PSHA) of nuclear power plants. In general, results of such studies are used as inputs for seismic PSA. Such approaches are nowadays considered as well established and come more and more used worldwide, generally in addition to deterministic approaches.
Nevertheless, some discrepancies have been observed recently in some PSHA, especially from studies conducted in areas with low to moderate seismicity. The lessons learnt from these results lead to conclude that, due to uncertainties inherent to such a domain, some deterministic choices have to be taken and, depending on expert judgment, may lead to strong differences in terms of seismic motion evaluation.
In that context, the objective of this paper is to point out some difficulties that may appear in the development of PSHA studies and to propose an approach that may be used to address epistemic uncertainties. The key point, which corresponds to the innovating point of the process, is the use of instrumental experience feedback to update the results of a PSHA. The method used here is based on a Bayesian updating technique including real observations as conditional events, with their own probabilistic distribution.
The results presented here point out that a PSHA must be conducted in a real probabilistic spirit that is totally different from a deterministic approach (the choice of "best-estimate" or "median" input data instead of "conservative" ones is one of the key points). In addition, logic tree procedure, which seems to be the most appropriate way to account for epistemic uncertainties, does not quantify the variability on the physical parameter itself but quantify variability on expert opinion. This may lead to an important bias in a PSHA.
Finally, results from PSHA may be strongly different from real seismicity, as recorded, especially depending on previous considerations. Then, the comparison to the instrumental experience data appears to be necessary to address such difficulties. In that context, the use of the Bayesian updating technique presented in this paper may become a necessary tool to address epistemic uncertainties in PSHA and its performances could allow to get PSHA more rugged and consistent with observations.

KEYWORDS: PSHA, Bayesian approaches, Instrumental experience feedback, Updating technique

1. OBJECTIVES OF THE STUDY

PSHA approaches are nowadays considered as well established and come more and more used worldwide. Nevertheless, some discrepancies have been observed recently in some PSHA [KLÜ-05], especially from studies conducted in areas with low to moderate seismicity. The lessons learnt from these results lead to conclude that, due to uncertainties inherent to such a domain, some deterministic choices have to be taken and, depending on expert judgment, may lead to strong differences in terms of seismic motion evaluation.
In that context, the objective of this paper is not to describe in detail the PSHA overall approach but its objective is to point out some difficulties that may appear in the development of PSHA studies, and to propose an approach that may be used to orient expert judgment and address epistemic uncertainties. The key point, which corresponds to the innovating point of the process, is the use of instrumental experience to update the results of a PSHA using observations that remain consistent with the real regional seismicity, as recorded.
This paper is divided in two parts. The first part presents some basic considerations on the current practice in PSHA studies and identifies some difficulties that one may face at different steps of the study, especially concerning input data selection. The second part describes an approach that may be used to update the results of the PSHA, which is based on a Bayesian updating technique including real observation as conditional events, with their own probabilistic distribution.

2. PART 1 - BASIC CONSIDERATIONS ON THE CURRENT PRACTICE IN PSHA AND ASSOCIATED DIFFICULTIES

As the objective of this paper is not to describe in detail the PSHA overall approach, we only focus hereafter on some specific aspects of a PSHA that could lead to difficulties (implying deterministic choices) and we try to quantify their consequences.

2.1. Seismic Motion Characterization : Random Uncertainty ?

We would like to discuss first on the characterization of the variability of the seismic motion. In particular, two question may arise related to this factor :

- Q1 : Where does sigma (i.e. standard deviation associated with a given attenuation relationship) come from ?
- Q2 : How many sigma around the median value of the attenuation relationship to integrate in the PSHA (1, 2, infinite …) ?

<u>Elements of answer to Q1</u>

Concerning Q1, it seems to be obvious that sigma comes from the (random) variability of the seismic motion itself. Nevertheless, it is also obvious that in strong motion databases that are used to determine attenuation relationships, the characteristics (magnitude, location, depth) of the events are known with an uncertainty which may be significant. [HUM-08-1] presents a detailed assessment on that topic and clearly shows that a significant part of sigma comes from epistemic uncertainty instead of random. As this value of sigma has a direct (and important) impact on the PSHA, the rigorous separation between random and epistemic uncertainties should be done, especially in attenuation relationships parameters. The usual "deterministic way" induces a systematic bias in PSHA which can be important especially for the long term periods. In that context, it appears urgent that a real discussion on epistemic uncertainties and random variability should be introduced in PSHA.

<u>Elements of answer to Q2</u>

In a pure probabilistic spirit, the first answer should be to integrate the sigma to infinite. Nevertheless, the fact that seismic motion distribution follows a lognormal law is only an assumption and should always be checked on the basis of real data. Depending on strong motion databases, this verification may not be possible over 2 to 3 sigma (see Figure 1).

This is an important fact to point out. Even if it may be shown that over 3 sigma, the impact may not be so important, the range of integration should be carefully assessed in PSHA.

Another point we would like to discuss now concerns the equivalence of magnitudes that may be necessary to use in PSHA, due to historical and instrumental seismicity, and magnitude used in attenuation relationships.

This leads to choices that can be taken in a probabilistic spirit or in a deterministic spirit. For instance, the choice to take a "conservative" relationship between M_S and M_L(LDG) (see Figure 1), which may be an appropriate choice for a deterministic hazard evaluation for a NPP for instance, may lead to a bias in a probabilistic approach and will directly over-estimate the median value of the ground motion.

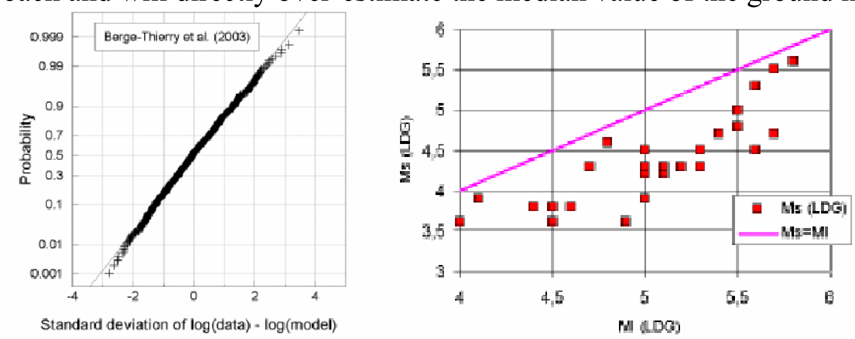

Figure 1 : Comparison between lognormal distribution, records and model (left)
Relationship between different types of magnitudes (right)

According to the PSHA methodology, best-estimate data should be kept at each stage of the process in order to keep the "median" estimation spirit of the PSHA, as expected.

Conclusion on the characterization of the seismic motion
Concerning variability of seismic motion, some choices have to be made and may have a strong impact on the results of a PSHA. The most important request is to keep "best estimate" data, to remain consistent with the PSHA philosophy. In that situation, the value of sigma associated with attenuation relationships may be carefully estimated in order to account for the random variability of the seismic motion only, its range of integration should also be carefully assessed (typically between 2 and 3 sigma), and the potential relationship use to transform different types of magnitudes should be also a best-estimate one.
This fact should be considered in a PSHA in order to get a "real" median value, as expected.

2.2. Selection Of Attenuation Relationship For A Given Area : Epistemic Uncertainty ?

As one can expect, attenuation relationships come from area with moderate to high seismicity (where data are available). However, they may be used for PSHA in area with low to moderate seismicity. In that situation especially, their applicability (even if there may not be another way to proceed) is to be seriously assessed. In addition, it may be obvious to say that very important differences may be observed from different attenuation relationships (see one illustration in Figure 2).

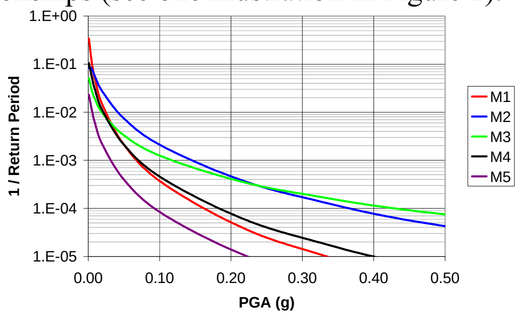

Figure 2 : Illustration of the results of a PSHA depending on attenuation relationships (M1 to M5)

In a PSHA, this aspect is accounted for based on expert judgment. The knowledge of (i) the real seismicity of the given area and (ii) the construction of attenuation relationships is then a key factor. Practically, this is usually accounted for by mean of logic trees that allow to take into consideration all possible judgments. Nevertheless, one must keep in mind that this logic tree approach allow to quantify variability on experts judgment only but not the variability on the physical parameter itself.
This epistemic uncertainty is expected to be reduced, with increase of knowledge.

Conclusion on the choice of an attenuation relationship
The choice of attenuation relationships is one of the most important choices in a PSHA and may lead to high differences in the results. In that situation, it is important to point out that the logic trees procedure which seems to be the most appropriate way to account for epistemic uncertainties does not quantify the variability on the physical parameter itself but quantify variability on expert opinion. This may lead to an important bias in a PSHA study.
The current practice should then be improved in order to reduce the epistemic uncertainty on the physical parameter itself.

2.3. Conclusion Of Part 1

Although well established, some basic steps of PSHA are still under discussion. But most of the difficulties come from the choice of input data and the way to account for uncertainties, which is different from the way to address them in a deterministic approach.

- The most important request is to keep "best estimate" data, at each stage, to remain consistent with the PSHA philosophy, and to obtain a "real" median estimation.
- It is also important to propagate variability but trying to separate random uncertainties from epistemic ones,
- Finally, it must be kept in mind that logic trees allow to quantify the variability on expert judgment but not the variability on the physical parameter itself, this should be improved.

The next part of this paper is to propose a method to address this last point.

3. PART 2 – A BAYESIAN UPDATING TECHNIQUE TO GET PSHA MORE RUGGED

3.1. Bacground : Impact of deterministic choices on PSHA and comparison with observations

The previous part has pointed out some problems that may occur in PSHA. These problems were also confirmed in a different way by using real seismic activity from a given area.
A detailed study was performed by [HUM-08-2] comparing PSHA results and actual observation. This study which includes all sources of uncertainties (epistemic and random) clearly shows that if PSHA results are to far from real seismic activity of a given territory, even a low period of observation may be sufficient to identify such inconsistency.
As an example of this, figure 3 shows the range of inconsistency of a typical (well-performed) PSHA when compared to observation (see [HUM-08-2] for details).

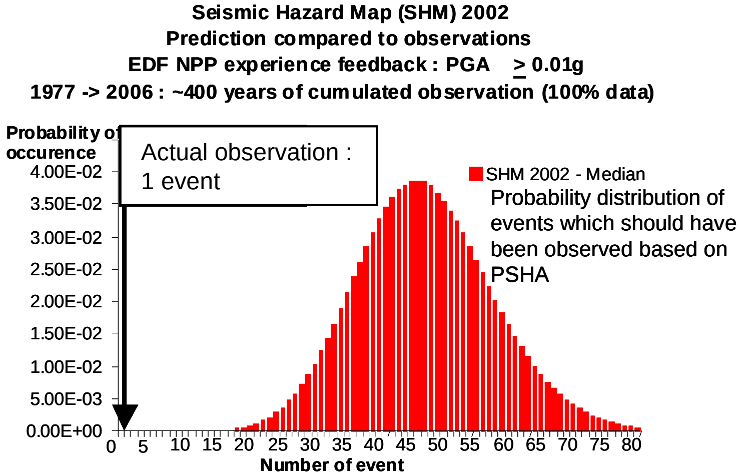

Figure 3 : Comparison of PSHA prediction (including all sources of uncertainties and random occurrence model for earthquake events for a given period of observation) to observation

Conclusion on the Impact of deterministic choices on PSHA and comparison with observations
The lesson learnt from this part is that results from PSHA may be strongly different from real seismicity, as recorded, especially depending on choices on input data that always depend on expert judgment.
Then, the use of instrumental experience data appears to be necessary to address such difficulties.

3.2. A Way To Address Uncertainties : The Use Of A Bayesian Updating Technique

Based on previous results, the objective here is to use records to update a PSHA in order to get results more consistent with observation (which can be trust as reality).

Principle of the method

The method used is based on a classical Bayesian updating technique, used for many years in reliability field, especially in mechanics.
The basic consideration is simple :

"Based on a probabilistic evaluation of a given parameter, what is the most probable values that this parameter could take (associated with a given confidence level) considering the available observations (including their uncertainties) as conditional events ?"

The technique used here is based on the one initially developed by Madsen [MAD-85] and its performance and pertinence have been already confirmed [HEI-99].
It uses the Bayesian theorem of conditional probability :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Where :

- A is the predicted value of the parameter (random variable) of interest, according to a probabilistic model,
- B is the observed value of the parameter (random variable) of interest.

Implementation in the present study

In our application here (PSHA), the Bayesian theorem is applied as follows :

- P(A) is the probability of A (predicted seismic event) according to the initial prediction
  P(A) can be obtained for each branch "i" of the logic tree (for PSHA) : P(Ai)
  -> This is usually the weight affected to the branch under consideration (i.e. equal weight w or weight given by experts w_i)
- P(B|A) is the conditional probability of the observed event B according to the predicted one A
  P(B|A) can be calculated for each branch "i" of the logic tree : P(B|Ai)
  -> This step is based on an occurrence model (the process will be explained later)
- P(B) is the total probability of the observed event B
  P(B) can be calculated based on the total probability theorem : P(B) = Σj P(Aj) . P(B|Aj)
- P(Ai|B) is the updated probability of A considering B
  P(Ai|B) is therefore the updated weight of each branch of the logic tree (updated based on observation)

The key point of the process is then to determine the so-called "likelihood function" which is defined by the following expression :

P(A|B)/ P(B)

Parameter of interest considered in this study

To apply Bayesian theorem, a parameter of interest has to be defined. This parameter is to be determined in a way that it could be easily predicted based on a given PSHA result and determined based on actual observations for a given territory.

For this study, this parameter is defined as a number of observed events with a PGA higher than a given value. This parameter can be cumulated among a certain number of accelerometric stations distributed among a given area.

The characteristics selected in our case are those defined in [HUM-08-2] :

- Territory under consideration : French metropolitan territory,
- Station of observation : 20 RAP [RAP] and 19 EDF NPP accelerometric stations (39 in total),
- Seismic event selected : Total number of observed events with PGA > 0.01 g
- Parameter used for Baysian theorem application : Cumulated number of observed events with PGA > 0.01 g among the whole set of stations, accounting for potential correlation between stations.

See [HUM-08-2] for more detail.

Determination of the likelihood function P(A|B)/ P(B)

This likelihood function is calculated based on the Poisson's occurrence model, as stated in many PSHA. This model is used to calculate the probability of occurrence of a given seismic event considering a given time of observation.

Finally, this process allows to account for all sources of uncertainties included in seismic motion occurrence:

- Random and epistemic uncertainties in seismic motion prediction, as accounted for in the PSHA model,
- Random occurrence of seismic events in a given territory and duration of observation by the mean of Poisson's model.

In addition, the following additional sources of bias or uncertainties are accounted for :

- Site effect : rock condition for PSHA and soil conditions for accelerometric stations in some cases,
- SSI effect : free field for PSHA and accelerometer located on a building foundation in some cases,
- Potential correlation between observation sites (1 seismic event may be observed by 2 stations).

The procedure used here is described in detail in [HUM-08-2]. This allows to avoid any bias in the process and to account for all sources of uncertainties.

3.3. Application

Initial prediction

The initial prediction is a full PSHA study, as describe below.

- Seismotectonic models
  Based on previous experience, 2 different seismotectonic models were used, as shown in figure 4 :

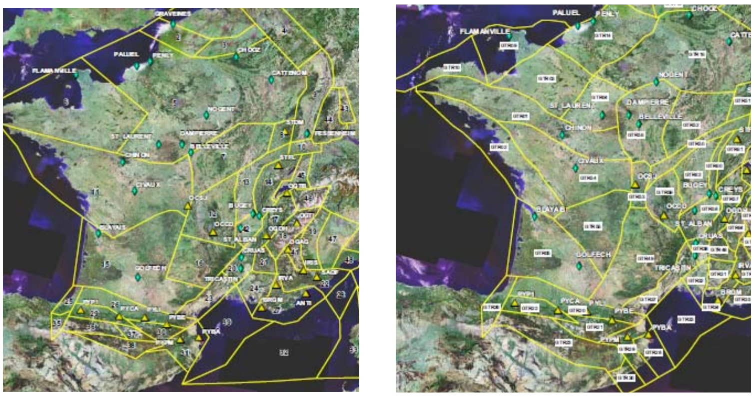

Figure 4 : Seismotectonic models used for the PSHA

The first model was developed in a previous study [MAR-02]. The second one was specifically developed for the present study.

The source model used is based on diffused seismicity (as usually assumed for French metropolitan territory), including up to date catalogs for French territory.

<u>Strong motion attenuation relationships</u>

An extensive work was done by [HUM-08-1] in order to quantify and take into account all the sources of uncertainties (random and epistemic ones) included in strong motion databases used to build attenuation relationships.

This work, performed on a European Strong Motion Data Base [ESMDB] clearly shows that the parameters of the attenuation relationship are not perfectly fitted by usual regressions and include epistemic uncertainty. A "deterministic" regression underestimates this epistemic uncertainties. Consequently, the σ parameter of a "determinist" regression overestimates the random variability by confusing a part of epistemic uncertainty in the random part. This the reason why an attenuation relationships built in a deterministic way induces a systematic bias in PSHA.

For this study, the attenuation relationships built by [HUM-08-1] on the basis of fuzzy data regression was used. The generic equation has the following form :

$$\log(PGA) = aM + a_2M^2 + bR + (b_1 + b_2M)\log\left(\sqrt{R^2 + b_3{}^2}\right) + c \pm \sigma$$

One of the innovating part of the process is that all the parameters (a, b, c, σ …) are characterized by a median value and an epistemic uncertainty, accounted for by mean of a standard deviation around the median value.

Finally, different attenuation relationship were used in a logic tree process, as shown in fig. 5.

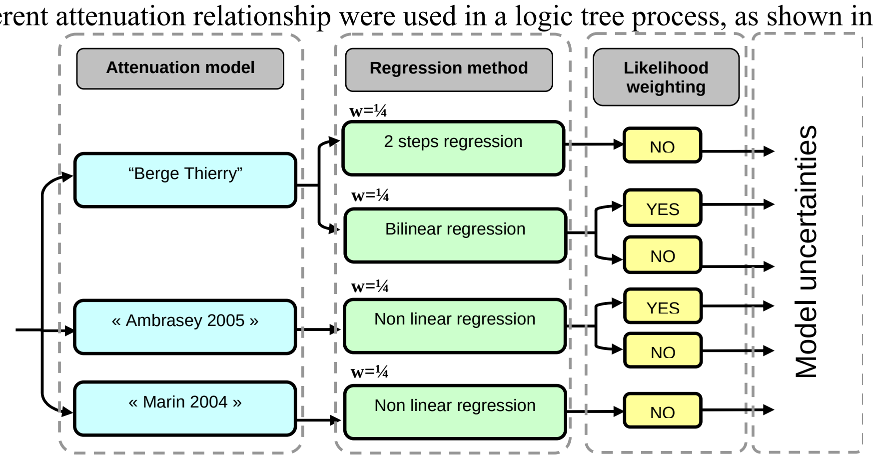

Figure 5 : Attenuation relationship

<u>Integration of the random variability of the seismic motion</u>

Due to the previous process which allow to separate epistemic and random uncertainties, the random part of the seismic motion (σ value of the attenuation relationships) is integrated to infinite in the PSHA.

Full logic tree

The logic tree finally defined for this study is composed of 50 epistemic branches, as shown in figure 6.

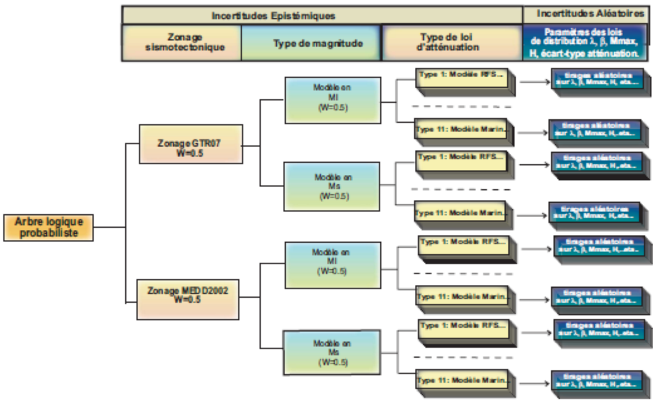

Figure 6 : Logic tree defined for the PSHA

Weighting process

Due to the way the PSHA was defined and considering the purpose of the study, which is the application of Bayesian theorem, no expert weighting was applied. Consequently, each branch of the initial prediction was equally weighted.

3.4. Results of the initial prediction

The results of the initial prediction are obtained for different points of the French metropolitan territory and expressed for different return period. The resulting hazard map for 100 and 500 return period are shown in figure 7.

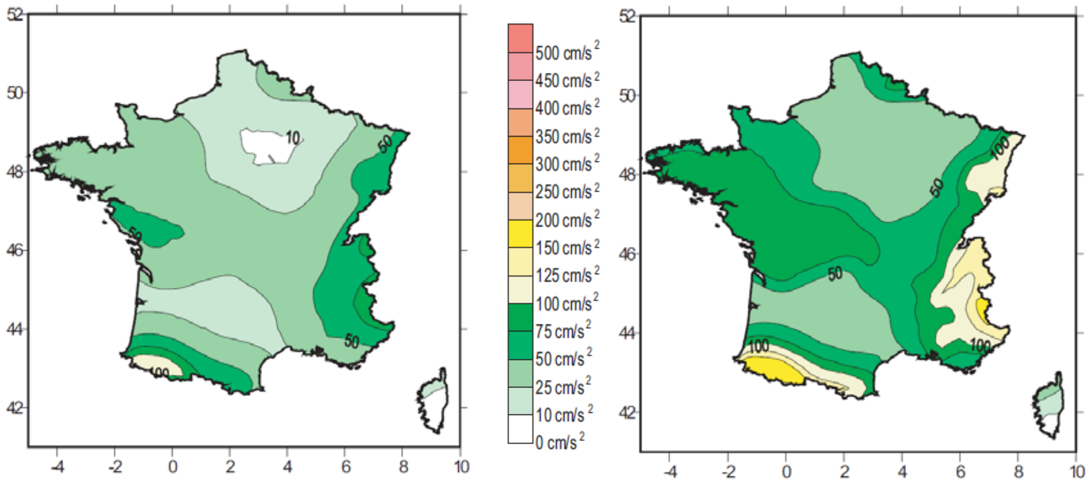

Figure 7 : Initial prediction 100 year (left) and 500 year (right) return period hazard maps (median)

3.5. Observations used for updating

The updating process is performed as described in paragraph 3.2. The location of the accelerometric stations is presented in figure 8.

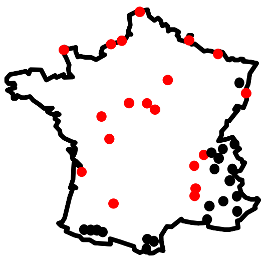

Figure 8 : Location of accelerometric stations (EDF NPP in red, RAP in black) for the updating process

This set of accelerometric stations lead to approximately 500 years of cumulated observation.

3.6. Results of the updating

The results of the updating process are obtained in the same way than for the initial prediction. The resulting hazard map for 100 and 500 return period are shown in figure 9.

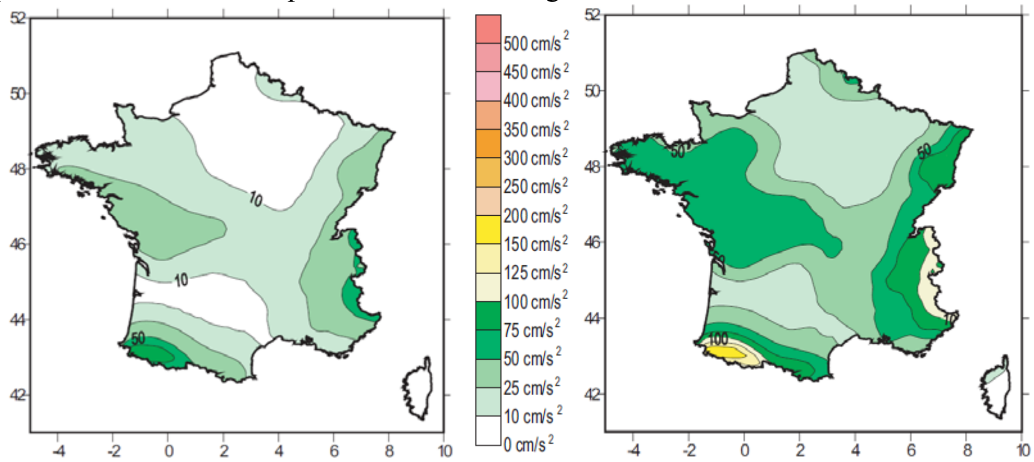

Figure 9 : Updated hazard map for 100 year (left) and 500 year (right) return period (median)

3.7. Comparison of the initial prediction and the updated one

The comparison between initial prediction and updated one in term of hazard map is shown in figure 10.

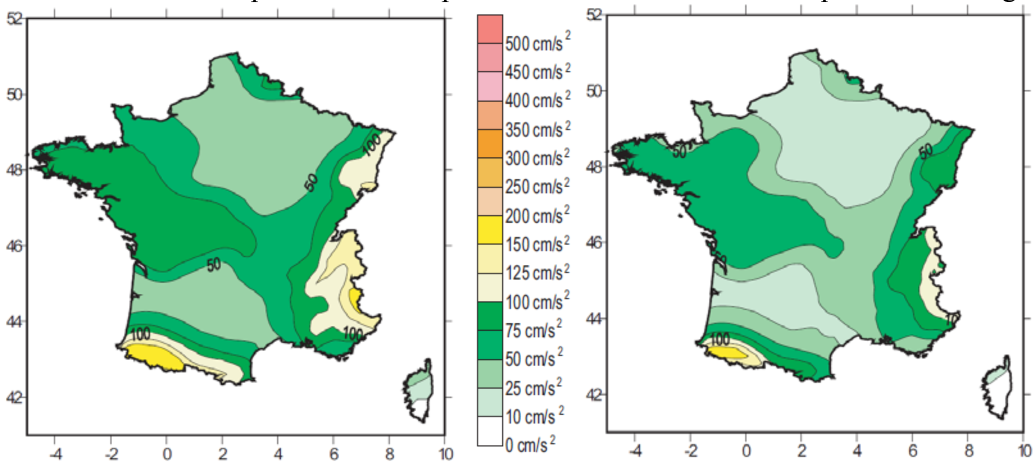

Figure 10 : Comparison between initial prediction (left) and updated one (right) in term of 500 year return hazard map (median)

The comparison between initial prediction and updated one in term of hazard curve is shown in figure 11.

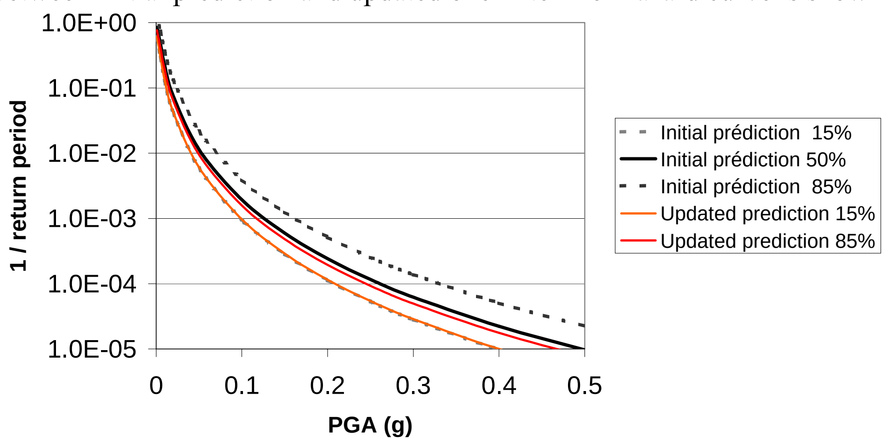

Figure 11 : Comparison between initial prediction and updated one in term of hazard curve

One can observe the potential effectiveness of the updating process which reduces significantly the scattering of the initial prediction and also “re-center” the prediction around fractiles which are not the median value of the initial prediction. These results are discussed in the next paragraph.

3.8. Discussion Around The Updating Technique

First of all, it is important to indicate that the updating technique does not modify any of the input data or assumptions of the initial prediction. The technique only accounts for observations in order to identify the most likely branches among the initial ones.
Then, one can notice that the techniques accounts for random uncertainties as far as these random uncertainties are included in the observations (this is the case by considering different stations in different areas and different time of observation). It also account for the random occurrence of earthquakes by Poisson’s occurrence model.
This is one of the reasons that lead to an updated prediction with still variability (in other mechanical studies dealing with other parameters with a low random uncertainty, such as delayed strain in nuclear reinforced concrete containment for instance [HEI-05], the updated technique leads to results with a low variability).
Anyway, some questions may arise as for instance :

- How does the updating process behave depending on the amount of observation used ?
- Is the updating process sensitive to random occurrence of earthquakes ?

In order to get an idea of the sensitivity of the method to some of its parameters, 2 tests were conducted.

- The first one uses a reduced set of observation only, in term of cumulated period of observation
- The second assume that 1 or 2 additional event would have been observed in the same period of observation.

The results are presented in figure 11.

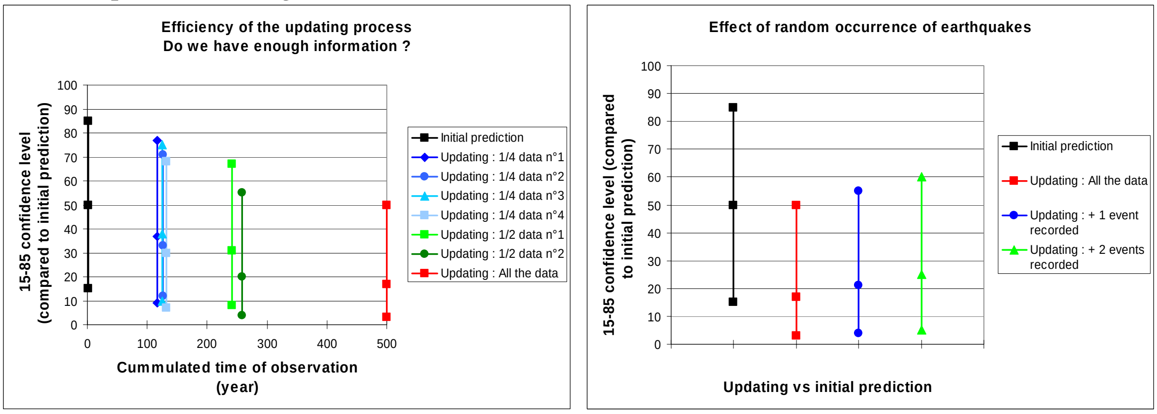

Figure 11 : Quantification of the impact of a reduced period of observation (left) and 1 or 2 additional event observed in the same period of observation (right)

These results clearly show that the method is robust. Concerning the period of observation, the trend is homogeneous and the precision is getting better when the period of observation is increasing. In addition, in the case when the period of observation is to low, the updating becomes not effective. Then the process reveals its limits by its own. Finally, the process is not so sensitive to random occurrence of seismic events, as observed in our case.

Conclusion on the use of the Bayesian updating technique
Our conclusion concerning the use of the Bayesian updating technique is that it is a real interesting tool to address epistemic uncertainties in PSHA and its performances could allow to get PSHA more rugged and consistent with observation.
Then, updating technique may become one necessary step in PSHA methodologies.

6. CONCLUSIONS

The objective of this study was to point out some difficulties in PSHA studies and to propose an approach that may be used to orient expert judgment and address epistemic uncertainties.

The most important conclusions that we would like to point out are the following :

Although well established, some basic steps of PSHA are still under discussion. Most of the difficulties come from the choice of input data and the way to account for uncertainties, which is different from the way to address them in a deterministic approach.

- The most important request is to keep "best estimate" data, at each stage, to remain consistent with the PSHA philosophy, and to obtain a "real" median estimation.
- It is also important to propagate variability but trying to separate random uncertainty from epistemic ones, especially in attenuation relationship, which is not done systematically at present.
- Finally, it must be kept in mind that logic trees allow to quantify the variability on expert judgment but not the variability on the physical parameter itself, this should be improved.

Consequently, the lesson learnt is that results from PSHA may be strongly different from real seismicity, as recorded, especially depending on choices on input data that always depend on expert judgment.

- In that situation, the comparison to the instrumental experience as observed appears to be necessary to address such difficulties.

In that context, the use of the Bayesian updating technique has clearly shown its effectiveness. This technique should become a necessary step in PSHA methodology to reduce epistemic uncertainties in order to obtain results more rugged and consistent with observations.

As a final statement, we would like to emphasis the effectiveness of sharing experience between seismologists and structural engineers (as it was done here), which should be systematized in future for PSHA but also for other fields such as characterization of seismic motion and its damaging potential for instance.

REFERENCES

[COR-68], A. Cornell, "Engineering Seismic Risk Analysis", BSSA – 1968

[KLÜ-05], Jens-Uwe Klügel, "On The Use Of Probabilistic Seismic Hazard Analysis As An Input For Seismic PSA", *SMiRT-18, KM02-2, 2005*

[HUM-08-1], N. Humbert, E. Viallet, "An evaluation of epistemic and random uncertainties included in attenuation relationship parameters", *14th World Conference on Earthquake Engineering (07-0117), Beijing, China*

[HUM-08-2], N. Humbert, E. Viallet, "A method for comparison of recent PSHA on the French territory with experimental feedback", *14th World Conference on Earthquake Engineering (07-0118), Beijing, China*

[SCO-03], Oona Scotti & al., «Révision du zonage sismique de la France métropolitaine – Discussion des hypothèses et de leur impact sur l'aléa», *6e colloque national AFPS – France – 2003*

[MAR-02], Martin Ch., Combes Ph., Secanell R., Lignon G., Carbon, D., Fioravanti A., Grellet, B. (2002) «Révision du zonage sismique de la France, Etude probabiliste» *GTR/MATE/0701-150.*

[RAP], Réseau accélérométrique permanent, http://www-rap.obs.ujf-grenoble.fr

[MAD-85], H. O. Madsen, "Model updating in First-Order Reliability Theory with application to fatigue crack growth", *2nd International Workshop on Stochastic Methods in Structural Mechanics, Pavia, Italy, 1985*

[HEI-99], G. Heinfling, I. Petre-Lazar, P. Hornet, "Use of Probabilistic Tools for the Material and Structural Analyses of Ageing Reinforced Concrete Structures", *RILEM 8 – Bratislava - Slovakia – 1999*

[HEI-05], G. Heinfling, A. Courtois, E. Viallet, "Reliability Based Approach for the Ageing Management of Prestressed Concrete Containment Vessel", *SMiRT-18, H06-5, 2005*

[ESMDB], European Strong-Motion Database, Imperial College (London, UK), Ente Nazionale per l'Energia Ellettrica (Roma, Italy), Ente per le Nuove Tecnologie, l'Energia e l'Ambiente (Roma, Italy), Institut de Protection et de Sûreté Nucléaire (Fontenay-aux-Roses, France), European Commission, Directorate General XII, Science, Research and Development, Environment and Climate Program, 2000.