The 14th World Conference on Earthquake Engineering
October 12-17, 2008, Beijing, China

An evaluation of epistemic and random uncertainties included in attenuation relationship parameters

N. Humbert[1] and E. Viallet[2]

[1]*Engineer, EDF SEPTEN, Dept. of Dynamics & Earthquake Engineering, Lyon France*
[2] *Team Manager Dept. Of Dynamics & Earthquake Engineering, EDF SEPTEN, Lyon France*
Email: nicolas.humbert@edf.fr, emmanuel.viallet@edf.fr

ABSTRACT :

Attenuation relationships are generally built in a deterministic way: based on physical considerations, the overall process (including stability tests) focuses on obtaining median values (median from epistemic and random variability). According to this point of view the magnitude, depth and localization of data base earthquakes are supposed to be perfectly estimated.
The objective of the paper is to present a stochastic method that allows to take into account these uncertainties of data base at the regression step of the attenuation relationship.
This method is applied on a European Strong Motion Database with three steps:
(i) an evaluation of data base uncertainties,
(ii) a regression taking into account the fuzzy data base,
(iii) an estimation of model uncertainties.
Finally this paper (i) presents a method which allows to build appropriate attenuation relationships for PSHA and clearly shows that (ii) the usual methods used to build attenuation relationships overestimate random variability and underestimate epistemic variability, introducing bias in PSHA.

KEYWORDS: Attenuation Relationship, random variability, epistemic uncertainties

1. INTRODUCTION

Due to large seismic motion database available nowadays, many attenuation relationships are available and used widely for deterministic or Probabilistic Seismic Hazard Assessments (PSHA).
Attenuation relationships are generally built in a deterministic way and focuses on obtaining the best evaluation of median values. With this goal in mind this process does not allow to take into account data base uncertainties. So magnitude, depth and localization of data base earthquakes are supposed to be perfectly estimated.
In the stochastic spirit of a PSHA, the epistemic and random uncertainties should be propagated at every step of the process in order to quantify each of the parameters in term of (i) median, (ii) epistemic and (iii) random variability.
This paper introduces a real discussion on epistemic uncertainties of attenuation relationships and on evaluation of random variability of seismic motion.
The argument is based on the work of Berge Thierry & Al. on a European Strong Motion Database (Ambraseys & Al.). This work uses an attenuation model based on the following equation:

$$\mathrm{Log}(PGA) = aM + bR - \mathrm{Log}(R) + c \qquad \pm \sigma \tag{1.1}$$

The σ parameter is the deviation representing the quality of regression. Commonly σ is used in PSHA to represent random uncertainties. This paper proves that this parameter overestimates the random variability by mixing a part of epistemic uncertainty in the random term of the attenuation relationship.

1. EVALUATION OF DATABASE UNCERTAINTIES

The first step consists in an evaluation of data base uncertainties. The selection of Ambraseys strong motion data base included by Berge Thierry in the regression of attenuation relationship is a set of 960 horizontal time histories recorded on 138 different earthquakes. These seismic events are Mediterranean and Californian ones recorded on site with Vs30 (average velocity of shearing waves in the 30 first meters). The hypocenter depths are spread from 0 to 30 kilometers and the magnitudes from 4 to 7.4.

1.1. Method

With the determination of data base uncertainties as an objective we compare Ambraseys evaluations for surface magnitude, depth and localization with the evaluations of the International Seismological Center (ISC). This comparison is done for the 138 events of the data base. So we get two evaluations for each magnitude, depth and epicenter localization of each event.
Because the recording and the method used by ISC and Ambrasey are different this method gives a good evaluation of epistemic uncertainties by mixing uncertainties due to model hypothesis and uncertainties due to the lack of recording.

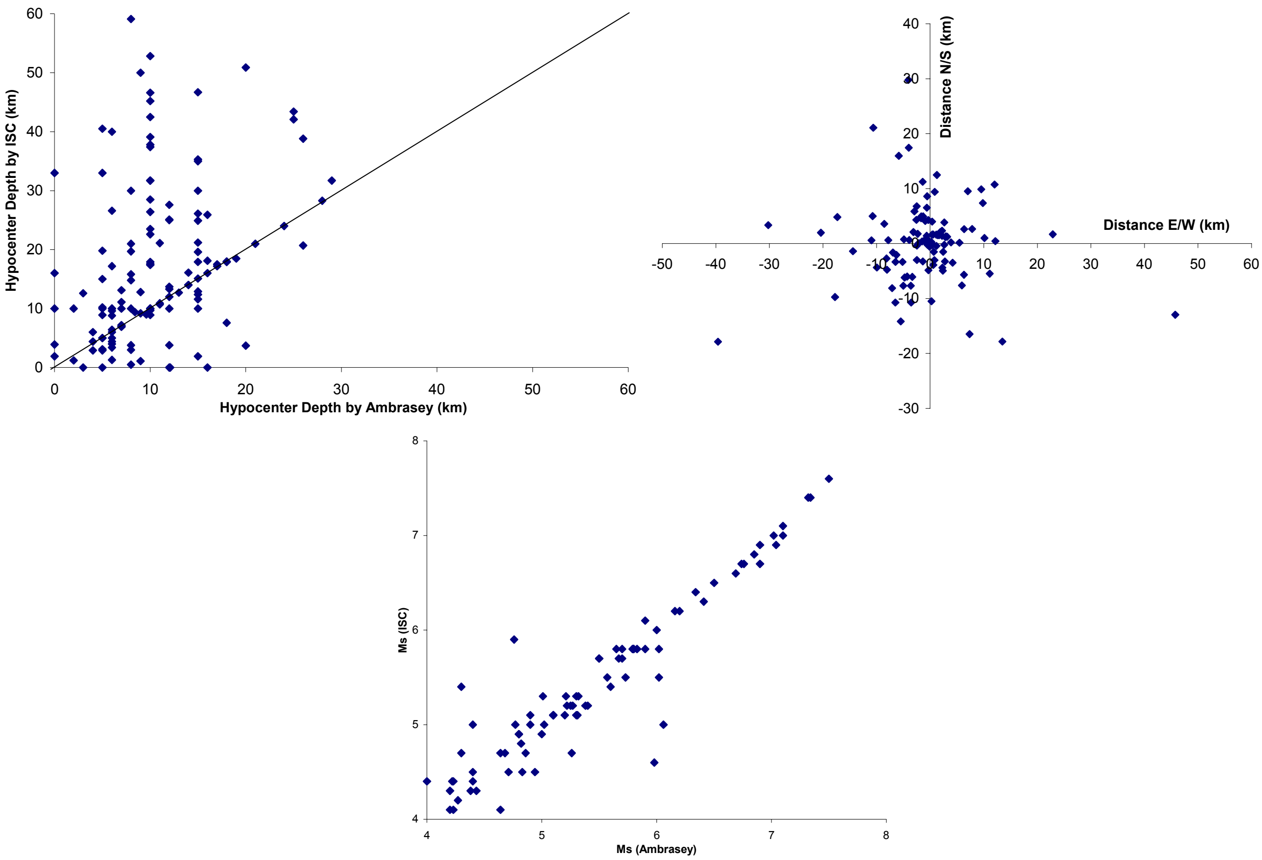

Figure 1 Comparison of Ambraseys and ISC evaluations for the same database

1.2. Results

Based on this double sample for each value of the database we build a statistical model of epistemic uncertainties (Table 1.1). These models are built with regression on the 138 events. We observe clearly that some events are well-defined and some other not. The effect of this reality on regression is a kurtosis higher than 3.
These models are accounted for in the attenuation relationship regression: this is the “regression on fuzzy data”.

Table 1.1 Results of regression for database uncertainties

Parameter		Standart Deviation	Skewness	Kurtosis
Magnitude	Centered	0.47	0	40
Depth	"LogNormal"	*1.7 /1.7	/	/
Localization	Centered	6 km	0	>>3

1.3. Comparison with uncertainties proposed by ISC

The International Seismological Center gives evaluation of magnitude, depth an localization uncertainties. Based on the model developed in 1.2, we can have an idea of the consistence between the two evaluations.
(i) Surface magnitude: the ISC gives an average value of 0.3 for this uncertainty. This is consistent with our evaluation.
(ii) Depth: the ISC gives an average uncertainty of 2 km (in standard deviation). This value is inconsistent we our evaluation at 1.2. The comparison between ISC and Ambraseys gives a much higher value of uncertainty. (cf. Figure 1)
(iii) Localization: the ISC gives an average uncertainty of 2.5 km. This is inconsistent with our evaluation (cf Figure 1).
In conclusion of this part we notice that this comparison is inconsistent: the uncertainties proposed by ISC are clearly underestimated. It is obvious that there is a lack of a real discussion on uncertainties in data base evaluation.

2. REGRESSION ON FUZZY DATA

Accounting for database uncertainties in the regression process allows us to describe the epistemic uncertainties on the attenuation parameters « a , b, c and σ ». The table 2.1 gives the results of the fuzzy regression on these four parameters. As shown the fuzzy regression gives the same median value for a, b and c parameters, but with non negligible uncertainties of calibration (in opposition to the deterministic approach which shows no uncertainties). The σ parameter used in PSHA to represent the random variability is clearly revised down in median value by the fuzzy regression, but with an epistemic variability of 13%.
In conclusion the σ parameter of a "determinist" regression overestimates the random variability by confusing a part of epistemic uncertainty in the random part. That the reason why an attenuation relationship built in a deterministic way induces an automatic bias in PSHA.

Table 2.1 Results of regression on fuzzy database

		a	b	c	σ
Deterministic regression		0.31	-0.00093	1.55	**0.29**
Fuzzy regression	Median	0.31	-0.00091	1.57	**0.23**
	Uncertainty	14%	40%	10%	13%

The figure 2 gives us a description of the epistemic uncertainty directly estimated on Pick Ground Acceleration (PGA). We observe that this uncertainty is not homogeneous and get a significant increase on the edge of magnitude and distance of the database.
But for small values of hypocentral distance the epistemic uncertainty does not increase as expected. This is a clue that the attenuation model lacks a degree of freedom impacting the near field values.

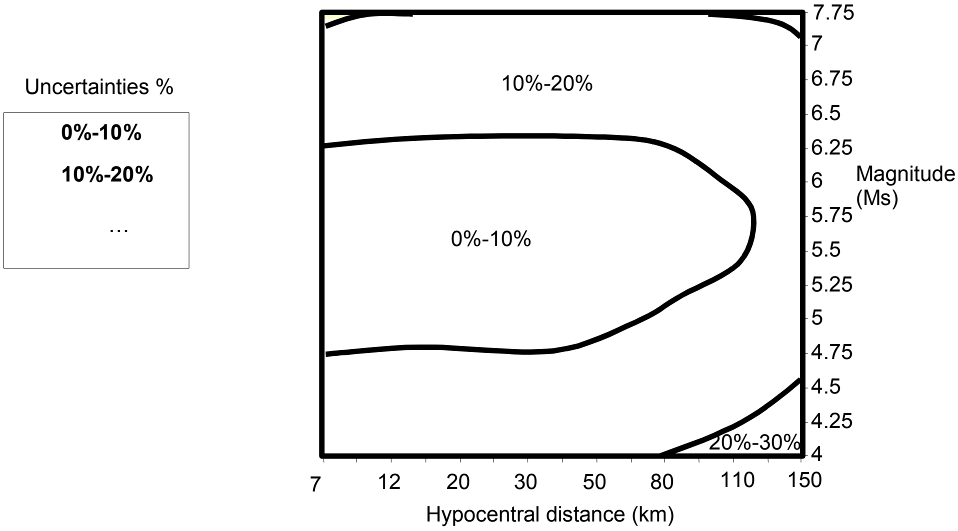

Figure 2: Effect of database fuzziness, epistemic uncertainties deaggregation

3. ATTENUATION MODEL UNCERTAINTIES

3.1. Method

Three attenuation models are developed :
(i) The attenuation model of Berge Thierry (equation 1.1)
(ii) An attenuation model developed by Ambraseys & Al. in 2005 :

$$\log(PGA)=aM+a2M^2+(b1+b2M)\log(\sqrt{R^2+d^2})+c \pm\sigma \tag{3.1}$$

(iii) An attenuation model developed by Marin & Al. in 2004:

$$\log(PGA)=aM+bR-b1*\log(R)+c \pm\sigma \tag{3.2}$$

Various regression techniques are developed on fuzzy data: a two step regression developed by Fukushima (i), a standard bilinear regression (ii) and a non linear regression (iii).
Moreover in order to increase the effect of well known events, a likelihood weighting is added on several branches. Based on these various options, a logical tree is specified on figure 3.
For each branch of the logic tree, epistemic uncertainties of the data base are accounted for by fuzzy data. So we compute more than 400 regressions for each branch.

3.2. Generic equation

The result of these multiple regressions is a set of more than 2000 attenuation relationships. In order to homogenize the result for the three attenuation models developed, a generic equation is proposed with 7 parameters.

$$\log(PGA) = aM + a_2M^2 + bR + (b_1 + b_2M)\log\left(\sqrt{R^2 + b_3^{\,2}}\right) + c \pm \sigma \tag{4.3}$$

For each attenuation model, some parameters are not effective, so they are kept equal to zero. But for the whole logic tree, we keep the same equation.

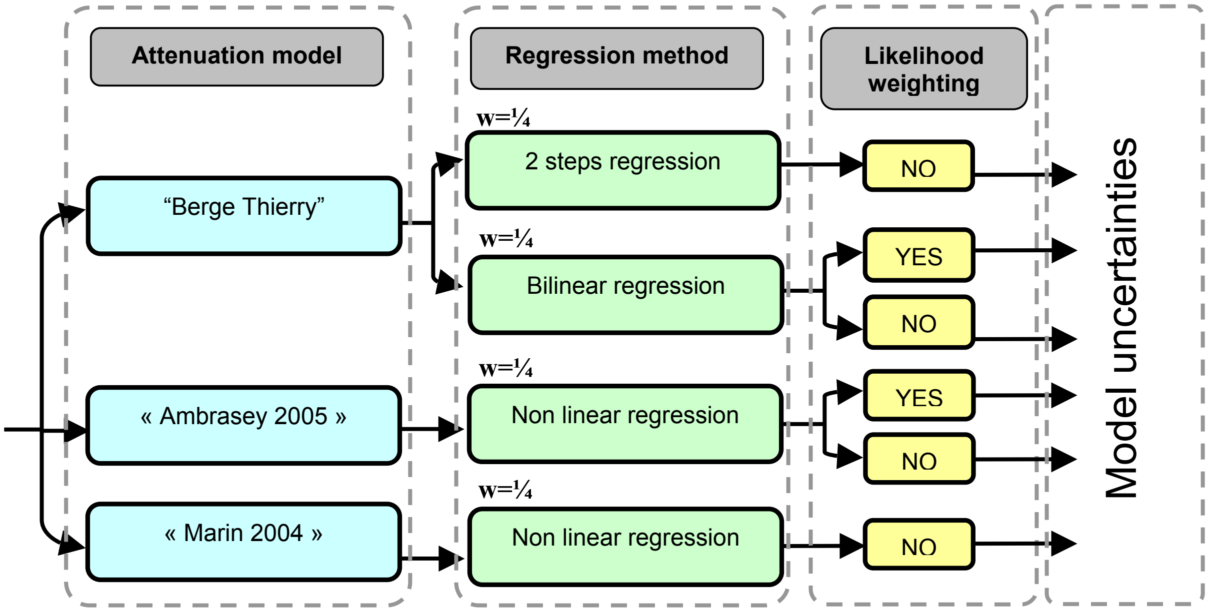

Figure 3 : Logic tree

3.3. Magnitude conversion study

The data base of this study is exclusively built with surface magnitude. However it is easy to build the same attenuation law for local magnitude by using a conversion equation between Ms and Ml. Introducing this conversion at this step of the study (before the regression) is especially interesting in that it allows taking into account uncertainties of such a conversion.

A second study is conducted developing an attenuation relationship in Ml. We use a conversion relationship recommended by Marin et al. in 2004.

$$Ml=0.64Ms+2.12\ (\pm 0.2) \tag{4.4}$$

The consequence of using this conversion is an addition of fuzziness on the data base. Consequently the relationship developed in Ml get more epistemic uncertainty than the relationship in Ms.

3.3. Results

2000 regressions are computed. Based on a statistical analysis of these 2000 relationships in term of epistemic fractiles (15% 50% and 85%) and in term of deaggregation on distance and magnitude we can write following conclusions :

3.3.1. Effect of model on uncertainties

On one hand epistemic uncertainty induced by the model is clearly higher than the one obtained with only one attenuation model (figure 4). Therefore the epistemic uncertainty induced by the model has a great impact on the final attenuation relationship. On the other hand there is absolutely no effect of model on random variability.

3.3.2. Consistence with validity domain

Epistemic uncertainties presented on figure 4 are consistent with the magnitude / distance in the data base. In particular, for the small values of hypocentral distance the epistemic uncertainty increases as expected. This observation makes the point that working on model uncertainties is fundamental when developing an attenuation law.

3.3.3. Uncertain domains

Epistemic uncertainties increase in two main zones of the validity domain. First the large and near earthquakes : this is rational because these events are extremely rare. Secondly the small and distant earthquakes : this is surprising because these events have the largest frequency of occurrence and have an significant impact on PSHA.

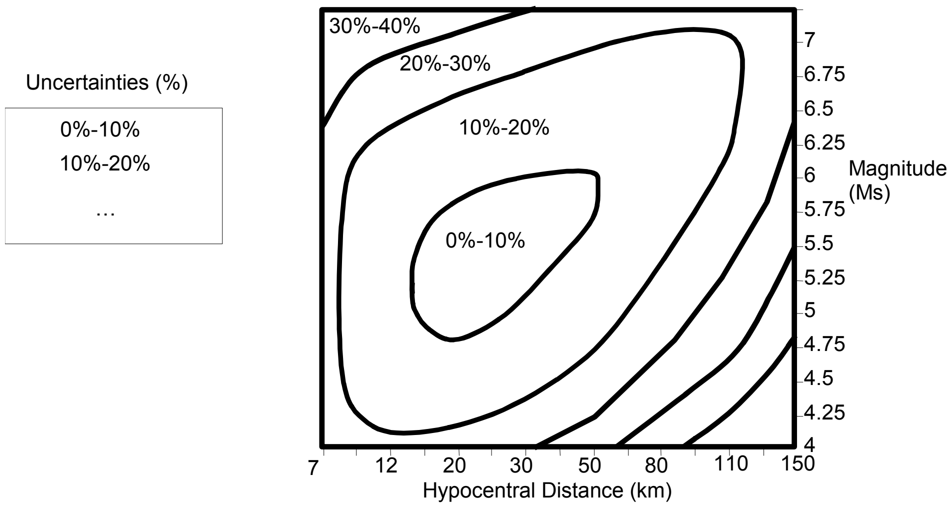

Figure 4 : Epistemic uncertainties, Data base + attenuation model : deaggregation on magnitude and distance.

3.3.4. Discussion on model uncertainties origin

Increase of epistemic uncertainties induced by attenuation model has two origins :

(i) The interval between median evaluations for two attenuation models can increase final epistemic uncertainties (cf. Figure 5).

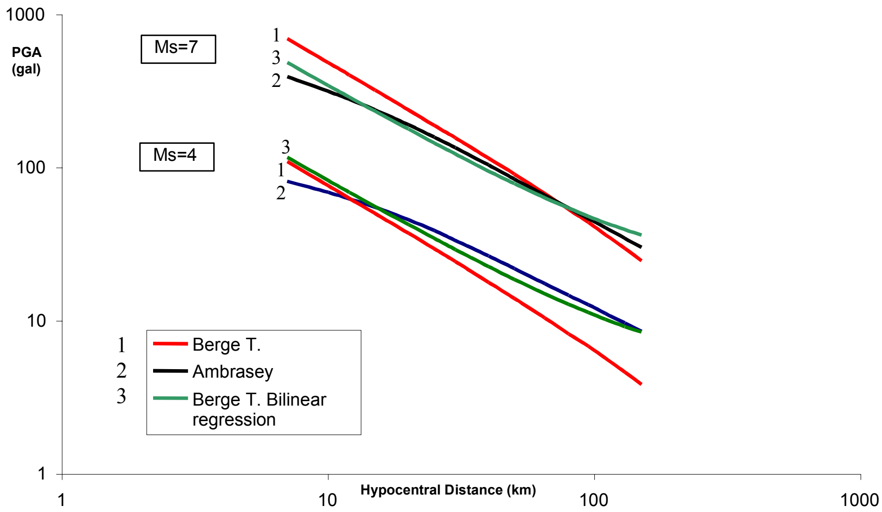

Figure 5 : Attenuation model deaggregation (median values).

(ii) For each branch the regression quality depends on the number of parameters in the model. The bias-variance trade-off (or "bias-variance dilemma") is a very important issue in data modelling: models with too few parameters are inaccurate because of a large bias (not enough flexibility) but models with too many parameters are inaccurate because of a large variance (too much sensitivity to the sample). On figure 6 we observe larger uncertainties for Ambraseys model than for Berge Thierry model.

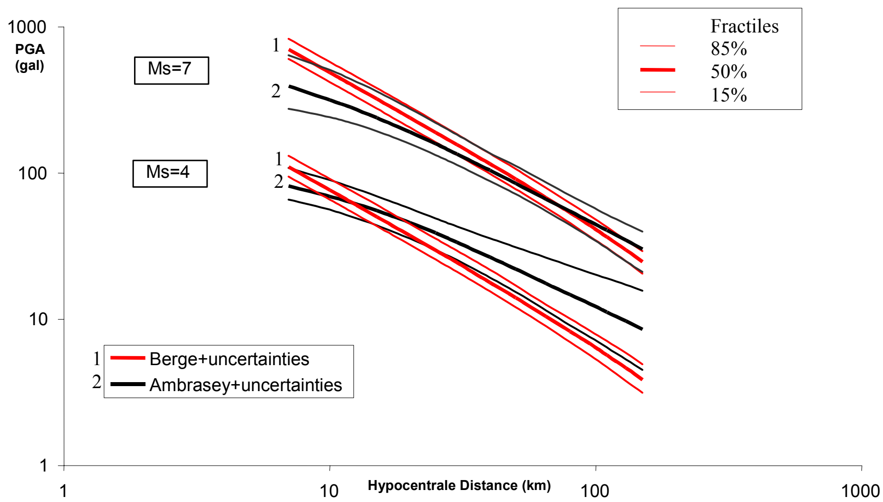

Figure 6 : Attenuation model deaggregation (median values+15% and 85% fractiles).

3.3.5. Discussion on model uncertainties origin

The magnitude conversion Ms to Ml has to impacts (cf. figure 7):

(i) An increase of epistemic uncertainties for the low magnitudes, (ii) An average deviation for the low magnitudes (median value). For the large magnitudes the conversion has no effect.

But magnitude conversion has absolutely no effect on random variability.

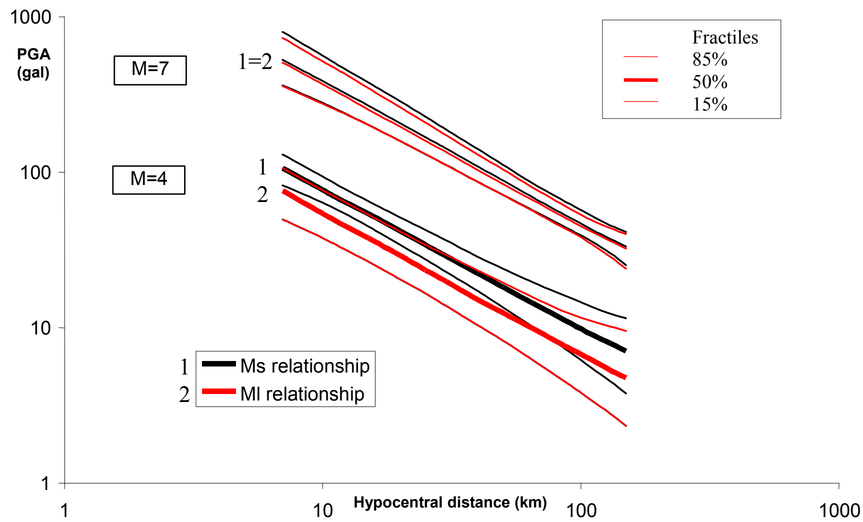

Figure 7 :Effect of magnitude conversion

4. CONCLUSIONS

4.1. On the use of attenuation relationship in PSHA

This study underlines that the effect of database uncertainties has to be accounted for in regression for two reasons:

(i) The parameters of the attenuation relationship are not perfectly fitted by the regression and include epistemic uncertainty. A “deterministic” regression underestimates this epistemic uncertainties.

(ii) The σ parameter of a "determinist" regression overestimates the random variability by confusing a part of epistemic uncertainty in the random part. That the reason why an attenuation relationships built in a deterministic way induces an automatic bias in PSHA. This bias can be important especially for the long term periods.

4.2. On the importance of attenuation model choice

(i) This study underlines that a large part of epistemic uncertainties is induced by the choice of attenuation models. We clearly observe that on the edges of the database validity domain the epistemic uncertainties increases sharply. That means that regression result becomes imprecise.
(ii) Using attenuation models with more degree of freedom enable to be sensitive to phenomenon that simple model occults, such as saturation of seismic motion in the near field. However these model are more instable. This is the bias-variance trade-off: model sophistication is limited by the size and the quality of databases.
(iii) Deaggregation of epistemic uncertainties on magnitude and distance reveals the point that large and near earthquakes are not the worse adjusted events in regression. The small and distant earthquakes are clearly badly represented in database although these events have a significant impact on PSHA. This deficiency could be avoided because these events have the largest frequency of occurrence. This observation is only a deficiency for a probabilistic approach in which small and distant events influence the hazard: at the very beginning this relationship was developed in a deterministic approach which does not care about the small and distant events.

4.3. Attenuation relationship in PSHA

A real discussion on epistemic uncertainties and random variability should be introduced in PSHA. Especially the attenuation logic tree presented figure 3 should be accounted for by the general PSHA logic tree in order to have a" best estimate" evaluation in epistemic uncertainties without bias on the median value of hazard.

REFERENCES

Fukushima Y., Tanaka T. (1990). A new attenuation relation for peak horizontal acceleration of strong earthquake ground motion in Japan. *BSSA*.

Berge-Thierry C., Cotton F., Cushing M., Griot-Pommera D.A., Joly J., Levet A., Scotti O. (2000). Méthode de détermination des spectres horizontaux et verticaux adaptés au site dans le cadre de la réglementation fondamentale de sûreté I.2.c. *Rapport IPSN/DPRE*.

Ambraseys N.N., Douglas J., Sarma S.K., Smit P.M. (2005). Equations for the Estimation of strong ground motion from shallow earthquakes using data from Europe and the middle East : horizontal peak ground acceleration and spectral acceleration. *Bulletin of Earthquake Engineering*.

Marin S., Avouac J.P., Nicolas M., Schlup A. (2004). A Probabilistic Approach to Seismic Hazard in Metropolitan France. *BSSA*.

Ambraseys N.N., Free M.W. (1997). Surface-wave magnitude calibration for European region earthquakes. *Journal of earthquake engineering* **Vol. 1, No. 1.** 1-22.

Saporta G. (2006), Probabilités analyse des données et statistique, Technip.

Douglas J. (2001). A comprehensive worldwide summary of strong-motion attenuation relationships for peak ground acceleration and spectral ordinates (1969 to 2000) *Engineering Seismology and Earthquake Engineering.*