Methodology and Testing of GMPEs based on ground-motion data from France and surrounding countries

Deliverable D2-12

AUTHORS			REVIEW			APPROVAL		
NOM	DATE	VISA	NOM	DATE	VISA	NOM	DATE	VISA
E. Delavaud	24/11/2011		PY. Bard ISTerre Grenoble Member of the SIGMA Scientific Committee	18/11/2011	In written review Attached document	G. Senfaute Leader of the SIGMA project		
			J. Savy Risk Consulting USA Member of the SIGMA Scientific Committee	18/11/2011	In written review Attached document			

***DISSEMINATION**: Authors; Steering Committee; Work Package leaders, Scientific Committee, Archiving.*

Summary

The goal of the task X2-4 of work package 2 is to select and rank ground-motion prediction equations (GMPEs) for seismic hazard assessment (SHA) in France and surroundings countries. It is needed to establish a logic tree for ground motion prediction that captures well epistemic uncertainty. A two years post-doc is dedicated to this task that involves three steps: the collection of data, the pre-selection of candidate GMPEs and the testing of the GMPEs using data driven methods.

The question associated with this task is the following: which GMPEs are appropriate in a specific region (selection) and with which degree of appropriateness (ranking)? This question is particularly critical in low seismicity regions such as France or Switzerland that do not have any indigenous strong motion prediction equation because of a lack of data. Moreover, many studies have identified a magnitude scaling of ground motion (stronger decrease at large magnitudes than at low magnitudes) that prevents to use GMPEs based on weak motion to predict strong motion (e.g., Bommer *et al.,* 2007, Cotton et al., 2008). In this context, we search which GMPEs derived for other regions can be applied to predict ground motion in France, and for which ranges in magnitude, distance and response spectrum period, having in mind that the whole magnitude range (from magnitude 4) has a contribution to the hazard in these regions.

1. Data

As a first step, we use data from the French Accelerometric Network (RAP) recorded between 1995 and 2007 (Péquegnat *et al.,* 2008). The maximum available magnitude is 4.6. The minimum magnitude we considered is 3.8 in order to limit the extrapolation of the GMPEs outside their magnitude validity range while having a sufficient amount of data to perform a robust GMPE testing. The response spectra have been visually checked. Records with the three components and a good signal to noise ratio were selected. Table 1 summarizes the characteristics of the 18 events used in the present study. They are also provided in an excel file (eventmetadata_RAP.xls).

For each event, the necessary metadata are: the moment magnitude, the focal mechanism, the location and the depth. Concerning the values assigned to these metadata, we decided to follow the criteria of Beauval *et al.* (2012) who worked on the testing of ground-motion prediction equations against small magnitude data.

- Moment magnitudes have been extracted from the catalogue of magnitude available on the RAP website (`http://www-rap.obs.ujf-grenoble.fr/spip.php?article38`). Moment magnitudes have been determined by Drouet et al. (2010) using an inversion scheme. Four of them were extracted from Beauval *et al.* (2012) who applied a magnitude conversion equation from local magnitudes determined by the "Réseau National de Surveillance Sismique" (RéNaSS: `http://renass.u-strasbg.fr/`) magnitude to moment magnitude.

- The focal mechanism is known only for 10 events. They have been extracted from Beauval *et al.* (2012). As magnitudes are small, the focal mechanism does not have a significant impact on the GMPEs prediction. In practice, the focal mechanism has been set to "reverse" for the events with an undefined focal mechanism. Another solution could have been chosen, used by Delavaud *et al.* (2012a). They computed predictions for each of the three possible mechanisms (reverse, normal and strike-slip) and determined the LLH value as the weighted mean of the three obtained LLH values. The weights can be estimated from the known proportion of each mechanism in the region of interest.

- Depth is a parameter that might have more impact on the GMPEs predictions as it directly influences the source-to-site distances. Depths have been extracted from the catalogue of magnitude available on the RAP website (`http://www-rap.obs.ujf-grenoble.fr/spip.php?article38`).

- Some GMPEs require additional metadata: the depth to top of rupture and the depth to a shear wave velocity of 2.5km/s (Abrahamson and Silva, 2008) or 1km/s (Chiou and Youngs, 2008) to take a basin effect into account. As we considered only small events, we set the depth to top of rupture equal to the hypocenter depth. The basin effect was not taken into account due to a large uncertainty of the associated parameters. The down dip rupture width required by some NGA models was determined using the relation of Wells and Coppersmith (1994).

For each record, the necessary metadata are: the location, the hypocentral and epicentral source-to-site distances and a Vs30 (shear velocity averaged over the upper 30 m) estimate. These metadata are provided in an excel file (obsmetadata_RAP.xls).

- As the magnitudes are small (<=4.6), the Joyner-Boore and rupture source-to-site distances are supposed to be equal to the epicentral and hypocentral distances respectively. Hypocentral distances up to 300km are considered.

- Both Vs30 estimates and EC8 classes are available for the RAP stations together with confidence levels (`http://www-rap.obs.ujf-grenoble.fr/IMG/txt/RAP2010-fiches_RAP-V1.txt`). We used Vs30 values with a "medium" to "high" confidence level. For a lower Vs30 quality, a mean value corresponding to the EC8 class is used: 1000m/s for "A" EC8 class, 600m/s for "B", 250m/s for "C" and 100m/s for "D, as Beauval *et al.* (2012) did.

- Some GMPEs need additional information about each site: whether it is located on a hanging wall region or not and how the associated Vs30 has been determined. We did not take the hanging wall effect into account, as the events are small. We set the Vs30 determination to "estimated" for all Vs30s due to the associated uncertainties. It has an influence on the standard variation of the GMPE's prediction. The missing parameters required by the NGA models can be also estimated according to Kaklamanos *et al.* (2010).

The dataset obtained is very similar to the one used in Beauval *et al.* (2012) to test GMPEs against small magnitude French data. The main difference resides in the events considered: their dataset does not contain 3 of our events (on the 22.02.03, the 05.12.04 and the 02.09.06) and they considered an event that our dataset does not includes, the event on the 25.02.01. For some common events, the number of observations is not the same.

Our dataset is composed of 18 earthquakes with 182 observations for response spectra ranging from 0.2Hz to 10Hz. A good homogeneity in terms of magnitude and distance is obtained, as shown in Figure 1. The location of the events and stations are shown in Figure 2. The events are located in the Pyrenees, Alps and Lower Rhine Embayment that have been considered as active shallow crustal regions by Delavaud *et al.* (2012b).

The influence of the uncertainties associated with the events and observations metadata on the ranking results will be addressed in another section. This concerns the moment magnitude, the mechanism and the depth of the hypocenter.

Table 1: Characteristics of the 18 events used for the GMPEs testing

Event date	Mag. (Mw)	Mechanism	Longitude (°)	Latitude (°)	Depth (km)	Number of obs.
31-10-1997	4.0	U	6.57	44.26	2	5
21-08-2000	4.4*	SS	8.44	44.86	10	10
16-05-2002	4.0	R	-0.16	42.94	10	9
11-12-2002	3.8	U	-0.33	43.04	5	5
12-12-2002	4.0	U	-0.28	43.11	10	9
21-01-2003	3.8	U	-0.36	43.05	10	12
22-02-2003	4.5	N	6.66	48.31	10	10
22-03-2003	3.9	U	8.91	48.19	5	3
11-04-2003	4.3*	SS	8.83	44.87	5	22
23-02-2004	4.2	SS	6.28	47.30	10	15
18-09-2004	4.6	N	-1.6	42.78	2	9
30-09-2004	4.1	N	-1.45	42.77	10	8
05-12-2004	4.1	U	8.0	48.11	10	1
08-09-2005	4.4	SS	6.87	46.01	10	17
02-09-2006	3.8	U	7.59	43.92	10	11
17-11-2006	4.5	N	0.01	43.08	11	18
30-07-2007	4.0*	U	9.71	44.92	10	4
15-11-2007	4.0*	N	0.0	43.01	8	14

* *Calculated using the magnitude conversion* $M_w = f\,(M_{L\_Renass})$

Mechanism U stands for undefined. In practice, we used a reverse mechanism when unknown.

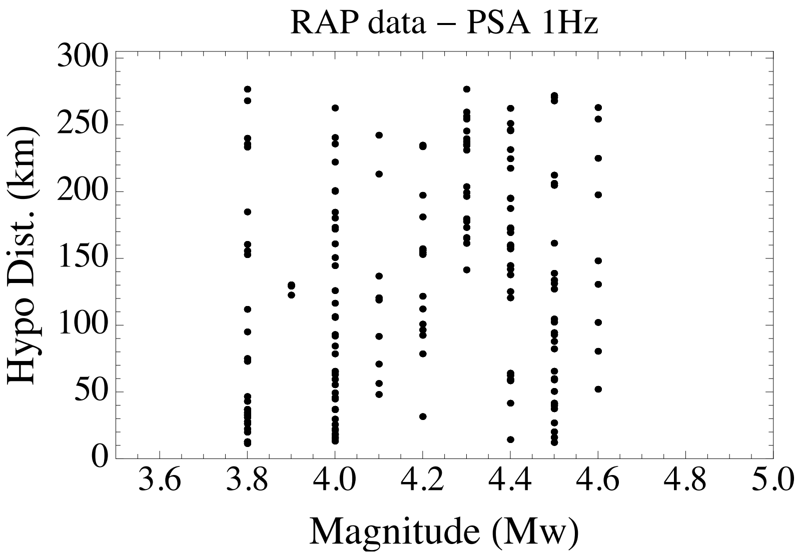


Figure 1: Distance - magnitude distribution of the French Accelerometric Network (RAP) data used for this study.

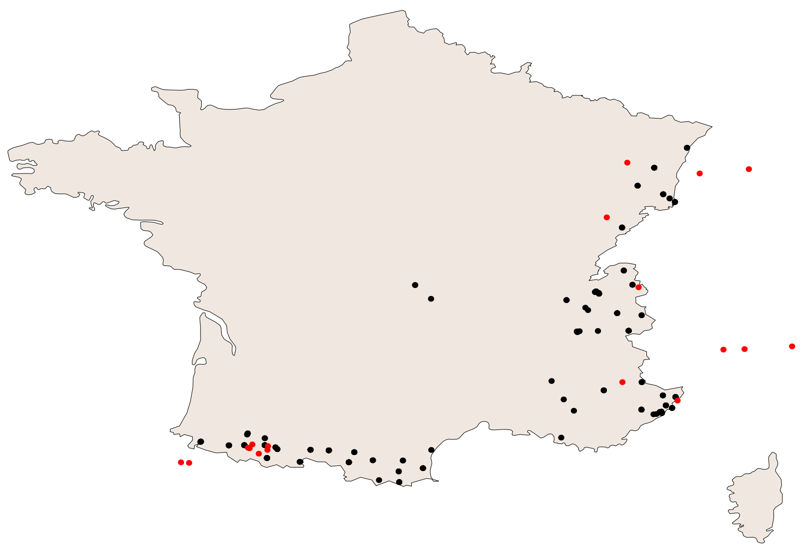


Figure 2: Epicenters (in red) of the earthquakes considered in this study and stations (in black) of the French Accelerometric Network where observations were recorded. Some events are on Germany, Italy and Spain.

2. Pre-selection of GMPEs

20 GMPEs have been pre-selected. Their characteristics are summarized in Table 2.

14 of them are developed for active shallow crustal regions (ASCR) (in red, blue and green in Table 2). Their datasets are composed of data from California, Europe and Middle East, Japan, Turkey, Italy or Spain (see last column of Table 2). In order to better capture epistemic uncertainties, 8

GMPEs derived for stable continental regions (SCR) are also considered. Most of them (6) are based on data from Eastern North America.

The report SIGMA WP2 D2.5 presents these GMPEs in details. Here are characteristics that particularly interest us:

- The models of Boore and Atkinson (2011), Atkinson (2011) and Atkinson and Boore (2011a) are revisions of the models of Boore and Atkinson (2008), Atkinson (2008) and Atkinson and Boore (2006) respectively following Atkinson and Boore (2011b). These revisions are simple modifications to take into account new data available for magnitudes below 5.75.

- Only 3 of the selected GMPEs are able to predict ground motion for magnitudes below 4: Atkinson and Boore (2011) [min Mw = 3.5], Akkar and Cagnan [min Mw = 3.5] and Bommer *et al.* (2007) [min Mw = 3]. These names are highlighted in Table 2.

- All the SCR GMPEs as well as the models of Zhao *et al.* (2006), Berge-Thierry et al. (2006) and Kanno et al. (2006) are valid for distances up to 300 km.

- All GMPEs are valid for frequencies up to 10Hz while large differences exist for the minimum response spectra frequency. 12 GMPEs are valid for frequencies down to 0.2Hz. Predictions will be calculated while respecting the GMPEs frequency validity. We will not extrapolate in the frequency domain.

- GMPEs are not adjusted but adjustments will be done in the next weeks, especially for the rock definition for the SCR GMPEs.

Table 2: Characteristics of the 20 candidate GMPEs

GMPE reference	GMPE Acronym	Magnitude range (Mw)	Distance range (km) and type	Frequency range (Hz)	Main region of the generating dataset
Campbell (2003)	Ca03	5.0 - 8.2	0 - 1000 (R_{RUP})	0.25 - 50	ENA
Toro (2002)	To02	5.0 - 8.0	0 - 1000 (R_{JB})	0.5 - 33.3	ENA
Atkinson and Boore (2011a)	AB11	3.5 - 8.0	0 - 1000 (R_{RUP})	0.2 - 40.0	ENA
Atkinson (2011)	At11	4.3 - 7.6	10 - 1000 (R_{JB})	0.2 - 10.0	ENA
Pezeshk *et al.* (2011)	Pe11	5.0 - 8.0	0 - 1000 (R_{RUP})	0.1 - 100	ENA
Raghu Kanth and Iyengar (2007)	RKI07	4.0 - 8.0	5 - 300 (R_{HYP})	0.25 - 100	India
Silva et al. (2002)	Si02	4.5 - 8.5	1 - 400 (R_{JB})	0.1 - 100	ENA
Somerville et al. (2009)	So09	5.0 - 7.5	1 - 500 (R_{JB})	0.1 - 100	Australia
Akkar and Bommer (2010)	AB10	5.0 - 7.6	0 - 99 (R_{JB})	0.33 - 20	EME
Berge-Thierry *et al.* (2003)	Be03	4.0 - 7.9	4 - 330 (R_{HYP})	0.1 - 34	EME
Bommer *et al.* (2007)	Bo07	3.0 - 7.6	0 - 100 (R_{JB})	2 - 100	EME
Bindi et al. (2009)	Bi09	4.0 - 6.9	2.8 - 100 (R_{JB})	0.5 - 33.3	Italy
Douglas et al. (2006) Spain	DoS06	4.5 - 7.5	1 - 1000 (R_{JB})	0.5 - 50	Southern Spain
Akkar and Cagnan (2010)	AC10	3.5 - 7.6	1 - 200 (R_{JB})	0.5 - 33.3	Turkey
Abrahamson and Silva (2008)	AS08	5.0 - 8.0	0 - 200 (R_{RUP})	0.1 - 100	California
Boore and Atkinson (2011)	BA11	5.0 - 8.0	0 - 200 (R_{JB})	0.1 - 100	California
Chiou and Youngs (2008)	CY08	4.0 - 8.0	0.2 - 200 (R_{RUP})	0.1 - 100	California
Kanno et al. (2006)	Ka06	5.2 - 8.2	1 - 300 (R_{RUP})	0.2 - 20	Japan
Zhao *et al.* (2006)	Zh06	5.0 - 8.3	0 - 300 (R_{RUP})	0.2 - 20	Japan
Faccioli *et al.* (2010)	Fa10	5.0 - 7.2	10 - 150 (R_{RUP})	0.05 - 20	Japan

ENA: Eastern North America; EME: Europe and Middle East
R_{RUP}: *Rupture distance;* R_{JB}: *Joyner-Boore distance;* R_{HYP}: *hypocentral distance*

Options selected for the following models:
For Raghu Kanth and Iyengar (2007): Peninsular India region
For Silva et al. (2002): Double corner source model
For Somerville et al. (2009): Non-cratonic region
For Chiou et al. (2010): Central California

3. Testing based on the LLH method

The LLH method

The pre-selected GMPEs are tested against the RAP data presented in Section 1. The goal of this testing is to judge the applicability of candidate models by evaluating their probability to have generated the available data. We use the data-driven method developed by Scherbaum *et al.* (2009) that implemented an information-theoretic approach for the selection and the ranking of GMPEs. The method derives a ranking criterion from the Kullback-Leibler (KL) divergence, which denotes the information loss when a model g defined as a distribution is used to approximate a reference model f (Burnham and Anderson, 2002). The KL divergence between two models represented by their probability density functions f and g is defined as:

$$D(f,g) = E_f\left[log_2(f)\right] - E_f\left[log_2(g)\right] \qquad (1)$$

where E_f is the statistical expectation taken with respect to f.

In the case of GMPE selection, f represents the data-generating process (nature) and is only known through observations. Consequently, the term $E_f\left[log_2(f)\right]$ called the self-information of f cannot be calculated. However, the second term, $-E_f\left[log_2(g)\right]$, can still be approximated via the observations.

This approximation is the negative average sample log-likelihood, noted LLH and defined by:

$$\mathrm{LLH}(g, \mathbf{x}) := -\frac{1}{N}\sum_{i=1}^{N} log_2\left(g(x_i)\right) \qquad (2)$$

where $\mathbf{x} = \{x_i\}, i = 1, \ldots, N$ are the empirical data and $g(x_i)$ is the likelihood that model g has produced the observation x_i. In the case of GMPE selection, g is the probability density function given by a GMPE to predict the observation produced by an earthquake defined by a magnitude M (and by other characteristics such as the style of faulting) at a site i that is located at a distance R from the source.

We use the LLH divergence as a criterion to rank the candidate GMPEs. Due to its negative sign, the negative average sample log-likelihood is not a measure of closeness but a measure of the distance between a model and the data-generating process. A small LLH indicates that the candidate model is close to the process that has generated the data while a large LLH corresponds to a model that is less likely of having generated the data.

In order to interpret the rankings, weights obtained from the LLH values can be compared to the uniform weight $w_{\mathrm{unif}} = \frac{1}{M}$, where M is the number of GMPEs. This comparison tells us to what degree the data support or reject a model with respect to the state of non-informativeness. It is expressed by the data support index (DSI) which gives the percentage by which the weight of a model is increased (positive DSI) or decreased (negative DSI) by data. The DSI of model g_i with LLH-value based weight w_i is:

$$\mathrm{DSI}_i = 100 \ \frac{w_i - w_{\mathrm{unif}}}{w_{\mathrm{unif}}}, \qquad (3)$$

where

$$w_i = \frac{2^{-\mathrm{LLH}(g_i, \mathbf{x})}}{\sum_{k=1}^{K} 2^{-\mathrm{LLH}(g_k, \mathbf{x})}} \qquad (4)$$

Even though the reference $E_f\left[log_2(f)\right]$ is not known, some confidence interval of LLH can be identified in practice. (Beauval *et al.*, 2012) made synthetic tests to better interpret the LLH absolute values. These tests show that if testing the distribution that simulated the dataset with this same dataset, mean LLH values obtained are close to 1.4-1.5. Then, if testing distributions that differ from the original one, mean LLH are increasing. For a distribution with a mean equal to the original mean plus one sigma, and a sigma twice the original sigma, LLH values are around 2.0. In the worst case considered in their example, the tested distribution has a mean equal to the original mean plus 2.5 sigma, and a sigma equal to 0.8 times the original sigma, producing mean LLH values as high as 9-10.

The LLH-based weights defined by eq. (4) cannot be automatically regarded as probabilities as the LLH values are independently determined for each model (Kolmogorov's axioms of mutual exclusiveness and collective exhaustiveness are not respected) and only subsequently made to sum up to one (see Scherbaum and Kuehn (2011) for more details about this subject). Therefore, we advise not to directly use them as logic tree weights but to use them in combination with expert judgment. The purpose of using empirical data was not to replace expert judgment but rather to help the judgment process by providing additional information about the applicability of GMPEs, especially in regions where no indigenous model exists.

We refer to Scherbaum *et al.* (2009) for a theoretical description of the method and to Delavaud *et al.* (2012a), Delavaud *et al.* (2012b) and Beauval et al. (2012) for applications of the method.

4. Results

We compute LLH values for a set of frequencies that respects the frequency validity of each GMPE (5th column in Table 2). The larger set of frequencies is: 0.2, 0.3, 0.4, 0.5, 1, 1.5, 2, 3, 4, 6, 8 and 10 Hz. On contrary, we allow extrapolations in magnitudes and distances. The moment magnitude range is 3.8 to 4.6. The distance range is 0 to 300 km.

Extrapolations outside the validity range of GMPEs in terms of magnitudes and distances are not recommended (Bommer *et al.*, 2007) but they can still be performed. It is an extrapolation in the functional form that can be appropriate if the functional form is robust and physical enough (Delavaud *et al.* (2012a) found that the model of Akkar and Bommer could be successfully extrapolated at larger distances). However, extrapolation in frequency should be excluded. To each spectral frequency predicted are associated coefficients that have been determined by regression. Such coefficients cannot be extrapolated for an other spectral frequency.

First, we show that we obtain consistent results with the study of Beauval *et al.* (2012). Figure 3 shows the LLH values in function of the frequency for the GMPEs selected by Beauval *et al.* (2012): Akkar and Bommer (2010), Cauzzi and Faccioli (2008), Chiou *et al.* (2010), Abrahamson and Silva (2008), Bindi *et al.* (2009), Zhao *et al.* (2006), Boore and Atkinson (2011) and Chiou and Youngs (2008).

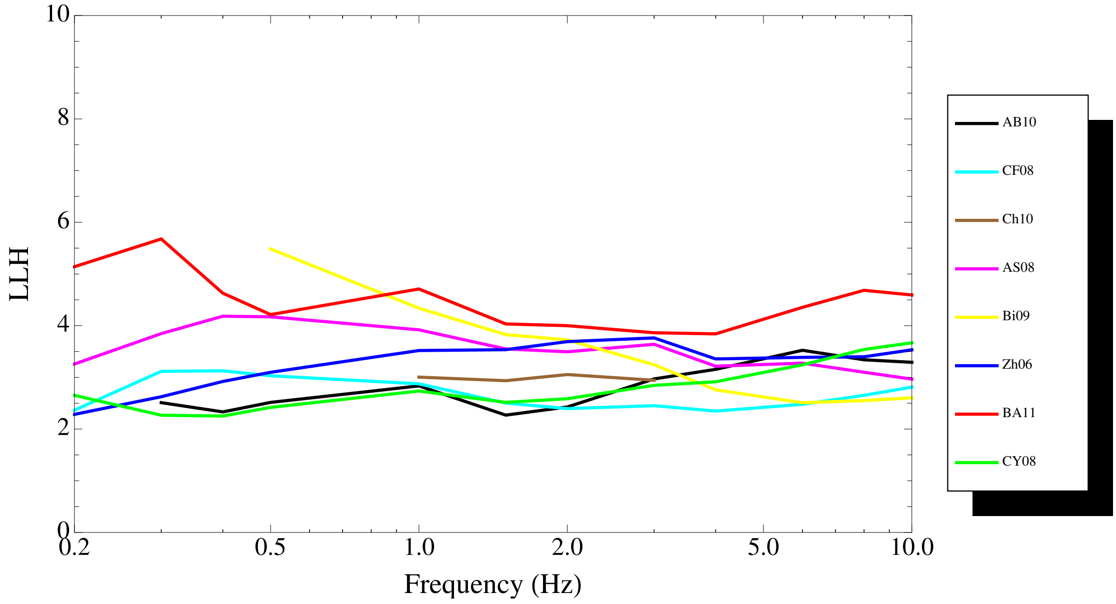


Figure 3: LLH values in function of frequency for the GMPEs selected by Beauval et al. (2012). Acronyms are given in Table 2.

Figures 4 shows the LLH values in function of frequency for the GMPEs that obtain the lowest LLH values for almost all frequencies. Figure 5 shows the corresponding DSI graph. Figure 6 shows the same graph for the GMPEs that obtain higher LLH values. Figure 7 shows the corresponding DSI graph.

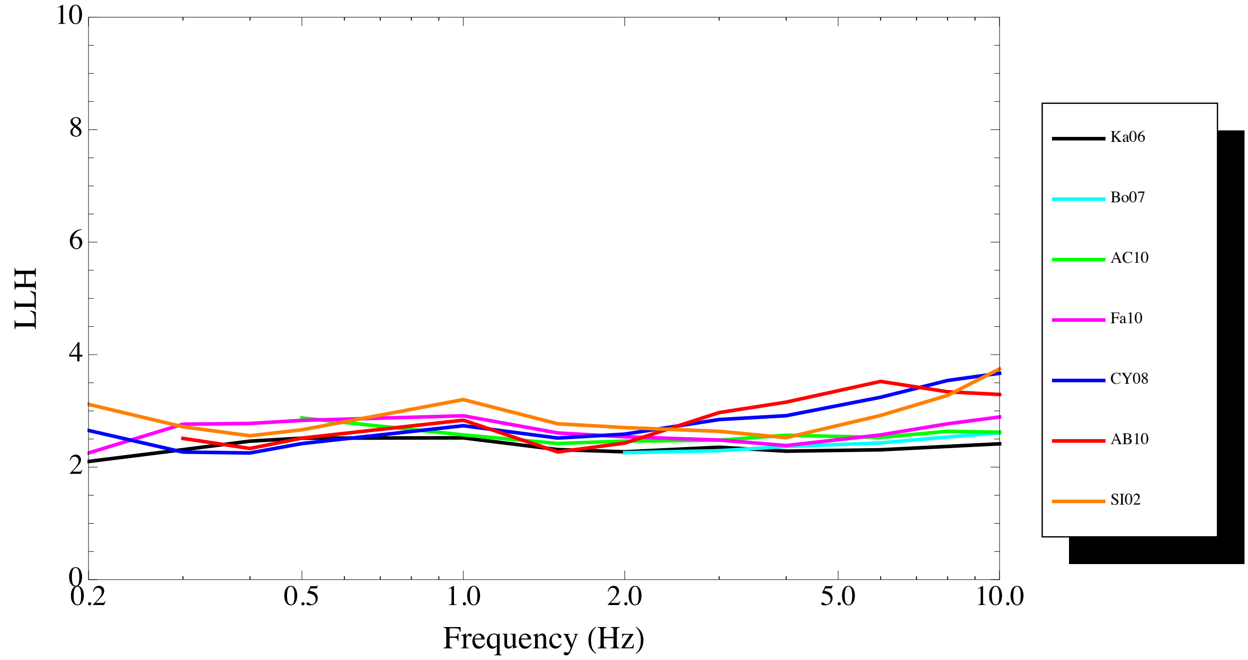


Figure 4: LLH values against response spectra frequency for the GMPEs that get low LLH values for all frequencies.

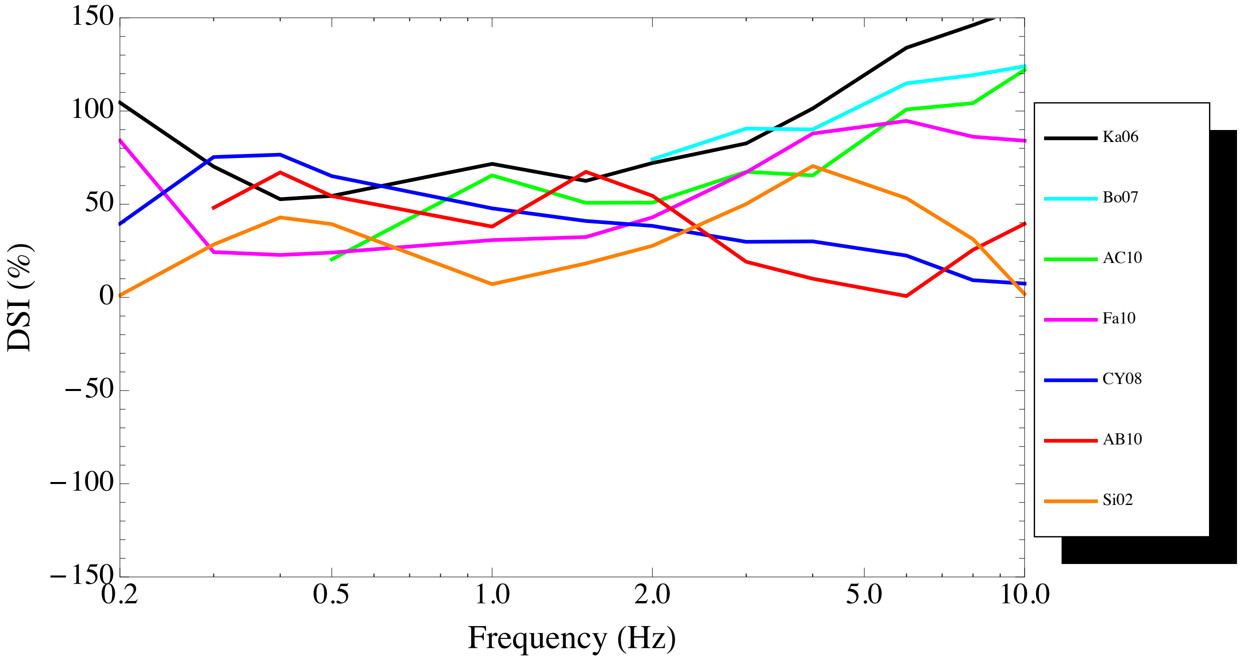


Figure 5: DSI values against response spectra frequency for the GMPEs that get low LLH values for all frequencies.

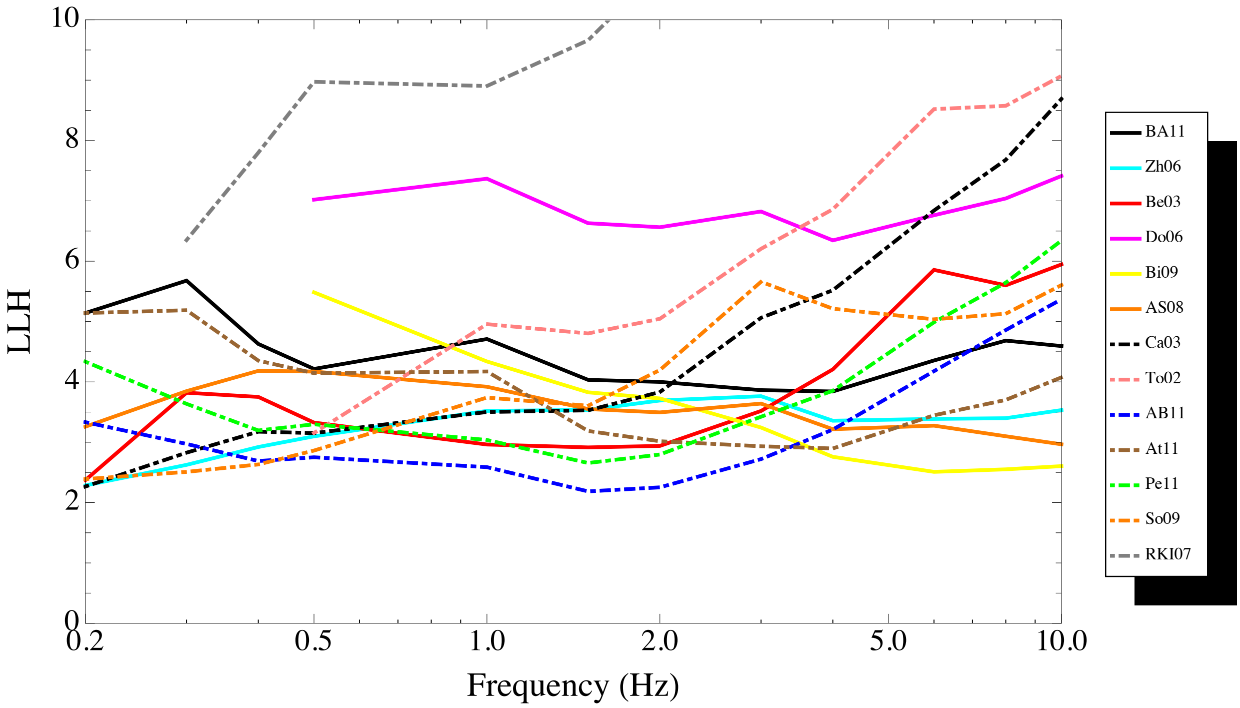


Figure 6: LLH values against response spectra frequency for the GMPEs that get higher LLH values.

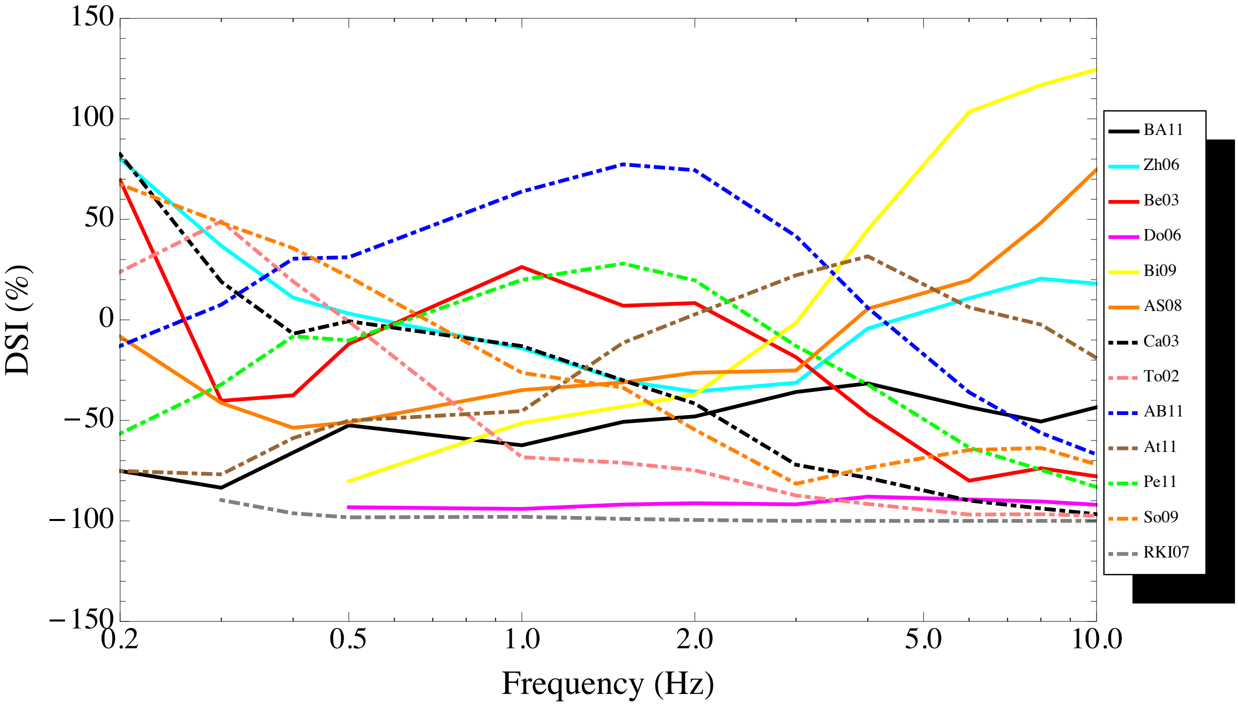


Figure 7: DSI values against response spectra frequency for the GMPEs that get higher LLH values.

The model that presents the lowest LLH values over a large range of frequencies is the Japanese model of Kanno *et al.* (2006), although its low boundary in magnitude is only 5.2. The other models that obtain low LLH values as well are the models of Bommer *et al.* (2007) [EME], Akkar and Cagnan (2010) [Turkey], Faccioli *et al.* (2010) [Japan], Chiou and Youngs (2008) [California], Akkar and Bommer (2010) [EME] and the model for SCR Silva *et al.* (2002). These models have been developed for different regions and only two of them (Bommer *et al.*, 2007 and Akkar and Cagnan, 2010) have been developed for magnitudes down to 3.8. The DSIs for these models all are positive, showing that data favor these models. Note that the LLH values obtained by these GMPEs range between 2 and 3. This means that they predict the observations only approximately well. The fact that GMPEs that differ quite a lot in terms of magnitude and distance ranges and also in region of derivation might tell that no model is able to predict well such low magnitudes in France. More work should be done to better understand the results. As a first step, one can compare predictions to observations in function of distance. Figure 18 shows PSA at 1Hz predicted by Kanno *et al.* (2006) and by Boore and Atkinson (2011) in function of distance, together with the observations. Kanno *et al.* (2006) predict in average well although it cannot reproduce the scatter in the observations. The model of Boore and Atkinson that has an LLH value of about 4.6 under-predicts observations especially for large distances.

Figure 6 shows that the 13 other GMPEs give predictions that are not consistent with the data. This is especially true for frequencies larger than 1.5Hz. The 7 models for SCR are not able to predict well the observations especially at large frequencies. However, their adjustment to the French rock definition should improve this result. The model of Raghu Kanth and Iyengar (2007) should not be taken into account because it systematically over-predict observations and has a very small standard variation.

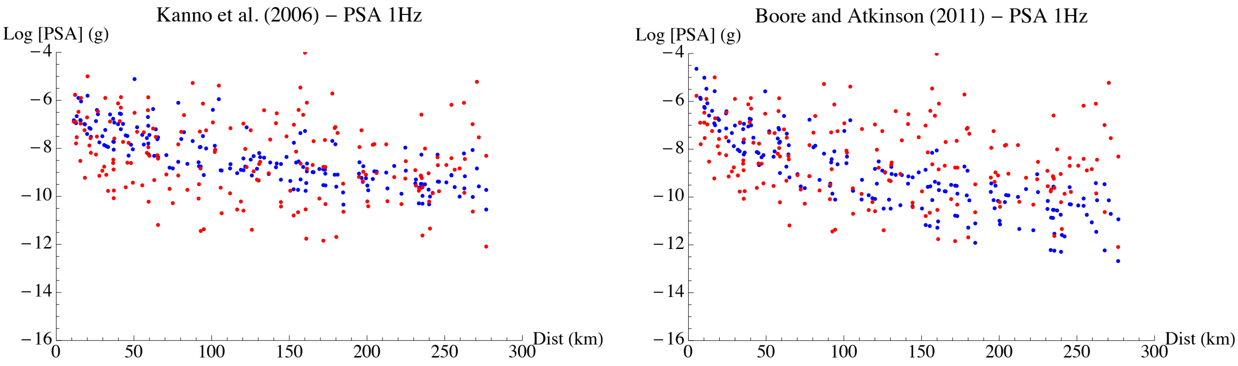
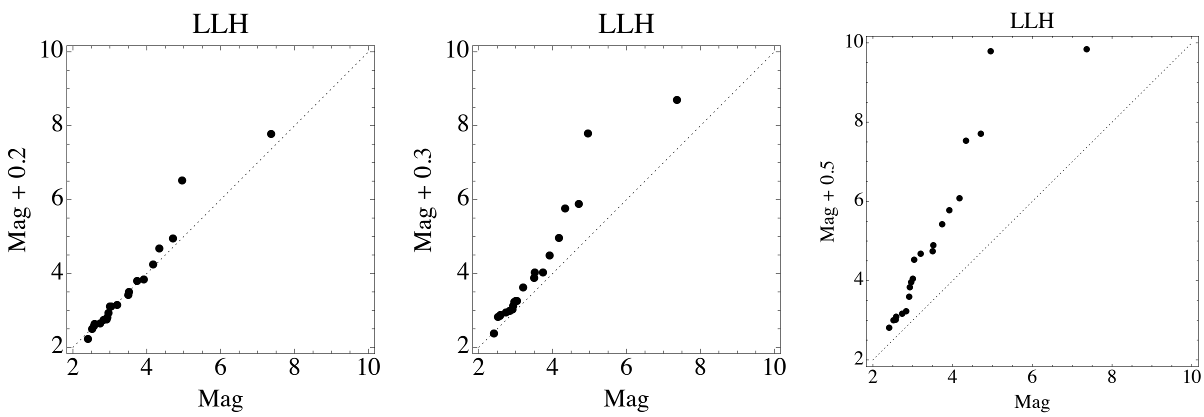


Figure 8: Logarithm of PSA at 1Hz predicted (in blue) and observed (red) against distance for the models Kanno et al. (2006) and Boore and Atkinson (2008).

5. Sensitivity of LLH-based results to source parameters uncertainties

We focus on predictions of PSA at 1Hz to conduct a sensitivity analysis.

- **Magnitude.** In order to assess the influence of magnitude uncertainties, we have computed LLHs by increasing or decreasing the events magnitudes by 0.1, 0.2, 0.3 and 0.5 for PSA at 1Hz. Figure shows the comparison of the LLH obtained by adding or subtracting 0.2, 0.3 or 0.5 to all magnitudes with LLH obtained with the original magnitudes. If no difference was observed, the points would be on the dotted line. What is observed is that by adding or subtracting 0.2 to the magnitudes, LLH values do not change significantly for GMPEs with a LLH lower than 4. Subtracting 0.3 gives significant differences for almost all GMPEs. With a change in 0.5 points, the results are very different. Note that these changes lead to an increase of the LLH values. This would mean that the original magnitudes that we are using are good estimates. Note also that the GMPEs that get the highest LLH values are more sensitive to magnitude uncertainties than the GMPEs that get low LLHs.



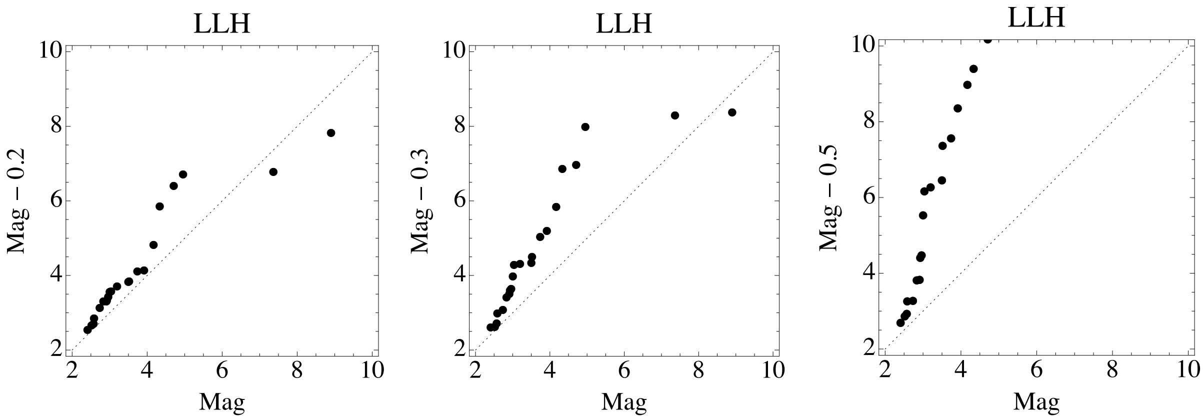


Figure 9: LLH values obtained by adding (up) or subtracting (down) 0.2, 0.3 or 0.5 to the events magnitudes in function of the LLH values with the original magnitudes.

- **Depth.** LLHs have been computed by increasing the depth of all events by 5 km and by 10 km for PSA at 1Hz. This modification has an influence only for GMPEs that use the hypocentral or the Joyner-Boore distance (which we suppose equal here). With an addition of 5 km and even of 10 km, rankings remain stable. This can be explained by the small amount of sites that are close to the events. No site is closer than 10 km to the sources, only 12 over 182 sites stand between 10 km and 20 km to the sources.

- **Focal mechanism.** LLHs have been computed for three cases: all events have a normal faulting or a reverse faulting or a strike slip faulting. No significant differences have been found between the three cases for PSA at 1Hz. The largest differences are obtained for the normal faulting.

- **Vs30 determination**. Two NGA GMPEs have an option that accounts for the way the Vs30 values have been derived, the models of Abrahamson and Silva (2008) and Chiou and Youngs (2008). They decided to make a distinction between estimated Vs30s and measured Vs30s because they observed a greater variability in the residuals for sites with Vs30 values inferred from geology than for measured Vs30 values. This distinction has an influence only on the standard deviation of the model. For the model of Chiou and Youngs (2008), we found no significant differences between LLHs derived from the option "estimated Vs30" and LLHs derived from the option "measured Vs30" for PSA at 1Hz. However, for the model of Abrahamson et al. (2008), we found an increase of 0.2 in LLH when using the option "estimated Vs30".

6. Conclusion and next steps

These first results show the discrepancies between candidate GMPEs to predict French ground motion at small magnitudes.

With this study, I have identified 7 GMPEs over 20 that are able to predict reasonably well French recordings for the set of frequencies 0.2Hz - 10Hz. These GMPEs differ a lot in terms of magnitude, distance and region. They also get a LLH value that is not particularly low (between 2 and 3). I still do not fully understand why such GMPEs get such similar results. Therefore, I will work on deriving tools associated with different approaches (LLH, residuals, Kullback-Leibler divergence, aggregation of models, ...) in order to gather as much information as possible about GMPEs prediction characteristics.

SCR models, except the model of Silva *et al.* (2002) appeared to be inappropriate to predict the present French ground motion. I will apply an adjustment to French rock definition. This should improve their prediction.

Magnitudes variations appeared to have a strong influence on LLH results. The uncertainties associated with the magnitude that can be large for historical events will have to be taken into account.

Because of magnitude scaling, GMPEs that are able to predict well ground motion for small magnitudes will not automatically be able to predict well ground motion at larger magnitudes. This is why the LLH method will be applied to macroseismic intensities that have been collected for large historical earthquakes.

References

Abrahamson N. and W. Silva (2008). Summary of the Abrahamson & Silva NGA Ground-Motion Relations. *Earthquake Spectra* 24, 1, 67–97.

Akkar, S. and Z. Cagnan (2010). A local ground-motion predictive model for Turkey and its comparison with other regional and global ground-motion models, *BSSA* 100(6), 2978-2995.

Akkar, S. and J.J. Bommer (2010). Empirical equations for the prediction of PGA, PGV and spectral accelerations in Europe, the Mediterranean and the Middle East, *Seismological Research Letters* 81(2), 195–206.

Atkinson, G. M. (2008). Ground-motion prediction equations for eastern North America from a referenced empirical approach: Implications for epistemic uncertainty. *Bull. Seism. Soc. Am.* 98,1304-1318.

Atkinson, G. M. and D. M. Boore (2006). Earthquake ground-motion prediction equations for eastern North America. *Bull. Seism. Soc. Am.* 96, 2181-2205.

Atkinson, G.M. and D.M. Boore (2011b). Modifications to existing ground-motion prediction equations in light of new data, *Bull. Seism. Soc. Am.* 101(3), 1121–1135.

Beauval, C., H. Oksuz, A. Laurendeau, E. Delavaud, F. Cotton, Ph. Guéguen and N. Kühn. On the testing of ground-motion prediction equations against small magnitude data. *Bull. Seism. Soc. Am.*, recommandation: minor revisions.

Berge-Thierry, C., F. Cotton, O. Scotti, D. A. Griot-Pommera, and Y. Fukushima (2003). New empirical response spectral attenuation laws for moderate European earthquakes. *J. Earthquake Eng.* 7(2), 193-222.

Bommer, J. J., P. J. Stafford, J. E. Alarcón and S. Akkar (2007). The influence of magnitude range on empirical ground-motion prediction. *Bull. Seism. Soc. Am.* 97(6), 2152–2170.

Bommer, J. J., J. Douglas, F. Scherbaum, F. Cotton, H. Bungum, and D. Fäh (2010). On the Selection of Ground-Motion Prediction Equations for Seismic Hazard Analysis. *Seism. Res. Lett.* 81, 783-793.

Bommer, J. J., P. J. Stafford, J. E. Alarcon, and S. Akkar (2007). The Influence of Magnitude Range on Empirical Ground-Motion Prediction, *Bull. Seismol. Soc. Am.* 97, no. 6, 2152–2170.

Boore, D. and G. Atkinson (2008). Ground Motion Prediction Equations for the Average Horizontal Component of PGA, PGV, and 5%-Damped PSA at Spectral Periods between 0.01 s and 10.0 s. *Earthquake Spectra*, 24, 99-138.

Burnham, K. P. and D. R. Anderson (2002). *Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach, 2nd edition*. Springer.

Campbell, K. W. (2003). Prediction of strong ground motion using the hybrid empirical method and its use in the development of ground-motion (attenuation) relations in eastern North America. *Bull. Seism. Soc. Am.* 93, 1012–1033.

Chiou, B., R.R. Youngs, N. A. Abrahamson, and K. Addo (2010). Ground-Motion Attenuation Model for Small-To-Moderate Shallow Crustal Earthquakes in California and Its Implications on Regionalization of Ground-Motion Prediction Models. *Earthquake Spectra* 26(4), 907-926.

Cotton, F., F. Scherbaum, J. J. Bommer, and H. Bungum (2006). Criteria for selecting and adjusting ground-motion models for specific target regions: Application to Central Europe and rock sites. *J. Seism.* 10, 137-156.

Cotton, F., G. Pousse, F. Bonilla, and F. Scherbaum (2008). On the Discrepancy of Recent European Ground-Motion Observations and Predictions from Empirical Models: Analysis of KiK-net Accelerometric Data and Point-Sources Stochastic Simulations. *Bull. Seism. Soc. Am.* 98, 2244-2261.

Delavaud E., F. Cotton, S. Akkar, F. Scherbaum, L. Danciu, C. Beauval, S. Drouet, J. Douglas, R. Basili, A. Sandikkaya, M. Segou, E. Faccioli, and N. Theodoulidis (2012b). Toward a Ground-Motion Logic Tree for Probabilistic Seismic Hazard Assessment in Europe. *J. Seism*. Accepted.

Delavaud, E., F. Scherbaum, N. Kühn, and T. Allen (2012a). Testing the Global Applicability of Ground Motion Prediction Equations for Active Shallow Crustal Regions. *Bull. Seism. Soc. Am.* Accepted.

Delavaud, E., F. Scherbaum, N. Kuehn and C. Riggelsen (2009). Information-Theoretic Ground-Motion Prediction Equation Selection for Seismic Hazard Analysis: An Applicability Study Using Californian Data. *Bull. Seism. Soc. Am.* 99(6), 3248-3263.

Faccioli, E., A. Bianchini and M. Villani (2010). New ground motion prediction equations for T > 1s and their influence on seismic hazard assessment. *Proceedings of the University of Tokyo Symposium on Long-Period Ground Motion and Urban Disaster Mitigation, Tokyo, March 17-18, 2010.*

Kaklamanos, J., L. G. Baise and D. M. Boore (2010). Estimating Unknown Input Parameters when Implementing the NGA Ground-Motion Prediction Equations in Engineering Practice. *Earthquake Spectra* 27(4), technical note.

Péquegnat, C., P. Guéguen, D. Hatzfeld, and M. Langlais (2008). The French Accelerometric Network (RAP) and National Data Center (RAP- NDC), *Seism. Res. Lett.* 79(1), 79-89.

Pezeshk, S., A. Zandieh and B. Tavakoli (2011). Hybrid Empirical Ground-Motion Prediction Equations for Eastern North America Using NGA Models and Updated Seismological Parameters, *Bull. Seism. Soc. Am.* 101(4), 1859-1870.

Scherbaum F., E. Delavaud and C. Riggelsen (2009). Model Selection in Seismic Hazard Analysis: An Information-Theoretic Perspective, *Bull. Seism. Soc. Am.* 99, 6, 3234–3247.

Toro, G. R. (2002). Modification of the Toro et al. (1997) Attenuation Equations for Large Magnitudes and Short Distances. *Risk Engineering Inc. report*. Unpublished.

Toro, G. R., N. A. Abrahamson, and J. F. Schneider (1997). Model of strong ground motions from earthquakes in central and eastern North America: best estimates and uncertainties. *Seism. Res. Lett.* 68, 41-57.

Wells and Copersmith (1994). New Empirical Relationships among Magnitude, Rupture Length, Rupture Width, Rupture Area, and Rupture Displacement. *Bull. Seism. Soc. Am.* 84(4), 974-1002.

Zhao, J. X., J. Zhang, A. Asano, Y. Ohno, T. Oouchi, T. Takahashi, H. Ogawa, K. Irikura, H.K. Thio, P.G. Somerville, Y. Fukushima, and Y. Fukushima (2006). Attenuation relations of strong ground motion in Japan using site classifications based on predominant period, *Bull. Seism. Soc. Am*. 96, 3, 898-913.

Project SIGMA

Review of Deliverable D2-12
Methodology and Testing of GMPEs based on ground-motion data from France and surrounding countries

(Ref : SIGMA-2011-D2-17, 24/11/2011)

Jean B. Savy
November 3, 2011

1. General comments

D2-12 is a very short document that, in the present states only describes a method of ranking of ground motion prediction equations. As I understand it the scope of this task goes beyond simply ranking, as it is expressed by the authors themselves in the reviewed document, as follows:

"The goal of the task X2-4 of work package 2 is to select and rank ground-motion prediction equations (GMPEs) for seismic hazard assessment (SHA) in France and surroundings countries. It is needed to establish a logic tree for ground motion prediction that captures well epistemic uncertainty"

(taken from document being reviewed for deliverable D2-12). It follows the work documented in D2-5. (Ref.SIGMA-2011-D2-16,version1)

The final product of the selection and evaluation of GMPEs for use in SHAs in project SIGMA is a set of equations that predicts ground motion in the form of a mean value (or mean log) and a characterization of the aleatory uncertainty, as function of magnitude, some measure of distance, and a few other parameters.
It is customary to identify independent classes of models, and different variations of models within a class, and to assign weights to each class, and to each model in a class, for the purpose of characterizing epistemic uncertainty, that has been chosen, as it is the common practice, to be represented by a logic tree in the project.

Although there is no unbiased way to effectively determine the weights objectively, there are objective ways to rank the models, such as with Sherbaum's method (Sherbaum et al. 2009) based on use of recorded data. There is still a need to go from ranks to weights in a manner that makes sense, probably a combination of information on ranking, limitations of the models and some expert judgment.

Aside for the form of the document which I find excellent, I have only the following comments or questions on this work, and I have suggestions for the presentation of November 17-18th.

1. First, I can only support the use of an objective method such as Sherbaum et al. method (2009) to inform analysts on the relative relevance of models with respect to observed data.

2. The ranking method is properly applied and it gives valuable results.

3. Since this is based on the pre-selection as documented in D2-5, why do we have a different number of equations. (Table 1 of D2-5 has 12 equations, and Table 2 has 10., but D2-12 only considers 14 equations)?

4. Which events are going to be used in the macro-seismic study, and what are the plans for it?

5. How are the weights to be assigned to the branches of the logic tree going to be determined? What are the detailed plans for this task?

2. Recommendations for the November 18 presentation

What is missing in the reviewed document is a presentation of a way forward. There is no explanation of how the ranking is going to be used. It may be that it is intended to be part of following work, but it is important to have a clear understanding as soon as possible of the directions to take.

1. For this reason I strongly suggest that an effort be made to present the plans for the determination of the weights on the logic tree of ground motion prediction equations at the coming meeting of CS2.
   If the plans are not fully formulated, it is still important that we can discuss general plans and alternatives methods at CS2.

2. Please give some details on the plans to use macro seismicity. What events will be used, what results, are expected, how those results will be used in the context of updating the ranking and determining the weights.

Respectfully submitted, November 3, 2011.

Jean Savy

733 Arimo Avenue, Oakland, CA 94610
(510) 502-3249 * (510) 834-1394 (Fax)
Jean.savy@att.net

Review of the SIGMA Deliverable D2.12 (Ref : SIGMA-2011-D2-17-V01)

" Methodology and Testing of GMPEs based on ground-motion data from France and surrounding countries "

(Author : E. Delavaud, ETHZ/SED, 01/11/2011)

This report presents a methodology and the first results for an "objective" ranking of a number of pre-selected GMPEs for their application to hazard assessment studies in the continental territory of France. This ranking is based on a measure of the fit between the considered GMPEs and the available accelerometric data in France.

The 11-page, concise report consists of four sections

- It starts with a short presentation of the used RAP data set
- It then lists the 14 pre-selected GMPEs
- The next and main section presents briefly the testing / ranking technique based of the "LLH" technique as proposed by Scherbaum et al. (2009), and its results when applied to the considered data and GMPE sets.
- The last section shortly wraps up the main results and lists a few open issues

The idea to have a method for an objective ranking is a major breakthrough compared with the previous practice where the weighting of GMPEs was based mainly on expert judgment. It thus deserves a careful presentation and discussion. The following comments and questions intend to improve the presentation of this report, which should be one of the key elements not only for the SIGMA PSHA studies in South-Eastern France, but also for all future hazard assessment studies in France.

The data set :

Testing GMPEs against a given data set requires the knowledge, for each earthquake and recording site, of the ***metadata*** that are used in the selected GMPEs. In order for the present work to be reproducible, it would be good to list, at least in an annex, the value of all the metadata associated with each recording or earthquake, and to discuss the procedure to estimate those which are not provided in standard seismological catalogs. It might also propose directions for the future enrichment of the RAP data base with appropriate metadata.

- Source parameters :
    - M_w was probably derived from the work by Drouet et al. (2010). I am wondering whether it is an issue to use Mw values which are not independent of the used data set.
    - The magnitude of the considered events is limited, and so is the corresponding rupture area. Therefore, some of the used parameters (such as Z_{TOR}) are tightly linked to the estimate of the hypocentral depth, which has a limited precision: a discussion on the sensitivity of used GMPEs to source parameter uncertainties would be welcome
    - Focal mechanism : about 45% (8/18) of the used events have unknown focal mechanisms : how does this affect the results of the LLH technique ?
- Site conditions :
    - the site classes and/or V_{S30} values were taken from the RAP site; there exists some concern about their reliability (there is a special working group dedicated to checking these values), and the stations used and the associated number of recordings should be at least listed in a Table to allow further checks when checked site conditions will be available.
    - Some of the used GMPEs include a distinction between measured and inferred V_{S30} values: how is this practically implemented for the LLH estimates?
    - How to estimate the value of $Z_{1.0}$ required by some GMPEs (AS2008)?
- Usable bandwidth: some of the used recordings have been obtained at large distances (several hundred kilometers) from moderate magnitude events. Even though the RAP instrumentation is sensitive, and long period response spectra may indeed be related to intermediate frequencies, it is needed in my opinion to check the reliability of such recordings especially in the long period range, for instance by investigating the signal to noise ratios over the whole frequency range.

The tested GMPEs :

- 14 candidate GMPEs are considered, 5 for stable continental regions and 9 for shallow active crustal regions. This pre-selection list is different from those considered in D2.5 (indeed it is mainly a subset, except for Faccioli et al. 2010), and updates of the two reports D2.5 and D2.12 should include either the same list, or explain why they are different (with a marked preference for the first option).
- In addition, as some of the selected GMPEs require some of the adjustments listed in Deliverable D2.5 (horizontal component, magnitude, possibly style-of-faulting, …), it is needed to know which value of sigma is considered : the original one or the adjusted (increased) one ?

The LLH method:

I have the – may be wrong - feeling that the present, very concise report is certainly meaningful for the highly specialized GMPE community, but is too short to allow a full understanding of this new technique by a wider community that is mainly interested in simply using them for hazard assessment purposes. A longer presentation and explanation, as for instance in Beauval et al. (2011), would certainly help the "naïve" reader in understanding the principle of the LLH approach, and get a feeling of the LLH and DSI values that are expected as a function of the difference between the actual distribution of data, and those of the tested GMPEs.

I have a few other, more technical comments and questions:

- A similar testing exercise is reported in Beauval et al (2011) with apparently the same data set and an overlapping GMPE set, with however somewhat different end results (despite some similarities), which emphasizes the need for a discussion about the robustness of LLH results.
- Some of the tested GMPEs are theoretically valid only for limited distance and frequency ranges. While extrapolating at larger distances seems straightforward, I do not understand how the Bommer et al. (2007) GMPE can be tested at frequencies below 2 Hz. Some further details are needed, especially as this GMPE turns out to be the "best" one.
- The Faccioli et al. (2010) also gets a good ranking score: this extension of Cauzzi & Faccioli (2008) should thus definitely be described in more detail in the companion report D2.5.
- Is it conceivable, and would it be useful, to include the partitioning between intra and inter event variability in the LLH ranking, to better qualify the nature of the fit ? In any case, the plots (in an annex) of the the observed residuals between each selected GMPE and the actual RAP data, as a function of magnitude, distance and may be site conditions, would be informative on the origin of the misfit.
- What is the "confidence interval" for LLH (or DSI) values? In other words, what may be considered as the minimum threshold for a statistically meaningful difference between the LLH values of two different GMPEs, which might then be used for assigning different weights to these two GMPEs ?
- The application of the same ranking method to two different data sets gives some hints on the robustness of the approach with respect to the data set, but does not tell much about the expectations for larger magnitude events, especially in the near source area where one expects damaging strong motion, for which the available data set is rather poor, as displayed in Figure 1. Some further developments / discussions on this issue would be useful to avoid any automatic, may be wrong, extrapolation of the present results to larger magnitude, less frequent events.

Main recommendation

I anticipate this report to be a key step for future hazard studies in France through its conclusions on GMPE ranking. I therefore recommend a significant update of the report in order to ensure a totally reproducible reasoning on the basis of fully transparent data and metadata, and address all the consistency issues with other SIGMA reports and parallel publications.

Grenoble, 14/11/2011

Pierre-Yves BARD