CEIDRE 1107307s 28/02/2013
D13A0227069

Σigma
Seismic **G**round **M**otion **A**ssessment

RANKING OF AVAILABLE GMPEs FROM RESIDUAL ANALYSIS FOR NORTHERN ITALY AND DEFINITION OF REFERENCE GMPEs.

AUTHORS			REVIEW			APPROVAL		
NOM	DATE	VISA	NOM	DATE	VISA	NOM	DATE	VISA
F. Pacor L. Luzi M. Massa (INGV Milano) D. Bindi (GFZ Potsdam)	21/12/2012		F. Scherbaum Membre of the SIGMA Scientific Committee	IN Written review attached		R. Figini Sigma Project Leader	27/02 2013	P.O.
			J. Savy Membre of the SIGMA Scientific Committee	IN Written review attached				

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

Summary

TABLE OF CONTENTS

List of acronymics

GMPE – Ground motion Prediction Equation

LLH- Log Likelihood parameter introduced by Scherbaum et al (2009) to quantify the goodness of fit of a GMPE with respect to observations

DBN – Strong motion data base for northern Italy

DBNG – a set of five strong motion data bases for northern Italy selected from DBN accordingly to the range of validity of the candidate GMPEs

DBN_E - subset of DBN including only records from 2012 Emilia sequence

DBNG_E – the same as DBNG but considering only the 2012 Emilia sequence

DBN_NE - subset of DBN excluding records from 2012 Emilia sequence

DBNG_NE – the same as DBNG but excluding the 2012 Emilia sequence

ITA10 – GMPE derived by Bindi et al (2010)

AB10 – GMPE derived by Akkar and Bommmer (2010)

BEA08 – GMPE derived by Boore and Atkinson (2008) including the modification of Atkinson and Boore (2011)

CF08 – GMPE derived by Cauzzi and Faccioli (2008)

MS08 – GMPE derived by Massa et al. (2008).

1. Introduction

The selection of the Ground Motion Prediction Equations (hereinafter GMPEs) appropriate to a specific regional geologic context is one of the crucial issues of any seismic hazard evaluation. This deliverable describes the activities carried out to select and rank the most appropriate GMPEs for northern Italy, with special emphasis on the Po plain area, which is one of the largest sedimentary basin in the world with an area of about 50.000 km^2 and a sediment thickness varying from few tens of meters to about 8 km. The Po plain geologic setting can potentially affect the ground motion, causing amplification of the long period motion and lengthening of the signals, as observed for other basins in the world, e.g. the Los Angeles basin (Joyner, 2000), the Osaka Basin in Japan (Kagawa et al., 2004) and the Kanto basin beneath the city of Tokyo (Hisada et al., 1993; Sato et al. 1999).

The seismicity of North Italy is generally characterized by the occurrence small energy events (about 200/year) and rare moderate earthquakes. For this reason, the ground motion of the area is poorly characterized and very few studies have been carried out on the ground motion attenuation (Massa et al., 2008; Castro et al., 2008). The occurrence of the 2012 Emilia sequence (Mw 6.1 on 20 May and Mw 5.9 on 29 May 2012) provided thousands of strong motion records relevant to moderate and strong magnitudes with 6 events with magnitude larger than 5 (Luzi et al., 2012; submitted). The Emilia dataset, together with the strong motion data available from past events occurred in northern Italy, hereinafter called DBN, offers the opportunity of investigating in detail the characteristic of the ground motion in the magnitude range 3.5 - 6.4 (1976 Friuli earthquake). DBN is used to select and rank a set of 5 candidate GMPEs, which are developed from global datasets or specifically derived for Italy. The data-driven method that has been recently developed by Scherbaum et al. (2009) is adopted to quantitatively judge the applicability of GMPEs in this region. A residual analysis is subsequently carried out in order to evaluate the features of ground motion attenuation in terms of magnitude and distance scaling.

2. Strong-motion dataset

To construct the DBN dataset, we select Italian stations and earthquakes which are in a geographical area defined by the latitude [43°30'N and 46° 30'N] and longitude [8°00'E and 13° 50'E] . The DBN dataset is composed of 1440 records relative to 137 events and 225 stations (Figure 1a) and about 50% of the records are relative to the Emilia sequence. The magnitude range is 3.5 – 6.4 (1976 Friuli earthquake) and the epicentral distances span from 3.5 to 250 km. The events occurred in the upper crust, at depths lower than 30 km.

The stations are mainly located in the northern Apennines, Eastern Alps and central Po plain, while the distribution is poor in the western Alps and western Po plain.

The events are mainly compressive events, with focal depths vary between 5 and 25 km. Two main seismic sequences have been sampled, the Friuli 1976-1977 (to the north-east in Figure 1) and the Emilia 2012 (in the centre of the Po plain, Figure 1); two minor sequences are the Mw 5.4 Parma 2008 (northern Apennines) and Mw 5.2 Garda lake 2004; other sparse events are localized in the northern Apennines.

The strong-motion data before 2006 mainly come from the National Accelerometric Network (Rete Accelerometrica Nazionale, RAN) operated by the Department of the Civil Protection (Dipartimento della Protezione Civile, DPC, www.protezionecivile.gov.it). The strong motion data recorded till 2009 are available at the site http://itaca.mi.ingv.it, for the following years, on request, at the website of the DPC.

After 2006, the Istituto Nazionale di Geofisica e Vulcanologia (INGV) started its accelerometric network and the strong motion data, including the Emilia seismic sequence, can be obtained at http://www.mi.ingv.it/ISMD/ and http://eida.rm.ingv.it

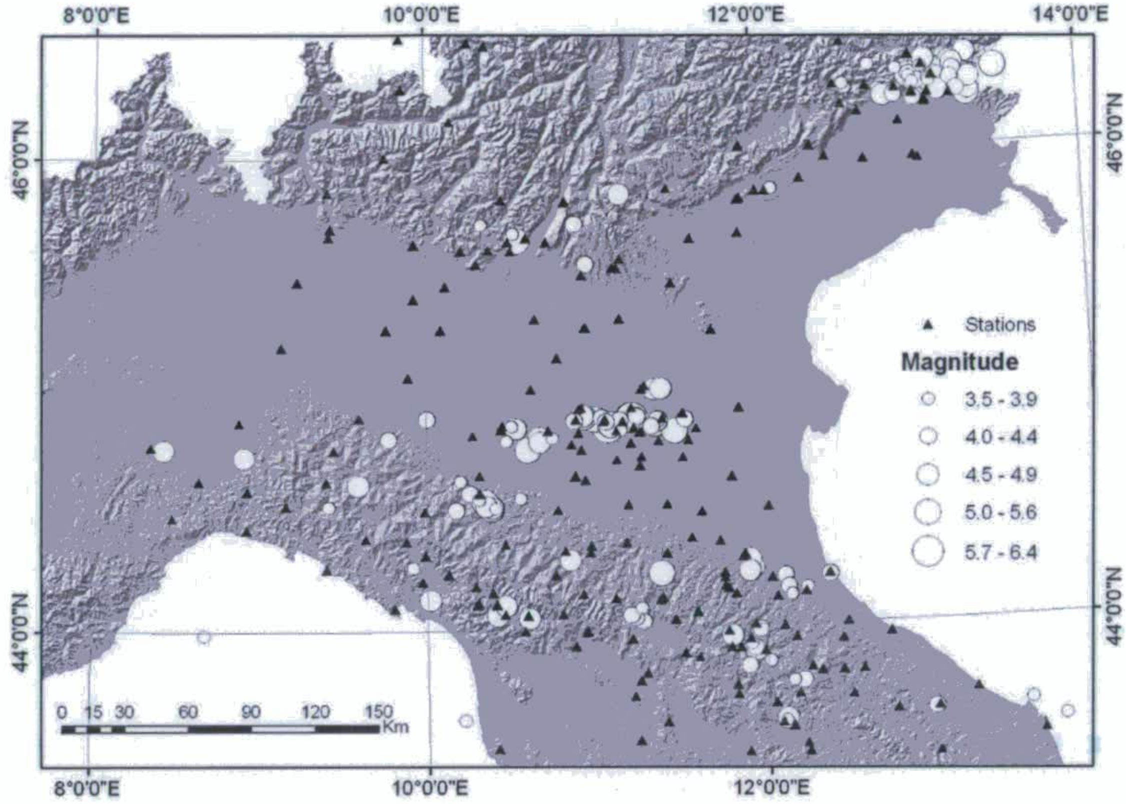


Figure 1. Map of the seismic stations (black triangles) and location of the epicenters (white circles) for DBN.

In Figure 2a, the distribution of records in time is plotted and clearly shows the transition between the analog and digital era, occurring around the year 2000, when digital accelerographs became capable of detecting small ground motion amplitudes.

The selected records are processed uniformly applying the procedure proposed by Pacor et al. (2011), which has also been adopted to process accelerometric data distributed through the ITACA database.

The basic steps of the processing procedure are (Pacor et al., 2011):

– baseline correction (constant de-trending);

– application of a cosine taper, the extension of which is based on the visual inspection of the record (typically between 2 and 5% of the total record length);

– visual inspection of the Fourier spectrum to select the band-pass frequency range; whenever feasible, the same range is selected for the 3-components;

– application of a 2nd order acausal time-domain Butterworth filter to the acceleration time-series;

– double-integration to obtain displacement time series;

– linear de-trending of displacement to avoid an unrealistic constant baseline shift in velocity;

– double-differentiation to get the corrected acceleration.

Before executing the filtering operation, zero pads are added at the beginning and end of the signals, as in Boore and Bommer (2005).

The typical band-pass frequency range is between 0.08 and 40 Hz, for the Emilia strong motion data and around 0.2 – 30 Hz for the rest of the data. Zero pads are then removed from the filtered signal and a procedure is applied in order to guarantee the compatibility between acceleration velocity and displacement by subsequent integration. The acceleration response spectra at 5% damping have been computed in the period range 0.04 – 4 s for each components. The geometric mean and the largest of the horizontal components have been also calculated.

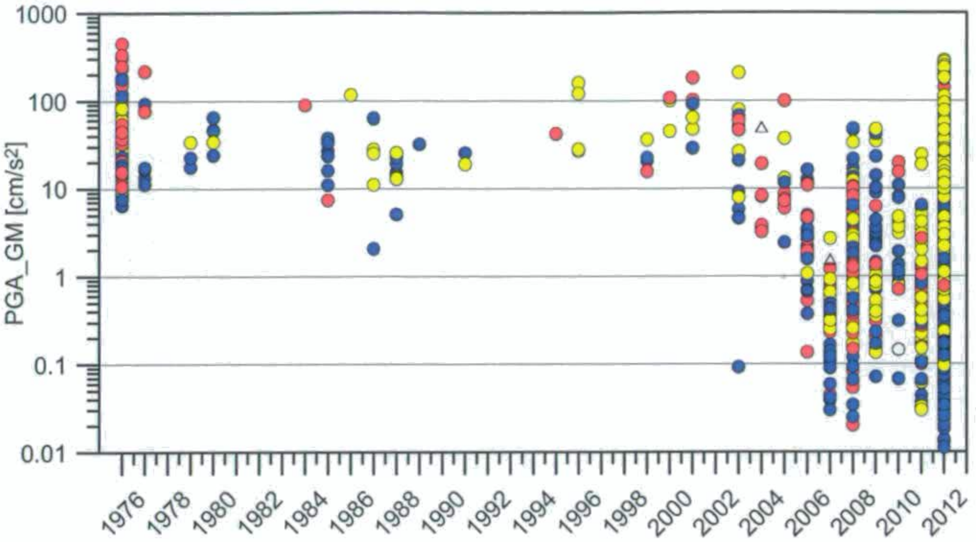


Figure 2a. Annual distributions of the peak ground acceleration (geometric mean of the horizontal components). The data are grouped for site conditions: blue: EC8-A class; red: EC8-B class; yellow: EC8-C class.

Figure 2b shows the distribution of the between-event errors of the DBN earthquakes with M > 4, evaluated with respect to the Italian ground motion equation (ITA10, Bindi et al., 2011). For each event, the error is estimated as the mean of the residuals, computed as the natural logarithm of the ratio between observed and predicted spectral ordinates. Most of the between-event errors varying in the range from −1 to 1 with some exceptions: the Emilia events (red squares) are generally overestimated (negative errors) at short periods, while the main shocks of the Friuli sequence are underestimated both at short and long periods .

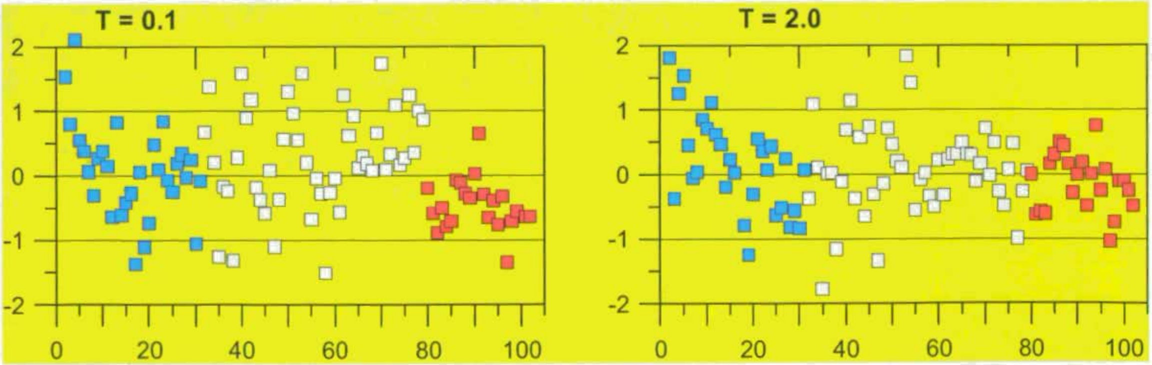


Figure 2b. Between event distribution of errors for acceleration responce spectra at T = 0.1s (left) and T = 2.0s (right) of DBN events with M > 4. The blue and red squares represent the Friuli and Emilia earthquakes, respectively.

2.1. Metadata

The metadata of the seismic events before 2010 were revised during the construction of the ITACA database, consulting the most reliable sources. In particular, event location and local magnitudes mainly come from Bollettino sismico Italiano, http://bollettinosismico.rm.ingv.it/ or Italian seismic catalog (CSI 1.1, http://csi.rm.ingv.it/). Moment magnitudes comes from specific studies or the Regional Centroid Moment Tensor project (http://www.bo.ingv.it/RCMT/). Event location and magnitudes or the major events and causative fault geometries are from specific studies (faults from DISS http://diss.rm.ingv.it/diss/ e.g., De Natale et al. (1987) for the events of the Friuli sequence). The most recent data (2011-2012) are from ISIDe (http://iside.rm.ingv.it/) as the Bollettin Sismico Italiano is going to be released in year 2012 (thus at the end of 2012).

About one third of the records of DBN is characterized by Mw and the rest by Ml, especially the low magnitude ones (< 5.0). Many of the selected GMPEs make use of moment magnitudes (see Table 1), therefore in order to apply the GMPEs consistently with the explanatory variables, we test the use of an empirical relationship to convert M_l in Mw.

The scatter between Ml and Mw, mainly at small magnitudes, is usually large and the resulting scaling relations are therefore uncertain. Empirical studies show that relation of Ml versus Mw is characterized by Ml ~ mMw, with the value of m being close to 1.5 for small (Ml ≤ 3:5) events and m ~ 1 for larger magnitudes. However the scale factor m in variable, depending on magnitude range, geographical region, and algorithm for the calculation of Mw (Bethmann et al., 2011).

In this study, we applied the equation proposed by Castello et al (2007):

$$M_w = 0.79(\pm 0.4) \cdot M_l + 1.20 \tag{1}$$

which is calibrated on an Italian data set and it is valid in the local magnitude range (3.5 – 5.8). As shown in Figure 3a, the difference between Ml and Mw converted with equation 1 decreases as the local magnitude increases and for values larger than 5.5, its effect is negligible.

Figure 3b and 3c shows the magnitude distribution before and after the conversion. The distribution is modified at lower magnitudes (M < 5.5): the number of records in the range 3.5 – 4.0 diminishes from about 200 to 50 while in the range 4.5 – 5.0, increases from 250 to 450. In the following analysis we use the dataset both with original and converted magnitude scale.

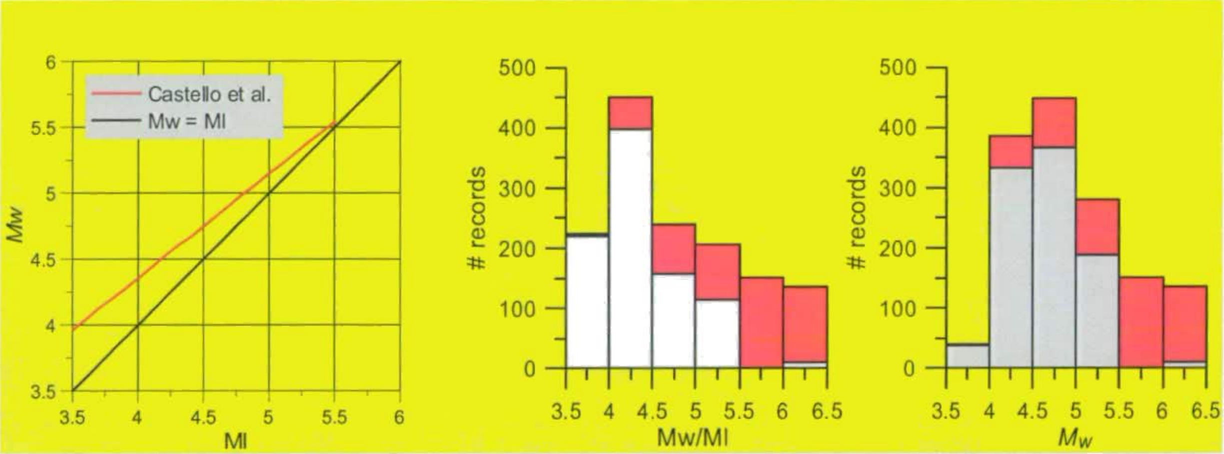


Figure 3. a) The Castello et al. (2007) relationship (red line) compared to the relation Ml = Mw (black line. b) Histogram of records included in included in the Northern Italy strong-motion dataset., grouped for magnitude as available in the dataset (Mw - red bar or Ml – white bar); c: histogram for Mw magnitude, obtained converting the Ml in Mw through the Castello et al. (2007) empirical relationship. Grey bars represent the number of strong-motion records of events associated to converted Mw.

Figure 4 displays the magnitude – distance distribution of the data set.

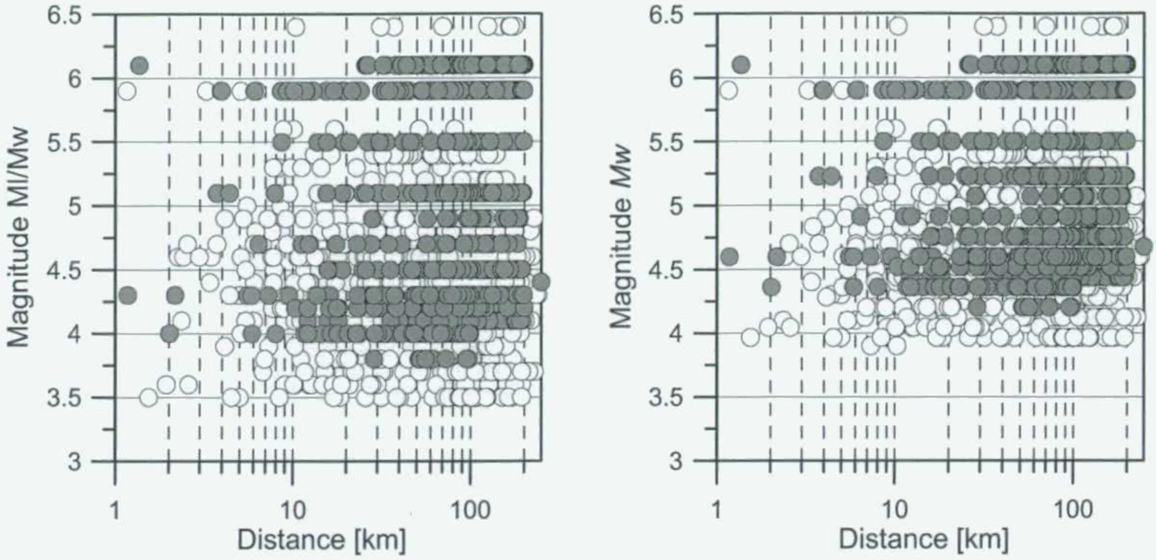
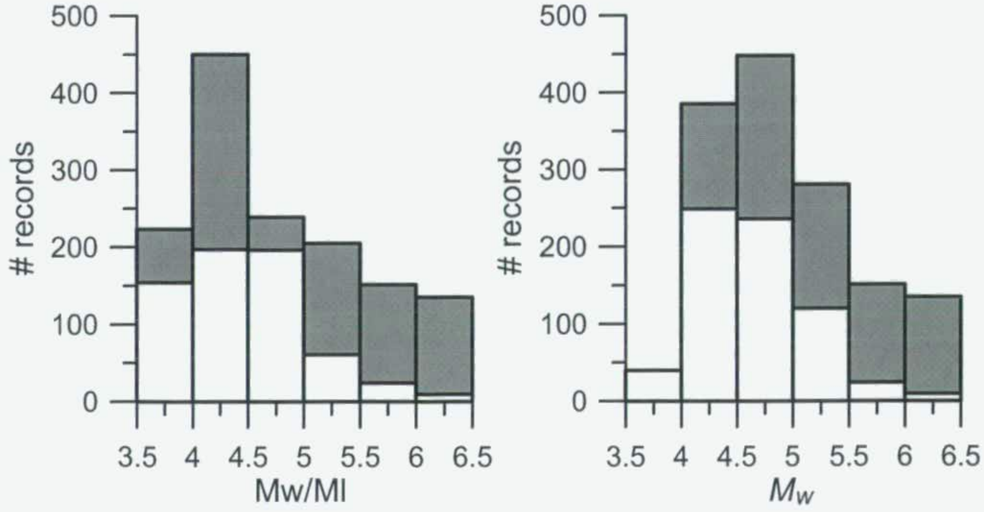


Figure 4. Magnitude-distance distribution for DBN. Left: Mw or Ml are used; Right: Mw or Mw converted in Ml are used.

Figure 5 shows the distribution of hypocentral depths, style of faulting and site classes according to the EC8 classification. The records come mainly from superficial seismic sources, in the uppermost 15 km of the crust and the prevalent focal mechanism is thrust fault. The majority of the sites have been classified according to a geological description and only a subset of them has a direct measure of the shear and compressive wave velocity.



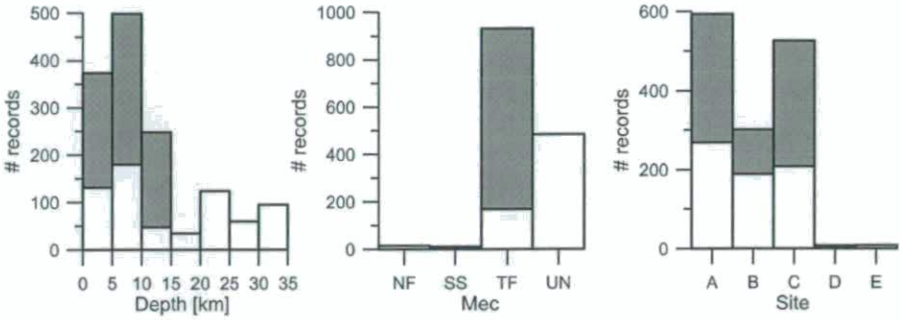


Figure 5. Depth, focal mechanism and site class distributions of the records included in the Northern Italy strong-motion dataset (DBN). Grey bars indicate the number of strong-motion records relative to the 2012 Emilia seismic sequence. NF: Normal fault; SS: Strike fault; TF: thrust fault; UN: unknown mechanism.

3. GMPE ranking and selection

To select a set of GMPEs to be evaluated with respect to the DBN, we exploit the results obtained in the SHARE project (2009) where, for each seismotectonic region of Europe, the smallest set of GMPEs to capture the uncertainty in ground-motion prediction was identified.

For active shallow crustal region (ASCR), the list includes (Delevauld et al., 2012) 4 models: the Akkar and Bommer (2010) model based on European dataset and 3 global models, i.e. Cauzzi and Faccioli (2008), the Zhao et al. (2006) and the Chiou and Youngs (2008) equations. While the first two were supported by both the experts and the empirical data testing, the other two were included based on expert judgment and data-testing on non-European data set.

From the set of GMPES identified for ASCR, we select the two models that could be directly applied to DBN dataset, i.e. the Cauzzi and Faccioli (2008) model, hereafter CF08; and the Akkar and Bommer (2010) equation, hereafter AB10. This choice has been done to avoid to use conversion factors for assign some explanatory variables (for example the fundamental period of the site, the rupture distance from the fault) of the more complex models and missing in DBN.

In addition to these GMPEs, we consider the Boore and Atkinson (2008) equation, including the modification by Atkinson and Boore (2011) equation, hereafter BEA_M, based on global dataset but extended toward the lower magnitudes and for this reason, suitable for the context of the Po plain region.

Finally, we include two regional GMPES developed for Italy, the Bindi et al. (2011), hereafter ITA10, based on the Italian strong motion dataset, recorded up to 2009 and the Massa et al. (2008), hereafter MS08, based on a data set specific for Northern Italy, recorded up to 2007.

Table 1 lists the five candidate GMPEs and the specific range of validity. Different distance metrics are used by the selected GMPEs: CF08 uses the hypocentral distance, MS08 the epicentral distance and ITA10, AB10 and BEA_M use the minimum distance from the surface projection of the fault, the so called Joyner-Boore distance. In general, all of them include similar definition of the style of faulting. With the exception of MS08, which includes deep events, the rest of the GMPEs is calibrated on crustal events. The site classification is based on the shear wave velocity in the uppermost 30m, although different classes can be accounted for: MS08 individuates only 2 classes, soil or rock, while ITA10 and CF08 define the entire range of the 5 EC8 classification. Only the BEA_M uses a continuous value of Vs30, although the average Vs,30 of the EC8 class has been adopted in this study.

The Global and European GMPEs use the moment-magnitude Mw, ITA10 uses both Ml and Mw, depending on which magnitude estimate is available in the dataset. Finally MS08 is defined for Mw, but it was originally calibrated on local magnitude from [3.5 – 6.3] and then converted in

Mw. The main characteristics of the functional forms of the candidate GMPEs are listed in Table 2.

The style of faulting terms are taken in account by all the GMPEs, except the MS08, as its calibration dataset is almost entirely derived from compressive events of Northern Italy.

Only the CF08, based on hyprocentral distance, does not assume the ground motion saturation at short distances, which is usually accounted by the pseudo depth. Finally CF08 and MS08 do not consider a magnitude dependent attenuation term and the anelastic attenuation coefficient.

Table 1- Characteristic of the 5 selected GMPEs. GM: geometric mean of horizontal components, Max: Maximum of the horizontal components, GMRotI50: median amplitude over all possible orientations of a horizontal component, N: normal fault, R: reverse fault T: thrust fault, S: strike fault, U: Unknown

GMPE	Region	Comp	Magnitude Range	Distance Range (km)	Site classification	Style of faulting	Depth range (km)
ITA10	Italian	GM	4.0– 6.9 Ms and Ml	≤ 200 Rjb or Repi	EC8	N,R/T,S,U	≤ 29
BEA_M	Global	GMRotI50/ GM	3.5 – 8 Mw	≤ 200 Rjb	Vs,30	N,R,S,U	≤ 31
AB10	Eurasia	GM	5 – 7.6 Mw	≤ 100 Rjb	Vs,30 < 360m/s; 360>Vs,30<760m/s Vs,30>760m/s	N,R/T,S	
CF08	Global	GM	5 - 7.2 Mw	≤ 150 Rhypo	EC8	N,R/T,S	≤ 22
MS08	Italian	Max	4.0 - 6.4 Mw	≤ 100 Repi	Vs,30<800m/s Vs,30>800m/s		≤ 60

Table 2- Characteristic of the functional forms of the 5 selected GMPEs.

Model	**Magnitude scaling**	**Saturation at short distances**	**Geometrical spreading M-dependent**	**Anelastic attenuation**	**Site classification**	**Style of Faulting**
ITA10	Quadratic	Through pseudodepth	*x*	*x*	Use categories	*x*
BEA_M	Quadratic	Through pseudodepth	*x*	Constrained using a subset of data	Use Vs,30 values	*x*
AB10	Quadratic	Through pseudodepth	*x*		Use categories	*x*
CF08	Linear	Hypocentral			Use categories or Vs,30 values	*x*
MS08	Linear	Through pseudodepth			Use categories	

3.1. Ranking results

Scherbaum et al. (2009) have recently proposed a data-driven method using an information-theoretic approach for the selection and ranking of GMPEs (Beauval et al., 2012). The starting point of their method is the definition of a meaningful distance measure between probabilistic models. This measure is given by the so called Kullback–Leibler (KL) divergence, which denotes the information loss when a model *g* 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:

$$D(f,g) = E_f[log_2(f)] - E_f[log_2(g)] \tag{1}$$

where *Ef* is the statistical expectation taken with respect to *f*.

In the case of GMPE selection, *f* represents the data-generating process and is only known through observations. Consequently, the term $E_f[log_2(f)]$, cannot be calculated. However, the second term, $-E_f[log_2(g)]$, can still be approximated via the observations (Beauval et al., 2012). This approximation is the negative average sample log-likelihood (LLH) defined by:

$$LLH(g,x) = -\frac{1}{N}\sum_{i=1}^{N} log_2(g(x_i)) \tag{3}$$

where x = {x_i} i = 1,...,N represents the data. Due to its negative sign, the negative average sample LLH is not a measure of closeness but a measure of the distance between a model and the data-generating distribution.

In our application, the GMPEs represent the models and the observations the data generating distribution: a small LLH will indicate that the candidate GMPE is close to the model that has generated the data, while a large LLH corresponds to a model that is less likely of having generated the data.

Following the recent works developed in the SHARE project (http://www.share-eu.org), to evaluate the performance of a set of GMPEs to characterize the ground motion in Europe (Segou and Akkar, 2010; Delavaud et al., 2012) and in the EDF-SIGMA project (2011-2016) to evaluate the GMPEs in France, we adopt the LLH measure defined by equation (3) as ranking criterion for GMPEs.

To rank the most appropriate GMPEs for northern Italy, different datasets are considered:

1. **DBN**: the total data set, where the Ml values are converted into Mw, in order to use the majority of the GMPE (Figure 5). The main characteristics are R < 250km; 4 < Mw < 6.4; depth < 35 km (number of records: 1440). It is worth noting that the minimum magnitude in terms of local magnitude is 3.5;

2. **DBNG** - five subsets of DBN, satisfying the range of validity of each GMPE according to Table 1. Note that by definition ITA10 uses Ml when Mw is not available;
3. **DBN_E** - a subset of DBN composed of the records of the Emilia events; the main characteristics are R < 200km; 4.2 < Mw < 6.1; depth < 15 km (number of records: 762 records),
4. **DBNG_E** - five subset of the DBN_E composed of the Emilia records satisfying the range of validity of each GMPE;
5. **DBN_NE** - a subset of DBN excluding the events of the 2012 Emilia sequence; the main characteristics are R < 250km; 3.9 < Mw < 6.4; depth < 35 km (number of records: 678);
6. **DBNG_NE** - five subsets of DBN, excluding the events of the 2012 Emilia sequence, satisfying the range of validity of each equation.

The LLH parameters have been calculated for 23 periods in the range 0.04 - 4s and then averaged either over the entire period range (hereafter referred to as LLH1) or considering 5 periods (0.1, 0.2, 0.3, 1.0 and 2.0s) selected according to Segou and Akkar et al. (2011) and Delavaud et al. (2012), (hereafter referred to as LLH2).

Figure 6 displays the LLH values in function of period for the 5 GMPES for DBN and DBNG. The corresponding LLH1 and LLH2 are reported in Table 3a and Table 4a. For DBN, we test the LLH values also considering Ml or Mw magnitude scale, as originally reported in the dataset. The results are reported in table 3b

When the DBNG is used, which means the GMPEs are used in their range of validity, the best performing GMPEs are the ITA10 and CF08, with LLH values in the range 2.0 – 3.0, although their different explaining variables and functional forms.

It is worth noting that the degree of extrapolation is different for the different GMPEs: while the minimum magnitude is Mw=3.5 for BEA_M and MS08, and Ml=4 for ITA10, both CF08 and AB10 are defined for Mw greater than or equal to 5. Regarding the distances, both ITA10 and BEA_M are defined up to 200km, CF08 up to 150km while MS08 and AB10 for distances up to 100km. It follows that DBN slightly extends the range of validity for both ITA10 and BEA_M, while a significant extrapolation is performed for AB10 in terms of both magnitude and distance range. For CF08, the extrapolation mainly affects the magnitude range. We also note that using a different Magnitude scale, the LLH values change but the ranking of the GMPEs is unvaried.

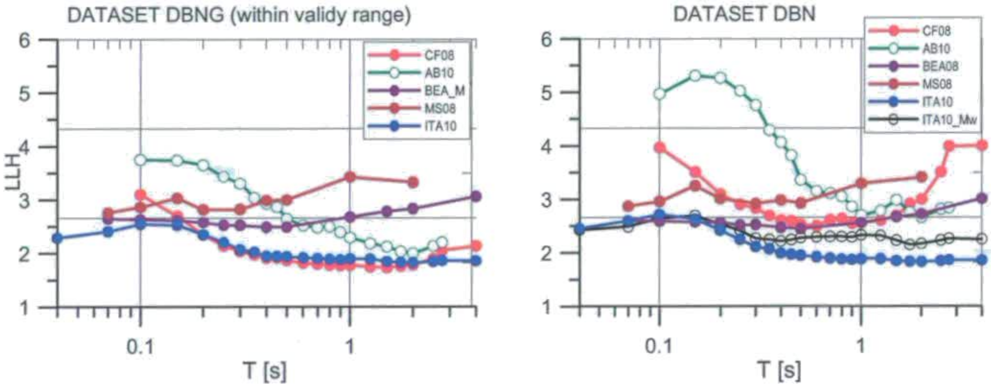


Figure 6 - LLH values plotted as a function of periods for DBN and DBNG. ITA10 was tested considering both Ml/Mw (blues curve) and Mw (black curve).

Table 3. Ranking of the GMPEs for DBN, considering Mw a) and Ml/Mw b)

a)

Rank	LLH1	LLH2	Model
1	2.08/2.34	2.17/2.40	ITA10_Mw
2	2.61	2.58	BEA_M
3	2.97	3.03	CF08
4	3.07	3.13	MS08
5	3.63	3.80	AB10

b)

Rank	LLH1	LLH2	Model
1	2.08	2.17	ITA10_Ml
2	2.25	2.40	BEA_M
3	2.31	2.48	CF08
4	2.73	3.14	AB10
5	4.37	4.55	MS08

If the Emilia data set is used to evaluate the LLH value, we notice, in Figure 7 and Tables 5 and 6, that the main differences with the DBN database are found at short periods and for the global models (AB10 and CF08). This can be attributable to the small amplitudes at short periods for the records of the Emilia events, as observed by Luzi et al. (2012) and explained as the high frequency attenuation of the waves propagating through the Po Plain sedimentary cover.

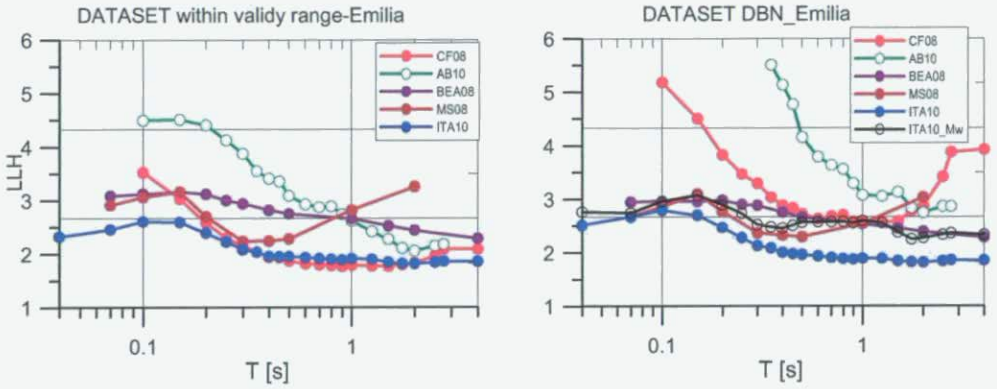


Figure 7 - LLH values plotted as a function of periods for DBN_E and DBNG_E. ITA10 was tested considering both Ml/Mw (blues curve) and Mw (black curve).

Table 4. Ranking of the GMPE for DBN_Emilia

Rank	LLH1	LLH2	Equation
1	2.09/2.57	2.19/2.66	ITA10/ITA_Mw
2	2.67	2.70	MS08
3	2.73	2.71	BEA_M
4	3.48	3.44	CF08
5	4.43	4.81	AB10

Table 5. Ranking of the GMPEs for DBNG_Emilia

Rank	LLH1	LLH2	Equation
1	2.05	2.14	ITA10
2	2.08	2.32	CF08
3	2.74	2.83	MS08
4	2.83	2.82	BEA_M
5	3.11	3.34	AB10

The results for the northern Italy datasets prior to the occurrence of the Emilia sequence are reported in Figure 8 and Table 6 and Table 7. The dataset in the validity range of AB10 and CF08, reduces to 14 events, whose 8 belonging to the Friuli seismic sequence, therefore the LLH values can be considered less reliable for testing the performance of the GMPEs in north Italy. On the other hand ITA10 well perform in describing the ground motion, confirming the results obtained for the other datasets.

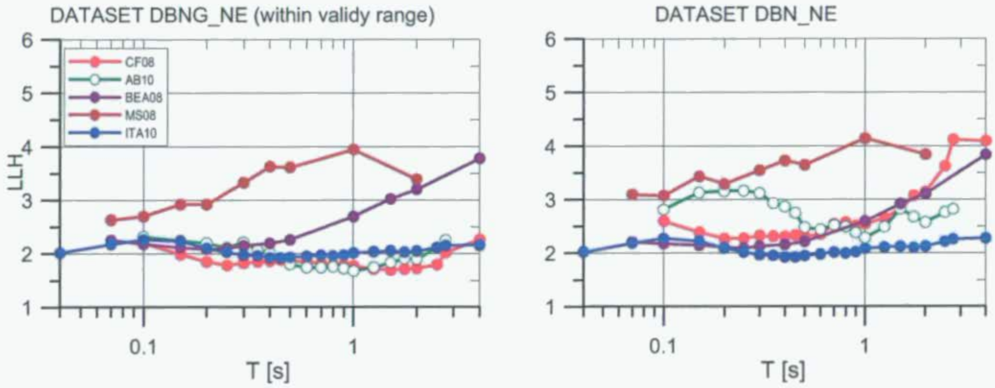


Figure 7 - LLH values plotted as a function of periods for DBN_NE and DBNG_NE.

Table 6. Ranking of the GMPE for DBN_NoEmilia

Rank	LLH1	LLH2	Equation
1	2.08	2.11	ITA10
2	2.48	2.44	BEA_M
3	2.72	2.57	CF08
4	2.74	2.66	AB10
5	3.53	3.60	MS08

Table 6. Ranking of the GMPE for DBNG_NoEmilia

Rank	LLH1	LLH2	Equation
1	1.86	1.88	CF08
2	1.97	1.98	AB10
3	2.05	2.08	ITA10
4	2.51	2.49	BEA_M
5	3.23	3.32	MS08

From the ranking analysis we can conclude that the two GMPEs that better explain the ground motion in Northern Italy in their range of validity are the ITA10 and CF08, where the former can be applied over the magnitude and distance ranges of the strong motion data base for northern Italy, since these ranges are closer to the GMPE range of validity and then a weak extrapolation is required.

4. Residual analysis

The residual analysis has been conducted with the aim of identifying the reason why some of the GMPEs are not appropriate to explain the data (i.e. large values of LLH) and identifying the peculiarity, if any, of the DBN dataset through the examination of the GMPEs which have low LLH values. The residual analysis has been carried out using the DBN data set defined in the validity range of each GMPE to avoid the introduction of errors due to the misuse of the predictive equations. For each GMPE, the residuals are computed as the natural logarithm of the ratio between observations and predictions and the mean residuals over fixed distance and magnitude bin are plotted (Figure 9).

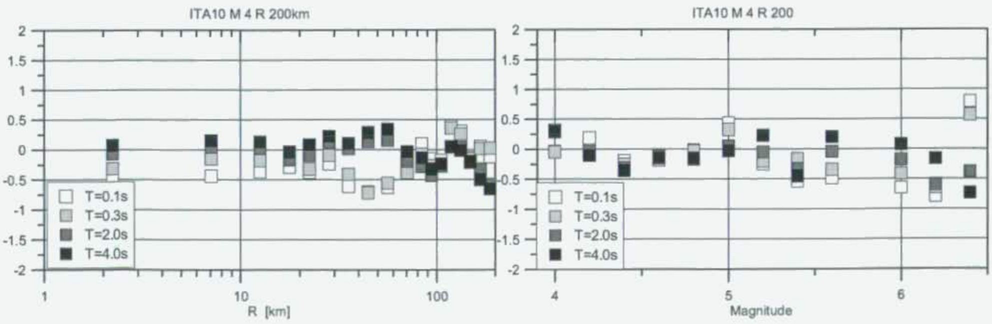
If we examine the ITA10 GMPE, we observe that the mean residuals are in the range -0.5/+0.5 and show a clear frequency dependence. At short periods, the residuals are negative (i.e. observations lower than predictions) in the range 1 – 100 km, while, at long periods the residuals are close to zero or positive, especially in the distance range 30 – 70 km. This could indicate a low frequency amplification and a high frequency deamplification of the ground motion, with respect to the mean predictions, due to the wave propagation through the Po plain sediments. At distances in the range 100 – 150km a bump in the residuals is observed, in agreement with the finding of Bragato et al (2011), interpreted as the reflection of S waves at the Moho (SmS phase). The residuals do not show a clear dependence on magnitude and the negative residuals at magnitude 6.4 at 4s might be attributed to the low quality of the analog records of the 1976 Friuli event.

The CF08 is not efficient in explaining the high frequency motion (mean residuals in the range -0.5/-1.5), also confirmed by the relatively high LLH at periods lower than 0.3 s, as the residuals are negative in the entire distance range. At period higher than or equal to 0.3s, the trend of the residuals with distance is similar to the one of ITA10 (with values in the range -0.5/+0.5).

BEA_M has, in general, negative residuals, in the range (0/-1.8) at all distance ranges for all periods and does not allow to detect the frequency dependence on distance in the range 30 – 100 km.

AB10 evidences the same relative attenuation trend with frequency observed by ITA10 and CF08, although the residuals are negative (range between 0 / -1.5) at distances larger than 10km for all periods and this explains the high values of the LLH parameters, especially at low periods.

The MS08 indicates a general under-prediction of the observations at periods larger than 0.1s (mean residuals in the range 0 / 1), also indicated by the high values of the LLH. In particular, the residuals at low-frequency (T = 2s), are larger than 0.5, probably due to the scarce sampling of moderate to high magnitude events of the data set with which has been calibrated. At 0.1s an underestimation is observed for distances lower than 30 km, while, in the range 30 – 80 km, the high-frequency motion is overestimated. Finally, we observe that the mean residuals decrease with increasing magnitude, suggesting that the linear magnitude scaling of the MS08 model is not appropriate to describe the DBN dataset.



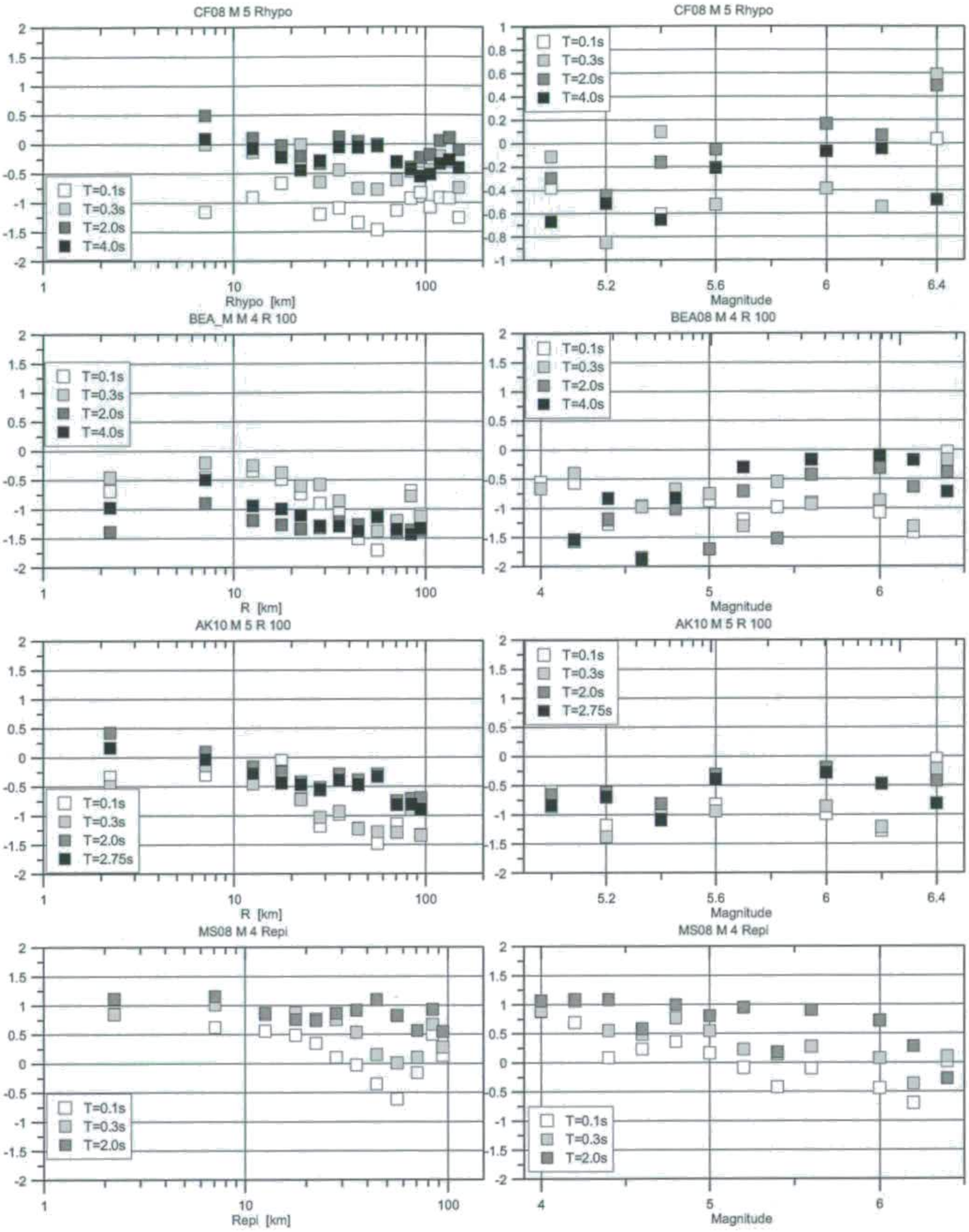


Figure 9 - Mean residuals plotted as a functions of distance (right) and magnitude (left) for DBNG, considering the 5 GMPEs. From the top to the bottom: ITA10; CF08; BEA_M; AB10 and MS08. Residuals equal 2 corresponds to a ratio between observations and predictions of about 7.5

5. Conclusions

In this work, five GMPEs, based on global (CF08 and BEA_M), European (AB10) and Italian (ITA10 and MS08) datasets are selected to evaluate their performance in describing the ground motion in northern Italy and especially in the Po plain region.

The northern Italy dataset, DBN, is composed by 1440 records from 224 stations, relative to seismic event from 1976 (Friuli sequence) to 2012 (Emilia sequence) in the magnitude range 3.5 – 6.4. Based on LLH and residuals analysis, we found that:

- ITA10 and CF08, are both applicable to predict the ground motion of the Northern Italy, providing the lower LLH values (Table 3 and 4) for a wide period range, in their range of applicability. It is worth to note that these GMPEs have been calibrated on different datasets (an Italian dataset with M from 4.0 – 6.9 and global dataset, with M larger than 5, respectively) and have different functional forms for the source and distance scaling.
- CF08 and global models, in general, when extrapolated outside their range of validity, cause an increase of LLH values, from 2.01 to 2.97 (CF08), and from 2.72 to 3.63 (AB10), therefore the extrapolation should be avoided unless corrective factors are introduced (see for example the results for BEA_M).
- In general we found that all the considered models are not able to represent high-frequency motion, predicting values which are larger than the observations, as shown from the LLH values and residual analysis.
- The DBN presents some peculiar features, that can be highlighted independently from the GMPES adopted: i) low amplitudes at short periods, ii) attenuation with distance strongly dependent on frequency; iii) amplification of spectral ordinates in the distance range from 80 to 100km, particularly evident at short periods (0.1 s).

On the base of the selection of the GMPEs that will be used for the hazard assessment of north Italy, the relative weights to be included into the GMPE logic tree branch could be evaluated.

Future developments of this work could concern i) the estimations of LLH using the different terms of the standard deviations (within-event and between-event varaibily)), to better understand which components of the ground motion (site, propagation, source) are scarcely captured by the selected GMPEs and ii) calibration of new attenuation relationships using an enlarged strong motion dataset, including records from temporary networks and exploiting non-parametric approaches.

Furthermore, specific studies could be devoted to investigate the variability of the ground motion at single site, which is extremely useful for hazard studies.

6. References

Akkar S. and J. J. Bommer (2010). Empirical Equations for the Prediction of PGA ,PG V, and Spectral Accelerations in Europe, the Mediterranean Region, and the Middle East, Seismological Research Letters Vol. 81, 195 – 206.

Atkinson G. and D Boore (2011). Modifications to Existing Ground-Motion Prediction Equations in Light of New Data, Bulletin of the Seismological Society of America, Vol. 101, 1121–1135

Beauval C., Tasan H., Laurendeau A., Delavaud E., Cotton F, Guéguen P, and N. Kuehn (2012). On the Testing of Ground-Motion Prediction Equations against Small-Magnitude Data. Bulletin of the Seismological Society of America, Vol. 102, No. 5, pp. 1994–2007

Bethmann F. Deichmann N and M. Mai (2011) Scaling relations of local magnitude versus moment magnitude of similar earthquakes in Switzerland, Bulletin of the Seismological Society of America vol. 101 515-534.

Bindi D., F. Pacor, L. Luzi, R. Puglia, M. Massa, G. ameri and R. Paolucci(2011). Ground motion prediction equations derived from the Italian strong motion database, Bulletin of Earthquake Engineering, Vol. 9, 1899-1920.

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, Vol. 24, 99–138.

Boore, D. M., and J. J. Bommer (2005). Processing of strong-motion accelerograms: Needs, options and consequences, Soil Dynamics and Earthquake Engineering, 25, 93-115.

Bragato, P. L., Sugan, M. Augliera, P. Massa, M., Vuan, A. and A. Sarao (2011). Moho Reflection Effects in the Po Plain (Northern Italy) Observed from Instrumental and Intensity Data. Bulletin of the Seismological Society of America, Vol. 101, 2142–2152.

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

Castello B., Olivieri M., Selvaggi G- (2007). Local and Duration Magnitude Determination for the Italian Earthquake Catalog, 1981–2002, Bulletin of the Seismological Society of America, Vol. 97, 128–139.

Castro R., Massa M., Augliera P. and F. Pacor (2008). Body-wave attenuation in the region of Garda, Italy. Pure and Applied Geophysics, Vol. 165, 1351-1366

Cauzzi C. and E. Faccioli (2008). Broadband (0.05 to 20 s) prediction of displacement response spectra based on worldwide digital records, Journal of Seismology, Vol. 12, 453–475.

Delavaud E., Cotton F., Akkar S., Scherbaum F., Danciu L., Beauval C., Drouet S. Douglas J., R.

Basili, Sandikkaya M.A., Segou M., E. Faccioli and N. Theodoulidis (2012). Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe, Journal of seismology, 16, 451–473

De Natale G. D., Madariaga R., Scarpa R. and A. Zollo (1987). Source Parameter Analysis from Strong Motion Records of the Friuli, Italy, Earthquake Sequence (1976-1977). Bulletin of the Seismological Society of America, Vol. 77, 1127–1987.

Hisada Y., Aki K., Teng T. (1993). Application of Nonstationary Ray Decomposition to Identifying Deep Seismic Bedrock of the Kanto Sedimentary Basin, Japan. Bulletin of the Seismological Society of America, Vol. 83, 1700-1720

Joyner W. B. (2000). Strong Motion from Surface Waves in Deep Sedimentary Basin. Bulletin of the Seismological Society of America, vol. 90, S95-S112.

Kagawa, T., B. Zhao, K. Miyakoshi, and K. Irikura (2004). Modeling of 3D Basin Structures for Seismic Wave Simulations Based on Available Information on the Target Area: Case Study of the Osaka Basin, Japan. Bulletin of the Seismological Society of America, Vol. 94, 1353-1368.

Luzi L., Pacor F., Ameri G., Puglia R., Burrato P., Massa M., Augliera P., Franceschina G., Lovati S. and R. Castro (2012). Overview on the strong motion data recorded during the May-June 2012 Emilia seismic sequence, submitted to Seismological Research Letter

Massa M., Morasca P., Moratto L., Marzorati S., Costa G. and D. Spallarossa (2008). Empirical Ground-Motion Prediction Equations for Northern Italy Using Weak- and Strong-Motion Amplitudes, Frequency Content, and Duration Parameters, Bulletin of the Seismological Society of America, Vol. 98, 1319–1342

Pacor, F., R. Paolucci, G. Ameri, M. Massa, and R. Puglia (2011). Italian strong motion records in ITACA: overview and record processing, Bulletin of Earthquake Engineering, Vol. 9, 1741-1759.

Sato, T., R. W. Graves, and P. G. Somerville (1999). Three-dimensional finite-difference simulations of long-period strong motions in the Tokyo metropolitan area during the 1990 Odawara earthquake (MJ 5.1) and the Great 1923 Kanto earthquake (MS 8.2) in Japan, Bulletin of the Seismological Society of America, Vol. 89, 579–607.

Scherbaum F., Delavaud E. and C. Riggelsen (2009). Model selection in seismic hazard analysis: An information-theoretic perspective, Bulletin of the Seismological Society of America, Vol. 99, 3234–3247.

Segou M. and S. Akkar (2010). A report on the evaluation of ground motion prediction equations within the context of share WP4 task, Deliverable Task WP4, August 2010, 40 pp.

S R C
Savy Risk Consulting

Project SIGMA

Review of:

RANKING OF AVAILABLE GMPEs FROM RESIDUAL ANALYSIS FOR NORTHERN ITALY AND DEFINITION OF REFERENCE GMPEs.

(Ref : SIGMA-2012-D2-53, version 1)

Jean B. Savy
November 8, 2012

1. Scope of the work reviewed

This is a review of the research work documented in EDF Ref: SIGMA-2012- D2-53, version 1. F. Pacor, L. Luzi, M Massa and D. Bindi are the joint authors of the document, and the work is to be presented at the CS4 of November 16th, 2012, under the title: "Selection of Italian and Worldwide records: Ranking of available GMPEs from residual analysis tests on Pô plain data and definition of reference".

The objective of this research was to identify and rank ground motion prediction models for use in PSHA analyses in the Pô plain.
The areas of investigations were split into three groups:

1. Select a data base
2. Select and describe a set of candidate ground-motion attenuation models
3. Rank and evaluate the candidate models
    - Rank models with F. Scherbaum's method
    - Perform a residual analysis to evaluate models

Formulate general conclusions and identify needed additional research to fulfill the goals of the study.

2. General review conclusions

This study is well constructed and achieves the goal of identifying the appropriate GMPEs for use in a PSHA for the Pô plain. It is well documented and in general the report has the right level of details and references. In all, I have only minor comments but for the exception of the issue of uncertainty in the

conversion of M_l to M_w, and for the missed opportunity to use the excellent DBN data base to get insights in the singe-event, and between-events uncertainties.
A brief residual analysis and discussion gives interesting and useful insights on the applicability ranges of the models.
The report stops short of attributing weights to the logic tree of GMPEs, but the authors state that this will be done in a continuing study. At this point, it would be useful to suggest methods that will be used to determine those weights, prior to actually start the study.

- *The database:*
  The starting point of such a study begins with the selection of a data base. It is comforting to see that the data for this region is of high quality, and sufficient in quantity, probably one of the best data base around. One interesting aspect of this data base, and an important one, is that it contains records for several well documented earthquakes, such as the Emilia and Friuli (sequences) earthquakes. I was expecting a discussion on this, leading perhaps to an analysis of single and inter-events properties. There was a segregation of the data for the Emilia sequence, but it was not pursued to any interesting conclusion, and seemed to be irrelevant since in the end, the models were ranked with the full database. Nevertheless, in some sense this allowed to show that with or without Emilia, the LLH did not change substantially (Tables 5 and 6).

- *The selection of models:*
  My biggest disappointment was in the pre-selection of candidate models, which is presented to us as a fait accompli. There is no description of the process (if any) that lead to the pre-selection of the 5 models analyzed. As part of a good approach, and to be consistent with the ongoing work in SIGMA, there should have been more of a rational and structured method, including identification of an exhaustive set of valid models, then application of pre-selection criteria (Cotton et al.,), and an attempt at covering the space of all possible models, leading to the selection of 4 or 5 models.

- *The conversion of M_l to M_w:*
  Use of the Castello et al. (2007) relationship between Ml and Mw is appropriate since it is the most recent one and is using only Italian data. However the fit (least squares) obtained by Castello et al. (2007) is less than perfect, with a R^2 of .76, and an uncertainty in the coefficients of +/- 0.4 for the coefficient of Ml, and +/-0.19 for the constant. This means that there is a fairly large uncertainty, and this uncertainty could affect the results of this study. By chance, ITA10 with and without the conversion ranks the same. But the LLH is larger by at least 25% for the conversion case. The other models analyzed, all use the conversion but they each could be affected differently. I suggest that an investigation should have been done to resolve this issue.

3. Detailed comments

- I do not identify the typos and other editorial and formatting details here, but I will transmit a copy of my hand-written notes to the authors in the hope that it will help.
- Page 8:
  Pacor et al. 2011, missing from bibliography.
- Page 8:
  "Zero pads are then removed….procedure is applied…"
  What procedure was applied? Can we give some detail?

- Page 10: Equation (1)
  There is no mention of the uncertainty in the estimation of M_w for a given M_l. This uncertainty is large, according to Catello et al. (2007), and should be taken in consideration, or at the least a sensitivity analysis should be performed.
- Page 8: "…almost unvaried."
  This is not a clear statement.
  First, I do not see, from the figure 3, that the distribution of magnitude is little affected by the conversion. Second, I believe that a large proportion of the data is within the 4.5 to 5.5 range, thus making any, even small, variation in the distribution of magnitude potentially important.
- Page 12:
  3. GMPE ranking and selection
  We are suddenly presented with a set of 5 candidate models without any explanation of how they came here. Elsewhere in SIGMA we have insisted on setting rational processes for this kind of selection, and I believe it should be similar here. There has to be some discussion, pre-selection process, etc. that support the choice of the 5 models.
- Page 15: middle of page
  It should clarified that it is the relative value of the LLH measure between models that is of interest, and not the absolute value.
- Page 15:
  Assuming that the uncertainty in the conversion $M_l>M_w$ is important for the ranking, can a modification of the LLH calculation be developed to account for these uncertainties?
- Page 20:
  "…to explain the data (i.e. large values of LLH)…"
  Values of LLH in the range of 3.0 are hardly large values. What we are seeing here is that, (without considerations of the results of the residual; analysis) all the models selected do an acceptable to good job at explaining the data. But as we see with the residual analysis they behave very differently in different ranges of magnitude and distance.
- Page 22:
  As a piece of information to help the reader understand the meaning of the amplitude of the residuals it would be useful to give an idea of what a 0.5 residual means in terms of variation in the ground motion value?

Respectfully submitted, November 9, 2012.

Jean Savy

Comments on deliverable D2-53

Ranking of available GMPEs from residual analysis for Northern Italy and definition of reference GMPEs

Ref SIGMA-2012-D2-53 Version 01

Frank Scherbaum
University of Potsdam
Potsdam, Nov 8, 2012

General remarks

The report SIGMA-2012-D2-53, Version 01 describes the work performed to quantitatively evaluate the applicability of five selected ground motion prediction equations (GMPEs) to predict ground motion in Northern Italy, in particular in the Po plain. Two methods have been chosen for this analysis
a) the information theoretical approach of Scherbaum et al. (2009) and
b) traditional residual analysis.

The data set used (DBN) consists of 1440 records from 137 events, recorded at 225 stations which to my knowledge is one of the largest datasets which so far have been used to perform such an analysis.

Overall I found this a very good piece of work which is well documented and clearly presented. My comments below address different issues of the report. Some of my comments address aspects where I believe further clarification is needed while others are merely meant as suggestions.

General comments

1) The pre-selection of the five GMPEs does not become clear and appears as rather ad-hoc. Since I assume that these GMPEs are going to be used within a logic tree framework, it should be traceable why the authors believe that these 5 GMPEs are able to cover the ground motion range relevant for hazard in the Po plain in other words why they believe these models to be collectively exhaustive.

2) A personal note and suggestion: Due to the excellent quality of the data set, the authors could consider to perform the information theoretic selection and ranking not only on the total sigma models but separately for within-event and between-event event variability. This would allow to test models much more specifically than just with total sigma. If data sets like the one in the present study would have been available at the time when we did the Scherbaum et al. (2009) study, we would have used within-event and between-event variability separately.

Specific comments

3) In order to convert all magnitudes to Mw, the authors use a correlation equation between Ml and Mw proposed by Castello (2007):

$$\mathrm{Mw} = 0.79\ \mathrm{Ml} + 1.20$$

and assume it to be valid in the local magnitude range from 3.5 - 5.8. I find this puzzling since (e.g. based on the work of Deichmann (2006)) I would assume that at least for the larger magnitudes there would be a 1:1 relation between Ml and Mw. The deviation from the 1:1 relationship, in the extreme Mw can become proportional to 2/3 Ml, is sometimes interpreted as being caused by low path filtering effects. A recent PhD at the ETh Zurich by Falko Bethmann (2011) entitled „Magnitude scaling relations and attenuation in thick sediments“, which is now partially published in BSSA (Bethmann et al., 2011), illustrates nicely how attenuation can affect the relationship between local magnitude and moment magnitude so that a linear relationship between Mw and Ml over a large magnitude range would no longer be justified.

2) The statement on page 10 that „ the records with local magnitude in the range 3.5 -4.0 are concentrated in the range 4.0-4.5“ but „ the distribution of the records characterized by larger magnitudes (>4.5) is almost unvaried“ is confusing because all magnitude tags should be affected in the same way. Once the magnitude conversion is applied the magnitude tag for a particular record becomes different.

3) On page 14, the authors refer to the work by Beauval et al. (2012) and their observation of LLH values in the range of 1.4-1.5 for cases where synthetic data are tested against the original distribution. From the report it did not become clear to me if the authors interpret absolute LLH values of individual GMPEs. I believe that one should only interpret **relative** information loss between GMPEs because the self-information of the process that produces the real observations (nature) and which cancels out in relative information-loss is only known in synthetic cases.

4) On page 23, the authors conclude that none of the chosen GMPEs are able to represent the high-frequency motions well. My interpretation here is that this might be caused by differences in the attenuation properties of the Po plain and the host regions of the model-generating data sets and could be addressed by some sort of Hybrid Empirical approach (Campbell, 2003).

References

Bethmann, Falko (2011). *Magnitude scaling relations and attenuation in thick sediments: application to the induced seismicity beneath the city of Basel, Switzerland.* PhD thesis, ETH Zürich.

Bethmann, Falko, Nicholas Deichmann, and P. Martin Mai (2011). Scaling Relations of Local Magnitude versus Moment Magnitude for Sequences of Similar Earthquakes in Switzerland, *Bulletin of the Seismological Society of America,* ***101***(2), 515-534.

Campbell, K. W. (2003). Prediction of strong ground motion using the hybrid empirical method and its use in the development of ground-motion (attentuation) relations in Eastern North America, *Bulletin of the Seismological Society of America,* **93**(3), 1012-1033.

Castello B., Olivieri M., Selvaggi G- (2007). Local and Duration Magnitude Determination for the Italian Earthquake Catalog, 1981–2002, Bulletin of the Seismological Society of America, Vol. 97, 128–139.

Deichmann, N. (2006). Local Magnitude, a Moment Revisited,*Bulletin of the Seismological Society of America,* **96**(4A): 1267-1277.