Bulletin of Earthquake Engineering

# The older the better? The strange case of empirical ground motion models in the near-source of moderate-to-large magnitude earthquakes

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | BEEE-D-21-00448R1 |
| **Full Title:** | The older the better? The strange case of empirical ground motion models in the near-source of moderate-to-large magnitude earthquakes |
| **Article Type:** | Original Article |
| **Keywords:** | empirical ground motion models, near-source conditions, analyses of residuals, strong motion dataset |
| **Corresponding Author:** | Angela Chiecchio, M.Sc.<br>Politecnico di Milano<br>Milan, ITALY |
| **Corresponding Author Secondary Information:** | |
| **Corresponding Author's Institution:** | Politecnico di Milano |
| **Corresponding Author's Secondary Institution:** | |
| **First Author:** | Roberto Paolucci |
| **First Author Secondary Information:** | |
| **Order of Authors:** | Roberto Paolucci |
| | Angela Chiecchio |
| | Manuela Vanini |
| **Order of Authors Secondary Information:** | |
| **Funding Information:** | |
| **Abstract:** | This paper aims at providing a quantitative evaluation of the performance of a set of empirical ground motion models (GMMs), by testing them in a magnitude and distance range ( $M_w = 5.5 \div 7.0$ and Joyner-Boore source-to-site distance $R_{jb} \leq 20$ km) which dominates hazard in the highest seismicity areas of Italy for the return periods of upmost interest for seismic design. To this end, we made use of the very recent release of the NESS2.0 dataset (Sgobba et al., 2021), that collects worldwide near-source strong motion records with detailed metadata. After selection of an ample set of GMMs, based on either their application in past seismic hazard assessment (SHA) studies or for their recent introduction, a quantification of between- and within-event residuals of predictions with respect to records was performed, with the final aim of shedding light on the performance of existing GMMs in the near-source of moderate-to-large earthquakes, also in view of their potential improvement by taking advantage of results from 3D physics-based numerical simulations. |
| **Suggested Reviewers:** | |
| | |
| | |
| | |
| | |

**Abstract**

This paper aims at providing a quantitative evaluation of the performance of a set of empirical ground motion models (GMMs), by testing them in a magnitude and distance range ($M_w$ = 5.5 ÷ 7.0 and Joyner-Boore source-to-site distance $R_{jb} \leq 20$ km) which dominates hazard in the highest seismicity areas of Italy for the return periods of upmost interest for seismic design. To this end, we made use of the very recent release of the NESS2.0 dataset (Sgobba et al., 2021), that collects worldwide near-source strong motion records with detailed metadata. After selection of an ample set of GMMs, based on either their application in past seismic hazard assessment (SHA) studies or for their recent introduction, a quantification of between- and within-event residuals of predictions with respect to records was performed, with the final aim of shedding light on the performance of existing GMMs in the near-source of moderate-to-large earthquakes, also in view of their potential improvement by taking advantage of results from 3D physics-based numerical simulations.

**Keywords**

empirical ground motion models, near-source conditions, analyses of residuals, strong motion dataset

**Acknowledgments**

This work has been partially supported by swissnuclear within the research activity "Development of advanced numerical approaches for earthquake ground motion prediction", in the framework of the Sigma2 project, and by the Department of Civil Protection within the ReLUIS project WP18 "Normative contributions related to seismic action". The authors are sincerely grateful to Sreeram Reddy Kotha for providing fruitful suggestions about the interpretations of results, and to Ezio Faccioli and Chiara Smerzini for their useful comments. Researchers of INGV are gratefully acknowledged for making the NESS2.0 dataset available. The authors also wish to thank two anonymous reviewers for their valuable and detailed remarks which contributed significantly to improve our manuscript.

**Declarations**

Conflict of interest

The authors acknowledge that there are no conflicts of interest.

Availability of data and material

The NESS2.0 dataset is freely available at the website http://ness.mi.ingv.it/ (Last access: June 6th 2021). The records from ITACA were downloaded from the ITalian ACcelerometric Archive, http://itaca.mi.ingv.it (Last access: March 12th 2021). The records from SIMBAD were provided by Chiara Smerzini (Politecnico di Milano).

Code availability

All the computations have been done using MATLAB (http://www.mathworks.com/products/matlab/). The authors don't give their consent to share their routines.

Consent for publication

The authors express their consent for publication of the manuscript in Bulletin of Earthquake Engineering, Springer

## 1. Introduction

There is a general consensus that empirical ground motion models (GMMs) are bound to improve their effectiveness and predictive performance as new records become available, although it is also recognized that, in spite of increasing data and more complex parametrizations, a significant reduction in the aleatory variability cannot be achieved (Douglas and Edwards, 2016).

However, the exponential increase of records has not populated uniformly the different ranges of magnitude and distances. As illustrated in Figure 1, with reference to the availability of strong-motion records in Italy, while, up to the 90s, records in the range $5.5 \leq M_w \leq 7$ and Joyner-Boore distance $R_{jb} \leq 20$ km (simply denoted in the following by *M-R* range) weighted about 10% of the whole dataset, now they constitute only about 1%. Although this is an obvious consequence of the seismicity recurrence laws, it clearly reminds that it takes a long time for the calibration dataset of the empirical GMMs to be significantly enriched by new records in the near-source region of moderate-to-strong earthquakes. Besides the Italian case, as shown in Figure 2, similar considerations apply to the calibration datasets of European and global GMMs.

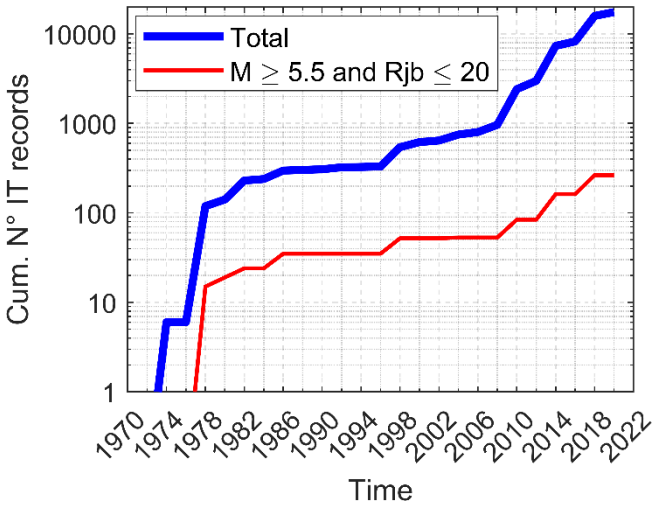


Figure 1 – Cumulative number of Italian records in the time period ranging from 1970 to 2020. (Thick line): total number of available Italian records; (thin line): number of Italian records with $M_W \geq 5.5$ and $R_{JB} \leq 20$ km. Records are taken from the Engineering Strong Motion database (ESM: https://esm-db.eu/#/home, Last access: 12/10/2021).

Our interest towards the selected *M-R* range of moderate to large magnitude earthquakes in near-source conditions stems from its dominance in the seismic hazard assessment (SHA) of the highest seismicity regions of Italy (e.g., Barani et al., 2009), in view of more accurate determinations of elastic spectra for design, of reliable scenarios for seismic risk evaluations in large urban areas, as well as for critical infrastructures, and, more generally, of a proper planning of risk mitigation strategies.

Although the selected *M-R* range may not be strictly considered as near-source, we will check the predictions of different GMMs with records included in NESS 2.0 (Sgobba et al., 2021, referred to in the following as NESS), consisting of a near-source dataset including worldwide records from active shallow crustal regions, with a careful documentation in terms of seismic source and site conditions metadata. Note that, in NESS, records are selected only based on *M-R* considerations and not on their impulsive/non-impulsive features.

For the purpose of this paper, a set of representative GMMs, listed in Table 1, was selected either because of their past use in the framework of SHA studies in Italy, or of their potential use in future studies. Therefore, this selection is neither intended to exhaustively cover the whole range of applicable GMMs nor to provide an absolute ranking for use for future SHA studies in Italy.

This paper is organized as follows. After introducing the selection of GMMs and of the records of NESS considered in this work, the performance of selected GMMs against NESS records is illustrated. Residuals, separated into their between- and within- event components, are computed with respect to each model, together with their standard deviation. Finally, general remarks on the adequacy of older and newer empirical GMMs to predict near-source ground motions of moderate-to-large earthquakes are provided, together with considerations on the possible support from 3D physics-based numerical simulations.

## 2. Selection of empirical earthquake ground motion models

Generation of GMMs, or recalibration of old ones, has nowadays become a continuous process, as new data or new regression procedures become available. In the framework of SHA studies in Italy, we selected several GMMs of interest either for their role in the derivation of the official SH maps of Italy (Ambraseys et al., 1996; Sabetta and Pugliese, 1996), or for their specific calibration with near-source records (Ambraseys and Douglas, 2003), or for their calibration in the Italian and European framework (Akkar and Bommer, 2010; Bindi et al., 2011; Akkar et al., 2014a; Bindi et al., 2014; Lanzano et al., 2019; Kotha et al., 2020), or for their worldwide calibration dataset and use (Boore et al., 2014; Campbell and Bozorgnia, 2014; Cauzzi et al., 2015), with, in the case of Chiou and Youngs (2014), the inclusion of a near-source term in the model formulation.

Table 1 gives an overview of the 13 selected models, in terms of area of calibration of its dataset, magnitude and distance ranges of applicability, distance metrics and intensity measure type. Note that the most recent models (ITA18 and K20) are based on the European Strong Motion dataset (ESM, Luzi et al. 2020), while two of them, ITA14 and ASB14, are derived from the European RESORCE dataset (Akkar et al. 2014b). Two of the selected models are derived from only Italian data (SP96 and ITA10). Anyway, all selected models are suitable to be used in the shallow active crustal regions that dominate seismicity of Italy, with few exceptions of volcanic areas and subduction zones offshore Calabria.

Table 1 GMMs selected for the analyses. For each model, reference details, area of calibration of the dataset, distance metric, distance ranges of applicability, earthquake magnitude and intensity measure type are given.

| GMM | acronym | Area of calibration | R metric | R [km] | M | IM type (‡) |
|---|---|---|---|---|---|---|
| Ambraseys et al., 1996 | AMB96 | Europe | $R_{jb}$ | 0 - 200 | $M_s$ = 4 - 7.5 | $H_{max}$ |
| Sabetta and Pugliese, 1996 | SP96 | Italy | $R_{jb}$, $R_{epi}$ | 0 - 100 | 4.6 - 6.8 | $H_{max}$ |
| Ambraseys and Douglas, 2003 | AD03 | Global | $R_{jb}$ | 0 - 15 | $M_s$ =5.8 - 7.8 | $H_{max}$ |
| Akkar and Bommer, 2010 | AB10 | Europe - Middle East | $R_{jb}$ | 0 - 99 | 5 - 7.6 | HGM |
| Bindi et al., 2011 | ITA10 | Italy | $R_{jb}$ | 0 - 200 | 4.1 – 6.9 | HGM |
| Akkar et al., 2014° | ASB14 | Europe - Middle East | $R_{jb}$, $R_{hypo}$, $R_{epi}$ | 0 - 200 | 4.7 – 7.6 | HGM |
| Campbell and Bozorgnia, 2014 | CB14 | Global | $R_{rup}$ | 0 - 300 | 3 – 7.9 | RotD50 |
| Boore et al., 2014 | BSSA14 | Global | $R_{jb}$ | 0 - 400 | 3 – 7.9 | RotD50 |
| Chiou and Youngs, 2014 | CY14 | Global | $R_{rup}$ | 0 - 300 | 3.5 – 8.0 | RotD50 |
| Cauzzi et al., 2015 | CEA15 | Global | $R_{rup}$ | 0 - 150 | 4.5 – 7.9 | HGM |
| Bindi et al., 2014 | ITA14 | Europe - Middle East | $R_{jb}$, $R_{hypo}$ | 0 - 300 | 4 – 7.6 | HGM |
| Lanzano et al., 2019 | ITA18 | Europe | $R_{jb}$, $R_{rup}$ | 0 - 200 | 3.5 - 8 | RotD50 |
| Kotha et al., 2020 | K20 | Europe | $R_{jb}$ | 0 - 545 | 3 – 7.4 | RotD50 |

(‡) Intensity Measures are defined as follows: HGM = horizontal geometric mean; RotD50 = median value of the IM distribution obtained from rotated waveforms (Boore, 2010); $H_{max}$ = max horizontal component.

Given that the aim of this work is the evaluation of performance of GMMs in the magnitude and distance ranges of engineering relevance in Italy (i.e., $M_w$ from about 5.5 to 7 and distances up to 20 km), Figure 2 highlights, for some of the selected models, the portion of records falling in that range (shown by the rectangle), with respect to the total.

It is clear that the records included in the selected *M*-*R* range are by far a minority of the total, especially for the most recent GMMs, despite the richness of their calibration dataset. Moreover, the fraction of the selected records with respect to the total decreases sharply moving towards recent years, because the number of small magnitude - large distance records increase with time much faster than the number of near-source records. Figure 4 shows the observations used in this analysis, within the same ranges of interest.

As a last remark, we point out that, despite the continuous improvement in regression models and metadata, the aleatory variability of the predictions of the GMMs tends to increase, most probably due to the large variability of the available calibration dataset. To support this sentence, Figure 3 illustrates the total standard deviation ($\sigma_{ln}$) of the selected models as a function of period, showing an increasing trend with time of $\sigma_{ln}$.

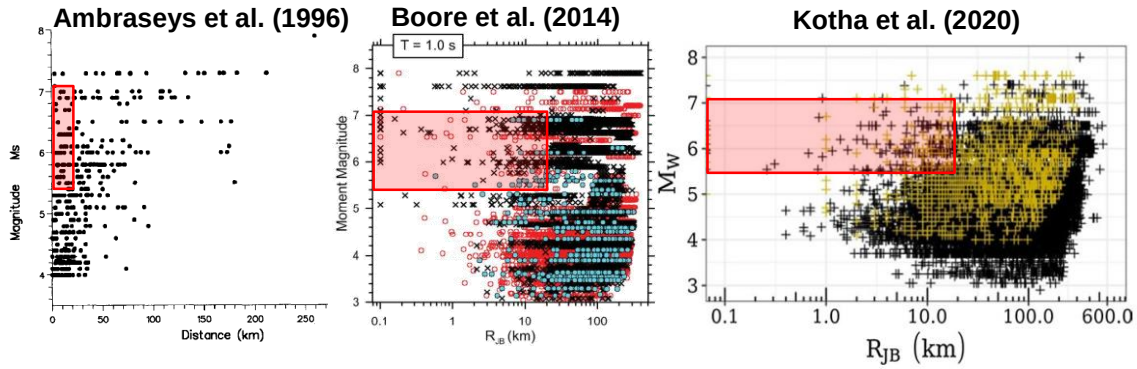


Figure 2 – Magnitude-distance distribution of calibration datasets for some of the selected GMMs, namely: AMB96, BSSA14 and K20 (see Table 1 for references). Distance metric is $R_{JB}$ for all plots. The rectangle highlights magnitude and distance ranges used for the selection of the data in this study.

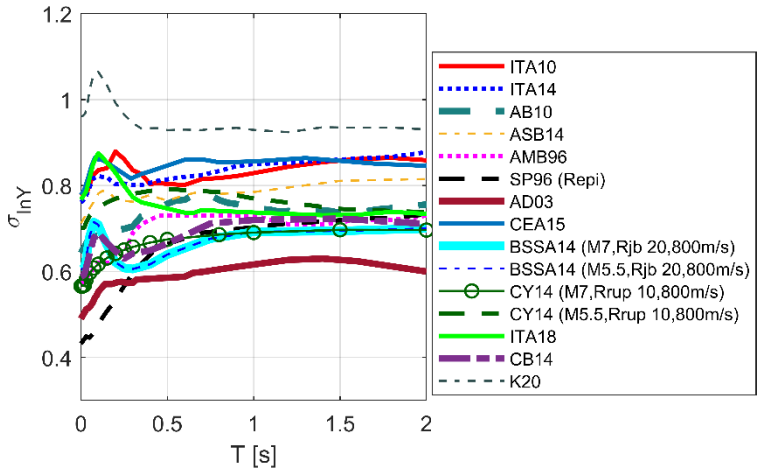


Figure 3 – Total standard deviation $\sigma_{ln}$ of selected GMM (used as a function of $R_{jb}$, when not specified). Note that for BSSA14 and CY14, $\sigma_{ln}$ depends on $M_w$, distance, and site conditions as well; adopted values are shown.

### 3. Overview of NESS near-source dataset

NESS is a dataset of high-quality metadata and intensity measures of near-source strong-motion records from active shallow crustal regions, with a careful documentation in terms of seismic source (e.g. geometries and rupture mechanisms) and site conditions. The last version of the dataset was released in January 2021 and it is available at the

website http://ness.mi.ingv.it/ (Sgobba et al., 2021). As explained by Pacor et al. (2018), records were selected according to the following criteria: $M_w \geq 5.5$, hypocentral depth $\leq$ 40km, distance from the source obtained from scaling relationships depending on magnitude, stress drop and fault dimensions. Figure 4 shows the distribution of data with respect to $M_w$, $R_{jb}$, site conditions and fault mechanism.

In the analyses of this study, only a selection of data from NESS was considered (Figure 4, rectangle). The selection has been done according to the following criteria:

- records on ground type B according to EC8 (stiff soil conditions, CEN 2004);
- records with $R_{jb} \leq 20$ km.

Selected data were classified in four $M_w$ classes, centered at 5.5, 6.0, 6.5 and 7.0 (±0.1), respectively, as shown in Figure 5. The class with the largest number of records is $M_w$ 6.0, with about 70 records. The same figure shows the distribution of the selected data with respect to focal depth as well. Note that most of the records are characterized by a focal depth of about 10 km.

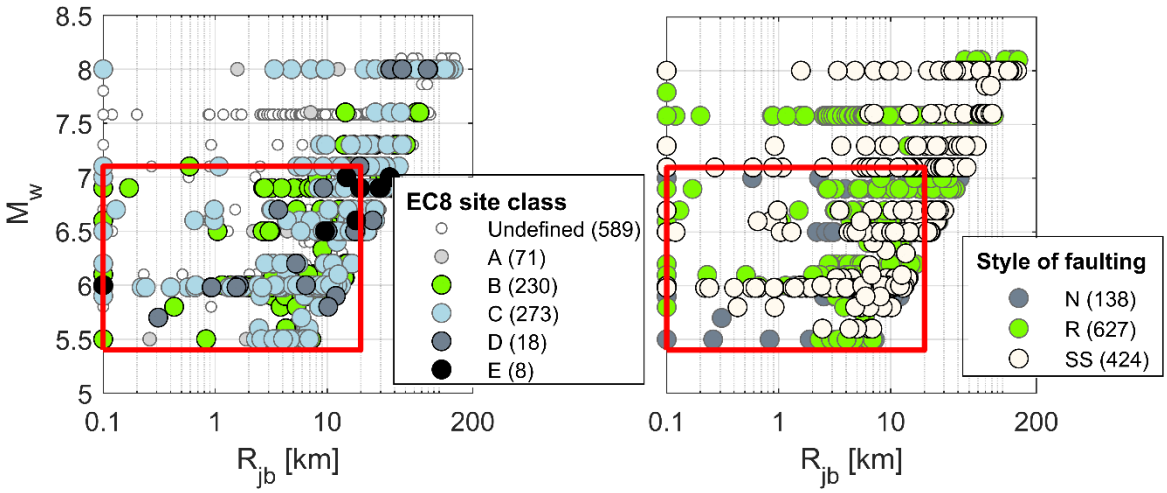


Figure 4 – $M_w$ and $R_{jb}$ distribution of NESS data with reference to EC8 site class (left panel) or Style of faulting (right panel, Normal, Reverse and Strike-Slip). The rectangle highlights the $M_w$-$R$ ranges used for the selection of the data in this study.

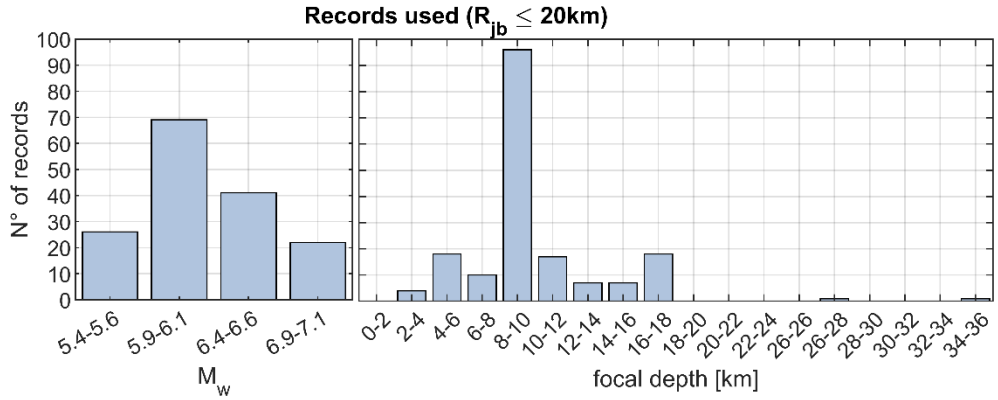


Figure 5 – Number of records used in the analyses with respect to different $M_w$ classes (left) or focal depths (right). Magnitude classes are centered respectively at 5.5, 6.0, 6.5 and 7.0. Note that only records at short distances from the fault (i.e. $R_{jb} \leq 20$ km) have been considered.

To increase data for the less populated $M_w$ classes (i.e., 5.5 and 6.5) and to have a statistically significant number of samples, some records from other databases have been used, namely: 12 records from the $M_w$ 5.4 Central Italy

earthquake of Oct 26, 2016 included in ITACA 3.0 (D'Amico et al. 2020) and 4 records from the $M_w$ 6.5 Jun 2000 Iceland sequence, included in SIMBAD v.06 (Smerzini et al. 2014).

## 4. Performance of selected GMMs against near-source records

The performance of the selected GMMs was tested for the horizontal components of motions, considering spectral acceleration values at 0 s (PGA), 0.5 s (SA05), 1 s (SA1), 2 s (SA2), in order to cover the short-to-long period range. Figure 6 shows some examples of comparison, for the four spectral accelerations, between the median GMM predictions for $M_w$ 6.0 and NESS records (dots) with $M_w$ = 5.9÷6.1; while Figure 7 shows the same comparison using the median prediction for $M_w$ 7.0 and NESS records with $M_w$ = 6.9÷7.1. In order for all the selected GMMs to have a proper basis of comparison, Figure 6 and Figure 7 show the predictions under the following assumptions: (i) 'unspecified' mechanism; (ii) $V_{s,30}$ = 600 m/s, representative of ground type B (for the models having $V_{s,30}$ as site proxy); (iii) suggested default values for any additional parameter required by the models.

In the case of SP96, the parametrized soil conditions of which cannot be directly associated to the ground type B, a default multiplication factor of 1.2 has been applied to the prediction on rock. For AMB96 and AD03 models, a conversion from $M_s$ to $M_w$ was applied, in agreement with that considered for the Italian seismic hazard map (MPS Working group, 2004). For SP96, as the magnitude scale adopted provides a good correspondence with $M_w$ (Idriss, 1991), no conversion was applied. Finally, note that for each GMM the comparison with records has been carried out considering the corresponding type of predicted intensity measure (Table 1), with no need of conversion.
Note that, following K20, we have distinguished the following focal depth (D) classes: 'shallow' (D < 10 km), 'intermediate' (D = 10÷20 km) and 'deep'(D ≥ 20 km), although, as shown in Figure 4, the NESS records are by far mostly populated in the 8÷10 km bin. The source-to-site distance was selected according to the specific GMM and the corresponding values provided in NESS.

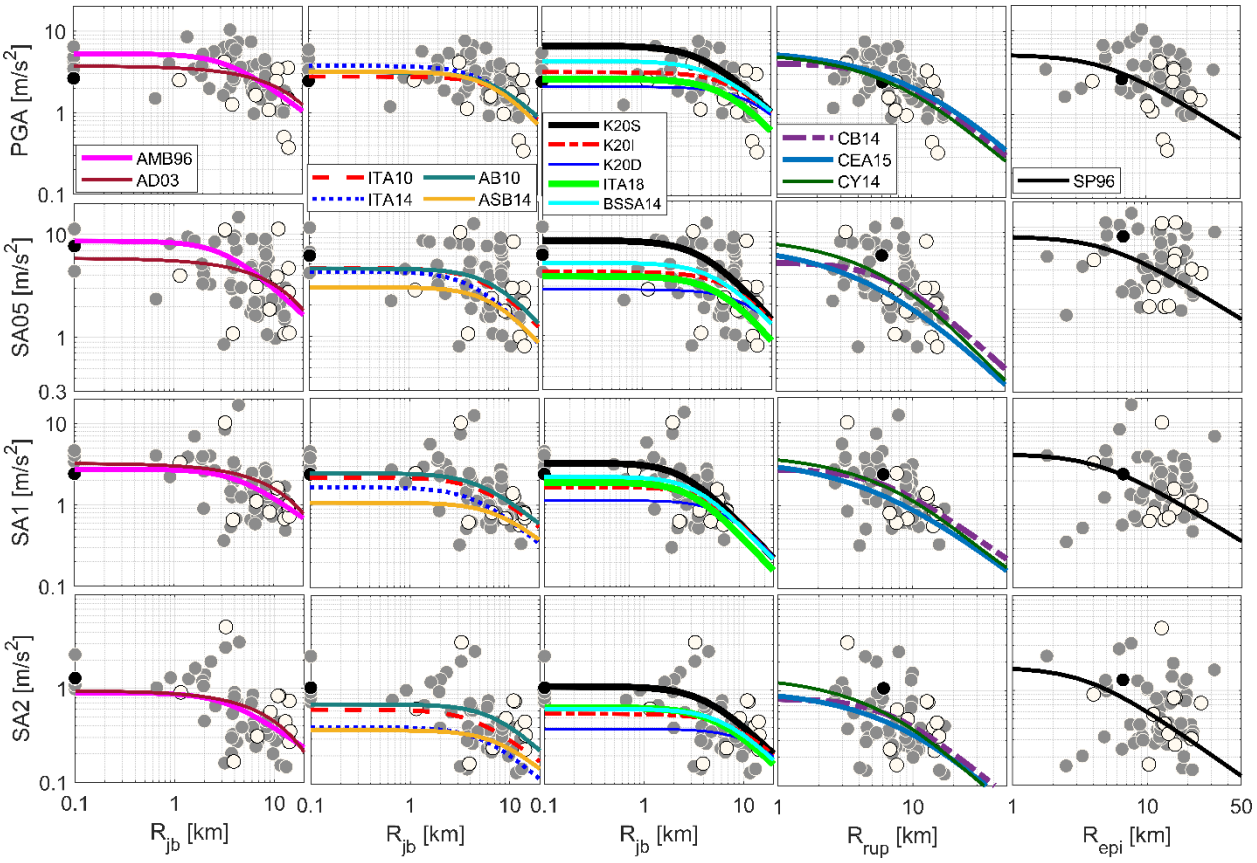

Figure 6 – Comparison between NESS records ($M_w$ = 5.9÷6.1) and median predictions of the selected GMMs for $M_w$ = 6.0. From top to bottom: PGA, SA05, SA1 and SA2. From left to right: AMB96 and AD03 ($R_{jb}$ and $H_{max}$); AB10, ITA10, ITA14, ASB14 ($R_{jb}$ and HGM); K20 'shallow' (K20S), K20 'intermediate' (K20I), K20 'deep' (K20D), ITA18, BSSA14 ($R_{jb}$ and RotD50); CB14, CY14 and CEA15 ($R_{rup}$ and HGM) and SP96 ($R_{epi}$ and $H_{max}$). Records symbols are distinguished by focal depth (grey dots, D < 10 km; white dots, D = 10÷20 km; black dots, D ≥ 20 km). The comparison between records and GMMs is shown for the corresponding type of predicted intensity measure (see Table 1). The only exception is in the fourth column (CB14, CY14 and CEA15), where records are shown in terms of HGM, while CB14 and CY14 predictions are in terms of RotD50.

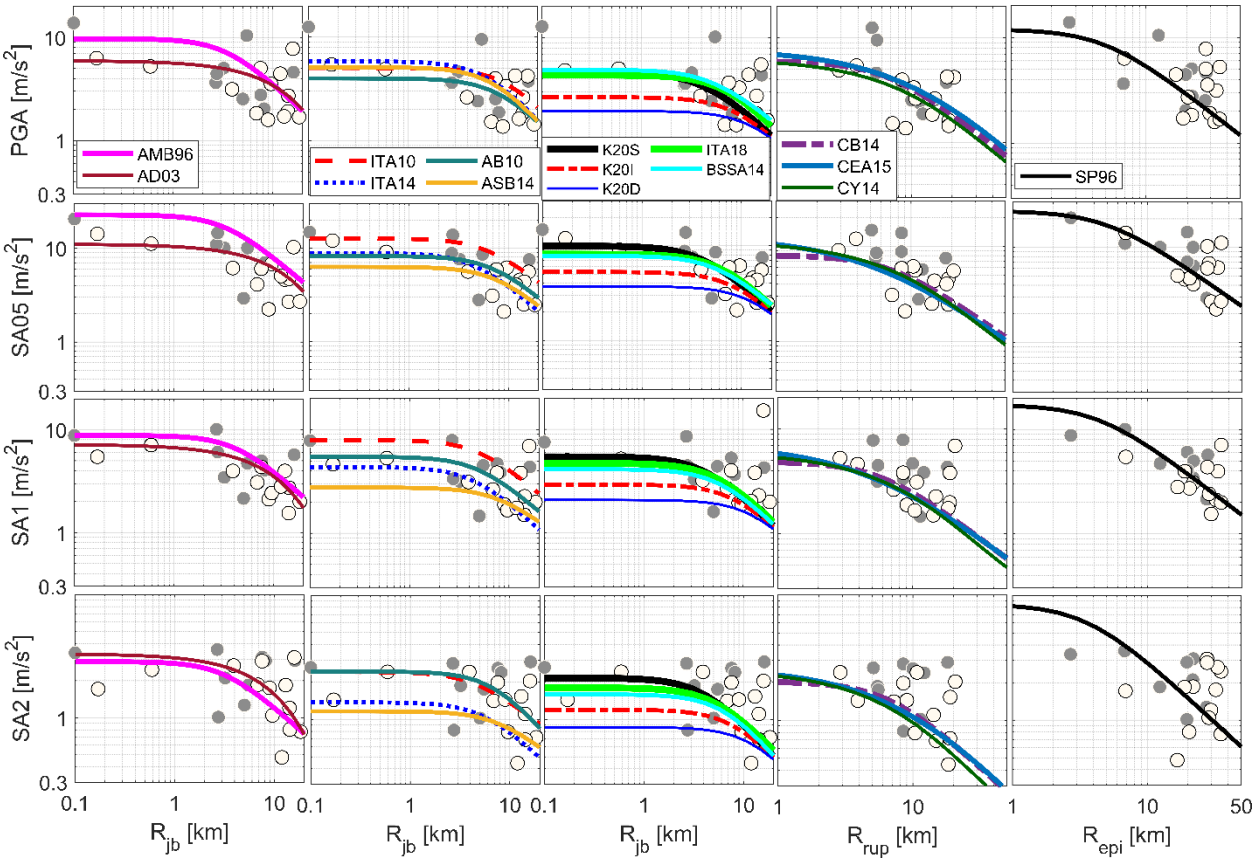


Figure 7 – Same as Figure 6, with NESS records ($M_w$ = 6.9÷7.1) and median GMMs predictions for $M_w$ = 7.0.

## 5. Calculation of residuals for different magnitude and distance bins for site class B

According to their general definition (Strasser et al., 2008) total residuals are the difference between the natural logarithm of records ($y_{es}$) and the corresponding prediction ($\mu_{es}$), for a given earthquake *e* at station *s*, obtained using a specific GMM. In order to analyze the performance of models as a function of either magnitude or distance, total residuals are separated into their between- and within -event components. Following the notation proposed by several authors (e.g., Al Atik et al., 2010; Rodriguez-Marek et al. 2011), the natural logarithm of the observed ground motion $y_{es}$ can be expressed as in Eq.1, where $\mu_{es}$ is the median ground motion value predicted by a specific GMM, $\delta B_e$ is the between-event residual, which corresponds to the average misfit of records from one particular earthquake with respect to the median ground motion model (Eq.2, where *NS* is the number of station recordings from one particular earthquake *e*), $\delta W_{es}$ is the within-event residual defined as the misfit between an individual observation at station *s* with respect to the event-corrected median estimate (Eq.3). Between- and within-event residuals are modelled using normal distributions and standard deviation τ and φ, respectively.

$$y_{es} = \mu_{es} + \delta B_e + \delta W_{es} \quad (1)$$

$$\delta B_e = \frac{1}{NS}\sum_{s=1}^{NS}(y_{es} - \mu_{es}) \quad (2)$$

$$\delta W_{es} = (y_{es} - (\mu_{es} + \delta B_e)) \quad (3)$$

In our analysis, we calculated the residuals by considering the proper $M_w$, style-of-faulting, $V_{s,30}$ (if not present in the database we used 600 m/s as representative of ground type B) and by taking extra parameters from the NESS database. In particular, for the GMMs requiring additional parameters, we made the following assumptions:

- for CB14, we took the hypocentral depth ($Z_{hyp}$), dip, rupture width ($W$), depth-to-top rupture ($Z_{tor}$), that were available from NESS; while we evaluated the depth to $V_s$ = 2.5 km/s horizon ($Z_{2.5}$) based on the correlations with $V_{s,30}$ provided in the NGAW2 spreadsheet (https://peer.berkeley.edu/peer-strong-ground-motion-databases);
- for CY14, we used the dip and depth-to-top rupture ($Z_{tor}$) that were available from NESS; we used zero for the directivity effect term (ΔDPP); $Rx$ was set negative so that no hanging wall effect was considered. Default values for the depth to $V_s$ = 1 km/s ($Z_{1.0}$) were calculated using eq. (1) and (2) of CY14;
- for BSSA14, we accounted for the specific regional anelastic attenuation parameter, although of minor relevance for near-source computations.

As an example, Figure 8 shows the computed $\delta B_e$ and $\delta W_{es}$ for the AMB96 model, together with the 16th, 50th and 84th percentiles of their normal distribution. Note that, as $\delta B_e$ can identify a systematic misfit of a specific GMM, with reference to a particular magnitude class, a positive $\delta B_e$ indicates a systematic underprediction of records, while a negative term indicates a systematic overprediction.

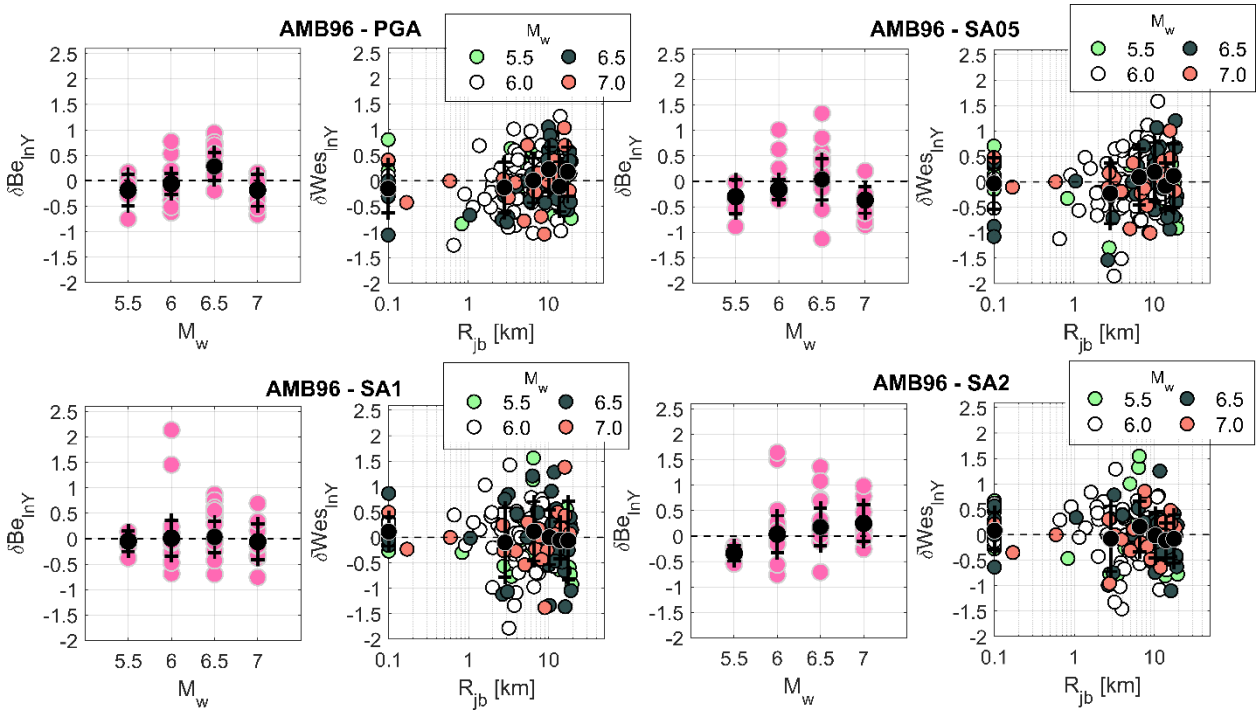


Figure 8 – PGA and SA05 (top), SA1 and SA2 (bottom) between- ($\delta B_e$) and within-event residuals ($\delta W_{es}$) computed for AMB96 model. For $\delta B_e$ four magnitude classes are shown: $M_w$ 5.5 ($M_w$ = 5.4÷5.6), $M_w$ 6.0 ($M_w$ = 5.9÷6.1), $M_w$ 6.5 ($M_w$ = 6.4÷6.6), $M_w$ 7.0 class ($M_w$ = 6.9÷7.1). For $\delta W_{es}$, percentiles are computed with bin of distances and residuals are shown for the four magnitude classes. Black dots and lines show the corresponding 16th, 50th and 84th percentiles of the corresponding normal distribution.

The whole set of between-event residuals calculated for all GMMs is shown in Figure 9 and Figure 10, for PGA and SA1. In addition, to give a more comprehensive overview of the performance of the analyzed GMMs, Figure 11 shows the mean computed $\delta B_e$ for all periods, magnitudes and GMMs. Note that the AD03 model was excluded from the calculations for $M_w$ 5.5 class, because its range of applicability is beyond $M_s$ = 5.8. Instead, the SP96 and ITA10 models were considered also for $M_w$ 7.0, because their calibration dataset includes records of the $M_w$ 6.9 1980 Irpinia earthquake, used for the validation.

In general, between-event residuals, and their corresponding scatter, show different patterns depending on the GMM under consideration. Considering for example the ITA18 model, PGA residuals suggest a systematic underprediction for almost all magnitude classes, with an overall positive mean $\delta B_e$ up to about +0.6 (in ln units) for $M_w$ 6.5 class (in PGA), that means an underestimation of recorded values by a factor of about 2. In some cases, such as for the previously noted PGA predictions at $M_w$ 6.5, all models tend to underestimate recorded values, that may be likely related to a specific trend of NESS records for this magnitude class.

Although the visual inspection of Figure 11 is sufficient to provide a meaningful overview of the performance of the different GMMs, a quantitative comparison of $\delta B_e$ values is shown in Table 2, in terms both of their average and of their maximum absolute values.

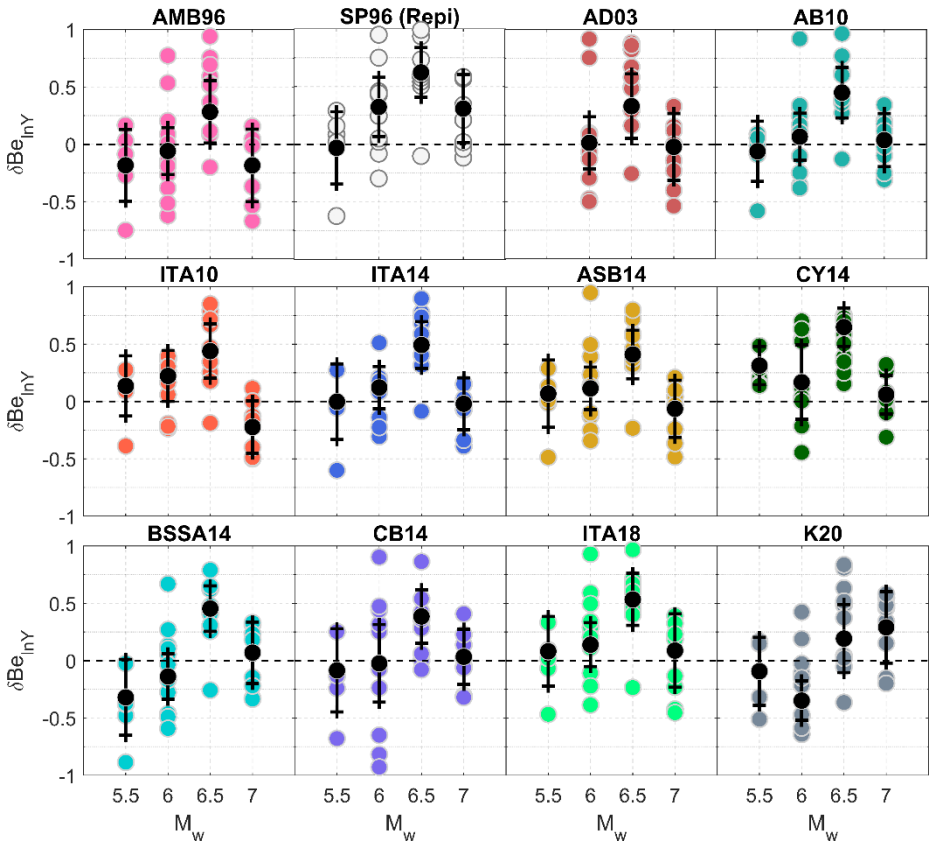


Figure 9 – PGA: computed between-event residuals ($\delta B_e$) for a selection of GMMs. Four magnitude classes are shown: $M_w$ 5.5 ($M_w$ = 5.4÷5.6), $M_w$ 6.0 ($M_w$ = 5.9÷6.1), $M_w$ 6.5 ($M_w$ = 6.4÷6.6), $M_w$ 7.0 class ($M_w$ = 6.9÷7.1). Black dots and lines show the corresponding 16$^{th}$, 50$^{th}$ and 84$^{th}$ percentiles of the respective normal distribution.

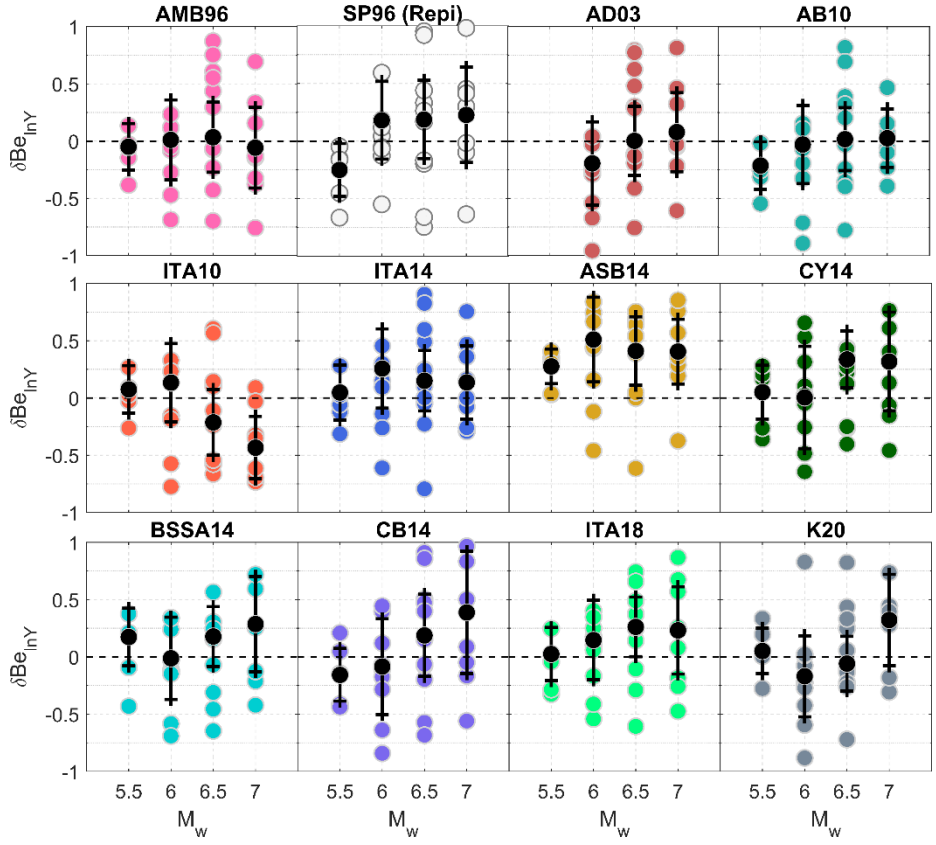


Figure 10 – Same as Figure 9 for SA1.

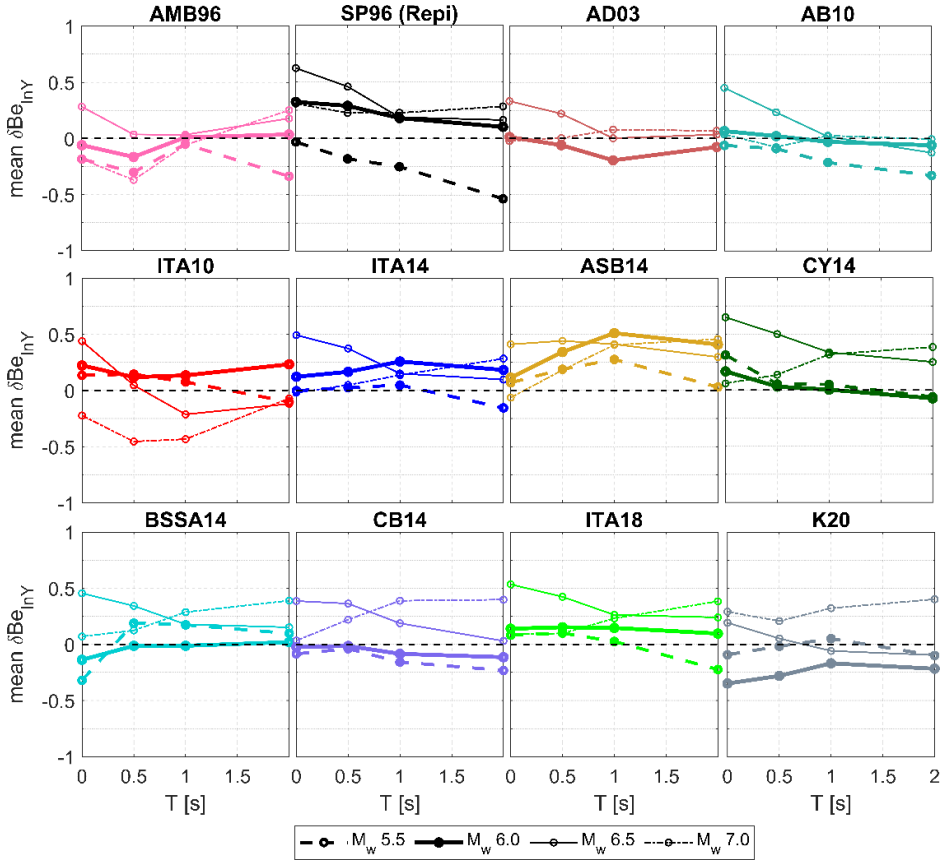


Figure 11 – Mean between-event residuals ($\delta B_e$) computed over the four spectral accelerations considered, for a selection of GMMs, and for the four adopted magnitude classes: $M_w$ 5.5 ($M_w$ = 5.4÷5.6), $M_w$ 6.0 ($M_w$ = 5.9÷6.1), $M_w$ 6.5 ($M_w$ = 6.4÷6.6), $M_w$ 7.0 class ($M_w$ = 6.9÷7.1).

## 6. Performance in terms of standard deviation

The performance check in the previous section was limited to the median prediction values. We aim in this section at checking if the models are suitable to quantify the observed variability of near-source records.
To this end, we computed the standard deviation of the within-event residuals $\delta W_{es\_j}$ of NESS records with respect to the *j-th* GMM, denoted by $\varphi_{NESS\_j}$, as follows:

$$\varphi_{NESS\_j} = \sqrt{\sum_{s=1}^{NS_e} \frac{\left(\ln(y_{es}) - \ln(a_{GMM_j}) - \delta B_{e_j}\right)^2}{NS_e - 1}} = \sqrt{\sum_{s=1}^{NS_e} \frac{(\delta W_{es\_j})^2}{NS_e - 1}} \quad (4)$$

where $y_{es}$ is the observed value at station $s$ from event $e$ (of either PGA or SA1), $a_{GMMj}$ is the median predicted acceleration from the *j-th* GMM, $\delta B_{e\_j}$ is the between-event residual computed with reference to the same model, for the same event and $NS_e$ is the number of recordings available for event $e$.

The $\varphi_{NESS\_j}$ values are compared in Figure 12 with the corresponding $\varphi_j$, that is the standard deviation of the within-event residuals of the *j-th* GMM. Note that for the older models (AMB96, SP96 and AD03) $\varphi_j$ was not computed, so the total standard deviation $\sigma_j$ was used for comparison.

Considering that $\varphi$ is generally slightly lower than $\sigma$, the performance of both AMB96 and AD03 models (first row of Figure 12) is good, while for SP96 ($R_{epi}$) the performance is fairly good for PGA, but not for the other periods, especially at higher magnitudes. For the more recent GMMs (second and third rows of Figure 12), there is a clear trend of $\varphi_{NESS\_j}$ values to be substantially lower than $\varphi_j$, suggesting that the variability implied by those GMMs tends to overestimate that observed in the near-source records, as expected since the size of the calibration dataset is higher. This is also visually confirmed in Figure 13, showing $\delta W_{es\_j}$ for the class with the highest number of records ($M_w$ 6.0), together with the standard deviation of AMB96, AD03, CB14, ITA10, ITA18 and K20 models.

Table 2 – Between-event ($\delta B_e$) residuals for all selected GMMs with respect to NESS dataset. Mean values for different magnitude classes are shown, for all periods. Each class includes magnitudes ± 0.1 around the central value. The 'Average $|(\delta B_e)|$' corresponds to the mean value of the between-event residuals (taken with their absolute values), for a given GMM, while 'Max $|(\delta B_e)|$' corresponds to its maximum absolute value.

| | | **$M_w$ class** | | | | | |
|---|---|---|---|---|---|---|---|
| **GMM** | **$\delta B_e$ mean** | **5.5** | **6.0** | **6.5** | **7.0** | **Average $\lvert\delta B_e\rvert$** | **Max $\lvert\delta B_e\rvert$** |
| **AMB96** | PGA | -0.183 | -0.059 | +0.283 | -0.184 | 0.158 | 0.369 |
| | SA05 | -0.301 | -0.165 | +0.036 | -0.369 | | |
| | SA1 | -0.050 | +0.011 | +0.034 | -0.058 | | |
| | SA2 | -0.337 | +0.035 | +0.178 | +0.251 | | |
| **SP96 ($R_{epi}$)** | PGA | -0.031 | +0.323 | +0.622 | +0.309 | 0.273 | 0.622 |
| | SA05 | -0.180 | +0.289 | +0.459 | +0.226 | | |
| | SA1 | -0.250 | +0.180 | +0.187 | +0.227 | | |
| | SA2 | -0.534 | +0.103 | +0.163 | +0.284 | | |
| **AD03** | PGA | | +0.013 | +0.332 | -0.023 | 0.092 | 0.332 |
| | SA05 | | -0.060 | +0.220 | +0.003 | | |
| | SA1 | | -0.195 | +0.002 | +0.078 | | |
| | SA2 | | -0.076 | +0.035 | +0.069 | | |
| **ITA10** | PGA | +0.135 | +0.222 | +0.439 | -0.223 | 0.198 | 0.455 |
| | SA05 | +0.144 | +0.118 | +0.046 | -0.455 | | |
| | SA1 | +0.075 | +0.134 | -0.213 | -0.434 | | |
| | SA2 | -0.103 | +0.232 | -0.120 | -0.071 | | |
| **AB10** | PGA | -0.061 | +0.067 | +0.451 | +0.037 | 0.115 | 0.451 |
| | SA05 | -0.090 | +0.022 | +0.234 | -0.076 | | |
| | SA1 | -0.213 | -0.029 | +0.017 | +0.026 | | |
| | SA2 | -0.329 | -0.060 | -0.126 | -0.005 | | |
| **ASB14** | PGA | +0.068 | +0.114 | +0.410 | -0.064 | 0.288 | 0.510 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | SA05 | +0.188 | +0.342 | +0.441 | +0.181 | | |
| | SA1 | +0.276 | +0.510 | +0.411 | +0.405 | | |
| | SA2 | +0.030 | +0.409 | +0.297 | +0.457 | | |
| **ITA14** | PGA | -0.002 | +0.122 | +0.492 | -0.021 | 0.160 | 0.492 |
| | SA05 | +0.024 | +0.165 | +0.373 | +0.048 | | |
| | SA1 | +0.048 | +0.257 | +0.150 | +0.136 | | |
| | SA2 | -0.156 | +0.182 | +0.096 | +0.283 | | |
| **BSSA14** | PGA | -0.320 | -0.138 | +0.454 | +0.070 | 0.185 | 0.454 |
| | SA05 | +0.188 | -0.013 | +0.342 | +0.124 | | |
| | SA1 | +0.173 | -0.013 | +0.178 | +0.287 | | |
| | SA2 | +0.097 | +0.019 | +0.152 | +0.389 | | |
| **CEA15** | PGA | +0.036 | +0.017 | +0.308 | -0.104 | 0.255 | 0.561 |
| | SA05 | +0.297 | +0.365 | +0.561 | +0.336 | | |
| | SA1 | +0.113 | +0.284 | +0.343 | +0.438 | | |
| | SA2 | -0.079 | +0.141 | +0.228 | +0.432 | | |
| **CB14** | PGA | -0.085 | -0.024 | +0.385 | +0.034 | 0.172 | 0.398 |
| | SA05 | -0.039 | -0.015 | +0.363 | +0.220 | | |
| | SA1 | -0.158 | -0.085 | +0.187 | +0.388 | | |
| | SA2 | -0.232 | -0.115 | +0.030 | +0.398 | | |
| **CY14** | PGA | +0.313 | +0.167 | +0.649 | +0.059 | 0.212 | 0.649 |
| | SA05 | +0.054 | +0.031 | +0.499 | +0.137 | | |
| | SA1 | +0.051 | +0.004 | +0.336 | +0.318 | | |
| | SA2 | -0.060 | -0.073 | +0.251 | +0.384 | | |
| **ITA18** | PGA | +0.081 | +0.139 | +0.534 | +0.088 | 0.201 | 0.534 |
| | SA05 | +0.100 | +0.152 | +0.423 | +0.098 | | |
| | SA1 | +0.025 | +0.146 | +0.262 | +0.231 | | |
| | SA2 | -0.224 | +0.096 | +0.240 | +0.383 | | |
| **K20** | PGA | -0.092 | -0.349 | +0.192 | +0.291 | 0.181 | 0.401 |
| | SA05 | -0.018 | -0.281 | +0.052 | +0.207 | | |
| | SA1 | +0.051 | -0.169 | -0.059 | +0.322 | | |
| | SA2 | -0.100 | -0.218 | -0.096 | +0.401 | | |

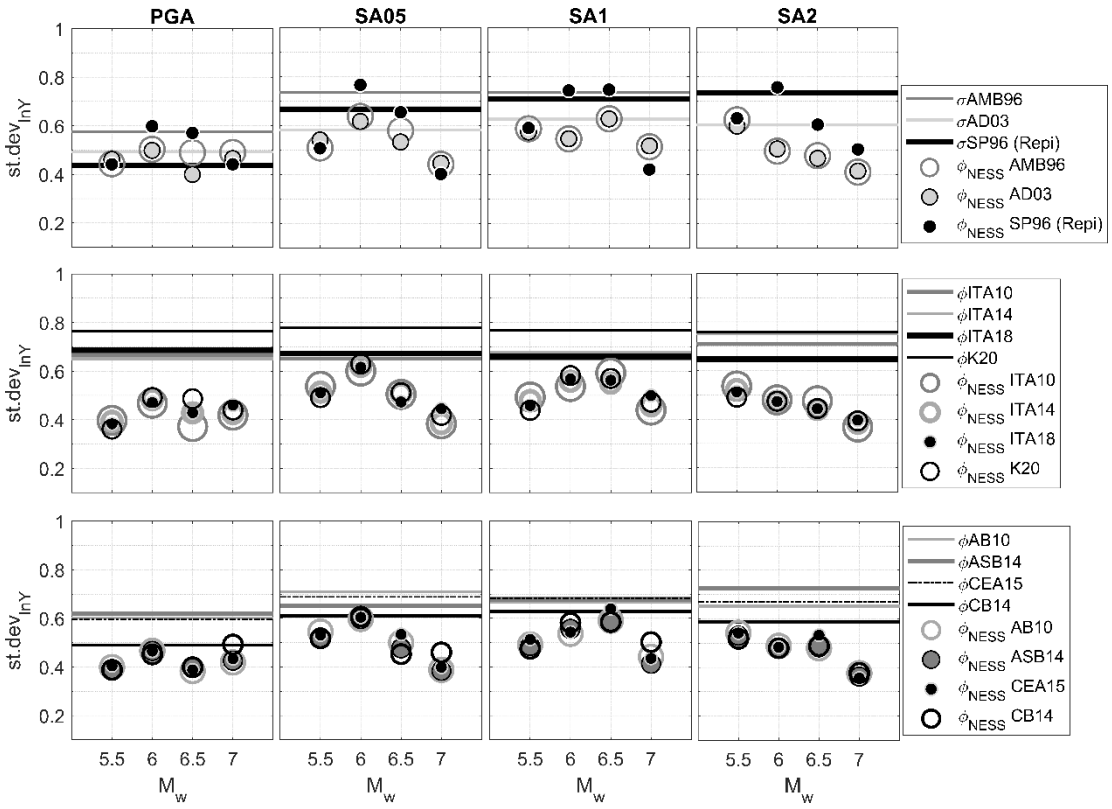


Figure 12 – Observed standard deviations $\varphi_{NESS\_j}$ for some of the GMMs under study, as a function of the four considered magnitude classes, compared to the standard deviations of each model, shown by thick lines. $\sigma$ is used for the total standard deviation, $\varphi$ for the within event standard deviation (whenever available). From left to right: PGA, SA05, SA1, SA2.

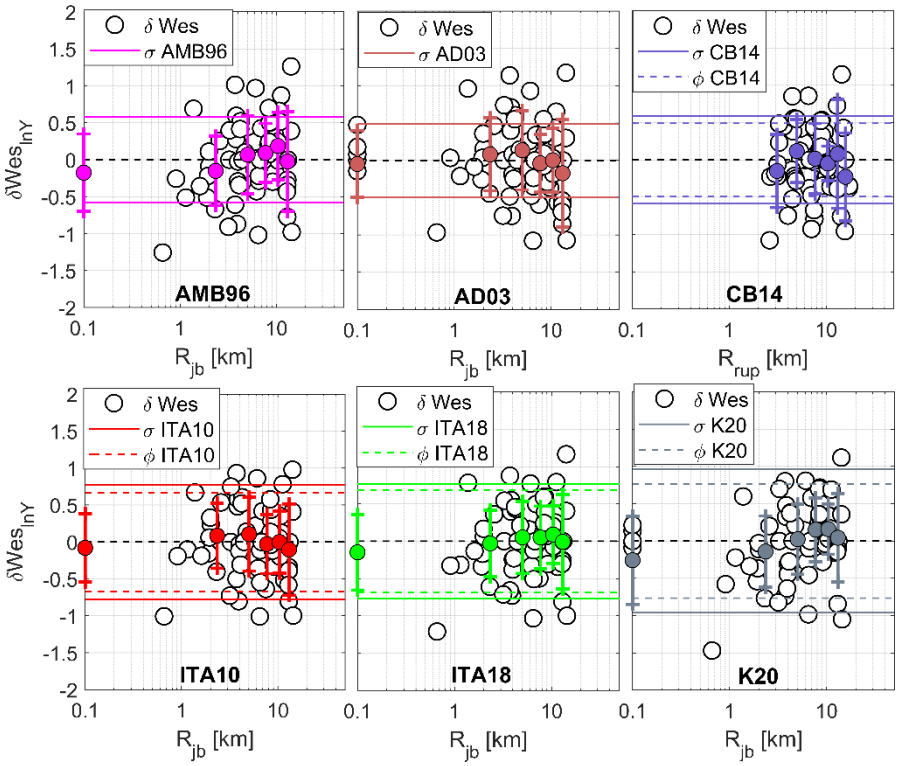


Figure 13 – PGA: Within-event residuals ($\delta W_{es}$) for $M_w$ 6.0 class, computed using different GMMs, with 16th, 50th and 84th percentiles (with bin of distances) of the corresponding normal distribution. Solid lines represent the total standard deviation (σ), dashed lines the within-event standard deviation (φ), when available, of the corresponding GMM.

## 7. Conclusions

In this paper, we provide a quantitative evaluation of the performance of a set of GMMs in the near-source area of moderate-to-large earthquakes ($M_w$ = 5.5÷7.0 and Joyner-Boore source-to-site distance $R_{jb} \leq 20$ km), selected because it dominates hazard in the high seismicity Italian regions, and, therefore it is crucial for accurate PSHA evaluations at long return periods, for reliable determination of seismic actions for design, and for a proper definition of seismic scenarios for risk analyses. GMMs were tested taking advantage of the NESS dataset, consisting of worldwide near-source records with high-quality metadata and intensity measures. Through an analysis of residuals extended to intensity measures representative of different period ranges (PGA, SA05, SA1, SA2), we investigated the misfit of the considered GMMs with respect to records and we compared the performance of the different models. The goal of this test is not to make an absolute ranking of GMMs in the selected *M-R* range, that would imply consideration of a much wider set of GMMs and the use of more advanced ranking approaches (e.g., Scherbaum et al., 2004; Kale and Akkar, 2013), but to check the actual predictive capability of GMMs and to identify possible systematic bias with respect to the newly available near-source records NESS dataset.

Our results are somehow unexpected, as it turned out that, according to the simple criteria considered in this paper, the AMB96, dating back about 25 years, and the AD03 models, dating about 20 years, are among the best performing GMMs in the selected *M-R* range, both in terms of median values and standard deviation. While the excellent performance of AD03 could be predicted, although not granted, because it was specifically devised for near-source conditions and it took advantage of the important sets of records from the 1999 Turkey and Taiwan earthquakes, that of AMB96 was not obvious, also considering that most of the NESS dataset does not contain records used for its calibration.

To explain these results, it should be considered that, although the calibration datasets of newer models are growing continuously and very fast in low amplitude motions, the number of near-source high amplitude records is increasing at a much lower rate, as previously shown in Figure 1. As a consequence, while the performance of new GMMs for low-magnitude large-distance earthquake ground motions is expected to steadily improve, a similar improvement is not granted in the near-source region of moderate-to-large earthquakes, that most often govern the seismic hazard assessment for the highest seismicity regions. Instead, the percentage of near-source records over the whole calibrations sets of older GMMs is larger than it is in the newer ones, and this may be a reason for the remarkably good performance of the AMB96 and AD03 models.

Although the title of this paper may be provocative, the main intent is of course not to claim that recent GMMs should be replaced by the old ones. Rather, it is to focus on the need for improvement of the tools of ground motion prediction in the near-source conditions, that are often overlooked in the development of empirical GMMs and that are instead the key, for most highly seismic regions, for a proper estimation of hazard and risk.

With this objective, it should be pointed out that the NESS dataset itself is still sparse and not sufficiently extended to cover the variety of fault slip distributions, fault geometry details, complex geological configurations that are expected to affect strong ground motion in near-source conditions. This variety is hardly sampled by the available near-source records, the number of which, as noted before, is not expected to grow significantly in the next few years. In these conditions, the most rationale way to progress towards advanced modelling of near-source ground motions seems to take advantage of the increasingly accurate broadband results of physics-based numerical simulations, as a numerical lab providing realistic realizations of future earthquake ground motions in realistic 3D geological configurations. For this purpose, simulated broadband accelerograms may successfully complement the sparse availability of records in near-source conditions, as in the BB-SPEEDset, recently introduced by Paolucci et al. (2021) based on several validated ground motion scenarios from Italian and worldwide earthquakes, or in the subset of CyberShake ground-motion time series, for the Los Angeles area, selected by Baker et al. (2021).

## 8. References


Akkar S, Bommer JJ (2010) Empirical equations for the prediction of PGA, PGV and spectral accelerations in Europe, the Mediterranean region and the Middle East. Seismol Res Lett 81:195–206. https://doi.org/10.1785/gssrl.81.2.195

Akkar S, Sandikkaya MA, Bommer JJ (2014a) Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. Bull Earthq Eng 12:359–387. https://doi.org/10.1007/s10518-013-9461-4

Akkar S, Sandıkkaya MA, Senyurt M, Azari AS, Ay BÖ (2014b) Reference database for seismic ground-motion in Europe (RESORCE). Bull Earthq Eng 12:311-339. https://doi.org/10.1007/s10518-013-9506-8

Al Atik L, Abrahamson NA, Bommer JJ, Scherbaum F, Cotton F, and Kuehn N (2010) The variability of ground-motion prediction models and its components. Seismol Res Lett 81: 794–801. https://doi.org/10.1785/gssrl.81.5.794

Ambraseys NN, Simpson KA, Bommer JJ (1996) Prediction of horizontal response spectra in Europe. Earthq Eng Struct Dyn 25: 371 – 400

Ambraseys NN and Douglas J (2003) Near-field horizontal and vertical earthquake ground motions. Soil Dyn and Earthq Eng 23:1–18. https://doi.org/10.1016/S0267-7261(02)00153-7

Baker JW, Rezaeian S, Goulet CA, Luco N, Teng G (2021) A subset of CyberShake ground-motion time series for response-history analysis. Earthq Spectra 37:1162–1176. https://doi.org/10.1177/8755293020981970

Barani S, Spallarossa D, Bazzurro P (2009) Disaggregation of probabilistic ground-motion hazard in Italy. Bull Seismol Soc Am 99:2638-2661. https://doi.org/10.1785/0120080348

Bindi D, Pacor F, Luzi L, Puglia R, Massa M, Ameri G and Paolucci R (2011) Ground motion prediction equations derived from the Italian strong motion database. Bull Seismol Soc Am 9:1899–1920. https://doi.org/10.1007/s10518-011-9313-z

Bindi D, Massa M, Luzi L, Ameri G, Pacor F, Puglia R, Augliera P (2014) Pan – European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5% - damped PSA at spectral periods up to 3.0s using the RESORSE dataset. Bull Earthq Eng 12:391– 430. https://doi.org/10.1007/s10518-013-9525-5

Boore DM (2010) Orientation – independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. Bull Seismol Soc Am 100:1830 – 1835. https://doi.org/10.1785/0120090400

Boore DM, Stewart JP, Seyhan E, Atkinson GM (2014) NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. Earthq Spectra 30:1057–1085. https://doi.org/10.1193/070113EQS184M

Campbell KW and Bozorgnia Y (2014) NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. Earthq Spectra 30: 1087 – 1115. https://doi.org/10.1193/062913eqs175m

Cauzzi C, Faccioli E, Vanini M and Bianchini A (2015) Updated predictive equations for broadband (0.01 – 10s) horizontal response spectra and peak ground motions, based on a global dataset of digital acceleration records. Bull Earthq Eng 13: 1587 – 1612. https://doi.org/10.1007/s10518-014-9685-y

Chiou BSJ, Youngs RR (2014) Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra. Earthq Spectra 30:1117-1153. https://doi.org/10.1193/072813EQS219M

Comité Européen de Normalisation (CEN) (2004) Eurocode 8: design of structures for earthquake resistance-Part 1: general rules, seismic actions and rules for buildings. Comité Européen de Normalisation, Brussels.

D'Amico M, Felicetta C, Russo E, Sgobba S, Lanzano G, Pacor F, Luzi L (2020) Italian Accelerometric Archive v 3.1 – Istituto Nazionale di Geofisica e Vulcanologia, Dipartimento della Protezione Civile Nazionale. https://doi.org/10.13127/itaca.3.1

Douglas J and Edwards B (2016) Recent and future developments in earthquake ground motion estimation. Earth Sci Rev 160:203-219. https://doi.org/10.1016/j.earscirev.2016.07.005

Idriss IM (1991) Selection of earthquake ground motions at rock sites. Report prepared for the Structures Division, Building and Fire Research Laboratory, Department of Civil Engineering, University of California, Davis

Kale Ö ve Akkar S (2013) A New Procedure for Selecting and Ranking Ground-Motion Prediction Equations (GMPEs): The Euclidean Distance-Based Ranking (EDR) Method. Bull Seismol Soc Am 103:1069-1084. https://doi.org/10.1785/0120120134

Kotha SR, Weatherill G, Bindi D et al. (2020) A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe. Bull Earthq En 18: 4091 – 4125. https://doi.org/10.1007/s10518-020-00869-1

Lanzano G, Luzi L, Pacor F, Felicetta C, Puglia R, Sgobba S, D'Amico M (2019) A revised ground-motion prediction model for shallow crustal earthquakes in Italy. Bull Seismol Soc Am 109:525–540. https://doi.org/10.1785/0120180210

Luzi L, Lanzano G, Felicetta C, D'Amico MC, Russo E, Sgobba S, Pacor F, and ORFEUS Working Group 5 (2020) Engineering Strong Motion Database (ESM) (Version 2.0). Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/ESM.2

MPS Working Group (2004) Redazione della Mappa di Pericolosità Sismica Prevista dall'Ordinanza PCM del 20 Marzo 2003 n.3274, All. 1, Rapporto Conclusivo per il Dipartimento della Protezione Civile, Istituto Nazionale di Geofisica e Vulcanologia (INGV), Milano-Roma, Italy, Aprile 2004

Pacor F, Felicetta C, Lanzano G, Sgobba S, Puglia R, D'Amico M, Russo E, Baltzopoulos G, Iervolino I (2018) NESS v1.0: A worldwide collection of strong-motion data to investigate near source effects. Seismol Res Lett. https://doi.org/10.1785/0220180149

Paolucci R, Smerzini C, and Vanini M (2021) BB-SPEEDset: A Validated Dataset of Broadband Near-Source Earthquake Ground Motions from 3D Physics-Based Numerical Simulations. Bull Seismol Soc Am, in press. https://doi.org/10.1785/0120210089

Rodriguez-Marek A, Montalva GA, Cotton F, Bonilla F (2011) Analysis of single-station standard deviation using the KiK – net data. Bull Seismol Soc Am 101:1242 – 1258. https://doi.org/10.1785/0120100252

Sabetta F and Pugliese A (1996) Estimation of response spectra and simulation of nonstationary earthquake ground motions. Bull Seismol Soc Am, 86: 337–352. https://doi.org/10.1785/BSSA0860020337

Scherbaum F, Cotton F, Smit P (2004) On the Use of Response Spectral-Reference Data for the Selection and Ranking of Ground-Motion Models for Seismic-Hazard Analysis in Regions of Moderate Seismicity: The Case of Rock Motion. Bull Seismol Soc Am 94:2164–2185. https://doi.org/10.1785/0120030147

Sgobba S, Felicetta C, Lanzano G, Ramadan F, D'Amico M, Pacor F. (2021) NESS2.0: an updated version of the worldwide dataset for calibrating and adjusting ground motion models in near-source. Istituto Nazionale di Geofisica e Vulcanologia (INGV) https://doi.org/10.13127/NESS.2.0

Smerzini C, Galasso C, Iervolino I, Paolucci R (2014) Ground Motion Record Selection Based on Broadband Spectral Compatibility. Earthq Spectra 30:1427–1448. https://doi.org/10.1193/052312EQS197M

Strasser FO, Bommer JJ and Abrahamson NA (2008) Truncation of the distribution of ground-motion residuals. J Seismol 12:79-105. https://doi.org/10.1007/s10950-007-9073-z