Near-field spectral displacements from observed and synthetic data: an attempt at merging two complementary approaches

C. Cauzzi & L. Dalguer
Swiss Seismological Service (SED-ETHZ), Zürich, Switzerland

SUMMARY:
As current empirical ground motion prediction equations (GMPEs) are often recognised to suffer from insufficient data coverage in the near-field, we investigate herein the complementary use of ground motion data originated from suites of physics-based synthetic earthquake ruptures. We use the recently compiled deterministic synthetic strong-motion databank ($T > 1$ s) of Dalguer and Mai (2011) and compare it with state-of-the-art GMPEs largely used at European level and worldwide, with special focus on the broadband spectral displacement model of Cauzzi and Faccioli (2008) and subsequent updates and the NGA model of Boore and Atkinson (2008).

Keywords: displacements, near-field, ground-motion prediction equations

1. INTRODUCTION

The recent earthquake-related urban disasters of L’Aquila (Italy) 2009 and Christchurch (New Zealand) 2011 have drawn increased interest of the earthquake engineering and seismological communities in the peculiar features of recorded near-source ground motions for moderate earthquakes ($M_W 6+$), exhibiting potentially destructive mixtures of high-frequency and relatively long-period energy content, with important implications for eventual update of current seismic building codes and for proper understanding of seismic source processes. Following the progressive worldwide introduction of high dynamic range strong-motion sensors and dataloggers and the intensive research carried out by the engineering community on performance based design approaches, recent trends in earthquake engineering practice propose displacements, rather than acceleration/forces, as the key parameter controlling damage to buildings (Priestley et al., 2007). Furthermore, displacements are the primary input to design of isolation systems for critical structures such as nuclear power plants. Computing long-period displacement spectral ordinates from strong-motion data poses some difficulties related to the long-period components of noise contaminating the records. This makes the results of numerical simulations ideal candidates to better constrain the levels of long-period spectral demands although, with modern strong-motion sensors writing on 24-bit dataloggers, long-period spectral ordinates up to $\sim 10$ s can be reliably computed from earthquake data without filtering and adopting very simple correction techniques (see Paolucci et al., 2008; Akkar and Boore, 2009; Cauzzi and Clinton, 2012). As current empirical ground motion prediction equations (GMPEs) are often recognised to suffer from insufficient data coverage in the near-field, we investigate herein the complementary use of ground motion data originated from suites of physics-based synthetic earthquake ruptures. We use the recently compiled deterministic synthetic strong-motion databank ($T > 1$ s) of Dalguer and Mai (2011) and compare it with state-of-the-art GMPEs largely used at European level and worldwide, with special focus on the very broadband spectral displacement model of Cauzzi and Faccioli (2008) and the NGA model of Boore and Atkinson (2008). The use of synthetic allows to easily overlooking the difficulties posed by the paucity of real data in the near-field. Aimed at creating a wide range of scenarios for evaluating near-source ground motion variability, the synthetic databank comprises $\sim 360$ rupture dynamic simulations in the range of $M_w \sim 5.5 – 7.0$, with strike-slip, reverse and normal style–faulting, as well as buried and surface rupturing. By means of the comparison with current GMPEs, we investigate the feasibility of
including synthetic data in empirical GMPEs to better constrain the latter in the near-source region and discuss the eventual development of a near-field GMPE entirely based on simulations. This contribution represents an attempt at studying the feasibility of calibrating a hybrid GMP model for the near-field, for shallow-crustal earthquakes in the magnitude range of interest for Europe, suitable to be used for the next generation of tools for seismic hazard analyses. The approach of the present study is similar to that of the Italian DPC-funded research project S2 (http://nuovoprogettoesse2.stru.polimi.it), where 3D numerical analyses of ground motion ($f < 2.5$ Hz) were performed with the spectral element code GeoELSE (http://geoelse.stru.polimi.it; Faccioli et al., 1997; Stupazzini et al., 2009) and used, after proper combination with those from finite-fault stochastic simulations carried out with the program EXSIM (Motazedian and Atkinson, 2005) to make the synthetic signals broadband, as the basic input to probabilistic seismic hazard assessment in the near-field.

2. THE SYNTHETIC DATASET

In this study we use the synthetic dataset of Dalguer and Mai (2011), constituted by 360 dynamic rupture models in the $M_W$ magnitude range 5.5 - 7, for three possible styles of faulting (reverse, normal and strike-slip with dip angle equal to 45°, 60° and 90° respectively), for both buried ruptures and capable faults (breaking the earth surface). Stress and frictional strength in the numerical models were taken as representative of two extreme cases of normal stress, i.e. depth dependent and depth-independent. The dynamic rupture of the synthetic scenarios occurs as dictated by the local stress conditions and follows the slip-weakening friction model of Andrews (1976). Calculations were developed using the Support Operator Rupture Dynamics code (SORD, Ely et al., 2008, 2009, based on a generalised Finite Difference scheme that can use computational meshes of arbitrary structure and incorporate irregular geometry, with the capability to model general fault geometry and topography. The code is parallelised, using Message Passing Interface (MPI), for multiprocessor execution, and is highly scalable, enabling large-scale earthquake simulations. The dynamic rupture model has been validated through the Southern California Earthquake Centre (SCEC) dynamic rupture code validation exercise, showing good agreement with semi-analytical boundary integral methods (Harris et al., 2009). The synthetic data distribution in the dataset with respect to magnitude, distance and focal mechanism is shown in Fig. 1.1. The distribution of data with respect to $R_{JB}$ (the distance of the receiver from the surface projection of the ruptured fault, Joyner and Boore, 1981) is uniform within 50 km from the earthquake source, while correctly no are data available for $M_W < 6$ and $R_{RUP} < 10$ km ($R_{RUP}$ is the closest distance of the receiver from the ruptured fault plane), due to the presence of earthquake scenarios where the fault, buried at 5 km, is not allowed to break the surface. The total amount of 3-component synthetics in the dataset is 60480. The source-receiver geometry of a typical synthetic earthquake scenario is shown given Fig. 1.2, for a $M_W$ 6.84 event with reverse style-of-faulting, breaking the surface. The rectangle with red dashed border in Fig. 1.2 is the surface projection of the ruptured fault plane, dipping towards the strike-normal direction. Note the densification of the receivers in the area where the fault breaks the surface of the computational model. Also shown in Fig. 1.2 is the epicentre (red star) and the contour lines of rupture time (black curves): the rupture starts at the hypocentre and propagates towards the surface with average rupture velocity $V_R = 1.95$ kms$^{-1}$. 
Figure 2.1. Magnitude, distance and style-of-faulting distribution for the records in the synthetic dataset used in this study.

Figure 2.2. Example of source-receiver geometry for a $M_W 6.84$ reverse fault scenario in the synthetic dataset. Note the locations of the receivers (blue symbols), the surface projection of the maximum allowed ruptured fault (rectangle with red dashed border), the epicentre (red star) and the rupture time contour lines (black curves).

Site-amplification corrections using the period dependent amplification coefficient of Borcherdt (1994, 2002) were applied to correct the amplitude of computed ground motions to $V_{S30} = 800 \text{ ms}^{-1}$ (the boundary between ground type A and B in the Eurocode 8 – CEN, 2004) from the minimum shear-wave velocity assumed in the crustal model used in the simulations ($V_{S30} = 2500 \text{ ms}^{-1}$). This simple correction is supposed to ease the comparison with existing empirically based ground motion prediction equations (GMPEs) that are based on data at recording sites with $V_{S30}$ hardly exceeding $1100 \text{ ms}^{-1}$. No filtering was applied to the synthetic accelerograms in order to preserve the long-period components of the synthetics, with particular reference to the permanent ground displacements. The
horizontal seismic action in the present study is represented by the simple geometric mean of the elastic 5%-damped relative displacement response spectra \( DRS(T; 5\%) \) of the two simulated horizontal components, computed over the period range \( 0.5 \, \text{s} < T < 20 \, \text{s} \). The same definition is used by Cauzzi and Faccioli (2008), while Boore and Atkinson (2008) used GMRot50 (Boore et al., 2006), independent on the sensor at the recording site. The two intensity measures are expected to exhibit maximum differences of \( \sim 3\% \), and close to 1% on average (Boore and Atkinson, 2007; Campbell and Bozorgnia, 2007).

3. REGRESSIONS ON SYNTHETIC DATA

The following functional form was initially chosen for the regressions on the synthetic dataset:

\[
\log_{10} DRS(T; \zeta) = c_1 + m_1 M_W + m_2 M_W^2 + (r_1 + r_2 M_W) \log_{10}(d + 3 \times 10^{0.6M_W}) + \epsilon \tag{3.1}
\]

where \( DRS(T; \zeta) \) is the geometric mean of the Displacement Response Spectra (in m) computed from the acceleration traces of the two orthogonal horizontal components of ground motion at each receiver. \( \zeta = 5\% \) is the damping ratio, \( T \) is the vibration period, \( M_W \) the moment magnitude, and \( d \) is either the fault distance \( (R_{RUP}) \) or the Joyner-Boore distance \( (R_{JB}) \). \( c_1, m_1, \ldots, 2, r_1, \ldots, 4 \) are coefficients to be determined through non-linear regressions on the available data, while \( \epsilon \) is a random error term assumed as normally distributed with zero mean and standard deviation \( \sigma_{\log DRS} \). The functional form of Eqn. 3.1 with distance saturation term dependent on magnitude (Fukushima and Tanaka, 1990; Kanno et al., 2006, Cauzzi et al., 2011) was preferred with respect to the one featuring a fictitious depth \( h \) adopted e.g. by Boore and Atkinson (2008) since Eqn. 3.1 can better fit strong-motion observations in the near-field region of large earthquakes (Faccioli et al., 2010a). Attenuation with distance was initially made to depend also on magnitude, to allow for the distance decay to be different for weak and strong events (Ambraseys et al., 2005; Akkar and Bommer, 2007a-b and 2010; Boore and Atkinson, cit.). Due to the limited distance range and the relatively long periods of interest, we did not explicitly model an additional dissipative term to account for anelastic decay with distance at long periods. Magnitude saturation was explicitly modelled by means of a \( M_W^2 \) term, with coefficient \( m_2 \) negative from regressions (Fukushima, 1996; Douglas, 2002). The results of the exploratory regressions carried out with Eqn. (3.1) led to assume a) \( m_2 = 0 \), b) \( r_2 = 0 \) and c) \( r_1 = 1 \) over the entire period range. a) and b) are consistent with the findings of Cauzzi and Faccioli (2008) and c) was dictated by the observed trade-off between \( r_1 \) and \( r_2 \) and the lack of physical meaning of the shape of \( r_2 \) as a function of \( T \). Eqn. (3.1) therefore simplified to:

\[
\log_{10} DRS(T; \zeta) = c_1 + m_1 M_W + r_1 \log_{10}(d + 10^{0.6M_W}) + \epsilon \tag{3.2}
\]

Regressions were carried out using the non-linear least-squares problem solver \textit{lsqcurvefit} available in Matlab®. Due to the uniform distribution of data with respect to magnitude and distance, we did not use the one-stage maximum-likelihood method of Joyner and Boore (1993 and 1994) that, tested on a few spectral ordinates, turned out to be extremely time-consuming due to the very large dimensions of the dataset. The dependence of the model parameters of Eqn. (3.2) on vibration period \( T \) is depicted in Fig. 3.1. Note in particular the scaling with moment magnitude, approaching values close to 1 at long periods. The distance saturation term (dependent on magnitude), was found to reach its maximum levels between 5 s and 7 s, equal to \( \sim 11 \, \text{km} \) and \( \sim 18 \, \text{km} \) for the \( R_{RUP} \)-based model and \( R_{JB} \) based model respectively, and \( M_W = 7 \).
The scaling with magnitude and distance determined as described above, we investigated the possibility of introducing style-of-faulting terms (Boomer et al., 2003) into the prediction model, by means of a source scaling term taking the form (Boore and Atkinson, cit.; Cauzzi and Faccioli, cit.):

\[ f_{\text{mech}} = c_1 + m_1 M_W + a_N E_N + a_R E_R + a_S E_S \]

(3.3)

where \( E_X \) (\( X=N,R,S \)) are dummy variables for normal, reverse, and strike-slip fault mechanisms, respectively. Data from all fault types were first grouped together to determine \( a_1 \) and \( m_1 \). Two-stage, maximum-likelihood regressions (Joyner and Boore, cit.) were then performed fixing the coefficients \( a_1 \) and \( a_2 \) and solving for the coefficients \( a_N, a_R, \) and \( a_S \) of the fault type explanatory variables. This leads to constraining the relative scaling of amplitudes with magnitude to be the same for all earthquakes, allowing an offset in the average predicted amplitude level according to the fault mechanism (Boore and Atkinson, cit.; Cauzzi and Faccioli, cit.). The effect of the introduction of style-of-faulting terms in the spectral predictions is shown in Fig. 3.2. As expected, while predicting spectral ordinates for normal fault earthquakes would always lead to a reduction of spectral values with respect to Eqn. (3.2), reverse fault motions are higher at periods less than 4 s, with maximum amplification of \( \sim 20\% \) for the \( R_{\text{JB}} \)-based model. The reduction in the standard error of the prediction obtained with Eqn. (3.3) over the whole period range explored, fully justifies the use of style-of-faulting terms in developing a prediction model based on the synthetic dataset at hand. Examples of the residuals of predictions given by combining Eqn. (3.2) and Eqn. (3.3) using \( R_{\text{RUP}} \) are depicted in Fig. 3.3 as a function of magnitude and distance for two selected spectral ordinates at short and long periods (\( T = 1 \) s and \( T = 10 \) s). While no clear trend can be recognised if the whole magnitude and distance range are considered, a slight underestimation of the synthetics spectra seems to occur at \( M_W \sim 7 \).
Figure 3.2. Effect of the explicit introduction of style-of-faulting terms in the prediction equations by means of Eqn. (3.3). Note the reduction in the standard error of the predictions (rhs plots).

Figure 3.3. Residuals of the prediction model Eqn. (3.2) and (3.3) as a function of magnitude and distance at short ($T = 1\, \text{s}$) and long periods ($T = 10\, \text{s}$). Red lines represent the $\pm \sigma$ bands of the prediction.

4. COMPARISON WITH EXISTING GMPEs

In the present section, the prediction equations derived using the synthetic dataset of Dalguer and Mai
(2011) are compared with the largely used empirical prediction equations of a) Cauzzi and Faccioli (2008), CF08, global model dominated by Japanese strong-motion data; b) Faccioli et al. (2010a), FA10, an update of the CF08 model with improved metadata and larger number of observations; c) Boore and Atkinson (2008), BA08, global model developed within the framework of the PEER-NGA research project. All the datasets, although being relatively uniform in the main predictor variables of the associated GMPEs, suffer from some paucity of data at rock sites for distance less than 15 km. In particular, the period dependent dataset of BA08, does not contain any observation at rock sites for \( T = 10 \) s and \( R_{JB} < 20 \) km. CF08 with the updates of FA10 is amongst the global GMPEs selected within the European project SHARE for contributing to the logic tree for shallow active crustal regions (Delavaud et al. 2012). In particular, at long periods between 3 s and 10 s, only two models were selected in SHARE, namely CF08 and Chiou and Youngs (2008). Fig. 4.1 shows the comparisons for unspecified fault mechanisms (Eqn. 3.2). \( R_{RUP} \) is used in the top panels while \( R_{JB} \) is used in the bottom panels, consistently with the distance metric of the GMPEs displayed in each subplot. The hypocentral distance used by Cauzzi and Faccioli (cit.) is converted into the distance from the ruptured fault following Faccioli et al. (2010b). While the median values of the synthetic predictions are in good agreement with the empirical GMPEs for \( M_W 6 \), especially if \( R_{RUP} \) is used, differences are strongly apparent for \( M_W 6.5 \) and \( M_W 7 \), suggesting stronger saturation with distance in the synthetics. For \( T < 5 \) s, the 84th percentile of the distribution of the synthetic prediction shows a remarkable agreement in both shape and amplitude with the median prediction of the empirical GMPEs (with the exception of \( M_W 6 \) for the \( R_{RUP} \)-based model), showing that the extrapolation of the GMPEs in the very near-field is compatible with the largest spectral values obtained from the synthetics.

![Figure 4.1](image-url)

**Figure 4.1.** Comparisons between the synthetic predictions of Eqn. (3.2) and the GMPEs of CF08, FA10 and BA08. The median values of the GMPEs are shown as thin coloured curves, while the median and the 84th percentile (for distance = 5 km only) of the synthetic predictions are shown with black thick solid and dashed curves respectively. Note the large variability of the predicted synthetic spectra in the near-field.

The corner periods of the displacement response spectra (hinge periods where the constant displacement plateau starts to be apparent) from synthetics and real data show a remarkable agreement. As anticipated in Fig. 3.2, the effect of style-of-faulting terms on the median synthetic prediction is significant: when reverse fault scenarios are concerned, as in Fig. 4.2, a better agreement
with CF08 and FA10 is found for $M_W 6$ and the remarkable agreement between the $84^{th}$ percentile of the synthetic predictions and the $50^{th}$ percentile of the reference GMPEs for $T < 5$ s is substantially confirmed.

**Ground type A – $R_{RUP} = 5, 15, 30$ km**

Figure 4.2. Same as Fig. 4.2, but for reverse style-of-faulting.

5. DISCUSSION AND CONCLUSIONS

In this contribution, use has been made of a large database (Dalguer and Mai, 2011) comprising 360 earthquake dynamic rupture models in the $M_W$ magnitude range 5.5 – 7 to explore the feasibility of developing synthetic GMPEs or including synthetic data into current GMPEs to better constrain the empirical predictions in the near-fault region of large earthquakes, with particular regard to the consolidated models of Cauzzi and Faccioli (2008) and Boore and Atkinson (2008). The results obtained in Section 3 clearly show that the parameter space of the synthetics based model can be explored with functional forms similar to those adopted in state-of-the-art GMPEs, with the additional advantage that data from numerical simulation can help in better understanding the significance of some model parameters, like saturation terms, that are often explicitly modelled in current GMPEs without being fully justified based on data distribution and sensitivity analyses. Future investigations may focus on the inclusion of a hanging wall term in the predictions and on the use of orientation independent intensity measures, like GMRotD and GMRotI (Boore et al., 2006). When the simple geometric mean is used, the comparisons presented in Section 4 suggest a stronger saturation with distance for the synthetics, with the $84^{th}$ percentile of the synthetic prediction clearly matching the $50^{th}$ percentile of the GMPEs adopted for comparisons. Recent earthquakes well recorded in the near-field can only partially help in deciding weather more weight should be put on the empirical or on the synthetic GMPEs. For example, median spectral displacements of ~ 10 cm were observed at rock sites and $T > 5$ s for the Christchurch (New Zealand) 2011 event and for the Olufus (Iceland) 2008 event, both with $M_W \sim 6.3$. The median near-field spectra of the L’Aquila (Central Italy) 2009 earthquake did not exceed 5 cm at long periods though, in close agreement with the synthetic predictions obtained herein. We therefore consider premature to derive a set of correction coefficients for the empirical
REFERENCES

Cauzzi, C. and Faccioli, E. (2008). Broadband (0.05 to 20 s) prediction of displacement response spectra based on worldwide digital records. J. Seismol. 12: 453-475.


