16th World Conference on Earthquake Engineering, 16WCEE 2017

Santiago Chile, January 9th to 13th 2017

Paper N° 2476

Registration Code: S-K1464729274

INTENSITY MEASURES AND STATISTICAL APPROACHES FOR THE ENGINEERING VALIDATION OF GROUND MOTION SIMULATIONS

A. Tsioulou(1), C. Galasso(2)

(1) *PhD student, Department of Civil, Environmental & Geomatic Engineering, University College London, UK, alexandra.tsioulou.14@ucl.ac.uk*
(2) *Lecturer, Department of Civil, Environmental & Geomatic Engineering, University College London, UK, c.galasso@ucl.ac.uk*

Abstract

The aim of this article is twofold: first, we propose a list of five ground motion intensity measures (IMs) that act as proxies for the (nonlinear) seismic response of more complex engineered systems, and can therefore be used to validate ground motion simulation methods for engineering applications. The proposed list of IMs include both spectral shape and duration-related proxies, shown to be the optimal IMs in several probabilistic seismic demand models of different structural types, within the framework of Performance-Based Earthquake Engineering. Second, we propose two quantitative approaches for the engineering validation of ground motion simulations, namely statistical hypothesis testing and information theory measures. We then demonstrate the application of these parameters and validation approaches to ground motion simulations computed using a variety of methods, including the Graves and Pitarka hybrid broadband method, the deterministic Composite Source Model (CSM) method and the stochastic white noise (EXSIM) finite-fault model. These types of validation exercises can highlight the similarities and differences between simulated and recorded ground motions for a given simulation method. The similarities should provide confidence in using the simulation method for engineering applications, while the discrepancies, should help in improving the generation of synthetic records.

Keywords: physics-based ground motion simulations; hypothesis testing; relative entropy

1. Introduction

Recent advances in high-performance computing and understanding of complex seismic source features, path effects and site effects, along with the scarcity or total absence of suitable recorded ground motions for specific earthquake scenarios (e.g., large magnitude crustal events recorded at close distance) have led to an increasing interest in physics-based ground motion simulation. Simulated (or "synthetic") ground motion signals (simply ground motions hereinafter) are now considered a valuable supplement to recorded ground motions, fulfilling a variety of engineering needs [1], such as seismic hazard assessment or assessment of seismic demand on structural and geotechnical systems through response history dynamic analysis, within the framework of Performance-Based Earthquake Engineering. Among engineers the general concern is that simulated records may not be equivalent to real records in estimating seismic demand, and hence, in estimating the induced damage and loss to structures. Moreover, synthetic ground motions are not yet widely available in engineering practice, especially in regions where seismogenic faults' locations and characteristics and the regional velocity structure are not well established. On the other hand, in California, the recently released Southern California Earthquake Center (SCEC) Broadband Platform (BBP) [2] provides scientists and engineers with a suite of open-source tools to compute and validate broadband synthetic ground motions by using several physics-based ground motion simulation models. A Technical Activity Group (TAG) focusing on Ground Motion Simulation Validation (GMSV) has been established by SCEC to develop and implement testing/rating methodologies via collaboration between ground motion modelers and engineering users. To this aim, a significant bulk of research has been developed in recent years, including: (1) the comparison of simulations and recordings in terms of waveforms (e.g., by visual inspection), intensity measures (IMs) and structural response for historical events, (2) the comparison in terms of IMs of simulations and predictions from empirical models (e.g., ground motion prediction equations, or GMPEs), and (3) the comparison in terms of structural response of sets of simulations and recordings with similar elastic response spectra, consistently with guidelines for ground motion selection and scaling for building code applications. As a recent example of (1), Galasso et al. [3,4] have investigated whether simulated ground motions are comparable to real records in terms of their nonlinear response in the domain of single degree of freedom (SDoF) systems and multiple degrees of freedom (MDoF) linear and nonlinear building systems. As a recent example of (2), ground motion simulations computed by using five different simulation methods implemented on the SCEC BBP v14.3 are compared with records from 12 earthquake events (western, central and eastern Unites States and Japan), and published GMPEs in the recent studies by Dreger et al. and Goulet et al. [5, 6], with focus on spectral accelerations. As a recent example of (3), Burks et al. [7] have investigated the validation of hybrid broadband simulations for use by structural engineers as input to nonlinear response history analysis following the American Society of Civil Engineers (ASCE) Standard ASCE/SEI 7-10 [8]. The authors consider a set of "appropriate" hybrid broadband simulations (computed by using different simulation methods) and a comparable set of recordings to analyze a building in Berkeley, CA, and compare the predicted structural performance due to the two sets. Finally, Burks and Baker [9] have developed a simulation validation framework combining the empirical models and similar spectra validation approaches (i.e., 2 and 3), proposing a list of parameters for the response of complex structural systems that can be used as proxies for the validation of ground motion simulations for engineering applications. The primary list of parameters includes correlation of spectral acceleration across periods, ratio of maximum to median spectral acceleration across all horizontal orientations, and the ratio of inelastic to elastic displacement, all of which have reliable empirical models against which simulations can be compared. The authors also describe secondary parameters, such as directivity pulse periods and structural collapse capacity, that do not have robust empirical models (so, the historical validation approach needs to be used) but are important for engineering analysis.

This article focuses on engineering validation of ground motion simulations in terms of spectral shape and duration-related IMs for past events. The novelty of this work is that it proposes the use of these advanced IMs as proxies for assessing the similarity of the expected nonlinear structural response and damage potential of simulated and recorded motions for many actual structural types. Nonparametric statistical hypothesis testing and information theory concepts are employed to quantitatively test a specific simulation method as well as to rate different simulation methods, consistently with the objectives of the SCEC GMSV TAG. The use of information theory measures as a tool for validation is another novel contribution of this study. For illustrative purposes, the proposed spectral shape and duration-related IMs, together with the proposed testing/rating methodologies, are derived for different systems (i.e., structural periods) considering three broadband simulation methods: Graves and Pitarka's (2010) hybrid broadband method [10], Composite Source Model (CSM) deterministic method and EXSIM stochastic simulation method. These methods are used to compute simulations for several past Californian earthquakes. In fact, past events provide an important opportunity to test the ability to use ground motion simulation methods to generate synthetic ground motions consistent (i.e., at the same locations) with those observed. Following a validation exercise, as the one presented in this article, end-users can make a decision regarding which model to use for their forward simulations of earthquake scenarios for which no observations exist.

2. Proposed intensity measures

An IM is a scalar ground motion parameter, which is considered to be representative of the earthquake damage potential with respect to a specific engineered system (e.g., a specific structure). Typical engineering applications (e.g., performance-based assessment and design) require the choice of an IM which is suitable to predict the response of the system with the smallest scatter ("efficiency") and providing a significant amount of information, downgrading the effect of other seismological parameters ("sufficiency") to predict the response quantities involved in the performance objectives. Conventional IMs, including the peak ground acceleration (PGA), peak ground velocity (PGV), peak ground displacement (PGD), and spectral (pseudo-) acceleration at the initial fundamental period (for a damping ratio of 5%), $S_a(T_1)$, are the most commonly used IMs. In general, PGA and $S_a(T_1)$ poorly predict the structural response of mid- to high-rise moment resisting frames, although the latter IM sufficiently captures the elastic behavior of first-mode dominated MDoF systems, especially in the case of low to moderate fundamental periods (e.g., [11]). However, the behavior of highly nonlinear structures (sensitive to periods greater than T_1 due to period lengthening) or structures dominated by higher-mode periods (less than T_1) are not very well represented by utilizing $S_a(T_1)$, due to the lack of information on the spectral shape provided by this IM. Therefore, it has become essential implementing advanced IMs that account for the elongated periods and/or consider nonlinear demand dependent structural parameters. Giovenale et al., Kazantzi and Vamvatsikos and Minas et al. [12, 13, 14] amongst others have investigated the adequacy of numerous advanced scalar IMs that take into consideration the aforementioned parameters.

The first advanced scalar IM considered in this article is $S_a{}^c$ (proposed by Cordova et al. [15]), which utilizes spectral-shape information (period elongation), and is expressed as in Eq. (1):

$$S_a^c = S_a(T_1)\left[\frac{S_a(cT_1)}{S_a(T_1)}\right]^{\alpha} \tag{1}$$

where c and α are coefficients conventionally taken to be $c = 2$ and $\alpha = 0.5$ respectively, based on the calibration carried out by the authors in the original study.

Bojórquez and Iervolino [16] also proposed the advanced scalar IM, I_{Np}, which is based on $S_a(T_1)$ and the parameter N_p, defined as in Eq. (2):

$$I_{N_p} = S_a(T_1) N_p^{\alpha} \tag{2}$$

where α parameter is taken as $\alpha = 0.4$ based on the tests conducted by the authors (similar findings are presented in [14]), and N_p is defined as in Eq. (3):

$$N_p = \frac{S_{a,avg}(T_1,...,T_N)}{S_a(T_1)} = \frac{\left[\prod_i^N S_a(T_i)\right]^{1/N}}{S_a(T_1)} \tag{3}$$

T_N corresponds to the maximum period of interest and lies within a range of 2 and 2.5T_1, as suggested by the authors. In this article, the advanced IMs described above (S_a^c and I_{Np}) are computed for four different fundamental periods T_1: 0.5s, 1s, 2s and 4s. For the N_p computation, 3 periods are considered: T_1, 1.5T_1 and 2T_1.

Integral (i.e., duration-related) IMs, as the Arias intensity or significant ground motion duration, are possible IMs, but they are considered to be related more to the cyclic energy dissipation rather than to the peak structural response. In fact, some studies (e.g., [17]) investigated how ground motion duration related parameters affect nonlinear structural response and particularly structural collapse (e.g., [18, 19]). It is widely acknowledged that, generally, spectral ordinates are sufficient (i.e., duration does not add much information) if one is interested in the ductility demand, while duration related measures do play a role only if the hysteretic structural response is to be assessed; i.e., in those cases in which cyclic deterioration and cumulative damage potential of the earthquake are of concern. Chandramohan et al. [19] highlight the need to consider ground motion duration, in addition to intensity and response spectral shape, in regions where significant hazard due to long duration shaking exists, such as locations susceptible to large magnitude, subduction zone earthquakes. Finally, integral IMs are also important for several other engineering applications, for example in geotechnical engineering, such as landslide and liquefaction risk assessment. Therefore, the engineering validation of simulated ground motions in terms of duration-related parameters is also of significant importance.

In particular, Arias intensity, I_A is one of the most commonly used integral IMs and is defined by the integral of ground acceleration as in Eq. (4):

$$I_A = \frac{\pi}{2g} \int_0^{t_E} a^2(t)dt \tag{4}$$

where $a(t)$ is the acceleration time history and t_E is the complete duration of the ground motion.

The term duration can also be used to identify only the portion of a record in which the ground motion amplitude can potentially cause damage to engineering and geotechnical structures. Several definitions are proposed to this aim; the most commonly used one is the significant duration, introduced by Trifunac and Brady [20], defined as the time interval over which the integral of the square of the ground acceleration (Husid plot) is within a given range of its total value. Usually this range is between 5 and 95% (as in this study), denoted as $D_{a5\text{-}95}$, or between 5 and 75%.

Finally, Cosenza and Manfredi [21] introduced the dimensionless I_D - factor defined in Eq. (5) that has proven to be a good proxy for cyclic structural response [22]:

$$I_D = \frac{\int_0^{t_E} a^2(t)dt}{PGA \cdot PGV} \tag{5}$$

where $a(t)$ is the acceleration time history, t_E is the complete duration of the ground motion and PGA and PGV are the peak ground acceleration and velocity respectively.

It is worth noting that the main objective of the BBP validation exercise presented in [5] was to validate elastic spectral response by using the BBP v14.3, and the parameters proposed in this study - as well as those introduced in [9] - are intended as a supplement, not a replacement, to that validation. It is understood that many other metrics would be necessary to fully assess the simulation methods' ability to produce reasonable ground motions as a whole. Also, for each of the proposed parameters, empirical models (i.e., GMPE) exist (e.g., [23] for I_A and I_D) or may be easily derived (e.g., [16] for I_{Np}) combining existing tools and can be used as a baseline comparison against simulations for a very broad range of conditions, including future earthquake scenarios.

3. Proposed validation methods

The validity of simulated ground motions is typically assessed based on methods that are used to quantitatively evaluate the similarity of simulated and recorded time histories in terms of IMs or structural response (e.g., EDPs). One common approach adopted by researchers involves the use of some goodness-of-fit criteria to compare how well the simulations match the ground motion records [24, 25, 5]. This article proposes hypothesis testing as well as validation approaches based on information theory as possible testing/rating methods for simulated ground motions to be used in engineering applications. The following sub-sections provide an overview of the aforementioned validation methods.

3.1 Hypothesis testing

Statistical hypothesis testing is a method of statistical inference used for testing scientific models and assumptions. The assumption to be tested herein is that the two datasets (simulations and recordings) are very similar. In particular, nonparametric hypothesis tests are proposed here to quantitatively assess the statistical significance of differences in terms of proposed IMs for recorded and simulated ground motions. A nonparametric hypothesis test does not require assumptions about the probability distributions of the variables being assessed, making it a more general and robust testing procedure. Moreover, since simulations and recorded ground motions are obtained at the same locations (i.e., seismic stations), it is more appropriate to use a paired test rather than a two sample test, so that each simulation is linked to the corresponding record at the same station. The null hypothesis (i.e., the theory we put forward) is that the median difference between paired IMs, recorded and simulated at a given station, is zero. To address this aim, we propose the Wilcoxon signed rank test [26] which is a nonparametric paired test. The test statistic is calculated by ranking the absolute differences between the paired data (ground motion simulations and recordings on the same site), calculate the sums of the positive and negative ranks by noting the sign of the ranks and calculate the test statistic as reported in Eq. (6):

$$w = \max(t_+, t_-) \tag{6}$$

where w is the test statistic and t_+ and t_- is the sum of the positive and negative ranks respectively.

Some recent validation exercises only focused on median values of the considered parameters (e.g., [6]), not on their aleatory variability (dispersion). As the variability of the ground motions is a very important problem in several engineering applications (e.g., assessing collapse risk of structures), the nonparametric Fligner-Killeen test by median [27] is chosen to test the null hypothesis that the variances of the proposed IMs for the simulated and recorded ground motions for a given event (i.e., intra-event variability) are the same. This is a powerful and robust test against departures from normality and overcomes some limitations of past studies employing a parametric test, the F-test for normally distributed data (e.g., [3, 4]).

3.2 Information theory measures

Information theory concepts can be employed to rate different ground motion simulation methods. In particular, the relative entropy also called the Kullback-Leibler divergence [28] or cross entropy, is proposed here to measure the difference between two probability distributions p and q. In our applications, p typically represents the "true" distribution of a given IM or EDP, i.e., the empirical distribution of the IM or EDP values derived from the recorded ground motions (for example, for a given past event or for a selected set); while q typically represents a model or approximation of p, i.e., the empirical distribution of the IM or EDP values derived from the simulated ground motions (for the given past event or selected set and by using a given simulation method). Specifically, the Kullback–Leibler divergence of q from p, denoted D_{KL}, is a measure of the amount of information lost when q is used to approximate p and is defined as in Eq. (7):

$$D_{KL} = \int_{-\infty}^{+\infty} p(x) \log_2 \left(\frac{p(x)}{q(x)} \right) dx \tag{7}$$

4. Illustrative application

4.1 Description of considered synthetic and recorded ground motion datasets

Ground motion simulations computed by Graves and Pitarka (2010), referred to as G&P (2010) method hereinafter, as implemented on BBP v10.9.0, are evaluated first. Graves and Pitarka [10] developed a hybrid broadband (0-10 Hz) ground motion simulation method which combines a physics-based deterministic approach at low frequency ($f \leq 1$ Hz; i.e., $T \geq 1$s) with a semistochastic approach at high frequency ($f > 1$ Hz; i.e., $T < 1$s). The low- and high-frequency waveforms are computed separately and then combined to produce a single time history through a matching filter. At frequencies below 1 Hz, the method contains a theoretically rigorous representation of fault rupture and wave propagation effects and attempts to reproduce recorded ground motion waveforms and amplitudes. At frequencies above 1 Hz, waveforms are simulated using a stochastic representation of source radiation combined with a simplified theoretical representation of wave propagation and scattering effects. The use of different simulation approaches for the different frequency bands results from the seismological observation that source radiation and wave propagation effects tend to become stochastic at frequencies of about 1 Hz and higher, primarily reflecting the relative lack of knowledge about these phenomena's details at higher frequencies. For both short and long periods, the effect of relatively shallow site conditions, as represented by shear wave velocity in the upper 30 m (V_{s30}), is accounted for using Campbell and Bozorgnia's [29] empirical site amplification model. In particular, the four past earthquakes used for validation of ground motion simulations in the first part of the study are: 1979 M_w 6.5 Imperial Valley, 1989 M_w 6.8 Loma Prieta, 1992 M_w 7.2 Landers, and 1994 M_w 6.7 Northridge. For each simulated event, the model region covers a wide area surrounding the fault, including many strong motion recording sites available in the NGA database: 33 for Imperial Valley, 71 for Loma Prieta, 23 for Landers, and 133 for Northridge. These sites are shown with triangles in Fig.1. This study uses a limited number of sites mentioned in the previous paragraph, considering only those that have real recordings with a usable bandwidth larger than 0.1s-8s. This limitation yields a total of 126 sites for the entire study. These sites are marked with filled triangles in Fig.1. Such large bandwidth for recorded motions provides a justifiable means of covering a good range of nonlinear structural systems where nonlinear response is sensitive to spectral ordinates beyond the maximum considered fundamental period (i.e., 4s).

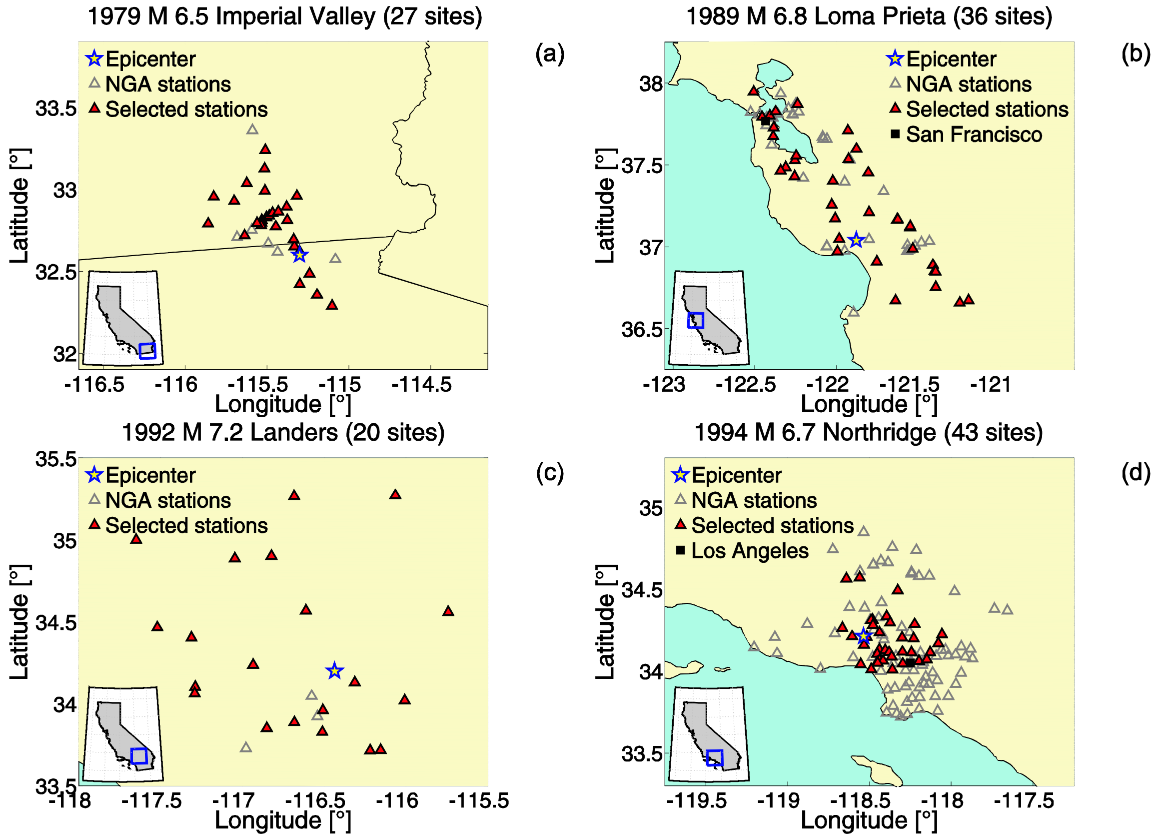

Fig. 1 - Maps of the earthquakes considered. The star is the epicenter and the triangles are the recording stations in the NGA database for which the simulations are available. The filled triangles are the recording stations considered in this study. San Francisco (b) and Los Angeles (d) are also indicated on the map (squares).

The second part of the study evaluates ground motion simulations generated by the SCEC BBP v13.5 and 13.6 using three broadband, finite-source simulation methods: the G&P (2010) hybrid approach described above, the deterministic Composite Source Model approach described by Yu and Zeng [30,31], herein referred to as CSM, and a band-limited stochastic white-noise method called EXSIM developed by Assatourians and Atkinson [32] based on previous work by Motazedian and Atkinson [33] and Boore [34]. The CSM method uses a kinematic source model for rupture on a finite fault. This source is propagated to the station using a flat-layered velocity model, scattering, and attenuation that can be measured from independent seismological observations. The objective is to reproduce the wave propagation entirely within the constraints of the measured velocity and Q structure [35]. As described in [36], EXSIM divides the fault plane in an array of subsources, each of which is treated as point source. The ground motion from each subsource is treated as random Gaussian noise of a specified duration. The duration of motion for each subsource comes from the source duration plus the path duration.

The simulations are compared against four past events and are computed by the G&P (2010) and EXSIM methods as implemented on SCEC BBP v13.6 and the CSM method implemented on SCEC BBP v13.5, as CSM v13.6 was validated against GMPEs and not against recorded events [37]. The four past events considered herein are: 1989 M_w 6.8 Loma Prieta, 1992 M_w 7.2 Landers, 1986 M_w 6.1 North Palm Springs and 1994 M_w 6.7 Northridge. For each simulation methodology and each earthquake event, 50 realizations of the kinematic of the source (e.g., amount of slip, slip velocity, rise time) are simulated, yielding a total of 50 realizations of ground motion simulations per station. The validation is performed on the average results from those 50 realizations. Moreover, as explained in [6], the simulation methods do not focus much on near-surface effects coming from nonlinear site response. In fact, a single generic site profile with a V_{s30} value of 863m/s was used for all the simulations. To make the simulations comparable to the as-recorded site conditions, empirical site effect models should be applied. However, this would add a further layer of epistemic uncertainty to the problem in selecting an appropriate site effect model. To avoid this, we only include recordings from sites with V_{s30} close enough to the V_{s30} used for the BBP simulations (863m/s) to reduce the uncertainties arising from applying site amplification factors. Stations with V_{s30} values greater than 700 m/s are identified to be of “similar” V_{s30} to the reference value used in the simulations, yielding a total number of 25 stations for the four past events.

It should be noted that the ground motion simulations generated by G&P (2010) implemented on BBP v10.9.0 include the effects of shallow site conditions as described above and thus, simulations on stations with different site conditions and shear wave velocities (i.e., not just V_{s30} values greater than 700 m/s) can be directly compared against the recordings at those sites. Therefore, this older version of the G&P (2010) simulation method is particularly useful, as it yields a large number of stations in total and per event, allowing the statistical analysis of the results and in particular the estimation of the intra-event variability, as discussed above. On the other hand, test for homogeneity of variances can’t be applied to the second part of the validation exercise studied herein, as this yields datasets of small number (less than 10 or 6) of simulated ground motions per earthquake event and decreases the power of the hypothesis test.

4.2 Validation results

All ground motions (recorded and simulated) selected for each earthquake event are used as input to compute the selected IMs described above. Only the horizontal components of ground motions (i.e., north-south [NS], and east-west [EW]) are used, while the vertical component is neglected, consistently with other studies. The IM values for the two horizontal components at each station are combined into an “average value” using the geometric mean. For the G&P (2010) BBP v10.9.0 simulation, there is a considerably large number of sites for each earthquake event where simulations are available and directly comparable to records. This is because, as explained above, in this version of the G&P (2010) simulation, the local site effects are explicitly modeled and incorporated in the simulation procedure. This large number of simulations per event allows for the use of hypothesis testing to compare variances for each IM, for the two datasets (recorded and simulated) corresponding to each earthquake event.

To summarize the results of the hypothesis tests and draw conclusions, the *p*-values of the hypothesis tests are reported in Tables 1 and 2 for each IM and earthquake event for the G&P (2010) BBP v10.9.0 simulation method. More specifically, Tables 1 and 2 present the *p*-values for the nonparametric paired Wilcoxon signed rank test and Fligner-Killeen test for the variance respectively, for both the spectral shape and duration-related proxies. For the hypothesis tests yielding a *p*-value less than 0.01 (1%), there is strong evidence to reject the null hypothesis and thus, the differences in the IMs (or their intra-event variability) from simulations and real records are statistically significant. These cases are highlighted with the red color in Tables 1 and 2. For the hypothesis tests yielding a *p*-value greater than 0.01 (1%) and smaller than 0.05 (5%) there is some evidence to reject the null hypothesis; these cases are highlighted with the orange color in Tables 1 and 2. Lastly, for *p*-values greater than 0.05 (5%), there is not sufficient evidence to reject the null hypothesis, meaning that the differences in the IMs (or their intra-event variability) from simulations and real records are not statistically significant. These cases are highlighted with light green color in Tables 1 and 2.

Table 1 – *p*-values for the Wilcoxon signed rank test for spectral shape and duration-related IMs

IM	**Event**	***p*-value for Wilcox signed rank test**			
		T_1=0.5 s	**T_1=1 s**	**T_1=2 s**	**T_1=4 s**
S_a^c	Landers	0.8813	0.1454	0.0124	0.2790
	Loma Prieta	0.1483	0.4796	**0.0049**	0.7415
	Imperial Valley	0.2905	0.0926	0.0517	0.0488
	Northridge	0.8563	0.8943	0.3716	0.1343
I_N	Landers	0.7938	0.4553	0.0304	0.0793

	Loma Prieta	0.3459	0.2714	**0.0022**	0.6374
	Imperial Valley	0.1864	0.1785	0.0643	0.3130
	Northridge	0.6638	0.4837	0.3340	0.3047
I_A	Landers	0.2043			
	Loma Prieta	0.8876			
	Imperial Valley	0.2116			
	Northridge	0.4255			
I_D	Landers	0.1354			
	Loma Prieta	0.5505			
	Imperial Valley	0.4279			
	Northridge	0.1801			
D_{a5-95}	Landers	**0.0040**			
	Loma Prieta	**0.0057**			
	Imperial Valley	0.0926			
	Northridge	0.0490			

Based on these Tables, tests have shown a statistical significance in the bias of the spectral shape-related IMs estimation using simulated records, only at periods around 2 s for Landers and Loma Prieta events. The differences in this period range reveal the large differences in both absolute and relative amplitudes (i.e., the shape) of the elastic response for the Loma Prieta and Landers events as discussed in [3]. For elastic periods of 0.5 and 1 s, the simulated spectral shape-related IMs match the recorded for all the events, whereas for periods of 4 s there is a sparse rejection in terms of S_a^c for Imperial Valley. For the duration-related IMs, the results show a statistical significance of the differences in terms of significant duration D_{a5-95} for most earthquake events. The differences for the rest of the duration-related IMs (I_A, I_D) are not statistically significant, indicating that the IMs estimated from the simulations match the ones estimated from the records. As shown in Table 2, the differences in the intra-event variability estimated using simulated and real ground motions appear not to be statistically significant for the cases of spectral shape and duration-related IMs with only some sparse rejections observed for Imperial Valley events in terms of I_{Np} around 2 s and D_{a5-95} for Northridge event. This reveals that the simulated intra-event variability matches well the observed variability for all the earthquake scenarios considered.

Table 2 – *p*-values for the Fligner-Killeen test for spectral shape and duration-related IMs

IM	**Event**	***p*-value for the Fligner-Killeen test**			
		T_1=0.5 s	**T_1=1 s**	**T_1=2 s**	**T_1=4 s**
S_a^c	Landers	0.5061	0.7525	0.7250	0.9266
	Loma Prieta	0.0996	0.2213	0.8210	0.1904
	Imperial Valley	0.1972	0.0549	0.1017	0.1282
	Northridge	0.2113	0.5476	0.9344	0.6151
I_{N_p}	Landers	0.5948	0.4489	0.8830	0.8741
	Loma Prieta	0.2960	0.0694	0.7047	0.2950
	Imperial Valley	0.2407	0.1338	0.0406	0.2274
	Northridge	0.2055	0.5755	0.8322	0.6886
I_A	Landers	0.8718			
	Loma Prieta	0.1503			

	Imperial Valley	0.2984
	Northridge	0.3767
I_D	Landers	0.4773
	Loma Prieta	0.5500
	Imperial Valley	0.3852
	Northridge	0.9860
$D_{a5\text{-}95}$	Landers	0.4975
	Loma Prieta	0.1388
	Imperial Valley	0.8404
	Northridge	0.0333

Table 3 summarizes the Kullback–Leibler divergence D_{KL} values of the recorded IM distribution from the simulated IM distribution for each of the three simulation methods implemented on BBP v13.5 and 13.6. For each method, the D_{KL} values can be estimated by grouping the simulations from all the earthquake events together and compare them with the records. The mean of the 50 realized values of IMs for the two horizontal components at each station is computed and then combined into an "average" value using the geometric mean. The probability distributions for each IM and simulation method are constructed from the available data through Kernel Density Estimation (KDE) and the D_{KL} values are subsequently estimated using numerical integration to compute the one-dimensional integral in Eq. (7). This allows for the comparison of the performance of the three simulations methods in estimating the probability distributions of spectral shape and duration-related IMs. As discussed above, the estimated D_{KL} value is a measure of the amount of information loss incurred from using the distribution of simulated IMs to approximate the "true" distribution of recorded IMs. Thus, when comparing two or more ground motion simulation methods, the method yielding the smallest D_{KL} value performs best in matching the distribution of recorded IMs; these cases are shown in bold font in Table 3. The results presented in Table 3 reveal that the performance of the simulation methods in estimating spectral shape proxies greatly depends on the advanced IM and period considered. In particular, CSM method performs worse than the other two in estimating S_a^c and I_{Np} across all periods, except for S_a^c distribution for 0.5 s. G&P (2010) method performs best in estimating S_a^c and I_{Np} distributions for 2 s period. EXSIM method gives most accurate predictions for 1 and 4 s for S_a^c and 0.5, 1 and 4 s periods for I_{Np}. Overall, EXSIM method outperforms the other two, having the highest number of best performances for the spectral shape-related IMs considered. On the other hand, there is a single best performing simulation method for all the duration-related IMs examined. Based on the results in Table 3, the G&P (2010) method results in most accurate predictions of the I_A, I_D and $D_{a5\text{-}95}$ distribution.

Table 3 – D_{KL} values for spectral shape and duration-related IMs for each simulation method

IM	**Simulation Method**	**D_{KL} value**			
		T_1=0.5 s	**T_1=1 s**	**T_1=2 s**	**T_1=4 s**
S_a^c	CSM	**0.18**	0.40	0.19	0.42
	EXSIM	0.42	**0.30**	0.35	**0.12**
	G&P(2010)	0.49	0.32	**0.06**	0.45
I_{N_p}	CSM	0.33	0.37	0.19	0.27
	EXSIM	**0.23**	**0.18**	0.17	**0.24**
	G&P(2010)	0.31	0.22	**0.06**	0.33
I_A	CSM	0.33			

	EXSIM	0.67
	G&P(2010)	**0.19**
I_D	CSM	0.56
	EXSIM	0.54
	G&P(2010)	**0.40**
$D_{a5\text{-}95}$	CSM	1.42
	EXSIM	3.42
	G&P(2010)	**0.14**

5. Conclusions

The design of new structures or assessment of existing ones may be complicated by the inherent rareness or total absence of suitable recorded ground motions for the earthquake scenarios that dominate the seismic hazard at a given site. Therefore, broadband synthetic records may be an attractive option as input to nonlinear dynamic analysis, if an accurate and transparent engineering validation is carried out. To this aim, the main focus of this article was on the design of such a validation exercise by 1) proposing a list of five ground motion intensity measures that act as proxies for the (nonlinear) seismic response of actual buildings and geotechnical systems; and 2) proposing two robust quantitative approaches for testing/rating simulation methods, namely statistical hypothesis testing and information theory measures. The application of the proposed parameters and evaluation criteria was demonstrated using two groups of ground motion simulations: 1) those computed by using Graves and Pitarka's (2010) simulation method implemented on BBP v10.9.0 for four past earthquakes: 1979 M_w 6.5 Imperial Valley, 1989 M_w 6.8 Loma Prieta, 1992 M_w 7.2 Landers, and 1994 M_w 6.7 Northridge; and 2) those computed by using Graves and Pitarka's (2010), CSM and EXSIM simulation methods implemented on v13.5 and 13.6 of BBP for four past earthquakes: 1989 M_w 6.8 Loma Prieta, 1992 M_w 7.2 Landers, 1986 M_w 6.1 North Palm Springs and 1994 M_w 6.7 Northridge. This study is part of a larger, longer-term, and broader ongoing plan for the validation of simulated ground motions for engineering applications. The first validation exercise yields results that are consistent with past studies (e.g., [3]) indicating that the proposed advanced IMs are good proxies for validation of ground motion simulations in terms of peak inelastic and cyclic structural response. Results from the second validation exercise can be used to rank the performance of the considered ground motion simulation methodologies. Such validation approaches are necessary when ranking of different simulation methods is desired and can be applied as supplement to statistical hypothesis testing.

6. Acknowledgements

We thank Christine Goulet for help in accessing the simulated ground motions used in this study and the Pacific Earthquake Engineering Research (PEER) Center for providing the recorded data.

7. References

[1] Baker JW, Luco N, Abrahamson NA, Graves RW, Maechling PJ, Olsen KB (2014): Engineering uses of physics-based ground motion simulations. *10th U.S. National Conference on Earthquake Engineering*, Anchorage, Alaska.

[2] Maechling PJ, Silva F, Callaghan, S, Jordan, TH (2015): SCEC broadband platform: System architecture and software implementation. *Seismological Research Letters*, **86** (1), 27-38.

[3] Galasso C, Zareian F, Iervolino, I, Graves RW (2012): Validation of ground-motion simulations for historical events using SDoF systems. *Bulletin of the Seismological Society of America*, **102** (6), 2727-2740.

[4] Galasso C, Zhong P, Zareian F, Iervolino I, Graves RW (2013): Validation of ground-motion simulations for historical events using MDoF systems. *Earthquake Engineering and Structural Dynamics*, **42** (9), 1395-1412.

[5] Dreger DS, Beroza GC., Day SM, Goulet CA, Jordan TH, Spudich PA, Stewart JP (2015): Validation of the SCEC broadband platform v14.3 simulation methods using pseudospectral acceleration data. *Seismological Research Letters*, **8** (1), 39-47.

[6] Goulet CA, Abrahamson NA, Somerville PG, Wooddell KE (2015): The SCEC Broadband Platform validation exercise: Methodology for code validation in the context of seismic hazard analyses. *Seismological Research Letters* **86**(1), 17-26.

[7] Burks LS., Zimmerman RB., Baker JW (2016): Evaluation of hybrid broadband ground motion simulations for response history analysis and design. *Earthquake Spectra,* **31** (3), 1691-1710.

[8] American Society of Civil Engineers (ASCE) 2010: Minimum design loads for buildings and other structures, *ASCE/SEI 7-10,* Reston, VA, USA.

[9] Burks LS, Baker JW (2014): Validation of ground motion simulations through simple proxies for the response of engineered systems. *Bulletin of the Seismological Society of America,* **104** (4), 1930-1946.

[10] Graves RW, Pitarka A (2010): Broadband ground-motion simulation using a hybrid approach. *Bulletin of the Seismological Society of America,* **100** (5A), 2095-2123.

[11] Shome N, Cornell CA, Bazzurro P, Carballo JE (1998): Earthquakes, records, and nonlinear responses. *Earthquake Spectra,* **14** (3), 469–500.

[12] Giovenale P, Cornell AC, Esteva L (2004): Comparing the adequacy of alternative ground motion intensity measures for the estimation of structural responses. *Earthquake Engineering and Structural Dynamics,* **33** (8), 951–979.

[13] Kazantzi AK, Vamvatsikos D (2015): Intensity measure selection for vulnerability studies of building classes. *Earthquake Engineering and Structural Dynamics*, **44** (15), 2677-2694.

[14] Minas S, Galasso C, Rossetto T (2015): Spectral-shape proxies and simplified fragility analysis of mid-rise reinforced concrete buildings. *12th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP)*, Vancouver, Canada.

[15] Cordova PP, Deierlein GG, Mehanny SS, Cornell CA (2000): Development of a two-parameter seismic intensity measure and probabilistic assessment procedure. *2nd U.S.–Japan Workshop on Performance-Based Earthquake Engineering for Reinforced Concrete Building Structures*, Sapporo, Japan.

[16] Bojórquez E, Iervolino I (2011): Spectral shape proxies and nonlinear structural response. *Soil Dynamics and Earthquake Engineering,* **31** (7), 996-1008.

[17] Iervolino I, Manfredi G, Cosenza E (2006): Ground-motion duration effects on nonlinear seismic response. *Earthquake Engineering and Structural Dynamics,* **35** (1), 21-38.

[18] Raghunandan M, Liel AB (2013): Effect of ground motion duration on earthquake-induced structural collapse. *Structural Safety*, **(41)**, 119-133.

[19] Chandramohan R, Baker JW, Deierlein GG (2016): Quantifying the influence of ground motion duration on structural collapse capacity using spectrally equivalent records, *Earthquake Spectra,* **32** (2), 927-950.

[20] Trifunac MD, Brady AG (1975): A study on the duration of strong earthquake ground motion. *Bulletin of the Seismological Society of America,* **65** (3), 581– 626.

[21] Cosenza E, Manfredi G (1997): The improvement of the seismic-resistant design for existing and new structures using damage criteria. *Seismic Design Methodologies for the Next Generation of Codes,* Fajfar P., Krawinkler H. (eds), Balkema, 119 –130.

[22] Manfredi G (2001): Evaluation of seismic energy demand. *Earthquake Engineering and Structural Dynamics,* **30** (4), 485-499.

[23] Iervolino I, Giorgio M, Galasso C, Manfredi G (2010): Conditional hazard maps for secondary intensity measures. *Bulletin of the Seismological Society of America,* **100** (6), 3312–3319.

[24] Anderson JG (2004): Quantitative measure of the goodness-of-fit of synthetic seismogram. *13th World Conference on Earthquake Engineering*, Vancouver, Canada.

[25] Olsen KB, Mayhew JE (2010): Goodness-of-fit criteria for broadband synthetic seismograms, with application to the 2008 M5.4 Chino Hills, California, earthquake. *Seismological Research Letters,* **81** (5), 715-723.

[26] Wilcoxon F (1945): Individual comparisons by ranking methods. *Biometrics Bulletin,* **1** (6), 80-83.

[27] Fligner M, Killeen T (1976): Distribution-free two-sample tests for scale. *Journal of the American Statistical Association*, **71** (353), 210-213.

[28] Kullback S (1959): *Information and statistics*. Wiley, NY.

[29] Campbell KW, Bozorgnia Y (2008): NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. *Earthquake Spectra*, **24** (1), 139–171.

[30] Yu G (1994): Some aspects of earthquake seismology: Slip partitioning along major convergent plate boundaries; composite source model for estimation of strong motion; and nonlinear soil response modeling. Ph.D. Thesis, University of Nevada, Reno.

[31] Zeng Y, Anderson JG, Yu G (1994): A composite source model for computing realistic synthetic strong ground motions. *Geophysical Research Letters*, **21** (8), 725–728.

[32] Assatourians K, Atkinson GM (2012): *EXSIM12: A Stochastic Finite-Fault Computer Program in FORTRAN*, software, http://www.seismotoolbox.ca.

[33] Motazedian D, Atkinson GM (2005): Stochastic finite-fault modeling based on a dynamic corner frequency. *Bulletin of the Seismological Society of America*, **95** (3), 995–1010.

[34] Boore DM (2009): Comparing stochastic point-source and finite-source ground-motion simulations: SMSIM and EXSIM. *Bulletin of the Seismological Society of America*, **99**(6), 3202–3216.

[35] Anderson JG (2015): The composite source model for broadband simulations of strong ground motions. *Seismological Research Letters*, **86** (1).

[36] Atkinson GM, Assatourians K (2015): Implementation and validation of EXSIM (a stochastic finite-fault ground-motion simulation algorithm on the SCEC Broadband Platform. *Seismological Research Letters*, **86** (1), 48-60.

[37] Goulet CA (2014). Personal communication.