Received: 23 October 2017 | Accepted: 24 November 2017
DOI: 10.1002/eqe.3015

SHORT COMMUNICATION

WILEY

Information theory measures for the engineering validation of ground-motion simulations

Alexandra Tsioulou | Carmine Galasso

Department of Civil, Environmental and Geomatic Engineering and Institute for Risk and Disaster Reduction, University College London, London, UK

Correspondence
Carmine Galasso, Associate Professor, Department of Civil, Environmental and Geomatic Engineering and Institute for Risk and Disaster Reduction, University College London, Gower St., WC1E 6BT, London, UK
Email: c.galasso@ucl.ac.uk

Summary
This short communication introduces a quantitative approach for the engineering validation of ground-motion simulations based on information theory concepts and statistical hypothesis testing. Specifically, we use the Kullback-Leibler divergence to measure the similarity of the probability distributions of recorded and simulated ground-motion intensity measures (IMs). We demonstrate the application of the proposed validation approach to ground-motion simulations computed by using a variety of methods, including Graves and Pitarka hybrid broadband, the deterministic composite source model, and a stochastic white noise finite-fault model. Ground-motion IMs, acting as proxies for the (nonlinear) seismic response of more complex engineered systems, are considered herein to validate the considered ground-motion simulation methods. The list of considered IMs includes 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. The proposed validation exercise (1) can highlight the similarities and differences between simulated and recorded ground motions for a given simulation method and/or (2) allow the ranking of the performance of alternative simulation methods. The similarities between records and simulations should provide confidence in using the simulation method for engineering applications, while the discrepancies should help in improving the tested method for the generation of synthetic records.

KEYWORDS
ground-motion simulations, relative entropy, SCEC Broadband Platform (BBP), spectral-shape and duration-related IMs

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-motion signals (simply ground motions hereinafter) for specific earthquake scenarios (eg, large-magnitude crustal events recorded at close distance), have led to an increasing interest in ground-motion simulation. Simulated (or "synthetic") ground motions 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 ground motions 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 the engineering practice, especially in regions where seismogenic faults 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 methods. 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. A similar effort is also being made in Italy, through a recently released web-repository (SYNTHEtic SeISmograms database) containing synthetic waveforms for Italian scenario earthquakes coming from different simulation techniques[3].

The study presented here proposes the use of information theory concepts and statistical hypothesis testing to quantitatively test a specific simulation method as well as to rate different simulation methods, consistently with the objectives of the SCEC GMSV TAG. We focus on the engineering validation of ground-motion simulations in terms of spectral-shape and duration-related intensity measures (IMs). These metrics are common proxies for assessing the similarity of the expected nonlinear structural response and damage potential of simulated and recorded ground motions for many actual structural types. For illustrative purposes, the proposed testing/rating methodology is applied to the considered spectral-shape and duration-related IMs, obtained for different systems (ie, structural periods) considering three broadband simulation methods: Graves and Pitarka's hybrid broadband method, the composite source model (CSM) deterministic method, and the 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 (ie, at the same locations) with those observed. Following a validation exercise, as the one presented in this article, end-users can decide regarding which model to use for their forward simulations of earthquake scenarios for which no observations exist. The confidence in using simulation methods beyond the tested limits must also be assessed considering the science behind each method[4].

The next section briefly reviews some recent approaches and studies aiming at the engineering validation of ground-motion simulation. This is followed by an introduction to the proposed validation approach. An illustrative implementation of the proposed approach is then presented and results of the application are finally discussed.

2 | ENGINEERING VALIDATION OF GROUND-MOTION SIMULATIONS

A significant bulk of research has been developed in recent years to validate ground-motion simulation methods for engineering applications, including (1) the comparison of simulations and recordings in terms of waveforms (eg, by visual inspection), IMs, and structural response for historical events; (2) the comparison in terms of IMs of simulations and predictions from empirical models (eg, ground-motion prediction equations, or GMPEs), for both historical events and future scenarios; 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[5,6] have investigated whether simulated ground motions are comparable to real records in their nonlinear response in the domain of single degree of freedom systems and multiple degrees of freedom linear and nonlinear building systems. The authors consider four historical earthquakes modeled by using the hybrid broadband ground-motion simulation method by Graves and Pitarka.[7] The validation exercise using various engineering demand parameters (EDPs) and formal statistical hypothesis testing indicates that, in most cases, the differences found in seismic demands produced by real and synthetic records are not significant, increasing the trust in the use of simulated motions for engineering applications. Rezaeian et al[8] propose a validation framework at the waveform level and considering three time-dependent validation metrics capturing the nonstationary features of intensity and frequency contents of earthquake ground motions. The proposed validation methodology is demonstrated by using example of recorded and simulated ground motions from the Northridge event computed with the method by Graves and Pitarka.[7]

As a recent example of (2), ground-motion simulations computed by 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 Goulet et al[4] and Dreger et al.[9] The validation is performed in these studies with a focus on spectral accelerations. Four generic strike-slip and reverse scenarios for which GMPEs are considered to be well constrained by data are considered to compare spectral accelerations produced by simulation with predictions from the selected GMPEs. The results from this study indicate that the simulation methods provide reasonable estimates of spectral acceleration; however, it is suggested that additional research work is necessary to validate ground-motion simulations for other applications by using different proxies and methods. As a recent example of (3), Burks et al[10] 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.[11] 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 using the two sets. Results show that the structural behavior resulting from recordings and simulations is similar, and most discrepancies are explained by differences in directional properties such as orientation of the maximum spectral response. These results suggest that when simulations meet the criteria outlined for recordings in ASCE/SEI 7-10 and properties such as directionality are realistically represented, simulations provide useful results for structural analysis and design. Finally, Burks and Baker[12] have developed a simulation validation framework combining the empirical models and similar spectra validation approaches (ie, 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. The authors demonstrate the application of the proposed framework to example simulations computed by using a variety of simulation methods. Results show that each simulation method matches empirical models for some parameters and not others, indicating that all relevant parameters need to be carefully validated.

3 | PROPOSED VALIDATION APPROACH

As discussed, the validity of simulated ground motions is typically assessed based on criteria that are used to quantitatively evaluate the similarity of simulated and recorded time series in terms of IMs or structural response (ie, 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.[9,13,14] We propose the use of a novel validation approach based on information theory as a possible testing/rating methodology for simulated ground motions to be used in engineering applications. Information theory concepts can be used to test the similarity of two datasets, which herein refers to the considered IMs (or EDPs) for simulated and recorded ground motions. Specifically, the relative entropy, also called the Kullback-Leibler divergence[15] or cross entropy, is proposed here to measure the difference between two probability distributions p and q. In our applications, $p(\mathrm{IM})$ represents the "true" distribution of a given IM (or EDP|IM), ie, the empirical distribution of the IM (or EDP|IM) values derived from the recorded ground motions (for example, for a given past event or for a selected hazard-compatible ground-motion set), while $q(\mathrm{IM})$ represents a model or approximation of $p(\mathrm{IM})$, ie, the empirical distribution of the IM (or EDP|IM) values derived from the simulated ground motions (for the given past event or selected set and by using a given simulation method). The Kullback-Leibler divergence of $q(\mathrm{IM})$ from $p(\mathrm{IM})$, denoted D_{KL}, is a measure of the amount of information lost when $q(\mathrm{IM})$ is used to approximate $p(\mathrm{IM})$ and is defined as

$$D_{\mathrm{KL}} = \int_{-\infty}^{+\infty} p(\mathrm{IM}) \log_2 \left(\frac{p(\mathrm{IM})}{q(\mathrm{IM})} \right) d\mathrm{IM} \tag{1}$$

If the logarithm is calculated in base 2, D_{KL} is expressed in bits of information. D_{KL} has been used in earthquake engineering applications to compare the relative sufficiency of alternative IMs in predicting structural response.[16]

In the context of ground-motion simulation validation, the empirical distribution of the observed IMs estimated from the records, $p(\mathrm{IM})$, and the empirical distribution of the IMs calculated from the simulated ground motions, $q(\mathrm{IM})$, are constructed by using kernel density estimation (KDE) based on n available IM samples as

$$p(\mathrm{IM})\ \text{or}\ q(\mathrm{IM}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}K\left(\frac{\mathrm{IM}-\mathrm{IM}_i}{h}\right) \tag{2}$$

where $K(.)$ is the chosen kernel and h is the kernel bandwidth. In this study, the Epanechnikov kernel is used, which is given by the following expression:

$$\begin{aligned} K(t) &= \frac{3}{4}\left(1-t^2\right)\ \text{if-}1\leq t\leq 1 \\ &= 0\ \text{else} \end{aligned} \tag{3}$$

with bandwidth h chosen to minimize the asymptotic mean-integrated squared error between the KDE and the target distribution to be approximated.[17] The entropy in Equation 1 can be then approximated by using the KDE estimates of p(IM) and q(IM) in Equation 2, with the 1-dimensional integral calculated through numerical integration; for example, by using the trapezoidal rule.

Given that the estimated D_{KL} values are not standardized nor do they have an upper bound, it may be challenging assessing how extreme the calculated D_{KL} value is and drawing conclusions regarding the similarity of the two datasets. To overcome this, a procedure using the bootstrapping technique to construct an empirical distribution of D_{KL} is proposed and statistical hypothesis testing is used to assess the similarity of the two datasets from the observed D_{KL} value. This procedure is summarized below, where samples of IMs estimated from real records are called X and samples of IMs from simulations are referred to as Y for simplicity.

In the first step of the proposed procedure, we compute the Kullback-Leibler divergence D_{KL} between X and Y, referred to as $D_{\mathrm{KL,obs}}$. In statistical hypothesis testing, the P-value for $D_{\mathrm{KL,obs}}$ is the probability that when the null hypothesis is true, D_{KL} would be the same as or more extreme than the actual observed value. In this case, the null hypothesis is that X and Y have the same probability distribution. If this is true, then X and Y can be merged into a single sample and be treated as being one larger draw from the same distribution. This is the second step of the proposed procedure. The bootstrapping technique is then used in the third step of the proposed procedure to get the empirical distribution of D_{KL} for each considered IM. To achieve this, two new vectors, X_{boot} and Y_{boot}, that have the same length as X and Y are drawn, by sampling observations at random from the combined X and Y data with replacement, so that observations from the original X sample may end up in the bootstrapped Y_{boot} sample and vice versa. For each set of new vectors, X_{boot} and Y_{boot}, the Kullback-Leibler divergence, $D_{\mathrm{KL,boot}}$, can be calculated. The third step is performed many times, 1000 in this specific exercise, yielding 1000 samples of $D_{\mathrm{KL,boot}}$. Finally, in the fourth step of the proposed procedure, the P-value for the observed $D_{\mathrm{KL,obs}}$ is computed by finding the proportion of the 1000 $D_{\mathrm{KL,boot}}$ samples that are more extreme (ie, larger) than the $D_{\mathrm{KL,obs}}$ value computed by using the original X and Y vectors. The obtained P-value represents the level of statistical significance in assuming that X and Y have the same probability distribution. Reasonable pass/fail thresholds can be applied to the obtained P-value results, for instance 95%, as in traditional hypothesis testing and in the illustrative application presented below.

The proposed validation approach distinguishes itself from the past studies and other proposed goodness-of-fit criteria by assessing the overall similarity of the probability distributions of the studied IMs for recorded and simulated GMs. Thus, it does not just provide a paired comparison (ie, at the same recording locations, for historical events) between the recorded and simulated IM datasets in mean and standard deviation of their distributions. This represents a useful tool for the engineering validation of simulated ground motions in terms of the nonlinear structural demands or expected loss for a portfolio of structures (or infrastructure) where an overall as opposed to a paired comparison of the records and simulations is of interest, for example, for catastrophe risk modeling purposes.[18] The proposed approach can also be used to measure the similarity of the distributions of seismic response to sets of simulations and recordings matching a target (elastic) response spectrum mean and variance, consistently with the current practice in ground-motion selection and scaling for building code applications.[19]

4 | ILLUSTRATIVE APPLICATION

The illustrative implementation of the proposed validation approach considers ground-motion simulations generated by the SCEC BBP v13.5 and 13.6 by using three broadband, finite-source simulation methods: the hybrid approach by Graves and Pitarka,[7] referred to as G&P (2010); the deterministic CSM approach[20], herein referred to as CSM; and a band-limited stochastic white-noise method called EXSIM[21] based on previous work by Boore.[22] G&P (2010), widely used in past validation studies, is a hybrid broadband (0-10 Hz) ground-motion simulation method that combines a physics-based deterministic approach at low frequency ($f \leq 1$ Hz) with a semistochastic approach at high frequency ($f > 1$ Hz). The low- and high-frequency waveforms are computed separately and then combined to produce a single time history through a matching filter. 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. The CSM method uses a kinematic source model for rupture on a finite fault. This source is propagated to the station by 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.[23] As described in Atkinson and Assatourians,[24] 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 used here 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 the CSM method on BBP v13.6 is only available for validation against GMPEs and not against recorded events (personal communication with C.A. Goulet, 2016). The four historical 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 method and each earthquake event, 50 different simulations were obtained based on the same number of realizations of different kinematic source models (eg, amount of slip, slip velocity, and rise time), 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 Goulet et al,[4] 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 863 m/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 increasing the epistemic uncertainty of the problem. Therefore, to reduce the uncertainties arising from applying site amplification factors, this article only includes recordings from sites with V_{s30} close enough to the V_{s30} used for the BBP simulations (863 m/s). 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 (SCEC GMSV TAG). This leads to datasets of relatively small size (less than 10 ground motions) for each considered earthquake event. The estimation of empirical probability distributions from such small datasets can result in unreliable values of D_{KL}. In this case, because of the purely illustrative nature of the application presented here, all the events for each simulation method were combined, focusing on assessing the overall performance of a given simulation method rather than the specific performance for a given earthquake event.

4.1 | Considered intensity measures

An IM is a single ground-motion parameter (scalar IM) or set of ground-motion parameters (vector IM), which are representative of the earthquake damage potential with respect to a specific class of engineered systems. Typical engineering applications (eg, 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. In addition, many researchers have investigated other IM selection criteria, related, for example, to "hazard computability," "proficiency," and "practicality."[25] Conventional IMs, including the peak ground acceleration (PGA), peak ground velocity (PGV), peak ground displacement, and (pseudo-) spectral 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-rise to high-rise moment resisting frames, although the latter IM sufficiently captures the elastic behavior of first-mode dominated multiple degrees of freedom systems, especially in the case of low to moderate fundamental periods.[26] 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 $S_a(T_1)$, because of 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. Kazantzi and Vamvatsikos [27] and Kohrangi et al[28] among several others have investigated the adequacy of numerous advanced scalar IMs that take into consideration the aforementioned parameters.

For the illustrative application presented here, we then use the advanced scalar IM proposed by Bojórquez and Iervolino.[29] This IM, denoted as I_{Np}, is based on $S_a(T_1)$ and the parameter N_p, and is defined as

$$I_{N_p} = S_a(T_1) N_p^{\alpha} \tag{4}$$

where the parameter α is taken as $\alpha = 0.4$ based on the tests conducted by the authors and N_p is defined as

$$N_p = \frac{S_{a,\text{avg}}(T_1, ..., T_N)}{S_a(T_1)} = \frac{\left[\prod_i^N S_a(T_i)\right]^{1/N}}{S_a(T_1)} \tag{5}$$

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 study, I_{Np} is computed for four different fundamental periods T_1: 0.5, 1, 2, and 4 s. For the N_p computation, 3 periods are considered: T_1, $1.5T_1$, and $2T_1$.

Integral (ie, duration-related) IMs, such as the Arias intensity or significant ground-motion duration, are commonly used, but they are considered to be related more to the cyclic energy dissipation rather than to the peak structural response. In fact, some studies[30] investigated how ground-motion duration-related parameters affect nonlinear structural response and particularly structural collapse.[31,32] It is widely acknowledged that, generally, spectral ordinates are *sufficient* (ie, 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; ie, in those cases in which cyclic deterioration and cumulative damage potential of the earthquake are of concern. Chandramohan et al[32] 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.

The term duration is typically 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,[33] defined as the time interval over which the integral of the square of the ground acceleration (Husid plot[34]) is within a given range of its total value. Usually, this range is between 5% and 95% (as in this study), denoted as $D_{s5\text{-}95}$, or between 5% and 75%.Finally, Cosenza and Manfredi[35] introduced the dimensionless I_D—factor defined as

$$I_D = \frac{\int_0^{t_E} a^2(t)dt}{\text{PGA}\cdot\text{PGV}} \tag{6}$$

which has proven to be a good proxy for cyclic structural response.[36] Here, $a(t)$ is the acceleration time history, and t_E is the complete duration of the ground motion (length of the record).

It is worth noting that the main objective of the BBP validation exercise presented in Dreger et al[9] was to validate elastic spectral response by using the BBP v14.3. The parameters proposed in our study—as well as those introduced in Burks and Baker[12]—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. An important property of the proposed validation parameters is that they are hazard computable, ie, empirical models (ie, GMPEs) exist (eg, Iervolino et al.[37] for I_A and I_D) or may be easily derived (eg, Bojorquez and Iervolino[29] for I_{N_p}) combining existing tools and can be used as a baseline comparison against simulations for a very broad range of conditions, including future earthquake scenarios.

4.2 | Validation results

All ground motions (recorded and simulated) selected for each simulation method are used as input to compute the selected IMs described above. Only the horizontal components of ground motions (ie, north-south and east-west) are used, while the vertical component is neglected, consistently with other studies.

Table 1 summarizes the D_{KL} values for all considered IM distributions for each of the three simulation methods implemented on BBP v13.5 and 13.6, as discussed above. The mean of the 50 values of IMs obtained from the same number of realizations for the two horizontal components at each station is computed and then combined into an "average" value by using the geometric mean. In this case, as explained above, the number of data per event is limited, and thus, the D_{KL} values are estimated for each simulation method by grouping the simulations from all the earthquake events together and compare them with the records. This allows the comparison of the performance of the three simulation

TABLE 1 D_{KL} values for spectral-shape and duration-related IMs for each simulation method

<table>
<tr><th rowspan="2">IM</th><th rowspan="2">Simulation Method</th><th colspan="4">$D_{KL,obs}$ Value</th></tr>
<tr><th>$T_1 = 0.5$ s</th><th>$T_1 = 1$ s</th><th>$T_1 = 2$ s</th><th>$T_1 = 4$ s</th></tr>
<tr><td>I_{N_p}</td><td>CSM</td><td>0.33</td><td>0.37</td><td>0.19</td><td>0.27</td></tr>
<tr><td></td><td>EXSIM</td><td>0.23</td><td>0.18</td><td>0.17</td><td>0.24</td></tr>
<tr><td></td><td>G&P (2010)</td><td>0.31</td><td>0.22</td><td>0.06</td><td>0.33</td></tr>
<tr><td>I_D</td><td>CSM</td><td colspan="4">0.56</td></tr>
<tr><td></td><td>EXSIM</td><td colspan="4">0.54</td></tr>
<tr><td></td><td>G&P (2010)</td><td colspan="4">0.40</td></tr>
<tr><td>$D_{s5\text{-}95}$</td><td>CSM</td><td colspan="4">1.42</td></tr>
<tr><td></td><td>EXSIM</td><td colspan="4">3.42</td></tr>
<tr><td></td><td>G&P (2010)</td><td colspan="4">0.14</td></tr>
</table>

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 1.

As explained in a previous section, statistical hypothesis testing can be performed by using the bootstrapping technique to assess how large the observed D_{KL} values are in each case and draw conclusions regarding the similarity of the two datasets, recorded and simulated, for a given simulation method. For the hypothesis tests yielding a P-value less than 0.05 (5%), there is strong evidence to reject the null hypothesis, and thus, the differences in the IM probability distributions from simulations and real records can be considered statistically significant. This means that the observed D_{KL} value lies above the 95th percentile of the empirical cumulative distribution function for D_{KL}. These cases are underlined in Table 1. For P-values greater than 0.05 (5%), there is not sufficient evidence to reject the null hypothesis, meaning that the differences in the IM probability distributions from simulations and real records are not statistically significant. In this case, the observed D_{KL} values fall below the 95th percentile of the cumulative distribution function for D_{KL}.

The results in Table 1 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 methods in estimating I_{Np} across all periods. G&P (2010) method performs best in estimating I_{Np} only for 2 s period. EXSIM method gives the most accurate predictions for 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 1, the G&P (2010) method results in the most accurate predictions of the I_D and $D_{s5\text{-}95}$ distribution.

With respect to the results of the hypothesis testing, all the observed D_{KL} values are within the region of non rejection established through bootstrapping, except for the D_{KL} values for $D_{s5\text{-}95}$ calculated from CSM and EXSIM methods. To shed further light on this result, histograms of the $D_{s5\text{-}95}$ samples from recorded and simulated ground motions are plotted in Figure 1 for the CSM and EXSIM methods. The white bars correspond to the simulated IMs, whereas the grey bars refer to the IMs from recorded time histories. A histogram plot of I_{Np} samples at 2s period for G&P (2010) that corresponds to a case of no rejection is presented in Figure 2 for comparison. It is evident that large differences exist between the resulting histograms (and derived probability density functions from KDE) for the simulated and recorded $D_{s5\text{-}95}$ values for the CSM and EXSIM methods as shown in Figure 1. By contrast, the histograms of the simulated and recorded I_{Np} samples at 2 s period for G&P (2010) are very similar, making it the best performing method for this specific validation metric, as shown in Figure 2.

In addition to the visual comparison in Figures 1 and 2, the proposed validation approach was compared with the more commonly used hypothesis testing method. Standard t-tests for the equality of means of the studied IMs were performed,[5,6] and the results were consistent with the results of the proposed approach based on D_{KL} values. However, since the latter approach evaluates the match between the full distributions (ie, probability density functions) of the two data sets, it is also necessary to perform hypothesis tests on higher moments of the distribution and not just the mean (first moment) to compare the results. Hence, F-tests for the equality of variances of the two datasets were performed in addition to the t-tests. The results showed that there were several rejections for the F-test that did not seem to be justified by

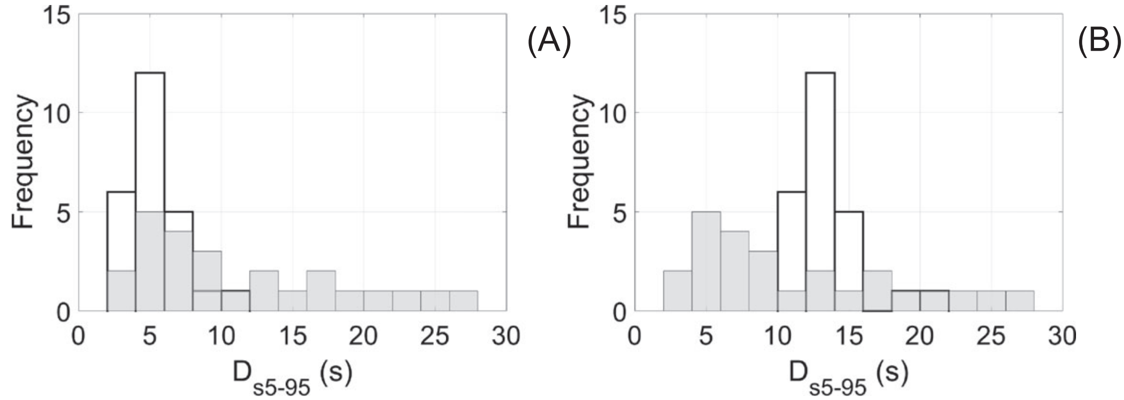

FIGURE 1 Histograms of the D_{s5-95} samples from recorded (grey bars) and simulated (white bars) ground motions for (A) the CSM and (B) the EXSIM methods

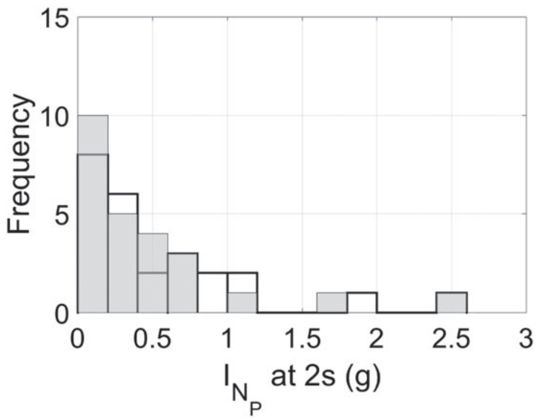

FIGURE 2 Histograms of the I_{Np} samples at 2 s period from recorded (grey bars) and simulated (white bars) ground motions for the G&P (2010) method

the empirical IM distributions (histogram plots of the IMs for these rejection cases were created to visually assess the equality of variances, similar to Figures 1 and 2). On the other hand, the method proposed here seems to accurately detect differences in the full distribution of the IMs as shown in Figures 1 and 2.

5 | CONCLUSIONS

The design of new structures or the 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 for the considered simulation method is carried out. To this aim, the focus of this note was on the design of such a validation exercise by proposing a novel quantitative approach for testing/rating ground-motion simulation methods, based on information theory measures coupled with statistical hypothesis testing. The proposed approach assesses the overall similarity of the probability distributions of the recorded and simulated IMs and uses the relative entropy to quantify their distance. Statistical hypothesis testing relying on the bootstrapping technique is then used to test the significance of the estimated distance. Ultimately, the approach can be used to rank the performance of different ground-motion simulation methods and it is part of a larger, longer-term, and ongoing plan for the validation of simulated ground motions for engineering applications.

The application of the proposed evaluation criteria was demonstrated by using a group of ground-motion simulations computed by Graves and Pitarka's (2010), CSM, and EXSIM simulation methods implemented on v13.5 and 13.6 of the SCEC 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. The illustrative application considers three ground-motion IMs: one spectral-shape and two duration-related that have been shown to be optimal proxies for the (nonlinear) seismic response of actual buildings and geotechnical systems. The proposed validation metrics are hazard computable, and their empirical models can be used as baseline for comparison for future earthquake scenarios. The list of IMs considered in this study is not exhaustive and can be used to supplement other validation metrics encountered in the literature. Finally, for the specific simulated ground-motion data considered in the illustrative example, the EXSIM and Graves and Pitarka's (2010) ground-motion simulation methodologies perform best in predicting the probability distributions of the spectral-shape and duration-related IMs, respectively.

It is worth noting that since broadband simulation methods evolve very fast, the intent here is not to provide a definite judgment about the specific simulation methods, but rather to illustrate the proposed validation metrics and approaches and discuss possible outcomes. Indeed, these types of validation exercises can highlight the similarities and differences between simulated and recorded ground motion 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.

ACKNOWLEDGEMENTS

The simulated ground-motion waveform data used in this study were obtained from the SCEC Broadband Platform validation exercise v13.5 and 13.6 available at https://scec.usc.edu/it/June_26_SCEC_BBP_Panel_Review_Meeting. The NGA strong-motion data came from the PEER Ground Motion Database, available at https://ngawest2.berkeley.edu/ (last accessed December 2015). The Southern California Earthquake Center (SCEC) Broadband Platform is available at https://github.com/SCECcode/bbp. 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. The review from Prof. Jack W. Baker (Stanford University, USA) is also gratefully acknowledged.

ORCID

Carmine Galasso http://orcid.org/0000-0001-5445-4911

REFERENCES

1. Bradley BA, Pettinga D, Baker JW, Fraser J. Guidance on the utilization of earthquake-induced ground motion simulations in engineering practice. *Earthq Spectra*. 2017;33(3):809-835.
2. Maechling PJ, Silva F, Callaghan S, Jordan TH. SCEC broadband platform: System architecture and software implementation. *Seismol Res Lett*. 2015;86(1):27-38.
3. D'Amico M, Puglia R, Russo E, Maini C, Pacor F, Luzi L. SYNTHESIS: a web repository of synthetic waveforms. *Bull Earthq Eng*. 2017;15(6):2483-2496.
4. Goulet CA, Abrahamson NA, Somerville PG, Wooddell KE. The SCEC Broadband Platform validation exercise: Methodology for code validation in the context of seismic hazard analyses. *Seismol Res Lett*. 2015;86(1):17-26.
5. Galasso C, Zareian F, Iervolino I, Graves RW. Validation of ground-motion simulations for historical events using SDoF systems. *Bull Seismol Soc Am*. 2012;102(6):2727-2740.
6. Galasso C, Zhong P, Zareian F, Iervolino I, Graves RW. Validation of ground-motion simulations for historical events using MDoF systems. *Earthq Eng Struct Dyn*. 2013;42(9):1395-1412.
7. Graves RW, Pitarka A. Broadband ground-motion simulation using a hybrid approach. *Bull Seismol Soc Am*. 2010;100(5A):2095-2123.
8. Rezaeian S, Zhong P, Hartzell S, Zareian F. Validation of simulated earthquake ground motions based on evolution of intensity and frequency content. *Bull Seismol Soc Am*. 2015;105(6):3036-3049.
9. Dreger DS, Beroza GC, Day SM, et al. Validation of the SCEC broadband platform v14.3 simulation methods using pseudospectral acceleration data. *Seismol Res Lett*. 2015;8(1):39-47.
10. Burks LS, Zimmerman RB, Baker JW. Evaluation of hybrid broadband ground motion simulations for response history analysis and design. *Earthq Spectra*. 2016;31(3):1691-1710.
11. American Society of Civil Engineers (ASCE). *Minimum design loads for buildings and other structures*. Reston, VA: *ASCE/SEI 7-10*; 2010.
12. Burks LS, Baker JW. Validation of ground motion simulations through simple proxies for the response of engineered systems. *Bull Seismol Soc Am*. 2014;104(4):1930-1946.

13. Anderson, JG, Quantitative measure of the goodness-of-fit of synthetic seismogram. in *Proc.* of *13th World Conference on Earthquake Engineering*, 1-6 August, Vancouver, Canada. 2004.
14. Olsen KB, Mayhew JE. Goodness-of-fit criteria for broadband synthetic seismograms, with application to the 2008 M5.4 Chino Hills, California, earthquake. *Seismol Res Lett*. 2010;81(5):715-723.
15. Kullback S. *Information and statistics*. NY.: Wiley; 1959.
16. Jalayer F, Beck JL, Zareian F. Analyzing the sufficiency of alternative scalar and vector intensity measures of ground shaking based on information theory. *ASCE J Eng Mech*. 2012;138(3):307-316.
17. Scott DW, Sain SR. Multivariate Density Estimation. In: *Handbook of Statistics, Data Mining and Data Visualization*. Vol.24 Amsterdam: Elsevier; 2005:229-263.
18. Sørensen MB, Lang DH. Incorporating simulated ground motion in seismic risk assessment: Application to the Lower Indian Himalayas. *Earthq Spectra*. 2015;31(1):71-95.
19. Jayaram N, Lin T, Baker JW. A computationally efficient ground-motion selection algorithm for matching a target response spectrum mean and variance. *Earthq Spectra*. 2011;27(3):797-815.
20. Zeng Y, Anderson JG, Yu G. A composite source model for computing realistic synthetic strong ground motions. *Geophys Res Lett*. 1994;21(8):725-728.
21. Motazedian D, Atkinson GM. Stochastic finite-fault modeling based on a dynamic corner frequency. *Bull Seismol Soc Am*. 2005;95(3):995-1010.
22. Boore DM. Comparing stochastic point-source and finite-source ground-motion simulations: SMSIM and EXSIM. *Bull Seismol Soc Am*. 2009;99(6):3202-3216.
23. Anderson JG. The composite source model for broadband simulations of strong ground motions. *Seismol Res Lett*. 2015;86(1):68-74.
24. Atkinson GM, Assatourians K. Implementation and validation of EXSIM a stochastic finite-fault ground-motion simulation algorithm on the SCEC Broadband Platform. *Seismol Res Lett*. 2015;86(1):48-60.
25. Padgett JE, Nielson BG, DesRoches R. Selection of optimal intensity measures in probabilistic seismic demand models of highway bridge portfolios. *Earthq Eng Struct Dyn*. 2008;37(5):711-725.
26. Shome N, Cornell CA, Bazzurro P, Carballo JE. Earthquakes, records, and nonlinear responses. *Earthq Spectra*. 1998;14(3):469-500.
27. Kazantzi AK, Vamvatsikos D. Intensity measure selection for vulnerability studies of building classes. *Earthq Eng Struct Dyn*. 2015;4(15):2677-2694.
28. Kohrangi M, Bazzurro P, Vamvatsikos D, Spillatura A. Conditional spectrum-based ground motion record selection using average spectral acceleration. *Earthq Eng Struct Dyn*. 2017;46(10):1667-1685.
29. Bojórquez E, Iervolino I. Spectral shape proxies and nonlinear structural response. *Soil Dyn Earthq Eng*. 2011;31(7):996-1008.
30. Iervolino I, Manfredi G, Cosenza E. Ground-motion duration effects on nonlinear seismic response. *Earthq Eng Struct Dyn*. 2006;35(1):21-38.
31. Raghunandan, M., and, A.B. Liel, Effect of ground motion duration on earthquake-induced structural collapse. *Struct Safety*, 2013. 41: 119-133.
32. Chandramohan R, Baker JW, Deierlein GG. Quantifying the influence of ground motion duration on structural collapse capacity using spectrally equivalent records. *Earthq Spectra*. 2016;32(2):927-950.
33. Trifunac MD, Brady AG. A study on the duration of strong earthquake ground motion. *Bull Seismol Soc Am*. 1975;65(3):581-626.
34. Husid LR. *Características de terremotos, Análisis general, Revista del IDIEM 8*. Santiago del Chile; 1969:21-42.
35. Cosenza E, Manfredi G. The improvement of the seismic-resistant design for existing and new structures using damage criteria. In: *Seismic design methodologies for the next generation of codes*. Rotterdam: Balkema; 1997:119-130.
36. Manfredi G. Evaluation of seismic energy demand. *Earthq Eng Struct Dyn*. 2001;30(4):485-499.
37. Iervolino I, Giorgio M, Galasso C, Manfredi G. Conditional hazard maps for secondary intensity measures. *Bull Seismol Soc Am*. 2010;100(6):3312-3319.

How to cite this article: Tsioulou A, Galasso C. Information theory measures for the engineering validation of ground-motion simulations. *Earthquake Engng Struct Dyn*. 2018;47:1095–1104. https://doi.org/10.1002/eqe.3015