Validation of Ground Motion Simulations for Historical Events using Skewed Bridges

C. Galasso[a)], P. Kaviani [b)], A. Tsioulou[a)] and F. Zareian[b)]

[a)] *Department of Civil, Environmental & Geomatic Engineering and Institute for Risk & Disaster Reduction, University College London, UK.*

[b)] *Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA.*

Abstract

This paper provides a statistical comparison between seismic demands of reinforced concrete (RC) bridges with skew-angled seat-type abutments (or simply 'skewed bridges') subjected to past events using simulations and actual recordings. Three short bridges located in California are selected as seed bridges, from which different models are developed by varying key bridge structural parameters such as column-bent height, symmetry of span arrangement, and abutment skew angle. Through extensive nonlinear dynamic analysis conducted using hybrid broadband simulated ground motions and real records for two historical earthquakes; i.e., the 1989 M 6.8 Loma Prieta earthquake and the 1994 M 6.7 Northridge earthquake, it is demonstrated that the distributions of column drift ratios, deck rotations and displacements produced by simulations agree reasonably well with those produced by recorded ground motions. Statistical hypothesis testing and information theory measures are proposed to quantitatively assess the statistical significance of the results for all the considered demand parameters. Finally, ground motion intensity measures (IMs) related to ground motion directionality and directivity, particularly affecting seismic response of skewed bridges, are

compared for both simulations and recorded waveforms. These types of validation exercises can highlight the similarities and differences between simulated and recorded ground motions. The similarities should provide confidence in using the simulation method for bridge engineering applications, while the discrepancies, should help in improving the generation of synthetic records.

KEY WORDS: hybrid broadband simulation; time-history analysis; skewed bridges; ground motion directionality; statistical hypothesis testing.

Introduction and motivation

In current model seismic codes and standards, nonlinear dynamic analysis (NLDA) represents the tool for assessing inelastic structural response in the most explicit and accurate manner. The use of NLDA for seismic performance assessment of existing structures and design for target performance of new ones, requires the availability of hazard-consistent ground motion signals (or simply 'ground motions'). Hazard disaggregation (e.g., [Bazzurro and Cornell, 1999]) in terms of causative magnitude (M) and source-to-site distance (R), is typically identified as a way to establish design earthquakes in terms of M and R. Specifically, for a given mean return period of the seismic hazard at a given site, the level of spectral acceleration at the fundamental structural period, or any other relevant intensity measure (IM) for the specific case-study structure, is disaggregated. This procedure leads to a joint probability distribution of M-R providing the frequency of occurrence of each M-R pair given the exceedance (or occurrence) of the IM being disaggregated. It is possible, then, identifying the mean or modal values of M and R which, being those most contributing to the hazard, are assumed to represent the design or scenario event. Next, a ground motion database can be accessed and a number of records can be selected to match, within tolerable limit, the values of M-R from disaggregation, site

conditions and, eventually, any other parameters supposed to be important for an unbiased estimation of the structural response. The selected records are eventually scaled to match the target value (i.e., hazard level) for the considered IM.

The inherent scarcity or total absence of suitable real (e.g., recorded during past earthquakes) ground motions for some specific scenarios (e.g., large-magnitude, strike-slip events recorded at close source-to-site distances) makes the use of alternative options unavoidable. This is the case, for instance, for several seismically active regions (e.g., California), where the spectral accelerations of interest are often relatively large, and the hazard-controlling earthquake scenarios are typically large-magnitude events on nearby faults. There is significant practical need for realistic waveforms for use as input of NLDA for conditions not well-represented in empirical databases. To this aim, physics-based simulated (or "synthetic") ground motions capturing complex source features (such as spatially variable slip distributions, rise-time, and rupture velocities), path effects (geometric spreading and crustal damping), and site effects (wave propagation through basins and shallow site response) provide nowadays a valuable supplement to recorded ground motions, fulfilling a variety of engineering needs (e.g., [Bradley *et al.*, 2017]). Moreover, synthetic records may be simulated at fine grid spacing, representing an attractive option for many seismic loss problems, such as damage of distributed infrastructure (including transportation networks) and losses to portfolios of structures.

Although physics-based ground motion simulation procedures vary in their methodology and sophistication, they generally utilize deterministic procedures at frequencies below about 1 Hz. High frequency seismic waveforms are difficult to reproduce deterministically, in part because source radiation and wave propagation become increasingly incoherent – and then stochastic – at high frequencies. This results in hybrid broadband ground motion simulations combining the strengths of deterministic procedures at low frequencies and stochastic or semi-stochastic procedures at higher frequencies to generate broadband waveforms. Engineering validation is

an essential step toward the establishment of simulation procedures as a reliable tool for engineering applications. To this aim, a Technical Activity Group (TAG) focusing on Ground Motion Simulation Validation (GMSV) has been established by the Southern California Earthquake Center (SCEC) to develop and implement testing/rating methodologies via collaboration between ground motion modelers and engineering users.

Aligned with the broader objectives of the SCEC GMSV TAG, the study presented in this paper expands the recent authors' efforts (e.g., [Galasso *et al.*, 2012, 2013; Rezaeian *et al.*, 2015; Tsioulou and Galasso, 2017]) toward the engineering validation of the hybrid broadband ground-motion simulation method by Graves and Pitarka [2010]. Particularly, this study focuses on the engineering validation of ground motion simulation in terms of seismic demand of RC bridges with skew-angled seat-type abutments (or simply 'skewed bridges'). A 3-D analytical/numerical model developed by the authors, which is computationally efficient yet adequately detailed, is used [Kaviani *et al.*, 2012]. As the horizontal component of seismic motions propagate at arbitrary directions to bridge axes due to the complex nature of the three dimensional seismic waves, the directionality of earthquake ground motions is explicitly considered in the seismic analysis of selected bridges[1]. This represents a step forward in the current engineering validation of ground motion simulation where little focus has generally been placed on issues related to ground motion directionality, particularly when 3-D nonlinear structural models are used (rather than simplified 2-D structural models, as in most of the previous and concurrent studies). Through extensive nonlinear time-history analyses, the seismic demand of skewed bridges to simulated and recorded ground motions is statistically compared. Different configuration parameters and attributes are considered, including the global torsional resistance, skew angles, column height and span arrangements. Seismic response parameters that are examined include maximum planar deck rotations and

---

[1] Bridge design codes (e.g., Eurocode 8 – Part 2) prescribe seismic forces to be applied, independently and fully, in two orthogonal directions. The choice of orthogonal directions is arbitrary and is left to the discretion of designer.

displacements and maximum column-bent drift ratios. Parametric statistical hypothesis testing and information theory concepts are proposed to quantitatively compare the response measures obtained using the selected recorded and simulated ground motions. Finally, ground motion IMs related to ground motion directionality and directivity, particularly affecting seismic response of skewed bridges, are compared for both simulations and recorded waveforms.

Results from this type of validation exercise can help highlight similarities and differences between synthetic and real records. These similarities should provide confidence in the use of the simulation methodology for engineering applications while the discrepancies, if statistically significant, should help in improving the generation of synthetic records. In a broader perspective, this paper aims to propose a rigorous and transparent statistical approach for the engineering validation of ground motion simulations so that end-users can make an informed decision regarding which methods to use for their forward simulations of earthquake scenarios for which few observations exist.

Description of synthetic and real ground motion datasets

We utilized the hybrid broadband simulation method of Graves and Pitarka [2010]. This method was selected from among several simulation methods principally because of its recent utilization in high-profile scenario earthquake and loss estimation studies (e.g., [Aagaard *et al.*, 2008; Graves *et al.*, 2011]) as well as in similar, recent ground motion simulation validation studies (e.g., [Galasso *et al.*, 2012, 2013; Rezaeian *et al.*, 2015]). However, the proposed validation framework can be seamlessly applied to a variety of simulation methods (e.g., [Tsioulou and Galasso, 2017].

In particular, Graves and Pitarka [2010] developed a hybrid broadband (0-10 Hz) ground motion simulation method which combines a physics-based deterministic approach at low frequency ($\leq$ 1 Hz) with a semi-stochastic approach at high frequency ($>$ 1 Hz). 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 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 [2008] empirical site amplification model.

The present study uses two historical earthquakes modeled by Graves and Pitarka [2010]: 1989 M 6.8 Loma Prieta, and 1994 M 6.7 Northridge. The only earthquake-specific input parameters used in the simulation process are the seismic moment, the overall fault dimensions and geometry, the hypocenter location, and a smoothed representation of the final slip distribution. All other required source parameters (e.g., rupture propagation time, rise time, slip function, and fine-scale slip heterogeneity) are developed using the scaling relations presented by Graves and Pitarka [2010]. Thus, the method provides a reliable framework for generating rupture descriptions for future earthquakes, including extrapolation to magnitude and distance ranges beyond those considered in the current set of validation events, as demonstrated by Graves and Aagaard [2011][2]. Also, it is worth noting that the Graves and Pitarka's simulation method has been recently refined to account for several issues, including the tendency to overpredict the

---

[2] In particular, in the case of a future event, the input parameters for the simulation can either be reliably estimated (e.g., seismic moment, fault dimensions) or parametrically assessed using multiple realizations (e.g., hypocenter location, slip distribution). All other source parameters can be determined using the scaling relations described in Graves and Pitarka [2010].

level of longer period motions (2-5 s; Graves and Pitarka [2015]). However, due to the illustrative nature of the application presented here, we use ground motion simulations from the earlier method, as the new ground motion simulations were not available at the time of this study. For each simulated event, the model region covers a wide area surrounding the fault, including many strong motion recording sites available in the Next Generation Attenuation (NGA) database: 71 for Loma Prieta, and 133 for Northridge. Northridge and Loma Prieta have been selected because they are characterized by the largest number of stations (results for other Californian events are not shown to save space but similar observations can be drawn for those cases). These sites are shown with triangles in **Figure 1**. Similar to Galasso *et al.* [2012, 2013], this study uses a limited number of sites, considering only those that have real recordings with a usable bandwidth larger than 0.1s-8s.

The distribution of the considered sites in terms of source-to-site distance and site conditions is shown in **Figure 2**. Specifically, the considered variables are the closest distance to the fault (in km), and the shear wave velocity in the upper 30 m, V_{S30} (in m/s).

With respect to similar previous and concurrent studies, an effort is made here to investigate the sensitivity of each bridge response to ground motion incidence angle (θ_{GM}) by varying the incidence angle from 0° to 150°, with 30° increments (due to symmetry with respect to the longitudinal axis). The ground motion incidence angle is defined as the angle between the ground motion strike-normal direction and the bridge longitudinal direction. There is indeed significant scientific evidence of sensitivity of bridge response to the direction of seismic excitation with respect to the bridge axis. This is particularly true in the case of skewed bridges, as the direction of seismic excitation is strongly coupled with the contribution of the excited torsional modes of vibrations and the resulting overall response. Taskari and Sextos [2015] show that ground motion directionality has a significant effect on the individual fragility of specific bridge components depending on the structural system and the damage model

considered. Assessing the impact of seismic excitation directionality on the seismic response of skewed bridge may help in revealing issues regarding ground motion directionality in a specific ground motion simulation method (e.g., [Burks and Baker, 2014; Burks *et al.*, 2015]. It is worth noting that past events have also shown some significant incoherence in earthquake ground motions measured at different locations within the spatial dimensions of large horizontally expanded structures (e.g., bridges). The ability of physics-based hybrid simulations to capture the effects of incoherent ground motions, and more in general to capture the spatial variability of ground motions, is not investigated here. In fact, the considered synthetic ground motions are simulated at real recording stations (i.e., where recordings for historical events exist), with typical distances between stations of some kilometers, i.e., much larger than the typical overall length of the considered bridges. The aim is to have simulations and records at the same locations, i.e., at all the recording stations of the considered historical earthquakes. In fact, past events provide an important opportunity to test the ability to use hybrid broadband ground motion simulation to generate synthetic ground motion consistent with those observed.

Description of considered bridges and demand measures

Background

Spatially distributed civil infrastructure systems located in seismically active regions may be vulnerable to moderate-to-strong earthquake events causing system interruption over a long period of time and substantial costs for business interruption, post-event repair and restoration. For instance, the highway transportation network serving the State of California, consisting of a large number of bridges, typically represents such a system. Past earthquakes have shown that bridges are often the most vulnerable components in a transport network and highly susceptible to damage, ranging from loss of serviceability to partial or global collapse during seismic events. Particularly, bridges with skew-angled abutments are at higher risk than straight bridges with

identical features (e.g., mass and dimensions) due to effects engaged by the activation of coupled transversal and longitudinal modes. Skewed bridges, which are very common in the construction practice in California, are constructed to accommodate geometry constraints resulting from alignment of waterway or roadway crossings that occur at an angle that is different from 90°. Reconnaissance reports on the 1994 M 6.7 Northridge earthquake in California (e.g., Astaneh-Asl et al. [1994]) indicated significant damage experienced by skewed highway bridges. In fact, for these bridges, the seismic demand increases with the increase in skew angle, triggering severe damage. Bridges characterized by a high value of the skew angles (e.g., larger than 45°) are generally not recommended by international seismic codes for high seismic zones (e.g., [CEN, 2004; Eurocode 8]).

In general, observations from past earthquakes suggest that there are significant differences between the response of straight and skewed bridges; a manifestation of this difference is the inherent tendency of the decks of skewed bridges to rotate about their vertical axes under seismic excitation, which can lead to unseating from abutments, and ultimately, to collapse. This empirical trend is confirmed by the detailed analytical study of Kaviani *et al.* [2012] indicating that demand parameters for skew-abutment bridges – e.g., deck rotation, abutment unseating, and column drift ratio – and collapse fragility (i.e., likelihood of collapse as function of the ground motion intensity level) are generally higher than those for straight bridges.

The performance expectations for bridges in the United States vary by the importance of the bridge, which is affected by whether the bridge forms part of a lifeline. Lifeline bridges, such as the new San Francisco-Oakland Bay Bridge (2013) are expected to be operational (i.e., little damage) after very rare earthquakes (i.e., corresponding to high mean return periods); significant damage is acceptable in non-important bridges during rare seismic events [NIST, 2011]. The *Guidelines for Nonlinear Analysis for Bridge Structures in California* [Aviram *et al*., 2008] presents a collection of general recommendations for the modeling and analysis of

highway bridges and overpasses subjected to earthquake ground motions, required for the design or evaluation of the seismic capacity (and ductility) of critical bridge components and systems. Rigorous and advanced nonlinear modeling and analysis are generally required in the case of geometric irregularities of the bridge structure, including curves and skews, long spans or significant total length, multiple expansion joints, massive substructure components, or unstable soil conditions. Nonlinear modeling and analysis lead to a more accurate determination of seismic demands of critical components, results that can then be utilized for the final design of the bridge subsystems or for the performance-based assessment of an existing bridge, within the Performance-based Earthquake Engineering (PBEE) framework. To this aim, the accurate estimation of the peak seismic demand of a bridge structure under dynamic excitation requires the use of a suite of ground motion signals, and will therefore further increase the complexity level of the analysis process. *Article 4.7.4.3.4b* (and relative commentary) of the *AASHTO LRFD Bridge Design Specifications* [AASHTO, 2017] provides guidance on selecting and scaling ground motions for performing nonlinear dynamic analysis of bridges. Ground motion selection and modification procedures for NLDA of bridges are consistent with those for NLDA of building structures. Like the case of buildings, properly validated synthetic records appear to be a viable and attractive alternative to the very limited amount of recorded ground motions, particularly in the nearby-field from large earthquakes.

Selection of index bridges

In this study, similar to Kaviani *et al.* [2012], three bridges located in California are selected as 'seed' bridges from which different model variations are generated; see Kaviani *et al.* [2012] for details. The selected bridges have prestressed RC box-girder superstructures and seat-type abutments and primarily differ in their global torsional resistance. The considered geometrical and structural properties of these bridges aim to represent typical features of ordinary bridges in California designed and built after 2000 in regions with high seismicity. In particular, the

first bridge is the Jack Tone Road Overcrossing (denoted as Bridge A) with two spans supported on a single column. The second bridge is the La Veta Avenue Overcrossing (denoted as Bridge B) with two spans supported on a two-column bent. As such, it has a larger global torsional stiffness than Bridge A. The third bridge is the Jack Tone Road Overhead, (denoted as Bridge C) with three spans and two three column bents. To investigate the trends in key response parameters, a matrix of model bridges are generated by using the three bridges (A–C) as seeds and by creating reasonable variations in their configuration and geometric attributes (a sample classification table is shown in **Table 1**). The varied geometrical parameters are:

1) Abutment skew angle, i.e., the angle between an abutment's (or pier's) centerline and the line normal to the roadway centerline. The skew angle for most of the bridges in California vary between 0° and 60°. Given this, we consider the aforementioned variation range in 15° increments;

2) Column-bent height and reinforcement. We consider two column height variation, i.e., original column height (L), and an extended column height (H), corresponding to a 50% extension of the original column-bent height, i.e., H: 1.5 x L. Column reinforcement ratio is set to 1%.

3) Span arrangement. We consider two types span arrangement: symmetrical - equal span length (S), and asymmetrical - ratio of span length equals to 1.2 (A).

The ranges of variation in each parameter are chosen such that they remained true to values encountered in practice, and preserved the main features of the seed bridge. Specifically, they are based on discussion with practicing engineers and by consulting the National Bridge Inventory (NBI). The structural parameters are component material and backfill soil properties.

3-D (nonlinear) modeling of skewed bridges

Several researchers have developed computational models capturing the peculiar damage mechanisms of skewed bridges under seismic loading, up to collapse. In this study, the modeling approach proposed in Kaviani *et al*. [2012], which is computationally efficient yet adequately detailed, is employed. OpenSees is used to carry out the time-history analyses. A representative bridge model used in the simulation is shown in **Figure 3**. The model comprises seat-type abutments, shear keys, expansion joints, column-bents, and the superstructure; modeling details are provided in Kaviani *et al*. [2012]. Such a simplified modeling approach and structural analysis software tool adopted in this study are not capable of simulating every possible collapse mechanism.

Thus, predefined collapse criteria are used here to identify collapse cases based on the analysis output. To this aim, it is assumed that collapse occurs if one of the following criteria occur: (a) Column-bent maximum drift ratio is greater than 8%; and (b) Deck displacement relative to the abutment in the longitudinal unseating direction is greater than the seat length (of each seed bridge).

Three seismic response parameters, or engineering demand parameters (EDPs), are computed, which are the maximum total column drift ratio (θ_{col}), the maximum planar deck rotation (θ_{rot}), and the deck displacement relative to the abutment in the longitudinal unseating direction.

Proposed validation approach

The validity of simulated ground motions is typically assessed based on some quantitative criteria that are used to assess the similarity of simulated and recorded time-histories in terms of IMs or structural response (e.g., EDPs). One common approach adopted by researches involves the use of some goodness-of-fit criteria to compare how well the simulations match the ground motion records (e.g., [Anderson, 2004; Olsen and Mayhew, 2010]). This study

proposes a multi-level statistical approach, combining statistical hypothesis testing (at bridge/EDPs level) and more advanced validation methods based on information theory (at IMs level) as possible testing methods for simulated ground motions to be used in engineering applications. Specifically, ground motion IMs related to ground motion directionality and directivity, particularly affecting seismic response of skewed bridges, are considered here. The following sub-sections provide an overview of the aforementioned validation approaches.

Statistical hypothesis testing for bridge EDPs

Statistical hypothesis testing is a method of statistical inference used for testing scientific models and assumptions. In particular, parametric hypothesis tests are proposed here to quantitatively assess the statistical significance of differences in terms of the proposed EDPs (for a given bridge) due to recorded and simulated ground motions.

Parametric hypothesis tests are performed assuming a lognormal distribution for each EDP of interest, θ_{col}, θ_{rot}, and the deck displacement (relative to the abutment in the longitudinal unseating direction). These distribution assumptions can first be checked with the Shapiro-Wilk (Shapiro and Wilk, 1965) test for a 95% significance level. The null hypothesis (H_o, i.e., the theory we put forward) is that the median EDPs (the mean of the natural logs of each EDP) for simulated ground motions are equal to those from recorded ground motion. To address this aim, we selected a two-tails Aspin-Welch test (Welch, 1938) over the standard Student t-test, as the former does not require the assumption of equal, yet still unknown, variances of populations originating the samples; this could be an unreasonable assumption in some cases. The employed test statistic is reported in Equation (1), in which z_x and z_y are the sample means, s_x and s_y are the sample standard deviations, and m and n are the sample sizes (in this case always equal for each earthquake). The test statistic, under H_o, has a Student t-distribution with the number of degrees of freedom given by Satterthwaite's approximation (Satterthwaite, 1941).

13

$$t = \frac{z_x - z_y}{\sqrt{\dfrac{s_x^2}{n}} + \sqrt{\dfrac{s_y^2}{m}}} \qquad (1)$$

Some recent validation exercises only focused on median values of the considered parameters (e.g., [Goulet *et al.*, 2015]), 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), a similar parametric hypothesis test, the *F*-test (e.g., Mood et al., 1974) for normally distributed data, is performed to compare variances for each bridge and EDP (in logs terms), for the two datasets (recorded and simulated) corresponding to each earthquake (i.e., intra-event variability). In this case, the null hypothesis is that the variance of each EDP for simulated ground motions is equal to the variance from recorded ground motions. The test statistics and corresponding *p*-values for both tests are easy to compute using built-in functions in the R free software environment for statistical computing.

Although not discussed here, it is worth noting that similar nonparametric tests can be employed to compare the empirical distribution functions (and not just their moments) of IMs or EDPs from ground motion records and simulations.

Information theory measures for ground motion IMs

Information theory concepts can be employed 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 [Kullback, 1959] 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, i.e., the empirical distribution of the 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 typically represents a model or approximation of p, i.e., the empirical distribution of the IM 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 Equation (2):

$$D_{KL} = \int_{-\infty}^{+\infty} p(x)\log_2\left(\frac{p(x)}{q(x)}\right)dx \qquad (2)$$

If the logarithm is calculated in base two, D_{KL} is expressed in terms of bits of information. D_{KL} has previously been used in earthquake engineering to compare the relative sufficiency of alternative IMs in predicting structural response [Jalayer et al., 2012].

In the context of ground motion simulation validation, the empirical distribution of the observed IMs estimated from the records and the empirical distribution of the IMs calculated from the simulated ground motions are constructed from the available IM samples through Kernel Density Estimation (KDE); see [Tsioulou and Galasso, 2017] for details. KDE is a non-parametric way to estimate the probability density function of a random variable given a finite number of samples. The D_{KL} value for the two distributions can be subsequently estimated using numerical integration to compute the one-dimensional integral in Equation (2); 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 $D_{KL,obs}$. In statistical hypothesis testing, the p-value for $D_{KL,obs}$ 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, $D_{KL,boot}$, can be calculated. The third step is performed many times, 1,000 in this specific exercise, yielding 1,000 samples of $D_{KL,boot}$. Finally, in the fourth step of the proposed procedure, the p-value for the observed $D_{KL,obs}$ is computed by finding the proportion of the 1,000 $D_{KL,boot}$ samples that are more extreme (i.e., larger) than the $D_{KL,obs}$ value computed 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 95th percentile of the empirical distribution corresponds to a 5% statistical significance level as only $D_{KL,obs}$ values that lie above the 95th percentile are significant. The hypothesis test is an one-sided test in this case.

The proposed validation approach distinguishes itself from the hypothesis testing described above and other proposed goodness-of-fit criteria by assessing the overall similarity of the probability distributions of the studied IMs for recorded and simulated ground motion. Thus, it does not just provide a paired comparison (i.e., at the same recording locations, for historical events) between the recorded and simulated IM datasets in terms of 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 modelling purposes [Sørensen and Lang, 2015]. 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 (e.g., [Jayaram *et al*., 2011]).

Results and discussion

The two sets of simulated and recorded ground motions computed for Loma Prieta and Northridge are used as input for nonlinear time-history analyses of the case-study bridges. As discussed, these ground motions are applied to the bridges with varied incidence angles to investigate the bridges' response-sensitivity with respect to the direction of seismic load. Given the bridge variations and the number of ground motions in each dataset, 3,160 nonlinear time-history analyses have been carried out for each bridge [5 skew angles x 2 column heights x 2 span arrangements x (36 + 43) ground motions x 2 sets (recorded and simulated)]. Results are presented here only for selected bridge models (and in terms of θ_{col} and θ_{rot}), representing 'extreme' variations of the considered properties: Bridge A with symmetrical span arrangement and lower column height (referred hereinafter as Bridge 1) and skew angles equal to 0°, 30°,

and 60°; Bridge C with asymmetrical span arrangement and higher column height (referred hereinafter as Bridge 2) and skew angles equal to 0°, 30°, and 60°. These cases are highlighted in grey in **Table 1**. Results for the other bridge models and in terms of deck displacement relative to the abutment are not shown to save space but similar findings and conclusions can be drawn for those cases.

For each considered earthquake and bridge model, we first compare the recordings and simulations using the model bias and standard error for each considered EDP (θ_{col} and θ_{rot}), following a similar approach to Graves and Pitarka [2010]. Specifically, for the *j*-th station, the residual between the EDPs due to recorded and simulated ground motions with incidence angle ($\theta_{GM,i}$) is given by Equation (3):

$$r_j\left(\theta_{GM,i}\right) = \ln\left[EDP_{sim,j}\left(\theta_{GM,i}\right)/EDP_{rec,j}\left(\theta_{GM,i}\right)\right] \tag{3}$$

where $EDP_{sim,j}\left(\theta_{GM,i}\right)$ and $EDP_{rec,j}\left(\theta_{GM,i}\right)$ are the EDPs (θ_{col} and θ_{rot}) due to recorded and simulated ground motions with incidence angle ($\theta_{GM,i}$), respectively. The model bias is then given by Equation (4):

$$B\left(\theta_{GM,i}\right) = \frac{1}{N}\sum_{j=1,N} r_j\left(\theta_{GM,i}\right) \tag{4}$$

and the standard error is given by Equation (5):

$$\sigma\left(\theta_{GM,i}\right) = \left\{\frac{1}{N}\sum_{j=1,N}\left[r_j\left(\theta_{GM,i}\right) - B\left(\theta_{GM,i}\right)\right]^2\right\}^{1/2} \tag{5}$$

where N is the total number of considered stations for each event. **Figure 4** and **Figure 5** show the individual residuals as well as the model bias and standard error across the range of incidence angles being considered (for different bridges and skew angles) for the Loma Prieta and Northridge simulations, respectively.

A model bias above zero, if statistically significant, means that the synthetic records tend to produce, on average, systematically more damaging bridge seismic response than real records. Conversely, deviations below zero indicate that the simulated records tend to be, on average, more benign in producing bridge seismic responses than those in nature. A direct comparison of response statistics is acceptable, as the simulated datasets are developed to match exactly the same earthquakes and site conditions (i.e., at the same stations) of the real recordings. It is worth noting that model bias calculated with Equations (3) and (4) is sensitive to small $EDP_{rec,j}\left(\theta_{GM,i}\right)$ values, possibly resulting in high values of the model bias. However, this is not an issue for the considered ground motions sets.

It is worth noting that the model region for each considered event covers a wide area surrounding the fault, including many strong motion recording sites, in a large range of distances (e.g., **Figure 1** and **Figure 2**). To analyze the number of bridge responses in the nonlinear range, the ratio of the recorded spectral acceleration at the first period of the structure, $S_{a,rec}\left(T_1\right)$, divided by yield base shear coefficient, γ, is computed for each ground motion/station. The base shear coefficient is computed as the ratio between the base shear at yielding point (from a pushover analysis) to the seismic weight of each bridge. One can assume that ground motions characterized by $S_{a,rec}\left(T_1\right)\big/\gamma$ ration larger than unity (i.e., 50-60% of the total number of records) would possibly force the structure into inelastic response. However, all the ground motions have been retained in the analysis presented here. This allows to assess

the full intra-event variability of seismic demands induced by simulated and recorded ground motions for the two historical events. In the case of bridges behaving in the elastic range, the analysis can still highlight the ability of ground motion simulations to capture the effect of the higher modes on the overall bridge response.

The comparisons shown in **Figure 4** and **Figure 5** exhibit little systematic model bias (i.e., near zero) across the entire angle of incidence range: bridges seismic response to simulated waveforms agree reasonably well with that to observed ground motions, particularly in the case of θ_{col} (bias $\leq \pm 10\%$ in all cases). Generally larger values of the bias are observed in the case of θ_{rot}, with a few peaks around $\pm 30\%$. This is generally true for the all the considered events and bridges combinations. The standard error ranges from about 0.4 to 0.7 natural log units, with a very few cases above 0.7 natural log units.

The parametric, two-tails Aspin-Welch test (as discussed above) is performed to quantitatively assess the statistical significance of the results found in terms of bias of the bridges EDPs (for each ground motion incidence angle and skew angle) to recorded and simulated ground motions: i.e., to assess whether the considered EDPs (in logs terms) come from independent random samples from normal distributions with equal means and equal but unknown variances. The normal distribution assumptions are first checked with the Shapiro-Wilk test and are not rejected, at a 95% significance level. The parametric, two-tail *F*-test for equal variances is also performed and no rejections are observed at a 95% significance level (to save space, detailed results for this test are not shown).

To summarize the results of the two-tails Aspin-Welch hypothesis test and draw conclusions, the obtained *p*-values from the test are reported in **Tables 2-5** for the different EDPs and the two bridge models. Specifically, **Tables 2-3** present the results for the Loma Prieta simulations while **Tables 4-5** present the results for the Northridge simulations.

We assume that 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 discrepancies in the seismic demand distributions from simulations and real records are statistically significant. These cases are reported in bold type in **Tables 2-5**. For the hypothesis tests yielding a *p*-value greater than 0.05 and smaller than 0.10 (10%), there may be some evidence to reject the null hypothesis. These cases are reported in italic type in **Tables 2-5**. Finally, for the hypothesis tests yielding a *p*-value greater than 0.05 there is not sufficient evidence to reject the null hypothesis, meaning that the discrepancies in the seismic demand distribution from simulations and real records are not statistically significant.

Based on the results from **Tables 2-5**, the discrepancies in the seismic demands estimated using simulated and real ground motions appear not to be statistically significant for all the considered cases with only very few, sparse rejections observed. In general, these results confirm the considerations based on the visual inspection of **Figures 4-5**. In some cases, the limited sample size and the relatively large variability may not confirm the hypothesis that the median seismic response generated by simulated ground motions (particularly in terms of θ_{rot}) differ systematically from those produced by recorded ground motions at the two customary significance levels (i.e., 5% and 10%).

These generally very favorable comparisons between simulations and recorded data can support the simulation methodology's predictive capabilities, and are directly relevant to the engineering community, who may use the simulation methodology with confidence.

Ground motion directionality

Uncertainty related to the directionality of the incoming ground shaking, that is, the orientation of the propagating seismic wavefront with respect to a structure's axis, is often poorly investigated, despite the scientific evidence of significant sensitivity of structural response to

the direction of seismic loading (e.g., [Taskari and Sextos, 2015]). As discussed in Taskari and Sextos [2015], a few studies have essentially shown that there is not a critical angle of excitation that may simultaneously trigger the most unfavorable response in all considered structural components and EDPs, either in the linear or in the nonlinear range. This also applies to bridges, and particularly skewed ones, for which the direction of seismic loading is strongly coupled with the contribution of the excited torsional modes of vibration and the resulting bridge response. In this case, the current seismic design framework (e.g., *AASHTO LRFD Bridge Design Specifications*) only provides some general guidance regarding the direction of seismic loading, when nonlinear time history analysis is employed. In particular, it is required that at least seven sets of independent ground motions are applied in orthogonal direction at the bridge supports. Uncertainty in ground motion directionality is addressed by orienting the ground motions at different angles, i.e., 0°, 30°, 60°, and 90°. Hence, the bridge is designed for the average (across different orientations and ground motion records) of the recorded peak responses.

The improved characterization of ground motion directionality (and directivity) for engineering applications has been recently addressed by the NGA-West2 research program [Bozorgnia *et al.*, 2014], coordinated by the Pacific Earthquake Research Center (PEER). One of the objectives of the NGA-West2 project was to provide refined model for predicting direction-dependent ground motion spectra. Therefore, in this section we aim to investigate metrics and models for ground-motion directionality developed as part of that project [Shahi and Baker, 2013], which have also been proposed as tools for validation of ground motion simulation by Burks and Baker [2014]. These metrics can be considered effective proxies for the earthquake damage potential with respect to a specific class of engineered systems, for instance skewed bridges, possible explaining the few 'rejections' in the previous sections in terms of bridges EDPs for recorded and simulated ground motions.

For a multicomponent ground motion, a spectral acceleration value can be computed for the shaking in any horizontal direction. It is well known that the resulting spectral acceleration depends on the considered orientation. Variation of spectral acceleration with orientation is captured by the parameters $SA_{RotD100}$ and SA_{RotD50}. For a given period, spectral accelerations can be computed for rotation angles from 0° to 180° (because 180°–360° are redundant): SA_{RotDnn} is the nth percentile of the results. Thus, $SA_{RotD100}$ is the maximum and SA_{RotD50} is the median of the spectral accelerations of a ground motion when rotated over all horizontal orientations. $SA_{RotD100}$ and SA_{RotD50} are computed independently at each period, so it is highly unlikely to observe a single maximum or median orientation for a given ground motion. The ratio of $SA_{RotD100}$ to SA_{RotD50} at a specific period is a proxy for the polarization of a ground motion. If $SA_{RotD100}$ is approximately equal to SA_{RotD50} (i.e., $SA_{RotD100}/SA_{RotD50} \approx 1$), then the structural response is about the same in all orientations. If $SA_{RotD100}$ is much larger than SA_{RotD50} then the structural response is polarized in one orientation, with $SA_{RotD100}/SA_{RotD50} = \sqrt{2}$. The median $SA_{RotD100}/SA_{RotD50}$ ratio from a suite of ground motions has a very stable relationship with period and is not dependent on magnitude, distance, or local site conditions [Shahi and Baker, 2013]. This ratio is important for predictions of structural behavior, especially for 3D structural models, which respond in all orientations. Therefore, ground motion simulations should have ratios consistent with empirical models (e.g., [Burks and Baker, 2014]).

$SA_{RotD100}$ to SA_{RotD50} ratios from recordings and simulations of both Loma Prieta and Northridge earthquakes are compared in **Figure 6**; also, the median $SA_{RotD100}$ to SA_{RotD50} ratios from recordings and simulations are compared to the empirical model proposed by Shahi and Baker [2013] in **Figure 7**. Only the two periods corresponding to the two bridges (0.6s for Bridge 1 and 1.4s for Bridge 2) are considered here. The visual inspection of **Figure 6** and **Figure 7** confirm that the simulations match the empirical models reasonably well at the two considered periods, except at 1.4s for Loma Prieta (i.e., for Bridge 2). In this case, the $SA_{RotD100}$ to SA_{RotD50}

ratios for the simulations are on average larger than expected, indicating that the amplitude of the simulations varies strongly with orientation and that the simulations will tend to polarize structural response more than recordings.

This result is confirmed by quantitatively comparing the two distributions of $SA_{RotD100}$ to SA_{RotD50} in terms of the proposed information theory measure D_{KL} introduced above and shown in **Figure 8** and **Figure 9** for Loma Prieta and Northridge respectively. 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. In this case, 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. Specifically, **Figures 8-9** show the empirical CDF for the D_{KL} values for the $SA_{RotD100}$ to SA_{RotD50} ratios constructed from the bootstrapped samples as well as the observed D_{KL} values from the original samples. The red dotted line on the graph indicates the D_{KL} value corresponding to the 95th percentile of the empirical CDF.

With respect to the results of the hypothesis testing, all the observed D_{KL} values ($D_{KL,obs}$) are within the 95% confidence intervals established through bootstrapping, except for the D_{KL} values for Loma Prieta in the case of T = 1.4 (Bridge 2). In this latter case, the observed D_{KL} value is above the 95% confidence interval indicating statistically significant differences between the recorded and simulated $SA_{RotD100}$ to SA_{RotD50} distributions.

A similar result was found by Burks *et al*. [2015] for a real building located in Berkeley, CA. The authors concluded that this may be due to simplifications of the hybrid broadband method on the broadband platform, such as the use of 1-D Green's functions, 1-D velocity models, and single planar faults.

Presence of pulse-like records

Ground motions recorded at near-source sites may be subjected to rupture directivity effects which result in a low frequency full cycle velocity pulse at the beginning of the signal. The occurrence of this effect depends on the rupture process, the geometrical configuration of the fault and site as well as other factors, (e.g., [Somerville *et al.*, 1997]). Elastic demand amplitude of pulse-like signals is generally larger than that of ordinary recordings, particularly concerning the fault-normal (FN) direction; also the spectral shape is non-standard with an increment of spectral ordinates in the range around the pulse period (e.g., [Chioccarelli and Iervolino, 2013]). NGA records, were classified as *pulse-like* and *non-pulse-like* according to Baker [2007] via a wavelets-based algorithm, which assigns a score, a real number between 0 and 1, to each record and determines the pulse period (T_P). The larger the score the more likely the record is to contain a pulse. Only the FN ground motions having a pulse score equal or larger than 0.85 are, arbitrarily, counted as pulse-type records. In particular, according to Baker's [2007] study, Northridge event presents several ground motions classified as pulse-like (11) while Loma Prieta features only 4 pulse-like ground motions.

In this section, the algorithm proposed by Baker [2007] is applied to the recorded and simulated datasets for Loma Prieta and Northridge events in order to investigate possible differences in terms of pulse-like ground motions that possibly explain the differences found in the previous sections in terms of EDPs for skewed bridges. For each earthquake and dataset, horizontal components have been rotated into fault-parallel (FP; 128° azimuth for Loma Prieta, and 122° azimuth for Northridge) and FN (218° azimuth for Loma Prieta, and 212° azimuth for Northridge) orientations.

Table 6 reports, for each earthquake, the number of identified pulse-like records in the recorded dataset (for Northridge this number is slightly different from that reported in Baker [2007]

because here only a subset of NGA stations has been considered for each event), the simulated dataset and the number of records classified as pulse-like in both datasets.

Table 6 shows that simulations for Loma Prieta do not reproduce well the pulse-like ground motions in the recorded dataset (0/4) with 5 additional ground motions identified as pulse-like in the simulated dataset and not in the NGA. For Northridge only 3/6 ground motions are classified as pulse-like in both datasets; two additional ground motions are identified as pulse-like in the simulated dataset but not in the NGA. **Table 7** shows that some significant differences exist in the values of T_p even for the ground motions classified as pulse-like in both datasets: for Northridge T_p values for the simulated dataset are quite well captured for 2/3 ground motions (discrepancy less than 10%) while a much larger (double) T_p value is observed for one of the simulation.

These results seem to suggest that the simulations likely tend to produce stronger directivity effects on ground motion amplitudes than they should (because of too strong coherence in the rupture descriptions) with large variance. Generally, it is difficult to quantify/calibrate this since directivity effects seem to be underestimated by current empirical models due to the scarcity of recordings. The pulse-period classification of Baker [2007] gets at part of this issue, but it does not really address the amplification due to directivity. It is evident that modelling of strong directivity effects in the simulations requires further study to increase confidence in their predictive capabilities for future earthquake scenarios. Given the scarcity of near-fault recordings, ground motion simulations containing pulses can help refine existing directivity prediction models.

Conclusions

The need for ground motions closely matching a specific target scenario (e.g., a hazard-consistent design event in terms of a magnitude-distance pair) together with the lack of recorded

ground motions for some specific scenarios, is supporting the use of ground motion simulations by which one can produce synthetic records with nominal characteristics of interest. This has the potential to generate more accurate representations of the ground motions at a given site than those available using ground motion recordings from other geographical regions. In particular, hybrid broadband ground motion simulation procedures typically utilize a physics-based modeling of source and path effects at low frequencies coupled with semi-stochastic procedures at high frequencies. The use of hybrid broadband simulations in the engineering practice is contingent to a rigorous and transparent engineering validation of ground motion simulations so that end-users can finally use forward simulations of earthquake scenarios for which few observations exist.

In this study, a statistical approach for such an engineering validation of hybrid broadband ground motion simulations has been presented through an application to the seismic demands of RC box-girder bridges with seat-type abutments. Skewed bridges are often encountered in highway design when the geometry cannot accommodate straight bridges. Post-earthquake reconnaissance studies have reported that larger values of skew angles adversely affect the seismic performance of such a bridge typology.

Three actual bridges were used as seed models and a matrix of bridge models was created by varying their abutment skew angles, number of columns and column heights, number of spans and span arrangement. An extensive set of nonlinear response history analysis was conducted using hybrid broadband simulations and actual recordings for two past earthquakes; i.e., 1989 M 6.8 Loma Prieta earthquake and 1994 M 6.7 Northridge earthquake.

The illustrative application of the proposed validation approach shows that simulation matches well the seismic demands produced by recorded ground motions: simulated accelerograms produce seismic demands that appear statistically equivalent to demand estimates using corresponding recorded motions at the same stations for most of the considered case-study

structures and EDPs. Assessment of the results using formal statistical hypothesis tests indicates that in most cases the differences found are not significant, increasing the confidence in the use of simulated motions for engineering applications.

The statistical differences in the directional properties of the recorded and simulated datasets have been examined next. For one of the two considered events, the differences in the directional properties of the two datasets were found to be significant explaining some of the differences observed in the EDPs of the studied bridges.

It is worth noting that the intent here is not to provide a definite judgement about the specific simulation method, 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 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. The findings of this study, even though obtained based on limited sets of ground motion records and structures, are in good agreement with previous similar studies based on simplified structural systems (elastic and inelastic SDoF and MDoF systems).

References

Aagaard, B. T., Brocher , T. M., Dolenc , D., Dreger , D., Graves , R. W., Harmsen S., Hartzell , S., Larsen , S., and Zoback M. L. [2008] "Ground-motion modeling of the 1906 San Francisco earthquake, part I: Validation using the 1989 Loma Prieta earthquake", Bulletin of the Seismological Society of America, 98(2), 989–1011. doi: 10.1785/0120060409.

AASHTO [2017] AASHTO LRFD bridge design specifications (8th edition). Washington, D.C., USA.

Anderson, J. G. (2004) "Quantitative measure of the goodness-of-fit of synthetic seismograms", Proc. of the 13th World Conference on Earthquake Engineering, Vancouver, Canada, pp. 243.

Astaneh-Asl, A., Bolt, B., McMullin, K. M., Roupen, R. D., Modjtahedi, D. and Cho, S.-W. [1994] "Seismic performance of steel bridges during the 1994 Northridge earthquake," Report No. UCB CEE-Steel-94/01.

Aviram, A., Mackie, K. R. and Stojadinovic, B. [2008] "Effect of abutment modeling on the seismic response of bridge structures", Earthquake Engineering and Engineering Vibration, 7(4), 395–402. doi: 10.1007/s11803-008-1008-3.

Baker, J. W. [2007] "Quantitative classification of near-fault ground motions using wavelet analysis", Bulletin of the Seismological Society of America, 97(5), 1486–1501. doi: 10.1785/0120060255.

Bazzurro, P. and Cornell, C. A. [1999)] "Disaggregation of seismic hazard", Bulletin of the Seismological Society of America, 89(2), 501–520. doi: 10.1785/0120060093.

Bozorgnia, Y. *et al.* (2014) 'NGA-West2 research project', *Earthquake Spectra*, 30(3), 973–987. doi: 10.1193/072113EQS209M.

Bradley, B. A., Pettinga, D., Baker, J. W. and Fraser, J. [2017] "Guidance on the utilization of earthquake-induced ground motion simulations in engineering practice", Earthquake Spectra, 33(3),809-835.

Burks, L. S. and Baker, J. W. [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. doi: 10.1785/0120130276.

Burks, L. S., Zimmerman, R. B. and Baker, J. W. [2015] "Evaluation of hybrid broadband ground motion simulations for response history analysis and design", Earthquake Spectra, 31(3), 1691–1710. doi: 10.1193/091113EQS248M.

Campbell, K. W. and 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. doi: 10.1193/1.2857546.

CEN [2004] *Eurocode 2: Design of concrete structures*. CEN, Brussels, EN 1992-1-1:2004.

Chioccarelli, E. and Iervolino, I. [2013] "Near-source seismic hazard and design scenarios", Earthquake Engineering & Structural Dynamics, 42(4), 603–622. doi: 10.1002/eqe.2232.

Galasso, C., Zareian, F., Iervolino, I. and Graves, R. W. [2012] "Validation of ground-motion simulations for historical events using SDoF systems", Bulletin of the Seismological Society of America, 102(6), 2727–2740. doi: 10.1785/0120120018.

Galasso, C., Zhong P., Zareian, F., Iervolino, I. and Graves, R. W. [2013] "Validation of ground-motion simulations for historical events using MDoF systems", Earthquake Engineering & Structural Dynamics, 42(9), 1395–1412. doi: 10.1002/eqe.2278.

Goulet, C. A., Abrahamson, N. A., Somerville, P. G. and Wooddell, K. E. [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. doi: 10.1785/0220140104.

Graves, R. and Pitarka, A. [2015] "Refinements to the Graves and Pitarka (2010) Broadband Ground-Motion Simulation Method", Seismological Research Letters, 86(1), 75–80. doi: 10.1785/0220140101.

Graves, R. W., Aagaard, B. T. and Hudnut, K. W. [2011] "The ShakeOut earthquake source and ground motion simulations", Earthquake Spectra, 27(2), 273–291. doi: 10.1193/1.3570677.

Graves, R. W. and Pitarka, A. [2010] "Broadband ground-motion simulation using a hybrid approach", Bulletin of the Seismological Society of America, 100(5 A), 2095–2123. doi: 10.1785/0120100057.

Jalayer, F., Beck, J. L. and Zareian, F. [2012] "Analyzing the Sufficiency of Alternative Scalar and

Vector Intensity Measures of Ground Shaking Based on Information Theory", Journal of Engineering Mechanics, 138(3), 307–316. doi: 10.1061/(ASCE)EM.1943-7889.0000327.

Jayaram, N., Lin, T. and Baker, J. W. [2011] "A Computationally efficient ground-motion selection algorithm for matching a target response spectrum mean and variance", Earthquake Spectra, 27(3), 797–815. doi: 10.1193/1.3608002.

JRC European Commission and International Code Council [2004] *Eurocode 8 : Seismic Design of Buildings Worked examples*, *European Committee for Normalization, Brussels*.

Kaviani, P., Zareian, F. and Taciroglu, E. [2012] "Seismic behavior of reinforced concrete bridges with skew-angled seat-type abutments", Engineering Structures. Elsevier Ltd, 45(December), pp. 137–150. doi: 10.1016/j.engstruct.2012.06.013.

Kullback, S. [1959] Information and Statistics, Wiley, New York, USA.

Mood, M.A., Graybill, F. A. and Boes D. C. [1974] Introduction to the Theory of Statistics (3rd edition), McGraw-Hill, New York, USA.

NIST [2011] Selecting and Scaling Earthquake Ground Motions for Performing Response-History Analyses, NIST/GCR 11-917-15, Gaithersburg, Maryland.

Olsen, K. B. and Mayhew, J. E. [2010] "Goodness-of-fit Criteria for Broadband Synthetic Seismograms, with Application to the 2008 Mw 5.4 Chino Hills, California, Earthquake", Seismological Research Letters, 81(5), 715–723. doi: 10.1785/gssrl.81.5.715.

Rezaeian, S., Zhong, P., Hartzell, S. and Zareian, F. [2015] "Validation of simulated earthquake ground motions based on evolution of intensity and frequency content", Bulletin of the Seismological Society of America, 105(6), 3036–3049. doi: 10.1785/0120140210.

Satterthwaite, F. E. [1941] "Synthesis of variance", *Psychometrika*, 6(5), 309-316.

Shahi, S. K. and Baker, J. W. [2013] NGA-West2 Models for Ground-Motion Directionality. Technical Report PEER 2013/10, Berkeley, CA.

Shapiro, S. S. and Wilk, M. B. [1965] "An analysis of variance test for normality (complete samples)",

Biometrika, 52(3-4), 591-611.

Somerville, P. G., Smith, N. F., Graves, R. W. and Abrahamson, N. A. [1997] "Modification of Empirical Strong Ground Motion Attenuation Relations to Include the Amplitude and Duration Effects of Rupture Directivity", Seismological Research Letters, 68(1), 199–222. doi: 10.1785/gssrl.68.1.199.

Sørensen, M. B. and Lang, D. H. [2015] "Incorporating simulated ground motion in seismic risk assessment: Application to the lower Indian Himalayas", Earthquake Spectra, 31(1), 71–95. doi: 10.1193/010412EQS001M.

Taskari, O. and Sextos, A. [2015] "Multi-angle, multi-damage fragility curves for seismic assessment of bridges", Earthquake Engineering and Structural Dynamics, 44, 2281–2301. doi: 10.1002/eqe.2584.

Tsioulou, A. and Galasso, C. [2017] "Information theory measures for the engineering validation of ground motion simulations", Earthquake Engineering & Structural Dynamics, 47(4), 1095-1104. doi: 10.1002/eqe.3015.

Welch, B. L. [1938] "The significance of the difference between two means when the population variances are unequal", *Biometrika*, 29, 350-362.

Table 1. Classification matrix of 60 different bridge models generated from the three seed bridges.

Bridges	Col. Height	Symmetric					Asymmetric				
		0°	15°	30°	45°	60°	0°	15°	30°	45°	60°
Two-span, Single-col. (A)	Higher	AHS0	AHS1	AHS2	AHS3	AHS4	AHA0	AHA1	AHA2	AHA3	AHA4
	Lower	ALS0	ALS1	ALS2	ALS3	ALS4	ALA0	ALA1	ALA2	ALA3	ALA4
Two-span, Multi-col. (B)	Higher	BHS0	BHS1	BHS2	BHS3	BHS4	BHA0	BHA1	BHA2	BHA3	BHA4
	Lower	BLS0	BLS1	BLS2	BLS3	BLS4	BLA0	BLA1	BLA2	BLA3	BLA4
Three-span, Multi-col. (C)	Higher	CHS0	CHS1	CHS2	CHS3	CHS4	CHA0	CHA1	CHA2	CHA3	CHA4
	Lower	CLS0	CLS1	CLS2	CLS3	CLS4	CLA0	CLA1	CLA2	CLA3	CLA4

Table 2. Aspin-Welch test results in terms of θ_{col} for Loma Prieta earthquake; p-values lower than 0.05 are reported in bold.

Skew angle	Bridge	θ_{inc}					
		0°	30°	60°	90°	120°	150°
0°	1	0.78	0.35	0.68	0.76	0.75	0.69
	2	0.89	0.74	0.58	0.78	0.99	0.88
30°	1	0.57	0.59	0.94	0.95	0.64	0.65
	2	0.79	0.97	0.85	0.84	0.68	0.62
60°	1	0.67	0.55	0.57	0.72	0.47	0.56
	2	0.64	0.89	0.90	0.76	0.48	0.54

Table 3. Aspin-Welch test results in terms of θ_{rot} for Loma Prieta earthquake; p-values lower than 0.05 are reported in bold.

Skew angle	Bridge	θ_{inc}					
		0°	30°	60°	90°	120°	150°
0°	1	0.17	0.27	0.17	0.67	0.46	0.91
	2	0.22	**0.03**	**0.05**	*0.09*	0.97	0.60
30°	1	0.81	0.60	0.67	0.88	0.75	0.94
	2	0.79	0.95	0.72	0.43	0.26	0.91
60°	1	0.88	0.88	1.00	0.96	0.60	0.77
	2	0.78	0.61	0.80	0.99	0.65	0.67

Table 4. Aspin-Welch test results in terms of θ_{col} for Northridge earthquake; p-values lower than 0.05 are reported in bold.

Skew angle	Bridge	θ_{inc}					
		0°	30°	60°	90°	120°	150°
0°	1	0.40	0.29	0.42	0.70	0.44	0.36
	2	0.87	0.79	0.78	0.66	0.61	0.90
30°	1	0.35	0.31	0.71	0.90	0.43	0.30
	2	0.94	0.96	0.97	0.69	0.89	0.76
60°	1	0.34	0.31	0.66	0.77	0.41	0.30
	2	0.65	0.64	0.79	0.79	0.87	0.88

Table 5. Aspin-Welch test results in terms of θ_{rot} for Northridge earthquake; *p*-values lower than 0.05 are reported in bold.

Skew angle	Bridge	θ_{inc}					
		0°	30°	60°	90°	120°	150°
0°	1	0.88	0.96	0.45	0.62	0.67	0.43
	2	0.56	0.62	0.87	0.93	0.95	0.40
30°	1	0.86	0.34	0.37	0.81	0.81	0.57
	2	0.78	0.83	0.92	0.70	0.67	0.24
60°	1	0.49	0.36	0.49	0.57	0.65	0.69
	2	0.94	0.67	0.75	0.95	0.35	0.65

Table 6. Number of FN ground motions identified as pulse-like using the classification procedure proposed by Baker [2007]

	Simulated	Recorded	Both
Loma Prieta	5	4	0
Northridge	5	6	3

Table 7. T_p values for simulated and recorded for Northridge.

GM ID	T_p Simulated	T_p Recorded	Sim/Rec
2	2.38	2.34	1.02
8	3.05	1.51	2.02
31	1.96	2.13	0.92
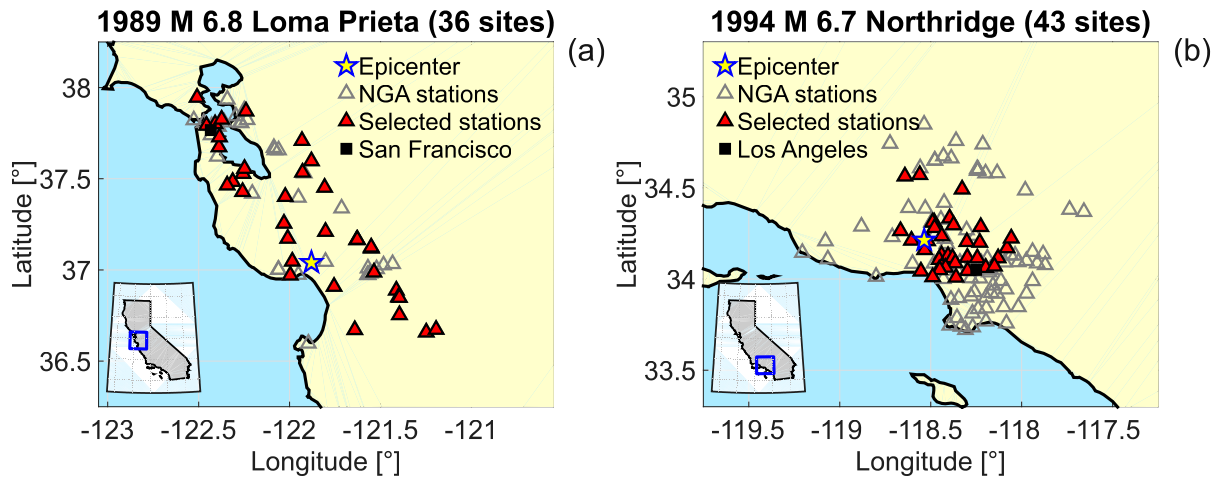
Figure 1. Maps of the earthquakes considered. The star is the epicenter and the triangles are the recording stations in the NGA database for which the simulations are available. The filled triangles are the recording stations considered in this study. San Francisco (b) and Los Angeles (d) are also indicated on the map (squares).
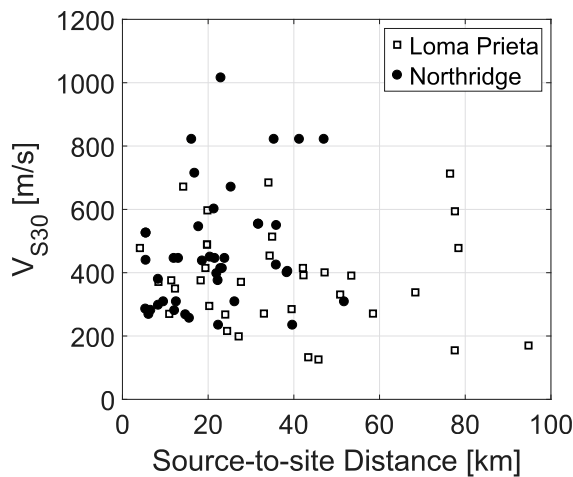

Figure 2. Distribution of the considered sites for the two events in terms of source-to-site distance and shear wave velocity in the upper 30m.
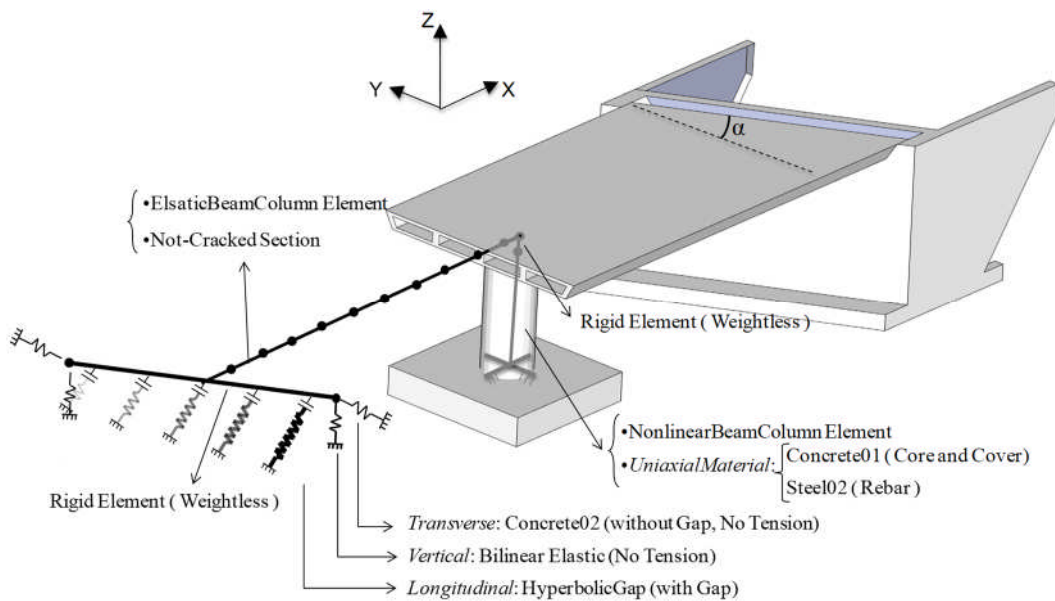

Figure 3. Generic model (α°-skew) used for nonlinear time-history analyses (adapted from Kaviani et al., 2012).
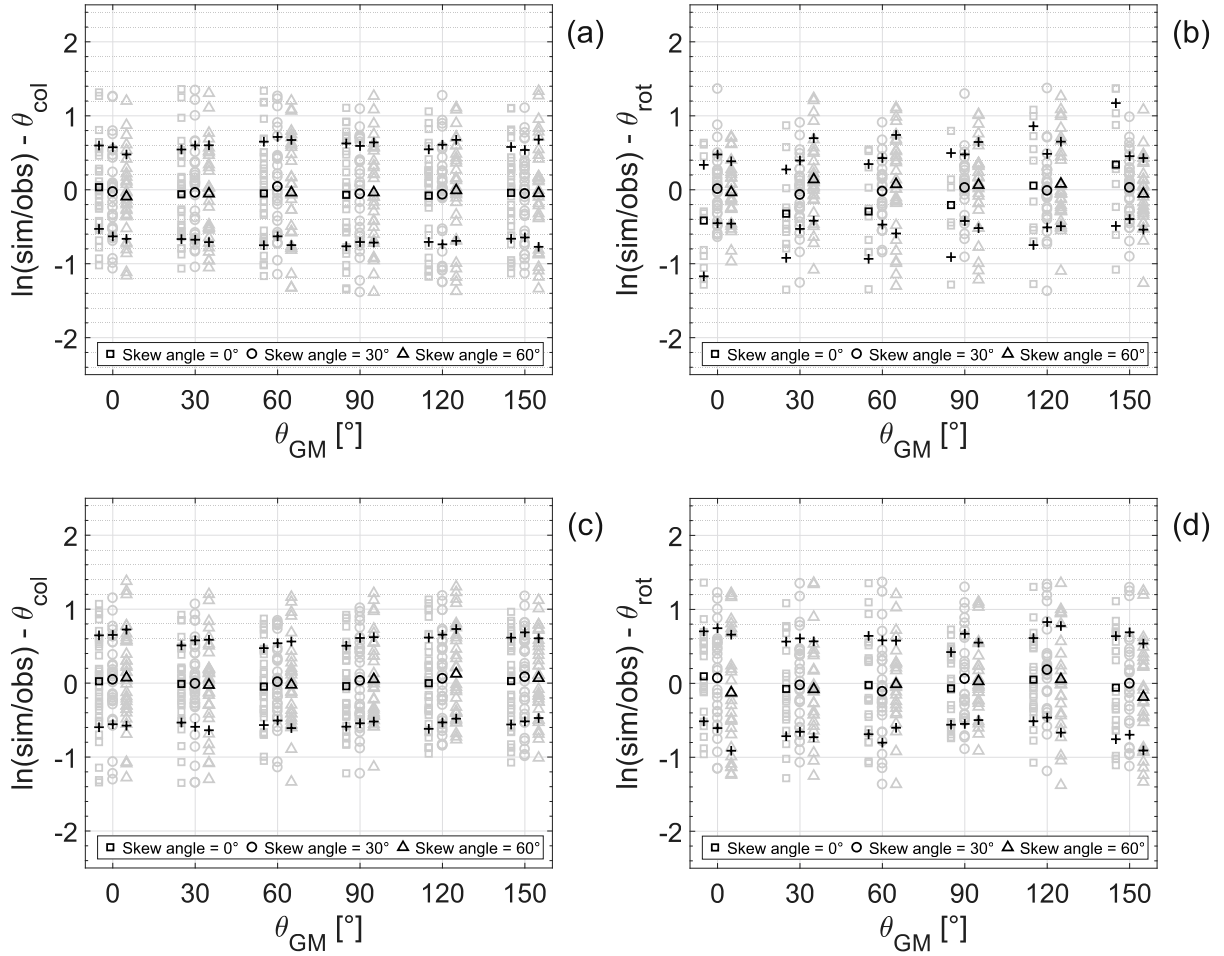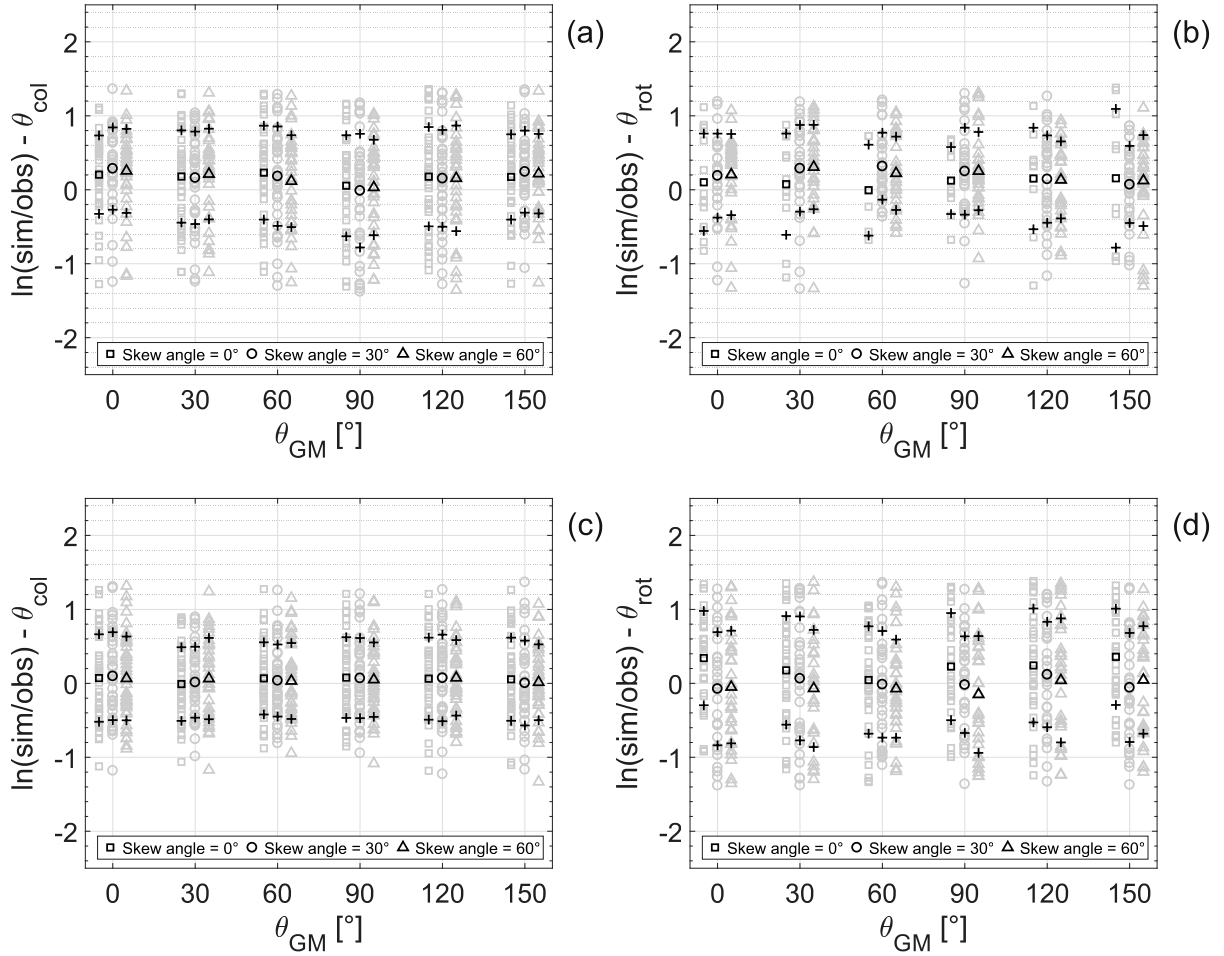
Figure 4. Model bias (heavy line markers) and standard error ('+' markers) for the Loma Prieta earthquake. (a) Panel is for Bridge 1 and θ_{col}; (b) panel is for Bridge 1 and θ_{rot}; (c) panel is for Bridge 2 and θ_{col}; and (d) panel is for Bridge 2 and θ_{rot}.

Figure 5. Model bias (heavy line markers) and standard error ('+' markers) for the Northridge earthquake. (a) Panel is for Bridge 1 and θ_{col}; (b) panel is for Bridge 1 and θ_{rot}; (c) panel is for Bridge 2 and θ_{col}; and (d) panel is for Bridge 2 and θ_{rot}.
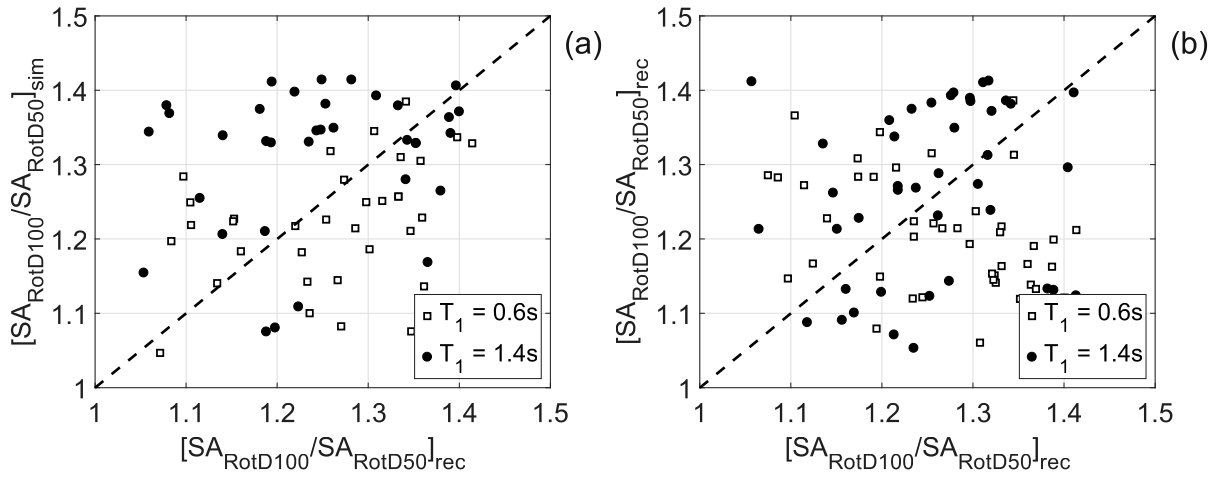
Figure 6. Comparison of ratio of maximum-to-medium response across orientations for recorded and simulated ground motions: Loma Prieta (a) and Northridge (b).
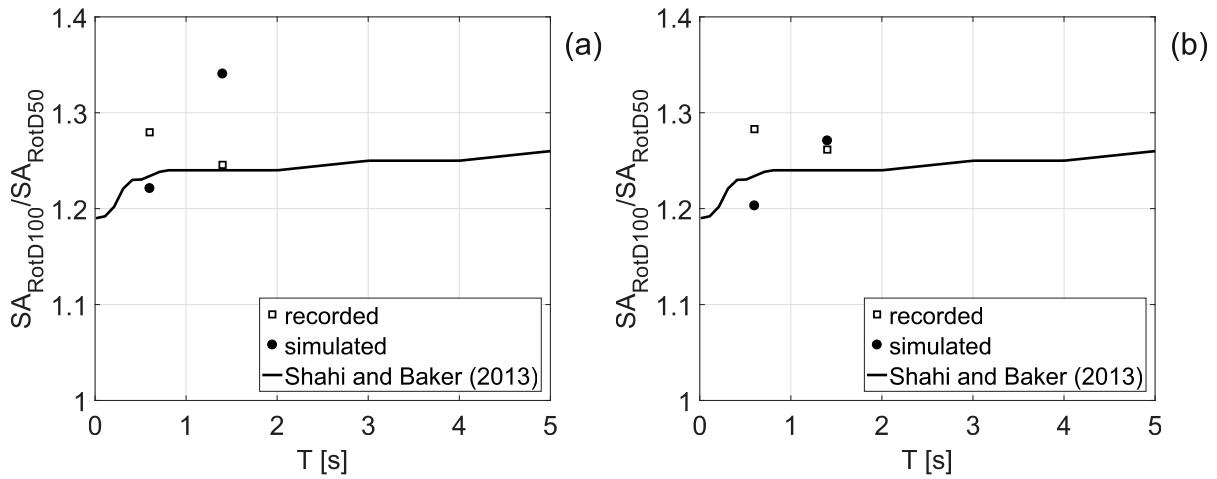

Figure 7. Comparison of empirical model for the median ratio of maximum-to-medium response across orientations and observed values for recorded and simulated ground motions: Loma Prieta (a) and Northridge (b).
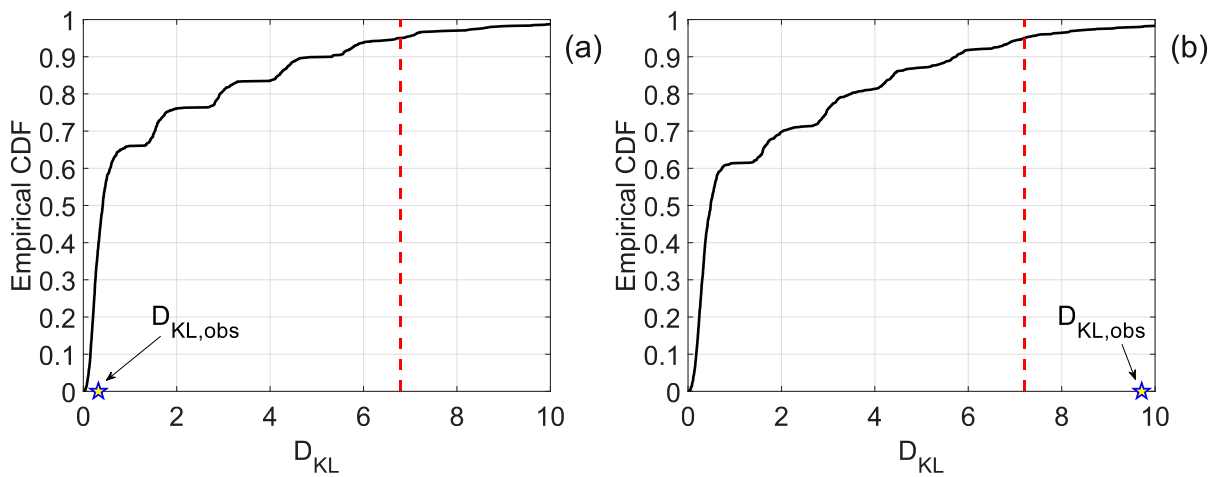

Figure 8. Results of comparison in terms of D_{KL} for Loma Prieta: T = 0.6s - Bridge 1 (a) and T = 1.4s - Bridge 2 (b).
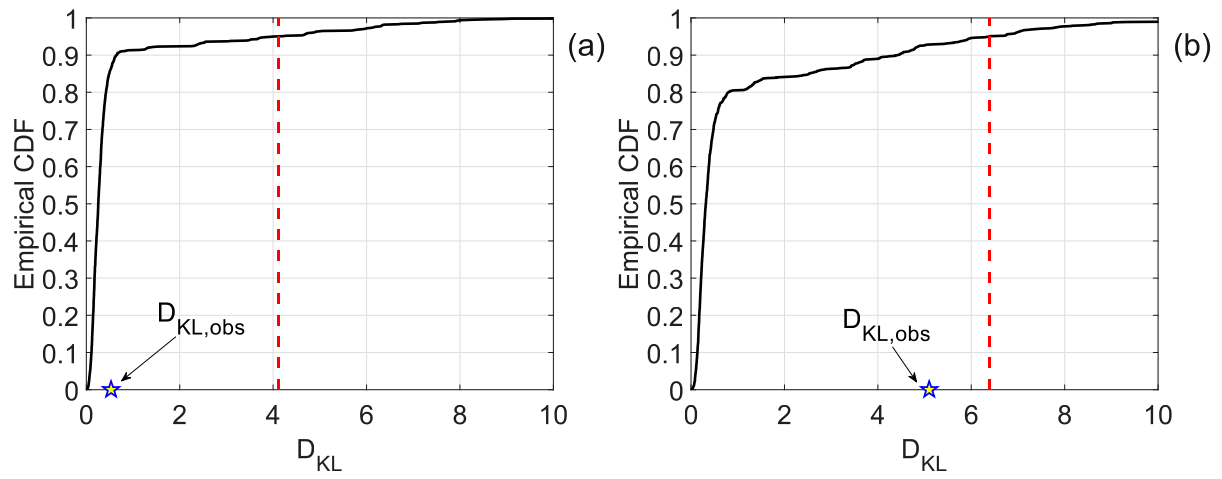
Figure 9. Results of comparison in terms of D$_{KL}$ for Northridge: T = 0.6s - Bridge 1 (a) and T = 1.4s - Bridge 2 (b).