Title

Validation of (not-historical) large-event near-fault ground-motion simulations for use in civil engineering applications

Journal

Earthquake Engineering & Structural Dynamics, 50(1)

ISSN

0098-8847

Authors

Petrone, Floriana

Abrahamson, Norman

McCallen, David

et al.

Publication Date

2021

DOI

10.1002/eqe.3366

Peer reviewed

Received: 16 April 2020 | Revised: 23 July 2020 | Accepted: 22 September 2020

DOI: 10.1002/eqe.3366

SPECIAL ISSUE PAPERS

WILEY

Validation of (not-historical) large-event near-fault ground-motion simulations for use in civil engineering applications

Floriana Petrone[1,2] | **Norman Abrahamson**[3] | **David McCallen**[1,2] | **Mamun Miah**[2]

[1] Department of Civil and Environmental Engineering, University of Nevada, Reno, Nevada

[2] Energy Geosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California

[3] Department of Civil and Environmental Engineering, University of California, Berkeley, California

Correspondence
Floriana Petrone, Department of Civil and Environmental Engineering, University of Nevada, Reno, Nevada.
Email: florianapetrone@unr.edu

Abstract

Ground-motion simulations generated from physics-based wave propagation models are gaining increasing interest in the engineering community for their potential to inform the performance-based design and assessment of infrastructure residing in active seismic areas. A key prerequisite before the ground-motion simulations can be used with confidence for application in engineering domains is their comprehensive and rigorous investigation and validation. This article provides a four-step methodology and acceptance criteria to assess the reliability of simulated ground motions of not historical events, which includes (1) the selection of a population of real records consistent with the simulated scenarios, (2) the comparison of the distribution of Intensity Measures (IMs) from the simulated records, real records, and Ground-Motion Prediction Equations (GMPEs), (3) the comparison of the distribution of simple proxies for building response, and (4) the comparison of the distribution of Engineering Demand Parameters (EDPs) for a realistic model of a structure. Specific focus is laid on near-field ground motions (<10km) from large earthquakes (M_w7), for which the database of real records for potential use in engineering applications is severely limited. The methodology is demonstrated through comparison of (2490) near-field synthetic records with 5 Hz resolution generated from the Pitarka et al (2019) kinematic rupture model with a population of (38) pulse-like near-field real records from multiple events and, when applicable, with NGA-W2 GMPEs. The proposed procedure provides an effective method for informing and advancing the science needed to generate realistic ground-motion simulations, and for building confidence in their use in engineering domains.

KEYWORDS

acceptance criteria, engineering applications, ground-motion simulation, nonlinear tall building response, statistical analysis, validation

1 | INTRODUCTION

Current approaches to earthquake engineering rely on the combined use of empirically calibrated ergodic ground-motion models (ie, GMPEs) to determine the shaking intensity for a specific site and appropriately scaled records from past events

to perform nonlinear dynamic analyses of engineering systems. Although well consolidated, this methodology has significant limitations for conditions without significant observation evidence[1] such as large magnitude ($M_w > 6.5$), near-field events on sedimentary basin sites with complex geology.

As anticipated by Aki,[2] advancement in High-Performance Computing (HPC) platforms and parallel computational ecosystems are creating the opportunity to address this limitation through the development of high-resolution physics-based models that can simulate ground motions for almost all conceivable scenarios (e.g., event magnitude, site conditions, geology structure, etc.) and provide site-specific input for infrastructure analysis at any location. ASCE/SEI 7–16 provisions[3] explicitly contemplate the possibility "to supplement the available records with simulation ground motions" when "the required number of recorded ground motions is not available," thus, substantiating the increasing relevance of ground-motion simulations in engineering domains.

Ground-motion simulation methods have evolved significantly over the past decades[4–7] to include techniques based on stochastic processes,[8–10] stochastic point source and finite fault,[11–13] and hybrid broadband approaches.[14–16] Parallel to the development and refinement of these methods, there have been many efforts to validate simulated ground motions to ensure realism of the simulations in multiple dimensions.[17–21] In general, the validations for specific application levels have been conducted through comparison of simulated ground motions from historical events with available real records from those events and through comparison of simulated ground motions with GMPEs for scenarios for which the GMPEs are reasonably well constrained. Although the empirical data for large magnitudes at short distances remain sparse, the extrapolation of the NGA-W2 GMPEs is constrained by the scaling from finite-fault simulations and from analytical site response studies. While the 3D simulations at a specific site are expected to differ from the GMPE estimates, the GMPEs provide a useful check for the average ground-motion amplitudes over a range of sites and distances for the M_w7 simulations evaluated in this study.

The first level of validation includes basic comparison of Intensity Measures (IMs) such as seismogram waveforms and response spectral parameters such as median linear response spectra, instantaneous spectral density, and good-of-fit measures.[22–24] More advanced validation methods comprise the analysis of simple proxies for building response, such as ground-motion duration and polarization, interperiod spectral correlation,[25] and the use of linear and nonlinear single-degree-of-freedom (SDOF) systems[26,27] to investigate and statistically compare the demand posed to structural systems as function of basic Engineering Demand Parameters (EDPs). The most application-specific validation methods entail the use of linear and nonlinear multiple-degree-of-freedom (MDOF) systems to examine the dynamic response of realistic structures subject to real and simulated records and statistically compare the distribution of the demand and the localization of damage at the structural level.[28]

Most of the previous simulation and validation efforts have focused on historical events,[24,27–31] for which available recordings (although in a very small number) allow a station-by-station inspection and comparison. While helpful to advance and refine simulation methodologies, results of these validation efforts cannot be generalized to the simulations of unknown (not historical) earthquake events, for which no recordings exist.[25] It would indeed be of high interest to be able to validate simulations of unknown events for performing risk-informed structural safety evaluations in urbanized areas where major earthquake events are expected.

With this aim, this work proposes a general methodology to evaluate the reliability of simulated ground motions from unknown (not-historical) earthquake events. The procedure is demonstrated for deterministic simulations up to 5 Hz for three M7 strike-slip earthquake scenarios that differ from each other for the location of the hypocenter only and are generated from the Pitarka et al.[32] kinematic rupture model.[33,34] The validation process includes four steps:

1. Selection of a population of real records collected from multiple events consistent with the simulated scenarios.
2. Comparison of the distribution of ground-motion IMs from the simulated records, real records, and GMPEs
3. Comparison of the distribution of simple proxies for building response.
4. Comparison of the distribution of EDPs for a realistic model of a structure.

For the evaluation of the effects of the simulated ground motions on realistic structure (step 4), the nonlinear dynamic response of a distributed plasticity Finite Element Model (FEM) of a 40-story building is analyzed and discussed on a statistical basis. The choice of a tall building is motivated by the interest of evaluating to what extent the discrepancies between real and simulated ground motions may affect the contribution of higher modes to the structural response and consequently influence the localization of damage. These are features that can be captured neither by SDOF systems nor by conceptual MDOF systems.

Overall, steps 2 though 4 represent three application-specific levels of ground-motion validation, which provide insight into ground-motion features pertaining to different levels of detail. IMs (step 2) provide understanding of the main features of the simulated motions and a first means for evaluating their realism. Simple proxies for building response (step 3) provide spectral correlation, interperiod correlation, and the ground-motion polarization parameters, that are relevant to the analysis of complex engineering systems and cannot be inferred from simple spectral analysis. Finally, the response of a realistic model of a structure (step 4) provides the most accurate illustration of how ground-motion features are reflected in the response of structural systems. Such step should be taken when simulated ground motions are used for design or assessment purposes, being either the design/assessment of a single structure or the definition of design/assessment procedures for structures located in the vicinity of a major active fault.

Acceptance criteria based on statistical estimates and, when applicable, on statistical and physical evaluations, are provided. Results of these analyses confirm that simulated ground motions can reliably capture features that have a significant effect on structural response, particularly in the bandwidth 0.2 –3 s.

2 | SELECTION OF REAL GROUND-MOTION DATASET—STEP 1

When interested in the use of ground-motion simulations of a non-historical earthquake for engineering applications, the validation should ideally be conducted against a sufficiently large population of real records having features statistically comparable with those of the simulated event (i.e., fault rupture mechanism, magnitude, distance from the fault rupture, site conditions, etc.). Such an approach, in fact, enables a detailed assessment of the structural response as obtained from the simulated and real records to address the question whether there are systematic dependencies of demand parameters on ground-motion model variables.[19]

In this study, the ground motions generated from three M7 strike-slip earthquake scenarios developed based on the rupture model of Graves and Pitarka[35] and Pitarka et al[2] are used to demonstrate the proposed validation procedure. The three scenarios differ from each other for the location of the hypocenter only,[33,34] have a geologic structure characterized by an average shallow shear-wave velocity of *Vs* 30 = 381 m/s and are resolved at frequencies up to 5 Hz. For each earthquake simulation, ground-motion time histories were stored at 1 km spacing intervals resulting in a total of 2490 near-field (within 10 km of the fault) simulated records.

The real records used as basis for comparison were obtained from the database of pulse-like motions developed by Baker and Shahi[36] and available at https://peer.berkeley.edu/Research/Transportation-Systems/Ground-Motion-Studies-Transporation-Systems.[37] The main features considered for the selection of the ground motions are summarized in Table 1 and include magnitude, rupture distance, and average shallow shear-wave velocity (*Vs* 30). A total of 38 records from seven events were selected, which are characterized by an average magnitude of 7.07, median distance of 2.88 Km, and median shallow shear-wave velocity of 365.7 m/s. Both the simulated records and real records were resolved in Fault Normal (FN) and Fault Parallel (FP) component, leading to a total of 2490*2 = 4980 simulated records and 38*2 = 76 real records utilized for the comparison of the structural response.

3 | GROUND-MOTION INTENSITY MEASURES (IMS)—STEP 2

3.1 | Pseudo-spectral acceleration across multiple periods

Elastic response spectra provide a concise representation of the peak response of linear SDOF systems and are largely used as IM for different purposes. As first step in the investigation of the simulated records, the pseudo-spectral acceleration across multiple periods is analyzed and compared with real records and ground-motion models from the literature.

The dataset of real records used for the validation of the simulated records collects ground motions from multiple events and multiple stations within the same event. Specifically, over a total of 7 events and 38 ground motions, there are extreme cases of events with one single record (ie, Landers 1992 and Morgan Hill 1994) and events with 16 records (ie, Chi-Chi, Taiwan 1999), see Table 1. When looking at ground-motion IMs (e.g., PGA, PGV, spectral acceleration, and velocity, etc.) in terms of median and logarithmic standard deviation, such a heterogeneity in the dataset, needs to be properly addressed, considering the broad variability of earthquakes, sites, and regions.

The approach used herein to homogenize the dataset follows the random-effects method proposed by Al Atik et al[38] for treating the correlation in the variability of the observed ground motions and accounting for the uneven sampling

TABLE 1 List and statistics of the real records considered in this study

No.	Earthquake name	Station name	Magnitude	Rupture distance (Km)	Vs 30 (m/s)
1	Imperial Valley-06 (1979)	EC County Center FF	6.53	7.31	192.1
2		EC Meloland Overpass FF		0.07	186.2
3		El Centro Array #4		7.05	208.9
4		El Centro Array #5		3.95	205.6
5		El Centro Array #6		1.35	203.2
6		El Centro Array #7		0.56	210.5
7		El Centro Array #8		3.86	206.1
8		El Centro Differential Array		5.09	202.3
9	Morgan Hill (1984)	Coyote Lake Dam (SW Abut)	6.19	0.53	597.1
10	Loma Prieta (1989)	Gilroy-Gavilan Coll.	6.93	9.96	729.7
11		LGPC		3.88	477.7
12	Landers (1992)	Lucerne	7.28	2.19	684.9
13	Northridge-01 (1994)	Jensen Filter Plant	6.69	5.43	373.1
14		Jensen Filter Plant Generator		5.43	525.8
15		Newhall-Fire Sta		5.92	269.1
16		Newhall-W Pico Canyon Rd.		5.48	285.9
17		Rinaldi Receiving Sta		6.50	282.3
18		Sylmar-Converter Sta		5.35	251.2
19		Sylmar-Converter Sta East		5.19	370.5
20		Sylmar-Olive View Med FF		5.30	440.5
21	Kobe, Japan (1995)	KJMA	6.90	0.96	312.0
22		Takarazuka		0.27	312.0
23	Chi-Chi, Taiwan (1999)	CHY028	7.62	3.14	542.6
24		CHY101		9.96	258.9
25		TCU049		3.78	487.3
26		TCU052		0.66	579.1
27		TCU053		5.97	454.6
28		TCU054		5.30	460.7
29		TCU068		0.32	487.3
30		TCU075		0.91	573.0
31		TCU076		2.76	615.0
32		TCU082		5.18	472.8
33		TCU087		7.00	473.9
34		TCU101		2.13	272.6
35		TCU102		1.51	714.3
36		TCU103		6.10	494.1
37		TCU122		9.35	475.5
38		WGK		9.96	258.9
		Median	7.07 (mean)	2.88	365.7
		Standard Deviation (ln units)	0.50 (linear)	1.16	0.42

of the earthquakes. According to this approach, the overall variability of a ground-motion IM (e.g., PGA, PGV, spectral acceleration Sa, etc.) is partitioned into between-events variability and within-event variability, which are zero-mean, independent, normally distributed random variables with standard deviation τ and ϕ, respectively. The between-events variability provides a measure of the shift between the median of observed ground motions from a single event and the median of the entire population of ground motions, whereas the within-event variability provides a measure of the misfit

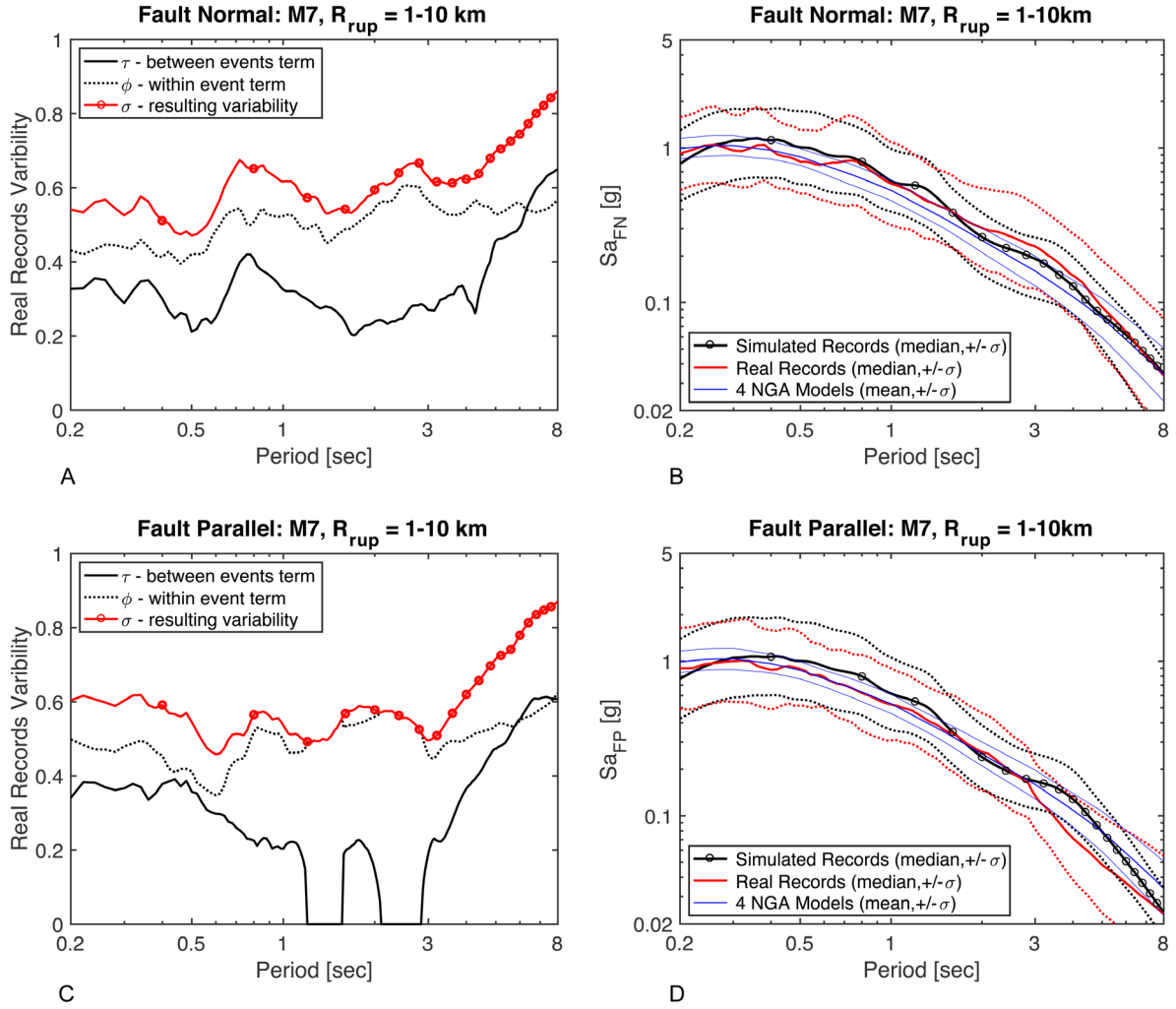


FIGURE 1 Real records variability: between-events (τ), within-event (ϕ), and resulting (σ) standard deviations for (A) the Fault Normal and (C) Fault Parallel components. Comparison of median acceleration spectra $\exp(\mu_{lnSa})$and standard deviation σ_{lnSa}across simulated records, real records and 4 NGA ground-motion models[41-44] for the (B) Fault Normal and (D) Fault Parallel

between an individual observation from a specific event and the median of that event. Since the between-events and within-event residuals are uncorrelated, the global standard deviation (σ) of a ground-motion IM is computed as $\sigma = \sqrt{\tau^2 + \phi^2}$.

In this study, this approach is adopted for the calculation of the mean and standard deviation of the natural logarithm of the pseudo-spectral acceleration $\ln(Sa)$ of the FN and FP components of the real records (referred to as weighted), as the two components will be used separately for the building analyses.

Figure 1A and C show the variation of τ, ϕ, and the resulting σ as function of the spectral periods ranging from 0.2 to 8 s for the FN and FP component, respectively. The zero values for τ for the FP component at periods between 1- and 3-s result from the small number of events for which the average ground motions happen to be similar in this period range. The higher values of the variability observed between 5 and 8 s are attributed to the inclusion of records characterized by particularly high values of the pseudo-spectral acceleration at longer periods. A separate analysis conducted by the authors showed in fact that Coyote Lake Dam (Morgan Hill, 1984) and Gilroy-Gavilan (Loma Prieta, 1989) ground motions are characterized by low pseudospectral acceleration at long periods (lower by a factor of ~3 compared to the median of the population of real records), and TCU052 and TCU054 (Chi-Chi Taiwan, 1999) records are characterized by high pseudospectral acceleration at long period (higher by a factor of ~6 compared to the median of the population of real records), see also (Ref.39 and 40).

Figure 1B and D show the comparison of the median pseudo-spectral acceleration $\exp(\mu_{lnSa})$ and standard deviation σ_{lnSa} with 5% damping of the FN and FP ground-motion components, respectively, across simulated records, weighted real

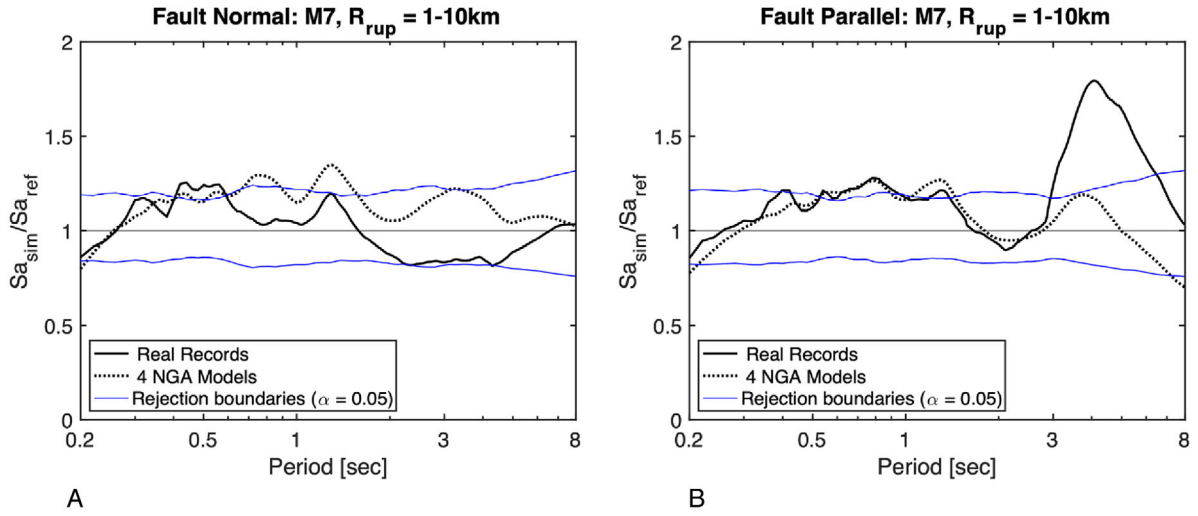


FIGURE 2 Ratio of the median pseudo-spectral acceleration of simulated ground motions (Sa_{sim}) to the median pseudo-spectral acceleration of real records and 4 NGA models (Sa_{ref}), and rejection boundaries for (A) Fault Normal and (B) Fault Parallel component

records, and the mean of four NGA ground-motion models (4 NGA GMPEs).[41–44] For the simulated records, the median and standard deviation are obtained directly from the 2490 ground motions (for each component), whereas for the real records, median, and standard deviation are derived following the random-effects method referenced above to account for the uneven sampling of the earthquakes. For the 4 NGA GMPEs, the standard deviation has been adapted to include the variability associated with a population of records resolved in the two horizontal components, the epistemic uncertainty deriving from the model-to-model differences, and an additional epistemic uncertainty that depends on magnitude, style-of-faulting, and spectral period as proposed by Al Atik and Youngs.[45] As a result, the standard deviation of the median of the observations considering the uncertainty in the median of four GMPEs is computed as:

$$\sigma_{lnSa(GMPEs)} = \sqrt{\frac{\sigma^2_{\ln\mu(RR)}}{n_{rec}} + \sigma^2_{ep} + \sigma^2_{add}} \tag{1}$$

where $\sigma_{\ln\mu(RR)}/\sqrt{n_{rec}}$ is the standard error of the mean of the natural logarithm of the pseudospectral acceleration as obtained from the original database of 38 real records (with no weighting), being $\sigma_{\ln\mu(RR)}$ the standard deviation of the natural logarithm of the real records pseudo-spectral acceleration evaluated separately for the two components, and n_{rec} the number of real records σ_{ep} is the model-to-model epistemic uncertainty calculated assuming that the predictions from the ground-motion models are normally distributed, and σ_{add} is the uncertainty in the median prediction of each model,[45] calculated as:

$$\sigma_{add} = \begin{cases} 0.072 \; for \; T < 1.0 \; s \\ \sigma_{add}\,(T<1) + 0.0217 \cdot \ln(T) \; for \; T \geq 1.0 \; s \end{cases} \tag{2}$$

A simple visual inspection of the pseudo-spectral acceleration indicates a general good agreement in terms of spectral shape and spectral amplitude across simulated records, real records, and ground-motion models for a bandwidth of periods ranging from 0.2 to 8 s. However, to better interpret and quantify how the spectra and their dispersion compare, the ratio of median pseudo-spectral acceleration (μ_{lnSa}) and standard deviation (σ_{lnSa}) is analyzed in the following and hypothesis testing is performed to assess the significance of the observed differences.

Figure 2A and B show the ratio of the median pseudo-spectral acceleration of simulated ground motions (referred to as Sa_{sim} in the plot notation) to the median pseudo-spectral acceleration of the weighted real records and the 4 NGA models (referred to as Sa_{ref} in the plot notation) as function of the spectral period for the (A) FN and (B) FP component, respectively. A ratio above unity indicates an overestimation of the median spectral values predicted by the simulation, whereas a ratio below unity denotes an underestimation. The plots also show the rejection boundaries of the hypothesis testing, here proposed as acceptance criteria, performed for a significance level $\alpha = 0.05$ (thin blue lines), that is,

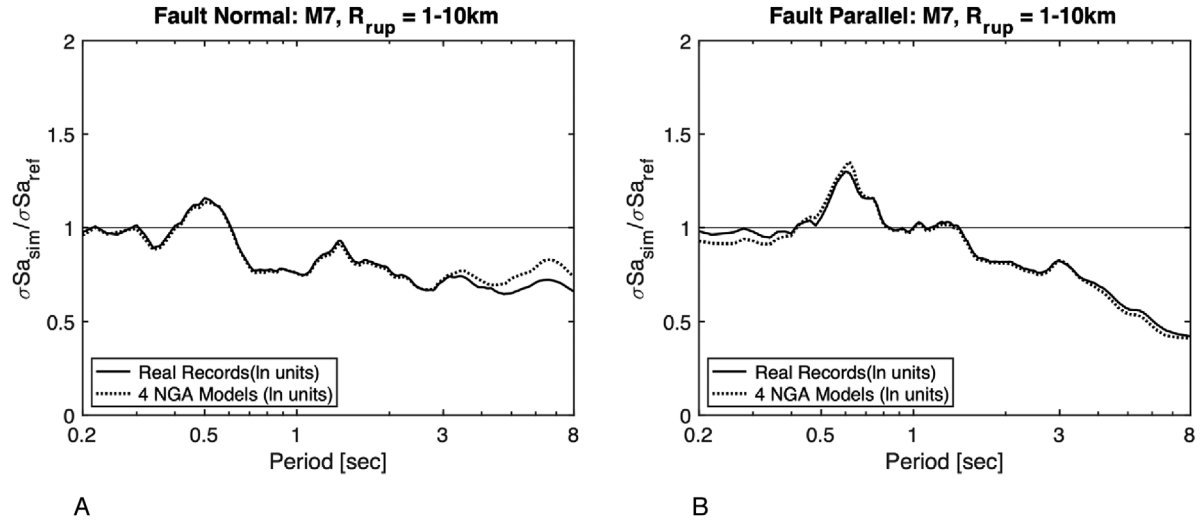


FIGURE 3 Ratio of the pseudo-spectral acceleration standard deviation (ln units) of simulated ground motions (σSa_{sim}) to the pseudospectral acceleration standard deviation (ln units) of real records and 4 NGA models (σSa_{ref}) for (A) Fault Normal and (B) Fault Parallel component

$\mu_{lnSa,rec} \pm 1.96\sqrt{\frac{\sigma^2_{lnSa,rec}}{n_{rec}} + \frac{\sigma^2_{lnSa,sim}}{n_{sim}}}$, where $\mu_{lnSa,rec}$ and $\sigma_{lnSa,rec}$ are the median pseudo-spectral acceleration and standard deviation of the (weighted) real records, respectively, $\mu_{lnSa,sim}$ and $\sigma_{lnSa,sim}$ are the median pseudo-spectral acceleration and standard deviation of the simulated records, respectively, and n_{sim} is the number of simulated records. A detailed explanation of how the hypothesis testing is conducted is presented in step 4. Such boundaries limit a region within which the differences between simulated records and weighted real records (the 4 NGA models are not included in the definition of the rejection boundaries) are attributed to the randomness associated with the finite size of the samples and not to inherent differences in the ground motions. This test is a pure statistics-based check performed at this stage to gain confidence in the significance of the datasets used in this study and does not include any physics-based consideration. A more comprehensive evaluation of the realism of the simulated records, derived on a statistical and physical basis, is provided at the end of this section.

From plots (A) and (B) it is seen that, for both the FN and FP component, the simulated ground motions tend to overestimate the spectral acceleration at shorter periods (from ~0.2 to ~2 s) with respect to both the real records and the 4 NGA models, with a model bias ranging between 10 and 25%, but with the simulation prediction falling mainly within the rejection boundaries, which is partially attributed to the characteristics of the source model. At longer periods, and for the FN component, a different trend is observed. An underestimate of ~20% is seen with respect to the weighted real records (but still within the rejection boundaries) and an overestimate of about 20% with respect to the 4 NGA models. For the FP component, the bias becomes more pronounced at longer periods, with a peak difference of ~75% observed in the comparison against the real records and ~20% in the comparison against the 4 NGA models, which is indicative of higher energy content at longer periods. It is also seen that in the bandwidth 3 to 8 s, the ratio of simulated records to real records falls outside the rejection boundaries. As it will be discussed later in the article, such differences have significant impact on the nonlinear response of tall structures and are reflected in the results obtained from the 40-story building simulations.

Figure 3A and B show the ratio of the standard deviation of the median pseudo-spectral acceleration (in ln units-σ_{lnSa}) of the simulated ground motions (referred to as σSa_{sim} for simplicity in the plot notation) to the standard deviation of the median pseudo-spectral acceleration of the weighted real records and the 4 NGA models, (referred to as σSa_{ref} for simplicity in the plot notation) as function of the spectral period for the FN and FP component, respectively. For both components, the simulated ground motions consistently show a smaller variability across all spectral periods with values that departure from the unit up to ~60% at longer periods, except for a bandwidth ranging between ~0.4 and 0.8 s where the standard deviation of the simulated records is greater than what predicted by the real records and 4 NGA models. This trend of lower variability in the simulations is attributed to the real records and 4 NGA models sampling a broader range of velocity structures and slip distributions than are included in this set of simulations. Recall, in fact, that the simulated

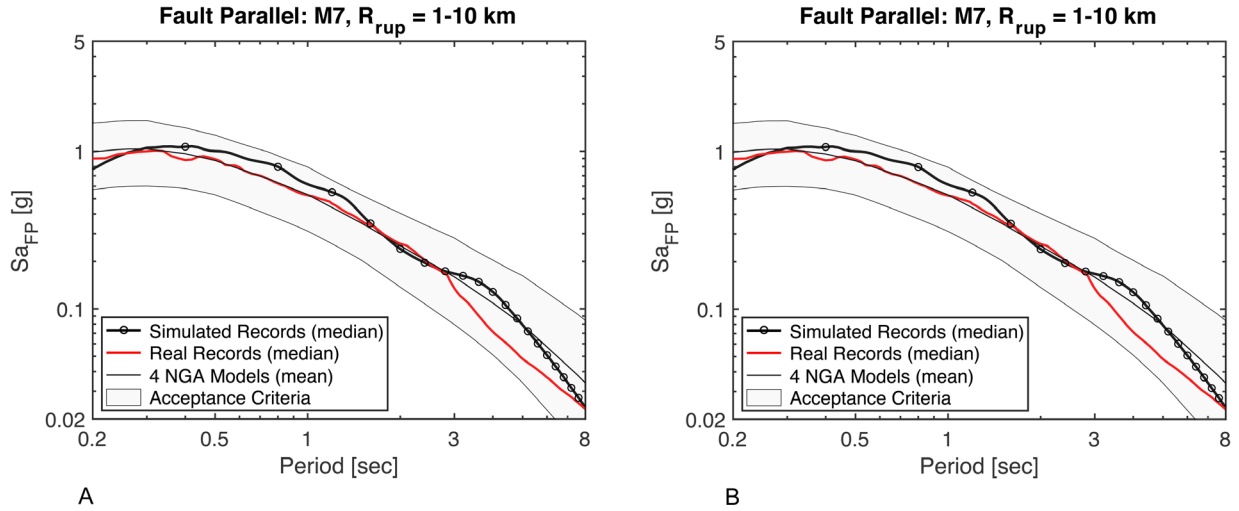


FIGURE 4 Median pseudo-spectral acceleration of simulated records and real records, and mean of 4 NGA models with acceptance criteria for (A) Fault Normal and (B) Fault Parallel component

motions are generated from three different scenarios that adopt the same geology structure, fault-rupture geometry and magnitude, and differ for the location of the hypocenter only.

To develop some judgment and quantitatively assess the ability of the simulated records to produce spectral accelerations consistent with real records and ground-motion models on a statistical and physical basis, the validation methodology proposed by Goulet et al[46] and Dreger et al[21] is here employed with some variants. The methodology defines acceptance criteria consisting in the definition of an upper boundary and a lower boundary for the median pseudo-spectral acceleration across multiple periods. In this study, the boundaries are defined by taking the largest absolute values of positive and negative differences of the individual 4 NGA ground-motion models median from the overall median across all periods, to which the total epistemic uncertainties are added with 95% confidence interval, that is $+1.96\sqrt{\sigma_{ep}^2 + \sigma_{add}^2}$(where σ_{ep} and σ_{add} are calculated as explained above). The shift in the spectral amplitude is applied across all periods, with differences in the effective range of the acceptance criteria deriving from the total epistemic uncertainties at each period. It should be also noted that in the original methodology the acceptance criteria were limited to 3 s, as the ground-motion models are not well constrained by real ground-motion data for longer spectral periods, whereas in this study, the acceptance criteria have been extended to 8 s. The passing criterion is met if the simulation median lies within the defined range. A departure of the simulation median from the acceptance bounds is seen as an indicator that the simulated motions are not consistent with the NGA models' predictions and a sign of potential issues with the simulation.

Figure 4A and B show the mean of the 4 NGA models with the acceptance criteria (shaded area) along with the median pseudo-spectral acceleration of simulated records and weighted real records for the FN and FP component, respectively. It is observed that the median spectrum from the simulated records falls within the acceptance boundaries over the entire bandwidth of spectral periods, with remarkable agreement observed at longer periods (> 4 s) for the FN component and at intermediate periods (2 to 3 s) for the FP component. The same plots also show the median spectrum from the weighted real records, which provides an indirect measure of the consistency between the selected population of real records and the 4 NGA models and corresponding acceptance criteria, supporting the use of such motions as dataset for the structural response analysis.

Evidence from these comparisons gives confidence in the realism of the simulated ground motions in terms spectral shape and amplitude. However, for use in engineering applications it is necessary to take a step forward and investigate additional parameters that are relevant to the response of complex structural systems whose nonlinear response is influenced by the contribution of higher modes. These additional parameters include spectral correlation, interperiod correlation, and ground-motion polarization, which will be addressed in the following sections.

4 | SIMPLE PROXIES FOR BUILDING RESPONSE—STEP 3

4.1 | Spectral correlation parameters and interperiod correlation

The IMs traditionally used for quantifying the strength of an earthquake and ultimately predicting the structural response are the peak ground acceleration (PGA) and the spectral acceleration at the fundamental period of the structure ($Sa(T1)$).[47] However, for structures whose response is expected to be strongly influenced by the contribution of higher modes, it is fundamental to evaluate the spectral amplitudes across different periods including the effective periods attained when the structure experiences nonlinearities. From the perspective of ground-motion simulation validation, this translates into the necessity of investigating the spectral correlation at multiple periods.

In this study, the normalized residual epsilon (ε), introduced by Baker and Cornell,[48,49] is employed to inspect the spectral shape of the simulated ground motions compared to the real records. For comparing the spectral shapes, ε_B (where the subscript B stands for "Bias") is computed relative to the median from the real records and is defined as:

$$\varepsilon_{Bi} = \frac{ln[Sa_i(T)_{sim}] - \mu_{lnSa(T)_{rec}}}{\sigma_{lnSa(T)_{sim}}} \tag{3}$$

where the subscript i refers to the i^{th} simulated record, $ln[Sa_i(T)_{sim}]$ is the natural logarithm of the pseudo-spectral acceleration of the i^{th} simulated record at period T, $\mu_{lnSa(T)_{rec}}$ is the mean of the natural logarithm of the real records pseudo-spectral acceleration at period T, and $\sigma_{lnSa(T)_{sim}}$ is the logarithmic total standard deviation of the simulated records at period T, computed accounting for both the between- and within-event variability, as discussed in the section 3.1. The ε_B parameter provides a measure of the difference (residuals) between the spectral acceleration of a simulated record and the mean logarithmic spectral acceleration of the (weighted) real records, expressed as number of simulated records standard deviations. Differently from what is usually done in literature, in this context ε_B is evaluated with respect to the real records and not to the ground-motion models. This is motivated by the objective of drawing a correlation between what is inferred from the analyses of spectral correlation parameters and what is observed in the nonlinear response of a tall building subject to both suites of simulated records and real records (as it will be discusses in the following sections of this article).

The interperiod correlation of epsilon, $\rho_{\varepsilon(T_1),\varepsilon(T_2)}$, has been demonstrated to be an essential feature of ground motions for capturing the variability of the structural response.[50] The interperiod correlation coefficient of ε_B is calculated with the Pearson's equation, as:

$$\rho_{\varepsilon_B(T_1),\varepsilon_B(T_2)} = \frac{\sum_{i=1}^{n}\left[\varepsilon_{B,i}(T_1) - \overline{\varepsilon_B(T_1)}\right]\left[\varepsilon_{B,i}(T_2) - \overline{\varepsilon_B(T_2)}\right]}{\sqrt{\sum_{i=1}^{n}\left[\varepsilon_{B,i}(T_1) - \overline{\varepsilon_B(T_1)}\right]^2 \sum_{i=1}^{n}\left[\varepsilon_{B,i}(T_2) - \overline{\varepsilon_B(T_2)}\right]^2}} \tag{4}$$

where $\varepsilon_{B,i}(T_1)$ is epsilon of the i^{th} simulated ground motion at period T_1, $\overline{\varepsilon_B(T_1)}$ is the average epsilon (across all simulated motions) at period T_1, and $\varepsilon_{B,i}(T_2)$ and $\overline{\varepsilon_B(T_2)}$ are the same quantities evaluated at period T_2, being T_1 and T_2 the spectral periods at which the correlation is defined.

Figure 5A shows the variation of the average ε_B (across 2490 simulated records) and the corresponding standard deviation centered around the mean for spectral periods ranging from 0.2 and 8 s, and for the FN and FP component. These curves, also known as "goodness-of-fit" plots,[46] provide a different way to look at the comparison between real records and simulated records, as shown in Figure 2A and B. Positive values of ε_B indicate that the average simulated ground motions at a specific period are stronger than the average real records and, likewise, negative values of ε_B imply that the average simulated ground motions at a specific period are weaker than the average real records. It is observed that ε_B is consistently around zero in natural logarithm units (corresponding to a difference $\ll \sigma_{lnSa(T)_{sim}}$) for the spectral periods ranging from 0.2 to ~3 s and attains a value around the unit in natural logarithm units (corresponding to a difference of $\sim\sigma_{\mathrm{lnSa(T)Sim}}$) around 4 s and for the FP component only, which is indicative of spectral amplitudes at longer periods debatably higher compared to the real records, see also Figure 2B. In general, the values of ε_B at neighboring periods are correlated, meaning that if the simulated ground motions are stronger than the real records in the average at spectral period T, they are likely to be stronger than expected at $T \pm \Delta T$; whereas for widely spaced periods the values of ε_B are

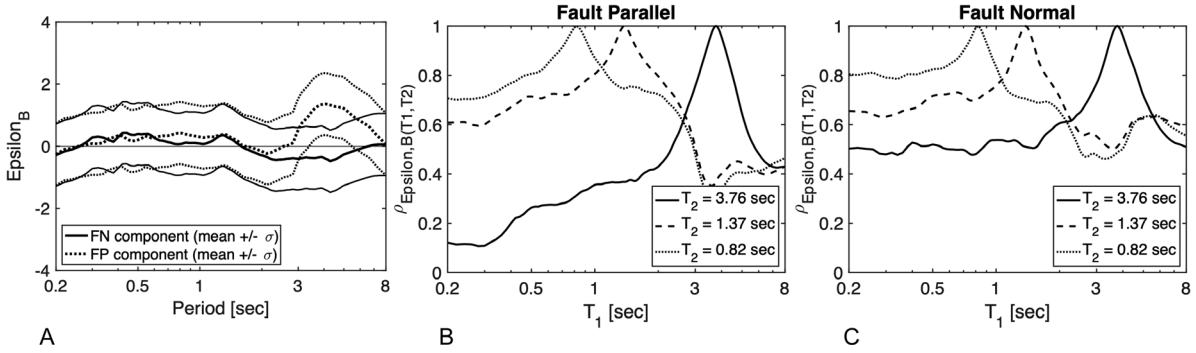


FIGURE 5 (A) Variation of ε_B as function of the period averaged over all (2490) simulated records. Interperiod correlation of ε_B for (B) Fault Parallel and (C) Fault Normal component

likely to be weakly correlated. A physical interpretation of the interperiod correlation of ε_B suggests that high values of $\rho_{\varepsilon_B(T_1),\varepsilon_B(T_2)}$ (close to one) for neighboring periods (eg, $T_1 \sim T_2$) are representative of smooth pseudo-acceleration spectra, whereas low values of $\rho_{\varepsilon_B(T_1),\varepsilon_B(T_2)}$ (close to zero) are representative of pseudo-acceleration spectra with pronounced (possibly unrealistic) local spectral peaks (or trough) at T_2.

Figure 5B and C show how $\rho_{\varepsilon_B(T_1),\varepsilon_B(T_2)}$ varies across spectral periods ranging between 0.2 and 8 s, for three conditioning periods ($T_2 = 3.76$, 1.37, and 0.82 s, corresponding to the natural vibration periods of the first three modes of the 40-story building that will be analyzed in the following sections), for the FN and FP component, respectively. Starting from the FP component, it is observed that at $T_2 = 3.76$ s (first mode) the interperiod correlation rapidly decays at neighboring periods, with values that drop to ~0.35 at $T_1 = 1.37$ s (second mode) and ~0.25 at $T_1 = 0.82$ s (third mode). Since at $T = 3.76$ s, ε_B attains a high-positive value (corresponding to a peak in the pseudoacceleration spectra), see Figure 5A, the lower values of ε_B at the nearby periods will affect the response of the structure by a small factor. This is expected to be reflected in the response of the 40-story building for the FP component, with a behavior influenced by the first mode and modest contribution from higher modes, with a tendency of the simulated records to pose a higher demand to the structure, compared to the real records. It is also seen that at longer periods ($T > 3.76$ s) ε_B is still positive and the interperiod correlation coefficient fluctuates around 0.6. This is relevant to the nonlinear response of the 40-story building, when the fundamental period of vibration elongates as the result of softening and reduced effective stiffness, and indicates that the structure is going to soften into a peak of the spectrum. At shorter conditioning periods, ε_B is still positive (although closer to zero) and the correlation appears to be stronger, with values of $\rho_{\varepsilon_B(T_1),\varepsilon_B(T_2)}$ ranging between 0.6 and 0.7. As a result, the response of stiffer structures is expected to be strongly influenced by higher modes contribution by a larger factor relative to the uncorrelated case.

As for the FN component, it is observed that the interperiod correlation is quite strong over the entire range of periods, with minimum values of $\rho_{\varepsilon_B(T_1),\varepsilon_B(T_2)}$ fluctuating around 0.5 for the conditioning period $T_2 = 3.76$ s. However, since at $T = 3.76$ s ε_B attains a negative value (corresponding to a trough in the spectra), see Figure 5A, and at shorter nearby periods ($T_1 = 1.37$ s, $T_1 = 0.82$ s) ε_B has positive values (corresponding to a peak in the spectra), the response of the 40-story building in the FN component is expected to be strongly influenced by higher modes with a tendency of the simulated records to underestimate the contribution of the first natural vibration mode.

These analyses provide insight into fundamental features of the simulated ground motions that affect the variability of the linear and nonlinear response of complex structural systems. Overall, results from the analyses show that the values of the inter-period correlation of ε_B fall in a range of realistic values that should ensure variability in structural response.

For this application, the epsilon is also computed separately for real and simulated records and relative to the median of the real and simulated records, respectively, to avoid differences between the median from the real records and the simulations affecting the correlation:

$$\text{a) } \varepsilon_{i,rec} = \frac{ln[Sa_i(T)_{rec}] - \mu_{lnSa(T)_{rec}}}{\sigma_{lnSa(T)_{rec}}} \qquad \text{and} \qquad \text{b) } \varepsilon_{i,sim} = \frac{ln[Sa_i(T)_{sim}] - \mu_{lnSa(T)_{sim}}}{\sigma_{lnSa(T)_{sim}}} \tag{5}$$

Figure 6A through F show the variation of the correlation coefficient (i.e., Pearson's coefficient), for spectral periods ranging from 0.2 to 8 s and for three conditioning periods ($T_2 = 3.76$, 1.37, and 0.82 s) for both the FN and the FP component.

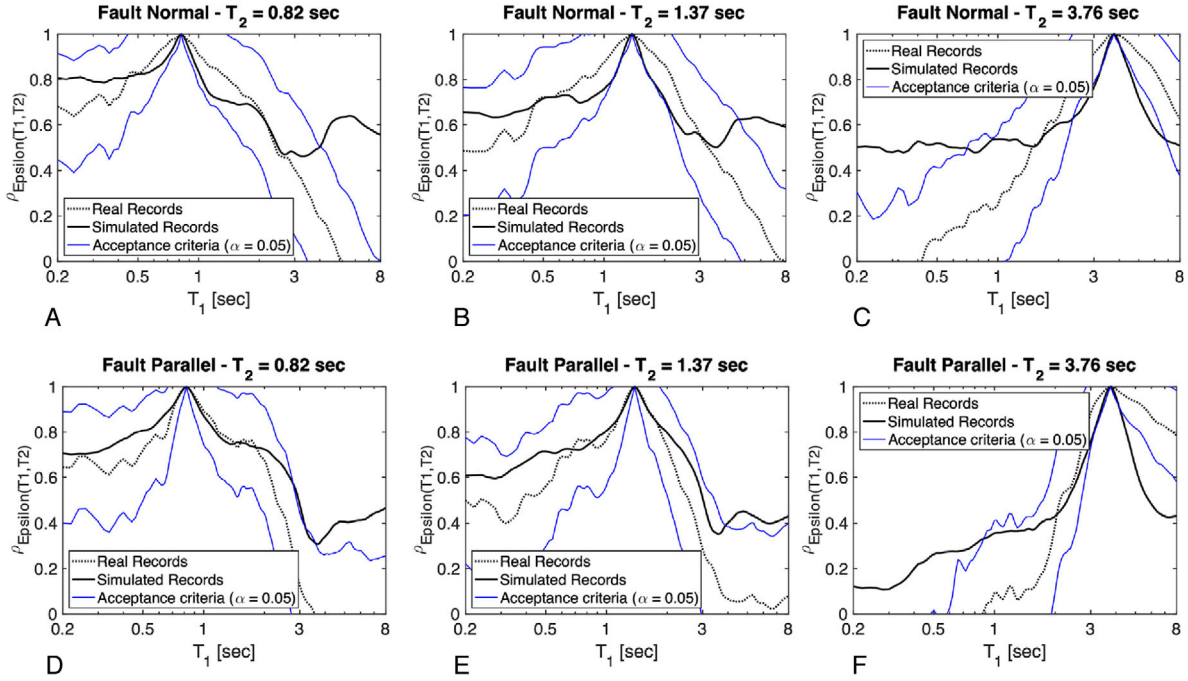


FIGURE 6 Interperiod correlation of $\varepsilon_{i,rec}$ and $\varepsilon_{i,sim}$ with corresponding acceptance criteria for three conditioning periods (A-D), $T_2 = 0.82$ s, (B-E), $T_2 = 1.37$ s and (C-F), $T_2 = 3.76$ s for both the Fault Normal (top row) and Fault Parallel (bottom row) component

The dotted black lines represent the correlation coefficients obtained with the real records, see Equations (4) and (5a), (i.e., reference values), the solid black lines represent the coefficients from the simulated records, see Equation (4) and (5b), and the thin blue lines provide the proposed acceptance boundaries derived considering a significance level $\alpha = 0.05$, that is $\rho_{rec(T)} \pm 1.96\, \sigma_{\rho(T)}$, where $\rho_{rec(T)}$ is the correlation coefficient derived from the real records at period T, and $\sigma_{\rho(T)} = \sqrt{\frac{\sigma^2_{\rho,rec(T)}}{n_{rec}} + \frac{\sigma^2_{\rho,sim(T)}}{n_{sim}}}$ is the resulting standard error at the same period T, being $\sigma^2_{\rho,rec(T)} = 1 - \rho^2_{rec(T)}$ and $\sigma^2_{\rho,sim(T)} = 1 - \rho^2_{sim(T)}$. Correlation coefficients of the simulated motions falling outside these acceptance boundaries are seen as indicator of a poor agreement between the interperiod correlation of real and simulated records.

In this study, a general consensus of the correlation is observed for lower conditioning periods and for both the FN and FP components. However, discrepancies are seen at longer spectral periods ($T_1 \geq 3$ s), for long conditioning periods T_2 and for the FP component, where the simulated records fail in satisfying the acceptance criteria. This is partially attributed to permanent ground displacements that are included in the simulated records, but not in the considered real records and it is reflected in the structural response of the 40-story building, as shown in step 4.

4.2 | Ground-motion polarization

A second proxy for the building response is the ground-motion polarization. In multi-component ground motions, the spectral acceleration varies depending on the orientation of interest. A measure of this variation is provided by the parameters Sa_{RotD50} and $Sa_{RotD100}$. For a given spectral period, Sa_{RotD50} and $Sa_{RotD100}$ represent the median and maximum value, respectively, of the pseudo-spectral acceleration of the resultant of the two horizontal components of the ground motions (i.e., FN and FP) as computed over each degree of rotation from 1° to 180°.[51] That is:

$$Sa_{RotD50}(T) = \underset{\theta}{\text{median}}\, Sa(T,\theta) \tag{6}$$

and

$$Sa_{RotD100}(T) = \underset{\theta}{\max} \text{Sa}(T,\theta) \tag{7}$$

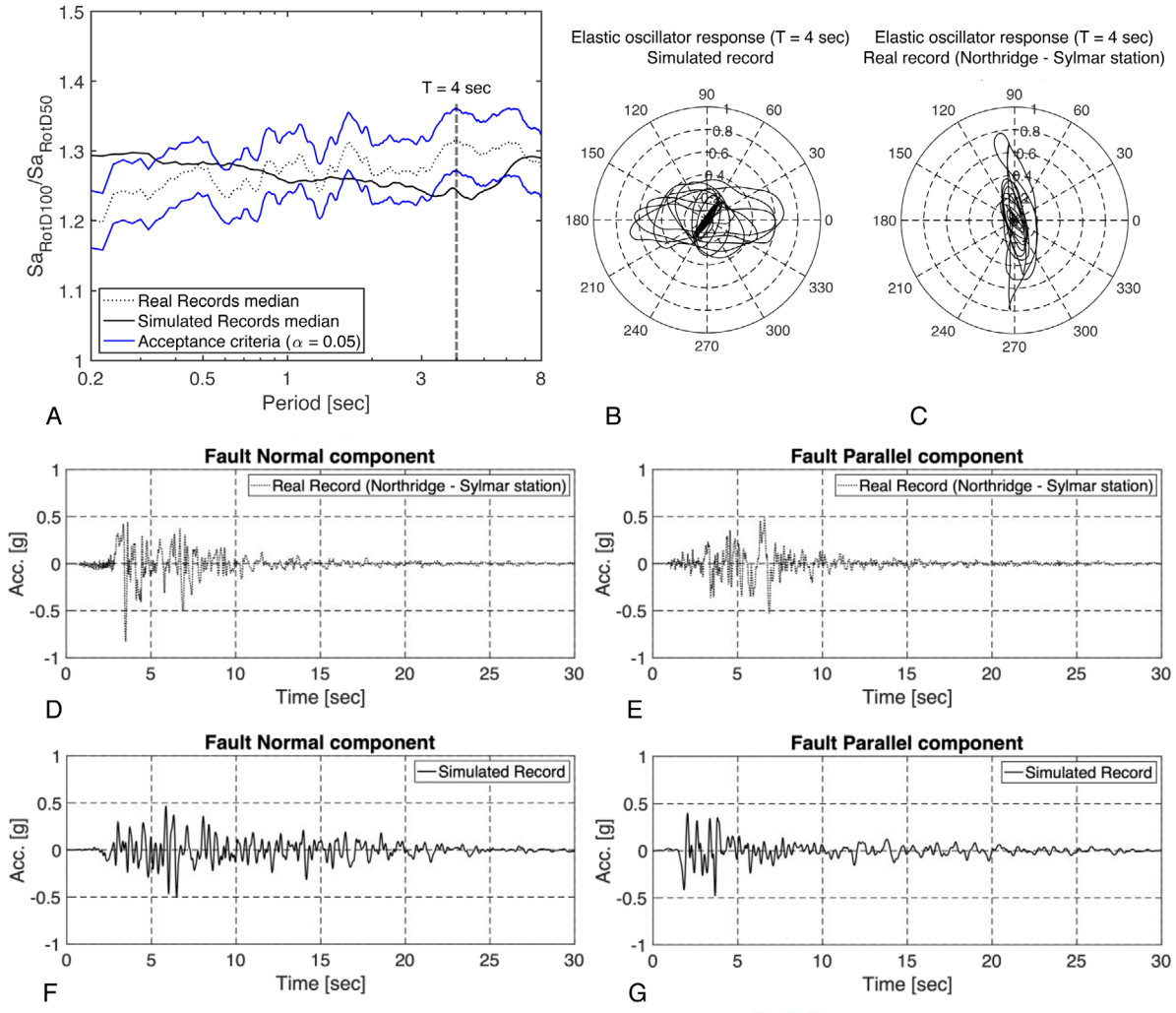


FIGURE 7 (A) Median of the ratio $Sa_{RotD100}/Sa_{RotD50}$ from real and simulated records and acceptance criteria. (B) Displacement trace for an elastic oscillator with $T_1 = 4$ *s*, computed using the simulated record. (C) Displacement trace for an elastic oscillator with $T_1 = 4$ *s*, computed using Northridge 20-Sylmar station record (D-G). Real and simulated records: acceleration time histories for Fault Normal (left) and Fault Parallel (right) component

where T is the spectral period and $1° \leq \theta \leq 180°$ is the orientation angle. The ratio $Sa_{RotD100}/Sa_{RotD50}$can vary between 1 and $\sqrt{2} = 1.41$ and is used as an indicator of the polarization of the ground motions.[25] Low values of this ratio are indicative of low polarized ground motions that lead to a response of three-dimensional systems equal in all directions. Conversely, high values of the ratio are indicative of polarized ground motions that lead to a structural response polarized in one direction.

Figure 7A shows the variation of the ratio $Sa_{RotD100}/Sa_{RotD50}$ as function of the spectral periods for the real records (dotted line) and the simulated records (solid line), where $Sa_{RotD100}$ is the mean of the natural logarithm of the maximum pseudo-spectral acceleration calculated across the 38 real records and 2490 simulated records , and Sa_{RotD50} is the mean of the natural logarithm of the median pseudo-spectral acceleration calculated across the 38 real records and 2490 simulated records. The plot also shows the acceptance boundaries derived based on the hypothesis testing for a significance level $\alpha = 0.05$, consistently with what done in step 2 and assuming that $Sa_{RotD100}$ and Sa_{RotD50} are correlated random variables. It is observed that the polarity of the simulated motions is within the acceptance boundaries in most of the 0.2 to 8 s bandwidth, with discrepancies seen at shorter periods (0.2-0.5 s), which may be attributed to a not well constrained geology structure of the geophysics model, and at longer periods. To better interpret the impact that this feature has on the structural response, Figure 7B and C show the displacement trace for an elastic oscillator with $T_1 = 4$ s in all orientations, computed applying simultaneously the two components (FN and FP) of the Northridge ground motion (Sylmar, Olive View Med FF station, see Table 1) and the two components (FN and FP) of one simulated record, respectively, which

TABLE 2 A 40-story building: Modal periods and frequencies

	*T*1		*T*2		*T*3	
	[s]	[Hz]	[s]	[Hz]	[s]	[Hz]
40-story building	3.762	0.266	1.366	0.732	0.815	1.227

are characterized by the same pseudospectral acceleration at $T = 4$ s. It is observed that the difference of about 5% in the ratio $Sa_{RotD100}/Sa_{RotD50}$ at spectral period of 4 s is strongly reflected in the response of the oscillators. The response of the oscillator excited by the simulated record, for which $Sa_{RotD100}/Sa_{RotD50}$ ~1.24, is far less polarized than the response of the oscillator to the real record, for which $Sa_{RotD100}/Sa_{RotD50}$ ~1.31. This demonstrates that the polarity is a key feature to investigate and assess prior to utilizing the simulated motions for the analysis of three-dimensional structures as it can lead to substantially different responses. For completeness, Figure 7D through G show the acceleration time histories of the selected real and simulated records, for both the FN and FP components, used to develop the elastic oscillator example.

5 | STRUCTURAL MODELS, EDPs AND RESPONSE HYPOTHESIS TESTING—STEP 4

The building adopted for this study is a prototype 40-story steel special moment frame,[52] designed according to the provisions of ASCE 7–16. The building is idealized as a two-dimensional analytical model in NEVADA, an implicit nonlinear finite element program.[53] The model utilizes displacement-based beam elements with fiber cross sections and multiple integration points for primary beams and columns, and includes P-Delta effects and geometrical nonlinearities. The steel follows a bilinear constitutive relationship, with yielding point at 65 ksi for columns and 50 ksi for beams and kinematic hardening of 0.0025. A consistent mass matrix is generated for the self-weight of the framing system through the definition of geometry and material mass density. The additional mass representing the superimposed dead loads and live loads at each floor is modeled as concentrated masses lumped at the beam nodes. Unmodeled energy dissipation is introduced through Rayleigh damping (ratio of 5%), anchored at the first ($T1$) and third ($T3$) modes of the structure. Table 2 summarizes the modal periods (and frequencies). The building was subjected to a total of 4980 simulated ground motions (2490 in each direction, FN and FP) and 76 real records (38 in each direction, FN and FP), through uniform excitation. All building analyses were performed on CORI supercomputer at National Energy Research Scientific Computing Center (NERSC).

Nonlinear response history analyses were performed and the Peak Interstory Drift (PID) ratio was selected as EDP to evaluate the response of the 40-story moment frame. Monotonic pushover analysis was employed to verify the effective nonlinear behavior of the structural model, and indicated that the structure exhibits nonlinear behavior for PID of 1%.[54]

The difference in the structural response to the simulated ground motions and real records is estimated in terms of max(PID) over the height of the building, which is indicative of the maximum demand posed to the structure, and PID envelopes, which provide indication on the localization of the demand in the structure.

The statistical significance of the difference in the building response is evaluated through the hypothesis testing, as proposed by Jayaram and Shome.[55] The objective of this test is to identify whether the differences in the response are due to the randomness associated with the finite size of the samples or to inherent divergences in the features of the ground motions. The test is conducted by introducing a "null hypothesis," which in this specific case corresponds to the condition of zero true difference between the median response of the buildings under the simulated records and under the real records. The population of the differences of median response is assumed to be normally distributed, with 0 mean and standard error $\sqrt{\frac{\sigma^2_{lnPID,rec}}{n_{rec}} + \frac{\sigma^2_{lnPID,sim}}{n_{sim}}}$, being $\sigma_{lnPID,rec}$ and $\sigma_{lnPID,sim}$ the standard deviation of the PID as obtained from the real records and simulated records, respectively, n_{rec} the number of real records, and n_{sim} the number of simulated records. User-defined percentiles of this normal distribution define the boundaries of the rejection region. Therefore, the null hypothesis is not rejected if the difference of the estimates of the median response falls within the region, meaning that the observed difference is due to the randomness associated with the finite size of the samples only. Whereas, the null hypothesis is rejected if the difference of the estimates of the median responses falls outside the region, meaning that the difference is unlikely to be caused by the sample size. In this study 2.5 and 97.5 percentiles are used, corresponding to a significance level $\alpha = 0.05$.

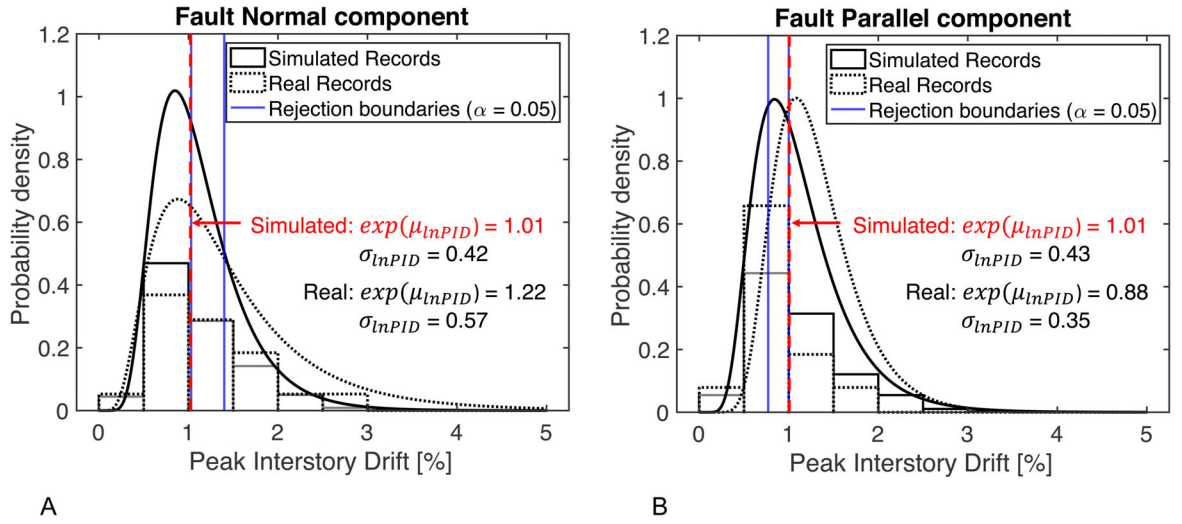
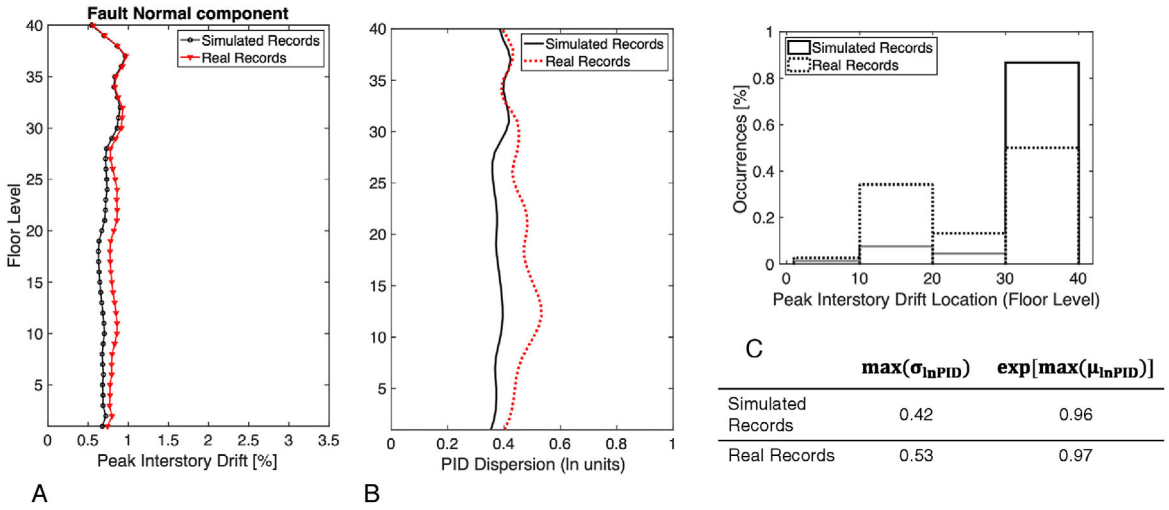


FIGURE 8 Normalized histogram of PID ratio, corresponding statistics, and acceptance criteria for (A) Fault Normal and (B) Fault Parallel component

5.1 | Results of structural response comparisons

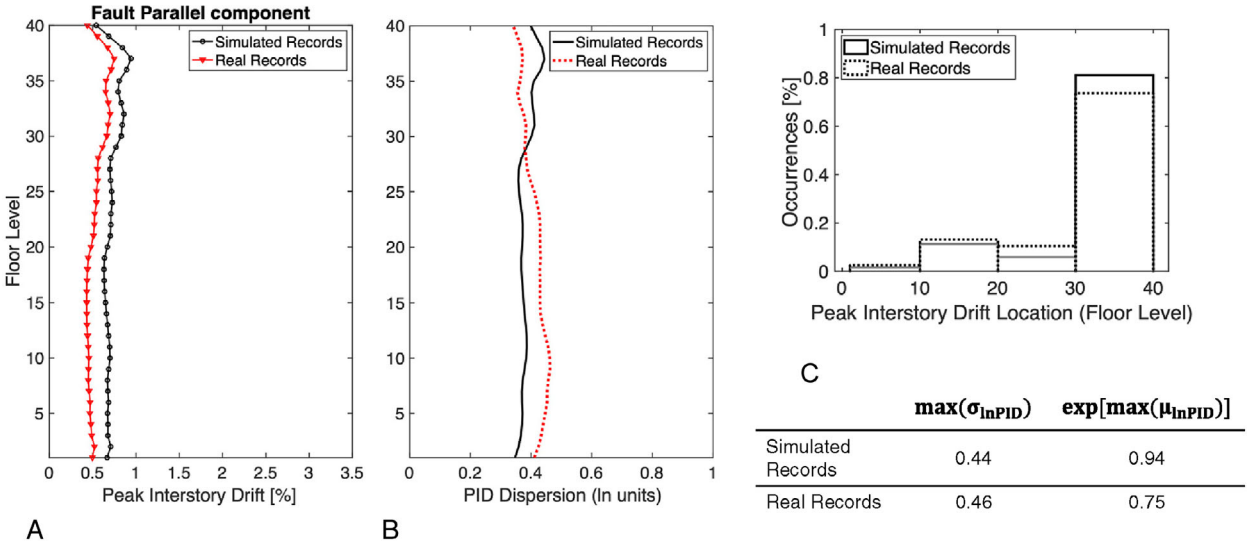
The nonlinear response of the 40-story building obtained from 38 real records (homogenized through the mixed-effects regression model, for consistency with the approach followed in treating the ground-motion database) and 2490 simulated records applied separately in the two horizontal directions (FN and FP components) is first statistically compared in terms max(PID) recorded over the entire duration of the ground motion. Figure 8A and B show the normalized histograms with overlain fitted lognormal distribution for PID resulting from real records (with dotted lines) and simulated records (with solid lines), for the FN and FP component, respectively. The agreement perceptible from visual inspection of the plots is verified against the statistics of the distributions. According to the hypothesis testing, the rejection boundaries of the median of the max(PID) - $\exp(\mu_{lnPID})$ - are 1.04% and 1.40% for the FN component, and 0.77% and 1.00% for the FP component (thin blue lines). As a result, the median of the max(PID) is right outside the lower bound of the rejection boundaries for the FN component with $\exp(\mu_{lnPID}) = 1.01$, whereas the median of the max(PID) is right outside the upper bound of the rejection boundaries for the FP component with $\exp(\mu_{lnPID}) = 1.01$, which is a first indicator of how the discrepancies observed in the ground-motion IMs, see Figure 2A, 2B and 5A, translate in differences in the structural response and, consequently, in the evaluation of the maximum demand posed to the structure. However, a more detailed analysis is conducted to investigate how the demand posed by real and simulated records is distributed at the structural level and whether a localization of the damage is observed.

Figure 9A and B show median PID and corresponding standard deviation at each floor level, obtained from the FN component of the 38 real records and 2490 simulated records. Again, for consistency in the analysis, the same approach (mixed-effects regression model) adopted to homogenize the database of real records is used to homogenize the structural response obtained from real records. It is seen that the median PID envelopes are in close agreement in the upper portion of the building, in terms of absolute value of the median PID and shape. This is also the portion of the structure where the demand is concentrated, which is usually indicative of a significant contribution of higher modes to the structural response.[56] Evidence from this plot suggests that the simulated motions have a frequency content comparable to the frequency content of the population of real records, confirming what inferred from Figure 5C. Moving to lower floors, the difference in the PID becomes more pronounced, with a tendency of the simulated records to underestimate the demand compared to the real records. As noted by Alavi and Krawinkler,[56] in tall buildings the demand tends to be concentrated in the first third of the building height when the exciting motion is characterized by long-period pulses. Therefore, evidence from this analysis may be indicative of a lack of long-period pulse-like motions in the FN component of the simulated ground motions, as inferred from on Figures 5A and 5C. However, a separate detailed study conducted on the real ground-motion database has shown that only a few motions, such as the FN component of the Imperial Valley-06 (1979) earthquake at stations ElCentro arrays #5, #6 and #7, pose a demand on the structure localized in the first third of the building height. This high variability is reflected in the standard deviation of PID derived from real records, that attains a max value $\sigma_{lnPID} = 0.53$ between the 10th and the 15th floor, differently from the standard deviation of PID from



	$\max(\sigma_{\ln PID})$	$\exp[\max(\mu_{\ln PID})]$
Simulated Records	0.42	0.96
Real Records	0.53	0.97

FIGURE 9 (A) Median, $exp(\mu_{lnPID})$ and (B) standard deviation, σ_{lnPID}, of PID envelope, and (C) frequency plot of PID location, for the Fault Normal component of simulated and real records



	$\max(\sigma_{\ln PID})$	$\exp[\max(\mu_{\ln PID})]$
Simulated Records	0.44	0.94
Real Records	0.46	0.75

FIGURE 10 (A) Median, $exp(\mu_{lnPID})$ and (B) standard deviation, σ_{lnPID}, of PID envelope and (C) frequency plot of PID location, for the Fault Parallel component of simulated and real records

simulated records that has an average value $\sigma_{lnPID} \sim 0.35$ over the entire building height, see Figure 9B. To better interpret, in a statistical sense, to what extent the demand posed to the structures differs between real and simulated records, the bar plot in Figure 9C shows how the PID is distributed along the building height. Each bar indicates the percentage of buildings exhibiting the PID between floors 1-10, 11-20, 21-30, 31-40. It is observed that in a comparable percentage of buildings subject to the FN component of the real records (dotted bars) the PID is located between floor 10 and 20 (~35%), and between floor 30 and 40 (~45%). Whereas the PID is located at the top ten floors in the absolute majority (~85%) of the buildings subject to the FN component of the simulated records (solid line). This distribution supports what has been observed in the median PID envelope and confirms what inferred from the analysis of the interperiod correlation.

Figure 10A shows the median PID envelope and corresponding variability (B) resulting from the application of the FP component of the 38 real records and 2490 simulated records. It is observed that over the full height of the building the shape of the PID envelopes is remarkably similar, and that the demand posed by the simulated records is consistently higher than the demand posed by the real records. This evidence suggests that the frequency content of the simulated records is comparable to the frequency content of the real records in the bandwidth of interest, ranging between 0.815 ($T3$) and 3.762 s ($T1$), see Table 2, as inferred from Figure 5A and B. The variability of the PID across all floors is similar over

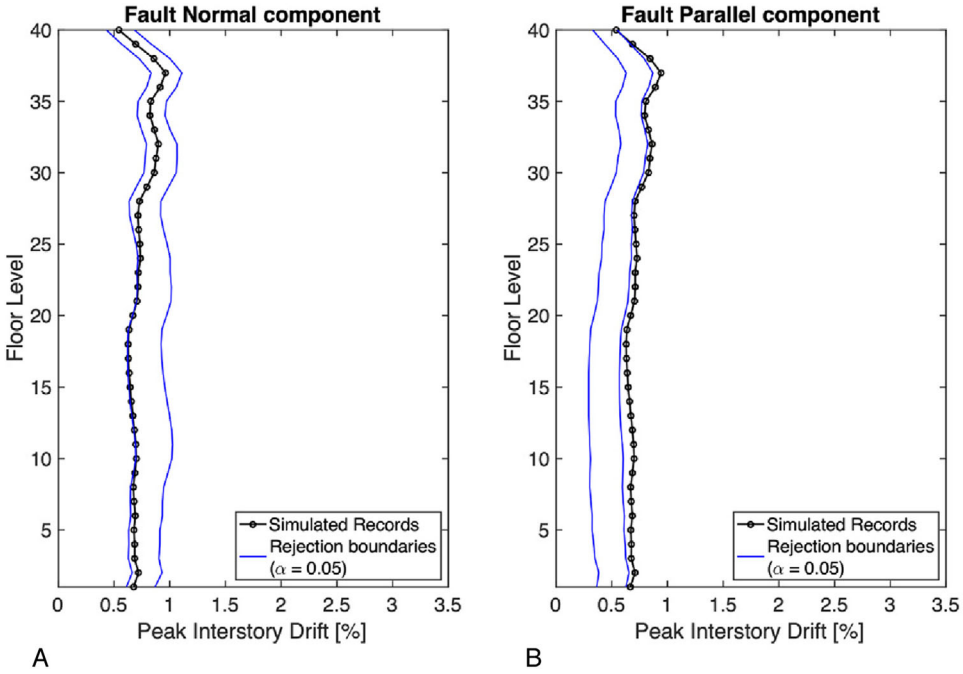


FIGURE 11 Median PID envelope from simulated records and acceptance criteria for the (A) Fault Normal and (B) Fault Parallel component

the building height and ranges between 0.35 and 0.45, see Figure 10B, which is indicative of a smaller heterogeneity of the real records in the FP component. as also reflected by the smaller standard deviation of the median spectral acceleration at longer periods in Figure 1C for spectral periods approximately between 1 s and 3 s. The similar shape of the PID envelopes obtained from the FP component of the real and simulated records is reflected in the bar plot of Figure 10C. It is in fact seen that in about ~80% of the buildings the demand is localized in the top ten floors of thfe structure and for a comparable percentage raging between 5 and 10% the demand is concentrated at the lower floors. This appears to be in contrast with what was inferred from Figure 5A and B, and may be attributed to the specific characteristics of the building used for this study.

Overall, for both the FN and FP components the median PID never exceeds the value of 1%, which marks the onset of the nonlinear behavior of the 40-story building. However, as shown in a separate study conducted by the authors, there is a high site-specific variability of the demand posed to the structures residing in the vicinity of the fault rupture for all three simulated scenarios, with PID values that go as high as 3.5% for stations located at 1 km from the fault,[33,34] which corresponds to major nonlinearities in the structural systems and short of collapse. Therefore, for studies conducted on single simulated records or a subset of simulated records, the relationship between spectral correlations parameters and observed structural response should focus on the spectral shape at periods longer than 3.76 s.

To help develop some judgment on the differences observed in the building response to simulated and real records, the null hypothesis described in the previous section is verified. Figure 11A and B show the median PID envelopes from simulated records along with the rejection region boundaries for $\alpha = 0.05$, for the FN and FP components, respectively. The null hypothesis is not rejected if the following condition is met:

$$(\mu_{lnPID})_{sim,i} \in (\mu_{lnPID})_{rec,i} \pm 1.96\sqrt{\frac{\sigma^2_{lnPID,rec}}{n_{rec}} + \frac{\sigma^2_{lnPID,sim}}{n_{sim}}} \tag{8}$$

where $(\mu_{lnPID})_{sim,i}$ is the median PID at the ith floor resulting from the simulated records, $(\mu_{lnPID})_{rec,i}$ is the median PID at the ith floor as obtained from the real records and 1.96 is the z-score corresponding to $\alpha = 0.05$.

It is seen that for the FN component the simulated response is well within the boundaries in the upper floors of the building and right on the edge toward the lower floors, confirming that the observed discrepancy may be attributed to features of the simulated motions which differ from the real records, see Figure 2A, 5A, and 5C, and not to the sampling error.

For the FP component the median PID envelope obtained from the simulated records consistently lies on the upper boundary of the rejection region at the top floors, and falls outside the rejection region at the lower floors, which is again indicative of discrepancies in the features of the FP component of the simulated ground motions, as noted in steps 2 and 3 of the proposed methodology, see Figure 2B, 5A, and 5B.

Overall, the outcome of the hypothesis testing conducted on the structural response aligns with what observed and inferred from the ground-motion IMs and EDPs, confirming their ability to provide useful, although partial, insight into the ground-motion features that mostly affect structural response variability. Future studies will address the impact of the polarization of the ground motion on three-dimensional buildings to evaluate to what extent this feature affects the demand posed to structures.

6 | CONCLUSIONS

This work proposes a four-step methodology for the investigation and validation of simulated ground motions of not-historical events for use in civil engineering applications. The methodology is conceived for the validation of earthquake scenarios that are outside the range of conventional validation methods (e.g., large unknown events in the near-field), for which the database of real records is limited. The objective is to provide engineers with an effective step-by-step procedure to follow before using the structural response from simulated ground motions to inform the design and assessment of infrastructure.

The methodology is demonstrated through a comparison of 2490 M_w 7 near-field (< 10 km) simulated records with 5 Hz resolution generated from the Pitarka et al. [32] kinematic rupture model with a selected population of 38 real records. The structure employed for the building analyses is a 40-story steel moment resisting frame building.

Results from the case study show how it is possible to draw insightful correlations between features of the ground motions identified in the first steps of the validation procedure (steps 2 and 3) and the structural response (step 4), and thus identify in what range of spectral periods the simulated ground motions reliably reproduce the features of the real records. For example, the adopted suite of simulated motions demonstrated to satisfyingly capture the spectral amplitude and interperiod correlation for the FN component, although with a lack in the pulse-like motions at longer periods, and to overestimate spectral amplitude at long spectral periods for the FP component, as reflected in the analysis of the building response. The acceptance criteria defined for each step (2 through 4) of the proposed methodology demonstrated to provide reliable metrics for the evaluation of the realism of the simulated motions.

Overall, the proposed procedure provides an effective method for informing and advancing the science needed to generate realistic ground-motion simulations, and for building confidence in their use in engineering domains. Results from the case study show that the simulated ground motions can reliably capture features that have a significant effect on the structural response, particularly in the bandwidth 0.2 to 3 s, demonstrating there is promise in the use of simulated motions for the evaluation of the risk posed by unknown (not-historical), yet expected, large earthquake events to infrastructure residing in the near-field of a major fault.

The long-term objective of this research is to develop a suite of validated earthquake scenarios (25-30 total) that can be used to estimate the aleatory variability of the ground motion for a specific site, to eventually provide engineers with design and assessment procedures that account for the variability of the demand at near-fault sites.

ACKNOWLEDGEMENTS

This research utilized simulated ground motions generated within the EQSIM regional-scale seismic simulation framework application development project. EQSIM is an application development supported by the Exascale Computing Project (ECP), Project Number 17-SC-20-SC—a jointly effort by the U.S. Department of Energy and the National Nuclear Security Administration. The authors gratefully acknowledge Drs. Arben Pitarka and Arthur Rodgers for providing the database of simulated ground motions necessary for developing the presented study.

ORCID

Floriana Petrone https://orcid.org/0000-0003-4545-6303

REFERENCES

1. Baker JW, Luco N, Abrahamson NA, Graves RW, Maechling PJ, Olsen KB. Engineering uses of physics-based ground motion simulations. *NCEE 2014-10th U.S. National Conference on Earthquake Engineering: Frontiers of Earthquake Engineering*. 2014; https://doi.org/10.4231/D3QF8JK2C.
2. Aki K. Presidential address: possibility of seismology in the 1980's. *Bull Seism Soc Am*. 1980;70:1969-1976.
3. ASCE/SEI. 2016. Minimum Design Loads and Associated Criteria for Buildings and Other Structures.
4. Komatitsch D, Vilotte J-P. The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bull Seism Soc Am*. 1998;88(2):368-392.
5. Bielak J, Loukakis K, Hisada Y, Yoshimura C. Domain reduction method for three-dimensional earthquake modeling in localized regions, part I: theory. *Bull Seismol Soc Am*. 2003;93:817-824.
6. Douglas J, Aochi H. A survey of techniques for predicting earthquake ground motions for engineering purposes. *Surv Geophys*. 2008;29(3):187-220.
7. Zhang L, Wang J-T, Xu Y-J, He C-H, Zhang C-H. A procedure for 3D seismicsimulation from rupture to structures by coupling SEM and FEM. *Bull Seism Soc Am*. 2020;110:1134-1148.
8. Jurkevics A, Ulrych TJ. Representing ans simulating strong ground motion. *Bull Seismol Soc Am*. 1978;68(3):781-801.
9. Der Kiureghian A, Crempien J. An evolutionary model for earthquake ground motion. *Struct Saf*. 1989;6(2–4):235-246.
10. Pousse G, Bonilla LF, Cotton F, Margerin L. Nonstationary stochastic simulation of strong ground motion time histories including natural variability: application to the K-Net Japanese database. *Bull Seismol Soc Am*. 2006;96(6):2103-2117.
11. Boore DM. Stochastic simulation of high-frequency ground motion based on seismological models of the radiated spectra. *Bull Seismol Soc Am*. 1983;73(6):1865-1894.
12. Beresnev IA, Atkinson GM. Modeling finite-fault radiation from the Ωn spectrum. *Bull Seismol Soc Am*. 1997;87(1):67-84.
13. Motazedian D, Atkinson GM. Stochastic finite-fault modeling based on a dynamic corner frequency. *Bull Seismol Soc Am*. 2005;95(3):995-1010.
14. Hartzell S, Harmsen S, Frankel A, Larsen S. Calculation of broadband time histories of ground motion: comparison of methods and validation using strong-ground motion from the 1994 Northridge earthquake. *Bull Seismol Soc Am*. 1999;89(6):1484-1504.
15. Graves RW, Pitarka A. Broadband ground-motion simulation using a hybrid approach. *Bull Seismol Soc Am*. 2010;100(5 A):2095-2123.
16. Mai PM, Imperatori W, Olsen KB. Hybrid broadband ground-motion simulations: combining long-period deterministic synthetics with high-frequency multiple s-to-s backscattering. *Bull Seismol Soc Am*. 2010;100(5A):2124-2142.
17. Somerville P. Engineering applications of strong ground motion simulation. *Tectonophysics*. 1993;218(1–3):195-219.
18. Somerville PG, 2002. Characterizing Near Fault Ground Motion For The Design And Evaluation Of Bridges. *Proceedings 3rd National Seismic Conference & Work Shop on Bridges & Highways*, Portland, Oregon, no. 01: 137-148.
19. Krawinkler H, Medina R, Alavi B. Seismic drift and ductility demands and their dependence on ground motions. *Eng Struct*. 2003;25(5):637-653.
20. Aagaard BT, Brocher TM, Dolenc D, et al. Ground-motion modeling of the 1906 San Francisco earthquake, Part I: validation Using the 1989 Loma Prieta earthquake. *Bull Seismol Soc Am*. 2008;98(2):989-1011.
21. 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;86(1):39-47.
22. Anderson JG. Quantitative measure of the goodness-of-fit of synthetic seismograms. *Proceedings of the 13th World Conference on Earthquake Engineering*. 2004;243:243.
23. Gersch W, Kitagawa G. A time varying AR coefficient model for modelling and simulating earthquake ground motion. *Earthq Eng Struct Dyn*. 1985;13:243-254.
24. Olsen KB, Mayhew JE. Goodness-of-fit criteria for broadband synthetic seismograms, with application to the 2008 Mw 5.4 ChinoHills, California, earthquake. *Seismol Res Lett*. 2010;81(5):715-723.
25. 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.
26. Bazzurro P, Sjoberg B, Luco N, 2004. Post-Elastic Response of Structures to Synthetic Ground Motions. *Report for Pacific Earthquake Engineering Research (PEER) Center Lifelines Program Project*, 65-112.
27. 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.
28. 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;42:1395-1412.
29. Krishnan S, Ji C, Komatitsch D, Tromp J. Performance of two 18-story steel moment-frame buildings in southern California during two large simulated San Andreas earthquakes. *Earthq Spectra*. 2006;22:1035-1061.
30. Jayaram N, Baker JW. Correlation model for spatially distributed ground-motion intesities. *Earthquake Eng Struct Dyn*. 2009;38:1687-1708.
31. Star L, Stewart JP, Graves RW. Comparison of ground motions from hybrid simulations to NGA prediction equations. *Earthq Spectra*. 2011;27(2):331-350.
32. Pitarka A, Graves R, Irikura K, Miyakoshi K, Rodgers A. Kinematic rupture modeling of GroundMotion from the M7 Kumamoto, Japan-Earthquake. *Pure Appl Geophys*. 2019;177:2199-2221.

33. McCallen D, Petersson A, Rodgers A, Pitarka A, Miah M, Petrone F, Sjogreen B, Tang H, 2020. EQSIM – A Computational Framework for Fault-to-Structure Earthquake Simulations on Exascale Computers Part I: Computational Models and Workflow. *Earthquake Spectra*. In press.
34. McCallen D, Petrone F, Miah M, Pitarka A, Rodgers A, Abrahamson N, 2020. EQSIM – A Computational Framework for Fault-to-Structure Earthquake Simulations on Exascale Computers Part II: Regional Simulations of Building Risk. *Earthquake Spectra*. In press.
35. Graves RW, Pitarka A. Kinematic ground motion simulations on rough faults including effects of 3D stochastic velocity perturbations. *Bull Seismol Soc Am*. 2016;106(5):2136-2153.
36. Baker JW, Shahi SK, 2011. New ground motion selection procedures and selected motions for the PEER transportation research program. Risk Management.
37. Baker JW, Jayaram N, Shahi SK, 2011.Ground Motion Studies for Transporation Systems. https://Peer.Berkeley.Edu/Research/Transportation-Systems/Ground-Motion-Studies-Transporation-Systems.
38. Al Atik L, Abrahamson NA, Bommer JJ, Scherbaum F, Cotton F, Kuehn N. The variability of ground-motion prediction models and its components. *Seismol Res Lett*. 2010;81(5):794-801.
39. Somerville PG. Magnitude scaling of the near fault rupture directivity pulse. *Phys Earth Planet Inter*. 2003;137:201-212.
40. Somerville PG, Pitarka A. Differences in earthquake source and ground motion characteristics between surface and buried faulting earthquakes. Bull. Earthq. Res. Inst., Univ. Tokyo, Vol. 81, 2006; pp. 259-266.
41. Abrahamson NA, Silva WJ, Kamai R. Summary of the ASK14 ground motion relation for active crustal regions. *Earthquake Spectra*. 2014;30(3):1025-1055.
42. Boore DM, Stewart JP, Seyhan E, Atkinson GM. NGA-West2 equations for predicting vertical-component PGA, PGV, and 5%-damped PSA from shallow crustal earthquakes. *Earthquake Spectra*. 2016;32(2):1005-1031.
43. Campbell KW, Bozorgnia Y. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. *Earthquake Spectra*. 2014;30(3):1087-1114.
44. Chiou BSJ, Youngs RR. Update of the Chiou and Youngs NGA model for the average horizontal component of peak groundmotion and response spectra. *Earthquake Spectra*. 2014;30(3):1117-1153.
45. Al Atik L, Youngs RR, 2013. Epistemic Uncertainty for NGA-West2 Models. *Peer Report*, no. May 2013, 86. https://doi.org/10.1193/062813EQS173M.
46. 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.
47. Shome N, Cornell CA, Bazzurro P, Carballo JE. Earthquakes, records, and nonlinear responses. *Earthquake Spectra*. 1998;14(3):469-500.
48. Baker JW, Cornell CA. A vector-valued groundmotion intensity measure consisting of spectral acceleration and epsilon. *Earthquake Eng Struct Dyn*. 2005;34(10):1193-1217.
49. Baker JW, Cornell CA. Spectral shape, epsilon and record selection. *Earthquake Eng Struct Dyn*. 2006;35(9):1077-1095.
50. Bayless J, Abrahamson NA. Evaluation of the interperiod correlation of ground-motion simulations. *Bull Seismol Soc Am*. 2018;108(6):3413-3430.
51. Boore DM. Orientation-independent, Nongeometric-mean measures of seismic intensity from two horizontal components of motion. *Bull Seismol Soc Am*. 2010;100(4):1830-1835.
52. Astaneh-Asl A, 2018. Designs of 3-, 9-, 20- and 40-Story Frames.
53. McCallen D, Larsen S, 2003. NEVADA—A Simulation Environment for Regional Estimation of Ground Motion and Structural Response. *Lawrence Livermore National Laboratory Directed Research and Development Report* UCRL-ID-15.
54. Wu S, Miah M, Mccallen D, 2018. Four Canonical Steel Moment Frame Buildings and Inter-Code Comparisons of Nonlinear Building Response.
55. Jayaram N, Shome N, 2012. A Statistical Analysis of the Response of Tall Buildings to Recorded and Simulated Ground Motions." *15th World Conference on Earthquake Engineering (15WCEE)*, 10p.
56. Alavi B, Krawinkler H. Behavior of moment-resisting frame structures subjected to near-fault ground motions. *Earthquake Eng Struct Dyn*. 2004;33(6):687-706.

How to cite this article: Petrone F, Abrahamson N, McCallen D, Miah M. Validation of (not-historical) large-event near-fault ground-motion simulations for use in civil engineering applications. *Earthquake Engng Struct Dyn*. 2021;50:116–134. https://doi.org/10.1002/eqe.3366