Bulletin of the New Zealand Society for Earthquake Engineering, Vol. 56, No. 4, December 2023

EVALUATION OF THE INTER-FREQUENCY CORRELATION OF NEW ZEALAND CYBERSHAKE CRUSTAL EARTHQUAKE SIMULATIONS

Jeff Bayless[1] and Scott Condon[2]

(Submitted *September 2022*; Reviewed *February 2023*; Accepted *May 2023*)

ABSTRACT

The inter-frequency correlation of ground-motion residuals is related to the width of peaks and troughs in the ground-motion spectra (either response spectra or Fourier amplitude spectra; FAS) and is therefore an essential component of ground-motion simulations for representing the variability of structural response. As such, this component of the simulations requires evaluation and validation when the intended application is seismic fragility and seismic risk. This article evaluates the CyberShake NZ [1] crustal earthquake ground-motion simulations for their inter-frequency correlation, including comparisons with an empirical model developed from a global catalogue of shallow crustal earthquakes in active tectonic regions, and with results from similar simulations (SCEC CyberShake; [2]). Compared with the empirical model, the CyberShake NZ simulations have a satisfactory level of total inter-frequency correlation between the frequencies 0.1 – 0.25 Hz. At frequencies above 0.25 Hz, the simulations have lower (statistically significant at 95% confidence level) total inter-frequency correlation than the empirical model and therefore require calibration. To calibrate the total correlation, it is useful to focus on the correlation of the residual components. The between-event residual correlations, physically related to source effects (e.g., stress drop) which drive ground motions over a broad frequency range, are low at frequencies greater than about 0.25 Hz. Modifications to the cross-correlation between source parameters in the kinematic rupture generator can improve the inter-frequency correlations in this range [3]. The between-site residual correlations, which represents the correlation between frequencies of the systematic site amplification deviations, are larger (statistically significant at 95% confidence level) than the empirical model for frequencies less than about 0.5 Hz. We postulate that this relates to the relative simplicity of site amplification methods in the simulations, which feature less variability than the amplification observed in the data. Additional insight would be gained from future evaluations accounting for repeatable path and basin effects, using simulations with refined or alternative seismic velocity models, and using simulations with a higher crossover frequency to deterministic methods (e.g., 1 Hz or higher).

https://doi.org/10.5459/bnzsee.1623

INTRODUCTION

The New Zealand CyberShake (CyberShake NZ; [1]) approach uses a hybrid of stochastic method and 3D wave propagation simulations, with finite-fault rupture descriptions, to forecast ground-motions that will be produced by scenario ruptures in NZ. CyberShake NZ generates kinematic ruptures using the Graves and Pitarka method [4] with the crustal fault sources from Stirling et al. [5]. The simulations use a detailed 3D velocity model and feature a total of 11,362 finite fault rupture simulations computed on a spatially variable grid of 27,481 locations.

There is increasing recognition in the seismological community that simulations, such as those from CyberShake NZ, can be utilized in future engineering applications such as the dynamic analyses of structures for specific site/source rupture geometries that are not well represented in empirical datasets. For these simulations to be used in forward applications, their predictive abilities must be validated. The validation process involves comparing the physics-based simulations to observations, when available, and to empirical ground-motion models [6,7]. Results of these comparisons, along with quantitative and qualitative acceptance criteria, are used to determine the magnitude, distance, and frequency ranges for which the simulations are deemed acceptable for forward use. Validation of the simulations should be carried out for the ground-motion parameters relevant to the intended application. For example, Goulet et al. [8] performed a complete ergodic (not region-specific) validation for a suite of simulation methods using median response spectra as the validation metric. As the simulation validation process matures, region-specific validations, such as those in Lee et al. [9], should be adopted in place of ergodic validations.

The inter-frequency correlation component of earthquake simulations require validation when the intended application is seismic fragility and seismic risk [10]. The inter-frequency correlation is related to the width of peaks and troughs in the Fourier amplitude spectra (FAS) of the simulations, which in turn impact the variability of the structural response when the simulated time histories are utilized in dynamic analyses. The significance of the inter-frequency correlation as a validation tool is described further in the *Background* section of this article.

This study builds upon Bayless and Abrahamson [10], which applied inter-frequency correlation validation methods to ground-motion simulations from the Southern California Earthquake Center (SCEC) Broadband Platform (BBP), and upon Bayless and Condon [11; BC20 hereafter] which extended the same validation methodology to the SCEC CyberShake [2] simulations. We evaluate the inter-frequency correlations of

[1] *Corresponding Author, Engineering Seismologist, AECOM, Los Angeles, CA, jeff.bayless@aecom.com*
[2] *Seismologist, AECOM, Los Angeles, CA, scott.condon@aecom.com*

FAS residuals from CyberShake NZ crustal earthquake ground-motion simulations, and include comparisons with an empirical model for the correlation, and with the correlation results from SCEC CyberShake simulations. This evaluation procedure contains three main components: 1) simulation data collection and processing, 2) residual analysis, and 3) the inter-frequency correlation analysis.

BACKGROUND

The purpose of evaluating the inter-frequency correlations of the FAS is to better identify areas of improvement for simulation methods, and to allow for comprehensive validation of inter-frequency correlations in ground-motion simulations. It is well understood by seismologists that the FAS provides a more direct representation of the frequency content of the ground-motions as compared with response spectra [12]. This is because the Fourier transform is a linear operation, whereas the response spectrum is not because it is derived from the peak response over time from a single degree of freedom system, which is influenced by a range of ground-motion frequencies. Using the FAS intensity measure rather than the traditionally adopted pseudo-spectral acceleration provides the developers of the simulation methods more meaningful feedback on how they can modify their methods both because it is better understood and because the FAS inter-frequency correlation models are less complex.

The appropriate inter-frequency correlations are required to correctly estimate structural seismic fragilities and risk, because the ground-motion inter-frequency correlation is related to the width of peaks and troughs in a spectrum (either a Fourier or response spectrum), and that the structural response variability can be under-estimated if the inter-frequency correlation is too low [10]. Low structural response variability leads to fragility curves that are too steep, and to un-conservative estimates of seismic risk [10]. These conclusions are applicable to structural risk assessments derived from ground-motion simulations, commonly referred to as "ruptures to rafters" simulations. Therefore, it is important to validate the inter-frequency correlation of simulations which are to be used for seismic risk.

Bayless and Abrahamson [10] evaluated six SCEC BBP simulation methods and compared the inter-frequency correlations with the Bayless and Abrahamson [13; BA18Corr hereafter] empirical model. BA18Corr was developed from the NGA-W2 database of recorded crustal earthquake ground-motions [14]. Bayless and Abrahamson [10] found that none of the six tested finite-fault simulation methods adequately represented the correlations over the entire frequency range evaluated, and although several of the methods showed promise at low frequencies, the total correlations were low compared with BA18Corr.

However, the conclusions from [10] were partly obscured by two procedural differences in the residual analyses performed on the simulations and on the NGA-W2 data used to derive BA18Corr. First, the SCEC BBP simulations are based on plane-layered (1-D) seismic velocity models so all sites for a given scenario have the same shear wave velocity at the surface, and there is no variability in the site response. As a result, the residual site term could not be distinguished from the residual constant term. In this context, the "site terms" in the 1-D simulations do not have the same meaning as they do in a residual analysis from a recorded database like NGA-W2, where each site has a unique velocity profile beneath it with characteristic effects.

Second, the "source terms" (analogous to event terms) in the BBP simulations were determined from alternative realizations of the same earthquake with different slip distributions, whereas in the NGA-W2 database each earthquake has unique properties (such as magnitude, location, dimensions, stress drop, slip, and hypocentre, etc.). These source terms also do not have the same meaning as those from NGA-W2, where the event terms represent systematic bias of the observed ground-motions from an individual earthquake.

Considering these limitations, the conclusions from [10] were limited to the inter-frequency correlation of the between-event and within-site residual components. In this study, by using CyberShake NZ simulations and by selecting a wider range of sources, sites, and site conditions, we are better able to match the distribution of data in the NGA-W2 database. This procedure allows for the separation of the FAS residuals into repeatable source, repeatable site, and remainder components, greatly improving the comparison between inter-frequency correlations calculated from the CyberShake ground-motions and from NGA-W2 ground-motions.

The remainder of this manuscript describes the subset of the CyberShake NZ database utilized (*Simulation Database*), the FAS residual analysis (*Residual Analysis*), the inter-frequency correlation analysis including comparison with the BA18Corr model and with SCEC CyberShake (*Inter-frequency Correlations*), an overview of calibration methods (*Calibration of Simulations*), and our *Summary and Conclusions.*

SIMULATION DATABASE

SCEC CyberShake is a computational study to calculate ground-motion hazard in the Los Angeles region using scenario-based earthquake simulations [e.g., 2, 15, 16, and subsequent unpublished updates]. CyberShake NZ employs similar simulation methodology to NZ, as summarised here. This study uses the CyberShake NZ version 20.11 simulations (v20.11; see *Data and Resources*). Bradley et al. [1] describes the v19.5 simulations, which are similar in methodology to v20.11. The main upgrades embodied in v20.11 are finer grid spacing of 200m and a higher transition frequency of 0.5 Hz. The QuakeCoRE CyberShake NZ wiki page contains a complete list of the project versions, including the computational and scientific components of each (see *Data and Resources*). Both the v19.5 and v20.11 CyberShake NZ adopt a 'forward' simulation approach, as opposed to the reciprocity approach implemented by SCEC [1]. The Graves and Pitarka [17] kinematic rupture generation method is used to define the earthquake ruptures, using the shallow crustal faults in [5]. For the full CyberShake NZ event set, a Monte Carlo scheme was used to sample variability in the seismic source parametrization by varying the hypocentre location along the strike and dip directions, and slip distribution per each hypocentre realization. The total number of rupture realizations for each fault was based on the corresponding rupture magnitude, with a minimum of 10 realizations per source [1].

The CyberShake NZ methodology adopts the hybrid broadband ground-motion simulation approach developed by Graves and Pitarka [17, 18, 19]. This approach computes the low-frequency (LF) and high-frequency (HF) ground-motion components separately using comprehensive and simplified physics, respectively. The LF simulation explicitly models 3D wave propagation using the finite difference method, and the HF model uses a finite fault version of the stochastic method (a stochastic source radiation pattern and simplified wave propagation with attenuation through a 1D layered velocity model). These two simulations are filtered and merged in the time domain at the transition frequency, set to 0.5 Hz for CyberShake NZ v20.11 (see *Data and Resources*).

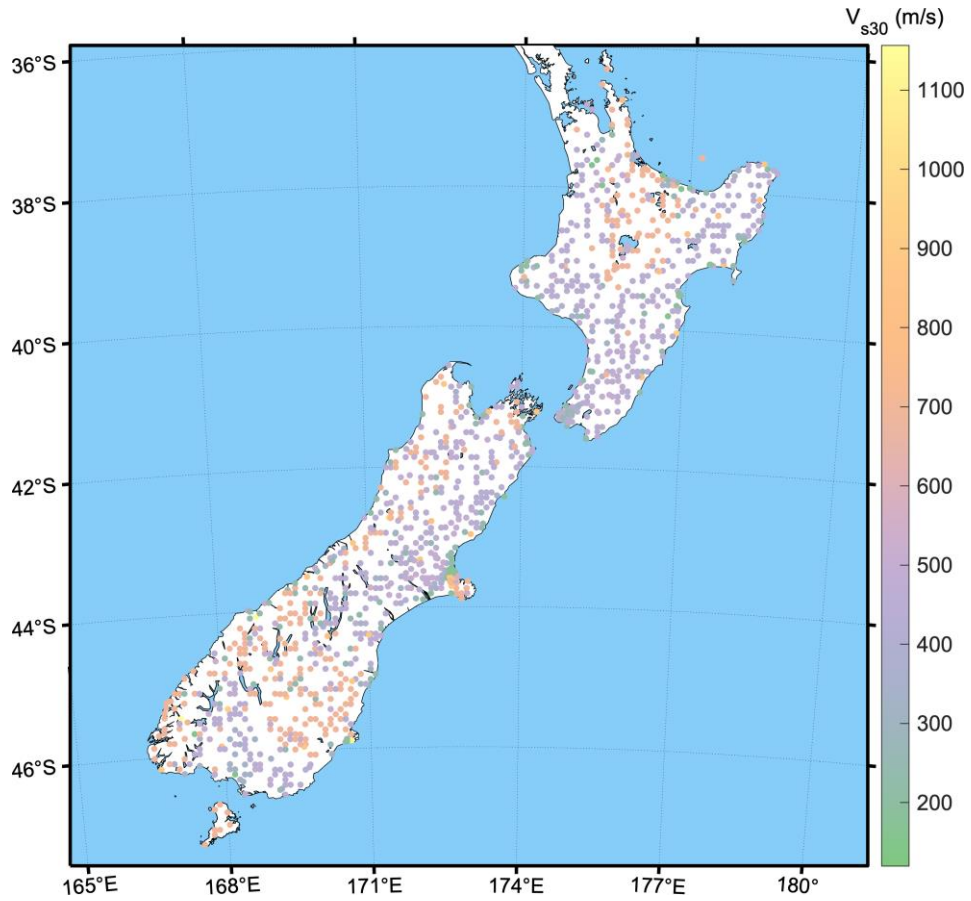

Figure 1: A map of the CyberShake NZ simulation stations selected for this study, colour coded by Vs30.

Lee et al. [9] describes two modifications to the Graves and Pitarka [18] method which have also been applied to the CyberShake NZ v20.11 simulations. These are changes to the HF path duration model to increase durations, and modification to the site amplifications. Details of the path duration model can be found in [9]. The site amplification model adjustments are described below.

In CyberShake NZ v19.5 and v20.11, the 3D simulations use a detailed velocity model with multiple sedimentary basins, NZVM2.03 [20]. The representation of the time averaged shear wave velocity in the upper 30m, Vs30, includes consideration for surface geology, topographic terrain, and direct Vs30 measurements including their uncertainty. The CyberShake NZ v20.11 results used in the current study, provided by University of Canterbury (see *Data and Resources*), are the result of 200m resolution simulations. These simulations have minimum shear wave velocity (Vs) of 500 m/s. Following the adjustment described in [9]), the Campbell and Bozorgnia [21; CB14] site response model is applied only to the HF component of the waveform before merging. This was due do inferred double-counting of long period site effects by Lee et al. [9] when applying the site corrections to the LF simulations.

The CB14 model is a spectral acceleration model, and the amplification was applied by University of Canterbury to the CyberShake NZ v20.11 simulations (see *Data and Resources*) in the frequency domain to the FAS of the HF simulation, and then transformed to the time domain before merging with the LF simulation to generate the broadband simulated waveforms. This process assumes that the response spectral amplification and the Fourier amplitude spectra amplification are the same.

CyberShake NZ Subset

A subset of the CyberShake NZ v20.11 simulations database was curated with the intention to approximate the distribution of earthquake magnitudes, site distances, and site conditions from a recorded database like NGA-W2. In this simulation set, there is a total of 25,785 simulated time series from 161 unique earthquakes and 1,233 unique sites. The simulation stations and earthquake scenarios span the North and South Islands of NZ. The scenarios span **M**5.4 – **M**8.0. For each earthquake source, one of the available rupture realizations (of source slip distribution and hypocentre) is randomly selected so that there is exactly one event term per source. Each earthquake has between 11 and 424 simulation stations with rupture distances ranging from 0-150 km. Each station has simulations from at least 8 scenario earthquakes. All stations have Vs30 between 120 and 1,156 m/s. Figure 1 maps the simulation stations in this database with colours corresponding to the assigned Vs30 values.

For the two databases, Figure 2 shows **M** vs distance scatterplots of the data and histograms of parameters **M**, Ztor (depth to top of rupture plane), Rrup (rupture distance), Vs30, and Z1.0 (depth to shear wave velocity horizon of 1 km/s). The CyberShake NZ database for this study is shown in Figure 2a and the NGA-W2 database used to develop BA18Corr is shown in Figure 2b. The minimum **M** in CyberShake NZ is notably larger than the smallest **M** from the NGA-W2 database. This is not expected to impact the inter-frequency correlations because there is not a statistically significant dependence of the inter-frequency correlation on **M** or distance [13]. The range of rupture distances in the CyberShake NZ database matches NGA-W2 well, although the Rrup from the NGA-W2 data are skewed towards smaller Rrup values. A future iteration of this analysis could improve the match between these distributions, but the differences are not expected to have an impact on the resulting correlations.

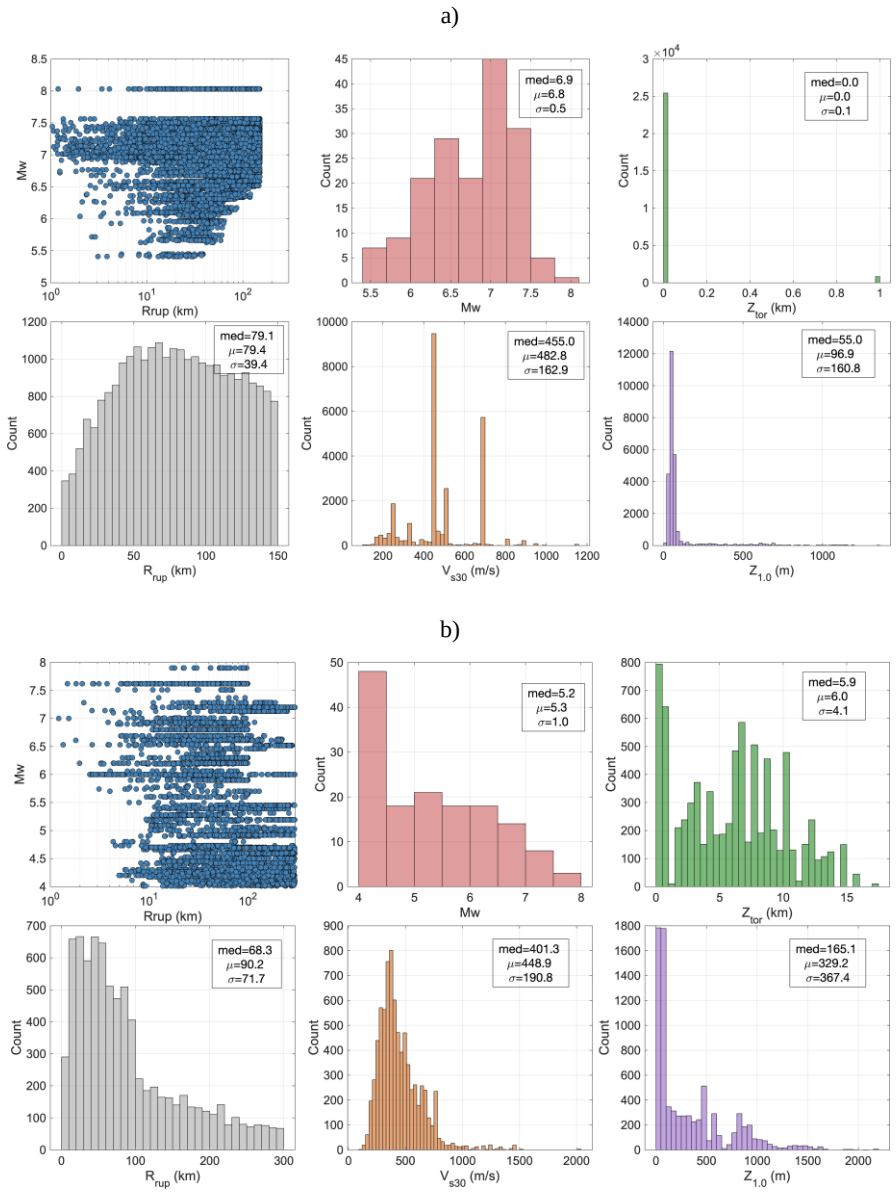

Figure 2: Distributions of the metadata in (a) the Cybershake NZ v20.11 database used in this study and (b) the subset of the NGA-W2 database used to develop BA18Corr .

The most significant difference between our CyberShake NZ database and the NGA-W2 is in the site conditions. The CyberShake NZ v20.11 simulations are calculated for a minimum Vs of 500 m/s in the 3D velocity model. The stations have an assigned Vs30 (from geology, terrain, and direct measurements) ranging from about 180 to 1200 m/s, and the CB14 site amplification is applied for the assigned Vs30 value to the HF simulations in the frequency domain, as described previously. It is likely that the relatively small variability in CyberShake NZ Vs profiles (compared with the profiles for the NGA-W2 recorded data) increases the LF inter-frequency correlation of the between-site residuals, as discussed further in *Inter-frequency Correlations.*

EAS Intensity Measure

The residual and inter-period correlation analyses performed on the simulations are based on the smoothed Effective Amplitude Spectra [EAS; 22] component of the Fourier amplitude spectra (FAS). To do so, the FAS for both horizontal components, the Effective Amplitude Spectra, and the smoothed EAS, are calculated following the procedure established by NGA-East (22, 23) and described in Bayless and Abrahamson [24; BA19 hereafter].

The BA19 empirical EAS ground-motion model used for calculating residuals (described in *Residual Analysis*) is valid over frequencies 0.1 – 24 Hz and is extended from 24 to 100 Hz using a kappa-based extrapolation. The CyberShake NZ v20.11 simulations are broadband hybrid (approximately 0.1 – 100 Hz) with transition frequency of 0.5 Hz between the low frequency (semi-deterministic) and high frequency (semi-stochastic) components. Considering the ranges of both the simulations and the empirical model, the residual analysis is performed for the frequency range 0.1 – 10 Hz. The CyberShake NZ v20.11 transition frequency of 0.5 Hz is important to consider when evaluating the results, as discussed later in this article. Figure 3 shows one example smoothed EAS from the CyberShake NZ v20.11 simulations (**M**5.66, Rrup = 46 km) compared with the BA19 median model prediction for the same scenario.

RESIDUAL ANALYSIS

The BA19 EAS ground-motion model was developed for crustal earthquakes using a database of ground-motions recorded primarily in California and Nevada. Ground-motions from crustal earthquakes recorded globally were used to develop the magnitude scaling component of the model. Using BA19, EAS residuals are calculated for the CyberShake NZ

v20.11 database. Following [25] and [10], the residuals take the form of Equation 1:

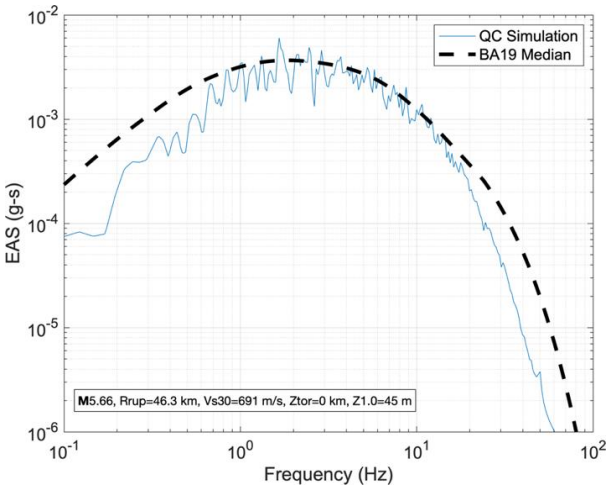

Figure 3: Smoothed EAS from an example CyberShake NZ v20.11 simulation (blue line), with the median BA19 prediction for the same scenario (black dashed line).

$$\delta_{total,es}(f) = Y(f) - g(X_{es}, \theta(f)) = \delta B_e(f) + \delta S2S_s(f) + \delta WS_{es} + C(f) \quad (1)$$

where $Y(f)$ is the natural log of the CyberShake NZ smoothed EAS at frequency f, $g(X_{es}, \theta(f))$ is the natural log of the median BA19 GMM, X_{es} is the vector of explanatory seismological parameters (magnitude, distance, site conditions, etc.), $\theta(f)$ is the vector of GMM coefficients, and $\delta_{total,es}(f)$ is the total residual for earthquake e and site s. The residual components $\delta B_e(f)$, $\delta S2S_s(f)$, and $\delta WS_{es}(f)$ represent the between-event, site-to-site, and single station within-site residuals, respectively. $C(f)$ represents the mean total residual, or the mean bias.

The mean bias exists because the median EAS from the simulations is different from the empirical model for a given scenario. The mean bias represents the average of this difference from all scenarios. The overall bias between the simulations and the empirical model is accounted for by removing $C(f)$. As shown in Figure 4, the mean bias is negative for frequencies less than about 1 Hz and is nearly zero for frequencies larger than 1 Hz. The crossover frequency between the low- and high-frequency simulation methods is at 0.5 Hz, as indicated by the red dashed line. The dip in $C(f)$ is most extreme at about -0.8 natural log units near 0.4 Hz. In this frequency range, the BA19 median model predicts on average larger EAS than the simulations by a factor of about 2.3. At 0.1 Hz, the difference is about a factor of 1.4.

In the frequency range of about 0.4 – 0.8 Hz there is a smooth transition in the mean bias from the from most extreme value to zero bias. This smooth transition is the result of the 4th order Butterworth applied to the LF and HF components of the simulation before merging them into the broadband waveforms. Additional causes of the mean bias in the LF simulations are an interesting topic for a future study aimed at validation of the median EAS.

Once the mean bias is removed from the residuals, the event terms ($\delta B_e(f)$, specific to each earthquake) and ($\delta S2S_s(f)$, specific to each site) are calculated using a mixed effects regression. Histograms of δB_e, $\delta S2S_s$, and δWS_{es}, for all frequencies analyzed, are shown in Figure 5. The residual components δB_e, $\delta S2S_s$, and δWS_{es} are well represented as zero mean, independent, approximately-normally distributed random variables with standard deviations τ, ϕ_{S2S}, and ϕ_{SS}, respectively.

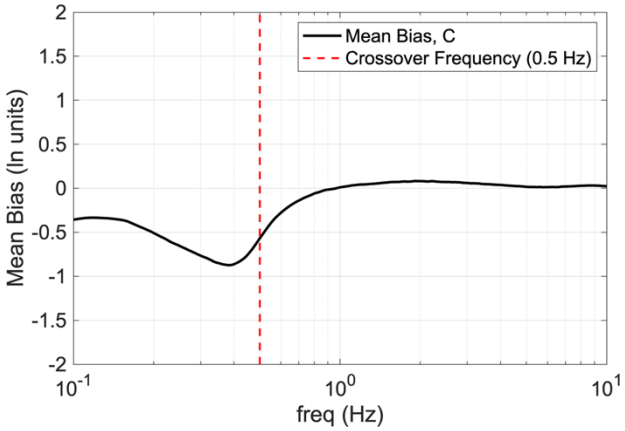

Figure 4: The mean bias term, $C(f)$.

The residuals from the regression analysis are visually inspected as functions of the main model parameters to check for errors and strong trends. The presence of obvious strong trends in the residuals versus predictor variables would indicate that the simulations do not agree with that component of the reference model. Figure 6 shows an example of this inspection at f = 0.2 Hz. In Figure 6a, the event and site terms are evaluated versus **M**, Ztor, Vs30, and Z1.0. In the right column, the within-site residuals are evaluated versus parameters **M**, Rrup, Vs30, and Z1.0. The events with Ztor = 0 km are plotted at the value Ztor = 0.01 km so they appear on the logarithmic-scale axis.

At f = 0.2 Hz, the event terms versus **M** and Ztor do not have strong trends, although trends with Ztor would be difficult to observe since the CyberShake simulations use two Ztor values: 0 and 1 km. The site terms with basin parameter Z1.0 do not appear to have strong overall trends, although at very deep sites the BA19 model tends to under-predict the simulations (positive residual). The most noticeable trend in Figure 6a is in the site terms versus Vs30. The linear trend in these site terms indicates that the Vs30 scaling of the BA19 GMM does not fully agree with the simulations, at f = 0.2 Hz. The BA19 GMM is predicting on average larger site terms for the low Vs30 values (below about 350 m/s) and slightly smaller site terms for Vs30 > 500 m/s conditions. This represents disagreement between the BA19 Vs30 scaling model and the CyberShake NZ simulations. The BA19 Vs30 scaling model is derived from ground-motions recorded in California, and therefore it represents the scaling associated with average shear wave velocity profiles in California. At low frequencies such as 0.2 Hz, where the CyberShake NZ v20.11 simulation is semi-deterministic and the CB14 site response model is not applied, the velocity profile specific to the region controls the site amplification.

At f = 0.2 Hz the within-site residuals (Figure 6b) do not have strong trends with **M**, Vs30, or Z1.0. However, there is disagreement in the distance scaling of the simulations with BA19. The bias is most pronounced for Rrup values less than about 10 km, and these residuals are negative, meaning the BA19 model over-predicts the simulations in this range. The BA19 model is based on relatively sparse data for these distances (Figure 7). Additionally, at less than about 10km distance, the empirical model saturates due to finite fault rupture dimension effects. The result shown in Figure 6 implies that the near-source saturation of f = 0.2 Hz ground-motions in BA19 should be stronger to match CyberShake NZ v20.11.

At frequencies greater than 0.2 Hz, there are no strong trends in the within-site residuals, although the same bias exists, to a lesser degree, in the distance scaling at small distances. The event terms do not exhibit strong trends with **M** or Ztor and the site terms do not show strong trends with Vs30 or Z1.0, hence these residuals are omitted here. These residuals, and the similar

results at other frequencies in the range 0.1 – 1.0 Hz, are suitable for the purposes of calculating the inter-frequency correlation in the next section. Several interesting observations discussed previously are excellent topics for a future validation study: differences in Vs30 scaling, distance scaling, and the mean bias (all at low frequencies).

Figure 7 shows the frequency dependence of the standard deviations of each residual component, along with the total standard deviation (σ). Figure 7b is replicated from BA19 (their Figure 3) and shows the same standard deviation components calculated from the NGA-W2 data. The variability from the HF (semi-stochastic method, f < 0.5 Hz) portion of the simulations is low compared with the LF portion. From this study, the τ is largest at frequencies less than about 0.35 Hz and has a peak value of about 0.45 natural log units. The low frequency peak is broadly consistent with the period dependence of τ in Figure 7b. The ϕ_{SS} between about 0.4 and 0.5 natural log units for f < 1 Hz is also broadly consistent with BA19.

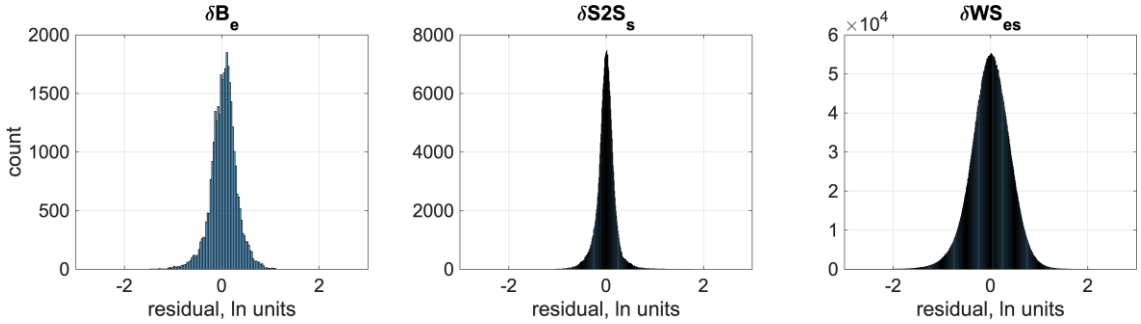

Figure 5: Histograms of δB_e, $\delta S2S_s$, and δWS_{es} from the residual analysis .

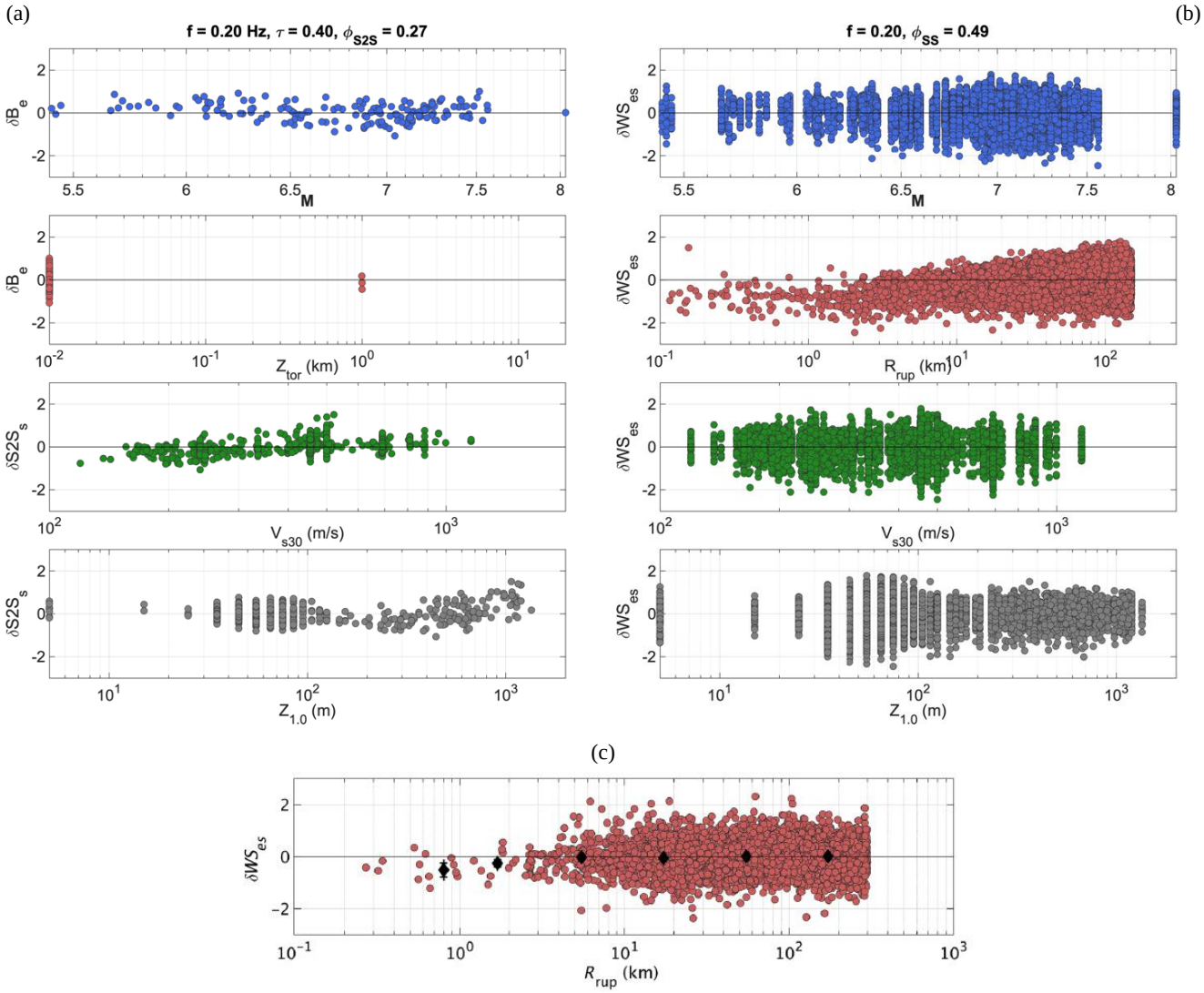

Figure 6: (a) Between-event, site-to-site and (b) within-site CyberShake NZ v20.11 EAS residuals at f = 0.2 Hz versus various predictor variables. (c) Example residuals from BA19 versus distance, illustrating the scarcity of NGA-W2 data with small rupture distances

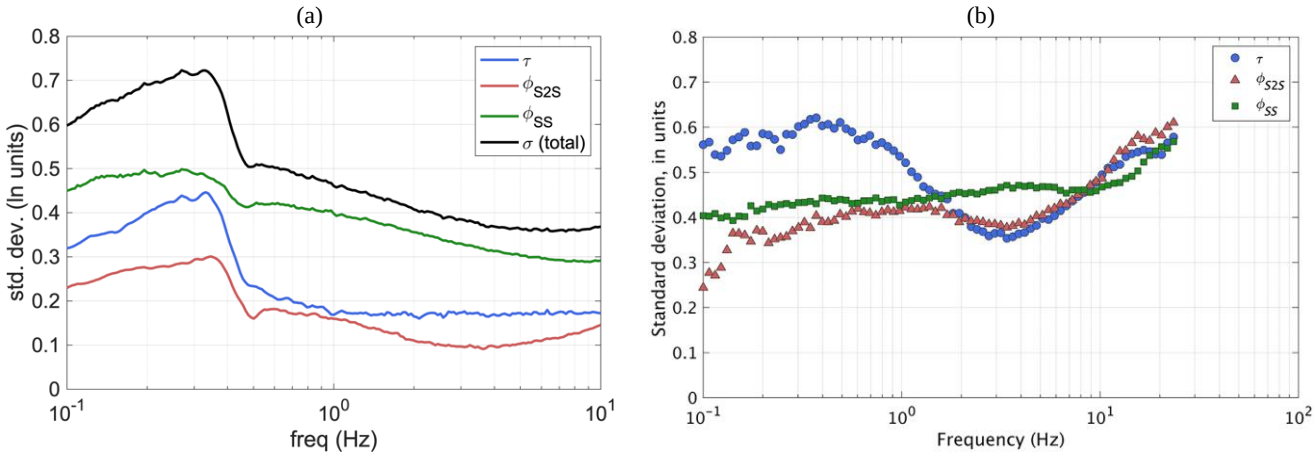

Figure 7: Frequency dependence of the standard deviations of the residual components from (a) this study and (b) BA18Corr .

INTER-FREQUENCY CORRELATIONS

The EAS residual components are converted to epsilon (ϵ) values by normalizing the residuals by their respective standard deviations, e.g., Equation 2:

$$\epsilon_B(f) = \frac{\delta B_e(f)}{\tau(f)}$$
$$\epsilon_{S2S}(f) = \frac{\delta S2S_s(f)}{\phi_{S2S}(f)} \quad (2)$$
$$\epsilon_{WS}(f) = \frac{\delta WS_{es}(f)}{\phi_{SS}(f)}$$

Because of the normalization, the random variables ϵ_B, ϵ_{S2S}, and ϵ_{WS} are well-represented by standard normal distributions (mean = 0 and variance = 1). The Pearson-product-moment correlation coefficient [ρ; 26] of ϵ between two different frequencies is calculated for each ϵ component ($\rho_{\epsilon,EAS}$), as well as the total correlation using Equations 3 and 4 of [13]. All correlations presented in this article are for the smoothed EAS, and for notational brevity the EAS subscript is dropped hereafter. Similarly, if not stated explicitly, the term "inter-frequency" is implied in all uses of the word "correlation" in this article.

The ρ_ϵ calculation can be repeated for every frequency pair of interest and the resulting correlation coefficients for each pair of frequencies can be saved as tables and displayed as contours (e.g., Figure 8).

Comparison with NGA-W2

Figure 8 shows contours of the total ρ_ϵ for (a) this CyberShake NZ v20.11 study and (b) BA18Corr (derived from NGA-W2 data). Both panels of this figure have the same color scale, and the horizontal and vertical axes are frequencies ranging from 0.1 – 10 Hz. These figures are symmetric about the 1:1 line because the correlation coefficient between two frequencies is the same regardless of which frequency is the conditioning frequency.

The total ρ_ϵ contours in Figure 8 are helpful for making broad comparisons between the total correlations from this study and those from BA18Corr, which is based on NGA-W2 recorded data. In this sense, the total correlation from the CyberShake NZ residuals below the transition frequency (f < 0.5 Hz) have contours running roughly parallel to the diagonal, and in this sense compare well with BA18Corr. At frequencies above 0.5 Hz, the total correlation is very different from BA18Corr, with much too steep of decay in the correlation at frequencies very close to the conditioning frequency.

Low inter-frequency correlation is a characteristic of the semi-stochastic simulation technique, which is used by the GP16 hybrid simulation method at frequencies above the transition frequency (0.5 Hz in CyberShake NZ v20.11). The stochastic method, outlined in [27], begins using the FAS of tapered white noise with random phase angles, which is then normalized to an acceleration amplitude spectrum. This results in no correlation between frequencies [10]. Therefore, the low total correlation at frequencies higher than the transition frequency, as evident in Figure 8, is an inherent limitation of hybrid-method simulations. The remainder of this evaluation focuses on the narrower frequency range 0.1 – 1.0 Hz. This range covers the semi-deterministic portion, for which the inter-frequency correlations can be calibrated, and the transition to the semi-stochastic portion of the simulations.

Previously, Bayless and Abrahamson [10] evaluated six different simulation methods on the SCEC BBP, and those conclusions varied by method, but the authors also generally found a poor match at frequencies greater than about 0.5 Hz with a rapid decay of the correlation at frequencies away from the conditioning frequency. Relative to that study, the result shown in Figure 8 is an improvement at f < 0.5 Hz. Both studies show a similar correlation pattern at f > 0.5 Hz. At these higher frequencies, the GP16 finite-fault implementation of the stochastic method is the same as used by several of the SCEC BBP methods.

To analyse these results in more detail, it is helpful to focus on the ρ_ϵ of the individual residual components (between-event, between-site, within-site) in addition to the total ρ_ϵ. To do so, the ρ_ϵ contours are deconstructed into cross sections at select conditioning frequencies; this is equivalent to taking "slices" of the contours. Figure 9 shows the total residual ρ_ϵ cross sections at conditioning frequencies 0.15 Hz (a) and 0.33 Hz (b). In this figure, solid lines are the total ρ_ϵ from this study, and the dashed lines are from BA18Corr. The darker and lighter shaded regions represent the 95% confidence intervals of ρ_ϵ from these studies respectively [28]. When the 95% confidence intervals don't overlap, there is a statistically significant difference between the ρ_ϵ (at the 0.05 level of significance). At 0.15 Hz conditioning frequency, Figure 9a, these confidence intervals overlap for most other frequencies below about 0.25 Hz. At 0.33 Hz conditioning frequency, Figure 9b, these confidence intervals do not overlap except in the frequencies immediately surrounding 0.33 Hz.

In Figure 10, ρ_ϵ cross sections are shown at conditioning frequencies 0.15, 0.33, and 1.0 Hz. To increase readability, only the 95% confidence bounds on ρ_ϵ are shown. Panels (a) through (d) of Figure 10 show the ρ_ϵ cross sections for the between-event terms, between-site terms, within-site terms, and the total correlation, respectively. As in Figure 9, the darker fill refers

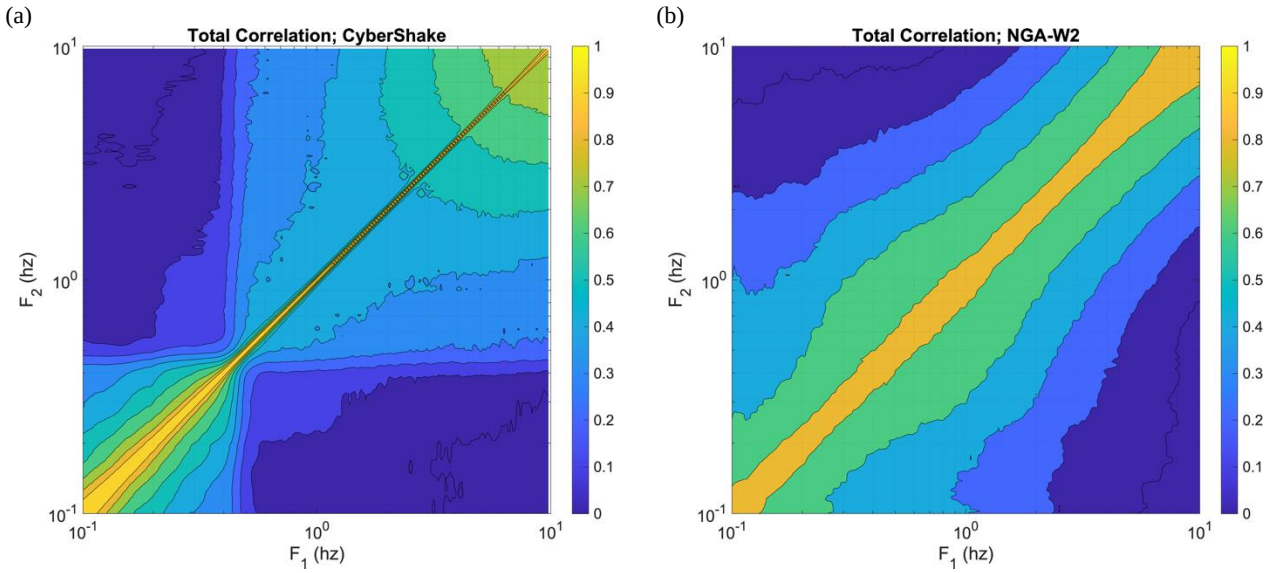

Figure 8: Contours of total ρ_ϵ from (a) this study and (b) BA18Corr.

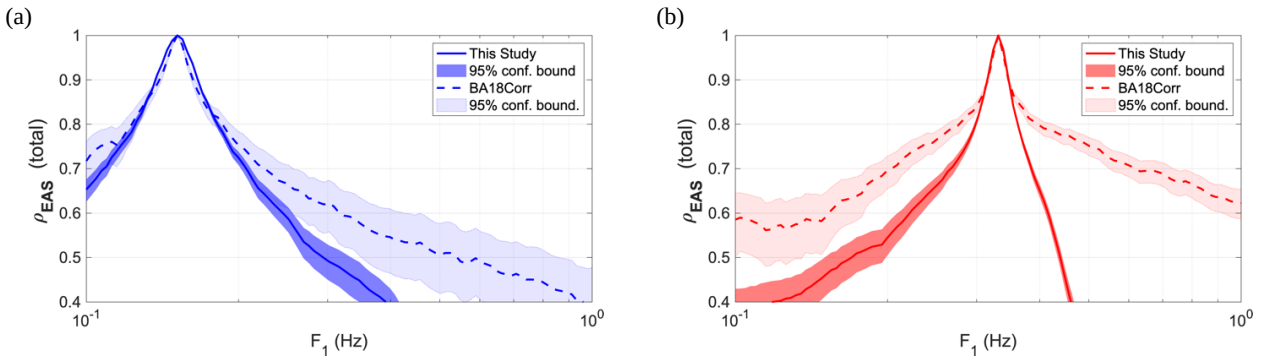

Figure 9: Total ρ_ϵ cross sections at conditioning frequency 0.15 Hz (a) and 0.33 Hz (b). The results from this study are the solid heavy line, with 95% confidence bound shown by the heavy fill. The results from BA18Corr are the dashed line, with light fill representing the 95% confidence bounds.

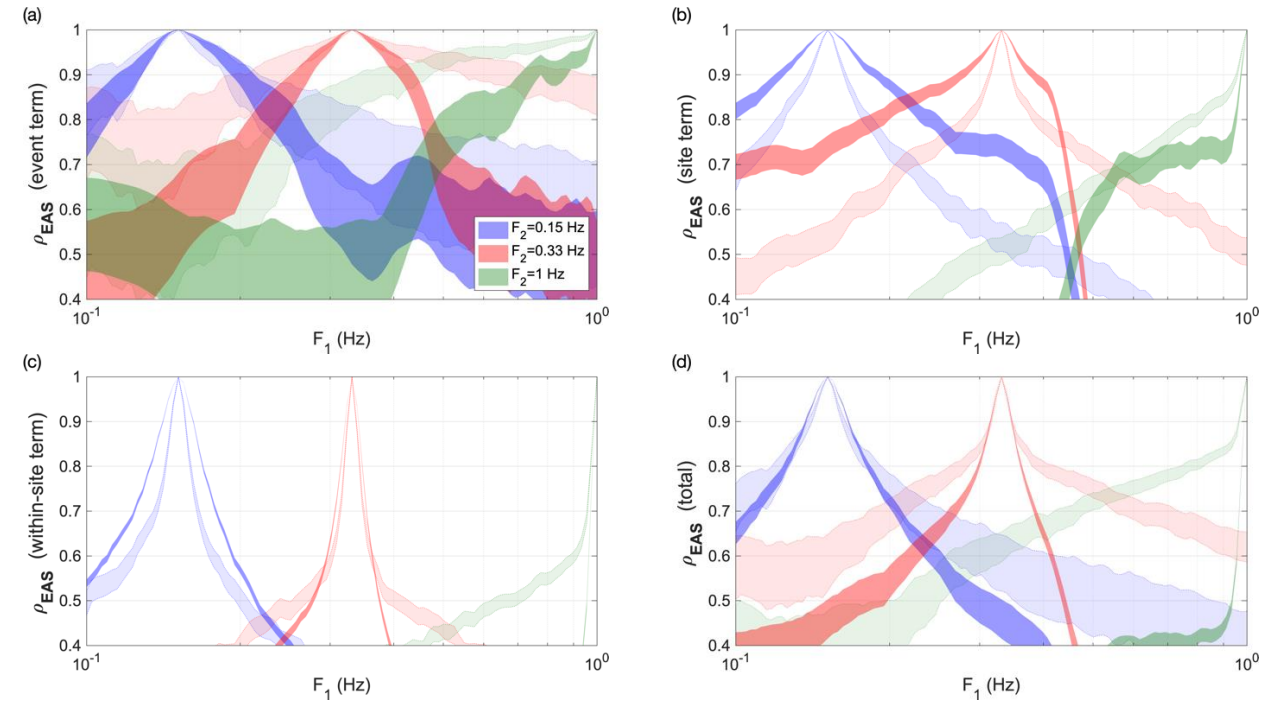

Figure 10: Cross sections of ρ_ϵ 95% confidence bounds at three conditioning frequencies. The darker fill corresponds to this study, and the lighter fill corresponds to BA18Corr. For (a) the between-event component, (b) the site-to-site component, (c) the within-site component, and (d) the total correlation.

to this study, and the lighter, transparent fill corresponds to BA18Corr. Each panel of this figure is discussed in the following paragraphs.

Figure 10a compares the between-event ρ_ϵ cross sections. Out of the three residuals components, these have the widest confidence intervals because there are the fewest samples of the between-event terms (earthquakes) for calculating ρ_ϵ. The between-event ρ_ϵ physically relates to source effects (e.g., stress drop), which drive ground-motions over a broad frequency range and thus lead to relatively broad ρ_ϵ [10]. At conditioning frequency 0.15 Hz, the 95% confidence bands from this study match BA18Corr well over the full frequency range. At conditioning frequency 0.33 Hz, the 95% confidence bands mostly do not overlap and the CyberShake NZ ρ_ϵ are lower than BA18Corr. The match deteriorates as frequencies increase to and beyond the 0.5 Hz transition frequency of the simulations; this is an expected feature of the "stochastic" part of the stochastic method, as discussed previously.

Figure 10b compares the site-to-site ρ_ϵ cross sections. This between-site residual represents the systematic deviation of the observed amplification at a site from the median amplification predicted by the model. Therefore, the site-to-site ρ_ϵ represents the inter-frequency correlation of the systematic site amplification deviations. Of the three residual components, between-site residual component has the largest differences to BA18Corr. Below the 0.5 Hz transition frequency, there is generally higher correlation than BA18Corr. Above 0.5 Hz, there is a steep decline to lower correlation than BA18Corr; again, this is expected of the stochastic method.

The BC20 evaluation of SCEC CyberShake, described further in the next section, found a similar over-prediction of the site-to-site ρ_ϵ at low frequencies. BC20 hypothesized that this was caused by the near surface velocity model, or geotechnical layer (GTL), used in the simulations. In SCEC, the GTL is based on averaged and spatially smoothed geotechnical profiles from borehole measurements, and so the site amplification inherent in the simulations represents these average (smoothed) profiles. Conversely, in a database of recorded ground-motions, the between-site residual represents the characteristic amplification at the site due to amplification of the seismic waves produced by variations of the material properties in these layers. These characteristic amplifications will have spectral peaks over frequency bands which correspond to the resonant frequencies of the profiles. The variability in velocity profiles of the recorded ground-motions is much larger than for the simulations.

The large between-site ρ_ϵ in the cross sections shown in Figure 10b could be the result of similar behavior for CyberShake NZ v20.11. With minimum Vs of 500 m/s the relatively small variability in Vs profiles (compared with the profiles for the NGA-W2 recorded data) increases the LF inter-frequency correlation of the between-site residuals. Another potential source of the mismatch observed in Figure 10b is the effect of low-frequency basin waves. These are surface waves usually produced by the conversion of body waves at the edge of basin into surface waves that propagate across the basin and have very long periods. The basin depth parameter (Z1.0) scaling is meant to represent these effects in an average sense, but if these effects are systematic, they could be mapped into the site terms and impact the low frequency ρ_ϵ.

Figure 10c compares the within-site ρ_ϵ cross sections. The within-site residual component represents the remaining residual after partitioning the random effects for the event and the site. These cross sections are characterized by a steep decay at frequencies very close to the conditioning frequency followed by a relatively flat slope at frequencies farther away from the conditioning frequency. There is generally less inter-frequency correlation from this component of the residuals and the confidence intervals are narrow due to the large number of samples (simulation or recording stations). In Figure 10c, the similarity between this study and BA18Corr is fair at frequencies below 0.4 Hz. At higher frequencies, the within-site ρ_ϵ is very low compared to BA18Corr. The match between the simulations and the recorded data at low frequencies is valuable because, unlike the source- and site- term components of the correlation, it is less clear how to calibrate the within-site ρ_ϵ in the CyberShake NZ v20.11 simulations.

Finally, in Figure 10d the total ρ_ϵ cross sections are compared. The total ρ_ϵ is calculated from Equation 4 of [13]; it is the combination of all the component ρ_ϵ weighted by their respective variances. It follows that the total ρ_ϵ exhibits similar trends as with the other components; with generally larger, satisfactory correlation at frequencies less than about 0.25 Hz, and with lower, inadequate correlation at higher frequencies.

Comparison with SCEC CyberShake

BC20 performed the same inter-frequency correlation analysis described in this article for a subset of the SCEC CyberShake v15.4 simulations [2]. SCEC CyberShake is a computational study analogous to CyberShake NZ, but for the Los Angeles region and with some modelling differences summarized in Table 1. SCEC CyberShake v15.4 is a low-frequency only study ($f < 1$ Hz) using the UCERF2 [29] kinematic sources with the Graves and Pitarka [19] rupture generator. SCEC CyberShake uses an elastic wave propagation simulation to calculate Strain Green tensors around the site of interest, and seismic reciprocity is used to obtain synthetic seismograms [2]. Additionally, the 3D seismic velocity model used in SCEC CyberShake is the CVM-S4.26.M01; this is the tomography improved southern CA model with a 500m resolution, is tri-linearly interpolated, and has minimum Vs of 500 m/s [30]. The CVM-S4.26.M01 near-surface material properties (upper 350m of the model) are based on averages of geotechnical profiles smoothed and interpolated to larger areas [31].

In comparison, CyberShake NZ v20.11 does not use reciprocity and the hybrid simulation method is broadband. Both SCEC and NZ have minimum Vs of 500 m/s, similar velocity model resolutions, and utilize very similar kinematic rupture generators (Table 1). The adoption of forward modelling as opposed to reciprocity will not impact the ground-motions or inter-frequency correlations since both CyberShake methods solve the wave propagation equations for 3D linear, isotropic elastic media. And because our focus is on evaluation of the low frequencies ($f < 1$ Hz), comparisons between the results stemming from both studies are appropriate, noting that the CyberShake NZ v20.11 transition frequency is 0.5 Hz and the SCEC CyberShake v15.4 simulations are fully deterministic (e.g., a finite difference solution to the 3D wave equation) for $f < 1$ Hz with no semi-stochastic method for the higher frequencies.

Figure 11 shows the BC20 results for SCEC CyberShake in a format equivalent to Figure 10, where the darker fill corresponds to BC20, and the lighter, transparent fill corresponds to BA18Corr. The between-event ρ_ϵ are shown in the (a) panels of Figures 10 and 11. At conditioning frequency 0.15 Hz, the 95% confidence bands from both studies match BA18Corr well over the full frequency range. At conditioning frequency 0.33 Hz, the SCEC CyberShake ρ_ϵ is compatible with the empirical ρ_ϵ and is higher and broader than for CyberShake NZ v20.11. The between-event ρ_ϵ physically relates to source effects, and both CyberShake studies use similar low-frequency methodologies (Table 1), therefore the difference in between-event ρ_ϵ between methods may be due to differences in rupture generators. SCEC CyberShake v15.4 uses Graves and Pitarka (2015; GP15) [19] and CyberShake NZ v20.11 uses Graves and Pitarka (2016; GP16) [17].

Table 1: Summary of SCEC CyberShake and CyberShake NZ v20.11 modelling differences.

Technique or Module	**SCEC CyberShake v15.4 (Graves et al., 2011)**	**CyberShake NZ v20.11 (see *Data and Resources*)**
Kinematic sources	UCERF2 (Field et al., 2009)	Stirling et al. (2012) crustal earthquake sources
Kinematic earthquake rupture generator	Graves and Pitarka (2015); GP14.3	Graves and Pitarka (2016); GP15.4
Low-frequency simulation	Comprehensive representation of source and wave propagation physics by solving the 3D viscoelastic wave equation using a staggered-grid finite difference scheme; f < 1.0 Hz; uses seismic reciprocity	Comprehensive representation of source and wave propagation physics by solving the 3D viscoelastic wave equation using a staggered-grid finite difference scheme; f < 0.5 Hz; uses forward simulation
High-frequency simulation	n/a	Stochastic method (a stochastic source radiation pattern and simplified wave propagation with attenuation through a 1D layered velocity model); Parameter values from Graves and Pitarka (2010) for active shallow crustal regions; f > 0.5 Hz
Seismic velocity model	SCEC CVM-S4.26.M01; Small et al. (2017). Minimum Vs of 500 m/s. 500 m resolution.	NZVM2.03; Thomson et al., (2020). Minimum Vs of 500 m/s. 200 m resolution.
Site amplification	n/a	HF amplification only, using Campbell and Bozorgnia (2014)
Additional modifications	n/a	Path duration model and site amplification model; described in Lee et al., (2022)

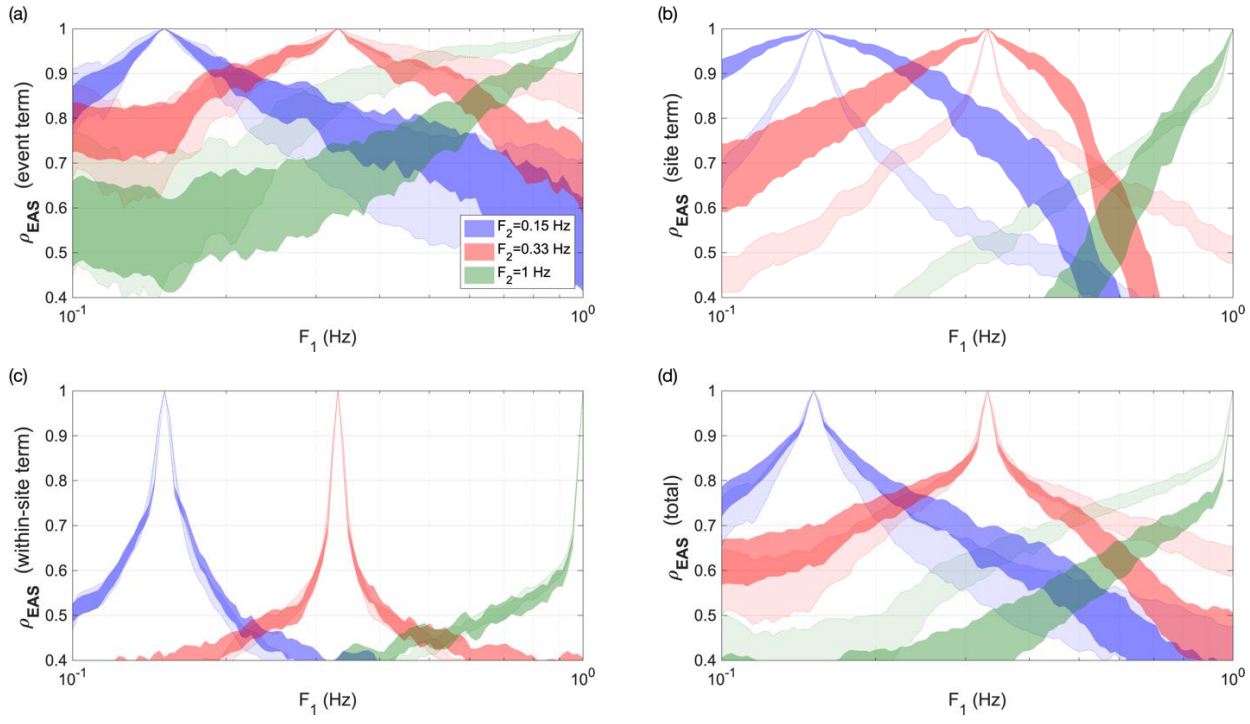

Figure 11: Cross sections of ρ_ϵ 95% confidence bounds at three conditioning frequencies. The darker fill corresponds to the SCEC CyberShake study [11], and the lighter fill corresponds to BA18Corr. For (a) the between-event component, (b) the site-to-site component, (c) the within-site component, and (d) the total correlation.

Both GP15 and GP16 are updates to Graves and Pitarka (2010; GP10) [18] and these updates involve modifications which reduce the coherency of the radiated higher frequency (f > 1 Hz) ground-motions. In both updates, this is achieved through introducing short length scale random perturbations to the rupture time and rise time of each subfault, so that these are not correlated 1:1 with local slip as in GP10. GP16 expanded upon GP15 by developing a more formal correlation structure among the perturbations to these rupture parameters and by including more variability in the background rupture speed (allowing lower rupture speeds additionally reduces the coherence of radiated energy). Although these modifications are designed to reduce the coherence of high-frequency ground-motions, these adjustments appear to also influence the ρ_ϵ at low-frequencies. Additionally, it is possible that path effects due to differences in the seismic velocity models could be influencing the between-event ρ_ϵ, since repeatable path effects were not separated in the residual analysis.

The between-site ρ_ϵ shown in Figure 10b and Figure 11b are comparable. Both the SCEC and NZ v20.11 CyberShake studies have higher correlation for this residual component than the NGA-W2 data. As described previously, the between-site residual represents the characteristic amplification at the site due to amplification of the seismic waves produced by variations of the material properties in these layers. BC20 hypothesized that the relatively small variability in CyberShake Vs profiles (compared with the profiles for the recorded data) increases the inter-frequency correlation of the between-site residuals.

The (c) and (d) panels of Figures 10 and 11 show the within-site and total ρ_ϵ, respectively. Because the total ρ_ϵ is the combination of all the component ρ_ϵ weighted by their respective variances, the total ρ_ϵ exhibits similar trends as with the other correlation components. The total ρ_ϵ from CyberShake NZ v20.11 are generally lower than SCEC CyberShake. There is satisfactory correlation at frequencies less than about 0.25 Hz, and with lower, inadequate correlation at higher frequencies. BC20 concluded that the SCEC CyberShake total ρ_ϵ did not require additional calibration over the frequency range 0.1-0.5 Hz. To improve the total ρ_ϵ for CyberShake NZ, it may be best to modify the source characterization (kinematic rupture generator) to increase the between-event and total ρ_ϵ.

CALIBRATION OF SIMULATIONS

A comprehensive validation of the inter-frequency correlations for a simulation method involves the following phases: (1) evaluation of the inter-frequency correlations to identify deficiencies, (2) refinement of the simulation methodology, (3) repeat of phases 1-2 as needed, and (4) comparison with a set of acceptance criteria to determine the magnitude, distance, and frequency ranges for which the simulations are deemed acceptable for forward use (e.g., [8] for median response spectra). The focus of this study is on phase 1 for the CyberShake NZ v20.11 simulations. The refinement phase in this validation process is challenging because the simulation parameters or techniques which drive the inter-frequency correlations are currently not well understood. This is a developing topic with ongoing research, as summarized here.

Burks and Baker [7] evaluated the inter-frequency correlations of response spectra using three of the SCEC BBP simulation methods from [8], and found that the Composite Source Method [32] had much higher inter-frequency correlation than the other two methods, which indicates that this method could provide some insight on the model features controlling the correlation.

Two studies, [33] and [34] made simple attempts at validation phase 2, simulation refinement. Both took similar approaches to imposing the inter-frequency correlation on simulations as a post-processing procedure. This technique involves using an unmodified (un-correlated) simulation method, imposing the inter-frequency correlation on the FAS of the simulated time history, and performing and inverse Fourier transform to return to the time domain. In our view, this approach may be practical in the short term, because it allows for full calibration of the inter-frequency correlation of the simulations, but it is not a desirable approach in the long term. The inter-frequency correlation observed in the data is an important property of ground-motions, and there are physical reasons for the existence of the correlation. Although the theoretical cause is currently not well understood, the correlation must be introduced in some combination of the earthquake source, the travel path, and the local site response. Therefore, the preferable approach is to incorporate the correlation into seismological models through these foundational elements so that the models most closely represent the earthquake process. When the post-processing method is applied, the physical process built into the finite-fault simulation is ignored.

As a step in the right direction, Song et al. [3] investigated the effect of pseudo-dynamic source models on the inter-frequency correlation of ground-motions by simulating the 1994 Northridge, California, earthquake, using the SCEC BBP. The authors found that the cross correlation between earthquake source parameters (slip, rupture velocity, and peak slip velocity) in the pseudo-dynamic source models significantly affected the inter-frequency correlation of ground-motions in the frequency around 0.5 Hz, whereas this effect was not visible in the other frequency ranges. Additional studies like [3] are needed to identify the causal features of the inter-frequency correlation in simulations, and these should identify techniques for future calibrations and comprehensive validations.

SUMMARY AND CONCLUSIONS

There is increasing recognition that simulations can be utilized in engineering applications, but the simulations require application-specific validation first. If ground-motion simulations are used in seismic fragility and seismic risk analyses, the inter-frequency correlation of normalized residuals (parameter ϵ) is an essential component of the ground-motion simulations for capturing the variability of structural response, and therefore should be validated thoroughly [10].

A comprehensive validation of the inter-frequency correlations for a specified simulation method involves the following phases: evaluation to identify deficiencies, refinement of the simulation methodology, and comparison with a set of acceptance criteria to determine the magnitude, distance, and frequency ranges for which the simulations are deemed acceptable for forward use. This article covers the evaluation phase of the CyberShake NZ v20.11 inter-frequency correlations. We focus on the evaluation component of validation because the parameters or techniques which drive the inter-frequency correlations are not well understood in simulations like CyberShake NZ v20.11. One promising pilot study investigated the effect of pseudo dynamic source models on inter-frequency correlations and found that cross-correlation between source parameters (slip, rupture velocity, and peak slip velocity) affects the inter-frequency correlation around 0.5 Hz [3]. Additional similar studies are needed to identify the causal features of the inter-frequency correlation in simulations, and these should identify techniques for future calibrations and validations.

Compared with the empirical BA18Corr model developed from NGA-W2 data, the 0.1 < f < 0.25 Hz CyberShake NZ v20.11 simulations have a satisfactory level of total inter-frequency correlation, which is a significant improvement from the conclusions of [10] about the SCEC BBP simulations. At frequencies above 0.25 Hz, the CyberShake NZ v20.11

simulations have lower total correlation than the empirical model.

Low inter-frequency correlation is a characteristic of the semi-stochastic simulation technique, which is used by the GP16 hybrid simulation method at frequencies above the transition frequency (0.5 Hz in CyberShake NZ v20.11). The stochastic method begins using the FAS of tapered white noise with random phase angles, which is then normalized to an acceleration amplitude spectrum. This results in no correlation between frequencies [10]. Therefore, low total correlation at frequencies higher than the transition frequency is an inherent limitation of hybrid-method simulations.

The between-event inter-frequency correlation for frequencies larger than about 0.25 Hz would benefit from calibration and would improve the total correlation. For tuning this component, it may be best to modify the source characterization (kinematic rupture generator) by modifying the cross-correlation between source parameters [3]. The observed differences between the SCEC and NZ v20.11 CyberShake between-event correlations identify the influence of minor changes in the rupture generator; it appears that recent methodology updates by [17] which reduce the coherency of the higher frequency (f > 1 Hz) ground-motions relative to previous versions may also reduce the between-event correlations at low frequencies.

The correlation from the between-site residual component (both SCEC and NZ v20.11 CyberShake) also requires calibration moving forward. The correlation from CyberShake NZ v20.11 is significantly higher than the empirical model at frequencies below 0.5 Hz. This may be due to the relative simplicity of the seismic velocity models in the simulations (with less variability in site amplification than the recorded data) and therefore will be more difficult to calibrate than the between-event correlations. Between-site correlations may also be affected by low frequency basin waves mapped into the site terms. In a future study, the cause of large correlation ρ_ϵ for the between-site component of the residuals could be investigated by evaluating the effect of low frequency basin waves on the analysis, or by utilizing results from refined or alternative seismic velocity models. The presence of repeatable path and basin effects could be evaluated through an in-depth residual analysis. Additionally, repeating this analysis with CyberShake NZ simulations using a higher crossover frequency (e.g., 1 Hz or larger) would be informative.

DATA AND RESOURCES

The CyberShake NZ v20.11 ground-motion data were provided by Brendon Bradley and Jason Motha at University of Canterbury (pers. comm., 2021). The QuakeCoRE CyberShake NZ project wiki describes versioning and the computational and scientific model components (https://wiki.canterbury.ac.nz/display/QuakeCore/Cybershake+NZ last accessed April 2023). Regression analyses and graphics production were performed using the numeric computing environment MATLAB (www.mathworks.com last accessed April 2023).

ACKNOWLEDGEMENTS

This research was supported by the QuakeCoRE NZ Centre for Earthquake Resilience and by the Southern California Earthquake Center under Award #20043. The authors are thankful to Dr Brendon Bradley and Dr Jason Motha from University of Canterbury for providing the CyberShake NZ Fourier amplitude spectra; and again to Dr Brendon Bradley and another anonymous reviewer for their thoughtful journal peer reviews of the manuscript.

REFERENCES

1 Bradley BA, Tarbali K, Lee RL, Huang J, Motha J, Bae SE, Polak V, Zhu M, Schill C, Patterson J and Lagrava D (2020). “Cybershake NZ v19.5: New Zealand simulation-based probabilistic seismic hazard analysis”. *NZSEE 2020 Annual Conference*, Paper #110.

2 Graves R, Jordan TH, Callaghan S, Deelman E, Field EH, Juve G, Kesselman C, Maechling P, Mehta G, Okaya D, Small P and Vahi K (2011). “CyberShake: A physics-based seismic hazard model for Southern California”. *Pure and Applied Geophysics*, **168**(3-4): 367–381.

3 Song SG, Causse M and Bayless J (2020). “Sensitivity of inter-frequency correlation of synthetic ground motions to pseudo-dynamic source models”. *Seismological Research Letters*, https://doi.org/10.1785/0220200181

4 Graves R and Pitarka A (2016). “Kinematic ground-motion simulations on rough faults including effects of 3D stochastic velocity perturbations”. *Bulletin of the Seismological Society of America*, **106**: 2136–2153.

5 Stirling M, McVerry G, Gerstenberger M et al. (2012). “National seismic hazard model for New Zealand: 2010 update”. *Bulletin of the Seismological Society of America*, **102**: 1514-1542.

6 Rezaeian S, Stewart J, Luco N and Goulet C. (2022). “Lessons learned from a decade of ground motion simulation validation (GMSV) exercises and a path forward”. *Proceedings of the 12th National Conference in Earthquake Engineering*, Earthquake Engineering Research Institute, Salt Lake City, UT.

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

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

9 Lee RL, Bradley BA, Stafford PJ, Graves RW and Rodriguez-Marek A (2022). “Hybrid broadband ground motion simulation validation of small magnitude active shallow crustal earthquakes in New Zealand.” *Earthquake Spectra*. https://doi.org/10.1177/87552930221109297

10 Bayless J and Abrahamson NA (2018). “Evaluation of the inter-period correlation of ground motion simulations”. *Bulletin of the Seismological. Society of America*, **10**(6): 3413–3430. https://doi.org/10.1785/0120180095

11 Bayless J and Condon S (2020). “*Evaluating the inter-frequency correlation of CyberShake simulations*”. Report to the Southern California Earthquake Center, Report #20043.

12 Bora SS, Scherbaum F, Kuehn N and Stafford, P (2016). “On the relationship between Fourier and response spectra: Implications for the adjustment of *Bulletin of the Seismological. Society of America*, **106**(3): 1235–1253. https://doi.org/10.1785/0120150129

13 Bayless J and Abrahamson NA (2018). “An empirical model for the interfrequency correlation of epsilon for Fourier amplitude spectra”. *Bulletin of the Seismological. Society of America*, **109**(3): 1058–1070. https://doi.org/10.1785/0120180238

14 Ancheta TD, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BS-J, Wooddell KE, Graves RW, Kottke AR, Boore DM, Kishida T and Donahue JL (2014). “NGA-West2 database”. *Earthquake Spectra*, **30**: 989-1005.

15 Cui Y, Poyraz E, Callaghan S, Maechling P, Chen P and Jordan TH (2013). "Accelerating CyberShake calculations on XE6/XK7 platforms of Blue Waters". *Blue Waters and XSEDE Extreme Scaling Workshop*, Aug 15-16, Bouldera.

16 Jordan TH, Callaghan S, Graves RW, Wang F, Milner KR, Goulet CA, Maechling PJ, Olsen KB, Cui Y, Juve G, Vahi K, Yu J, Deelman E and Gill D (2018). "CyberShake models of seismic hazards in Southern and Central California". *Eleventh US National Conference on Earthquake Engineering*, Los Angeles, CA, USA. 11ncee.org/images/program/papers/11NCEE-001458.pdf

17 Graves R and Pitarka A. (2016) "Kinematic ground-motion simulations on rough faults including effects of 3D stochastic velocity perturbations". *Bulletin of the Seismological Society of America*, **106**: 2136–2153.

18 Graves R and Pitarka A (2010) "Broadband ground-motion simulation using a hybrid approach". *Bulletin of the Seismological Society of America*, **100**: 2095–2123.

19 Graves R and Pitarka A (2015). "Refinements to the Graves and Pitarka (2010) Broadband ground-motion simulation method". *Seismological Research Letters*. **86**. https://doi.org/10.1785/0220140101

20 Thomson EM, Bradley BA and Lee RL (2020). "Methodology and computational implementation of a New Zealand Velocity Model (NZVM2.0) for broadband ground motion simulation". *New Zealand Journal of Geology and Geophysics*, **63**: 110-127.

21 Campbell KW and Bozorgnia Y (2014). "NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra". *Earthquake Spectra*, **30**(3): 1087–1115.

22 Kottke A, Abrahamson NA, Boore DM, Bozorgnia Y, Goulet CA, Hollenback J, Kishida T, Ktenidou OJ, Rathje EM, Silva WJ, Thompson EM and Wang X (2021). "Selection of random vibration theory procedures for the NGA-East project and ground-motion modelling". *Earthquake Spectra*, **37**(S1): 1420–1439. https://doi.org/10.1177/87552930211019052

23 Goulet CA, Bozorgnia Y, Kuehn N, Al Atik L, Youngs R, Graves RW and Atkinson G (2021). "NGA-East ground-motion characterization model part I: Summary of products and model development". *Earthquake Spectra*. **37**: 1231-1282. https://doi.org/10.1177/87552930211018723

24 Bayless J and Abrahamson NA (2019) "Summary of the BA18 ground-motion model for Fourier amplitude spectra for crustal earthquakes in California". *Bulletin of the Seismological Society of America*, **109**(5): 2088-2105. https://doi.org/10.1785/0120180238

25 Villani MA and Abrahamson N (2015). "Repeatable site and path effects on the ground-motion sigma based on empirical data from southern California and simulated waveforms from the CyberShake platform". *Bulletin of the Seismological Society of America*, **105**. https://doi.org/10.1785/0120140359

26 Fisher RA (1958). *Statistical Methods for Research Workers*, 13th Ed., Hafner, Edinburgh, London.

27 Boore DM (2003). "Simulation of ground motion using the stochastic method". *Pure Applied Geophysics*, **160**: 635–675.

28 Kutner M, Nachtsheim C, Neter J and Li Q (2005). *Applied Linear Statistical Models*. McGraw-Hill/Irwin, NewYork, 1396 pp.

29 Field E, Dawson T, Felzer K, Frankel A, Gupta V, Jordan T, Parsons T, Petersen M, Stein R, Weldon R and Wills C (2009). "Uniform California Earthquake Rupture Forecast, Version 2 (UCERF 2)". *Bulletin of the Seismological Society of America*. **99**. https://doi.org/10.1785/0120080049

30 Small P, Gill D, Maechling PJ, Taborda R, Callaghan S, Jordan TH, Olsen KB, Ely GP and Goulet C (2017). "The SCEC unified community velocity model software framework". *Seismological Research Letters*, **88**(6): 1539–1552.

31 Magistrale H, Day S, Clayton RW and Graves RW (2000). "The SCEC southern California reference three-dimensional seismic velocity model version 2". *Bulletin of the Seismological Society of America*, **90**(6B): S65–S76.

32 Anderson JG (2015) "The composite source model for broadband simulations of strong ground motions". *Seismological Research Letters*, **86**: 68-74. https://doi.org/10.1785/0220140098

33 Bayless J (2018). "*Inter-Period Correlations of Fourier Amplitude Spectra of Ground-Motions: Modeling, Calibration of Earthquake Simulations, and Significance in Seismic Risk*". PhD Dissertation, Department of Civil and Environmental Engineering, University of California, Davis.

34 Wang N, Takedatsu R, Olsen KB and Day SM (2019). "Broadband ground-motion simulation with inter-frequency correlations". *Bulletin of the Seismological Society of America*, **109**(6): 2437–2446. https://doi.org/10.1785/0120190136