



Eleventh U.S. National Conference on Earthquake Engineering
Integrating Science, Engineering & Policy
June 25-29, 2018
Los Angeles, California

CYBERSHAKE MODELS OF SEISMIC HAZARDS IN SOUTHERN AND CENTRAL CALIFORNIA

Thomas H. Jordan¹, Scott Callaghan², Robert W. Graves³, Feng Wang⁴, Kevin R. Milner², Christine A. Goulet², Philip J. Maechling², Kim Bak Olsen⁵, Yifeng Cui⁶, Gideon Juve⁷, Karan Vahi⁷, John Yu², Ewa Deelman⁷, and David Gill²

ABSTRACT

CyberShake is a high-performance computational platform developed by the Southern California Earthquake Center (SCEC) to produce seismic hazard models from large suites of earthquake simulations. Code optimization and the development of workflow tools on the NCSA *Blue Waters* and OLCF *Titan* supercomputers have substantially reduced the computational costs. A recent workflow efficiently utilized both machines to generate 285 million two-component seismograms for the central California region from which 46 billion intensity measurements were extracted. We summarize the current suite of CyberShake models and apply averaging-based factorization (ABF) to decompose each model into a hierarchy of site, path, directivity, stress-drop, and source complexity effects. We summarize how the strength of the directivity effect decreases with increasing source complexity, and how the site and path effects vary with differences in crustal structure. Comparisons quantified by the ABF variances indicate that simulation-based hazard models can potentially reduce σ_T , the total unexplained variability in current ground-motion prediction equations, by as much as one-third.

¹Dept. of Earth Sciences, University of Southern California, Los Angeles, CA 90089-0740 (email: tjordan@usc.edu)

²Southern California Earthquake Center, University of Southern California, Los Angeles, CA 90089-0742

³U.S. Geological Survey, Pasadena, CA 91106

⁴AIR Worldwide Inc., Boston, MA 02116

⁵San Diego State University, San Diego, CA 92182

⁶San Diego Supercomputer Center, San Diego, CA 92093

⁷USC Information Sciences Institute, Los Angeles, CA 90292

CyberShake Models of Seismic Hazards in Southern and Central California

Thomas H. Jordan¹, Scott Callaghan², Robert W. Graves³, Feng Wang⁴, Kevin R. Milner², Christine A. Goulet², Philip J. Maechling², Kim Bak Olsen⁵, Yifeng Cui⁶, Gideon Juve⁷, Karan Vahi⁷, John Yu², Ewa Deelman⁷, and David Gill²

ABSTRACT

CyberShake is a high-performance computational platform developed by the Southern California Earthquake Center (SCEC) to produce seismic hazard models from large suites of earthquake simulations. Code optimization and the development of workflow tools on the NCSA *Blue Waters* and OLCF *Titan* supercomputers have substantially reduced the computational costs. A recent workflow efficiently utilized both machines to generate 285 million two-component seismograms for the central California region from which 46 billion intensity measurements were extracted. We summarize the current suite of CyberShake models and apply averaging-based factorization (ABF) to decompose each model into a hierarchy of site, path, directivity, stress-drop, and source complexity effects. We summarize how the strength of the directivity effect decreases with increasing source complexity, and how the site and path effects vary with differences in crustal structure. Comparisons quantified by the ABF variances indicate that simulation-based hazard models can potentially reduce σ_T , the total unexplained variability in current ground-motion prediction equations, by as much as one-third.

Introduction

Advanced applications of probabilistic seismic hazard analysis (PSHA) in California combine fault-based earthquake rupture forecasts (ERFs) with site-specific ground motion prediction equations (GMPEs) to estimate long-term seismic shaking probabilities. Both PSHA components have been refined through comprehensive, iterated studies, including the Uniform California Earthquake Rupture Forecast (UCERF) Project [1,2] and the Next Generation Attenuation (NGA) Project [3,4]. These collaborative efforts have improved our understanding of seismic hazards, but the uncertainties in PSHA forecasts remain high. For example, when GMPEs are applied to recordings of well-characterized earthquake sources, the logarithmic residuals between the observed and predicted values of peak ground acceleration (PGA) scatter with a total standard deviation σ_T of about 0.6 in natural-log units [5]. Accounting for these

¹Dept. of Earth Sciences, University of Southern California, Los Angeles, CA 90089-0740 (email: tjordan@usc.edu)

²Southern California Earthquake Center, University of Southern California, Los Angeles, CA 90089-0742

³U.S. Geological Survey, Pasadena, CA 91106

⁴AIR Worldwide Inc., Boston, MA 02116

⁵San Diego State University, San Diego, CA 92182

⁶San Diego Supercomputer Center, San Diego, CA 92093

⁷USC Information Sciences Institute, Los Angeles, CA 90292

unexplained discrepancies is a major goal of PSHA research, because even modest reductions in this measure of uncertainty can translate into order-of-magnitude differences in the hazard predicted at the high intensity thresholds relevant to critical-facility design [5,6]. As Strasser et al. [5] have emphasized, however, this unexplained variability has not declined during four decades of GMPE development, despite the substantial increases in strong-motion recordings, the sophistication of the data analysis, and the statistical complexity of the models. The Next Generation Attenuation-West (NGA-W) projects exemplify this persistence. In the 2008 study (NGA-W1), σ_T for large-event ($M \geq 7$) PGA varied among the models from 0.52 to 0.56 [7]; in the 2014 study (NGA-W2), it varied from 0.57 to 0.65 [8]. Other intensity measures, including the long-period response spectra relevant to this study, show residuals of similar magnitude.

Much of the unexplained variability in ground motions can be attributed to unmodeled complexities in the source-excitation and wave-propagation processes [5,9]. One approach is to introduce additional explanatory variables into the GMPEs, such as anomalies specific to individual sources, sites, and paths, and then attempt to estimate these parameters directly from observations [10-12]. A difficulty with this approach—indeed, with any purely empirical GMPE methodology—is that the proliferation of parameters can quickly overwhelm the information gain from new observations, so that any decrease in a model’s rendition of aleatory variability is offset by an increase in its epistemic error. This aleatory-epistemic tradeoff limits the net reduction in σ_T , which is a measure of both types of uncertainty.

The alternative, taken here, is to harness the explanatory power of rupture and wave physics through three-dimensional (3D) earthquake simulations. Full-3D simulations are capable of modeling much of what we know about earthquake processes, including source complexity and 3D path effects [13]. Kinematic rupture models have been developed that are more faithful to the space-time correlations derived from observations and dynamic rupture models [14-16]. Full-3D tomography (F3DT) has refined the 3D crustal models of Southern California [17,18], assimilating large datasets of earthquake waveforms and ambient-field correlograms. At low frequencies (< 1 Hz), these SCEC community velocity models (CVMs) replicate observations of seismic wave propagation from earthquakes not used in the inversions [19, 20].

Recent advances in seismology and high-performance computing now make it feasible to generate sets of full-3D earthquake simulations that are large enough ($> 10^8$ seismograms) to sample the probability distributions that describe the variability of the source-excitation and wave-propagation processes. In this paper, we present simulation-based PSHA models computed for the Los Angeles and Central California regions.

CyberShake Computational Platform and Workflow

CyberShake is a high-performance computational platform developed by the Southern California Earthquake Center (SCEC) for producing seismic hazard models from large suites of earthquake simulations [21]. The CyberShake workflow (Figure 1) is “site-oriented” in the sense that the computations are bundled by the PSHA location of interest. Each iteration synthesizes horizontal-component seismograms at a single site for many rupture variations, derived by varying hypocenter location and slip distribution for the Uniform California Earthquake Rupture Forecast, Version 2 (UCERF2) fault ruptures within 200 km of the site [21, 22]. Each seismogram is computed by a single quadrature of the space-time slip function with the strain-displacement Green’s tensor (SGT) for the site [23]. This formulation, based on seismic reciprocity, is computationally efficient when the number of rupture variations M is substantially

greater the number of sites N ; i.e., it requires only $2N$ simulations per CyberShake model, compared to the M simulations needed if each source were to be simulated separately [21].

In the CyberShake models described here, the number of rupture variations, $M \approx 400,000$, is much larger than the number of sites, $N \approx 400$; hence, reciprocity reduces the wave-propagation computation (which dominates the overall expense) by a factor of about 500. Multiple variations are needed to sample the conditional hypocenter and slip distributions for each of the $\sim 7,000$ UCERF2 ruptures with moment-magnitudes $M_W \geq 6.0$ [21, 24]. The conditional hypocenter distribution is assumed to be uniform along strike, and the conditional slip distribution is that sampled by Graves-Pitarka kinematic rupture realizations.

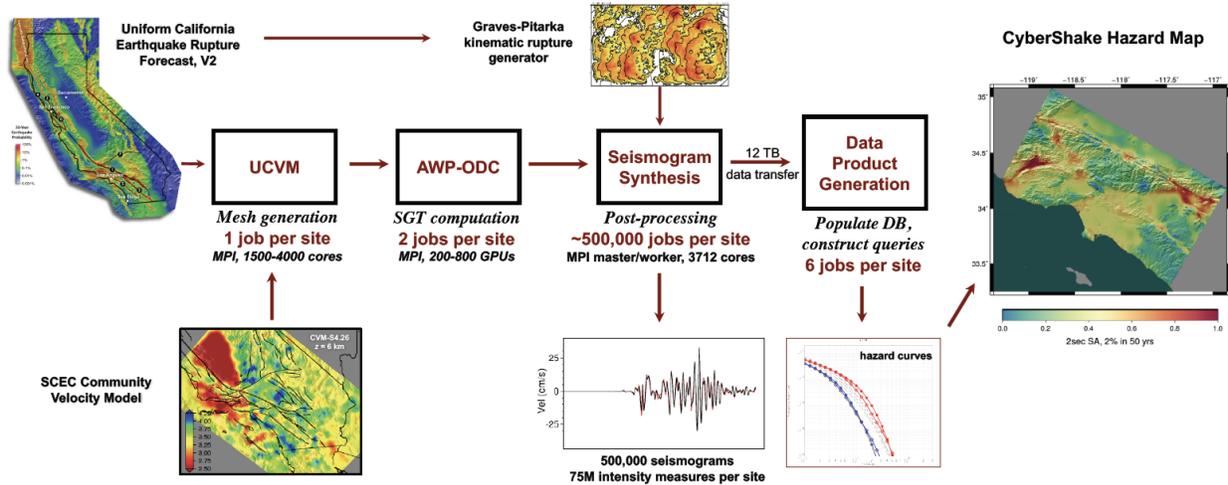


Figure 1. The CyberShake computational workflow, illustrated here for CS-LA15.4.

The UCERF2 fault model and the CVM are registered onto regular mesh using the Unified California Velocity Model (UCVM) software [25]. The mesh spacing is adjusted to sample the smallest wavelength at the maximum seismic frequency f_{\max} at about eight nodes per wavelength. In the current CyberShake implementation, the SGT for the horizontal components at each site is calculated by the finite-difference anelastic wave-propagation code, AWP-ODC, which has been highly optimized for massively parallel CPU and GPU machines [26, 27]. These wavefields are captured on all mesh points corresponding to the UCERF2 rupture surfaces, and seismograms are synthesized by quadrature of the SGT with realizations from the Graves-Pitarka conditional slip distribution. Various ground motion intensities, such as the RotD50 and RotD100 spectral response, are calculated from the seismograms and stored in a database. Using the OpenSHA toolkit [28], a user can then aggregate the CyberShake intensities and their UCERF2 rupture probabilities into hazard curves and hazard maps, or disaggregate the site-specific hazard into its dominant earthquakes, for which CyberShake produces entire time series.

Each regional CyberShake hazard model requires the synthesis of hundreds of millions of seismograms and the management of almost a petabyte of data. The CyberShake platform makes such large calculations feasible by employing workflow middleware [22], including Pegasus-WMS [29], HTCondor [30], and the Globus Toolkit [31], to automate remote job submissions, orchestrate millions of real-time job executions, manage data and provenance, and provide for error recovery. Using this workflow-based approach, CyberShake simulations have been performed on nine separate supercomputer systems, utilizing a total of more than 10^8 core-hours.

CyberShake hazard modeling has been enabled by many code enhancements and

workflow optimizations, including the parallelization of serial processing stages, the migration of the SGT code to GPUs, the refactoring of the seismogram synthesis code, and the development of new workflow technology that automates remote job submissions to systems with authentication restrictions [32]. Improvements to both the code and overall workflow throughput have enabled CyberShake to be run on the largest open-science systems, including the GPU-enabled supercomputers *Blue Waters*, operated by the National Center for Software Applications (NCSA), and *Titan*, operated by the Oak Ridge Leadership Computing Facility (OLCF).

CyberShake Models

Here we present CyberShake hazard models for the Los Angeles (CS-LA) and Central California (CS-CC) regions computed from three types of input models: UCERF2 [1]; the pseudo-dynamic rupture generators of Graves & Pitarka, denoted GP-07 [14], GP-10 [15], and GP-14 [16]; and the 3D CVMs curated by SCEC.

Table 1. CyberShake models for Los Angeles (LA) and Central California (CC) regions.

Study ID	Model ID	f_{\max} (Hz)	Rupture Generator	Velocity Model	SGT Code	# Sites
17.3	CS-CC17.3a	1.0	GP-14	CCA06-3D	AWP-ODC-SGT-GPU	438
	CS-CC17.3b	1.0	GP-14	CCA06-1D	AWP-ODC-SGT-GPU	438
15.12	CS-LA15.12	1.0, 10*	GP-14	CVM-S4.26	AWP-ODC-SGT-CPU	336
15.4	CS-LA15.4	1.0	GP-14	CVM-S4.26	AWP-ODC-SGT-GPU	336
14.2	CS-LA14.2a	0.5	GP-10	CVM-S4.26	AWP-ODC-SGT-GPU	286
	CS-LA14.2b	0.5	GP-10	CVM-BBP-1D	AWP-ODC-SGT-CPU	286
	CS-LA14.2c	0.5	GP-10	CVM-H11.9	AWP-ODC-SGT-GPU	286
	CS-LA14.2d	0.5	GP-10	CVM-S4.26	AWP-ODC-SGT-CPU	286
13.4	CS-LA13.4a	0.5	GP-10	CVM-S4.0	RWG v3.0.3	283
	CS-LA13.4b	0.5	GP-10	CVM-H11.9-GTL	RWG v3.0.3	283
	CS-LA13.4c	0.5	GP-10	CVM-S4.0	AWP-ODC-SGT-CPU	283
	CS-LA13.4d	0.5	GP-10	CVM-H11.9-GTL	AWP-ODC-SGT-CPU	283
1.0	CS-LA1.0	0.5	GP-07	CVM-S4.0	RWG v1.16.3	272

*1 Hz deterministic, 10 Hz stochastic

The first published CyberShake hazard model, CS-LA1.0, was computed on the *Ranger* supercomputer of the Texas Advanced Computing Center (TACC) and comprised horizontal-component synthetic seismograms for 415,000 UCERF2 rupture variations ($M_W \geq 6.0$) at 272 sites in the Los Angeles region at frequencies up to 0.5 Hz [21]. It used the GP-07 rupture generator, the CVM-S4.0 crustal structure [32], and a 4th-order staggered-grid finite-difference (FD) code [34]. Following a preliminary phase of experimentation with different models and codes, we conducted a series of CyberShake studies, each producing one or more hazard models, designated by the ‘year.month’ in which the study was initiated (Table 1).

Study 13.4. The scientific goals of this study, begun in April, 2013, were threefold. The first was to cross-verify the hazard simulations using different SGT codes on different supercomputers. We found that the models run using the original FD code and those run using

the highly optimized FD code AWP-ODC-SGT agreed to within the expected numerical accuracy. Numerical efficiencies allowed us to reduce the total wall-clock time (“makespan”) per CyberShake model by about a factor of four relative to the original CS-LA1.0 calculation. The second goal was to assess the differences in source directivity caused by substituting the GP-10 rupture generator for the original GP-07 version. GP-10 produces more complex rupture patterns than GP-07, in better agreement with dynamic rupture simulations [15]. Comparisons of model CS-LA13.4a, which used GP-10, with that of CS-LA1.0, which used GP-07, confirmed that the more complex ruptures reduced the constructive interference and thus decreased the amplitudes of the directivity pulses [24]. The third goal was to investigate the hazard differences obtained from two different 3D crustal structures, CVM-S4.0 and CVM-H11.9-GTL. In the latter model, the H11.9 structure was augmented with a shallow (300 m) geotechnical layer derived from maps of near-surface (v_{S30}) shear velocities [18]. The hazard maps for 3-s spectral acceleration at a probability of exceedance of 2% in 50 years are compared in Figure 2. The largest hazard differences are associated with the depth and extent of sediments in the major sedimentary basin of the Los Angeles region.

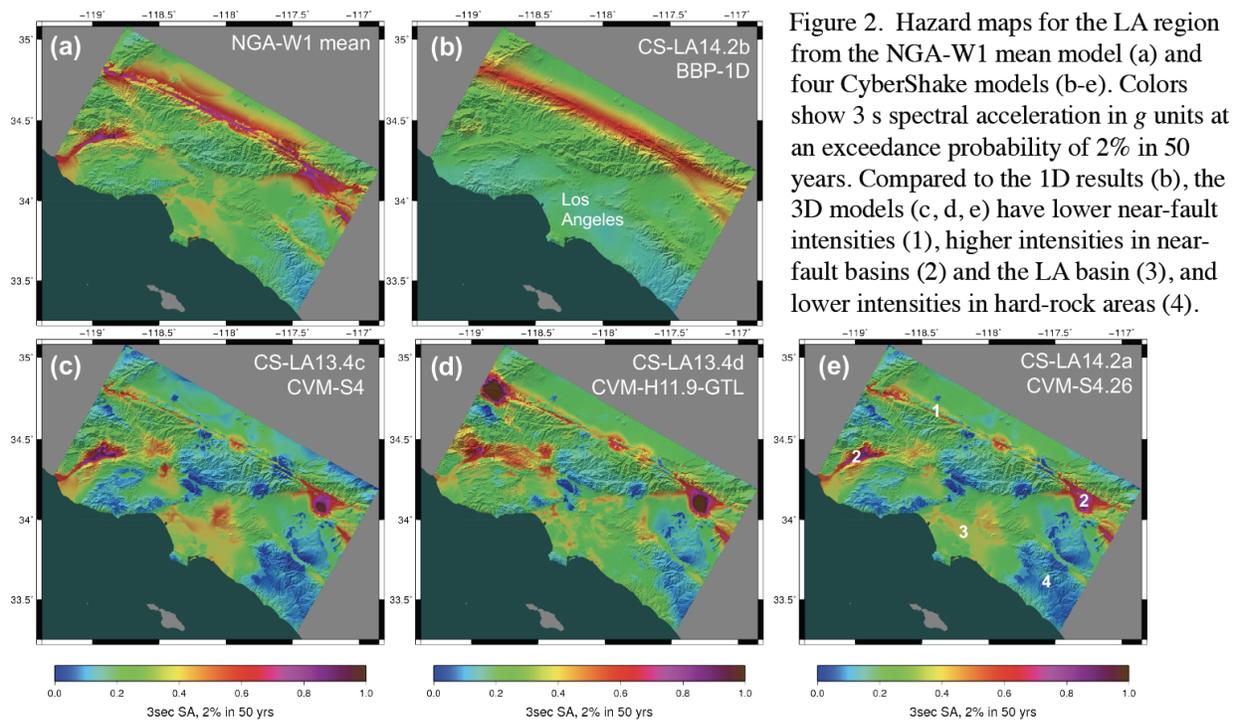


Figure 2. Hazard maps for the LA region from the NGA-W1 mean model (a) and four CyberShake models (b-e). Colors show 3 s spectral acceleration in g units at an exceedance probability of 2% in 50 years. Compared to the 1D results (b), the 3D models (c, d, e) have lower near-fault intensities (1), higher intensities in near-fault basins (2) and the LA basin (3), and lower intensities in hard-rock areas (4).

Study 14.2. In February, 2014, at our request, NCSA made a policy change regarding workflow management software that allowed us to exploit the heterogeneous architecture of *Blue Waters* much more effectively, reducing the CyberShake makespan per model to about 86 hours (~3.6 days). The four CyberShake runs of Study 14.2 were designed to verify a GPU implementation of the AWP-ODC-SGT code [27], which proved to be 6.5 times more efficient than the CPU implementation, and to compare hazard models based on three velocity structures, the original CVM-H9.11 model (without the GTL), the new F3DT model CVM-S4.26 [17], and the 1D depth-variable structure CVM-BBP-1D. Hazard maps from the latter two are compared in Figure 2 with the other CS-LA 3D models, as well as the mean of four NGA-W1 GMPEs used in the 2008 update of the National Seismic Hazard Model [35].

Study 15.4. In April 2015, we extended the frequency range to $f_{\max} = 1$ Hz. We retained the CVM-S4.26 structure but updated the rupture generator to GP-14. In this version, which had been tested on SCEC’s Broadband Platform [16], Graves & Pitarka added stochastic perturbations to the correlation structure for rise time and rupture speed parameterization, further increasing slip complexity and reducing the coherence of radiated energy at 1 Hz. A redesign of the seismogram synthesis code to handle GP-14 and more complex conditional hypocenter distributions reduced I/O by 99.9%. Study 15.4 was the largest to date, utilizing 37.6 million core-hours in 39 days on *Blue Waters* and *Titan* and generating over a petabyte of data.

Study 15.12. This study, performed on *Blue Waters*, augmented the CS-LA15.4 time series with stochastic components in the band 1-10 Hz. These high-frequency components were generated using the Graves-Pitarka methodology implemented on the SCEC Broadband Platform [16]. Owing to their broad bandwidth (0-10 Hz), the seismograms and intensity measures derived from CS-LA15.12 may be of particular interest to earthquake engineers. In addition to the acceleration measures, we also computed and stored duration measures based on Arias Intensity.

Study 17.3. In this study, we migrated CyberShake to Central California (Figure 3). We used the GP-14 rupture generator and two velocity structures: a three-dimensional model, CCA06-3D, and a laterally homogeneous, one-dimensional model, CCA06-1D, obtained by averaging CCA06-3D over its land area. CCA06-3D was derived from a 3D starting model by six F3DT iterations to fit about 12,000 ambient-field correlograms [36]. These inversions were executed on the *Mira* supercomputer of the Argonne Leadership Computing Facility (ALCF). A new workflow tool developed for this study, rvGAHP [32], enabled execution of end-to-end CyberShake workflows on *Titan* for the first time. Using a total of 21.6 million core-hours on both *Blue Waters* and *Titan*, we simulated 285 million two-component seismograms at 476 sites in a makespan of 31 days, extracting 46 billion intensity measures.

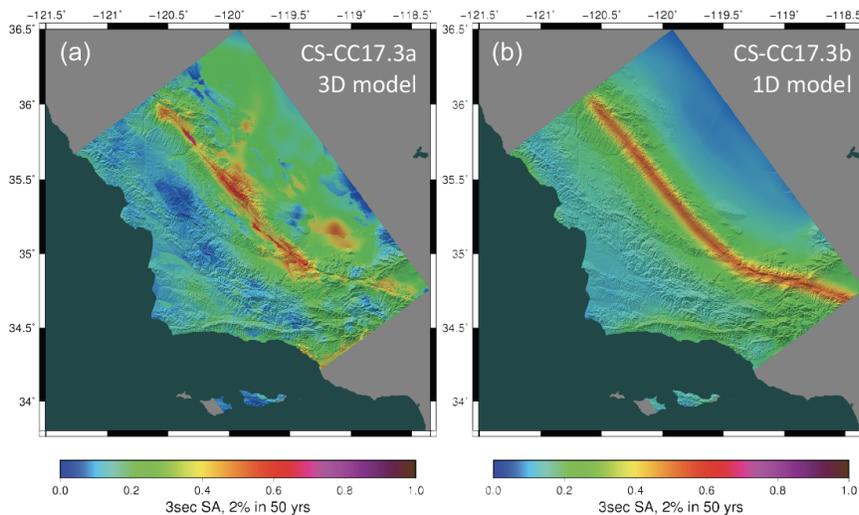


Figure 3. Hazard maps for the Central California region from CyberShake Study 17.3. Colors show 3 s spectral acceleration in g units at an exceedance probability of 2% in 50 years. (a) is based on the CCA06-3D model [36] and (b) is based on the land-area average of CCA06-1D.

Averaging-Based Factorization of CyberShake Models

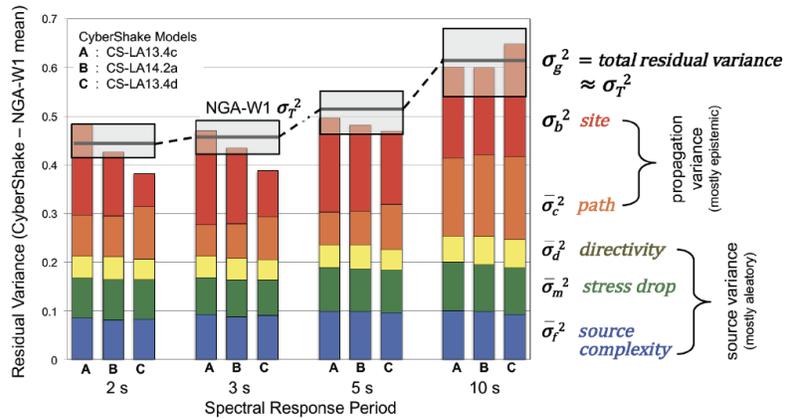
GMPEs comprise factors representing the dependence of the shaking intensity Y on explanatory variables such as magnitude, distance, site conditions, basin depth, and rupture directivity [3, 4]. This type of model-based factorization is not available for CyberShake, but the variability of ground motions can nevertheless be separated into well-defined components using the technique

of averaging-based factorization (ABF) [24]. CyberShake organizes the simulations into a five-level tree structure: each site is characterized by many ruptures, each rupture has many rupture variations, and each rupture variation first samples magnitude and hypocenter distributions, which are conditional on the rupture, and then samples a slip distribution, which is also conditional on the hypocenter. ABF averages the logarithms of the shaking intensities, $G = \ln Y$, at each level of this simulation hierarchy, starting at the lowest (slip distribution), and it moves the mean values up to the next level of the hierarchy. This process uniquely and exactly separates G into six terms: a constant mean hazard level, $A \equiv \langle G \rangle$, and five terms that are stochastic samples of site (B), path (C), directivity (D), stress-drop (M), and source-complexity (F) effects.

These effects are, by construction, mutually uncorrelated, so we can simply sum the average variance at each level to get the total variance $\sigma_g^2 \equiv \langle (G - A)^2 \rangle = \sigma_B^2 + \bar{\sigma}_C^2 + \bar{\sigma}_D^2 + \bar{\sigma}_M^2 + \bar{\sigma}_F^2$ (see [24] for precise definitions of these terms). The individual variances measure the effect size. For example, at 3-s period, the size of the directivity effect in CS-LA1.0 ($\bar{\sigma}_D^2 = 0.210$) is much larger than in CS-LA13.4a ($\bar{\sigma}_D^2 = 0.073$). The only difference between the two models is the rupture generator; therefore, the variance reduction (65%) quantifies how the strength of the directivity effect decreases with increasing source complexity, in this case from GP-07 (less complex) to GP-10 (more complex). A similar comparison between GP-10 and the even more complex GP-14 rupture generator, obtained from CS-LA14.2a ($\bar{\sigma}_D^2 = 0.082$) and CS-LA15.4 ($\bar{\sigma}_D^2 = 0.076$), gives a much smaller variance reduction ($\sim 7\%$), indicating that the GP-14 refinements to the source complexity are mainly at scales smaller than the seismic wavelengths at 3-s period.

If a reference model and a target model share the same simulation hierarchy, we can subtract the logarithmic intensities of the first from the second, $g \equiv G - \tilde{G}$, and apply ABF to the residual g . Owing to linearity, the residual variances (denoted by lower-case letters; e.g. σ_b^2) remain uncorrelated. In particular, a reference model can be constructed by GMPE simulations of the UCERF2 rupture set; the residual variances then represent the misfit of the GMPE to the CyberShake model. This procedure allows GMPE basin-effect terms to be directly compared with the CyberShake basin effects, for example. Refinements of the 3D velocity structures have reduced the magnitude of the basin effects from the CyberShake models based on CVM-S4.0 [24], but models with the more accurate CVMs, such as CS-LA15.4, still show basin amplifications that are larger than the NGA-W2 GMPEs at periods greater than about 3 s [37].

Figure 4. Residual variance (natural-log units) from the subtraction of the NGA-W1 mean model from three CyberShake models at periods of 2, 3, 5, and 10 s. Total residual variances σ_g^2 (colored bars) approximate the NGA-W1 values of σ_T^2 (gray-shaded range). ABF partitions the total residual variance into independent contributions from site (red), path (orange), directivity (yellow), stress-drop (green), and source-complexity (blue) effects.



In Figure 4, we plot the residual variances obtained by subtracting the NGA-W1 mean model from three CS-LA models; the latter were computed using the same rupture simulator (GP-10) but different 3D velocity models. The total residual variance σ_g^2 increases with period, and its mean value at each period is approximately equal to the mean value of σ_T^2 for the NGA-W1 GMPEs. In other words, the shaking intensities predicted by NGA-W1 differ from those predicted by CyberShake to the same degree that they differ from the intensities observed on real seismograms. The near-equality $\sigma_g^2 \approx \sigma_T^2$ across multiple frequency bands indicates the realism of the CyberShake simulations.

ABF partitions σ_g^2 into its component variances, shown by the different colors in Figure 4. The residual variances corresponding to directivity (σ_d^2), stress drop (σ_m^2), and source complexity (σ_f^2) are largely aleatory in the sense that these types of variability are intrinsically difficult to predict by source modeling. On the other hand, about half of residual variance come from the two terms associated with site effects (σ_b^2) and path effects (σ_c^2). This variability is largely due to epistemic uncertainties in the deterministic modeling of seismic wave propagation. Reducing these uncertainties by improving 3D velocity structures thus has the potential for reducing σ_T^2 by as much as one-half and σ_T by almost one-third, which could modify the site-specific exceedance probabilities at high shaking intensities by orders of magnitude. This model-based inference is consistent with recent empirical studies [10-12].

Conclusions

CyberShake has been developed into an efficient computational platform capable of producing probabilistic seismic hazard models up to $f_{\max} = 1$ Hz from very large suites of earthquake simulations. Here we have summarized the CyberShake models for the Los Angeles and Central California regions (Table 1). Comparison of models with 1D and 3D velocity heterogeneities show how wave scattering acts to reduce shaking intensities near major faults and amplify them in sedimentary basins (Figures 2 & 3). Simulations with the Graves-Pitarka rupture generators quantify how the strength of the directivity effect decreases with increasing source complexity. Averaging-based factorization of the CyberShake models indicate that simulations can potentially reduce the NGA σ_T by as much as one-third, by accounting for path and site effects that are otherwise treated as ergodic (Figure 4). The SCEC Committee on the Utilization of Ground Motion Simulations is currently investigating how the CyberShake results can be combined with GMPEs to improve long-period hazard estimates in the LA region [37].

Acknowledgments

Software development was supported by NSF awards ACI-1148493, ACI-1450451, and EAR-1349180, the Blue Waters Sustained-Petascale Computing Project, and the Pacific Gas and Electric Company. Computational support was provided by NSF PRAC award OAC-1713792 and by the Department of Energy's INCITE program. We used resources of the Argonne Leadership Computing Facility, supported under DOE contract DE-AC02-06CH11357, and the Oak Ridge Leadership Computing Facility, supported under DOE contract DE-AC05-00OR22725. Computational support was also provided by the XSEDE program under NSF grant ACI-1053575 and by the Center for High Performance Computing of the University of Southern California. SCEC is funded by NSF Cooperative Agreement EAR-1033462 and USGS Cooperative Agreement G12AC20038. This is SCEC Publication Number 7994.

References

1. Field EH, Dawson TE, Felzer KR, Frankel AD, Gupta V, Jordan TH, Parsons T, Petersen MD, Stein RS, Weldon II RJ, Wills CJ. Uniform California Earthquake Rupture Forecast, Version 2 (UCERF 2). *Bull. Seismol. Soc. Am.* 2009; **99**: 2053-2107.
2. Field EH, Jordan TH, Page MT, Milner KR, Shaw BE, Dawson TE, Biasi GP, Parsons T, Hardebeck JL, Michael AJ, Weldon II RJ, Powers PM, Johnson KM, Zeng Y, Bird P, Felzer KR, van der Elst N, Madden C, Arrowsmith R, Werner MJ, Thatcher WR. A Synoptic View of the Third Uniform California Earthquake Rupture Forecast (UCERF3). *Seismol. Res. Lett.* 2017; **88**: 1-9.
3. Power, M, Chiou B, Abrahamson NA, Bozorgnia Y, Shantz T, Roblee C. An overview of the NGA project. *Earthq. Spectra* 2008; **24**: 3-21.
4. Bozorgnia, Y, Abrahamson NA, Al Atik L, Ancheta TD, Atkinson GM, Baker JW, Baltay A, Boore DM, Campbell KW, Chiou BS-J, Darragh R, Day S, Donahue J, Graves RW, Gregor N, Hanks T, Idriss IM, Kamai R, Kishida T, Kottke A, Mahin SA, Rezaeian S, Rowshandel B, Seyhan E, Shahi S, Shantz T, Silva W, Spudich P, Stewart JP, Watson-Lamprey J, Wooddell K, Youngs R. NGA-West2 Research Project, *Earthq. Spectra* 2014; **30**: 973-987.
5. Strasser FO, Abrahamson NA, Bommer JJ. Sigma: issues, insights, and challenges. *Seismol. Res. Lett.* 2009; **80**: 40-56.
6. Bommer JJ, Abrahamson NA. Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates? *Bull. Seismol. Soc. Am.* 2006; **96**:1967-1977.
7. Abrahamson N, Atkinson G, Boore D, Bozorgnia Y, Campbell K, Chiou B, Idriss IM, Silva W, Youngs R. Comparisons of the NGA Ground-Motion Relations. *Earthq. Spectra* 2008; **24**: 45-66.
8. Gregor N, Abrahamson NA, Atkinson GM, Boore DM, Bozorgnia Y, Campbell KW, Chiou BS-J, Idriss IM, Kamai R, Seyhan E, Silva W, Stewart JP, Youngs R. Comparison of NGA-West2 GMPEs. *Earthq. Spectra* 2014; **30**: 1179-1197.
9. Day SM, Graves RW, Bielak J, Dreger D, Larsen S, Olsen KB, Pitarka A, Ramirez-Guzman L. Model for basin effects on long-period response spectra in southern California. *Earthq. Spectra* 2008; **24**: 257-277.
10. Al Atik L, Abrahamson N, Bommer JJ, Scherbaum F, Cotton F, Kuehn N. The variability of ground-motion prediction models and its components. *Seismol. Res. Lett.* 2010; **81**: 794-801.
11. Anderson JG, Uchiyama Y. A methodology to improve ground motion prediction equations by including path corrections 2011; *Bull. Seismol. Soc. Am*; **101**: 1822-1846.
12. Baltay AS., Hanks TC, Abrahamson NA. Uncertainty, variability, and earthquake physics in ground-motion prediction equations. *Bull. Seismol. Soc. Am.* 2017; **107**: 1754-1772.
13. Day SM., Graves RW, Bielak J, Dreger D, Larsen S, Olsen KB, Pitarka A, Ramirez-Guzman L. Model for basin effects on long-period response spectra in southern California. *Earthq. Spectra* 2008; **24**: 257-277.
14. Graves R, Pitarka A. Broadband time history simulation using a hybrid approach, *Proc. 13th World Conference on Earthquake Engineering* 2004; Paper 1098, Vancouver, Canada.
15. Graves RW, Pitarka A. Broadband ground-motion simulation using a hybrid approach, *Bull. Seismol. Soc. Am.* 2010; **100**: 2095-2123.
16. Graves R, Pitarka A. Refinements to the Graves and Pitarka (2010) Broadband ground-motion simulation method. *Seismological Research Letters* 2015; **86**: 75-80.
17. Lee E-J, Chen P, Jordan TH, Maechling PB, Denolle MAM, Beroza GC. Full-3D tomography for crustal structure in Southern California based on the scattering-integral and the adjoint-wavefield methods. *J. Geophys. Res.* 2014; **119**: 6421-6451.
18. Shaw JH, Plesch A, Tape C, Suess MP, Jordan TH, Ely G, Hauksson E, Tromp J, Tanimoto T, Graves R, Olsen K, Nicholson C, Maechling PJ, Rivero C, Lovely P, Brankman CM, Munster J. Unified Structural Representation of the southern California crust and upper mantle. *Earth Planet. Sci. Lett.* 2015; **415**: 1-15.

19. Lee E-J, Chen P, Jordan TH. Testing waveform predictions of 3D velocity models against two recent Los Angeles earthquakes, *Seismol. Res. Lett.* 2014; **85**: 1275-1284.
20. Taborda R, Azizzadeh-Roodpish S, Khoshnevis N, Cheng K. Evaluation of the southern California seismic velocity models through simulation of recorded events. *Geophys. J. Int.* 2016; **205**: 1342-1364.
21. Graves R, Jordan TH, Callaghan S, Deelman E, Field E, Juve G, Kesselman C, Maechling P, Mehta G, Milner K, Okaya D, Small P, Vahi K. CyberShake: A physics-based probabilistic hazard model for Southern California. *Pure Appl. Geophys.* 2011; **167**: 367-381.
22. Callaghan S, Deelman E, Gunter D, Juve G, Maechling P, Brooks C, Vahi K, Milner K, Graves R, Field E, Okaya D, Jordan T. Scaling up workflow-based applications. *J. Comp. System Sci.* 2010; **76**: 428-446.
23. Zhao L, Chen P, Jordan TH. Strain Green's tensors, reciprocity, and their applications to seismic source and structure studies. *Bull. Seismol. Soc. Am.* 2006; **96**: 1753-1763.
24. Wang F, Jordan TH, Comparison of probabilistic seismic hazard models using averaging-based factorization. *Bull. Seismol. Soc. Am.* 2014; **104**: 1230-1257.
25. Small P, Gill D, Maechling PJ, Taborda R, Callaghan S, Jordan TH, Ely GP, Olsen KB, Goulet CA. The SCEC Unified Community Velocity Model Software Framework. *Seismol. Res. Lett.* 2017; **88**: 1539-1552.
26. Cui Y, Olsen KB, Jordan TH, Lee K, Zhou J, Small P, Roten D, Ely G, Panda DK, Chourasia A, Levesque J, Day SM, Maechling P. Scalable earthquake simulation on petascale supercomputers, *Proc. 2010 ACM/IEEE Int. Conf. High Performance Computing Networking, Storage and Analysis*, New Orleans, Nov. 13-19.
27. Cui Y., Poyraz E, Callaghan S, Maechling P, Chen P, Jordan, TH. Accelerating CyberShake calculations on XE6/XK7 Platforms of Blue Waters. *Blue Waters and XSEDE Extreme Scaling Workshop* 2013, Aug 15-16, Boulder.
28. Field EH, Jordan TH, C.A. Cornell CA. OpenSHA: A developing community-modeling environment for seismic hazard analysis. *Seismol. Res. Lett.* 2003; **74**:406-419.
29. Deelman E, Vahi K, Juve G, Rynge M, Callaghan S, Maechling PJ, Mayani R, Chen W, da Silva RF, Livny M, Wenger K. Pegasus: a Workflow Management System for Science Automation, *Future Generation Computer Systems* 2015; **46**: 17-35.
30. Thain D, Tannenbaum T, Livny M. Distributed computing in practice: the Condor experience, *Concurrency and Computation: Practice and Experience* 2005; **17**: 323-356.
31. Foster I, Kesselman C, Tuecke S. The anatomy of the Grid: enabling scalable virtual organizations, *International Journal of High Performance Computing Applications* 2001; **15**: 200-222.
32. Callaghan S, Vahi K, Juve G, Maechling P, Jordan TH, and Deelman E. rvGAHP – Push-based job submission using reverse SSH connections. *Proceedings of the 12th Workshop on Workflows in Support of Large-Scale Science* 2017; Article 3, doi:10.1145/3150994.3151003.
33. Kohler M, Magistrale H, Clayton R. Mantle heterogeneities and the SCEC three-dimensional seismic velocity model version 3. *Bull. Seismol. Soc. Am.* 2003; **93**: 757-774; <https://scec.usc.edu/scecpedia/CVM-S4>.
34. Graves RW. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite-differences. *Bull. Seism. Soc. Am.* 1996; **86**: 1091-1106.
35. Petersen MD, Frankel AD, Harmsen SC, Mueller CS, Haller KM, Wheeler RL, Wesson RL, Zeng Y, Boyd OS, Perkins DM, Luco N, Field EH, Wills CJ, Rukstales KS. *Documentation for the 2008 Update of the United States National Seismic Hazard Maps*, USGS Tech. Rept. OFR-2008-1128.
36. Lee EN, Jordan TH, Chen P, Maechling PJ, Boué P, Denolle M, Beroza GC, Eymold WK, 2015. Full-3D tomography of crustal structure in central California, *Seismol. Res. Lett.*, (abstract) 2015.
37. Crouse CB, Jordan TH. Development of new ground-motion maps for Los Angeles based on 3-D numerical simulations and NGA West2 equations, *Proceedings of the SMIP17 Seminar on Utilization of Strong Motion Data*, University of California, Irvine, October 6, 2016.