CyberShake: A Physics-Based Seismic Hazard Model for Southern California

ROBERT GRAVES,[1] THOMAS H. JORDAN,[2] SCOTT CALLAGHAN,[2] EWA DEELMAN,[3] EDWARD FIELD,[4] GIDEON JUVE,[2] CARL KESSELMAN,[3] PHILIP MAECHLING,[2] GAURANG MEHTA,[2] KEVIN MILNER,[2] DAVID OKAYA,[2] PATRICK SMALL,[2] and KARAN VAHI[2]

Abstract—CyberShake, as part of the Southern California Earthquake Center's (SCEC) Community Modeling Environment, is developing a methodology that explicitly incorporates deterministic source and wave propagation effects within seismic hazard calculations through the use of physics-based 3D ground motion simulations. To calculate a waveform-based seismic hazard estimate for a site of interest, we begin with Uniform California Earthquake Rupture Forecast, Version 2.0 (UCERF2.0) and identify all ruptures within 200 km of the site of interest. We convert the UCERF2.0 rupture definition into multiple rupture variations with differing hypocenter locations and slip distributions, resulting in about 415,000 rupture variations per site. Strain Green Tensors are calculated for the site of interest using the SCEC Community Velocity Model, Version 4 (CVM4), and then, using reciprocity, we calculate synthetic seismograms for each rupture variation. Peak intensity measures are then extracted from these synthetics and combined with the original rupture probabilities to produce probabilistic seismic hazard curves for the site. Being explicitly site-based, CyberShake directly samples the ground motion variability at that site over many earthquake cycles (i.e., rupture scenarios) and alleviates the need for the ergodic assumption that is implicitly included in traditional empirically based calculations. Thus far, we have simulated ruptures at over 200 sites in the Los Angeles region for ground shaking periods of 2 s and longer, providing the basis for the first generation CyberShake hazard maps. Our results indicate that the combination of rupture directivity and basin response effects can lead to an increase in the hazard level for some sites, relative to that given by a conventional Ground Motion Prediction Equation (GMPE). Additionally, and perhaps more importantly, we find that the physics-based hazard results are much more sensitive to the assumed magnitude-area relations and magnitude uncertainty estimates used in the definition of the ruptures than is found in the traditional GMPE approach. This reinforces the need for continued development of a better understanding of earthquake source characterization and the constitutive relations that govern the earthquake rupture process.

1. Introduction

Numerical simulations of the strong ground motions caused by earthquakes have improved to the point where it is worth investigating the predictive power of these physics-based methods in seismic hazard analysis. Southern California provides a suitable natural laboratory. Researchers working together through the Southern California Earthquake Center (SCEC) have developed community velocity models, CVM-S (http://www.data.scec.org/3Dvelocity) and more recently CVM-H (http://sger5.harvard.edu/cvm-h), that include detailed representations of sedimentary basins and other near-surface structures, which influence ground motions. Numerical simulations of anelastic wave propagation through these three-dimensional (3D) structures have been tested against the ground motions recorded by the California Integrated Seismic Network (CISN; e.g., GRAVES, 2008; MAYHEW AND OLSEN, 2010), and efforts are underway to improve the CVMs using the earthquake waveform data (CHEN *et al.*, 2007; TAPE *et al.*, 2010). The plate-boundary fault system has been well described in a Community Fault Model (PLESCH *et al.*, 2007), and long-term earthquake rupture forecasts based on the CFM are now available (FIELD *et al.*, 2009).

These developments have motivated considerable research on the prediction of strong ground motions from the large, as-of-yet unobserved, fault ruptures that will someday occur. Source directivity and basin excitation effects have been studied systematically

[1] URS Corporation, Los Angeles, CA, USA. E-mail: robert_graves@urscorp.com
[2] USC, Los Angeles, CA, USA.
[3] USC/ISI, Los Angeles, CA, USA.
[4] USGS, Golden, CO, USA.

(e.g., OLSEN et al., 2006; GRAVES et al., 2008; AAG-AARD et al., 2010), dynamical rupture models have been used to calibrate kinematic rupture models (GUATTERI et al., 2004; OLSEN et al., 2008, 2009; SONG et al., 2009; SCHMEDES et al., 2010), and different simulation codes have been cross-verified (BIELAK et al., 2010). One goal of this research is to improve the ground motion prediction equations (GMPEs) commonly used in seismic hazard analysis.

Ground motion prediction equations specify the conditional probability of exceeding a ground motion intensity measure at a particular geographic site for a particular source represented in an earthquake rupture forecast (CORNELL, 1968). The intensity measure commonly used is SA(T), the 5% damped spectral acceleration at a frequency T. In probabilistic seismic hazard analysis (PSHA) terminology, a "source" is the spatial locus of a rupture, usually a delineated area of a fault surface. By combining the GMPE probabilities with the rupture probabilities of all considered sources, one can compute the unconditional probability of exceeding the intensity measure during a future time interval. A plot of this probability of exceedance, PoE, as a function of SA is called a hazard curve for the specified site. In general, the hazard curve will depend on the site location, as well as the ground shaking period and the time interval being considered. However, most GMPEs are attenuation relationships in which the site location is parameterized by a relative location with respect to the source (which depends on the fault geometry), the site conditions (often represented by V_{s30}, the local average of the shear velocity in the upper 30 m) and, in some cases, the local depth of the sedimentary basin (e.g., FIELD, 2000; ABRAHAMSON AND SILVA, 2008; CAMPBELL AND BORZOGNIA, 2008; CHIOU AND YOUNGS, 2008). These parameterizations are determined by empirical regressions of assumed functional forms to the available data. Additionally, the application of GMPEs in the probabilistic framework typically assumes that the measured variability of ground motions (encompassing multiple earthquakes observed at spatially distributed sites over the last few decades) accurately represents the variability of motions expected at a single site over many earthquake cycles (potentially thousands of years). This is the so-called ergodic assumption, and it has been

suggested that it can have an unintended upward bias in the estimated hazard level (e.g., ANDERSON AND BRUNE, 1999; O'CONNELL et al., 2007).

Empirical GMPEs can potentially be improved by supplementing the direct observations of ground motions with simulation data that use the physics of wave propagation to extrapolate to unobserved conditions. For example, the simulated basin responses of DAY et al. (2008) were used by some of the GMPE modelers in the Next Generation Attenuation (NGA) project. The simulated motions provided key constraints to the functional form and period dependence of the basin response that would not have been possible using empirical observations alone.

This paper outlines progress towards a more ambitious goal: entirely replacing the attenuation relationship with simulation-based ground motion predictions. The computational platform for such a calculation must be able to efficiently simulate the ground motions at each site for an ensemble of rupture variations. The ensemble must be sufficiently large to characterize all sources in the earthquake rupture forecast. In particular, it must be large enough to properly represent the expected variability in the source parameters—e.g., hypocenter location, stress drop, and the slip heterogeneity. In the implementation described here, called the CyberShake platform, the ground motion time series at a given site are calculated using seismic reciprocity for an ensemble of "pseudo-dynamic" rupture variations that sample the Uniform California Earthquake Rupture Forecast, Version 2 (UCERF2) (FIELD et al., 2009). Results are presented for approximately 415,000 rupture scenarios simulated at each of 250 sites in the Los Angeles region and for ground shaking periods of 2 s and longer. The entire computational process has been automated using scientific workflow tools developed within the SCEC Community Modeling Environment using TeraGrid high-performance computing facilities (JORDAN AND MAECHLING, 2003; DEELMAN et al., 2006).

2. Cybershake Computational Platform

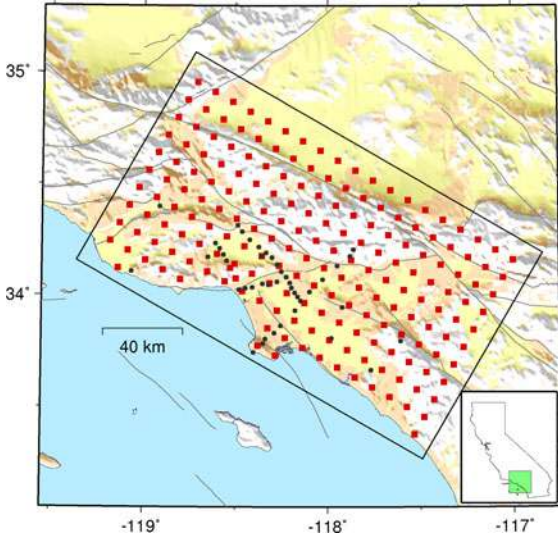Figure 1 shows a map of the Southern California region where CyberShake is being implemented. To

Figure 1

Map of Southern California showing target region for CyberShake. *Red squares* indicate uniform grid of sites where CyberShake hazard curves have been computed and *black circles* denote additional sites computed along selected profiles

date, hazard curve computations have been completed at about 250 sites. The eventual goal is to cover the greater Los Angeles basin region with a grid of sites having an average spacing of about 5 km. This will provide a basis for generating full hazard maps of this region.

For a typical site in the Los Angeles region, the latest Earthquake Rupture Forecast available from the USGS (UCERF 2.0) identifies more than 10,000 earthquake ruptures with moment magnitude (M_w) greater than 6 that might affect the site. For each rupture, we must capture the possible variability in the earthquake rupture process. So we create a variety of hypocenter and slip distributions for each rupture to produce over 415,000 rupture variations, each representing a potential earthquake. In CyberShake processing, there is a fairly technical, but important, distinction between ruptures ($\sim$7,000) and rupture variations ($\sim$415,000), which impacts our computational workflows. The ruptures as defined in UCERF2.0 can be generically thought of as faults that generate an earthquake of a certain magnitude. However, UCERF2.0 contains no information about the details of the rupture process, that is, where the rupture initiates (hypocenter) or what the slip distribution might be. In the CyberShake workflow, the

SGT calculations generate Green's functions for all the faults of interest, and then post-processing must be done in order to generate the ground motions for each individual rupture variation on each fault.

Once we define the ruptures and their variations, CyberShake uses an anelastic wave propagation simulation to calculate strain Green tensors (SGTs) around the site of interest. Seismic reciprocity is used to post-process the SGTs and obtain synthetic seismograms (e.g., GRAVES AND WALD, 2001; ZHAO *et al.*, 2006). These seismograms are then processed to obtain peak spectral acceleration values, which are combined into a hazard curve. Figure 2 contains a workflow illustrating these steps.

The CyberShake platform must run many jobs and manage many data files. An outline of computational and data requirements for each site of interest is given in Table 1. To compute the SGTs, a mesh of about 1.5 billion points must be constructed and populated with seismic wave velocity information. The velocity mesh is then used in a wave propagation simulation for 20,000 time steps.

Once the SGTs are calculated, the ground motion waveforms for each of the approximately 415,000 rupture variations are computed. For each individual rupture variation, the SGTs corresponding to the location of the rupture (fault) are extracted from the volume and convolved with the specific rupture variation to generate synthetic seismograms, which represent the ground motions that would be produced at the site we are studying. Next the seismograms are processed to obtain the IM of interest, which, in our current study, is peak spectral acceleration at periods ranging from 3 to 10 s. Each execution of these post-processing steps takes no more than a few minutes, but SGT extraction must be performed for each

Table 1

Data and CPU requirements for the CyberShake computational components, per site of interest

Component	Data	CPU hours
Mesh generation	15 GB	150
SGT simulation	40 GB	10,000
SGT extraction	680 GB	250
Seismogram synthesis	10 GB	6,000
PSHA calculation	90 MB	100
Total	755 GB	17,000

rupture, and seismogram synthesis and peak spectral acceleration processing must be performed for each rupture variation. On average, 7,000 ruptures and 415,000 rupture variations must be considered for each site, which requires approximately 840,000 executions, 17,000 CPU-hours and generates about 750 GB of data.

Considering only the computational time, performing these calculations on a single processor would take almost 2 years. In addition, the large number of independent post-processing jobs necessitates a high degree of automation to help submit jobs, manage data, and provide error recovery capabilities should jobs fail. The velocity mesh creation and SGT simulation are large MPI jobs which run on a cluster using spatial decomposition. The post-processing jobs have a very different character, as they are loosely coupled—no communication is required between jobs.

These processing requirements indicate that the CyberShake computational platform requires both high-performance computing (for the SGT calculations) and high-throughput computing (for the post-processing). To make a Southern California hazard map practical, time-to-solution per site needs to be short, on the order of 24–48 h. This emphasis on reducing time-to-solution pushes the CyberShake computational platform into the high productivity computing category, which is emerging as a key capability needed by science groups. The challenge is to minimize overhead and increase throughput to reduce end-to-end wall clock time. A thorough discussion of the software tools and engineered solutions to enable CyberShake to run at scale can be found in (CALLAGHAN *et al.*, 2010).

3. *Earthquake Rupture Characterization*

Uniform California Earthquake Rupture Forecast, Version 2, utilizes the rectangularized fault definitions given by the SCEC Community Fault Model version 3.0 (CFM3). Magnitudes for ruptures of these faults are estimated using magnitude-area scaling relations. Four scaling relations are currently implemented within UCERF2: Ellsworth-B (WGCEP, 2003), HANKS AND BAKUN (2007), SOMERVILLE (2006) and WELLS AND COPPERSMITH (1994), although currently

only Ellsworth-B and Hanks-Bakun are given non-zero weights. For magnitudes larger than about 7, the Ellsworth-B and Hanks-Bakun relations predict magnitudes about 0.2 units larger for the same fault area compared to Somerville and Wells-Coppersmith. While this has little impact on calculations utilizing GMPEs, we have found that this has a significant impact on physics-based simulations. The 0.2 unit increase in magnitude corresponds to a factor of 2 increase in seismic moment. At long periods and for a fixed fault area, the ground motion amplitudes of the numerical simulations scale almost directly with increasing seismic moment (M_o). However, the GMPE has a built-in magnitude-area relation that implicitly adjusts the area for the prescribed magnitude. Thus, the ground motions for the empirical model scale more like $M_o^{1/3}$ as the magnitude is changed.

For this reason, it is imperative that we specify the fault areas and magnitudes for CyberShake using the most appropriate scaling relation. The Somerville relation is based on fault rupture characterizations that were developed using waveform inversions of well recorded earthquakes. These inversions utilize the same wave propagation physics and ground motion representation theorems as implemented in CyberShake. Thus, for consistency in the CyberShake implementation, we have modified the weighting of the scaling relations in UCERF2 to give full weight to Somerville. In doing this, we retained the magnitude estimates given by the original UCERF2 definitions, and simply adjusted the fault areas (by increasing their down-dip widths) such that the resulting set of ruptures correspond, on average, to the Somerville relation. Figure 3 plots all the UCERF2 ruptures and compares the original and corrected rupture areas for each magnitude.

In order to perform the numerical simulations, we need to develop a full kinematic description of fault rupture for each scenario. This includes slip distribution, hypocenter location, rupture propagation and slip time function. For a given fault and M_w, the rupture model is generated in the wavenumber domain by constraining the amplitude spectrum of slip to fit a K^{-2} falloff with random phasing (SOMERVILLE *et al.*, 1999; MAI AND BEROZA, 2002). The slip velocity function is constructed using two triangles as shown in Fig. 4, with the rise time scaling with increasing
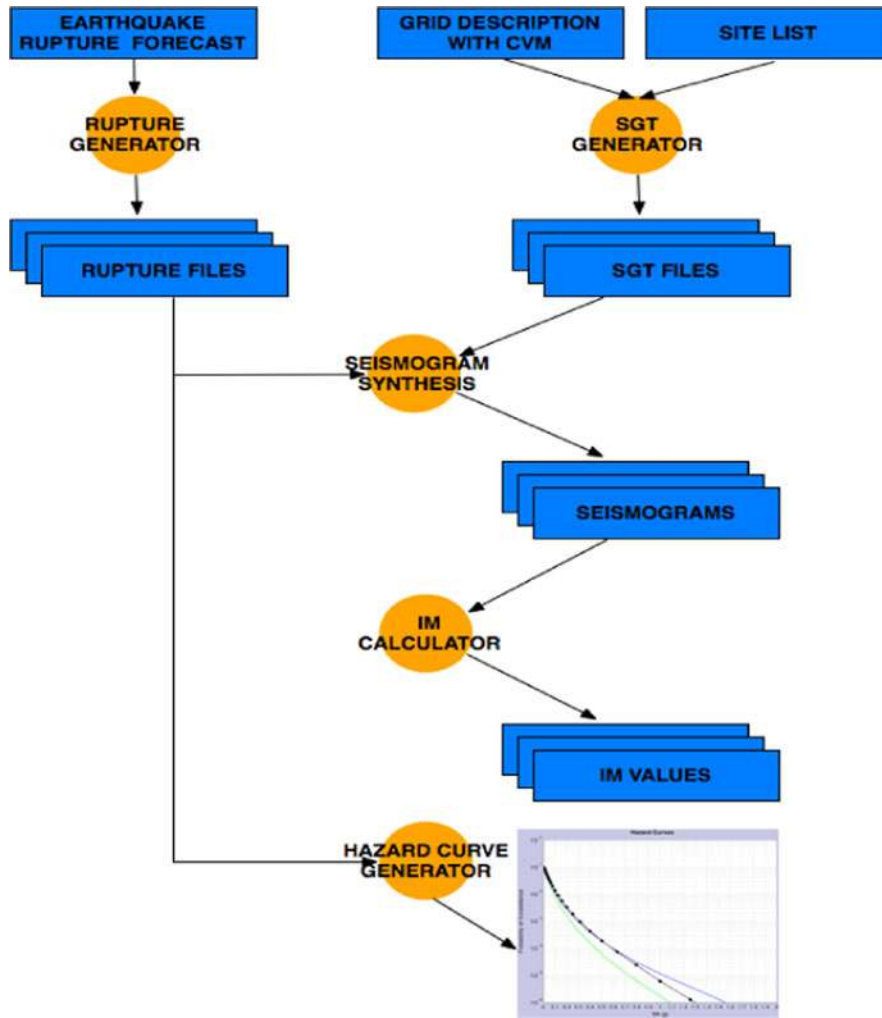
Figure 2
A high level CyberShake workflow. *Gold circles* indicate computational modules and *blue rectangles* indicate files and databases

magnitude (SOMERVILLE *et al.*, 1999). For each fault, multiple hypocenters and slip distributions are considered. Hypocenters are placed every 20 km along strike and two slip distributions are run for each hypocenter. The current implementation only considers median values for rise time and rupture velocity (80% of local V_s). Figure 4 displays representative rupture models for 3 different magnitudes.

4. Verification

For each site of interest, we first compute a full set of SGTs for all ruptures within 200 km of the site. The SGTs are calculated via reciprocity within the prescribed 3D velocity structure using a parallelized anelastic FD algorithm (GRAVES, 1996; GRAVES AND WALD, 2001). We set the minimum shear wave velocity at 500 m/s and, using a grid spacing of 200 m, we obtain a maximum frequency resolution of 0.5 Hz. Two calculations are required, one each to obtain the SGTs for each of the two orthogonal horizontal components of motion. Each fault surface is sampled at a 1 km spacing, and the SGTs are saved for each point on each of these fault surfaces. In total, there are about 420,000 point SGT locations, and the resulting set of SGTs requires about 40 GB of storage for each site.
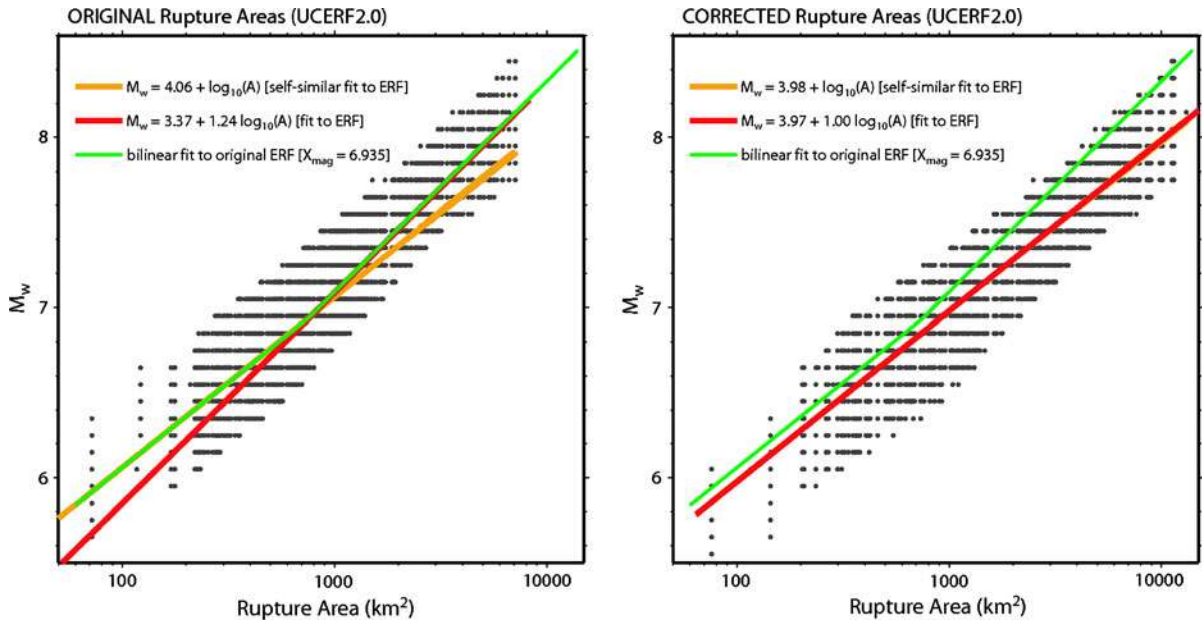
Figure 3
Magnitude vs area plots for all UCERF2 ruptures. *Left panel* shows original rupture set based on the Ellsworth-B and Hanks-Bakun scaling relations. *Right panel* shows corrected rupture set based on Somerville scaling relation. *Lines* are regression fits through the roughly 10,000 data points each of which represents a rupture from UCERF2.0

The advantage of pre-computing and saving the SGTs is that it requires only two large-scale 3D simulations per site (roughly 15 h per simulation using 400 CPUs). Once computed, the convolution of the SGTs with each of the approximately 415,000 rupture variations is quite fast and easily parallelizable (a few hours total runtime for all ruptures). Running a forward simulation for each rupture would have the advantage of providing ground motions for a great number of locations, but this approach would not be currently practical for the large number of rupture scenarios that need to be considered. Theoretically, these two approaches produce exactly the same results. Figure 5 compares ground motions simulated at a site in Long Beach (LBP) for a M_w 7.8 rupture of the San Andreas fault using both approaches. The agreement between the two calculations is excellent, with the very slight differences in this case due to small differences in the spatial discretization of the forward and reciprocal models.

The bulk of recorded strong motion data are from events of magnitude from about 6 to 7 and for distances out to about 70 km. Within these ranges,

the GMPEs are quite well constrained. Figure 6 compares CyberShake simulations at Pasadena (PAS) against GMPE predictions for the subset of UCERF2 rupture scenarios within this magnitude and distance range. We chose PAS as a reference site for this comparison because it is not strongly affected by the deep sediments of the surrounding basins. For this comparison, we have averaged the results of four recent GMPEs; ABRAHAMSON AND SILVA (2008), BOORE AND ATKINSON (2008), CAMPBELL AND BOZORGNIA (2008), and CHIOU AND YOUNGS (2008). On average, the agreement between the CyberShake and empirical predictions is quite good and supports the validity of the CyberShake approach. In computing the CyberShake values, we have averaged over all hypocenters and slip models for a given source. Thus, the scatter of these results is due primarily to the magnitude (i.e., static stress drop) variability of the sources as defined by UCERF2. Additionally, the scatter is not isotropic, but is skewed vertically along the CyberShake axis indicating that the magnitude variability produces a much greater effect on the CyberShake ground motions than on the GMPE ground motions. These results also support the
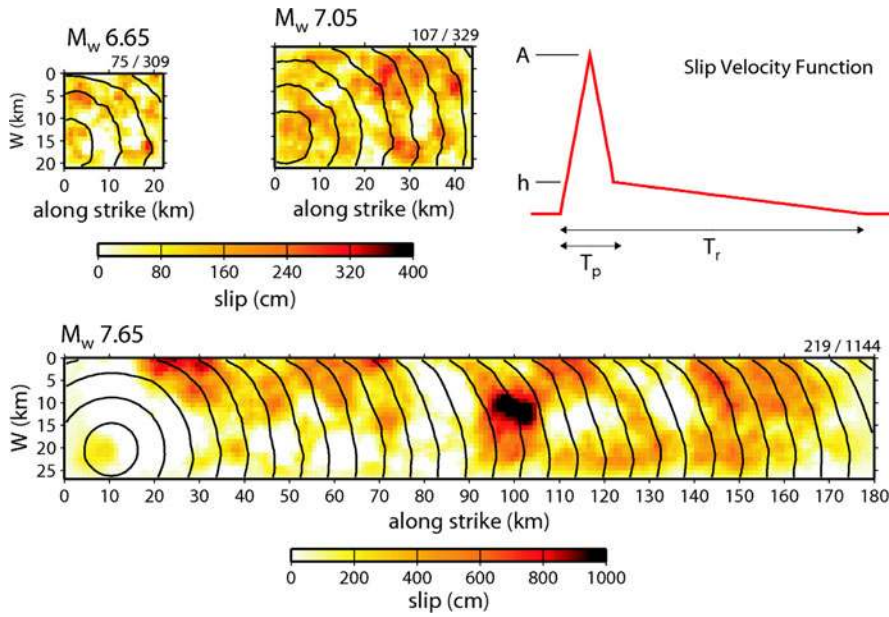
Figure 4

Kinematic rupture models for scenarios of magnitude 6.65, 7.05 and 7.65. Contours of rupture time are shown at 2 s intervals for each rupture. Upper right displays schematic view of slip velocity function used for the kinematic description
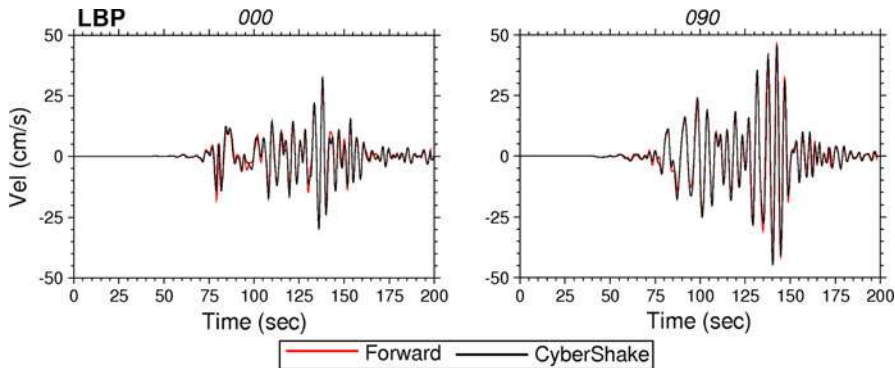

Figure 5

Comparison of waveforms simulated at Long Beach for a M_w 7.8 rupture of the San Andreas fault using a forward calculation (*red*) and the reciprocal SGT method of CyberShake (*black*). The bandwidth of the motions is $f < 0.5$ Hz

adjustment of fault rupture areas described in the previous section. If we had used the original UCERF2 areas, which on average are smaller by about 25% for these magnitudes, then our kinematic rupture models would require an increase in average slip of a comparable amount. The resulting Cyber-Shake simulations would then be about 25% higher on average, which would not be consistent with the GMPE predictions.

5. Hazard Curves

Figure 7 compares 3 s spectral acceleration (SA) hazard curves computed at four sites with the CyberShake Platform against standard curves derived from the GMPEs of BOORE AND ATKINSON (2008) and CAMPBELL AND BOZORGNIA (2008), hereafter referred to as BA08 and CB08, respectively. Background seismicity is excluded from both models and the

GMPE calculation is truncated at 3 sigma. Both GMPEs utilize V_{s30} (the travel-time averaged shear wave velocity in the upper 30 meters) to account for site response effects. Additionally, the Campbell-Bozorgnia relation incorporates a basin response effect based on the depth to $V_s = 2.5$ km/s beneath the site (referred to as Z2.5), which can have a significant impact at deep basin sites for the longer periods. Rupture directivity effects are not explicitly included in these GMPEs. However, all of these effects are naturally included in the CyberShake results through the use of the 3D velocity structure for the ground motion simulations. The hazard curves were generated using the resources and applications of OpenSHA (FIELD *et al.*, 2003, http://www.opensha.org) to combine the ground motion amplitudes

calculated by CyberShake with the rupture probabilities specified in UCERF2.0. The PAS site can be regarded as a "rock" site with a relatively high V_{s30} of 748 m/s and a relatively shallow Z2.5 of 0.31 km, whereas the USC, WNGC and STNI sites are "basin" sites with relatively low V_{s30} of 280 m/s and a thick accumulation of soft sediments (Z2.5 ranges from about 3–6 km).

At the rock site (PAS), both the CyberShake and GMPE approaches produce similar results, whereas for the basin sites, the hazard levels produced by CyberShake are at or above the GMPE results. At USC, which has a modest basin depth (Z2.5 of 3.9 km), the CyberShake curve tracks the CB08 curve quite closely, with both being at a somewhat higher level than BA08. A similar trend is seen at

PAS
$6 < M_w < 7$
0 km < Dist < 70 km
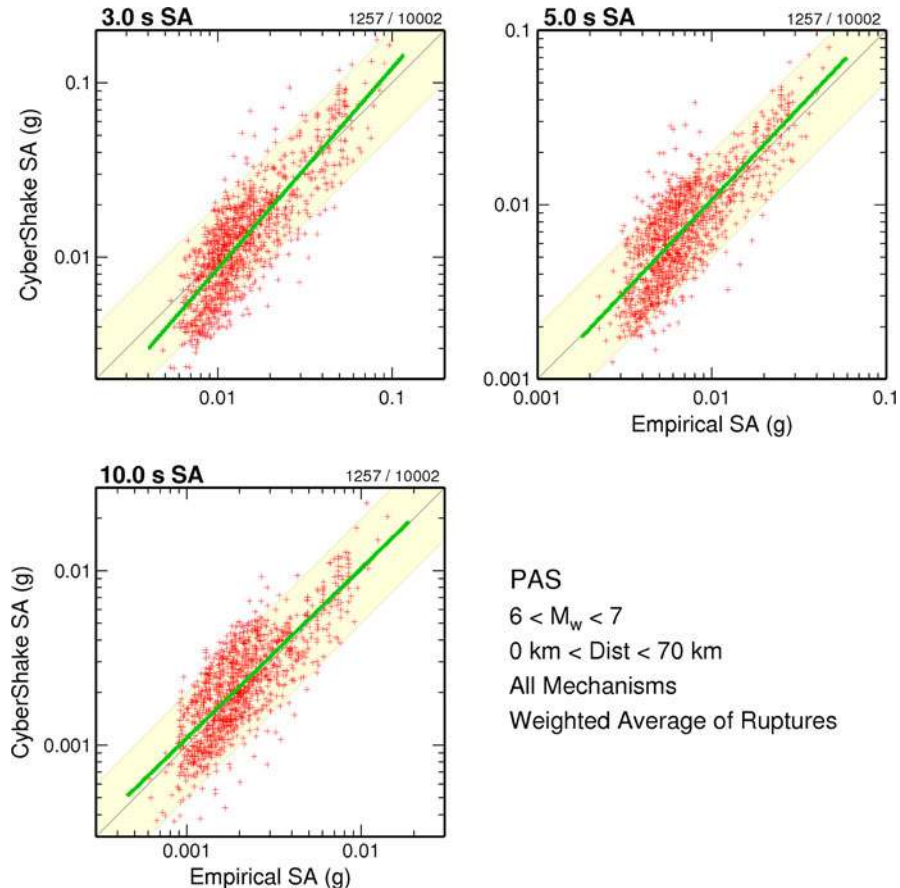All Mechanisms
Weighted Average of Ruptures

Figure 6

Comparison of CyberShake computed motions with empirical estimates at Pasadena for the subset of ruptures having magnitudes between 6 and 7, and within 70 km closest distance to the site. The empirical estimates are the average of the four GMPEs discussed in the text. The *three panels* show spectral acceleration at oscillator periods of 3 s (*upper left*), 5 s (*upper right*), and 10 s (*lower left*). The *green line* in each panel is a least squares fit to the points. *Yellow shading* is region within a factor of 2 of one-to-one correspondence

STNI; however due to the much greater basin depth (Z2.5 of 6.0 km), both CyberShake and CB08 predict much higher hazard levels than BA08. At 3 s period, the basin amplification term of CB08 increases the median ground motion level by about 25 and 75% for Z2.5 of 3.9 and 6.0 km, respectively. The GMPE of BA08 implicitly incorporates basin response effects through the V_{s30} site response term; however, this approach cannot distinguish between sites of different Z2.5 with the same V_{s30}, such as the case considered here. The similarity of the hazard levels produced by CyberShake (which naturally includes basin response effects) and CB08 (which has an explicit basin amplification term) for the USC and STNI sites demonstrate the potential importance of deep basin sediments on long period hazard levels.

In addition to basin response, the CyberShake motions can be further amplified by rupture directivity effects. This is particularly evident at the WNGC site which has a modest basin depth (Z2.5 of 2.81 km), but stills exhibit relatively high ground motion hazard from the numerical simulations. Previous TeraShake (OLSEN *et al.*, 2006, 2008) and ShakeOut (GRAVES *et al.*, 2008) ground motion modeling studies have shown that this site is susceptible to channeling and amplification of basin waves for larger ruptures on the southern San Andreas fault. This amplification effect represents a coupling of rupture directivity and basin response which cannot be accounted for using the existing GMPE parameterizations. At this point, it is not entirely clear whether this coupling of directivity and basin response might occur at other sites in and around the Los Angeles basin region. However, the development of hazard maps using the CyberShake approach will allow for the systematic analysis of this effect and the identification of additional sites that are susceptible to this type of ground motion amplification phenomenon.

6. Disaggregation

Disaggregation breaks down the hazard to find which sources most significantly contribute to the hazard at a particular site. Disaggregation of the hazard curves can be done on a probability or a ground

motion level. Figure 8 shows the magnitude/distance disaggregation of the 3 s SA hazard curves for an annual probability of exceedance of 2% in 50 years at the STNI site, which corresponds to a 3 s SA level of about 0.55 g. The results are plotted as a function of magnitude, distance and percent contribution to the hazard. For most sites in the Los Angeles basin region, the pattern of the disaggregations is similar for both the empirical and CyberShake results. This occurs because the overall hazard for sites within the Los Angeles region is typically controlled by nearby moderate-sized (M_w 6–7.5) events and more distant large magnitude events ($M_w > 7.5$) on the San Andreas fault. The fact that CyberShake reproduces this pattern is a key validation of the methodology. The main differences between CyberShake and the empirical results are (1) CyberShake shows a higher contribution from the nearby sources compared to the San Andreas events and (2) CyberShake shows significant contributions for negative epsilon values, particularly for the nearby sources.

For the STNI site, the main contributing nearby sources are ruptures of the Newport-Inglewood and Palos Verdes fault systems. Both of these fault systems lie south of the site, and generally form the south-western margin of the Los Angeles basin. Because these faults are immediately adjacent to the basin, they are particularly efficient at channeling energy into the basin as the rupture propagates along the shallow portion of the fault. This coupling of fault rupture and basin response leads to larger ground motions for these ruptures, and thus increases their percent contribution to the overall hazard compared to the San Andreas ruptures. Furthermore, this type of coupling is not possible to capture using the current set of empirical GMPEs.

Since the CyberShake calculations are explicitly site-based, the variability in ground motions estimated at a given site is controlled directly by variations in the source characterization model (hypocenter location and slip distribution) and deterministic wave propagation effects. This direct approach alleviates the need for the ergodic assumption that is implicit in traditional PSHA (ANDERSON AND BRUNE, 1999; O'CONNELL *et al.*, 2007). For moderate magnitude ruptures and for long ground shaking periods (e.g. SA at 3 s and greater), the contribution of source variability to
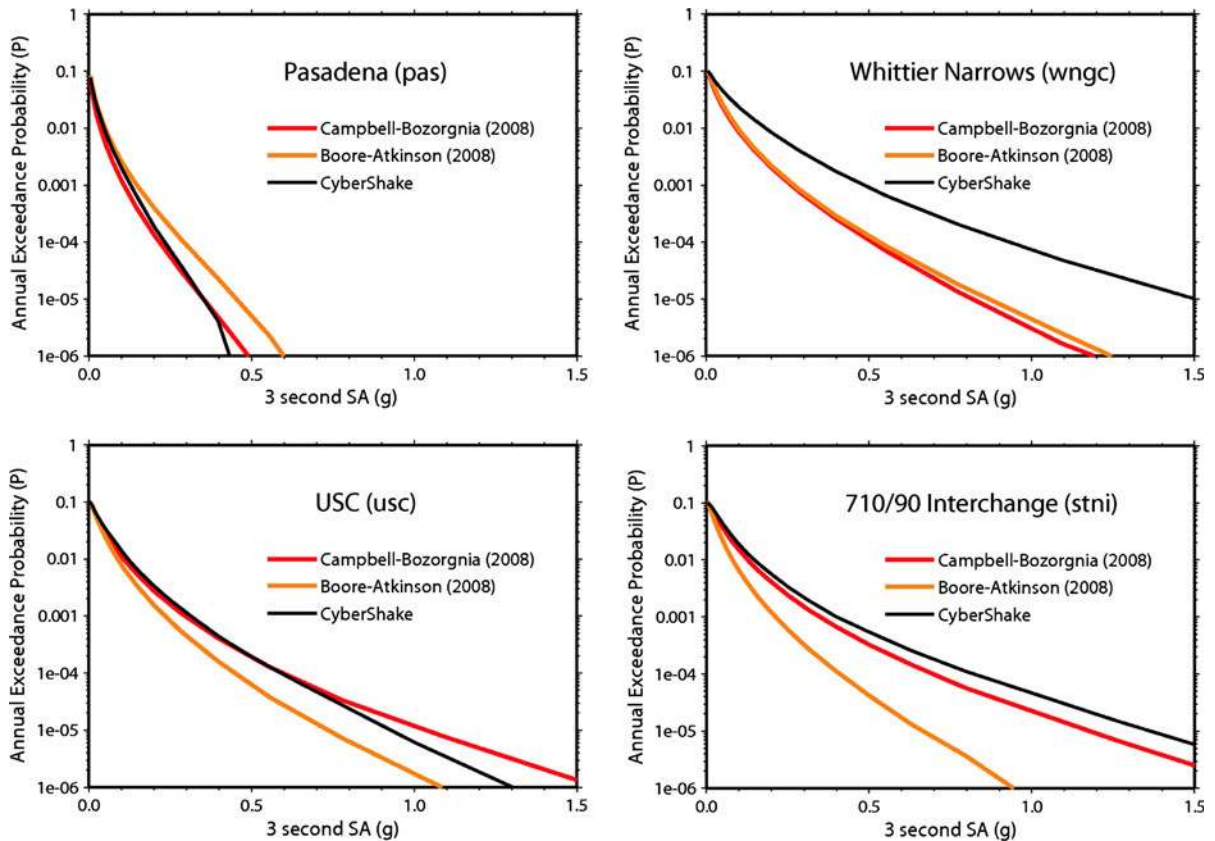
Figure 7
Hazard curves for spectral acceleration (SA) at 3 s period at four Los Angeles region sites for all ruptures of M_{w} 6 and greater excluding background seismicity. The empirical models use a truncation at 3 sigma

ground motion variability at a given site is relatively small. Hence the estimated distribution of ground motions for these ruptures is rather narrow, or stated in another way, the ground motions have a have a relatively low sigma compared to that predicted by empirical GMPEs. There is mounting evidence to support the idea that ground motion variability expected at a particular site for a particular set of ruptures is significantly less than would be estimated using the full sigma of the GMPEs (e.g. ATKINSON, 2006), which is consistent with the CyberShake results presented here.

7. Cybershake Hazard Maps

With the grid of CyberShake sites completed thus far (Fig. 1), we are able to construct a first generation CyberShake hazard map for the Los Angeles region.

We recognize that the density of sites (nominally at 10 km spacing) does not provide the resolution needed for detailed interpretation of these results. However, it does highlight the importance of effects such as rupture directivity and basin response on the hazard levels, and furthermore, it demonstrates the potential of using the CyberShake approach for hazard characterization on a regional scale (Fig. 9).

To produce a CyberShake hazard map, we first compute the residuals between the CyberShake prediction with those obtained from an empirical GMPE approach at each site. We then construct an interpolated version of this residual field that covers the region of interest. Finally, we construct the Cyber-Shake map by adding the interpolated residuals to the original GMPE based map. An example of this process is illustrated in Fig. 9 using the GMPE of CAMPBELL AND BOZORGNIA (2008). For this demonstration, we use 3 s SA and an annual exceedance
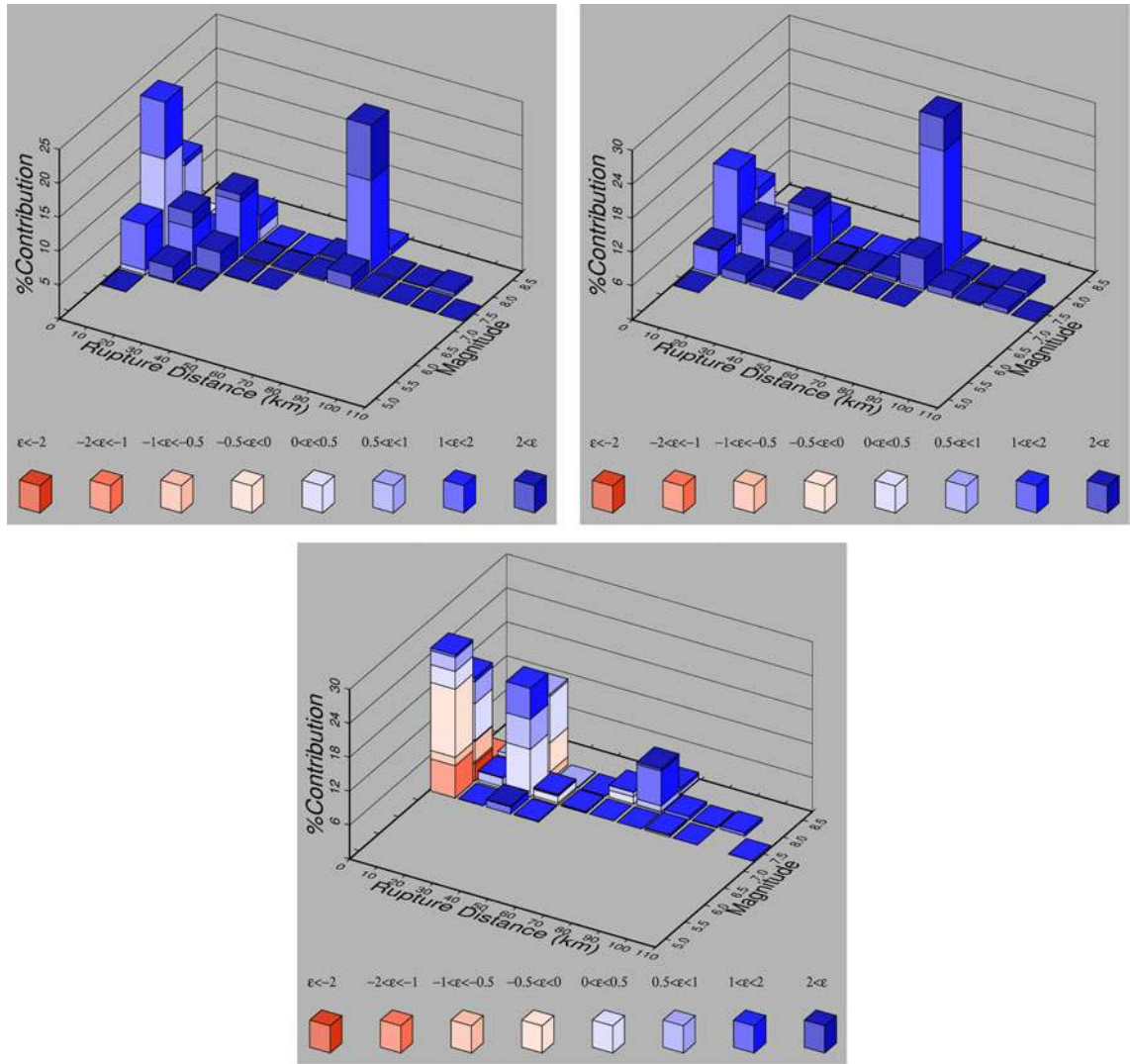
Figure 8
Magnitude/distance disaggregation at STNI for 3 s SA at an annual exceedance probability of 2% in 50 years. *Top left panel* is based on the GMPE of CAMPBELL AND BOZORGNIA (2008), *top right panel* is based on the GMPE of BOORE AND ATKINSON (2008), and *bottom panel* is derived from CyberShake

probability of 2% in 50 years. The resulting Cyber-Shake map shows generally elevated hazard for many of the deep basin sites, and a generally reduced hazard level along the San Andreas fault.

8. Conclusions

The SCEC CyberShake project has developed an approach for implementing physics-based waveform simulations in seismic hazard calculations. The advantage of the physics-based approach over the GMPE approach is that deterministic earthquake rupture and wave propagation effects are explicitly included in the ground motion response. Furthermore, since CyberShake is explicitly site-based, there is no need for the ergodic assumption when computing probabilistic hazard estimates. That is, the predicted variability in ground motions at a given site directly includes all of the 3D path and rupture characteristics specific and unique to that site for the prescribed earthquake rupture forecast. Thus, there is no need to
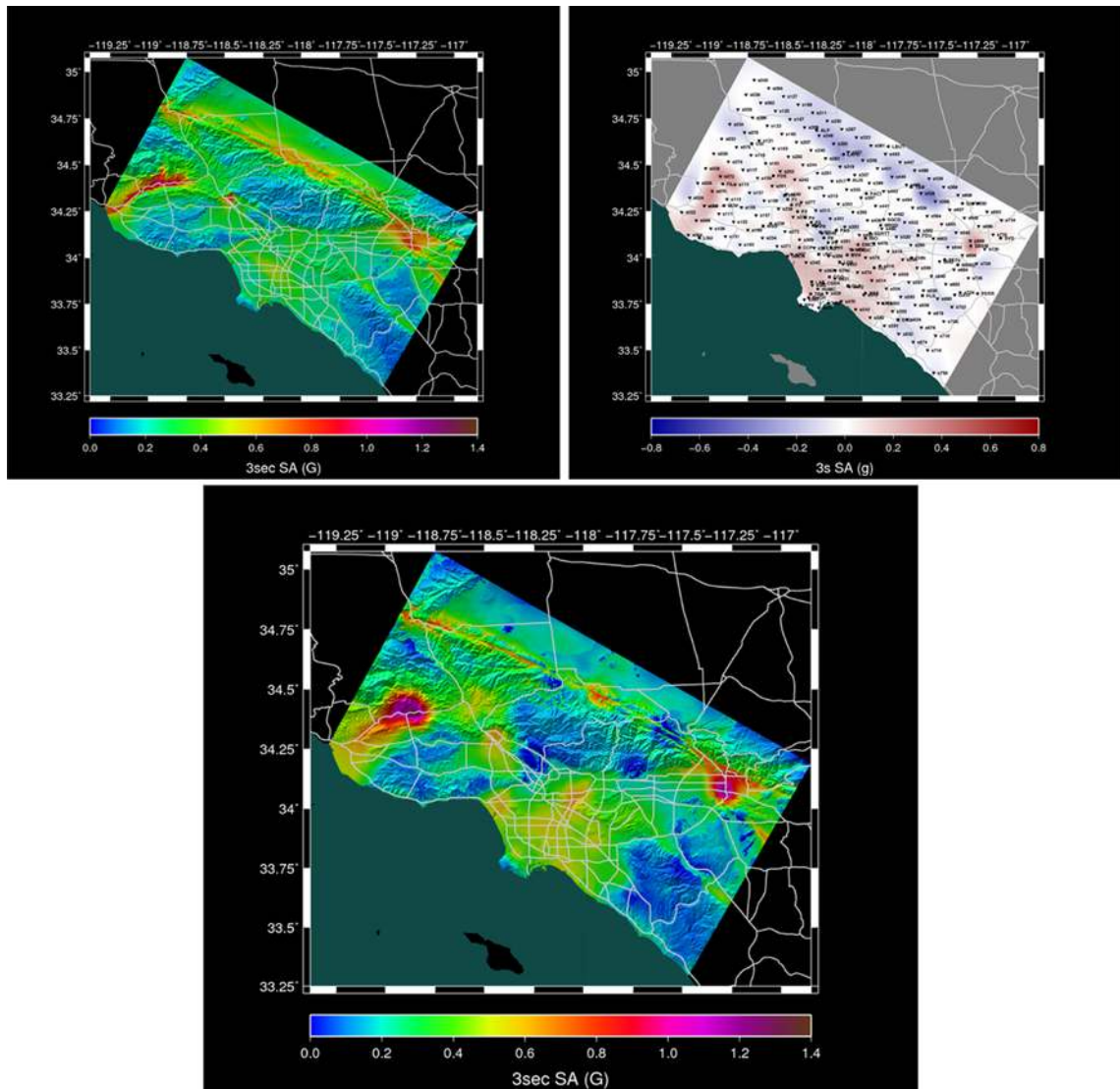
Figure 9
Top left panel shows hazard map for the Los Angeles region for 3 s SA at an annual exceedance probability of 2% in 50 years developed using the GMPE of CAMPBELL AND BOZORGNIA (2008). *Top right panel* shows the contoured residuals between the CyberShake hazard values and the CAMPBELL AND BOZORGNIA (2008) values. *Bottom panel* shows the CyberShake hazard map derived by interpolating the residual map onto the background map of CAMPBELL AND BOZORGNIA (2008)

utilize ground motions recorded in one region for application in another, or to assume that the variability of ground motions observed over a spatially distributed set of sites can be used as a surrogate for the temporal variability of ground motions expected at a given sites over many earthquake cycles. Application of the CyberShake approach requires significant computational resources which have been made available through the SCEC Community

Modeling Environment (http://www.scec.org/cme). Currently, the CyberShake wave form calculations are band-limited to ground shaking periods greater than 2 s (providing spectral accelerations for periods of 3 s and greater). The restriction to longer periods is primarily due to computational limitations, and is not an inherent limitation of the physics-based methodology. As the methodology is further developed, we intend to push the calculations to shorter periods. Our

preliminary results demonstrate this approach is viable and practical.

Incorporation of physics-based ground motions within a probabilistic framework requires careful consideration of how the earthquake ruptures are characterized. The current UCERF2.0 characterization uses an average of two magnitude-area scaling relations (Ellsworth-B and Hanks-Bakun), which systematically underestimate fault areas compared with physics-based rupture model inversions for event magnitudes larger than about 6.7, particularly for strike-slip faults (SOMERVILLE, 2006). This has little consequence on the traditional GMPE based hazard calculations because the GMPEs do not explicitly consider fault rupture area (or static stress drop) in determining ground motion levels. However, the physics based approach is quite sensitive to the magnitude-area scaling, because both magnitude and rupture area are required to fully characterize the rupture. Thus, using a fault area that is too small requires a corresponding increase in slip to preserve the target magnitude (seismic moment) in the physics-based simulations, which scales almost directly into ground motion amplitude at the longer periods. To circumvent this problem, we have modified the original UCERF2.0 fault descriptions by extending their down-dip widths such that the resulting fault rupture areas correspond, on average, to the SOMERVILLE (2006) scaling relation. Validation tests indicate that the modified rupture descriptions provide a much better match to recorded ground motion levels than the original descriptions.

The range of prescribed magnitude variability for the characteristic ruptures in UCERF2.0 is about 0.7 units or larger for a given fault rupture area. Since the fault rupture area is held fixed in these characteristic ruptures, the average ground motion levels in the numerical simulations scale almost directly with seismic moment. This 0.7 unit magnitude range corresponds to a change in seismic moment of over a factor of 10. However, as described above, rupture area is not a parameter that is used directly by the GMPEs. Thus, the GMPE implicitly adjusts the area (to maintain a constant median static stress drop) when the magnitude is changed, and consequently the ground motion levels predicted from the GMPEs are much less sensitive to changes in magnitude, scaling

roughly with seismic moment to the one-third power. The magnitude range of 0.7 units produces a variability in the median ground motion levels predicted by the GMPEs of only about a factor of two. This raises two important issues with respect to magnitude characterization the ERF. First, the strong sensitivity of the numerical simulation results to magnitude variability for a constant rupture area (combined with the relative lack thereof for the GMPEs) suggests that the prescribed range of magnitude variability defined by UCERF2.0 needs to be carefully examined. Second, it is possible that the large range of UCERF2.0 magnitude variability coupled with the use of GMPEs may result in a double counting of this effect due to its incorporation within the uncertainty estimates (sigma) of the GMPE.

Acknowledgments

REFERENCES

AAGAARD, B. T., GRAVES, R. W., MA, S., LARSEN, S. C., RODGERS, A., PETERSSON, N. A., JACHENS, R. C., BROCHER, T. M., SIMPSON, R. W., DREGER, D. (2010), *Ground motion modeling of Hayward fault scenario earthquakes II: Simulation of long-period and broadband ground motions*, submitted to Bull Seismol Soc Am

ABRAHAMSON, N. A., SILVA, W. J. (2008), *Summary of the Abrahamson and Silva NGA ground motion relations*. Earthq Spectra *24*(S1), 67–97

ATKINSON, G. (2006), *Single station sigma*. Bull Seismol Soc Am *96*, 446–455

ANDERSON, J. G., BRUNE, J. N. (1999), *Probabilistic seismic hazard analysis without the ergodic assumption*. Seismol Res Lett *70*, 19–28

BIELAK, J., GRAVES, R., OLSEN, K., TABORDA, R., RAMIREZ-GUZMAN, L., DAY, S., ELY, G., ROTEN, D., JORDAN, T., MAECHLING, P.,

URBANIC, J., CUI, Y., JUVE, G. (2010), *The ShakeOut earthquake scenario: Verification of three simulation sets*. Geophys J Int *180*, 375–404

BOORE, D. M., ATKINSON, G. M. (2008), *Ground motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 and 10.0 s*. Earthq Spectra *24*(S1), 99–138

CALLAGHAN, S. E., DEELMAN, D., GUNTER, G., JUVE, P., MAECHLING, C., BROOKS, K., VAHI, K., MILNER, R., GRAVES, E., FIELD, D., OKAYA, K., JORDAN, T. (2010), *Scaling up workflow-based applications*. J Comput Syst Sci, special issue on scientific workflows (in press)

CAMPBELL, K. W., BOZORGNIA, Y. (2008), *NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD, and 5%-damped linear elastic response spectra for periods ranging from 0.01 to 10 s*. Earthq Spectra *24*(S1), 139–172

CHEN, P., JORDAN, T. H., ZHAO, L. (2007), *Full three-dimensional waveform tomography: a comparison between the scattering-integral and adjoint-wavefield methods*. Geophys J Int *170*, 175–181. doi:10.1111/j.1365-246x.2007.03429.x

CHIOU, B. S. -J., YOUNGS, R. R. (2008), *Chiou and Youngs PEER-NGA empirical ground motion model for the average horizontal component of peak acceleration and pseudo-spectral acceleration for spectral periods of 0.01 to 10 seconds*. Earthq Spectra *24*(S1), 173–216

CORNELL, C. (1968), *Engineering seismic risk analysis*. Bull Seismol Soc Am *58*, 1583–1606

DAY, S. M., GRAVES, R. W., BIELAK, J., DREGER, D., LARSEN, S., OLSEN, K. B., PITARKA, A., RAMIREZ-GUZMAN, L. (2008), *Model for basin effects on long-period response spectra in southern California*. Earthq Spectra *24*, 257–277

DEELMAN, E., CALLAGHAN, S., FIELD, E., FRANCOEUR, H., GRAVES, R., GUPTA, N., GUPTA, V., JORDAN, T. H., KESSELMAN, C., MAECHLING, P., MEHRINGER, J., MEHTA, G., OKAYA, D., VAHI, K., ZHAO, L. I. (2006), *Managing large-scale workflow execution from resource provisioning to provenance tracking: The cybershake example*, 2ⁿᵈ. International conference on e-science and grid computing, Amsterdam.

FIELD, E. H. (2000), *A modified ground-motion attenuation relationship for Southern California that accounts for detailed site classification and a basin-depth effect*. Bull Seismol Soc Am *90*, S209–S221

FIELD, E. H., JORDAN, T. H., CORNELL, C. A. (2003). *OpenSHA: A developing community-modeling environment for seismic hazard analysis*. Seismol Res Lett *74*, 406–419

FIELD, E. H., DAWSON, T. E., FELZER, K. R., FRANKEL, A. D., GUPTA, V., JORDAN, T. H., PARSONS, T., PETERSEN, M. D., STEIN, R. S., WELDON II, R. J., WILLS, C. J. (2009), *Uniform California earthquake rupture forecast, version 2 (UCERF 2)*. Bull Seismol Soc Am *99*, 2053–2107. doi:10.1785/0120080049

GRAVES, R. (1996), *Simulating seismic wave propagation in 3D elastic media using staggered grid finite differences*, Bull Seismol Soc Am *86*, 1091–1106

GRAVES, R. W. (2008), *The seismic response of the San Bernardino basin region during the 2001 Big Bear Lake earthquake*. Bull Seismol Soc Am *98*, 241–252

GRAVES, R., AAGAARD, B., HUDNUT, K., STAR, L., STEWART, J., JORDAN, T. H. (2008), *Broadband simulations for Mw 7.8 southern San Andreas Earthquakes: Ground motion sensitivity to sputure

speed*. Geophys Res Lett *35*, L22302. doi:10.1029/2008GL035750

GRAVES, R. W., WALD, D. J. (2001), *Resolution analysis of finite fault source inversion using 1D and 3D Green's functions, part 1: strong motions*. J Geophys Res *106*, 8745–8766

GUATTERI, M., MAI, P., BEROZA, G. (2004), *A pseudo-dynamic approximation to dynamic rupture models for strong ground motion prediction*. Bull Seism Soc Am *94*, 2051–2063

HANKS, T. C., BAKUN, W. H. (2007), *M-log A observations for recent large earthquakes*. Bull Seismol Soc Am (in press)

JORDAN, T. H., MAECHLING, P. (2003), *The scec community modeling environment—an information infrastructure for system-level earthquake science*. Seismol Res Lett *74*, 324–328

MAI, M., BEROZA, G. (2002), *A spatial random field model to characterize complexity in earthquake slip*. J Geophys Res *107*, B11

MAYHEW, J. E., OLSEN, K. B. (2010), *Goodness-of-fit criteria for broadband synthetic seismograms, with application to the 2008 Mw 5.4 Chino Hills, CA, earthquake*. Seismol Res Lett (submitted)

O'CONNELL, D. R. H., LAFORGE, R., LIU, P. -C. (2007), *Probabilistic ground motion assessment of balanced rocks in the Mojave Desert*. Seismol Res Lett, *78*, 649–662

OLSEN, K., DAY, S., MINSTER, J., CUI, Y., CHOURASIA, A., FAERMAN, M., MOORE, R., MAECHLIN, P., JORDAN, T. (2006), *Strong shaking in Los Angeles expected from a southern San Andreas earthquake*. Geophys Res Lett *33*, L07305

OLSEN, K. B., DAY, S. M., MINSTER, J. B., CUI, Y., CHOUASIA, A., MOORE, R., MAECHLING, P., JORDAN, T. (2008), *TeraShake2: Simulation of Mw 7.7 earthquakes on the southern San Andreas fault with spontaneous rupture description*. Bull Seismol Soc Am *98*, 1162–1185. doi:10.1785/0120070148

OLSEN, K. B., DAY, S. M., DALGUER, L. A., MAYHEW, J., CUI, Y., ZHU, J., CRUZ-ATIENZA, V. M., ROTEN, D., MAECHLING, P., JORDAN, T. H., OKAYA, D., CHOURASIA, A. (2009), *ShakeOut-D: Ground motion estimates using an ensemble of large earthquakes on the southern San Andreas fault with spontaneous rupture propagation*. Geophys Res Lett *36*, L04303. doi:10.1029/2008GL036832

PLESCH, A., SHAW, J. H., BENSON, C., BRYANT, W. A., CARENA, S., COOKE, M., DOLAN, J., FUIS, G., GATH, E., GRANT, L., HAUKSSON, E., JORDAN, T. H., KAMERLING, M., LEGG, M., LINDVALL, S., MAGISTRALE, H., NICHOLSON, C., NIEMI, N., OSKIN, M., PERRY, S., PLANANSKY, G., ROCKWELL, T., SHEARER, P., SORLIEN, C., SÜSS, M. P., SUPPE, J., TREIMAN, J., YEATS, R. (2007), *Community fault model (CFM) for Southern California*. Bull Seismol Soc Am *97*, 1793–1802. doi:10.1785/0120050211

SCHMEDES, J., ARCHULETA, R. J., LAVALLE, D. (2010), *Correlation of earthquake source parameters inferred from dynamic rupture simulations*. J Geophys Res. doi:10.1029/2009JB006689 (in press)

SOMERVILLE, P., IRIKURA, K., GRAVES, R., SAWADA, S., WALD, D., ABRAHAMSON, N., IWASAKI, Y., KAGAWA, T., SMITH, N., KOWADA, A. (1999), *Characterizing crustal earthquake slip models for the prediction of strong ground motion*. Seismol Res Lett *70*, 199–222

SOMERVILLE, P. (2006), Review of magnitude-area scaling of crustal earthquakes, *Report to 2007 WGCEP*, URS Corporation, Pasadena, CA, pp 22

SONG, S. G., PITARKA, A., SOMERVILLE, P. (2009), *Exploring spatial coherence between earthquake source parameters*. Bull Seismol Soc Am *99*, 2564–2571

TAPE, C., LIU, Q., MAGGI, A., TROMP, J. (2010), *Seismic tomography of the southern California crust based on spectral-element and adjoint methods*. Geophys J Int *180*, 433–462

WELLS, D. L., COPPERSMITH, K. J. (1994), *New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement*. Bull Seismol Soc Am *84*, 974–1002

WORKING GROUP ON CALIFORNIA EARTHQUAKE PROBABILITIES. (2003), Earthquake Probabilities in the San Francisco Bay Region: 2002–2031, *USGS Open-File Report 2003-214*

ZHAO, L., CHEN, P., JORDAN, T. H. (2006), *Strain Green's tensors, reciprocity, and their applications to seismic source and structure studies*. Bull Seismol Soc Am *96*, 1753–1763