An Overview of the SCEC CyberShake Project

Thomas H. Jordan

University of Southern California


Meeting of the SCEC Committee for the Utilization of Ground Motion Simulations

03 Nov 2014
Probabilistic Seismic Hazard Model

Working Group on California Earthquake Probabilities (2007)
Uniform California Earthquake Rupture Forecast (UCERF2)

Earthquake Rupture Forecast \( P(S_k) \) \( \rightarrow \) Ground Motion Model \( P(Y_n / S_k) \) \( \rightarrow \) Intensity Measures \( P(Y_n) \)
**Probabilistic Seismic Hazard Model**

**Hazard Curve:**
- Shaking intensity: Peak Ground Acceleration (PGA)
- Interval: 50 years
- Site: Downtown LA

**Intensity Measures**

- **Ground Motion Model**
  - $P(S_k)$
  - $P(Y_n / S_k)$
  - $P(Y_n)$

**Earthquake Rupture Forecast**

**Hazard Curve**
- 2% in 50 yr
- 10% in 50 yr
Probabilistic Seismic Hazard Model

National Seismic Hazard Map

PGA (%g) with 2% Probability of Exceedance in 50 years

Earthquake Rupture Forecast \( P(S_k) \)  
Ground Motion Model \( P(Y_n / S_k) \)  
Intensity Measures \( P(Y_n) \)
Probabilistic Seismic Hazard Model

Few data epistemic uncertainty

High scatter aleatory variability

Boore et al. (1997) Empirical Ground Motion Prediction Equations (GMPEs)

Earthquake Rupture Forecast → Ground Motion Model → Intensity Measures

\[ P(S_k) \quad P(Y_n / S_k) \quad P(Y_n) \]
Much of the aleatory variability in the GMPEs comes from 3D heterogeneity in crustal structure.
NGA (2008) Attenuation Relations used in National Seismic Hazard Maps

**Epistemic Uncertainties**

- **near-fault amplitudes**
- **basin effects**

**SA-3s**

PE = 2%/50 yr

UCERF2, no background seismicity
ShakeOut Scenario
M7.8 Earthquake on Southern San Andreas Fault
CyberShake Model: Physics-Based PSHA

KFR = kinematic fault rupture model
AWP = anelastic wave propagation model
NSR = nonlinear site response
CyberShake Model: Physics-Based PSHA

- **Sites:**
  - 283 sites in the greater Los Angeles region

- **Ruptures:**
  - All UCERF2 ruptures within 200 km of site (~14,900)

- **Rupture variations:**
  - ~415,000 per site using Graves-Pitarka pseudo-dynamic rupture model

- **Seismograms:**
  - ~235 million per model
NGA (2008) Attenuation Relations used in National Seismic Hazard Maps

PE = 2%/50 yr
UCERF2, no background seismicity
**CyberShake Platform: Physics-Based PSHA**

**Essential ingredients**

1. **Extended earthquake rupture forecast**
   - probabilities of all fault ruptures (e.g., UCERF2)
   - conditional hypocenter distributions for rupture sets
   - conditional slip distributions from pseudo-dynamic models

2. **Three-dimensional models of geologic structure**
   - large-scale crustal heterogeneity
   - sedimentary basin structure
   - near-surface properties (“geotechnical layer”)

3. **Ability to compute large suites (> 10^8) of seismograms**
   - efficient anelastic wave propagation (AWP) codes
   - reciprocity-based calculation of ground motions

from SCEC CVMs
Uniform California Earthquake Rupture Forecast (UCERF2)

Probability gain estimated from date of last event according to a Reid-type stress renewal model (BPT model)

Ratio of time-dependent to time-independent participation probabilities for $M \geq 6.7$
CyberShake Rupture Models
Conditional Slip Distribution
Graves-Pitarka Pseudo-Dynamic Rupture Models

GP07 used in CS11
GP10 used in CS13
**SCEC Community Velocity Models (CVMs)**

Data sources
- Surface geology
- Well logs
- Refraction surveys
- Reflection surveys
- Seismic tomography
- Geologic models
Basin Structures of Three SCEC CVMs

CVM–S4

CVM–S4.26

CVM–H11.9

$Z_{2500}$ : iso-velocity surfaces at $V_S = 2.5\, \text{km/s}$
Rapid Simulation of Large Rupture Ensembles
Using Seismic Reciprocity

- To account for source variability requires very large sets of simulations
  - 14,900 ruptures from UCERF2; 415,000 rupture variations
- Ground motions need only be calculated at much smaller number of surface sites to produce hazard map
  - 283 in LA region, interpolated using empirical attenuation relations
- **Use of reciprocity reduces CPU time by a factor of ~1,000**

\[
\text{M sources to } N \text{ sites requires } M \text{ simulations}
\]
\[
\text{M sources to } N \text{ sites requires } 2N \text{ or } 3N \text{ simulations}
\]
CyberShake Workflow

Mesh generation
1 job per site
MPI, 320 cores

SGT computation
2 jobs per site
MPI, 10k CPUs or 100 GPUs

Post-processing
415,000 jobs per site
serial

Data Product Generation
Populate DB, construct queries
4 jobs per site

Uniform California Earthquake Rupture Forecast

Full-3D tomographic model of S. California

scismograms

hazard curves

hazard maps
CyberShake Hazard Map Interpolation

Campbell & Borzognia (2008) GMPE with CGS soil map

CyberShake (2011) differences

CyberShake (2011) map

3-s Spectral Acceleration (in g) at Probability of Exceedance = 2% in 50 yr
Comparison of 1D and 3D CyberShake Models for the Los Angeles Region

1. lower near-fault intensities due to 3D scattering
2. much higher intensities in near-fault basins
3. higher intensities in the Los Angeles basins
4. lower intensities in hard-rock areas
Seismological Hierarchy of CyberShake

\[ G(r, k, x, s) = \ln Y(r, k, x, s) \]

- **Site set:** \( r \) \( R \)
  - 283 sites in the greater Los Angeles region
  - Elastic structures: BBP-1D, CVM-S4, CVM-H11, or CVM-S4.26

- **Rupture set:** \( k \) \( K(r) \)
  - All UCERF2 ruptures within 200 km of site (~7000 total)

- **Conditional hypocenter distribution:** \( x \) \( X(r, k) \)
  - Uniform distribution along fault strike with \( \Delta x \approx 20 \text{ km} \)

- **Conditional slip distribution:** \( s \) \( S(r, k, x) \)
  - Approximately 415,000 rupture variations per site, 235 million synthetic seismograms per model (2 horizontal components)
Averaging-Based Factorization
(Wang & Jordan, BSSA, 2014)

• **Representation of excitation functionals**

Expected shaking intensities constructed by averaging over slip variations \((s)\), hypocenters \((x)\), sources \((k)\), and sites \((r)\)

\[
G(r, k, x, s) = A + B(r) + C(r, k) + D(r, k, x) + E(r, k, x, s)
\]

\[\uparrow \quad \text{level} \quad \uparrow \quad \text{site} \quad \uparrow \quad \text{path} \quad \uparrow \quad \text{directivity} \quad \uparrow \quad \text{slip complexity}
\]

• **Representation of excitation variance**

\[
\text{Var}[G] = \bar{\sigma}_G^2 \equiv \left\langle [G(r, k, x, s) - A]^2 \right\rangle_{S, X, K, R}
\]

\[
= \sigma_B^2 + \left\langle \sigma_C^2(r) \right\rangle_R + \left\langle \sigma_D^2(r, k) \right\rangle_{K, R} + \left\langle \sigma_E^2(r, k, x) \right\rangle_{X, K, R}
\]

\[
= \sigma_B^2 + \bar{\sigma}_C^2 + \bar{\sigma}_D^2 + \bar{\sigma}_E^2
\]
ABF Variance Analysis of the CyberShake Model

CS13.b - NGA08

Redducible (epistemic) variance

<table>
<thead>
<tr>
<th>Period (s)</th>
<th>Site Variance ($\sigma_b^2$)</th>
<th>Path Variance ($\sigma_c^2$)</th>
<th>Directivity Variance ($\sigma_d^2$)</th>
<th>Magnitude Variance ($\sigma_m^2$)</th>
<th>Source Complexity Variance ($\sigma_j^2$)</th>
<th>Total Variance ($\sigma_T^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td>3.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td>5.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td>10.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.6</td>
</tr>
</tbody>
</table>

NGA $\sigma_T^2$
Importance of Reducing Aleatory Variability

\[
\ln Y (r, k, x, m; \varepsilon) = \ln \bar{Y} (r, k, x, m) + \sigma_T \varepsilon
\]

\[
\sigma_T = 0.43, 0.48, 0.52, 0.57, 0.62
\]

SA-3s (g)

Annual Frequency of Exceedance

x80 reduction
NGA(2014)-CyberShake Hazard Curve Comparisons

Site LADT
(Los Angeles)
Dependence of Basin Effects on Velocity Structures
(FA corrected for $V_{S30}$ using BA08)

CVM-S4.26

CVM-H11

T=3.0s

T=5.0s

T=10.0s

Abrahamson & Silva (2008) NGA GMPEs

Abrahamson & Silva (2008) NGA GMPEs

CS14b

CS13b
CVM-S4.26
Full-3D tomography model of Southern California crustal structure

- CVM-S4 starting model
- 26th iterate of a full-3D tomographic (F3DT) inversion procedure (Lee et al., 2013).
- Data sets comprise ~ 550,000 differential waveform measurements at \( f \leq 0.2 \) Hz
  - 38,000 earthquake seismograms
  - 12,000 ambient-noise Green functions
- Nonlinear iterative process involved two methods:
  - adjoint-wavefield (AW-F3DT)
  - scattering-integral (SI-F3DT)
Full-3D Waveform Tomography

(Lee, Chen, Jordan, Maechling, Denolle & Beroza, JGR, 2014)
LARSE Profiles
Test of CVM-S4.26 synthetics against data from the 03/28/14 La Habra Earthquake (M5.1)

Data in black
Synthetics in red
03/28/14 La Habra Earthquake (M5.1)

Station MOP
Observed in black
Synthetic in red

CS11: CVM-S4
CS14.2: CVM-S4.26
CS13.4: CVM-H11.9
03/28/14 La Habra Earthquake (M5.1)

Station SDD
Observed in black
Synthetic in red
03/28/14 La Habra Earthquake (M5.1)

Station EDW2
Observed in black
Synthetic in red

CS11: CVM-S4
CS14.2: CVM-S4.26
CS13.4: CVM-H11.9
03/28/14 La Habra Earthquake (M5.1)
03/17/14 Encino Earthquake (M4.4)
03/18/14 La Habra Earthquake (M5.1)
(Taborda et al., 2014)
CyberShake: Initiative to Compute a Statewide Physics-Based Hazard Model

- Extend CyberShake models to 1400 sites across California
  - Develop statewide Unified Community Velocity Model (UCVM)
  - Compute site response to 1 Hz deterministic, 10 Hz stochastic

- Couple time-dependent UCERF3 to CyberShake
  - Provide frequently updated time-dependent seismic hazard maps

- Extend CSEP to prospectively test ground motion forecasts against observations throughout California

Statewide CyberShake

- Computational requirements for 1 Hz deterministic, 10 Hz stochastic:
  - Number of jobs: 23.2 billion
  - Storage: 2800 TB seismograms
  - Computer hours: 392 million
End