Societal Impact of Earthquake Simulations at Extreme Scales

Thomas H. Jordan

Director, Southern California Earthquake Center

SC15 Plenary Lecture, 19 Nov 2015


SCEC Community Modeling Environment – A National Collaboration

- National computing centers
- SCEC core institutions (18)
SCEC/CME Use of HPC Resources

362M core-hr in 2015 from NSF and DOE resources

- What has society actually gained from computational earthquake science?
- Has this public investment been worth it?

<table>
<thead>
<tr>
<th>Year</th>
<th>Service Units on NSF TeraGrid/XSEDE Resources</th>
<th>CPU Hours on Blue Waters</th>
<th>Service Units on NSF Yellowstone</th>
<th>Service Units on DOE INCITE Resources</th>
<th>Service Units on USC Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2004</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2006</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2008</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2012</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2014</td>
<td>362M</td>
<td>362M</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Los Angeles will have the nation's toughest earthquake safety rules

by Rong-Gong Lin II

October 9, 2015, 5:38 pm

In a stark recognition of Los Angeles’ vulnerability to a major earthquake, the city on Friday enacted the nation’s most sweeping seismic regulations, requiring an estimated 15,000 buildings be retrofitted so they will better withstand violent shaking.

The unanimous vote by the Los Angeles City Council caps decades of efforts to strengthen the two types of buildings that pose the most serious potential for loss of life in a big quake…
Los Angeles will have the nation's toughest earthquake safety rules

Soft first story buildings

Non-ductile concrete buildings

Oct 9, 2015

Lucy Jones celebrates with City Atty. Mike Feuer, center, as Mayor Eric Garcetti signs sweeping legislation to require earthquake retrofits on 15,000 buildings in Los Angeles.
Resilience by Design
Report of the Los Angeles Mayoral Seismic Task Force (Lucy Jones, chair)
Released Dec 8, 2014

An ambitious plan to
– strengthen buildings
– fortify water supply and distribution system
– enhance reliable telecommunications
Resilience by Design
Report of the Los Angeles Mayoral Seismic Task Force (Lucy Jones, chair)
Released Dec 8, 2014

“This Report’s approach to evaluating the severity of the risk relies on the ShakeOut Scenario... created by a multidisciplinary team convened by the Multi-Hazards Demonstration Project of the USGS...”

Team included USGS, CGS, FEMA, SCEC, and nearly 200 other partners in government, academia, emergency response, and industry.
**The ShakeOut Scenario**

M7.8 Earthquake on Southern San Andreas Fault

**Scenario Results**

- M7.8 mainshock
  - Broadband ground motion simulation (0-10 Hz)
- Large aftershocks
  - M7.2, M7.0, M6.0, M5.7...
- 10,000-100,000 landslides
- 1,600 fire ignitions
- $213 billion in direct economic losses
  - 300,000 buildings significantly damaged
  - Widespread infrastructure damage
  - 270,000 displaced persons
  - 50,000 injuries
  - 1,800 deaths
- Long recovery time

**Great Southern California ShakeOut**

November 13, 2008

**Exercise Results**

- Largest emergency response exercise in US history
  - Golden Guardian exercise
  - Public events involving 5.3 million registered participants
- Demonstrated that existing disaster plans are inadequate for an event of this scale
  - Motivated reformulation of system preparedness and emergency response
- Scientific basis for the LA Seismic Safety Task Force report, *Resilience by Design*
The ShakeOut Scenario
M7.8 Earthquake on Southern San Andreas Fault

Scenario Results

• Resilience study by Los Angeles Department of Water and Power (LADWP)

  • predicted severe mainshock damage to Los Angeles water supply and distribution systems, compounded by aftershock damage

  • promoted the planning of actions to mitigate risk and improve system resilience


Great Southern California ShakeOut
November 13, 2008
The ShakeOut Scenario

M7.8 Earthquake on Southern San Andreas Fault

Scenario Results

- Resilience study by Los Angeles Department of Water and Power (LADWP)
- predicted severe mainshock damage to Los Angeles water supply and distribution systems, compounded by aftershock damage
- promoted the planning of actions to mitigate risk and improve system resilience
The ShakeOut Scenario

M7.8 Earthquake on Southern San Andreas Fault

Scenario Results

- Resilience study by Los Angeles Department of Water and Power (LADWP)
  - predicted severe mainshock damage to Los Angeles water supply and distribution systems, compounded by aftershock damage
  - promoted the planning of actions to mitigate risk and improve system resilience

Much has been learned from this and other virtual earthquakes about how to reduce risk and improve resilience

- Beats waiting to learn tragically from the real thing!
2015 ShakeOut Earthquake Drills

2015 Official ShakeOut Regions
28 Regions worldwide
22 U.S. regions spanning 51 states & territories
65 additional countries with independent registrations (individuals, schools, etc.)

Participation History (worldwide)
2015: 43.8 million (+ TX, IA, LA, NE, global growth)
2014: 26.5 million (+ NM, KS, FL, Quebec, Yukon, more)
2013: 25.0 million (+ Southeast, Northeast, MT, WY, CO)
2012: 19.5 million (+ Japan, New Zealand, UT, WA, AZ)
2011: 12.5 million (+ Central US, BC, OR)
2010: 8.0 million (+ Nevada and Guam)
2009: 6.9 million (+ Northern California)
2008: 5.4 million (Southern California)

Key Facts
- Participants practice “Drop, Cover, and Hold On” and other aspects of their emergency plans.
- Register at www.ShakeOut.org
- Largest component “America’s PrepareAthon,” sponsored by FEMA

In 2015, more than 43 million people registered for ShakeOut to improve their earthquake preparedness
Societal Impact of Earthquake Simulations

- Earthquake simulations enabled by HPC
  - TeraShake simulations, 2006
  - ShakeOut scenario, 2008
  - M8 scenario, 2010

- Full-scale physics-based hazard models
  - Probabilistic seismic hazard analysis
  - CyberShake hazard models, 2011-2015
  - Building Seismic Safety Council Project 2017
  - Reducing epistemic uncertainty: the Central California Seismic Project

- Beyond petascale
  - Push to higher frequencies: the High-F Project
  - Towards more realistic simulations
Probabilistic Seismic Hazard Analysis

- SCEC-USGS-CGS Working Group on California Earthquake Probabilities has developed the Uniform California Earthquake Rupture Forecast (UCERF)
  - Used by engineers, emergency managers, urban planners, and insurance industry; supported by the California Earthquake Authority (CEA)
  - UCERF3, computed on TACC Stampede (Field et al., 2014), has been incorporated into the USGS 2014 National Seismic Hazard Model

- SCEC ground motion simulations have informed standardized ground motion prediction models developed under the NGA Project of the PEER Center
Probabilistic Seismic Hazard Analysis

- SCEC-USGS-CGS Working Group on California Earthquake Probabilities has developed the Uniform California Earthquake Rupture Forecast (UCERF)
  - Used by engineers, emergency managers, urban planners, and insurance industry; supported by the California Earthquake Authority (CEA)
  - UCERF3, computed on TACC Stampede (Field et al., 2014), has been incorporated into the USGS 2014 National Seismic Hazard Model

- SCEC ground motion simulations have informed standardized ground motion prediction models developed under the NGA Project of the PEER Center

The NGA and UCERF studies provided CEA with the science basis for insurance rate reductions of 22.1% (2006) and 12.5% (2012)

- Earthquake simulation and modeling have helped to reduce the cost of earthquake insurance in California by 1/3
Probabilistic Seismic Hazard Analysis

- PSHA, as currently practiced, is based on empirical statistical models
  - Persistent uncertainties due to unexplained variance between GMPEs and real datasets
- We seek to improve earthquake forecasting by incorporating more physics through numerical simulations

Diagram:
- **RSQsim** (Earthquake Rupture Simulator)
- **CyberShake** (Ground Motion Simulator)
- **Ground Motions**
- **Intensity Measures**
- **NGA-W2 GMPEs**
- **UCERF3**
- **Uniform California Earthquake Rupture Forecast (UCERF3)**
- **U.S. National Seismic Hazard Map**
**Essential Ingredients for Ground Motion Predictions**

1. **Earthquake source descriptions**
   - detailed representation of fault geometry
   - rupture models that capture the complexities of dynamic fault failure

2. **Three-dimensional models of geologic structure**
   - large-scale crustal heterogeneity
   - sedimentary basin structure
   - geotechnical layer based on $V_s 30$

3. **Calculation of wave propagation and attenuation**
   - efficient anelastic wave propagation (AWP) codes
   - nonlinear models of near-surface response
Essential Ingredients for Ground Motion Predictions

1. Earthquake source descriptions
   - detailed representation of fault geometry
   - rupture models that capture the complexities of dynamic fault failure

2. Three-dimensional models of geologic structure
   - large-scale crustal heterogeneity
   - sedimentary basin structure
   - geotechnical layer based on $V_s30$

3. Calculation of wave propagation and attenuation
   - efficient anelastic wave propagation (AWP) codes
   - nonlinear models of near-surface response

Anelastic Wave Propagation Code (Olsen, Day & Cui)
13-point asymmetric stencil computation
3D velocity-stress wave equations solved by explicit staggered-grid 4th-order FD.

$$\frac{\partial \mathbf{v}}{\partial t} = \frac{1}{\rho} \nabla \cdot \mathbf{\sigma}$$
$$\frac{\partial \mathbf{\sigma}^E}{\partial t} = \lambda (\nabla \cdot \mathbf{v}) \mathbf{I} + \mu (\nabla \mathbf{v} + \nabla \mathbf{v}^T)$$

Memory variable formulation of anelastic relaxation using coarse-grained representation
$$\frac{\partial \mathbf{\sigma}}{\partial t} = \frac{\partial \mathbf{\sigma}^E}{\partial t} - \delta M \sum_i \mathbf{\varepsilon}_i$$

$$\tau_i \frac{d\mathbf{\varepsilon}_i(t)}{dt} + \mathbf{\varepsilon}_i(t) = \lambda_i \frac{\delta M}{M_u} e(t)$$

$$Q^{\text{i}}(\mathbf{M}) = \left( \mathbf{M}_u \right)^N_{i=1} \left( \frac{e(t)}{\sum_{i=1}^N e_i(t) + 1} \right)$$

- Dynamic rupture by the staggered-grid split-node (SGSN) method
- Absorbing boundary conditions by perfectly matched layers (PML)
SCEC Extreme-Scale Earthquake Computations

Year


0.01 TF/s 0.10 TF/s 1 TF/s 10 TF/s 100 TF/s 1,000 TF/s 10,000 TF/s 100,000 TF/s 1,000,000 TF/s

Sustained SCEC

Top 500 No. 1 Peak Capability

Peak DOE Capability

Peak NSF

Sustained DOE

Sustained NSF

Sustained SCEC

High-F Project

near-fault plasticity

topography

small-scale heterogeneity

Q (f)
TeraShake Simulations of M7.7 Earthquake on the San Andreas Fault

Empirical GMPE

Physics-Based Simulation

Directivity-basin coupling

Peak ground velocity (PGV) map

TeraShake Simulations (M7.7) by Olsen et al. (2006, 2008) demonstrated the importance of directivity-basin coupling.
Simulations indicated strong 3D focusing of ground motions from strike-slip faults in Southern California

- Quantified the importance of source directivity and basin excitation effects in earthquake forecasting and PSHA
The San Andreas Fault System is a major transforms plate boundary in southern California. It is the most Active seismic zone in the United States. The fault runs from the Gulf of California to northern California. The fault has produced several large earthquakes, including the 1906 San Francisco earthquake and the 1857 Fort Tejon earthquake. The Pacific plate is moving northwest relative to North America at a rate of 5 meters (M7.8) per 100 years.
The southern San Andreas Fault is “locked, loaded and ready to roll”

Pacific plate is moving NW relative to North America at 5 meters (M7.8) per 100 years
ShakeOut Scenario – 2008

M7.8 earthquake simulation on Southern San Andreas Fault
(deterministic band $f = 0$-$1$ Hz; stochastic band $f = 1$-$10$ Hz)
Supershear Rupture of the San Andreas Fault

Southern California Earthquake Center
M8 Simulation – 2010

- Magnitude 8.0 “wall-to-wall” scenario on southern San Andreas Fault
  - Fault length: 545 km
  - Minimum wavelength: 200 m
- Dynamic rupture simulation on *Kraken*, 7.5 hours using 2160 cores
  - 881,475 subfaults, 250 sec of rupture
  - 2.1 TB tensor time series output
- Wave propagation simulation performed on *Jaguar*, 24 hours using 223,074 cores (220 Tflop/s sustained).
  - 436 billion grid points representing geologic model of dimension 810 x 405 x 85 km (40-m sampling)
  - 368 s of ground motions (160,000 steps of 0.0023 s)

Ground motions of outer-scale event (M8) computed deterministically up to 2 Hz
Performance of SCEC Large-Scale Simulations

4D outer/inner scale ratio

Sustained computational speed

- M8-2Hz
- M8-1Hz
- ShakeOut
- TeraShake

x3000 in 6 years

Computational Size (Mesh points x time steps)

- TeraShake
- ShakeOut
- M8-1Hz
- M8-2Hz
California’s San Andreas System

Probabilistic characterization of seismic hazard at a single site requires the consideration of ~500,000 rupture variations within this fault network.

Seismogram ensembles of this size can be efficiently calculated using seismic reciprocity.

“CyberShake Project”

Frequent large earthquakes on over 1000 major active faults threaten California’s population of 38 million people.
SCEC Computational Pathways

Structural Representation

1. Earthquake Rupture Forecast
   - FM
   - DM
   - ERM
   - PM

2. AWP
   - KFR

3. DFR
   - AWP

4. F3DT
   - Other Data
     - Geology
     - Geodesy

Ground Motions

Intensity Measures

Empirical GMPE

Other Data

- Geology
- Geodesy

TACC Stampede
- NCSA Blue Waters
- OLCF Titan
- ALCF Mira

T. H. Jordan
A complete CyberShake model for the Los Angeles region requires the execution of ~500 million jobs.

- We manage this workflow using Pegasus-WMS and other tools.

Pathway 1

- UCVM
  - Mesh generation
  - 1 job per site
  - MPI, 3840 cores

Pathway 2

- AWP-ODC
  - SGT computation
  - 2 jobs per site
  - MPI, 800 GPUs

Pathway 3

- SeisSynth
  - Post-processing
  - ~500,000 jobs per site
  - MPI master/worker, 3712 cores

Pathway 4

- Population DB, construct queries
  - 6 jobs per site

Data Product Generation

- 12 TB data transfer
- Populates the database
- Constructs queries
Comparison of 1D and 3D CyberShake Models for the LA 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
CyberShake 15.4 Study

Preliminary Los Angeles urban seismic hazard model

- Seismic frequency limit $f_{\text{max}}$ increased to 1 Hz
- Spatial sampling increased to 336 sites
  - > 330 million synthetic seismograms
  - > 22 billion shaking intensity measures
- Run on NCSA Blue Waters and OLCF Titan
- Computational makespan of 38 days
  - 795,000 GPU-hours, 19 million CPU-hours
    - 956,000 node-hours on NCSA Blue Waters
    - 428,000 node-hours on OLCF Titan
  - 3.2x code efficiency gain
  - post-processing file I/O reduced by 99.9%
  - > 1 petabyte of data managed
- Fully automated using workflow tools
  - 42 concurrent workflows running on average

Earthquake forecast: UCERF2
Structural model: CVM-S4.26
Community Velocity Model S4.26
Full-3D tomography (F3DT) model of Southern California Crustal Structure
(E.-J. Lee, P. Chen et al., 2014)

Seismic velocity structure developed on ALCF **Mira**

- Three-dimensional CVM-S4 starting model
- 26 F3DT iterations for $v_p$ and $v_s$ on 500-m grid
  - 3.8 million parameters
- 550,000 differential waveform measurements ($f \leq 0.2$ Hz)
- 38,000 earthquake seismograms
- 12,000 ambient-noise Green functions
- 954 full-wavefield simulations per iteration
- Inversion of centroid-moment tensors for 160 eqk sources
- $> 4$ PB of data managed
Community Velocity Model S4.26

Full-3D tomography (F3DT) model of Southern California Crustal Structure (E.-J. Lee, P. Chen et al., 2014)

Seismic velocity structure developed on ALCF *Mira*

- Three-dimensional CVM-S4 starting model
- 26 F3DT iterations for $v_p$ and $v_S$
- 550,000 differential waveform measurements ($f \leq 0.2$ Hz)
- 3.8 million parameters
- 38,000 earthquake seismograms
- 12,000 ambient noise Green functions
- 954 full-wavefield simulations per iteration
- Inversion of centroid-moment tensors for 160 eqk sources
- > 4 PB of data managed

Through data-intensive computing, F3DT has provided much better images of Earth structure

- New information of high scientific value
- High-fidelity predictions of ground motions from earthquakes
Test of CVM-S4.26 synthetics against data from the 03/28/14 La Habra Earthquake (M5.1)

Lee et al. (2014)
SCEC Committee for the Utilization of Ground Motion Simulations (C. B. Crouse, chair) is conducting a special study on use of CyberShake simulations in seismic design maps.
Project 17 of the Building Seismic Safety Commission

Development of Next Generation of Seismic Design Maps

The Committee on the Utilization of Ground Motion Simulations will review these maps on Nov 30 and make recommendations regarding their use in Project 17.
Uncertainties in the Ground Motion Prediction Equations

Few data epistemic uncertainty

Much scatter aleatory variability

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

Earthquake Rupture Forecast

Ground-Motion Prediction Eqn

Intensity Measures

UCERF
NGA GMPEs
Response Spectra
Uncertainties in the Ground Motion Prediction Equations

\[ \ln Y = \ln Y(x) + \sigma_T \varepsilon \]

At constant \( Y(x) \), increasing the aleatory variability \( \sigma_T \) increases the exceedance probabilities.
Uncertainties in the Ground Motion Prediction Equations
In 2015, PG&E, SCEC, and USGS initiated the “Central California Seismic Project” to test the ability of simulation-based PSHA to lower the epistemic uncertainties in PSHA.

- Achieving this reduction would have a broad impact on risk-reduction strategies for critical facilities such as large dams, nuclear power plants, and energy transportation networks.
CyberShake – A Layered Hazard Model

1. Hazard map

2. Hazard curves

3. Hazard disaggregation

4. Rupture model

5. Seismograms
CyberShake is a transformative step towards physics-based PSHA

- First complete PSHA model derived from simulations
- Leading model in USGS Urban Seismic Hazard Mapping Project
- Source of new design maps under consideration by BSSC Project 17 for revisions to building code
- Used as database for machine-learning of California Earthquake Early Warning system
- Pre-computed ground motions coupled to short-term UCERF3-ETAS model as prototype operational earthquake forecasting system
High-F Project: Push to Higher Frequencies

Seismic band

- low-order free oscillations
- mantle waves
- crustal waves
- basin waves
- strongly scattered waves

<table>
<thead>
<tr>
<th>Period (s)</th>
<th>Frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>.001</td>
</tr>
<tr>
<td>100</td>
<td>.01</td>
</tr>
<tr>
<td>10</td>
<td>.1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0.1</td>
<td>10</td>
</tr>
</tbody>
</table>

Earthquake engineering band

tall buildings  houses  stiff structures

SCEC simulations 2015
- physics-based deterministic
- High-F modeling
- fault roughness
- near-fault plasticity
- frequency-dependent attenuation
- validate
- topography
- new
- small-scale near-surface heterogeneity
- physics
- near-surface nonlinearity

CyberShake
- 1 Hz

SCEC simulations 2020
- physics-based deterministic
- 10 Hz

empirical stochastic

T. H. Jordan
11/19/15
Rough Fault Simulation

with and without small-scale crustal heterogeneity

• AWP-ODC FD code
• 443 billion mesh points
• 9.3 trillion degrees of freedom
• 100,000 time steps
• 2.3 Pflops sustained performance on OLCF Titan

Fractal fault roughness, 3D heterogeneous medium

Fractal fault roughness, 1D layered medium
ShakeOut Scenario: AWP with Drucker-Prager Plasticity

Societal Impacts of Earthquake Simulations

**Realized Impacts:**

- **ShakeOut scenario** was the hazard basis for “Resilience by Design”
  - 2015 LA city ordinances to retrofit dangerous buildings
  - Planning underway to fortify the water system and enhance reliable telecommunications

- **UCERF3** has become the PSHA standard for California
  - Used extensively by engineers, emergency managers, urban planners, and insurance industry; incorporated into the 2014 National Seismic Hazard Model

- **NGA and UCERF studies** were used by CEA to lower insurance rates
  - California homeowner insurance costs decreased by one-third

- **CyberShake** demonstrated feasibility and value of physics-based PSHA
  - 1-Hz model for Los Angeles region is a key component of the USGS Urban Seismic Hazard Mapping Program
  - Quantified the effects of source directivity and basin excitation

- **CyberShake MCE$_R$ maps** are under consideration by BSSC Project 17
  - For use in designing tall buildings and large structures
**Societal Impacts of Earthquake Simulations**

**Potential Impacts:**

- **Use simulations to evaluate risk to distributed infrastructure**
  - LADWP Water System Resilience Program

- **Use simulations to reduce PSHA uncertainties**
  - PG&E-sponsored Central California Seismic Project

- **Push deterministic simulations to 10 Hz**
  - Validate new physics codes for modeling fault roughness, near-fault plasticity, frequency-dependent attenuation, nonlinear near-surface response, and small-scale near-surface heterogeneities

- **Apply physics-based forecasting methods outside of California**
  - Urban seismic hazard models for Seattle, Salt Lake City, and Memphis

- **Apply pre-computed ground-motion simulations in time-dependent hazard estimation**
  - Operational earthquake forecasting
  - Earthquake early warning
  - Post-earthquake response
Future SCEC Computational Requirements

- **5-Year Target**
- **Current**

### Computational Size (mesh points X time steps)

- **10-Hz ShakeOut**
- **4-Hz LA basin**
- **1-m Fault resolution**
- **0.2-Hz South California**
- **M1.5+ 10 million years**
- **M4.5+ 1 million yrs**

### Onward to Exascale!

- **<24 hrs**
- **<2 weeks ensemble**
- **<24 hrs**
- **<10 days ensemble**
- **<24 hrs**

### Key Points
- **7 orders of magnitude more computation required**
Thank you!
Recommendations for the National Strategic Computing Initiative

1. Federal agencies should improve mechanisms to sustain deep collaborations among domain scientists and computer scientists to advance extreme-scale computing
   – Support software engineering teams dedicated to solving extreme-scale scientific problems

2. NSF and DOE should form a strategic alliance in exascale computing
   – Coordinate allocation of HPC resource and domain-science funding

3. NSF and DOE should sponsor co-design activities for adapting scientific production codes to new machine architectures
   – Qualify codes for production runs before full-scale machines are opened for production
   – Provide early feedback to optimize system design