An Overview of the SCEC Earthquake Simulation Program and the CyberShake Project

Thomas H. Jordan

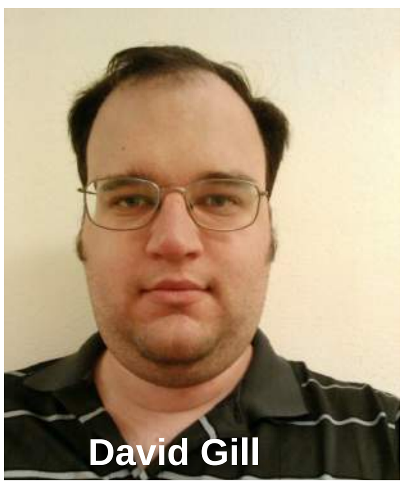
University of Southern California

CyberShake co-developers: S. Callaghan, R. Graves, F. Wang, K. Olsen, K. Milner, and P. Maechling, En-Jui Lee, Po Chen

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

12 May 2014

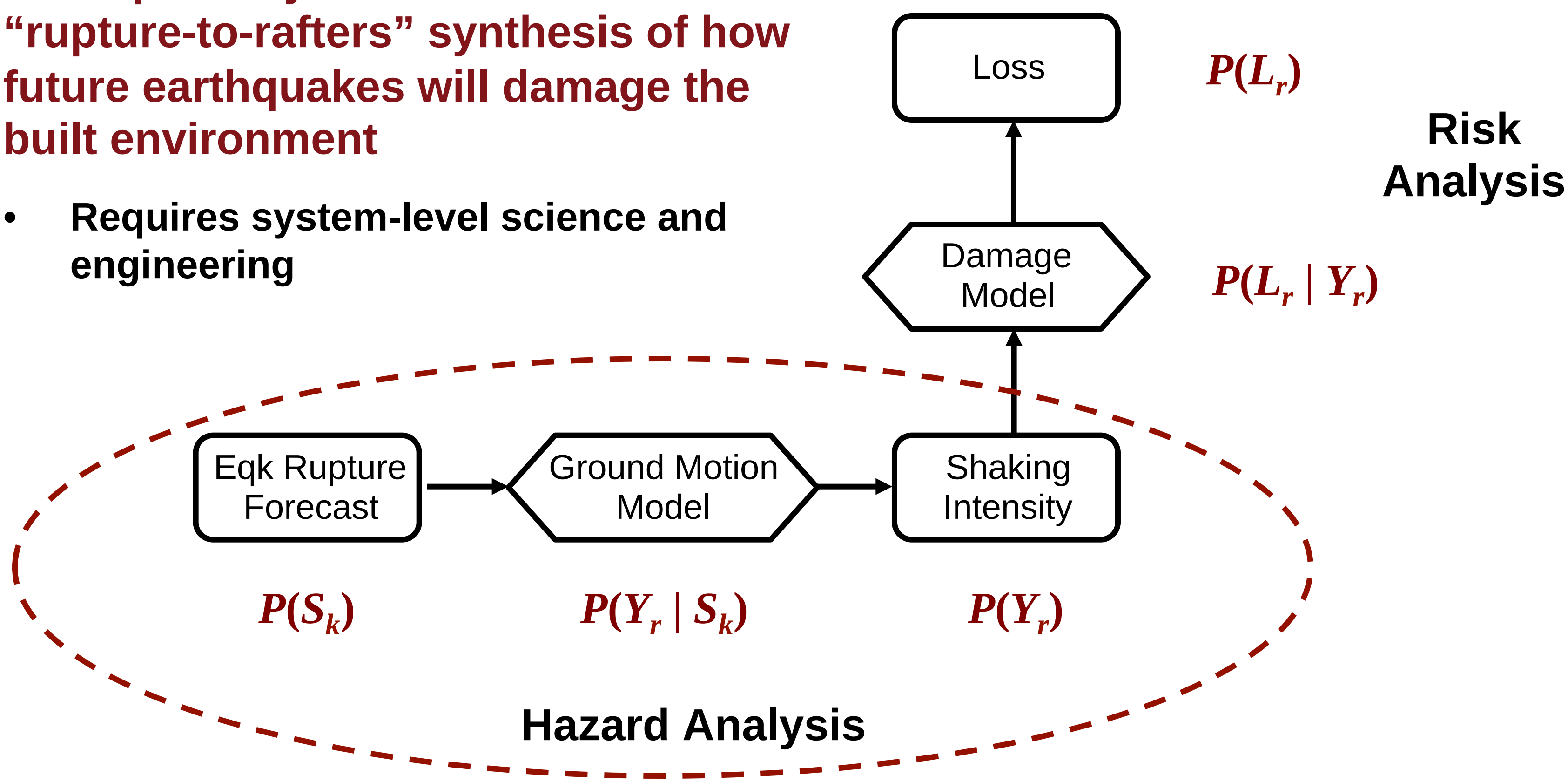
SCEC IT/CS Staff



Earthquake System Science

Earthquake system science seeks a "rupture-to-rafters" synthesis of how future earthquakes will damage the built environment

- **Requires system-level science and engineering**



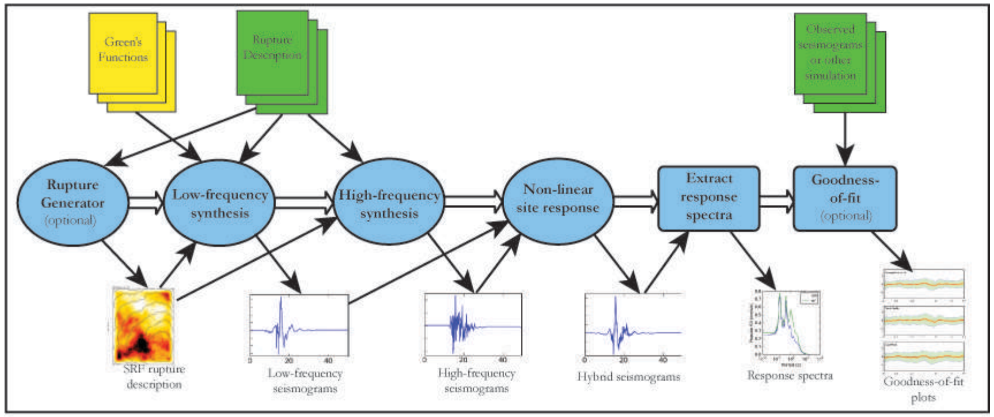
Presentation Outline

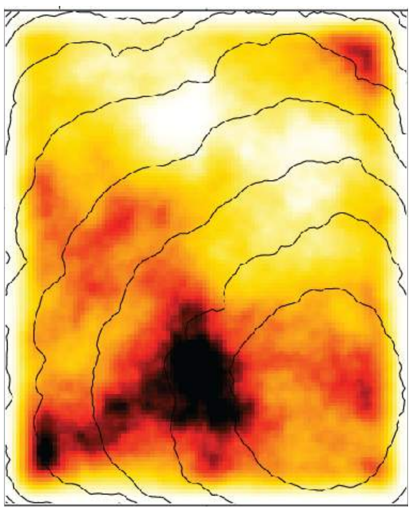
1. Overview of the SCEC earthquake simulation program
2. Formulation of simulation-based PSHA
3. Introduction to the CyberShake computational platform
4. 3D seismic velocity structure from full-3D inversion
5. Comparisons Among CyberShake Models and NGA GMPE
6. Plans for future CyberShake reserarch

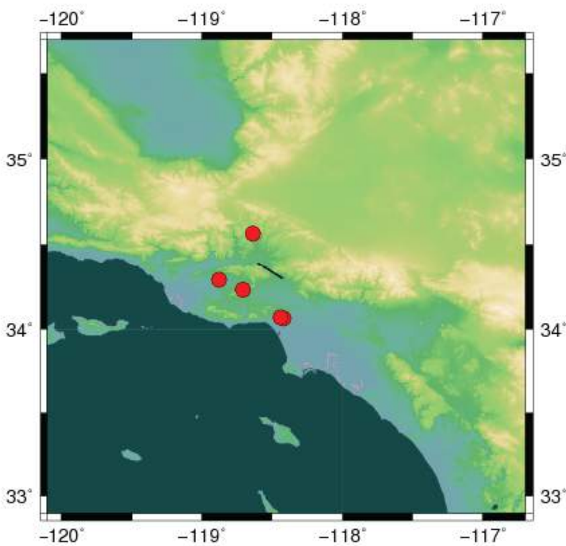
Overview of the SCEC Earthquake Simulation Program

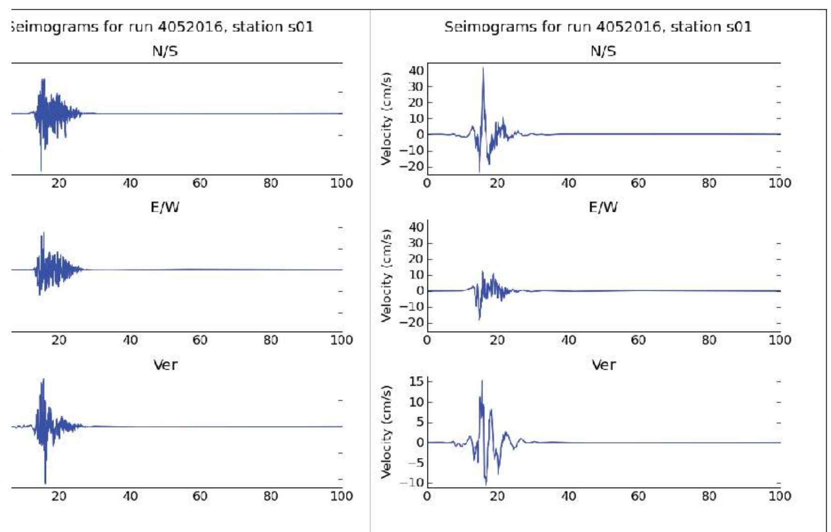
Broadband Platform

Ground motion modeling software for frequencies from 0.0 to 10Hz, designed for flexible use by earthquake engineers









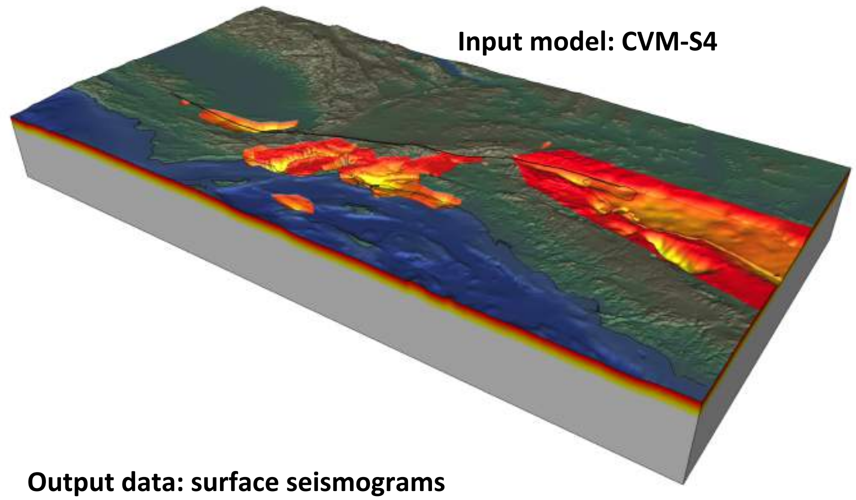
Rupture models Station and fault trace maps Acceleration and velocity seismograms

M8 Simulation

- **Magnitude 8.0 wall-to-wall scenario, worst-case for southern San Andreas Fault**
  - **Fault length: 545 km**
  - **Minimum wavelength: 200 m**

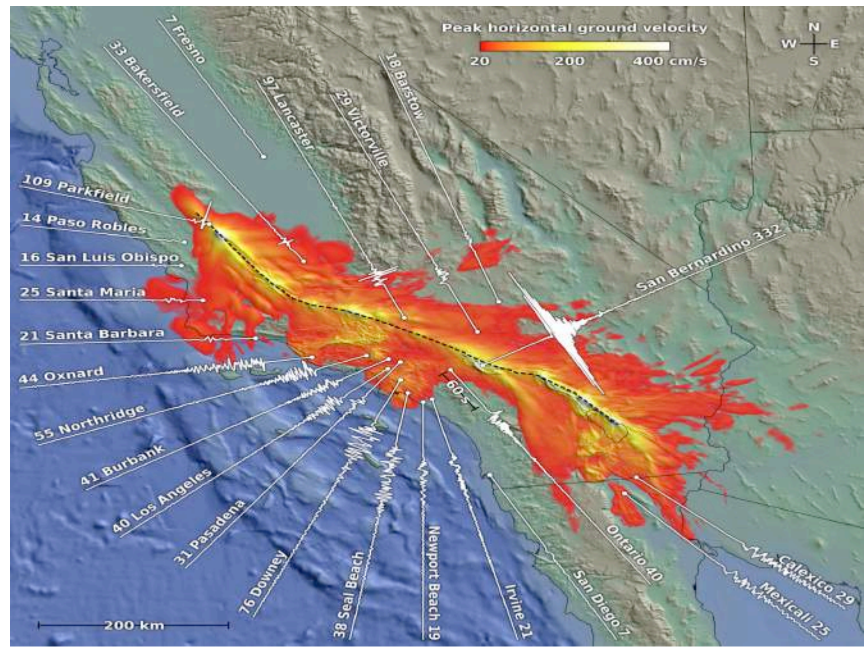
- **Dynamic rupture simulation (pathway 3) performed on *Kraken*, 7.5 hours using 2160 cores**
  - **881,475 subfaults, 250 sec of rupture**
  - **2.1 TB tensor time series output**

- **Wave propagation simulation (pathway 2) 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) representing seismic frequencies up to 2 Hz**



Input model: CVM-S4

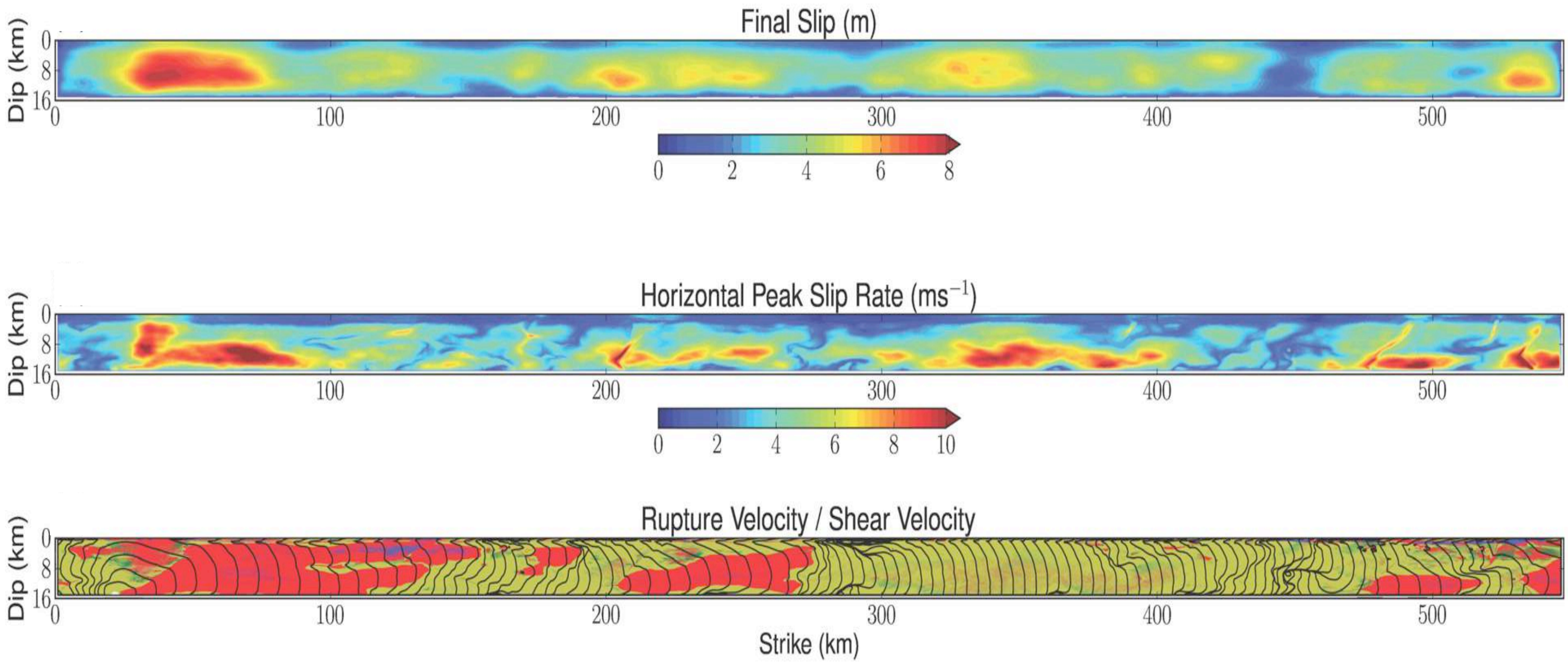
Output data: surface seismograms

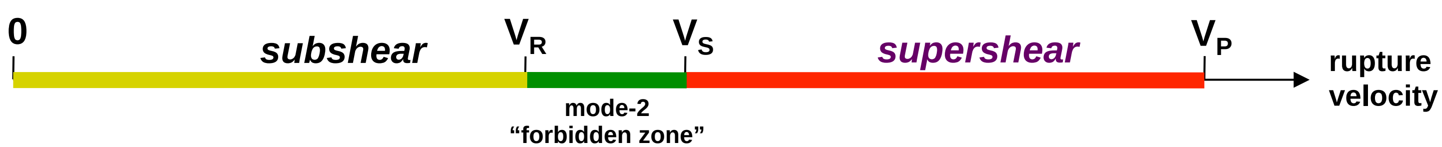


4D outer/inner scale ratio: 7×10^{16}

M8 Dynamic Rupture Simulation

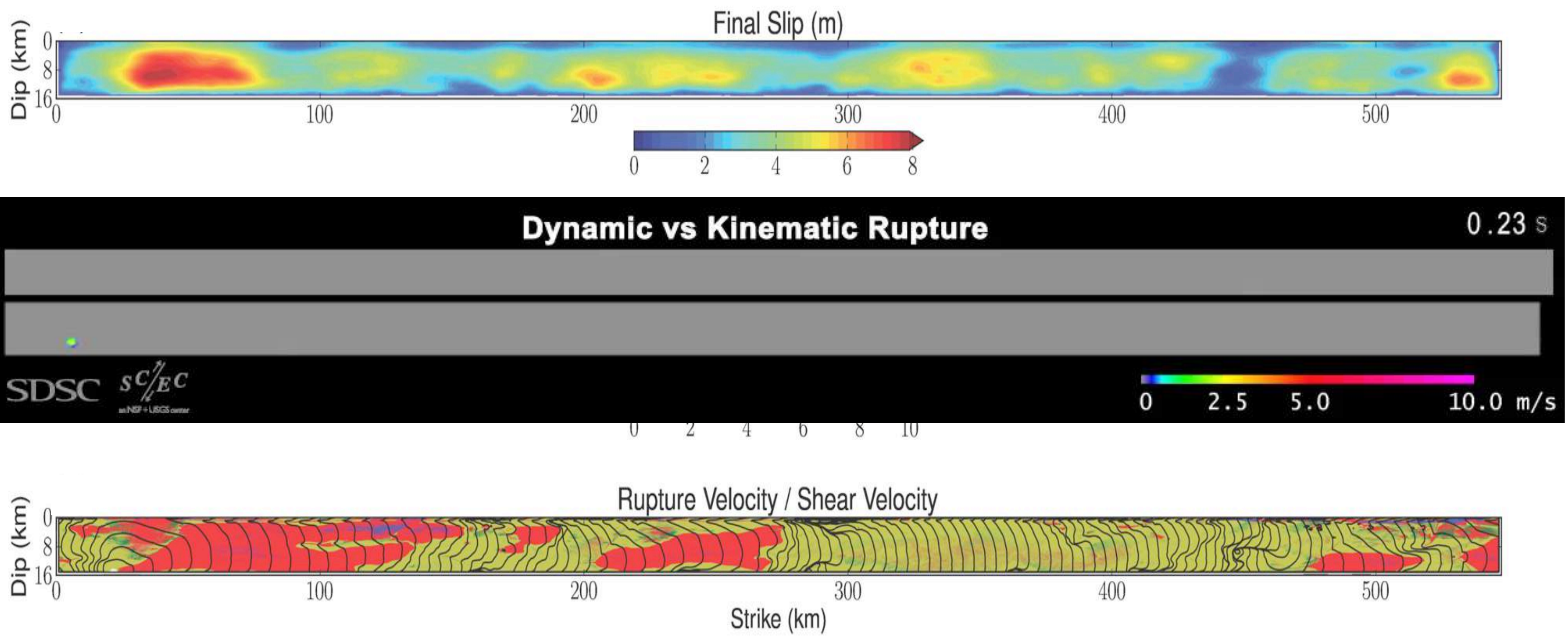
Northwest — **rupture direction** → **Southeast**

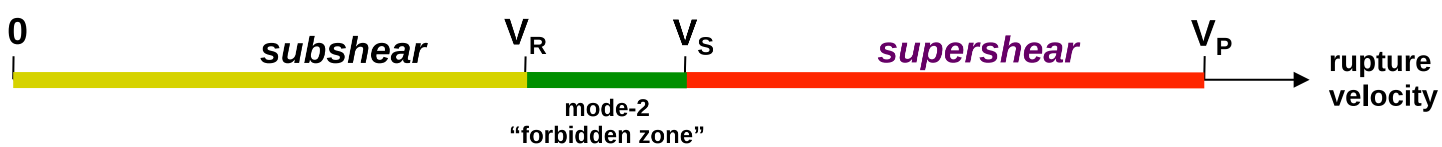




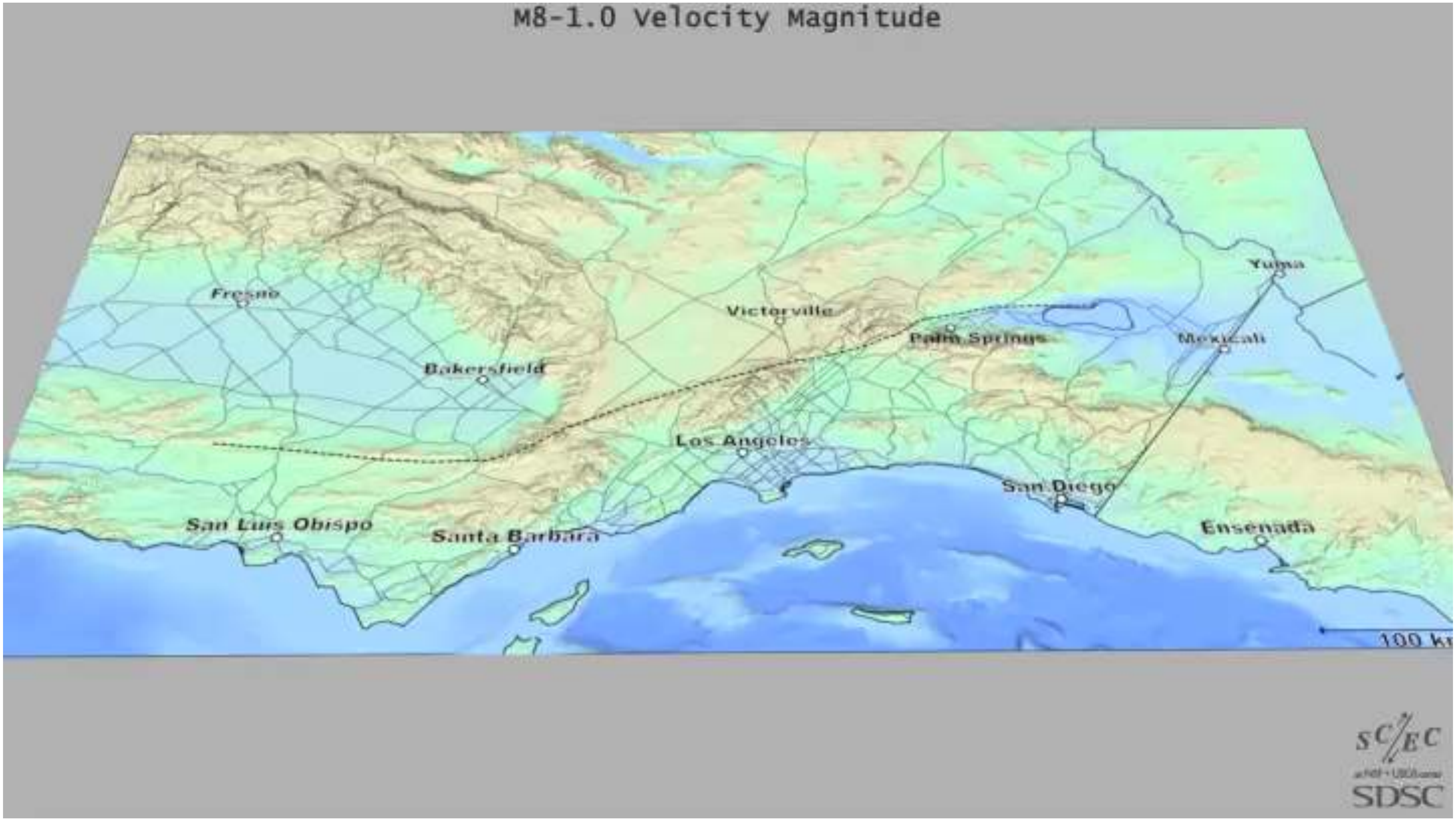
M8 Dynamic Rupture Simulation

Northwest — rupture direction → Southeast





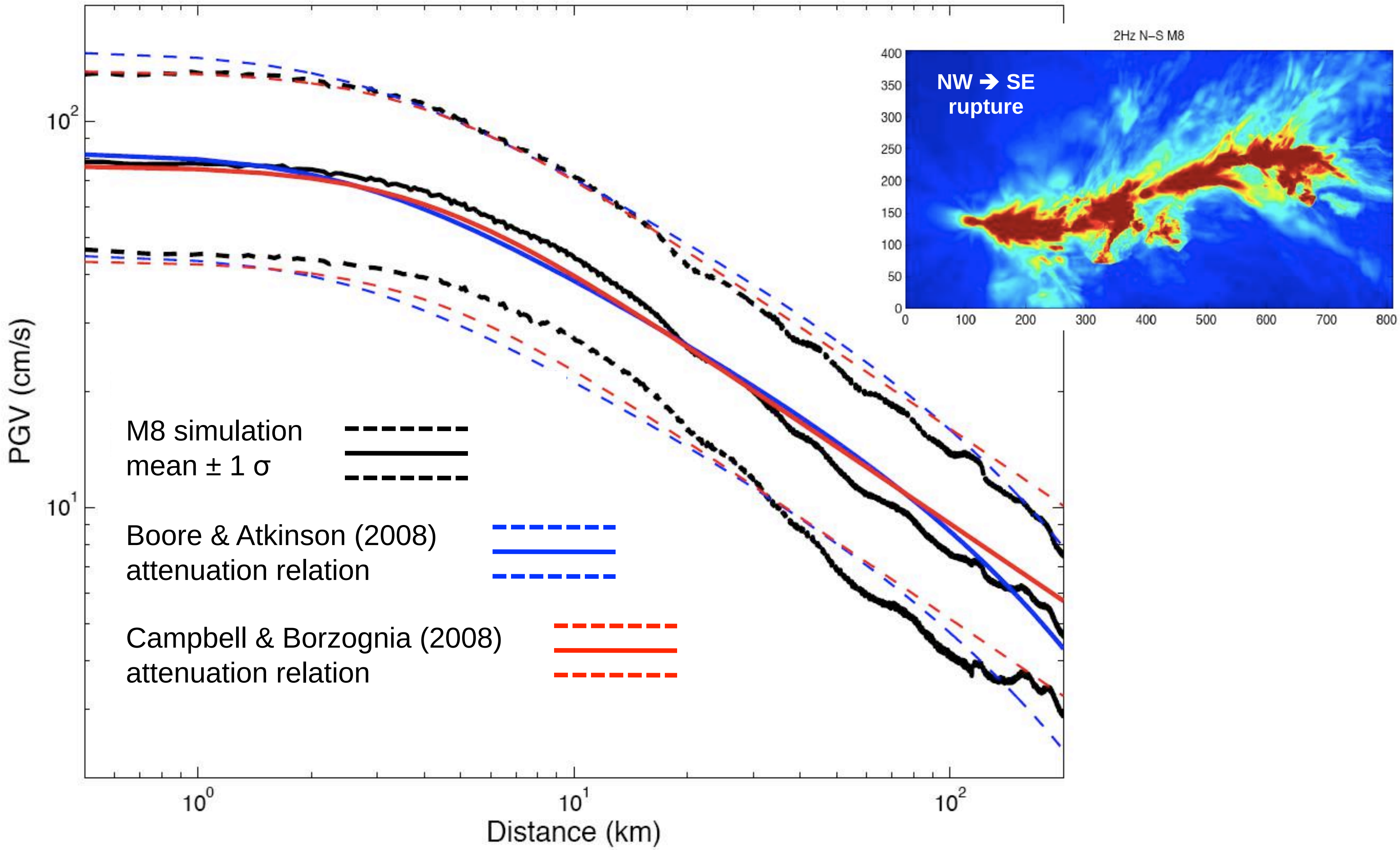
M8 Ground Motion Simulation



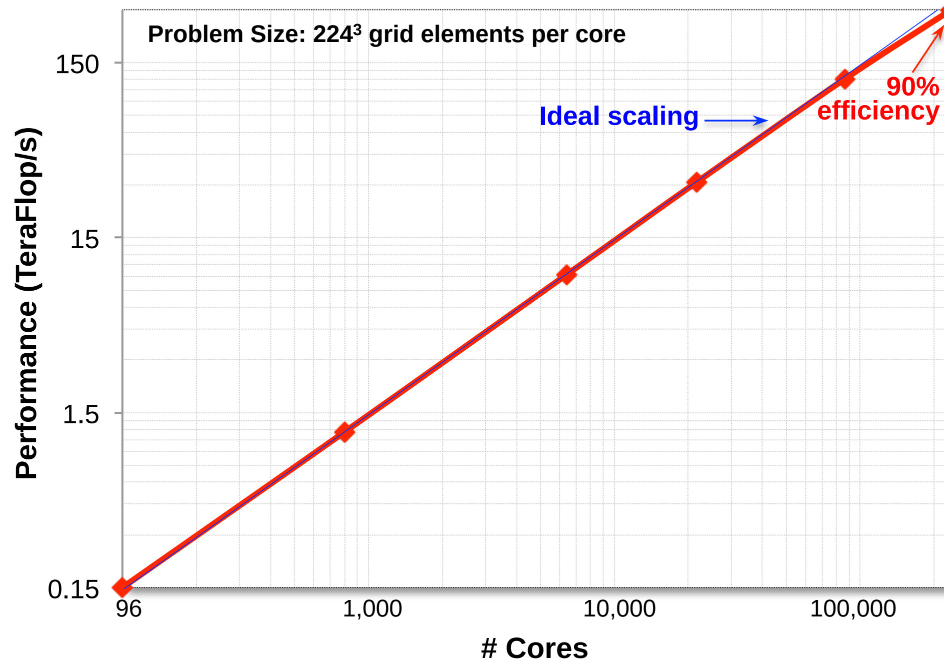
Simulation by Cui, Olsen, et al (2010), animation by A. Chourasia, SDSC

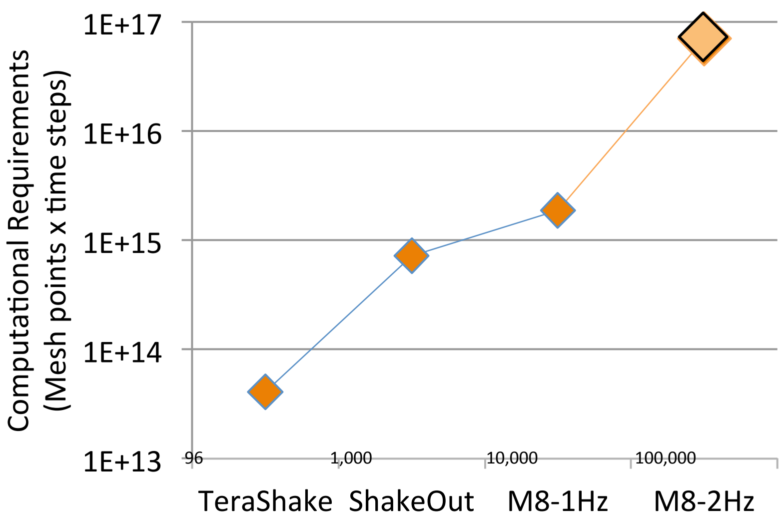
Comparison of M8 Ground Motions with NGA GMPEs

M8 Simulation of Cui et al. (2010)



SCEC Large-Scale Simulations



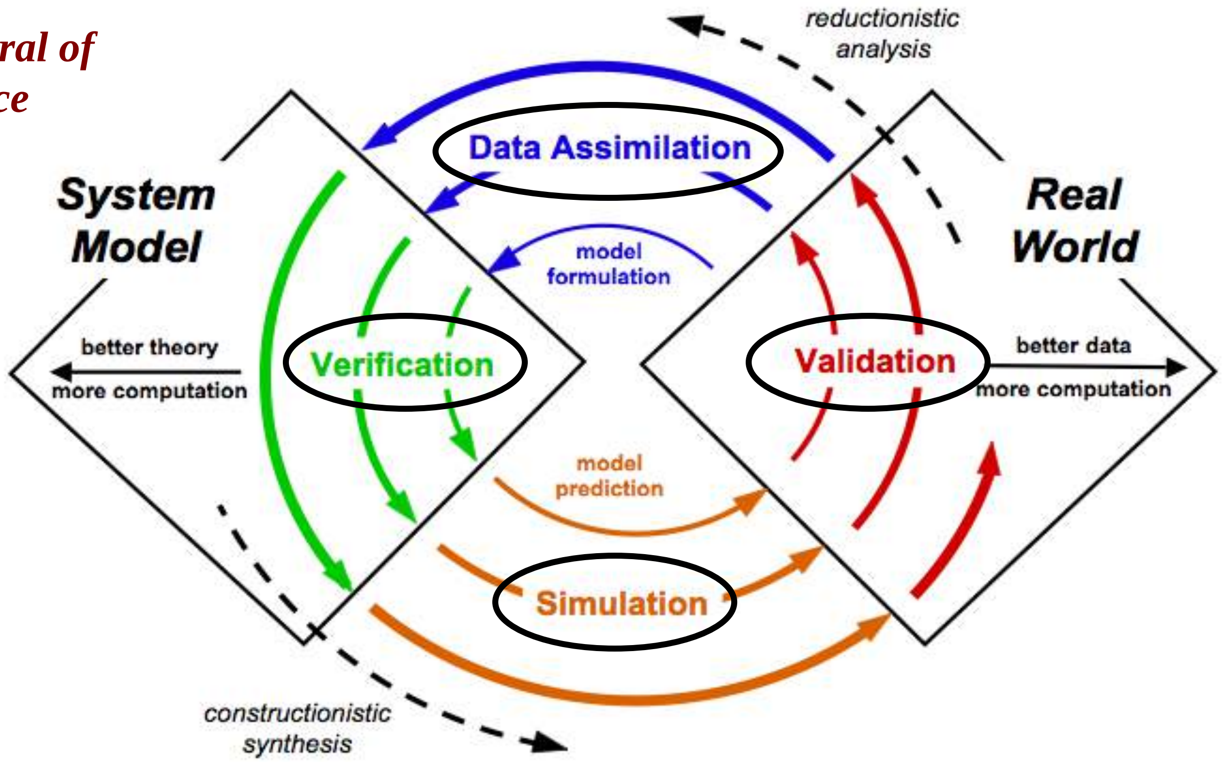


Increased computational performance has paved the way for the large suites of simulations needed for probabilistic seismic hazard analysis (PSHA)

Inference Spiral

- **Earthquake system science requires an iterative, computationally intense process of model formulation and verification, simulation-based predictions, validation against observations, and data assimilation to improve the model**

Inference Spiral of System Science



- **As models become more complex and new data bring in more information, we require ever increasing computational resources**

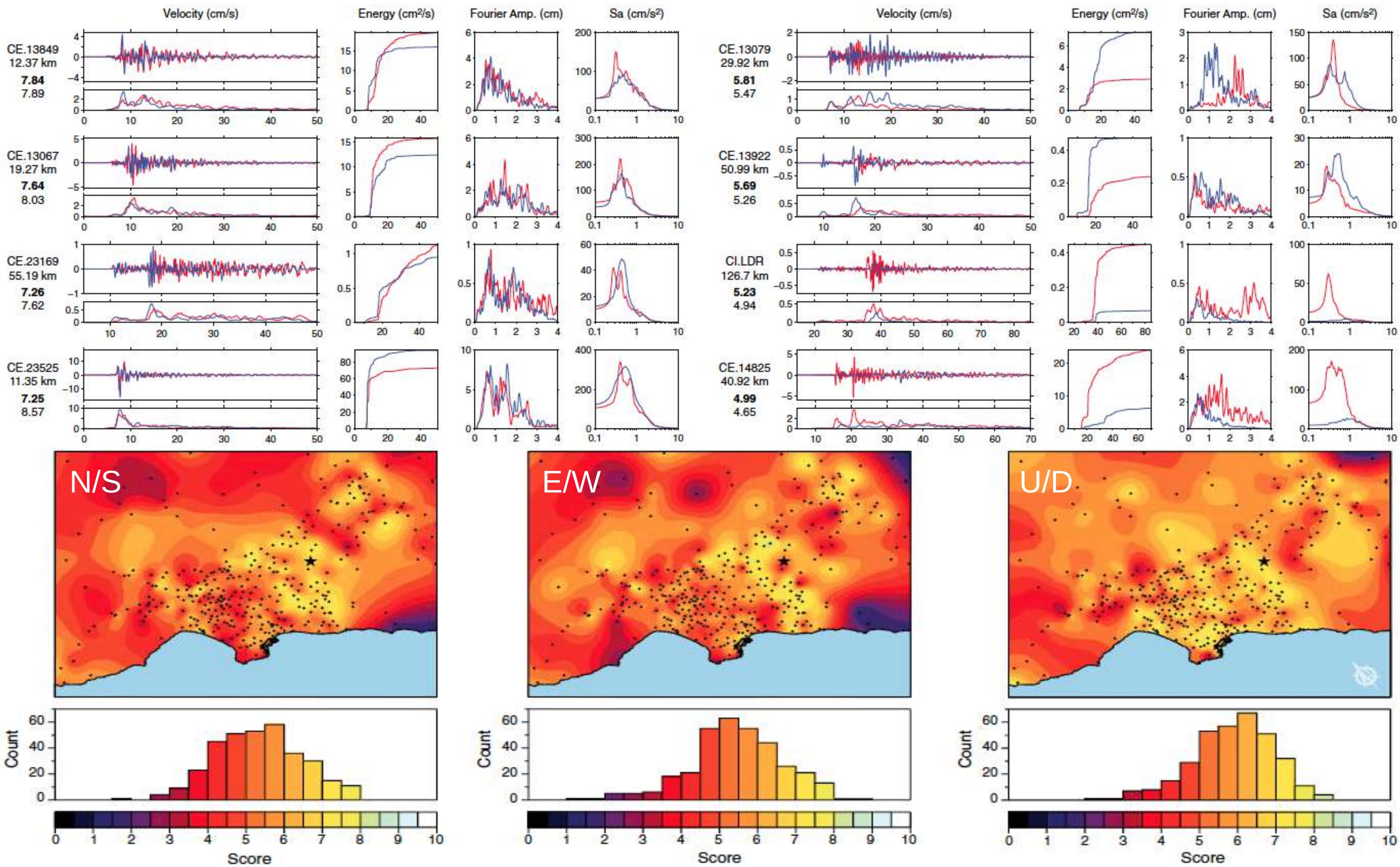
Cross-Verification of Simulations

SCEC ShakeOut Simulations

Bielak, J., R. W. Graves, K. B. Olsen, R. Taborda, L. Ramírez-Guzmán, S. M. Day, G. P. Ely, D. Roten, T. H. Jordan, P. J. Maechling, J. Urbanic, Y. Cui & G. Juve (2010)

Validation Using Small Earthquakes

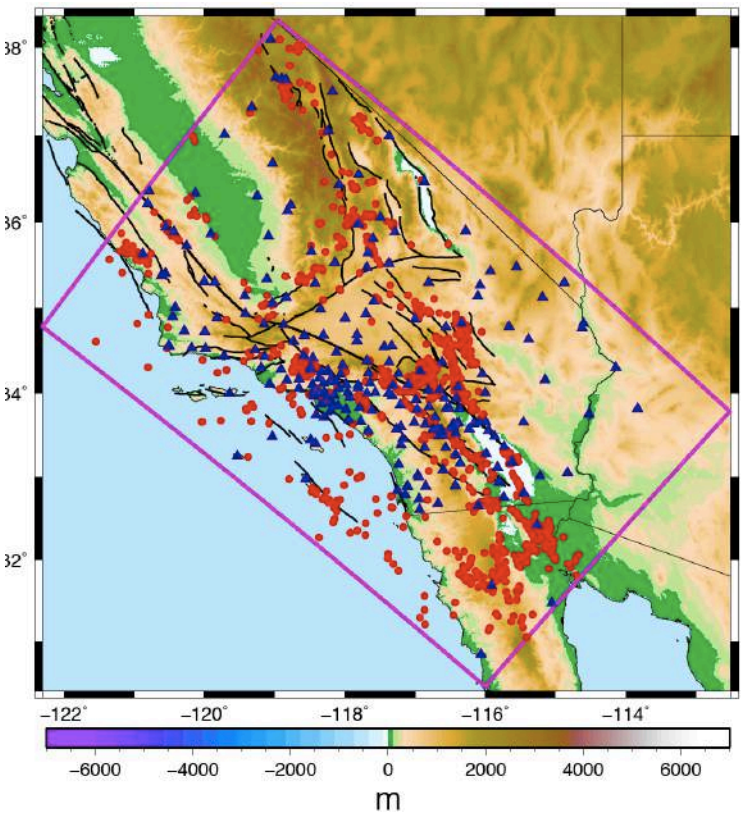
2008 Chino Hills, M5.4 (Taborda & Bielak, 2013)



SCEC

Southern California Earthquake Center

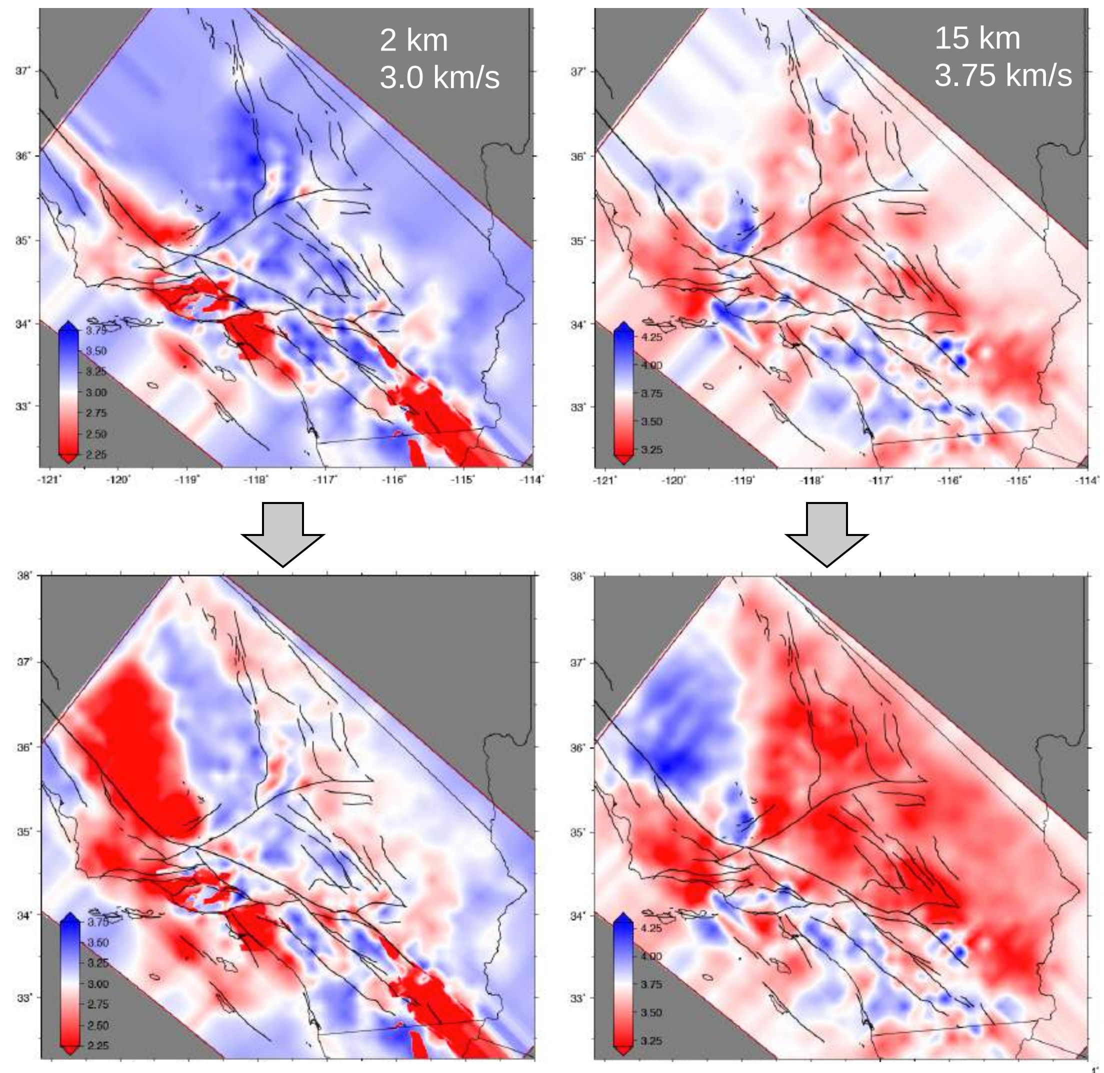
Data Assimilation Using Full-3D Waveform Tomography



E.-J. Lee, P. Chen, T. H. Jordan, P. Maechling, M. Denolle, G. Beroza (2012)

Inversion of Earthquake Waveforms and Ambient-Noise Green Functions

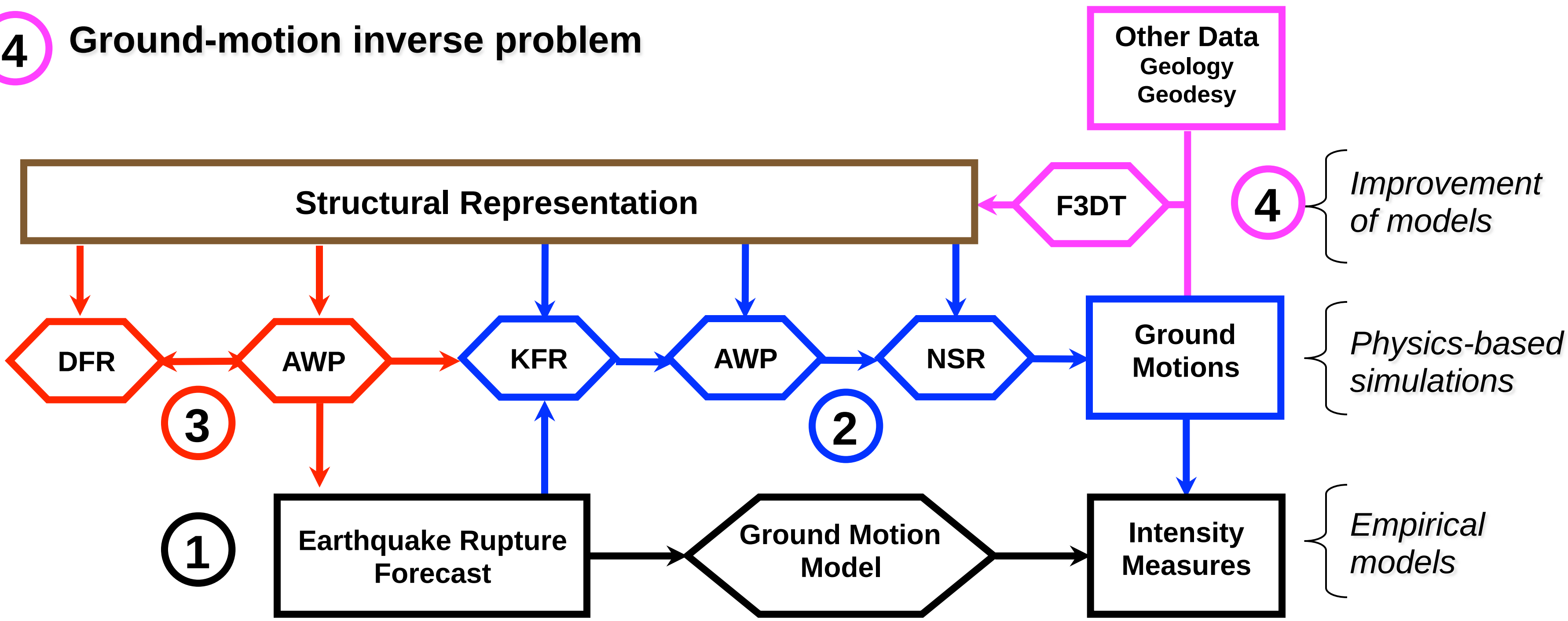
CVM-S4



CVM-S4.20

SCEC Computational Pathways

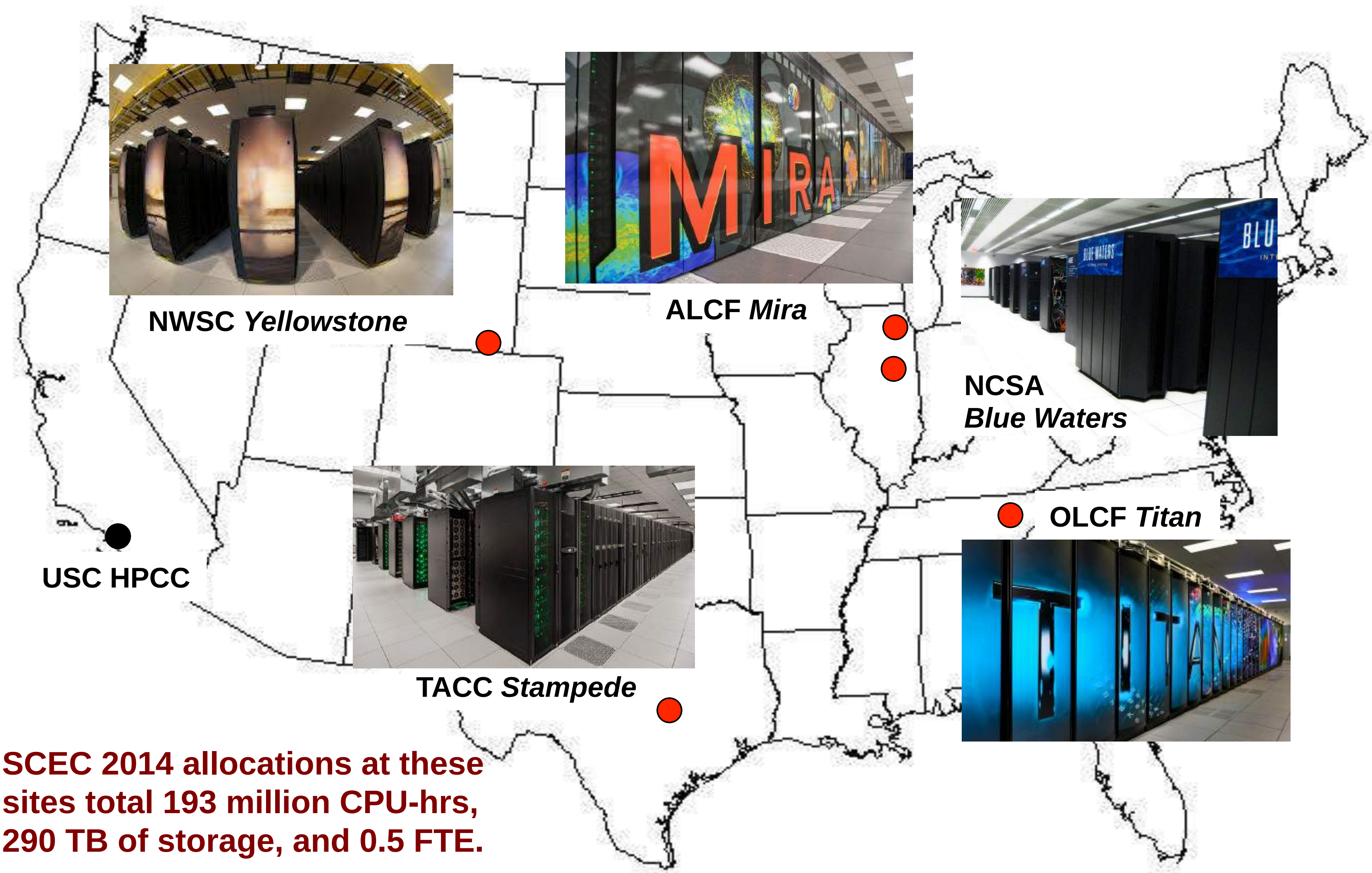
1. Earthquake rupture forecasting
2. Ground motion simulation
3. Dynamic rupture modeling
4. Ground-motion inverse problem



KFR = Kinematic Fault Rupture
DFR = Dynamic Fault Rupture

AWP = Anelastic Wave Propagation
NSR = Nonlinear Site Response

New HPC Resources Used by SCEC CME Collaboration

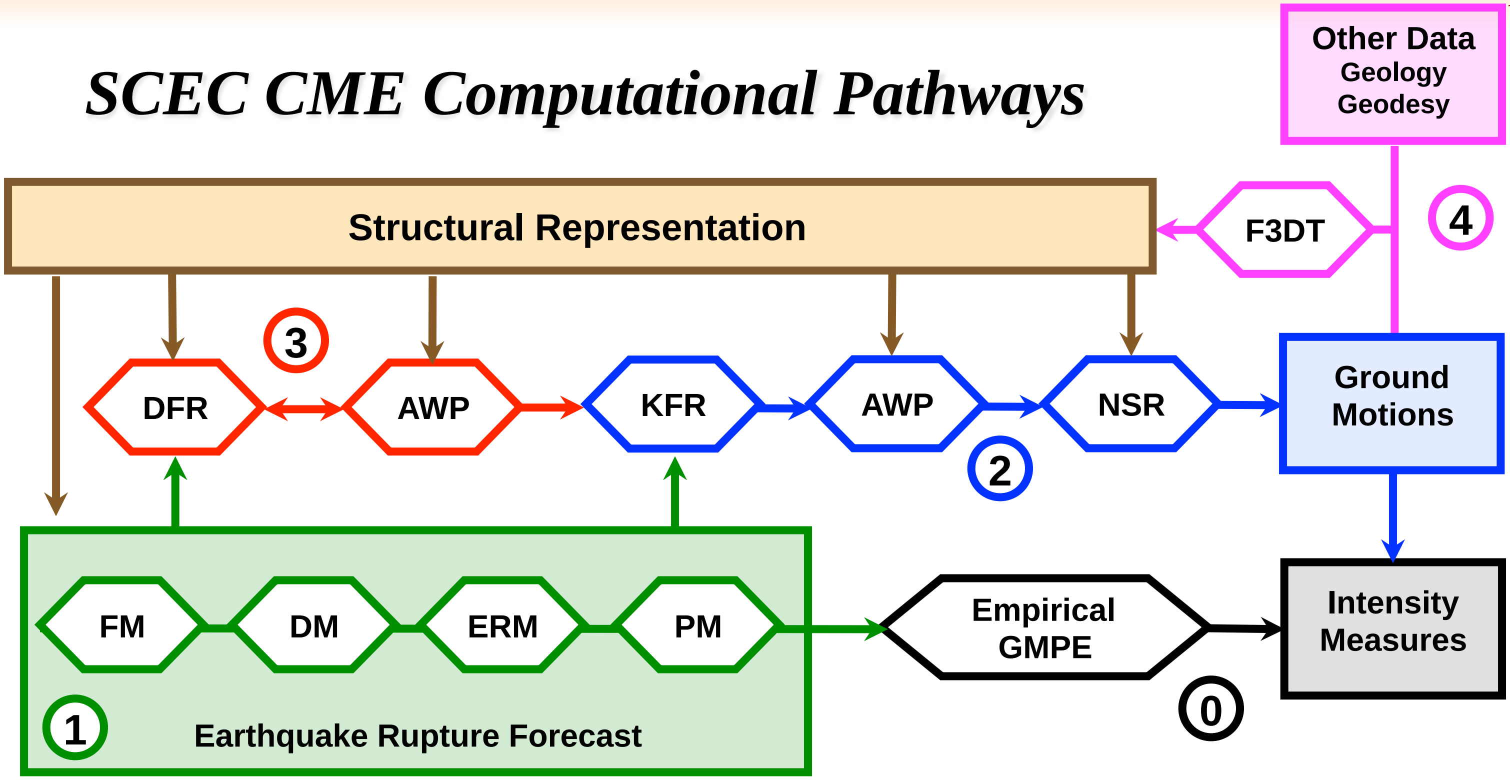


SCEC 2014 allocations at these sites total 193 million CPU-hrs, 290 TB of storage, and 0.5 FTE.

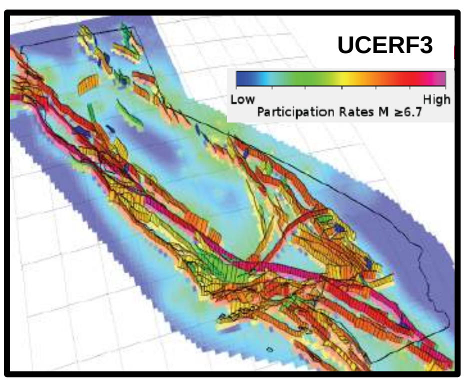
SCEC

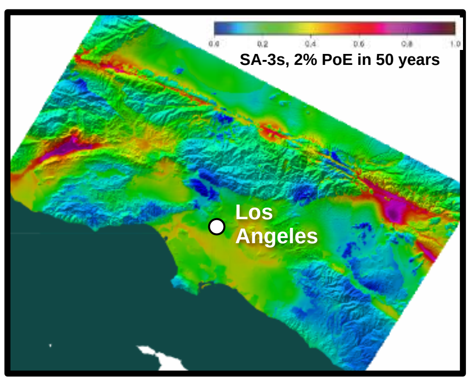
Southern California Earthquake Center

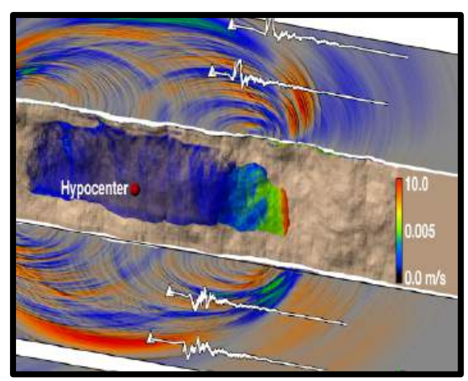
SCEC CME Computational Pathways

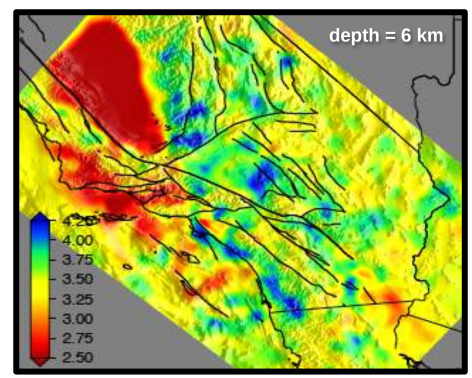


TACC *Stampede*	NCSA *Blue Waters*	ORNL *Titan*	ANL *Mira*









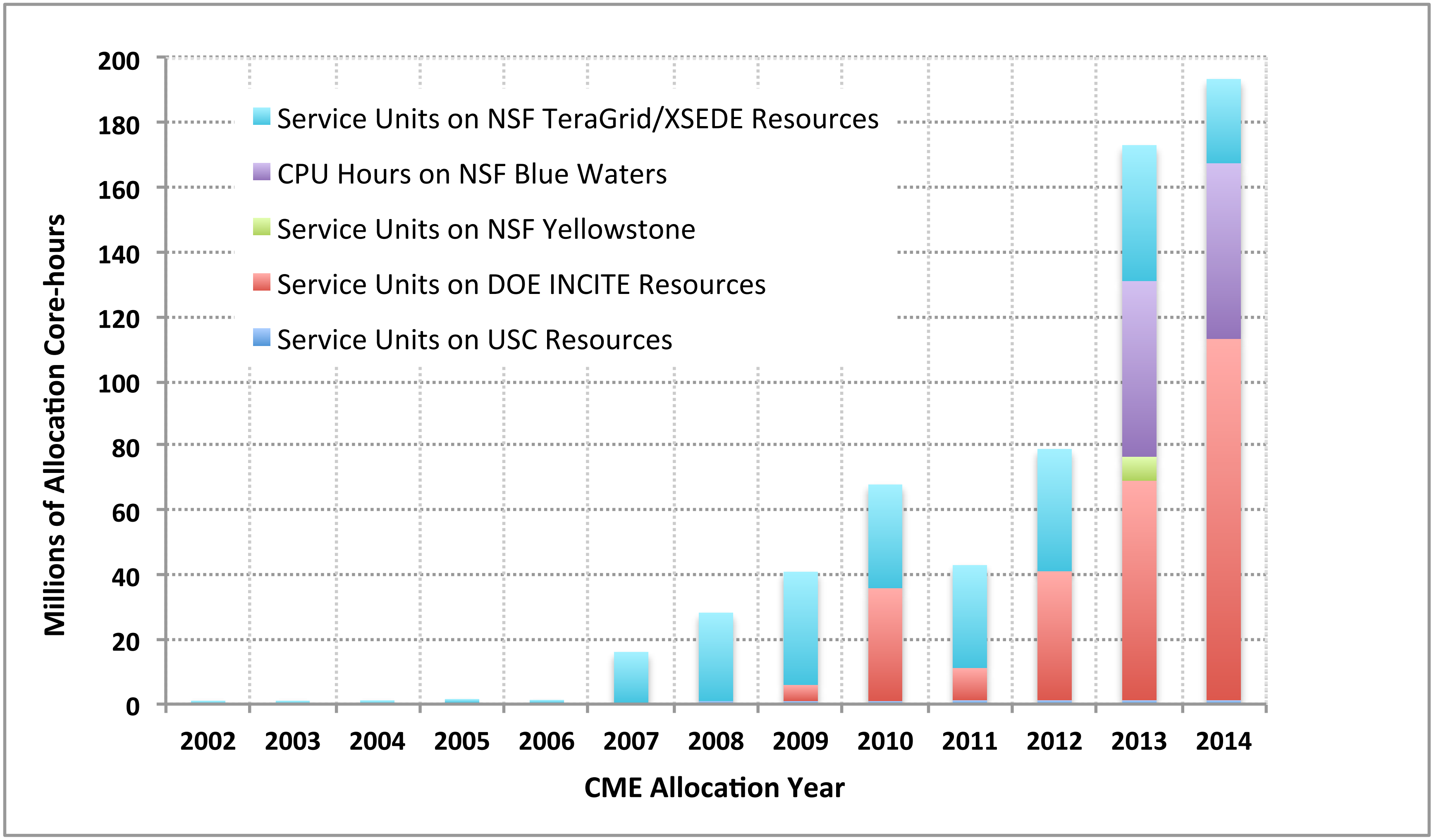
1 Uniform California Earthquake Rupture Forecast (UCERF3)

2 CyberShake 14.2 seismic hazard model for LA region

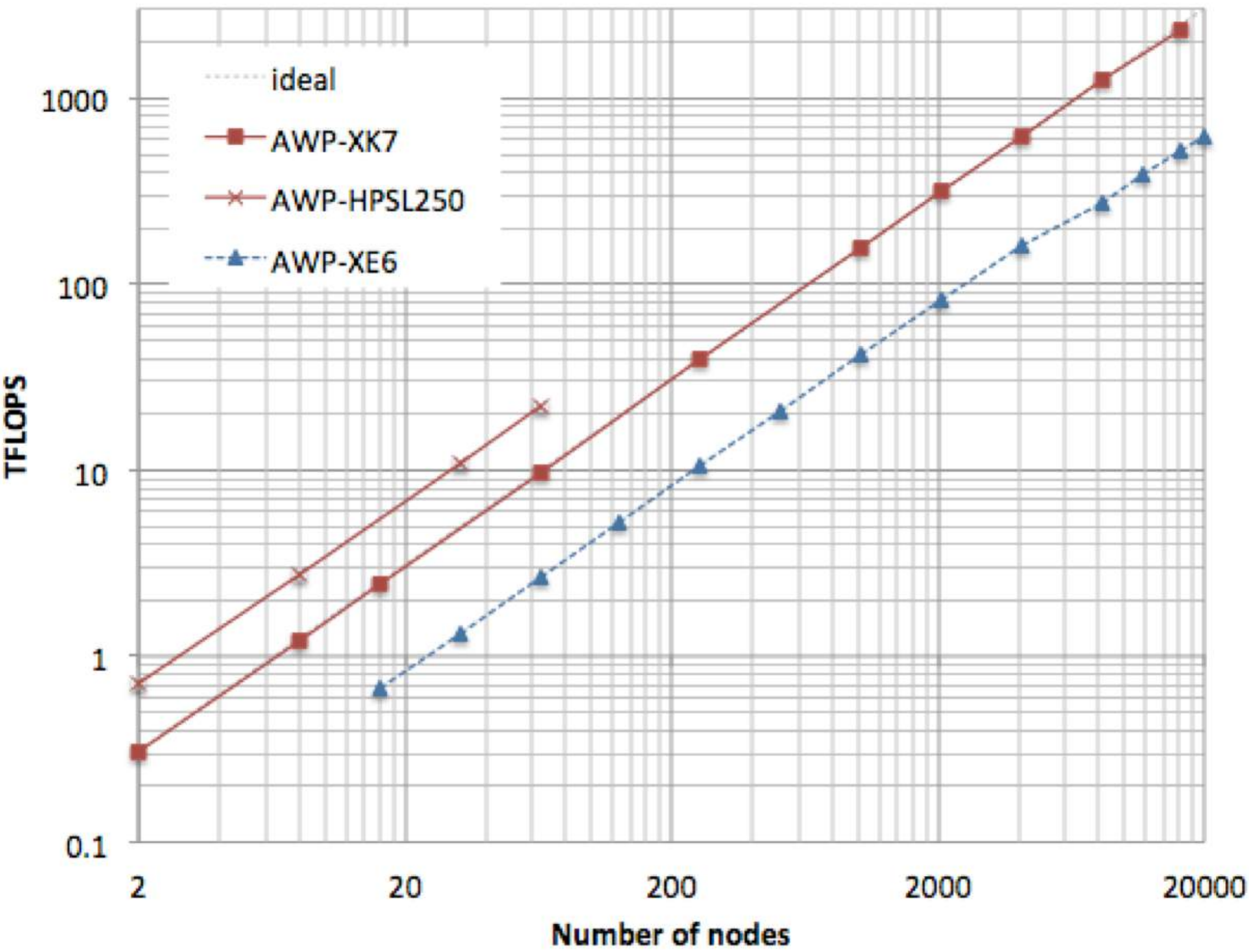
3 Dynamic rupture model of fractal roughness on SAF

4 Full-3D tomographic model CVM-S4.26 of S. California

SCEC use of HPC resources is growing rapidly…

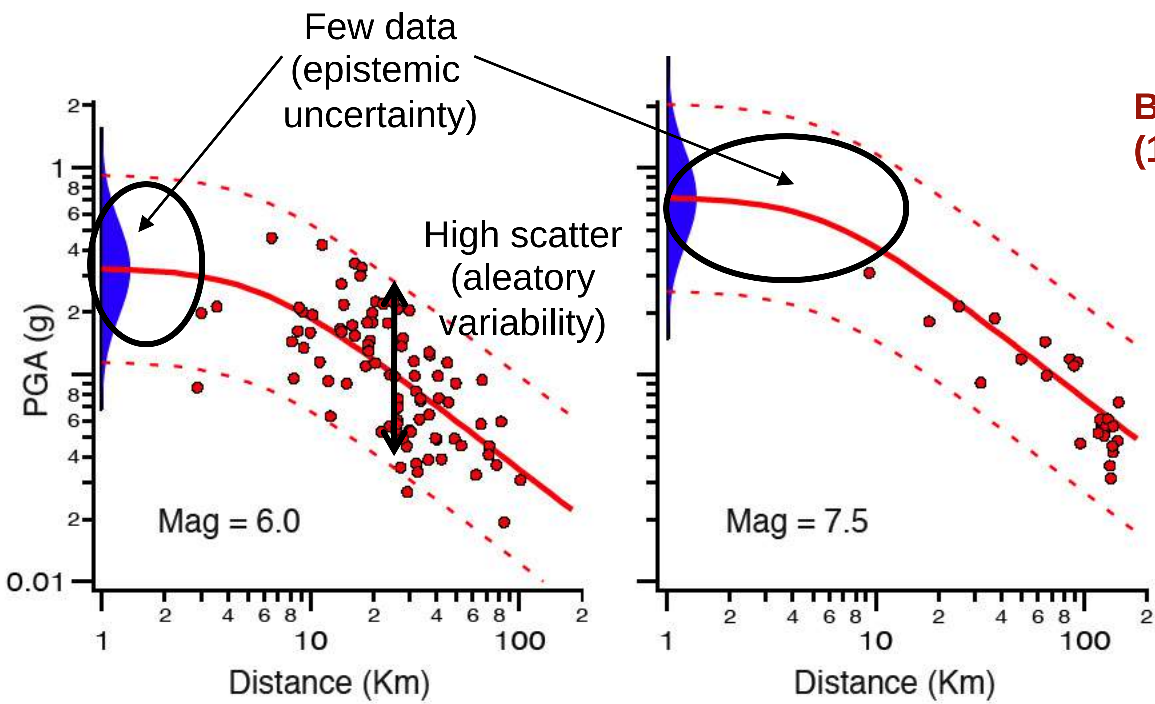


AWP-ODC Multi-GPU Performance



Formulation of Simulation-Based PSHA

Probabilistic Seismic Hazard Analysis

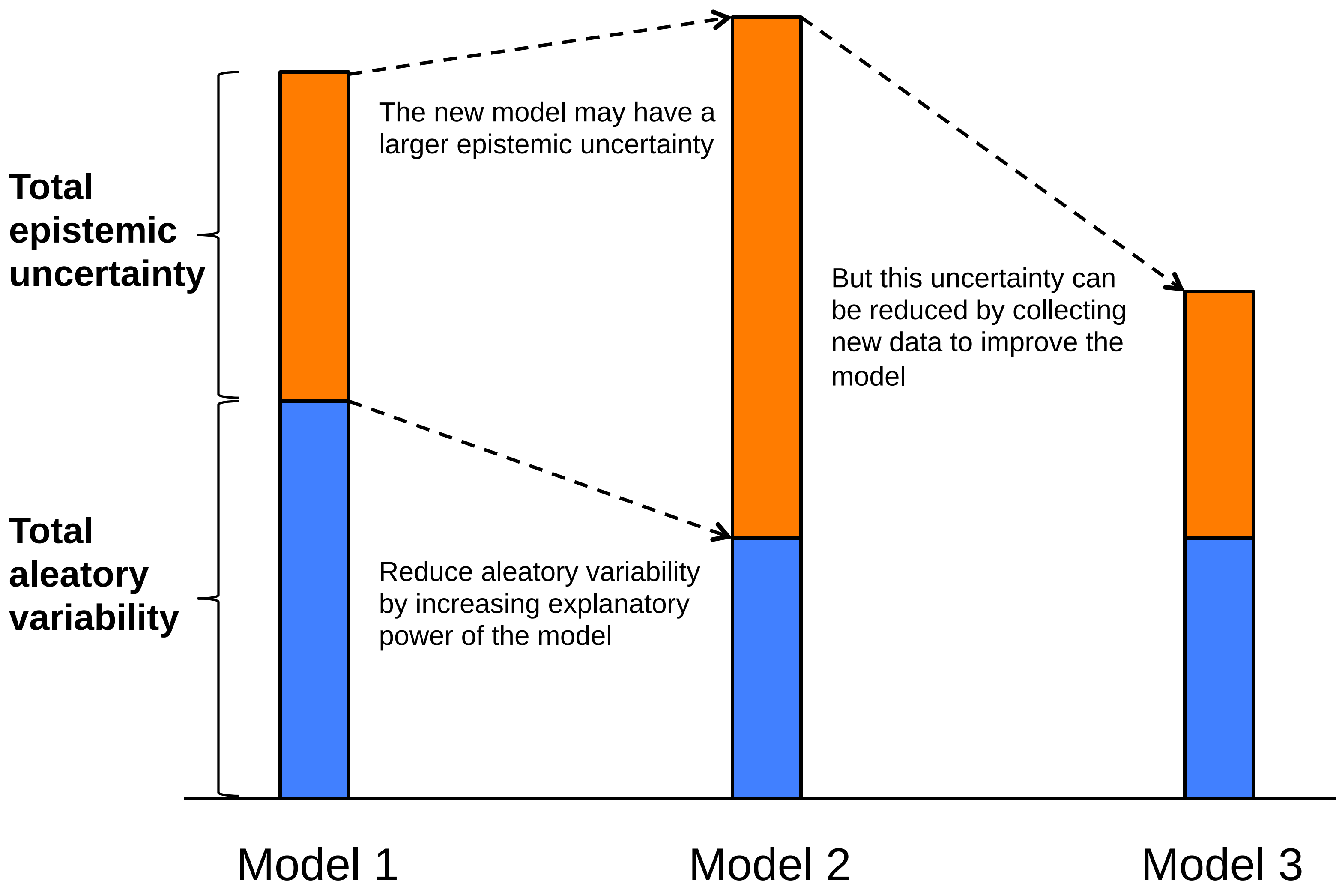


Boore, Joyner & Fumal (1997) GMPEs

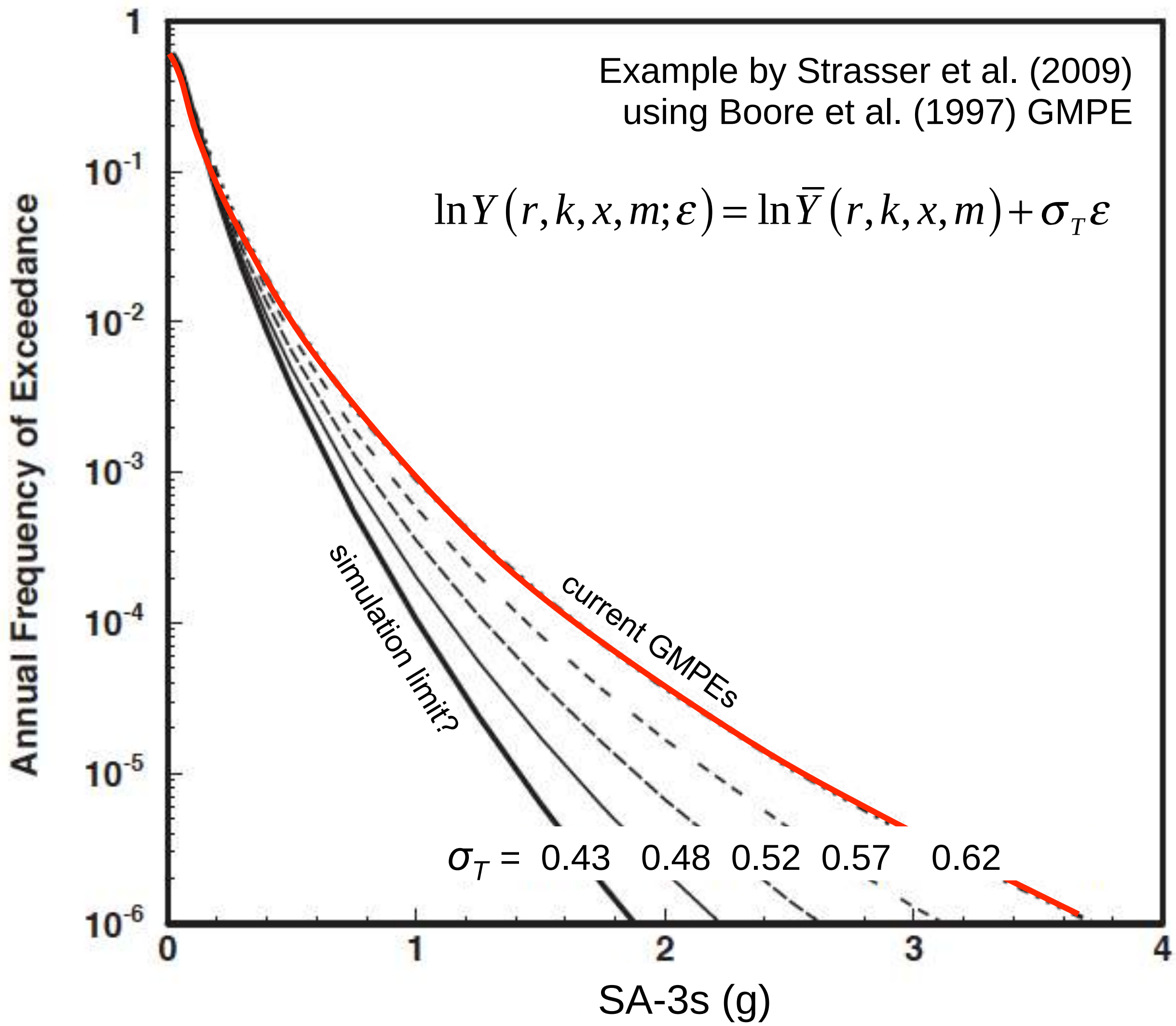
$P(S_n)$ $P(Y_k \mid S_n)$ $P(Y_k)$

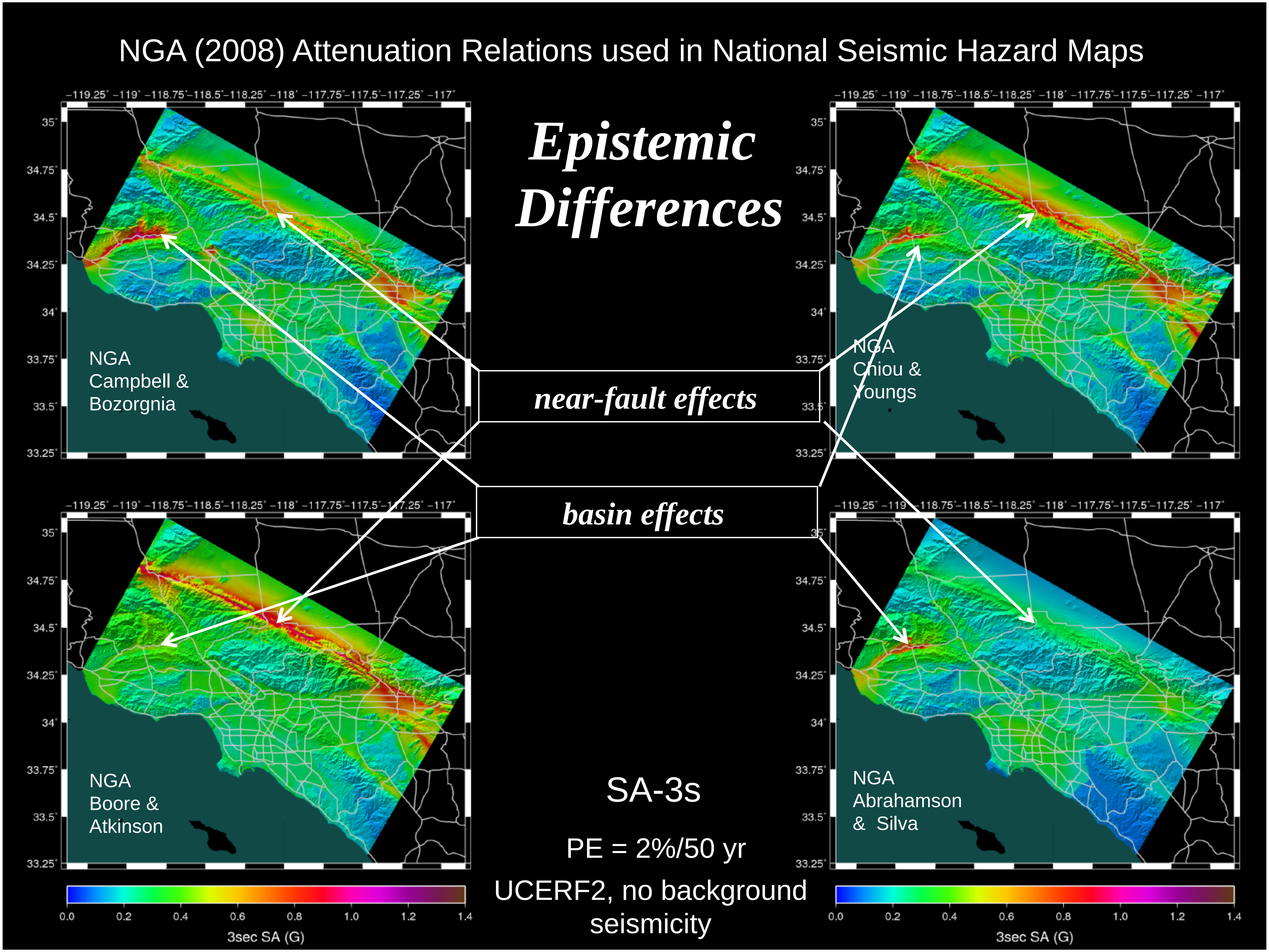
Probabilistic Seismic Hazard Analysis

Reduction of Aleatory Variability



Importance of Reducing Aleatory Variability




NGA (2008) Attenuation Relations used in National Seismic Hazard Maps
Epistemic Differences
near-fault effects
basin effects
NGA Campbell & Bozorgnia
NGA Chiou & Youngs
NGA Boore & Atkinson
NGA Abrahamson & Silva
SA-3s
PE = 2%/50 yr
UCERF2, no background seismicity
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4
3sec SA (G)
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4
3sec SA (G)

Seismological Hierarchy of CyberShake

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

- **Site set:** $r \in R$
  - 283 sites in the greater Los Angeles region
  - Elastic structures: BBP-1D, CVM-S4, CVM-H11, or CVM-S4.26
- **Rupture set:** $k \in K(r)$
  - All UCERF2 ruptures within 200 km of site (~7000 total)
- **Conditional hypocenter distribution:** $x \in X(r, k)$
  - Uniform distribution along fault strike with $\Delta x \approx 20$ km
- **Conditional slip distribution:** $s \in S(r, k, x)$
  - Pseudo-dynamic rupture models of Graves & Pitarka (2007, 2010)
  - Approximately 415,000 rupture variations per site, 235 million synthetic seismograms per model (2 horizontal components)

CHD and CSD define the “Extended ERF”

Ground Motion Prediction Equations

Formulation of time-independent PSHA for empirical GMPEs:

$$P\left(Y > y; r, T\right) = 1 - e^{-\lambda(Y > y|r)T}$$

conditional hypocenter distribution (CHD)

conditional magnitude distribution (CMD)

$$\lambda\left(Y > y \mid r\right) = \sum_{k \in K} \nu_k \sum_{x \in X(k)} p\left(x \mid k\right) \sum_{s \in S(k,x)} p\left(m \mid k\right) P\left(Y > y \mid r, k, x, m\right)$$

$$P(Y > y \mid r, k, x, m) \;=\; \int_{-\infty}^{+\infty} f(\varepsilon)\; H[\ln Y(r, k, x, m; \varepsilon) - \ln y]\, d\varepsilon$$

$$\ln \overline{Y}\left(r, k, x, m\right) = F_1\left(r\right) + F_2\left(r, k\right) + F_3\left(r, k, x\right) + F_4\left(k, m\right)$$

- F_1 site effect
- F_2 path effect
- F_3 directivity effect
- F_4 source-size effect

CyberShake Hazard Model

Formulation of time-independent PSHA for CyberShake simulations:

$$P\left(Y > y; r, T\right) = 1 - e^{-\lambda\left(Y > y | r\right) T}$$

conditional hypocenter distribution (CHD) ↓

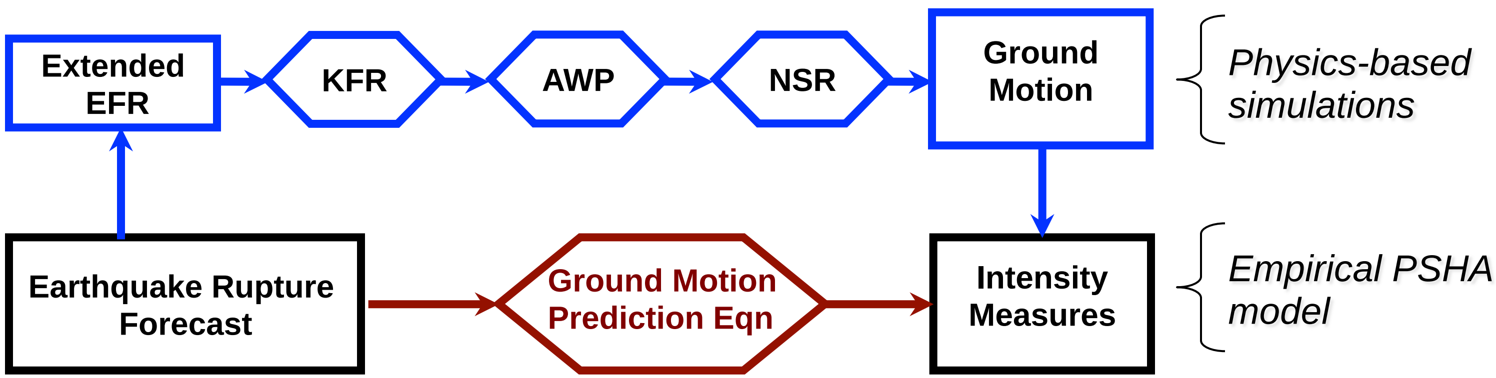
conditional slip distribution (CSD) ↓

$$\lambda\left(Y > y \mid r\right) = \sum_{k \in K} v_k P\left(Y > y \mid r, k\right) = \sum_{k \in K} v_k \sum_{x \in X(k)} p\left(x \mid k\right) \sum_{s \in S(k,x)} p\left(s \mid k, x\right) P\left(Y > y \mid r, k, x, s\right)$$

$$P\left(Y > y \mid r, k, x, s\right) = H\left[\ln Y\left(r, k, x, s\right) - \ln y\right]$$

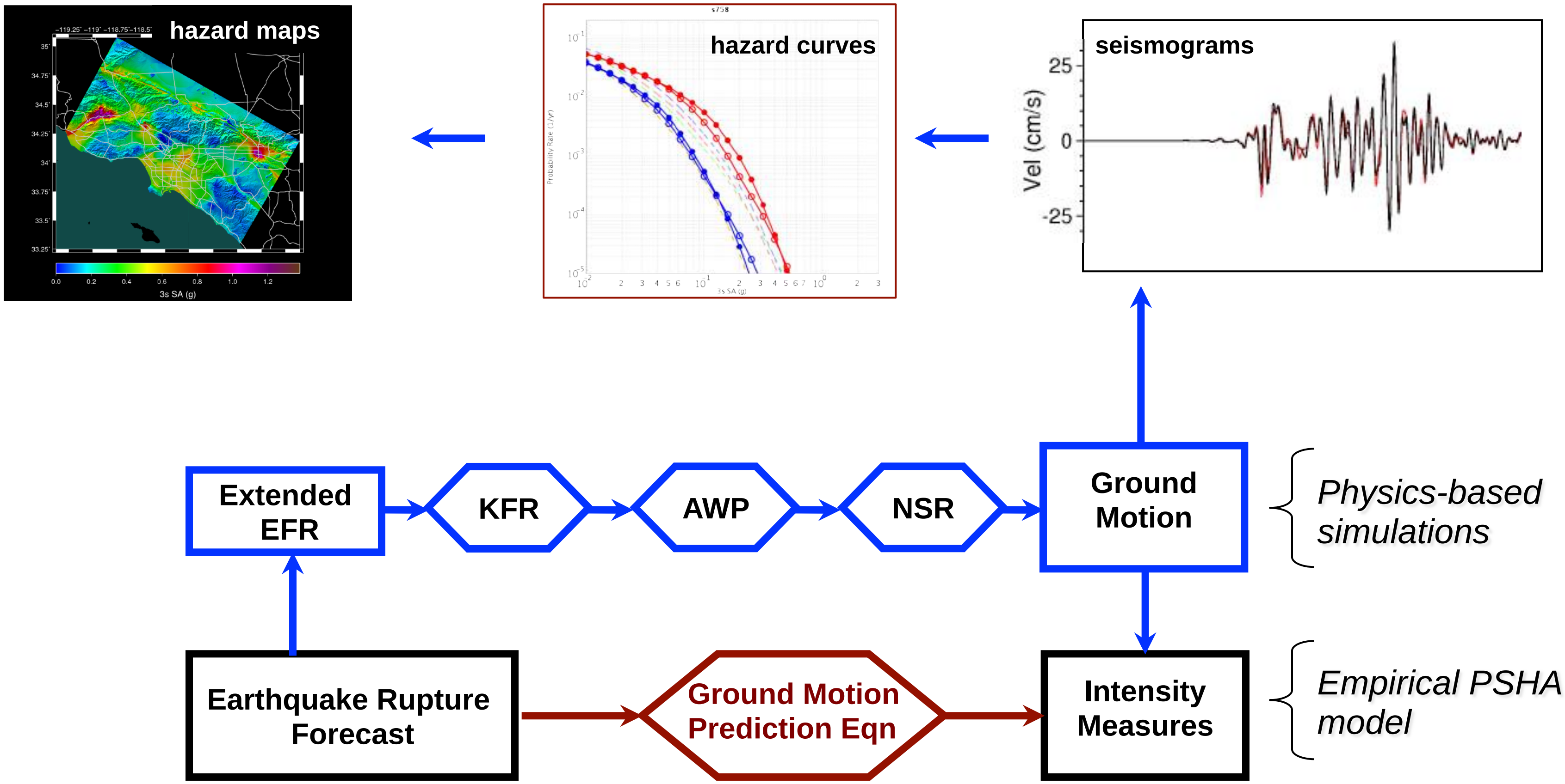
CSD magnitude functional:

$$m(s) \sim \log \left\| \int_{\Sigma_k} \int_{-\infty}^{+\infty} s\left(\xi, t; k, x\right) dt\, d\xi \right\|$$



KFR = kinematic fault rupture model
AWP = anelastic wave propagation model
NSR = nonlinear site response

CyberShake Hazard Model



KFR = kinematic fault rupture model
AWP = anelastic wave propagation model
NSR = nonlinear site response

CyberShake CS11 Hazard Model

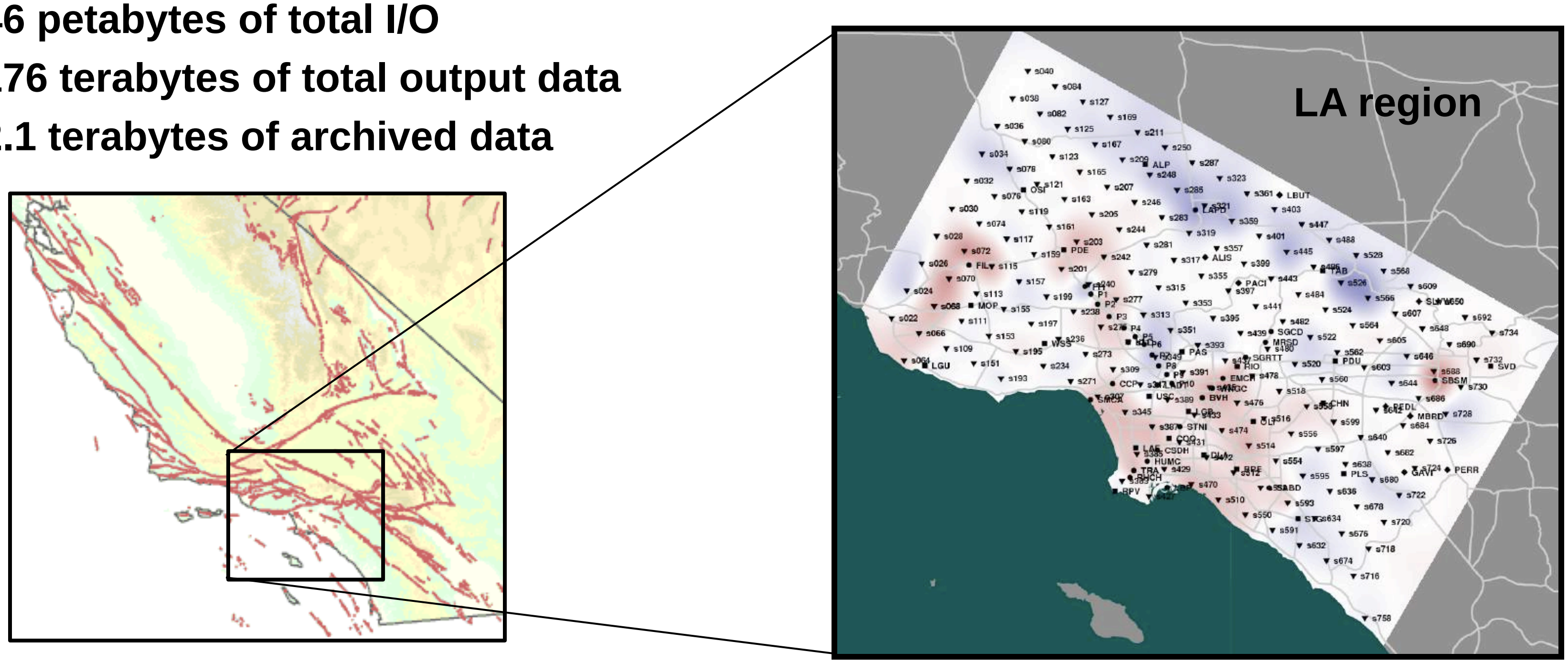
R. Graves, T. H. Jordan, S. Callaghan, E. Deelman, E. Field, G. Juve, C. Kesselman, P. Maechling, G. Mehta, K. Milner, D. Okaya, P. Small, and K. Vahi (2011)

225 sites in LA region (*f* < 0.5 Hz)

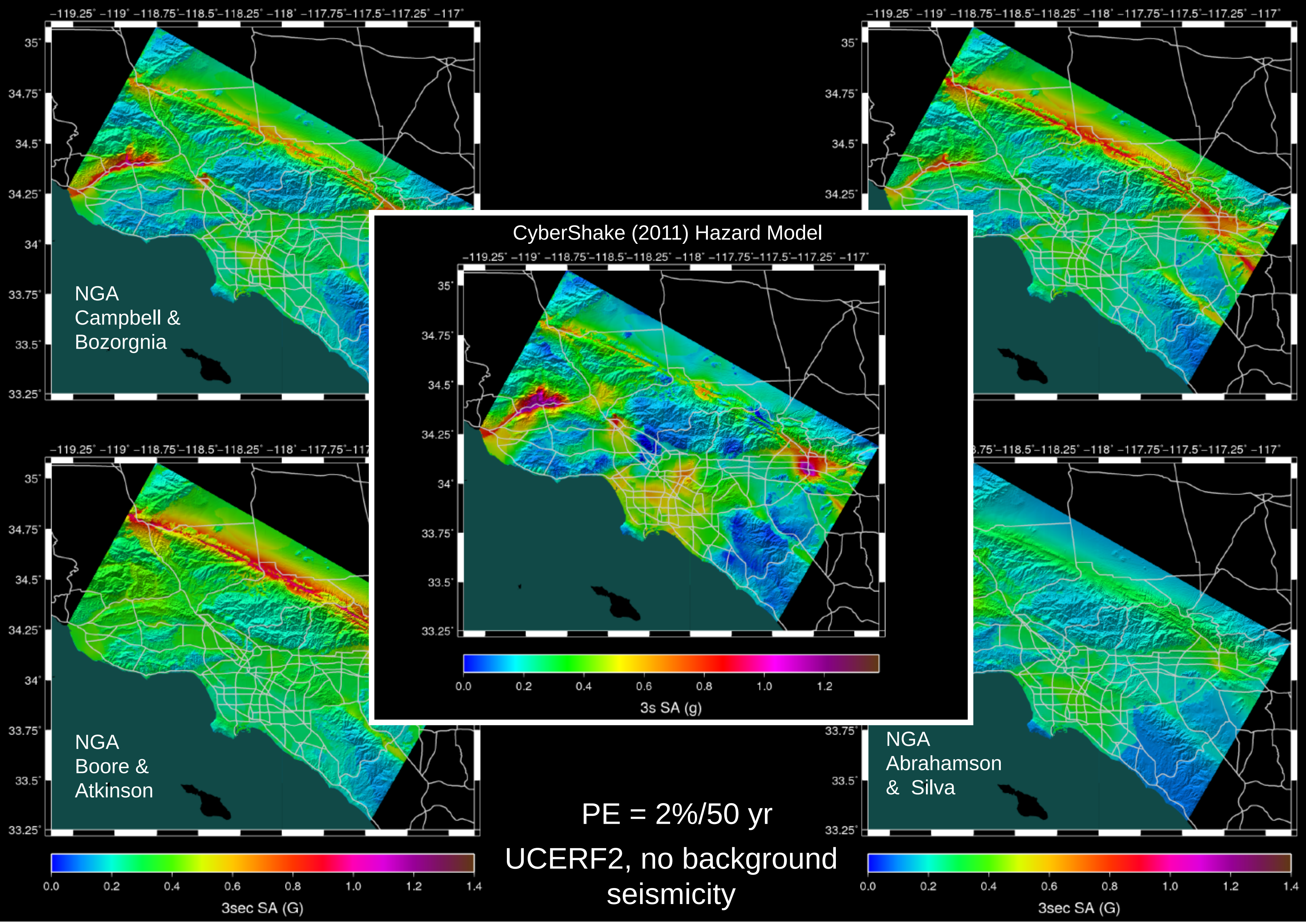
- **410,000 rupture variations per site**
- **total of 185 million seismograms (N & E components)**

Run on TACC *Ranger* (5.3 million hrs, 4,400 cores, 50 days)

- **189 million jobs**
- **46 petabytes of total I/O**
- **176 terabytes of total output data**
- **2.1 terabytes of archived data**



NGA (2008) Attenuation Relations used in National Seismic Hazard Maps



PE = 2%/50 yr

UCERF2, no background seismicity

Introduction to the CyberShake Computational Platform

CyberShake Platform: Physics-Based PSHA

Essential ingredients

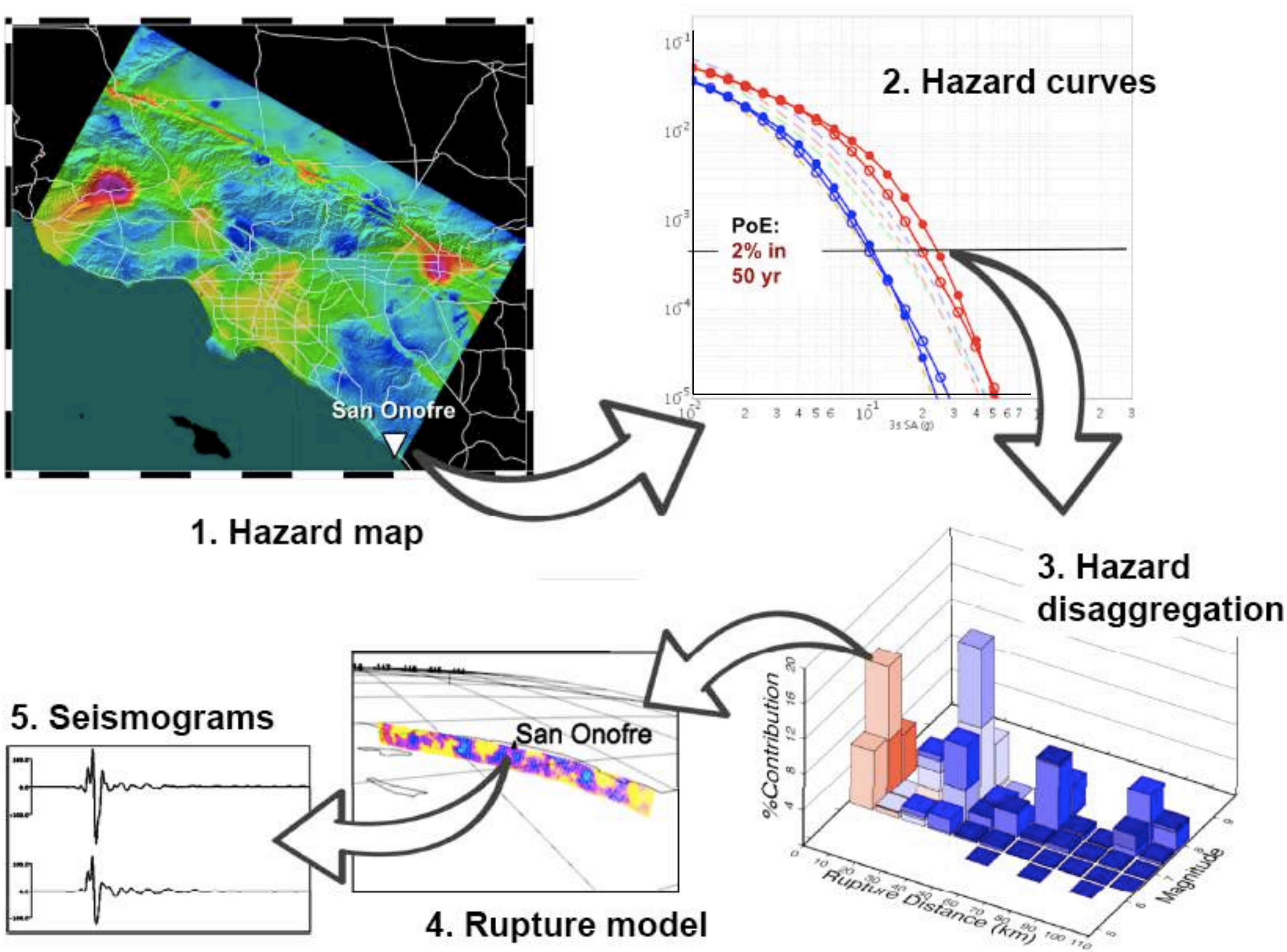
1. **Extended earthquake rupture forecast**
   - **probabilities of all fault ruptures (e.g., UCERF2)**
   - **conditional hypocenter distributions for rupture sets**
   - **conditional slip distributions from pseudo-dynamic models**

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

   } from SCEC CVMs

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

CyberShake Platform: Physics-Based PSHA



Computational Efficiency of Seismic Reciprocity

- **To account for source variability requires very large sets of simulations**
  - **7,000 ruptures from UCERF2; 415,000 rupture variations**
- **Ground motions can be calculated at much smaller number of surface sites to produce hazard map**
  - **283 in LA region, interpolated using empirical attenuation relations**

- **Elastodynamic representation theorem**

$$u_n(x,t) = \int_{-\infty}^{\infty} d\tau \int_{\Sigma} d\sigma(\xi) \frac{\partial}{\partial \xi_j} G_{ni}(x,t;\xi,\tau)\, \Gamma_{ij}(\xi, t-\tau)$$

- **Reciprocity**

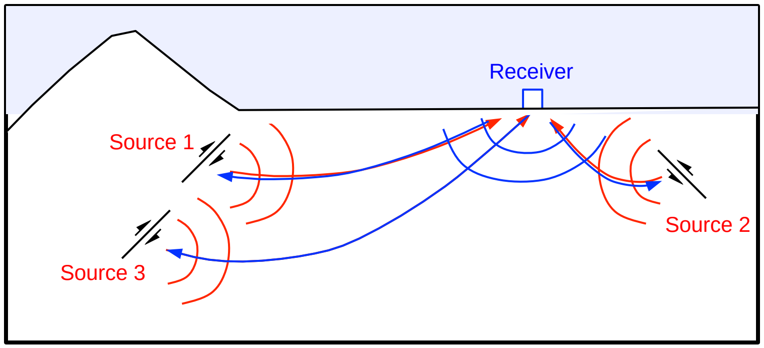
$$G_{ni}(x,t;\xi,\tau) = G_{ni}(\xi,-\tau;x,-t)$$

- **Strain Green tensor (SGT)**

$$H_{ijn}(x,t;\xi,\tau) = \frac{1}{2}\left[\frac{\partial}{\partial x_i} G_{jn}(x,t;\xi,\tau) + \frac{\partial}{\partial x_j} G_{in}(x,t;\xi,\tau)\right]$$

- **Site-oriented simulation**

$$u_n(x,t) = \int_{-\infty}^{\infty} d\tau \int_{\Sigma} d\sigma(\xi) H_{ijn}(\xi,\tau;x,t)\, \Gamma_{ij}(\xi, t-\tau)$$



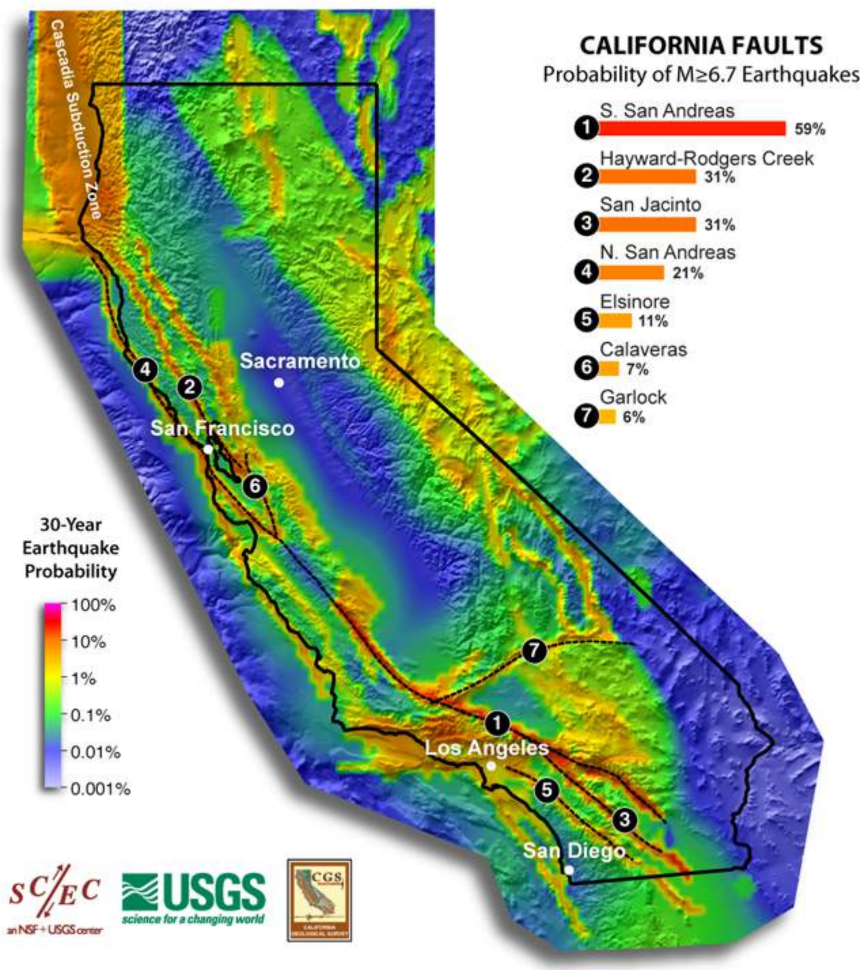
M **sources to** ***N*** **receivers requires** ***M*** **simulations**

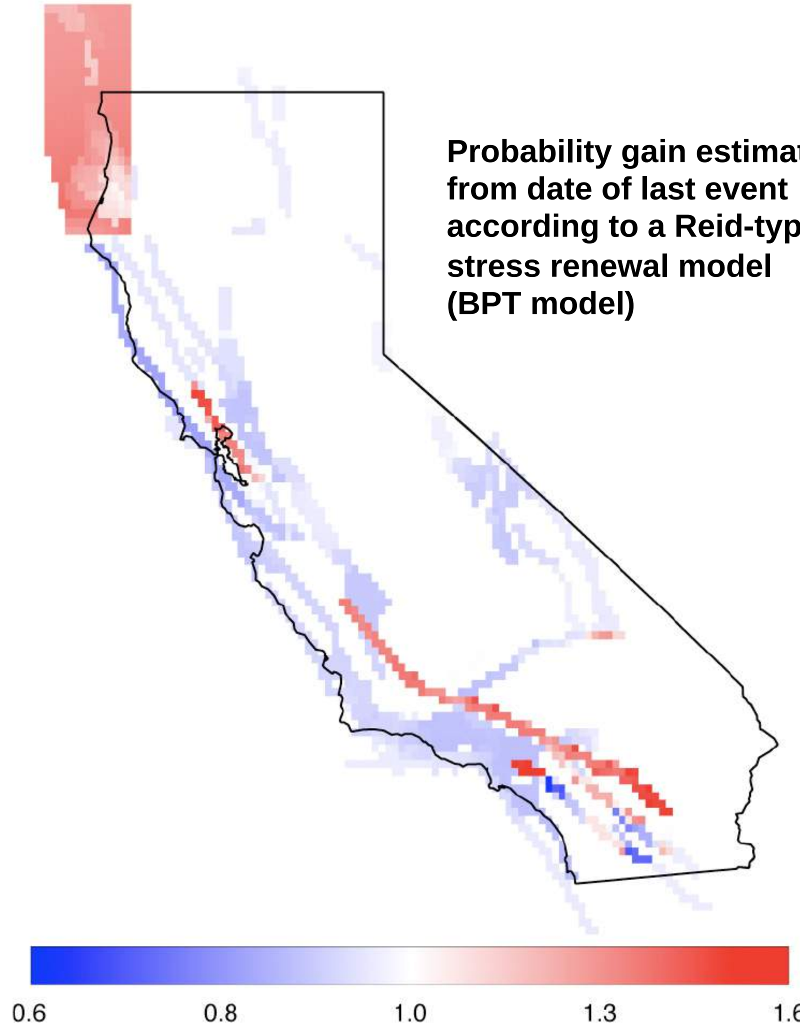
M **sources to** ***N*** **receivers requires 2*N* or 3*N* simulations**

- ***Use of reciprocity reduces CPU time by a factor of ~1,000***

Working Group on California Earthquake Probabilities (2007)

Uniform California Earthquake Rupture Forecast (UCERF2)





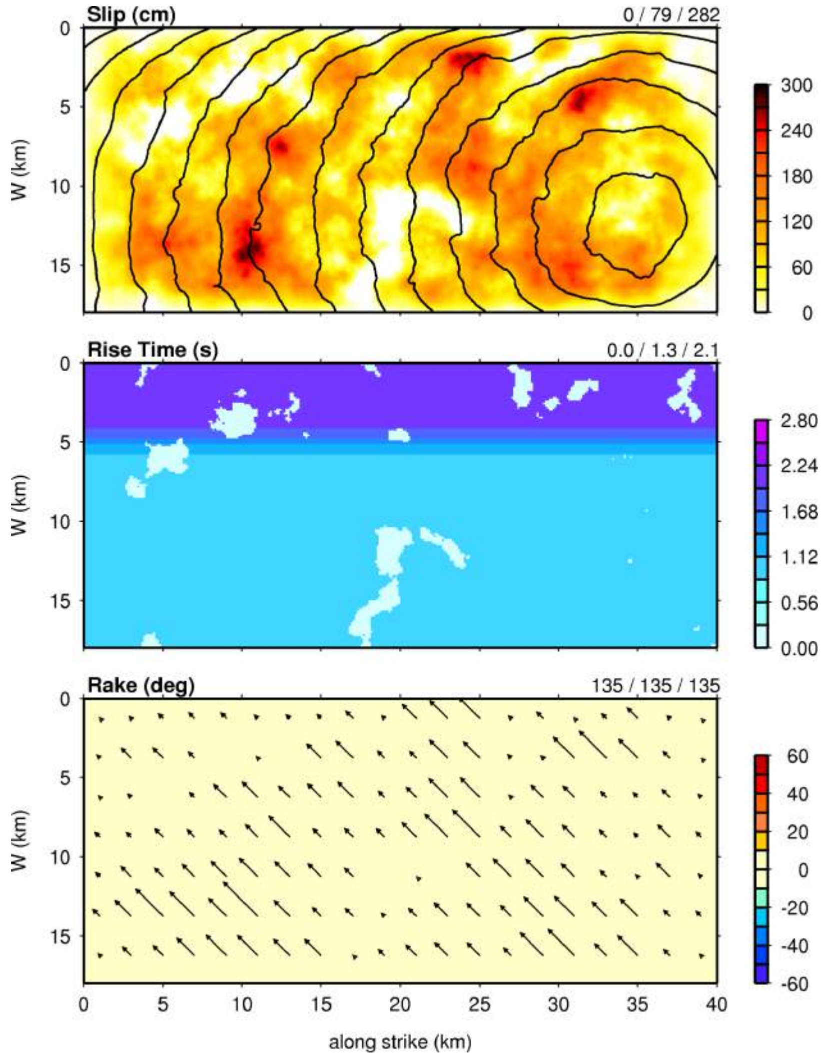
Probability gain estimated from date of last event according to a Reid-type stress renewal model (BPT model)

Ratio of time-dependent to time-independent participation probabilities for M ≥ 6.7

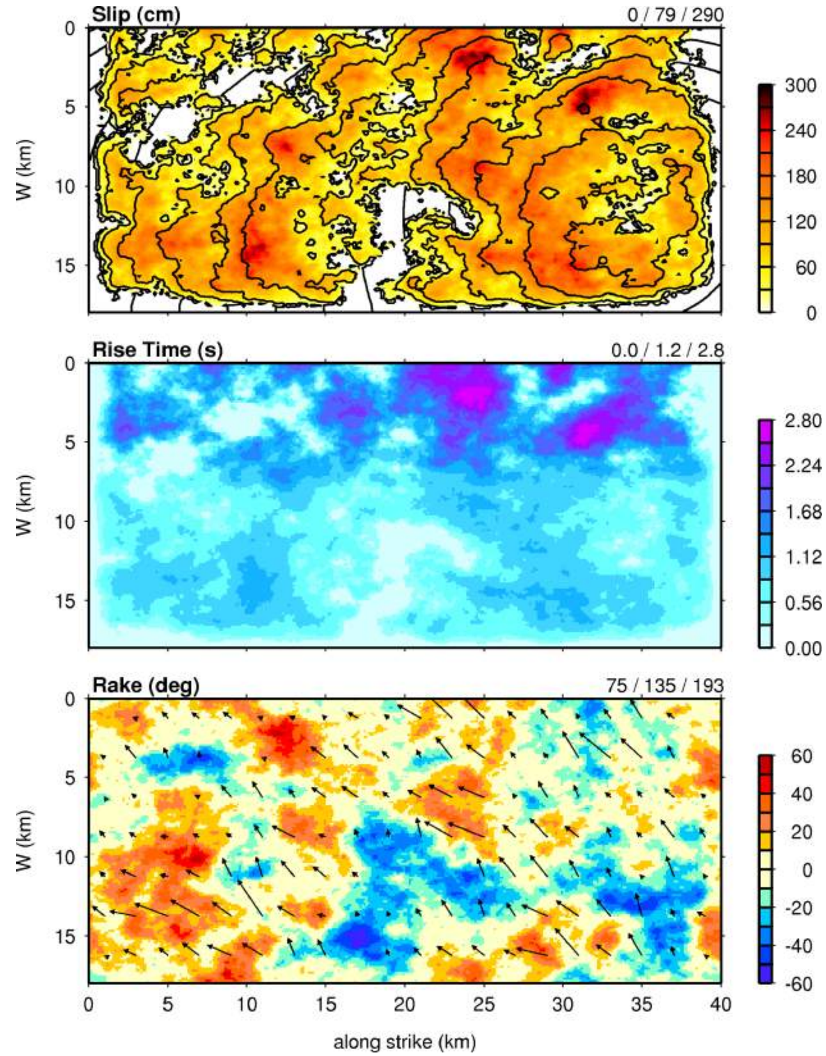
Conditional Slip Distribution

Graves-Pitarka Pseudo-Dynamic Rupture Models

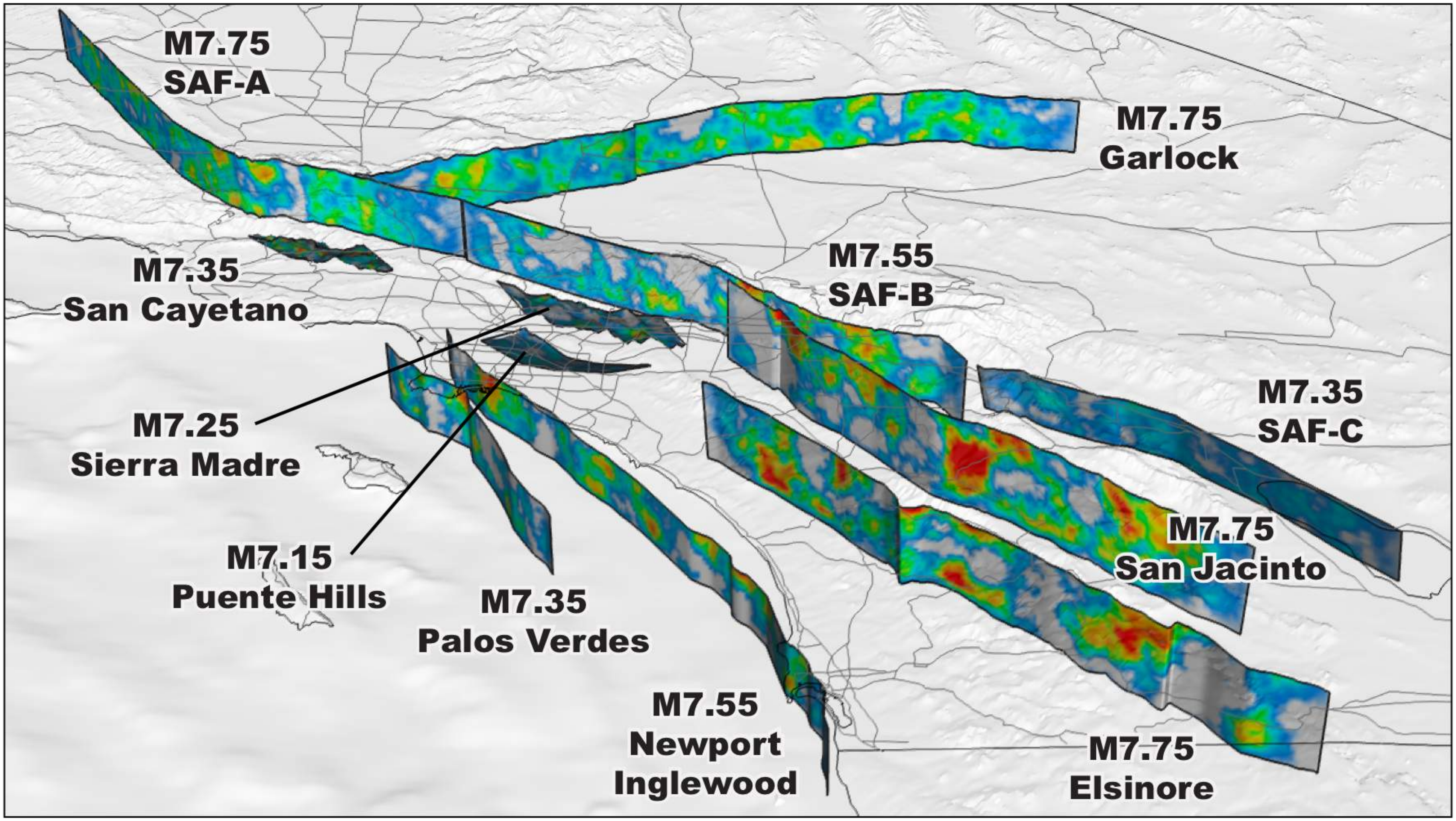
GP07
used in CS11



GP10
used in CS13

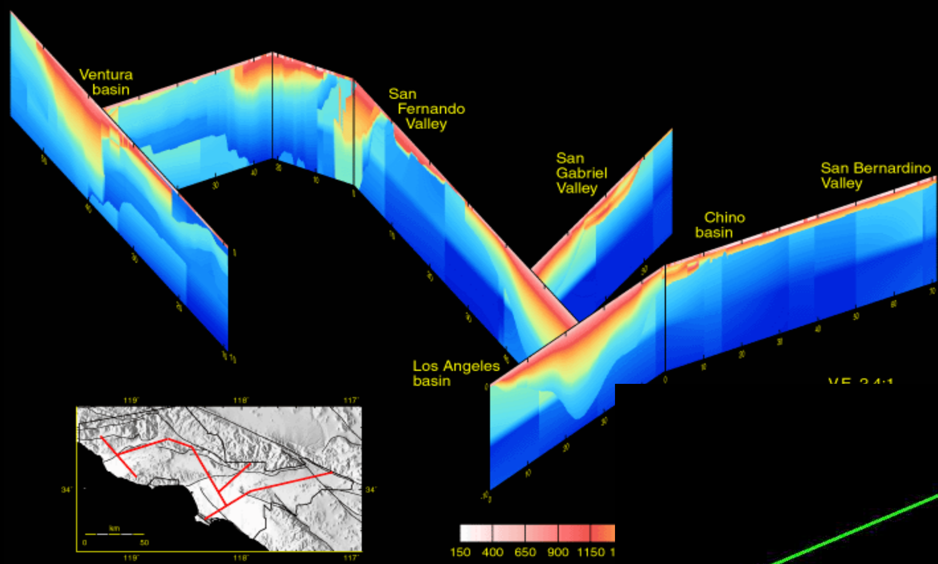


CyberShake Rupture Models

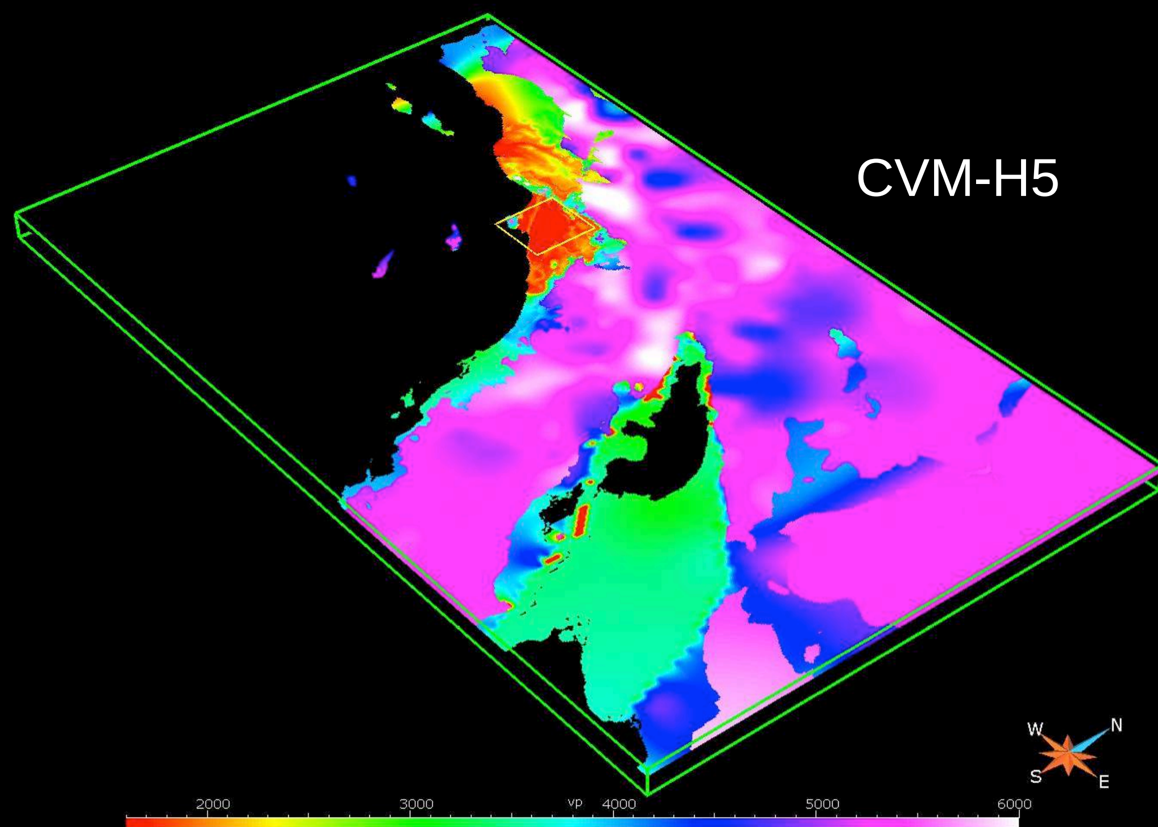


SCEC Community Velocity Models (CVMs)

CVM-S4



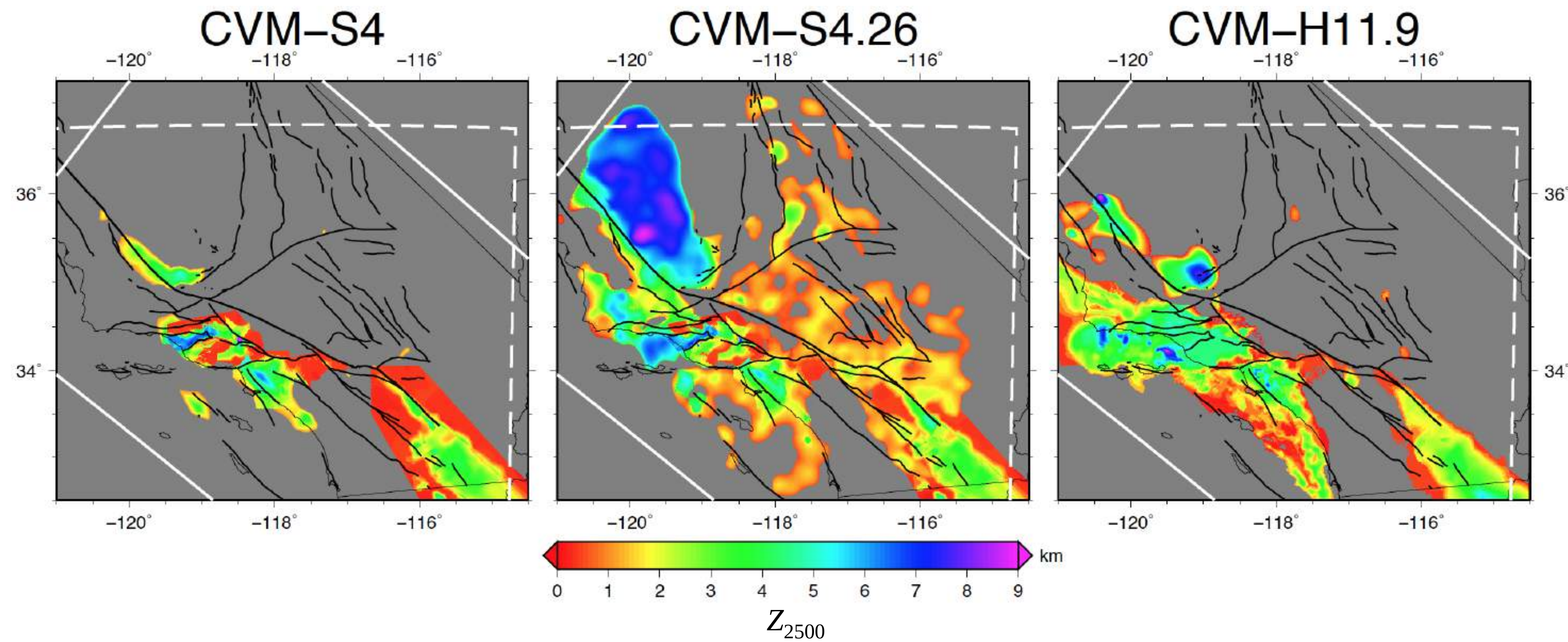
CVM-H5



Data sources

- **Surface geology**
- **Well logs**
- **Refraction surveys**
- **Reflection surveys**
- **Seismic tomography**
- **Geologic models**

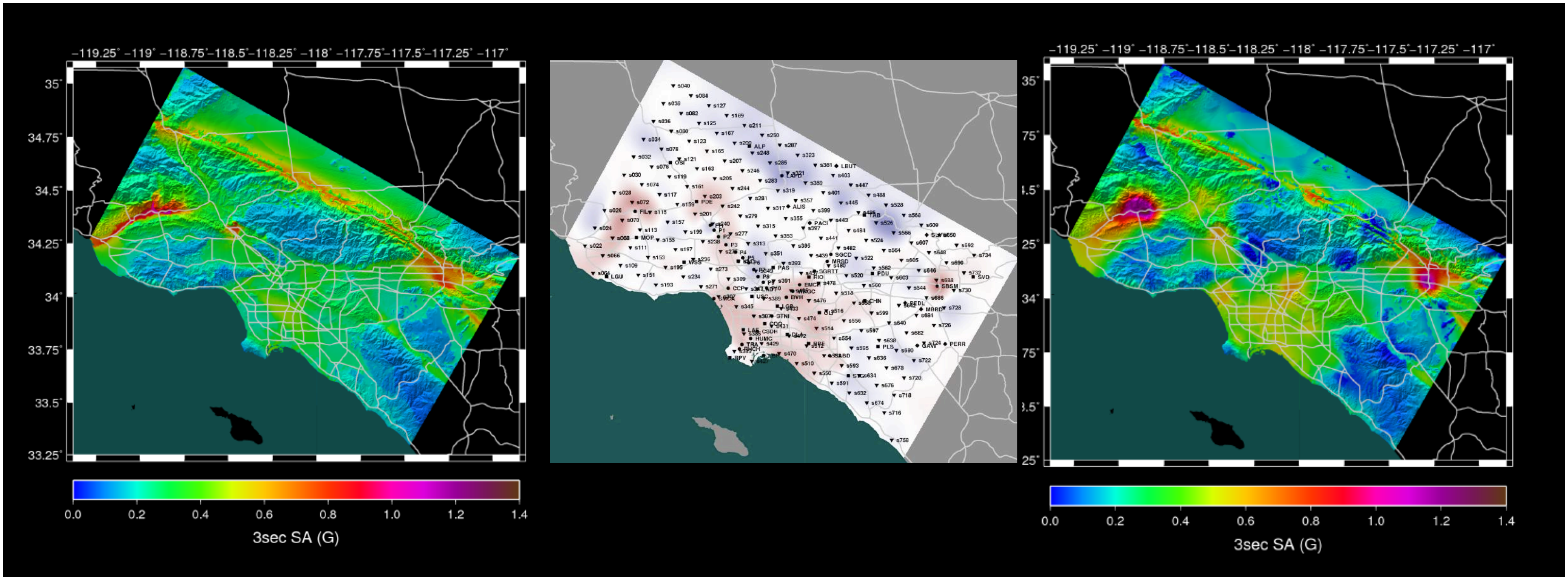
Basin Structures



Z_{2500} **:** iso-velocity surfaces at V_S = 2.5 km/s

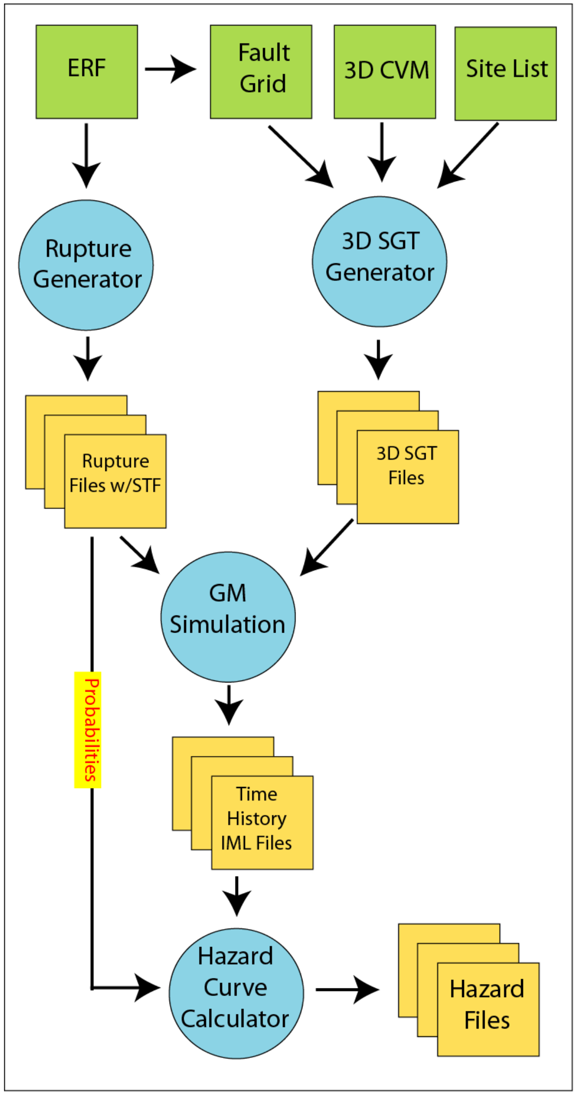
CyberShake Hazard Map Interpolation

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

CyberShake Workflow



ERF (Earthquake Rupture Forecast)
Defines fault surfaces, dimensions, segmentation, magnitudes and rupture probabilities

Rupture Generator
Generates full kinematic rupture description given fault geometry / dimension and magnitude. Includes slip distribution, hypocenter location, rupture propagation and slip time function (STF).

SGT (Strain Green's Tensor) Generator
3D finite difference simulation code to calculate SGTs at each site of interest. SGTs are computed using reciprocity by inserting a force at the site location and recording the strains on all potential fault surfaces. Storage of SGT files is on the order of 250 GBytes per site.

GM (Ground Motion) Simulation
Each rupture is simulated by convolving the SGTs appropriate for the specific fault with the kinematic rupture parameters. The output of the simulation is a time history of the ground motion at the site for the given rupture. Spectral acceleration (SA) values are also computed and stored for each time history.

Hazard Curve Calculator
Hazard curves are defined as the probability of exceeding a particular ground motion level (IML) in a specified time span. The curves are computed by combining the occurence probability and simulated ground motion for each rupture, and then integrating over all possible ruptures.

CyberShake Workflow

UCVM → **AWP-ODC** → **SeisSynth** → (data transfer) → **Data Product Generation**

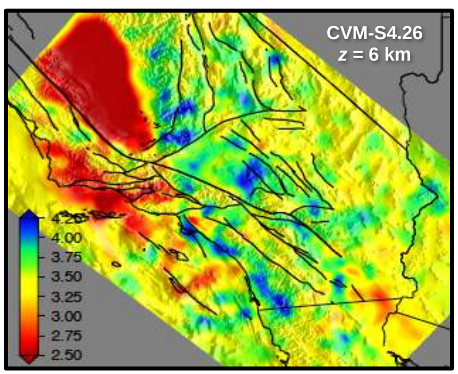
UCVM
Mesh generation
1 job per site
MPI, 320 cores

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

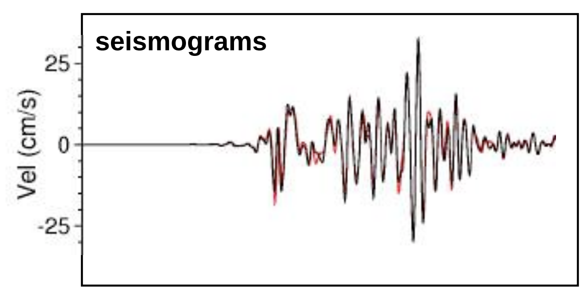
SeisSynth
Post-processing
415,000 jobs per site
serial

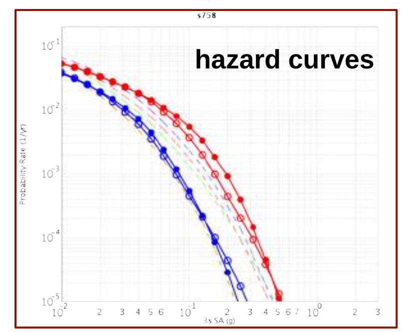
data transfer

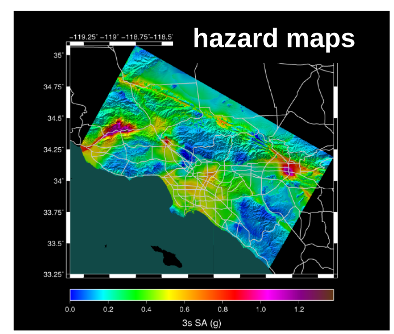
Data Product Generation
Populate DB, construct queries
4 jobs per site



Full-3D tomographic model of S. California

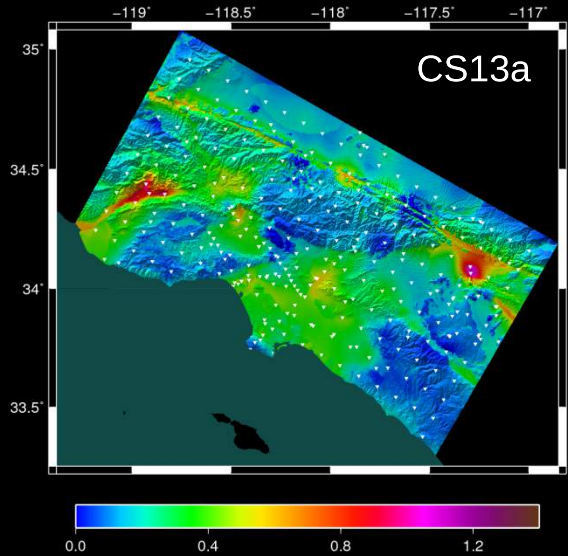


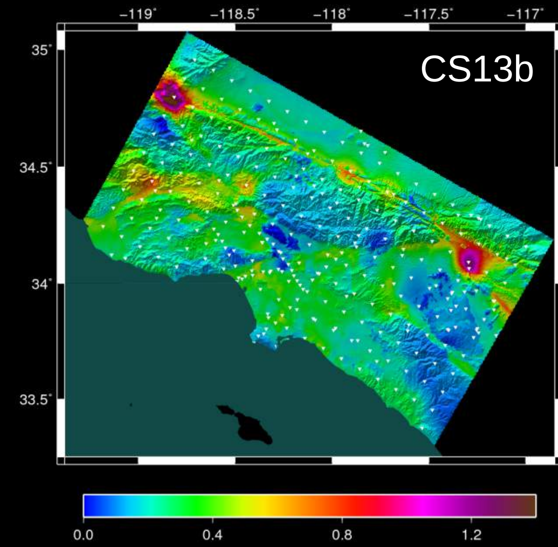




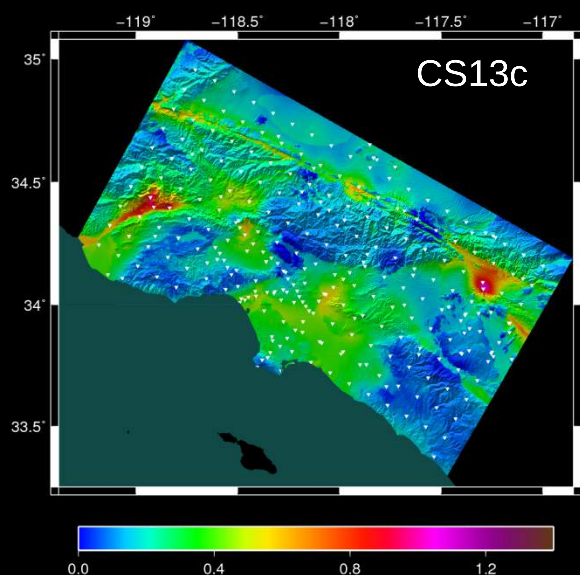
CyberShake Hazard Maps from CS13 Study

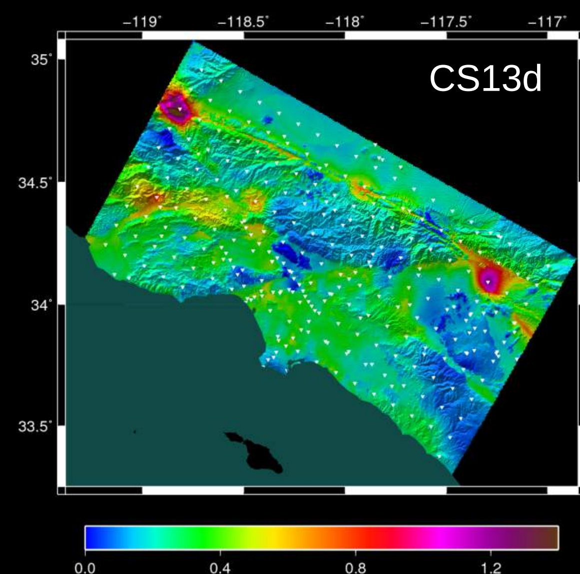
Graves SGTs





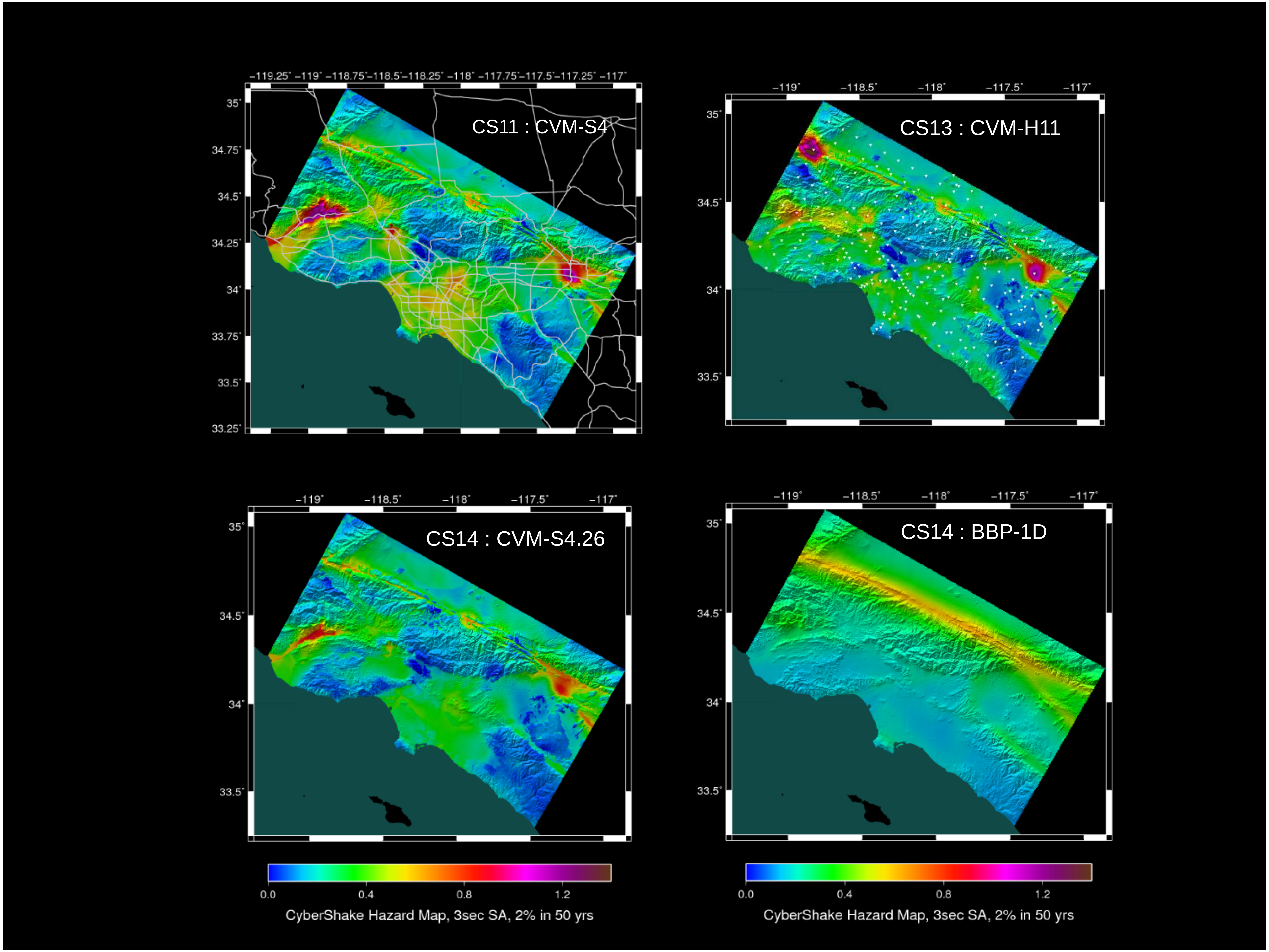
AWP-ODC SGTs





CVM-S4

CVM-H11



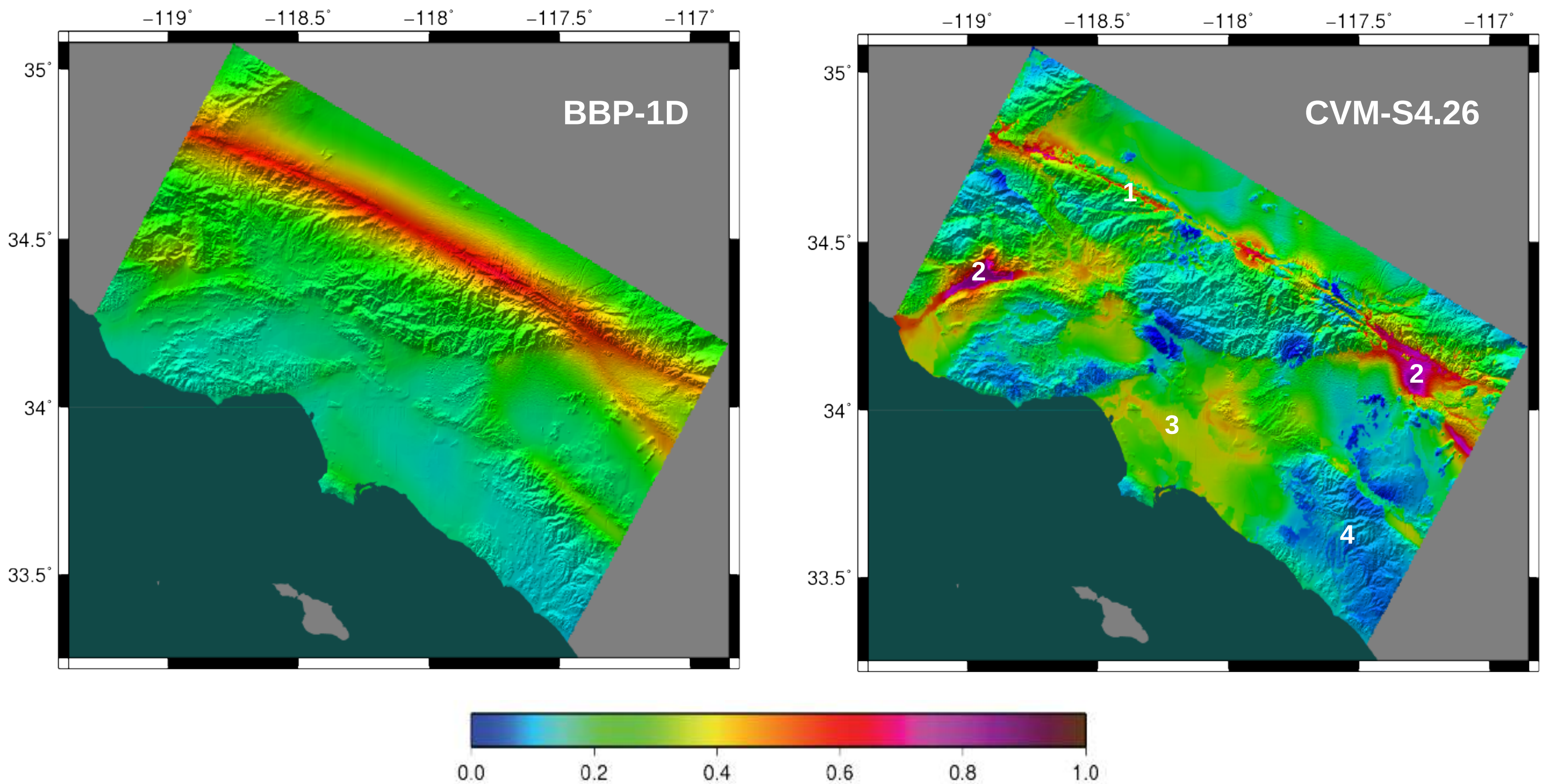
CS11 : CVM-S4
−119.25° −119° −118.75° −118.5° −118.25° −118° −117.75° −117.5° −117.25° −117°
35° 34.75° 34.5° 34.25° 34° 33.75° 33.5° 33.25°
CS13 : CVM-H11
−119° −118.5° −118° −117.5° −117°
35° 34.5° 34° 33.5°
CS14 : CVM-S4.26
−119° −118.5° −118° −117.5° −117°
35° 34.5° 34° 33.5°
CS14 : BBP-1D
−119° −118.5° −118° −117.5° −117°
35° 34.5° 34° 33.5°
0.0 0.4 0.8 1.2
CyberShake Hazard Map, 3sec SA, 2% in 50 yrs
0.0 0.4 0.8 1.2
CyberShake Hazard Map, 3sec SA, 2% in 50 yrs

CyberShake Workflow

Computational statistics for CS14.2 study:

- Reservation for 700 XE nodes, 200 XK nodes
- 1144 CyberShake sites
  - 568 with SGT CPU
    - 2792 sec/job x 313.8 nodes = 243.4 node-hrs
    - Queue time: mean 973 sec, median 191 sec
  - 568 with SGT GPU
    - 1338 sec/job x 100 nodes = 37.2 node-hrs (6.5x efficiency improvement)
    - Queue time: mean 2889 sec, media 731 sec
- 99.8 million tasks produced 470 million seismograms
  - 81 tasks/sec
- 31,463 jobs submitted remotely to the Blue Waters queue
- 860 TB of data managed
  - 57 TB output files
  - 12 TB staged back to SCEC storage

Comparison of 1D and 3D CyberShake Models for the Los Angeles Region



CyberShake Hazard Map, 3sec SA, 2% in 50 yrs

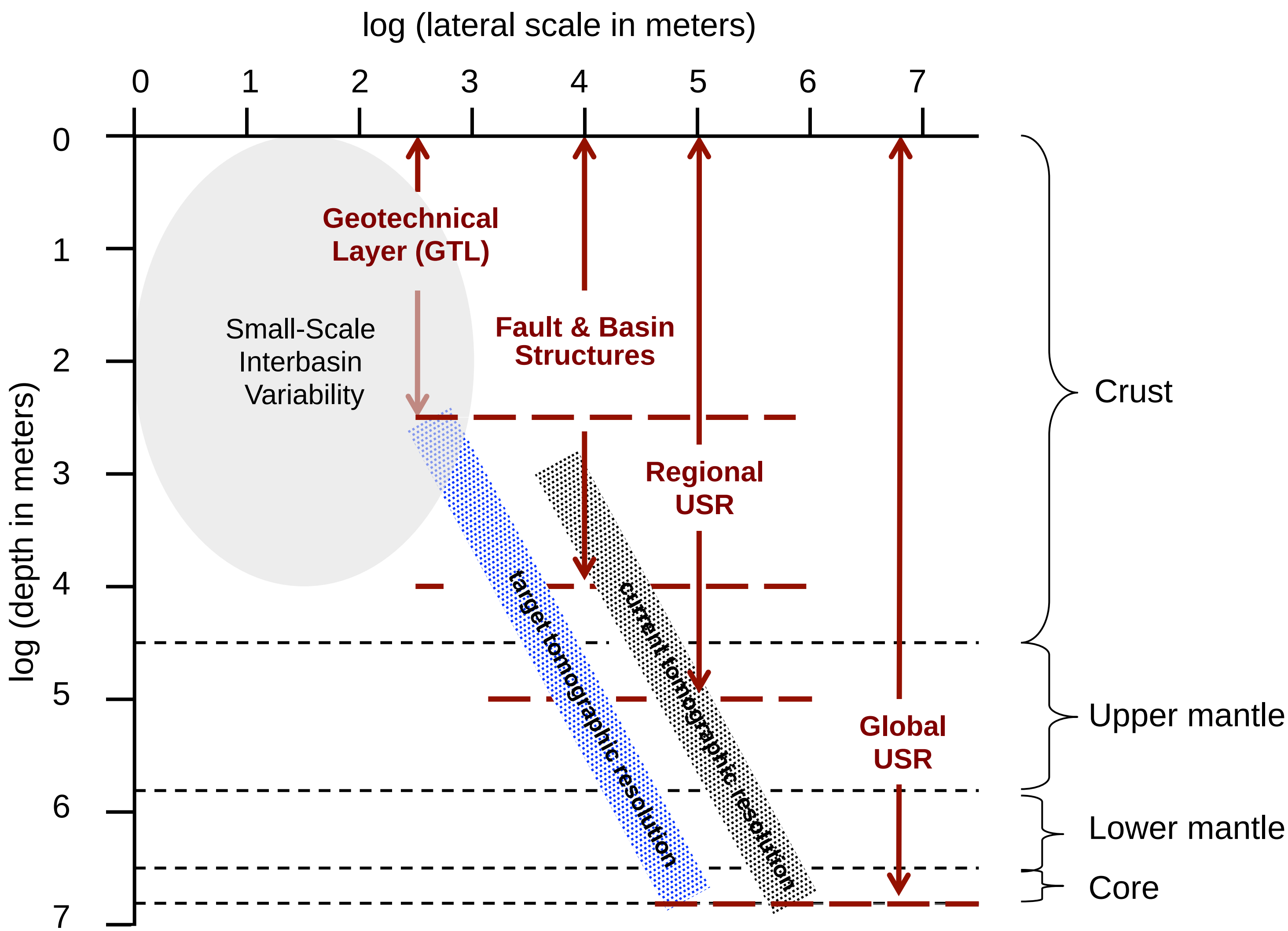
CyberShake Workflow

Los Angeles Region Hazard Models (1144 sites)

CyberShake Application Metrics (Hours)	2008 (Mercury, normalized)	2009 (Ranger, normalized)	2013 (Blue Waters / Stampede)	2014 (Blue Waters)
Application Core Hours:	19,488,000 (CPU)	16,130,400 (CPU)	12,200,000 (CPU)	15,800,000 (CPU +GPU)
Application Makespan:	70,165	6,191	1,467	342

Metric	2013 (Study 13.4)	2014 (Study 14.2)
Simultaneous processors	21,100 (CPU)	46,720 (CPU) + 160 (GPU)
Concurrent Workflows	5.8	26.2
Job Failure Rate	2.6%	1.3%
Data transferred	52 TB	12 TB

3D Seismic Velocity Structure from Full-3D Inversion


log (lateral scale in meters)
0
1
2
3
4
5
6
7
log (depth in meters)
0
1
2
3
4
5
6
7
Geotechnical
Layer (GTL)
Small-Scale
Interbasin
Variability
Fault & Basin
Structures
Regional
USR
Global
USR
target tomographic resolution
current tomographic resolution
Crust
Upper mantle
Lower mantle
Core

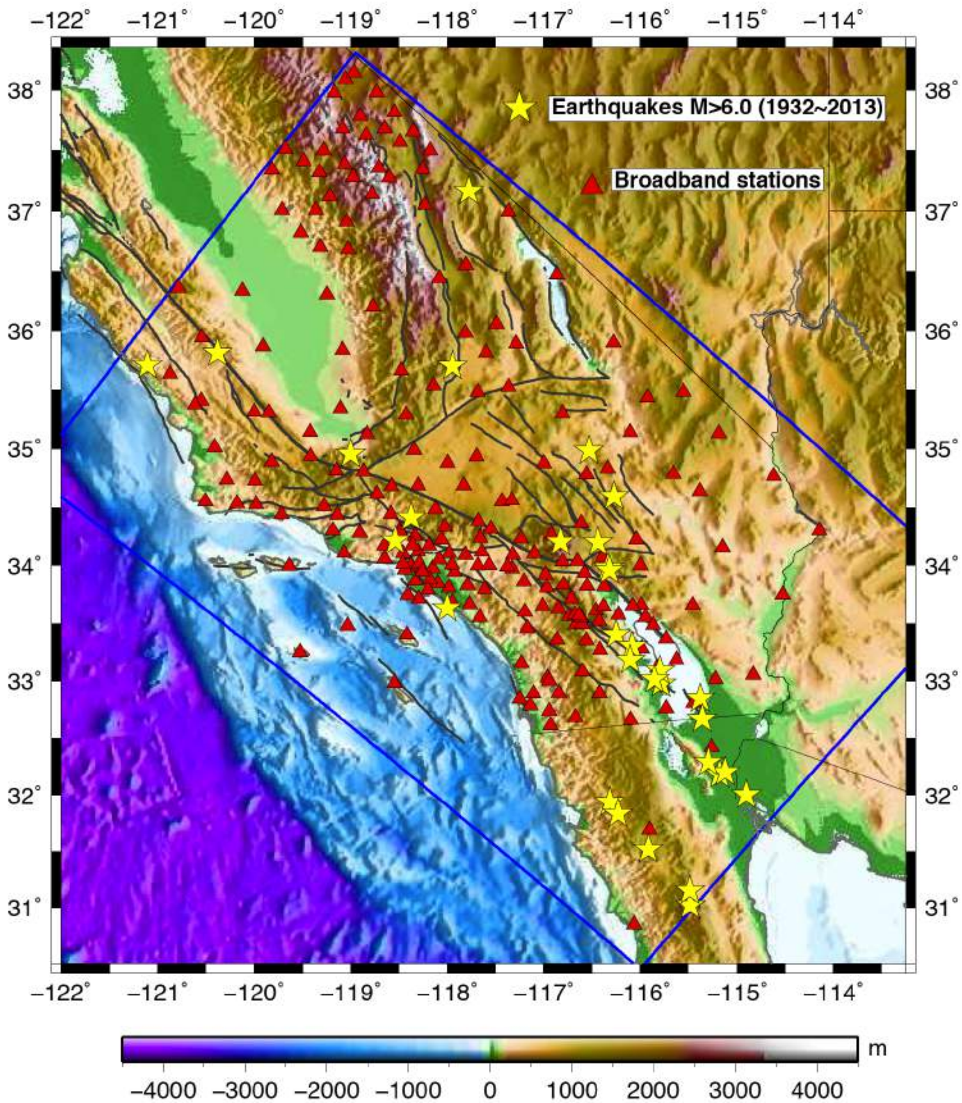
CVM-S4.26 Structure for Southern California

[1]En-Jui Lee, [2]Po Chen, [1]Thomas H. Jordan, [1]Philip Maechling, [3]Marine Denolle, [3]Greg Beroza

[1]University of Southern California,
[2]University of Wyoming,
[3]Stanford University

CVM-S4.26

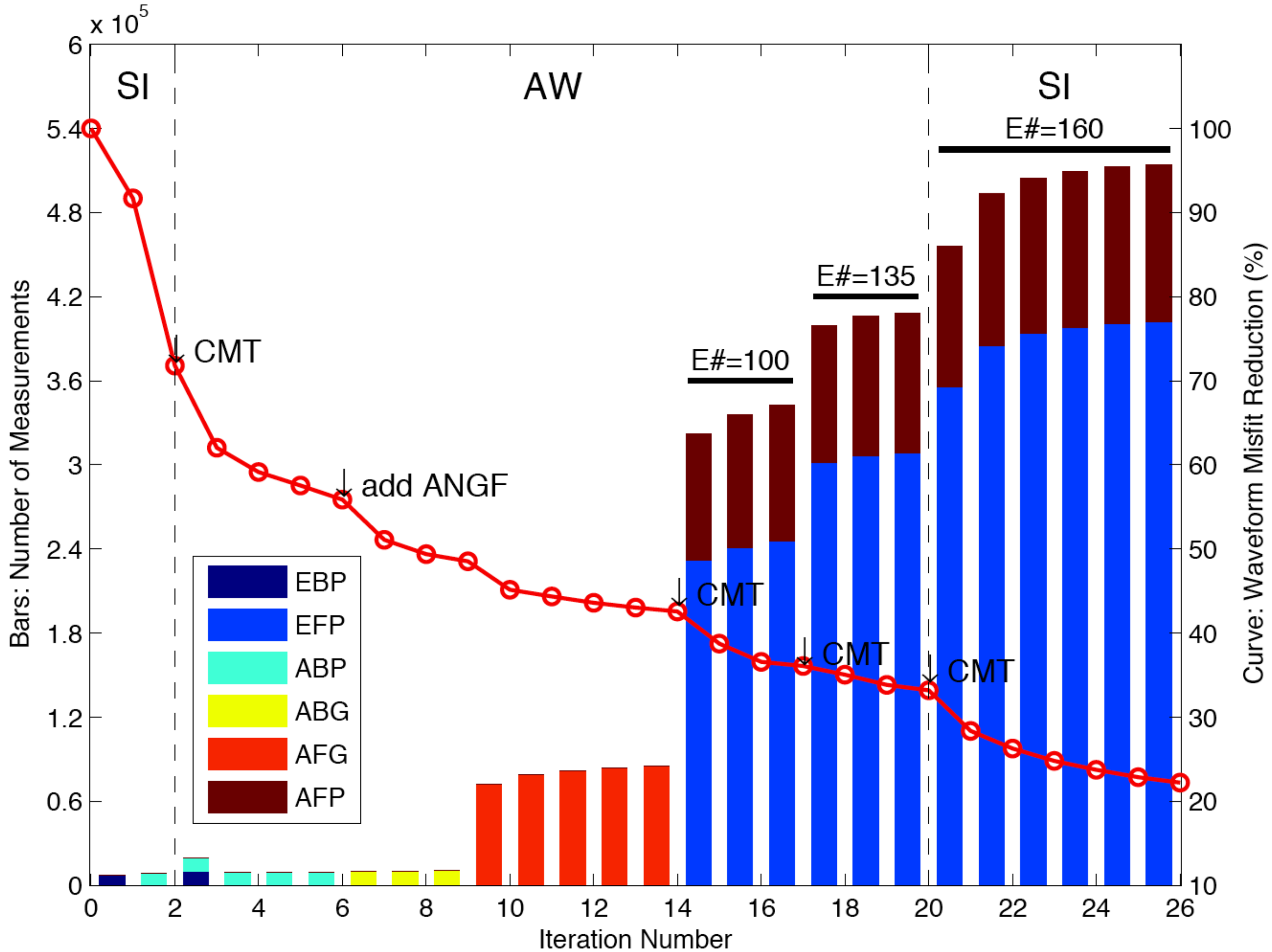
Full-3D tomography model of Southern California crustal structure

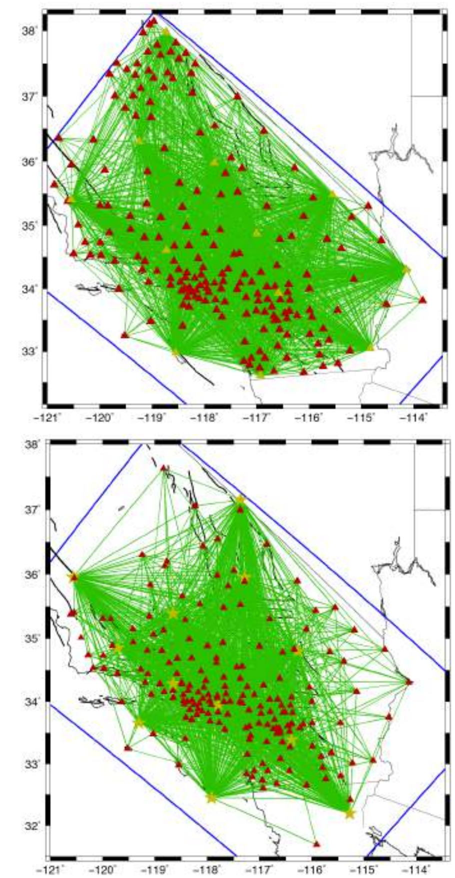


- **CVM-S4 starting model**
- **26th iterate of a full-3D tomographic (F3DT) inversion procedure (Lee et al., 2013).**
- **Data sets comprise ~ 550,000 differential waveform measurements at *f* ≤ 0.2 Hz**
  - **38,000 earthquake seismograms**
  - **12,000 ambient-noise Green functions**
- **Nonlinear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

CVM-S4.26

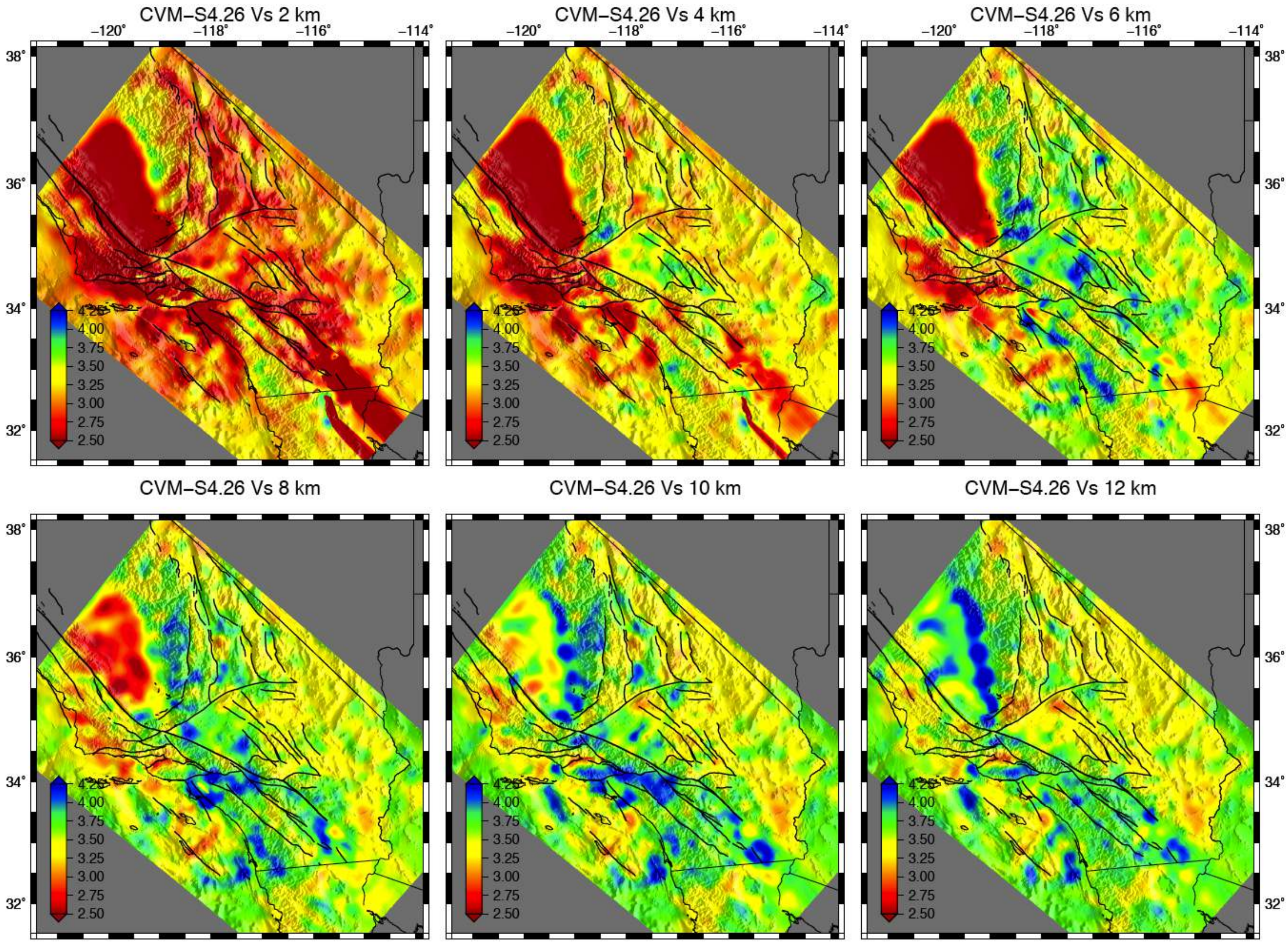
Full-3D tomography model of Southern California crustal structure





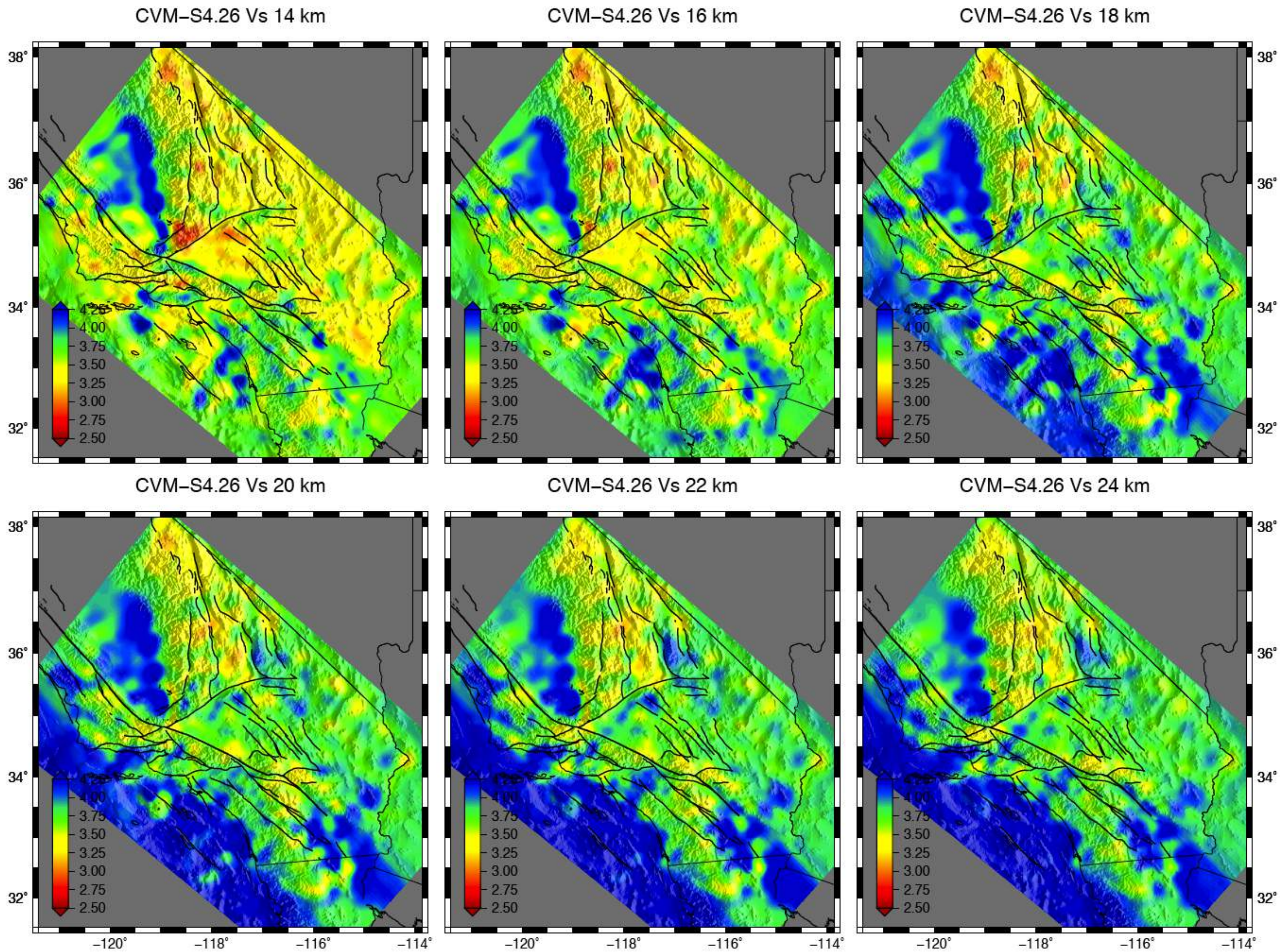
CVM-S4.26

Full-3D tomography model of Southern California crustal structure

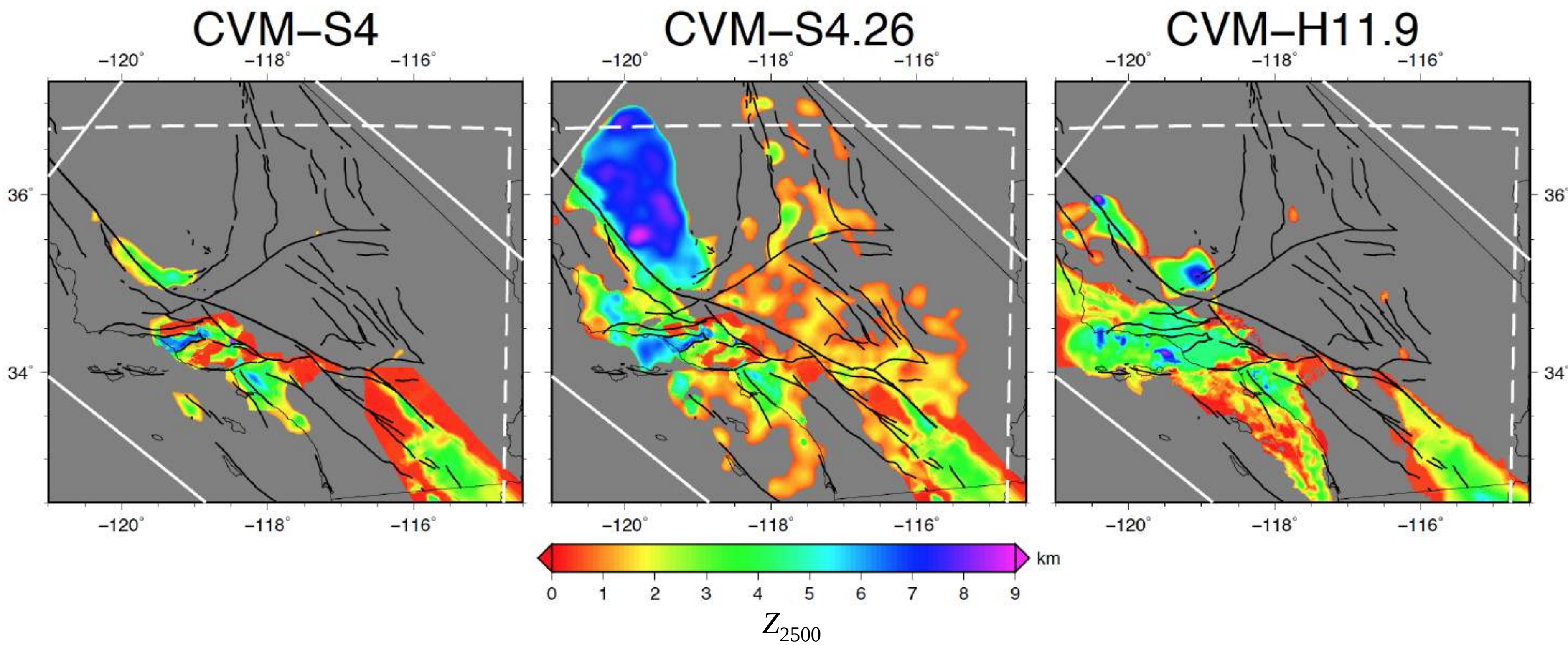


CVM-S4.26

Full-3D tomography model of Southern California crustal structure



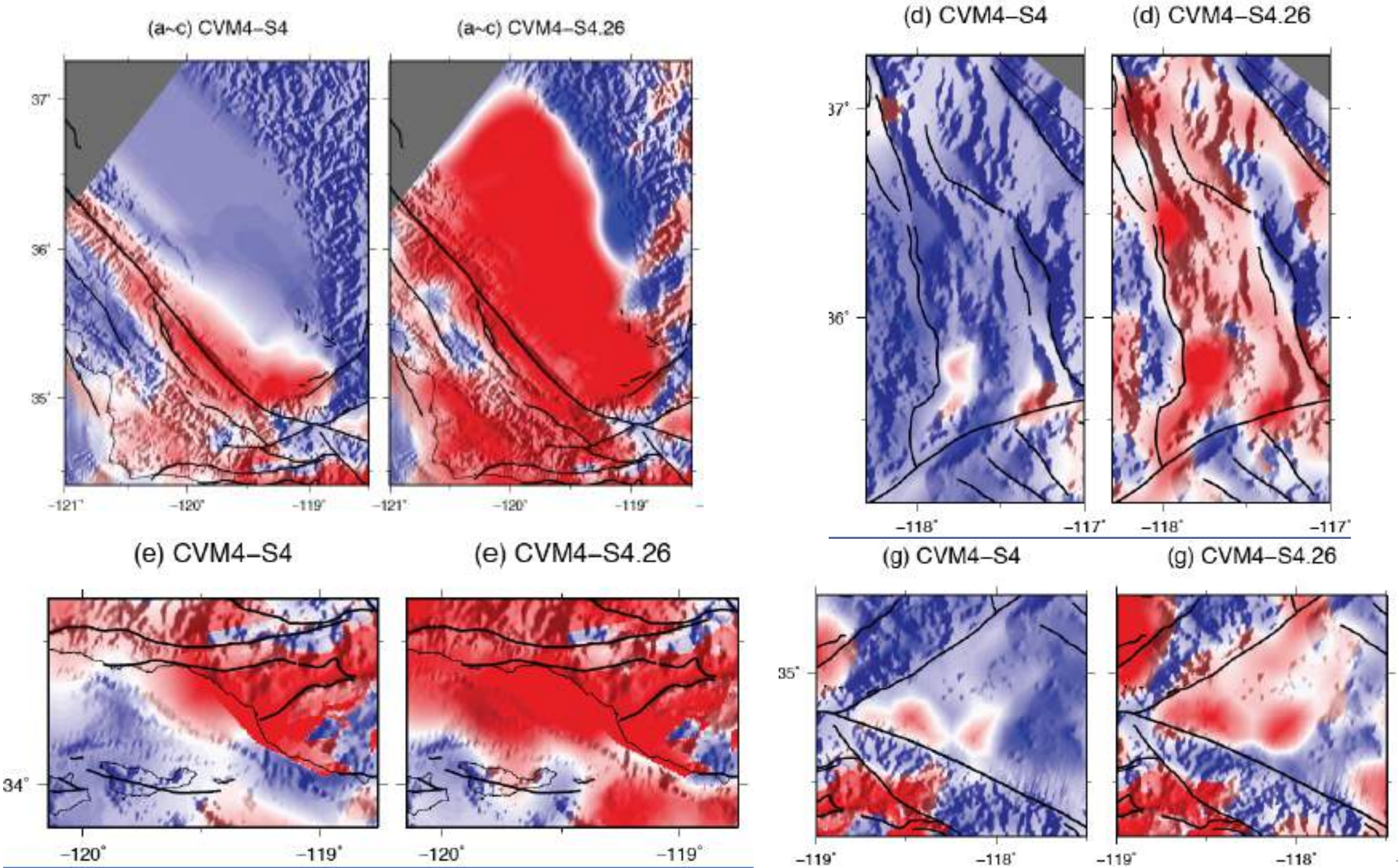
Basin Structures



Z_{2500} **:** iso-velocity surfaces at V_S = 2.5 km/s

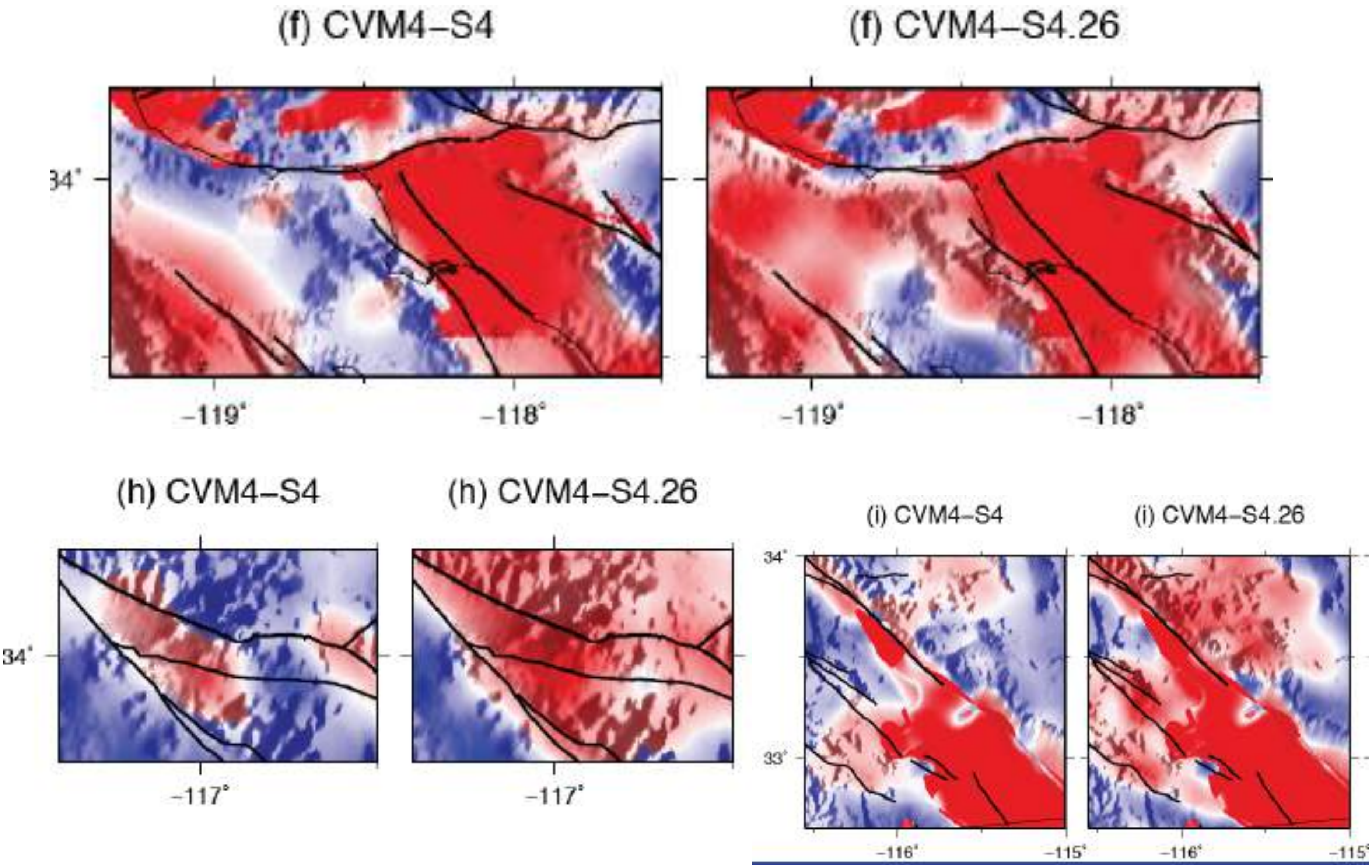
CVM-S4.26

Full-3D tomography model of Southern California crustal structure

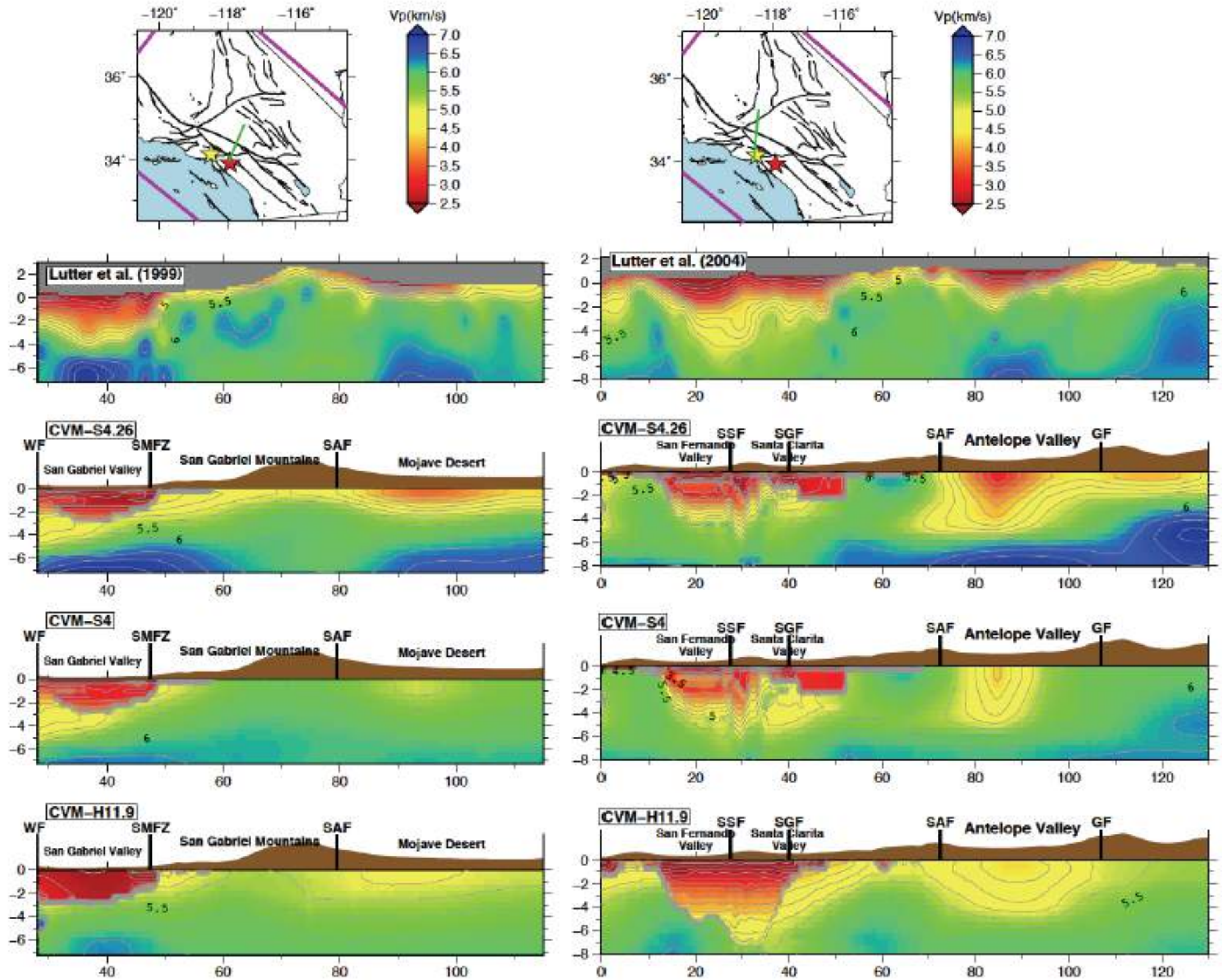


CVM-S4.26

Full-3D tomography model of Southern California crustal structure

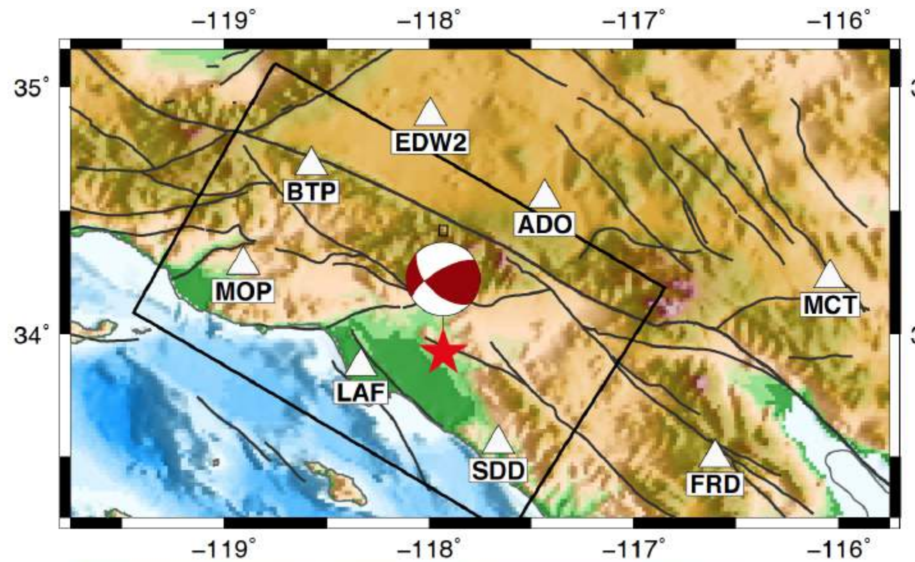


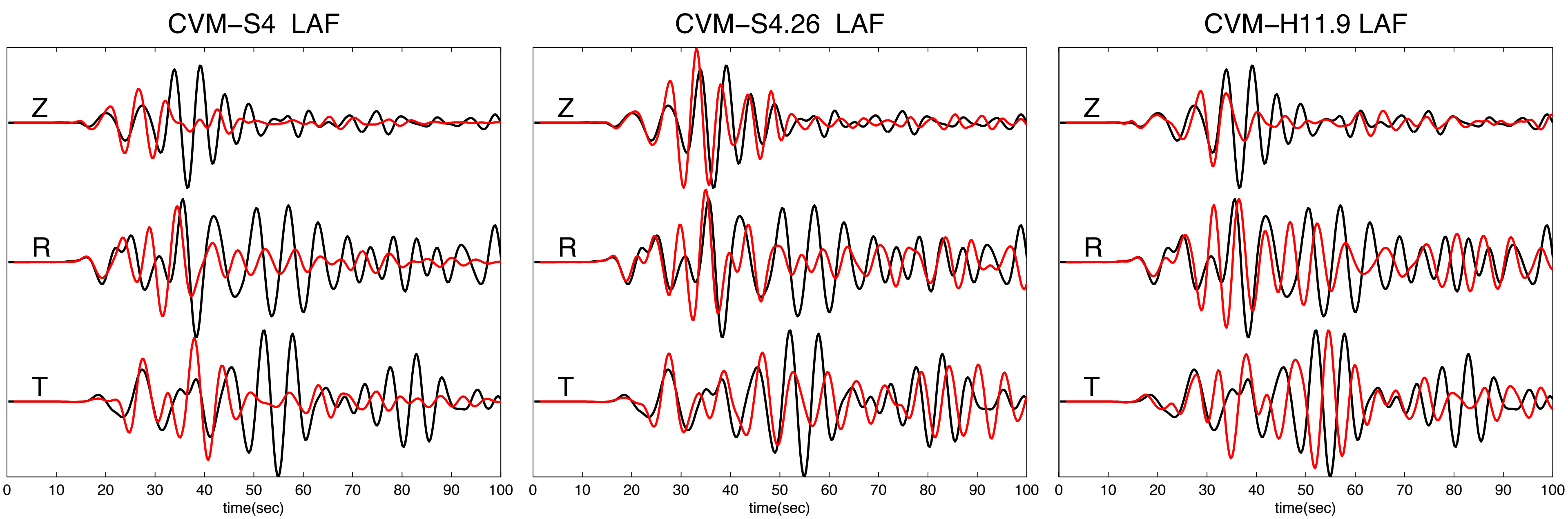
LARSE Profiles



03/28/14 La Habra Earthquake (M5.1)

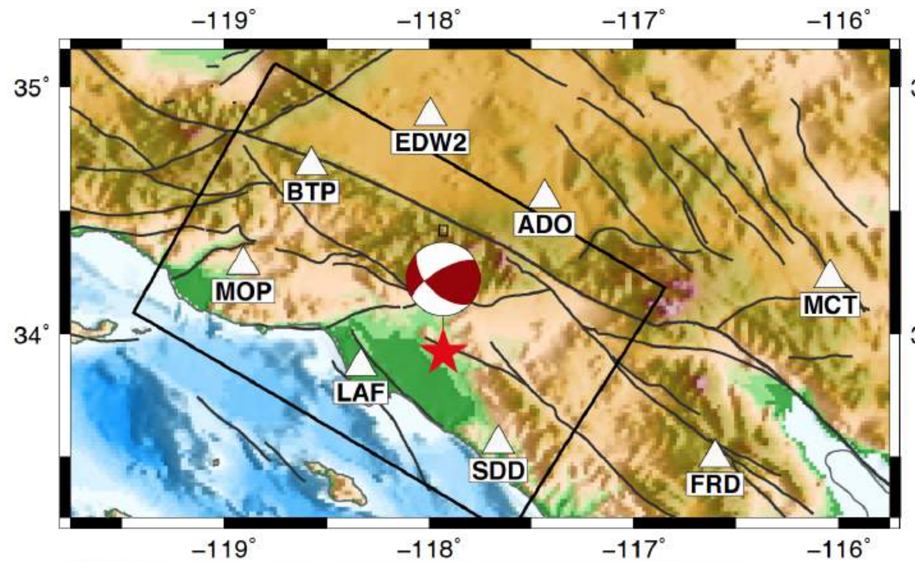
Observed: Black
Synthetic: Red

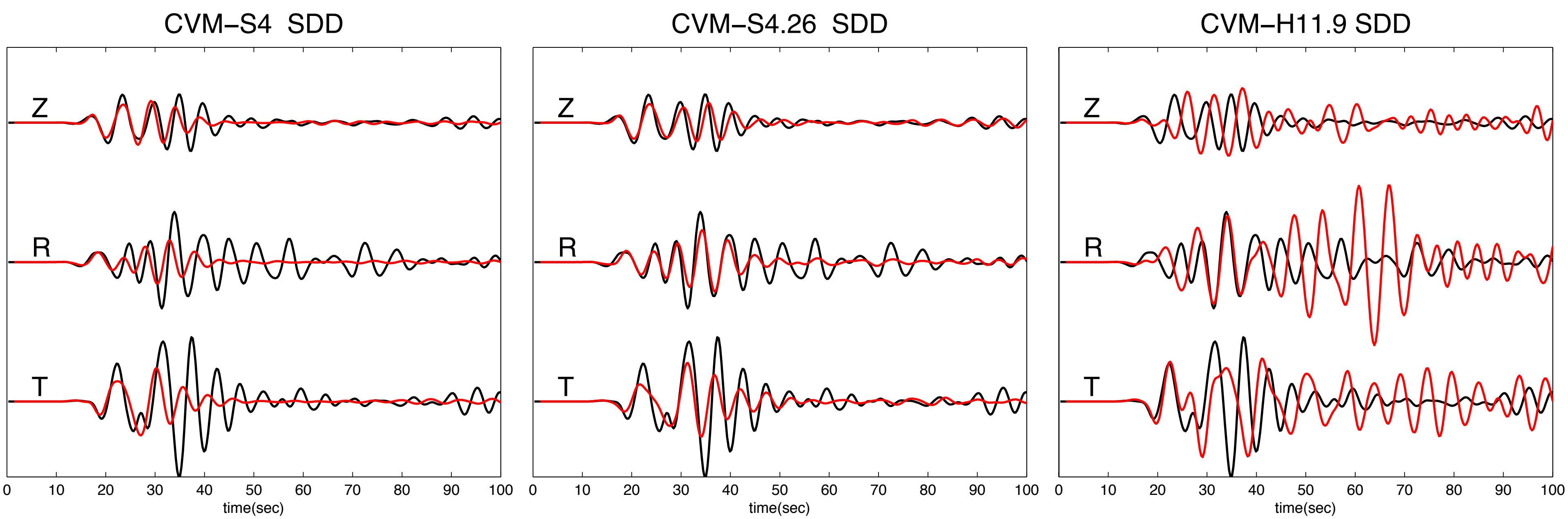




03/28/14 La Habra Earthquake (M5.1)

Observed: Black
Synthetic: Red



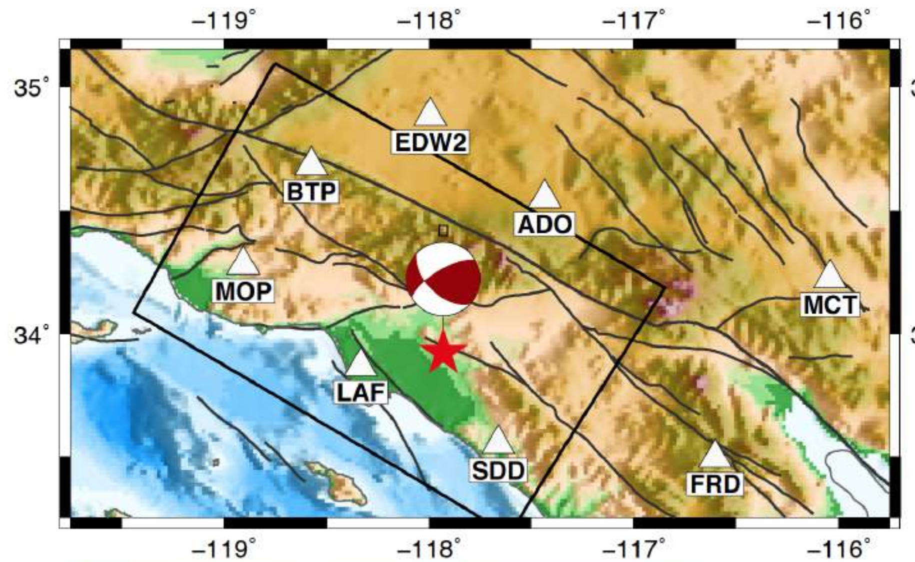


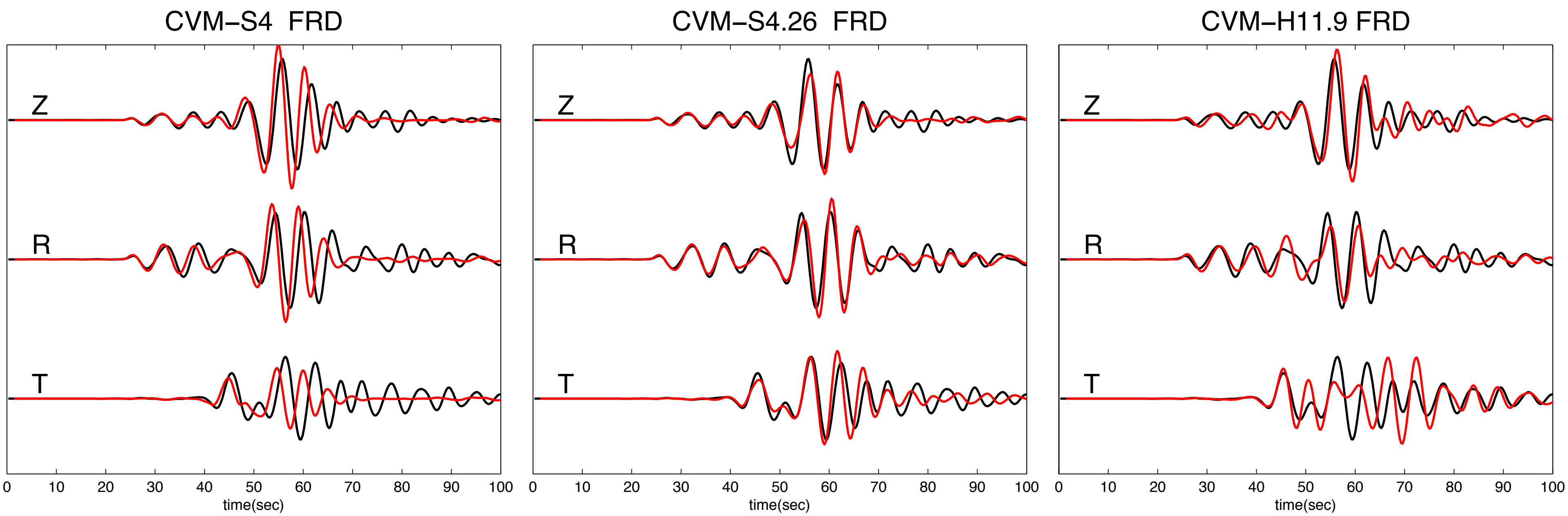
SCEC

Southern California Earthquake Center

03/28/14 La Habra Earthquake (M5.1)

Observed: Black
Synthetic: Red

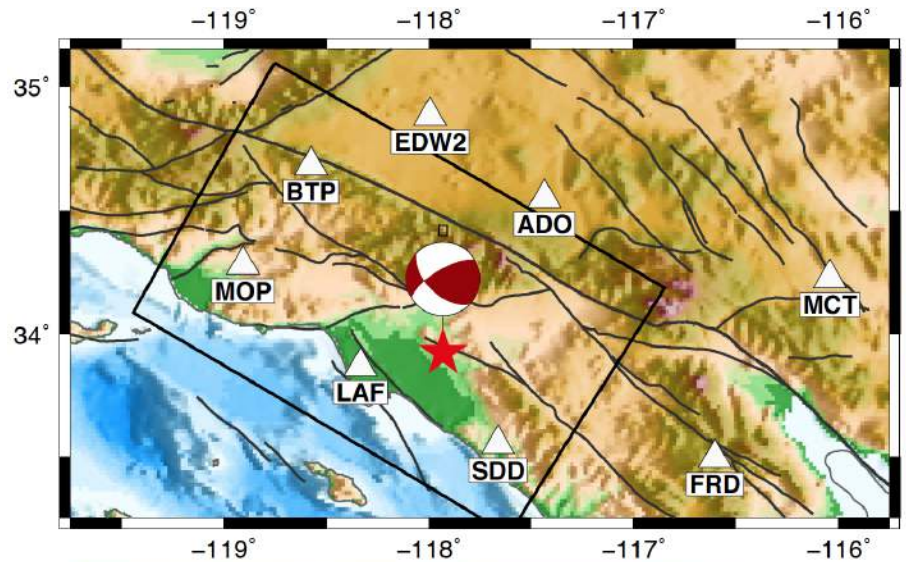


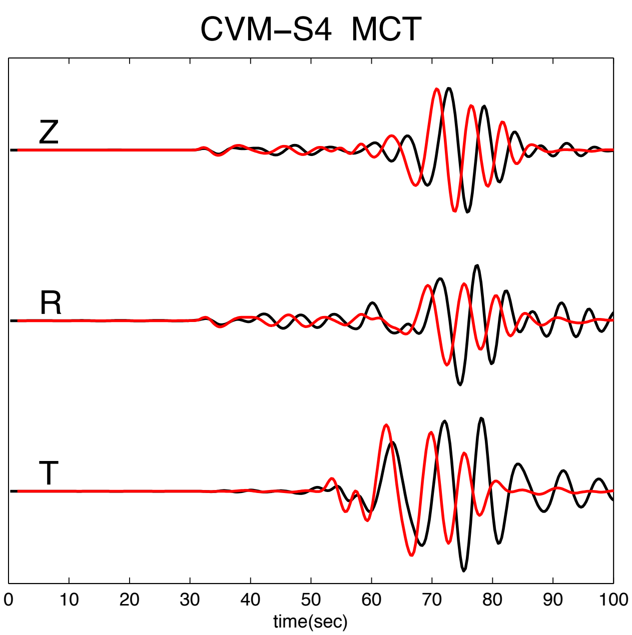


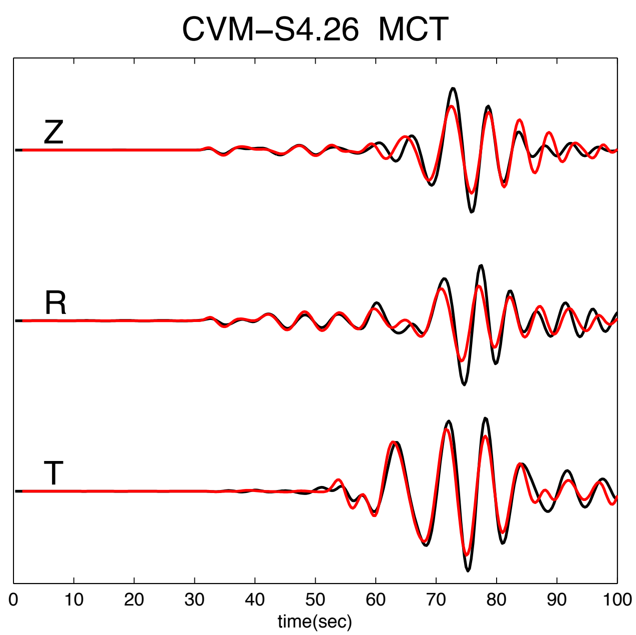
Southern California Earthquake Center

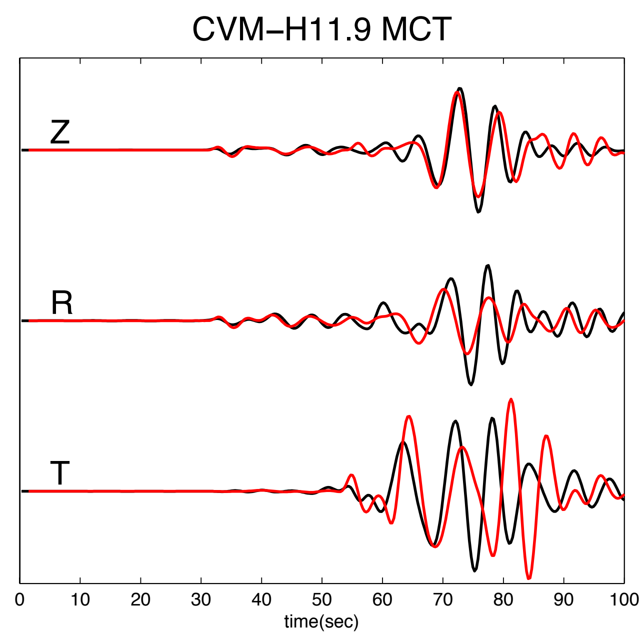
03/28/14 La Habra Earthquake (M5.1)

Observed: Black
Synthetic: Red





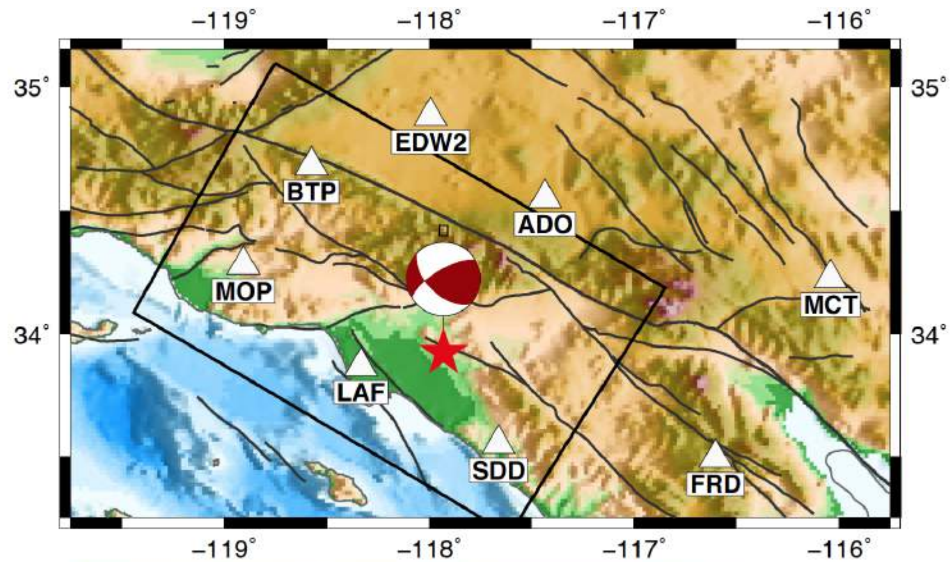


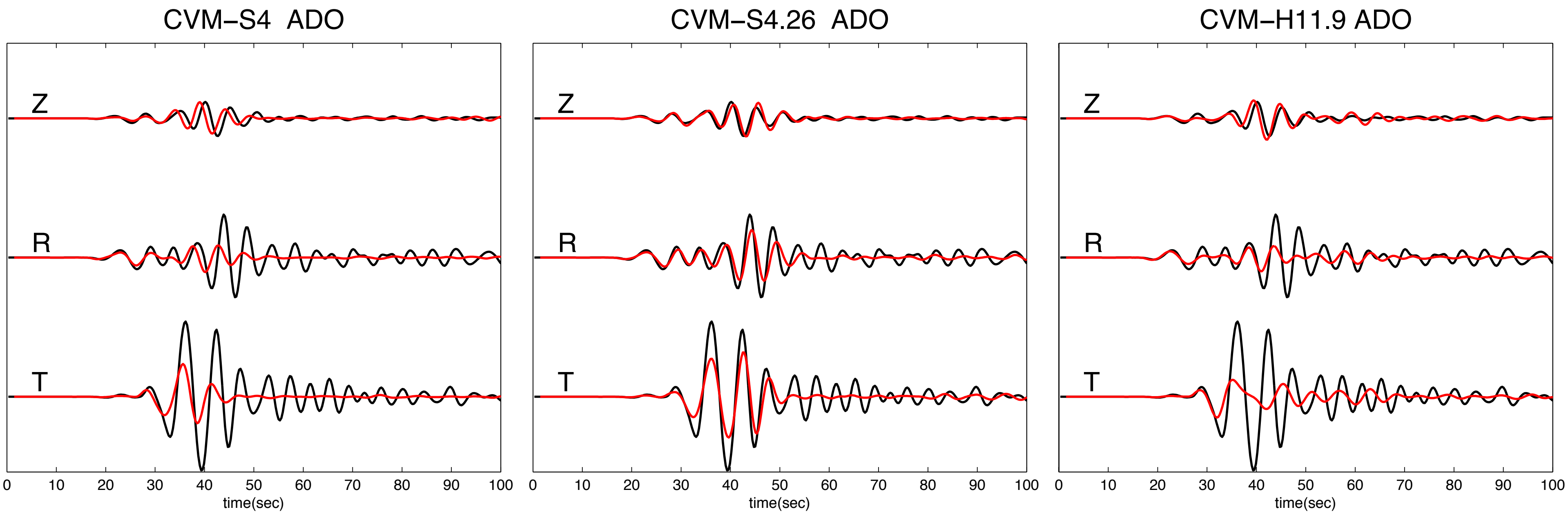


03/28/14 La Habra Earthquake (M5.1)

Southern California Earthquake Center

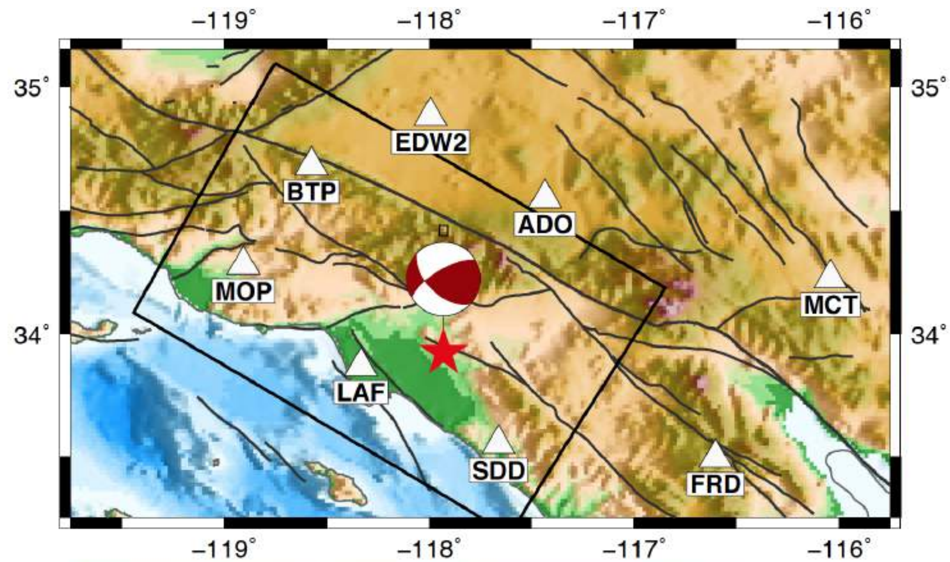
Observed: Black
Synthetic: Red

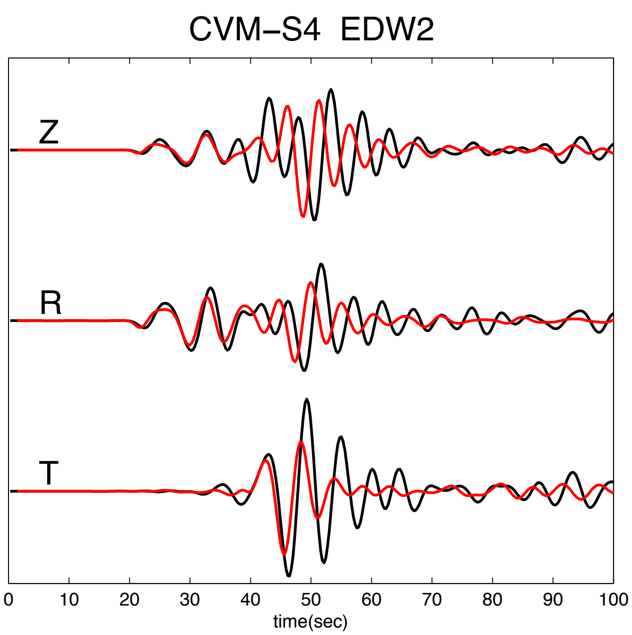


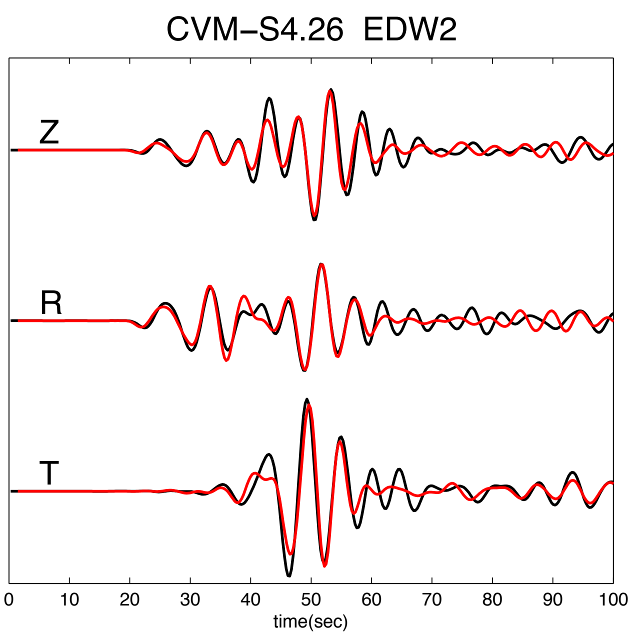


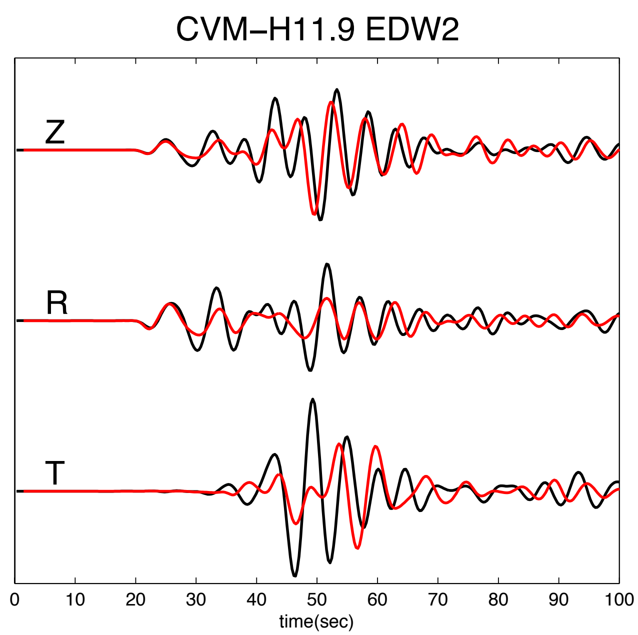
03/28/14 La Habra Earthquake (M5.1)

Observed: Black
Synthetic: Red





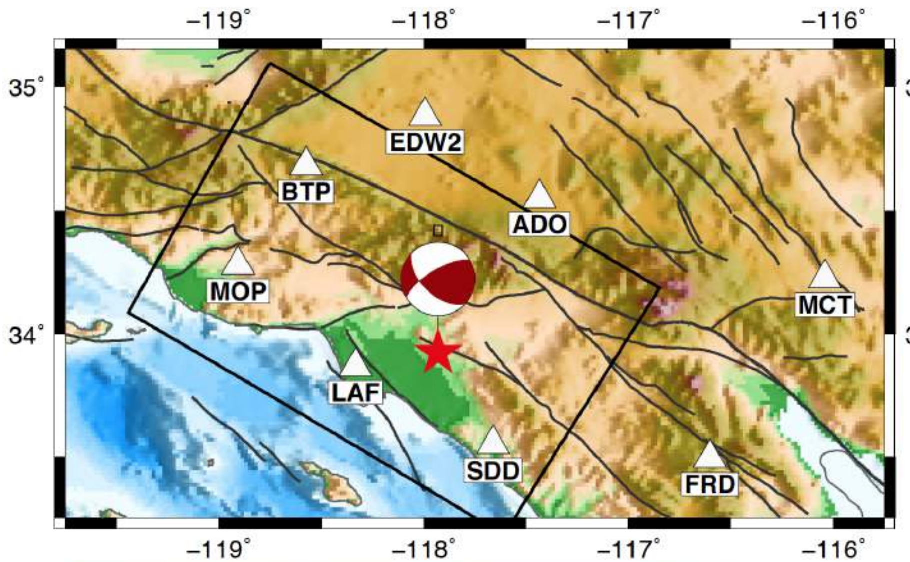


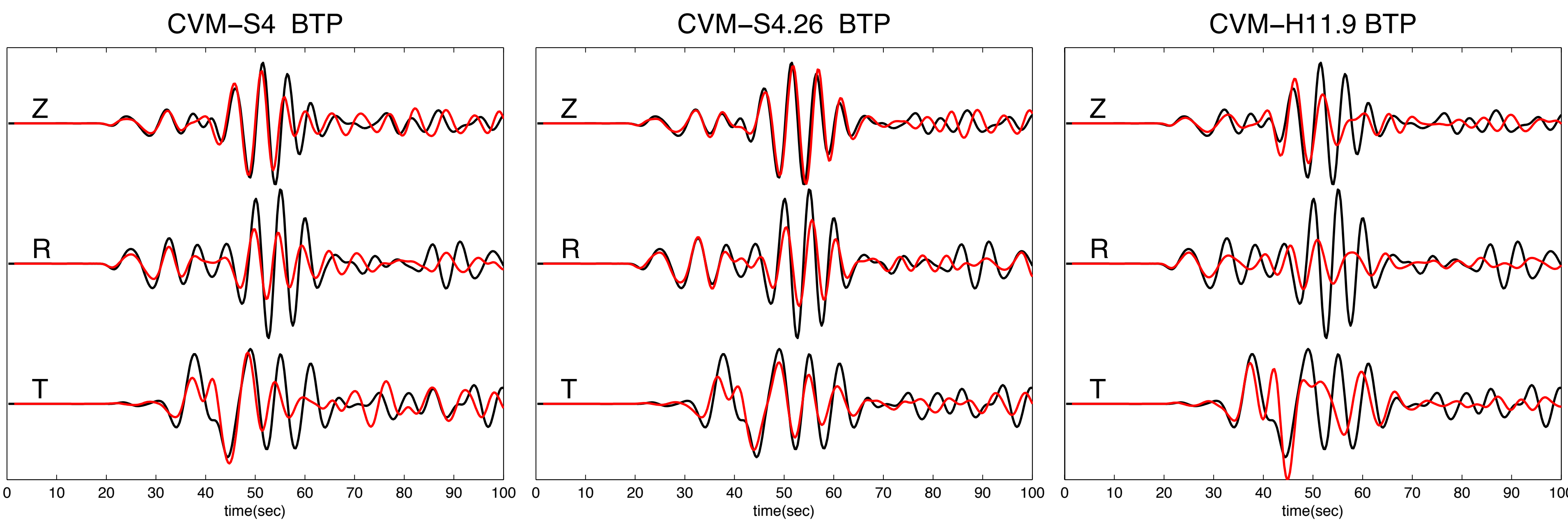


Southern California Earthquake Center

03/28/14 La Habra Earthquake (M5.1)

Observed: Black
Synthetic: Red

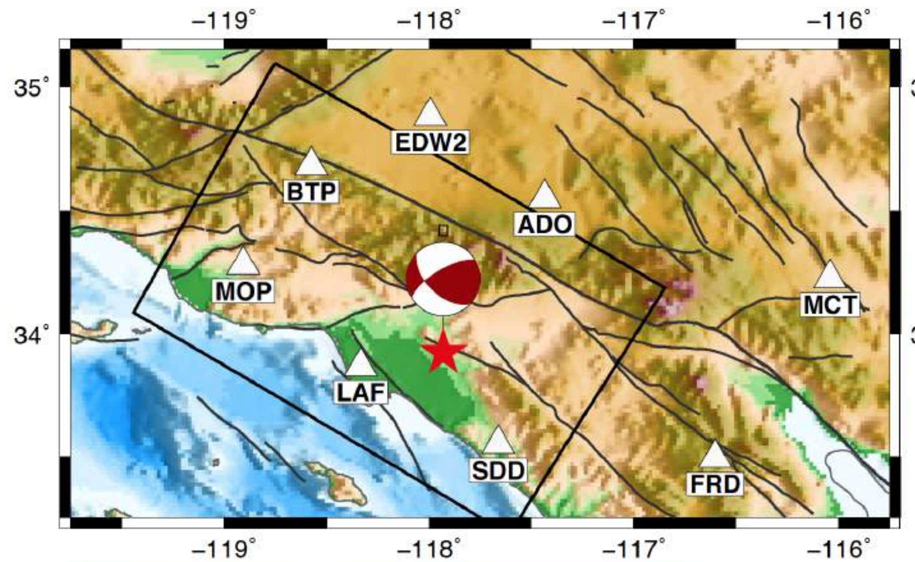


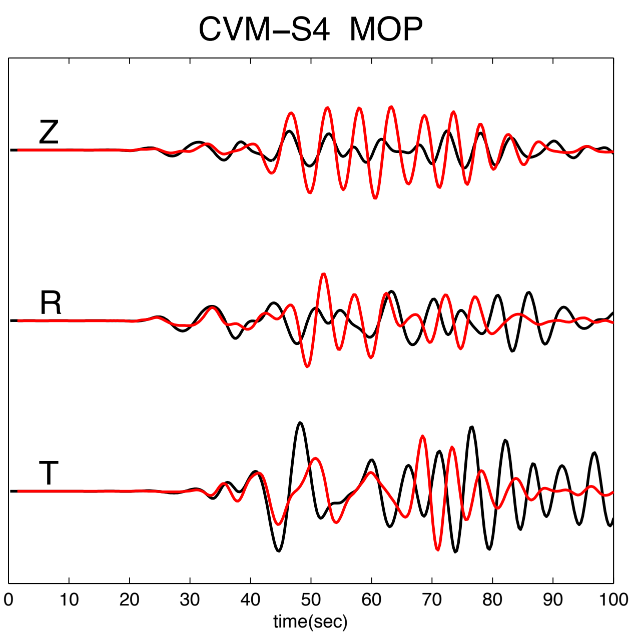


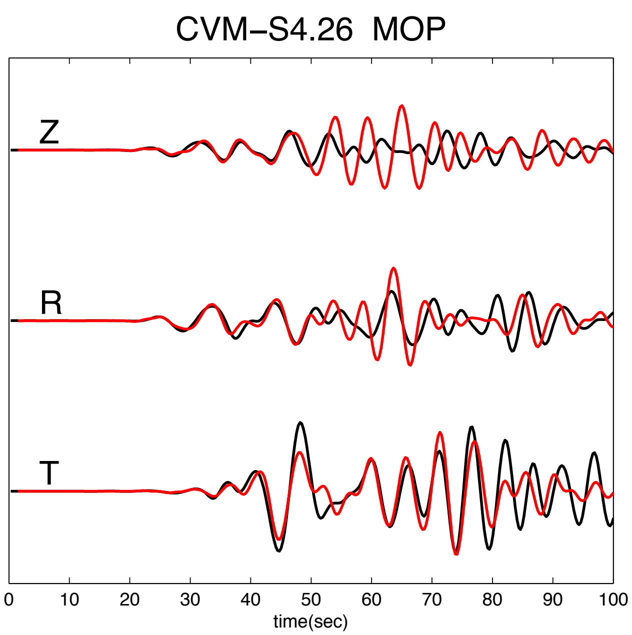
Southern California Earthquake Center

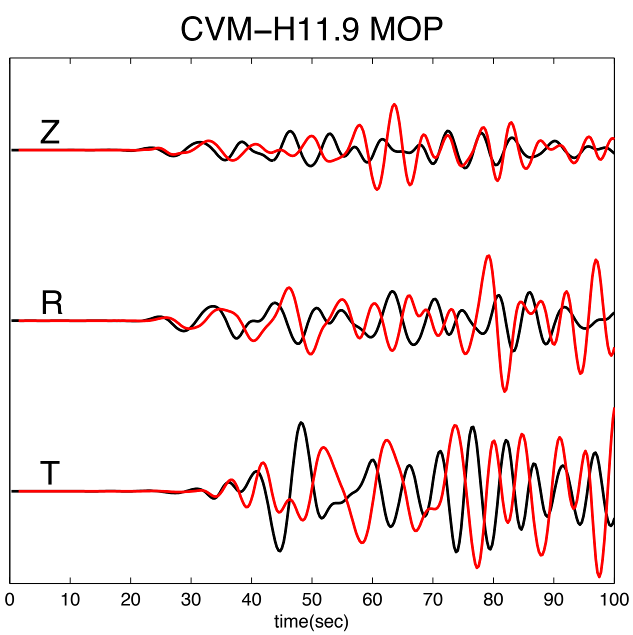
03/28/14 La Habra Earthquake (M5.1)

Observed: Black
Synthetic: Red

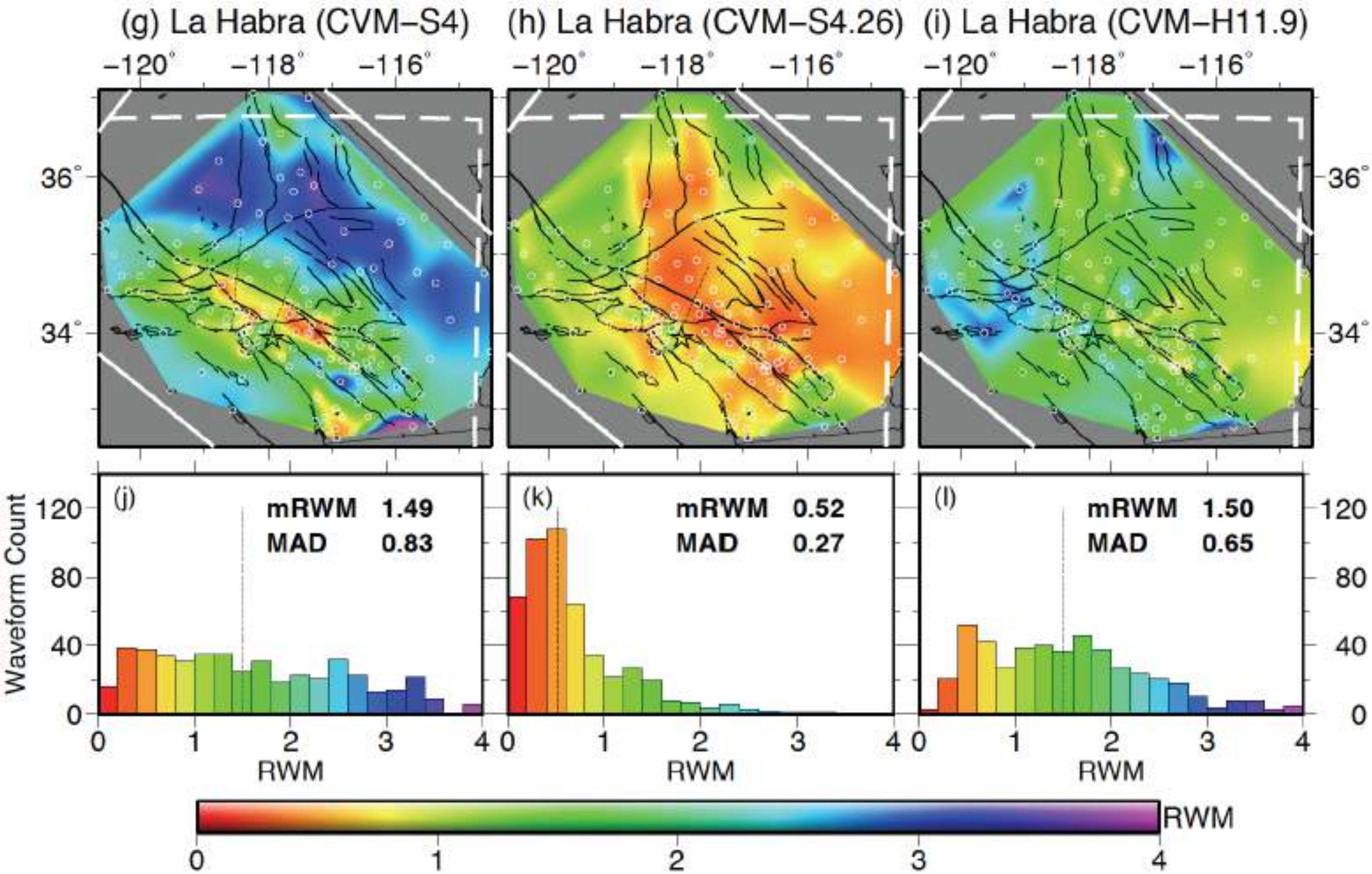




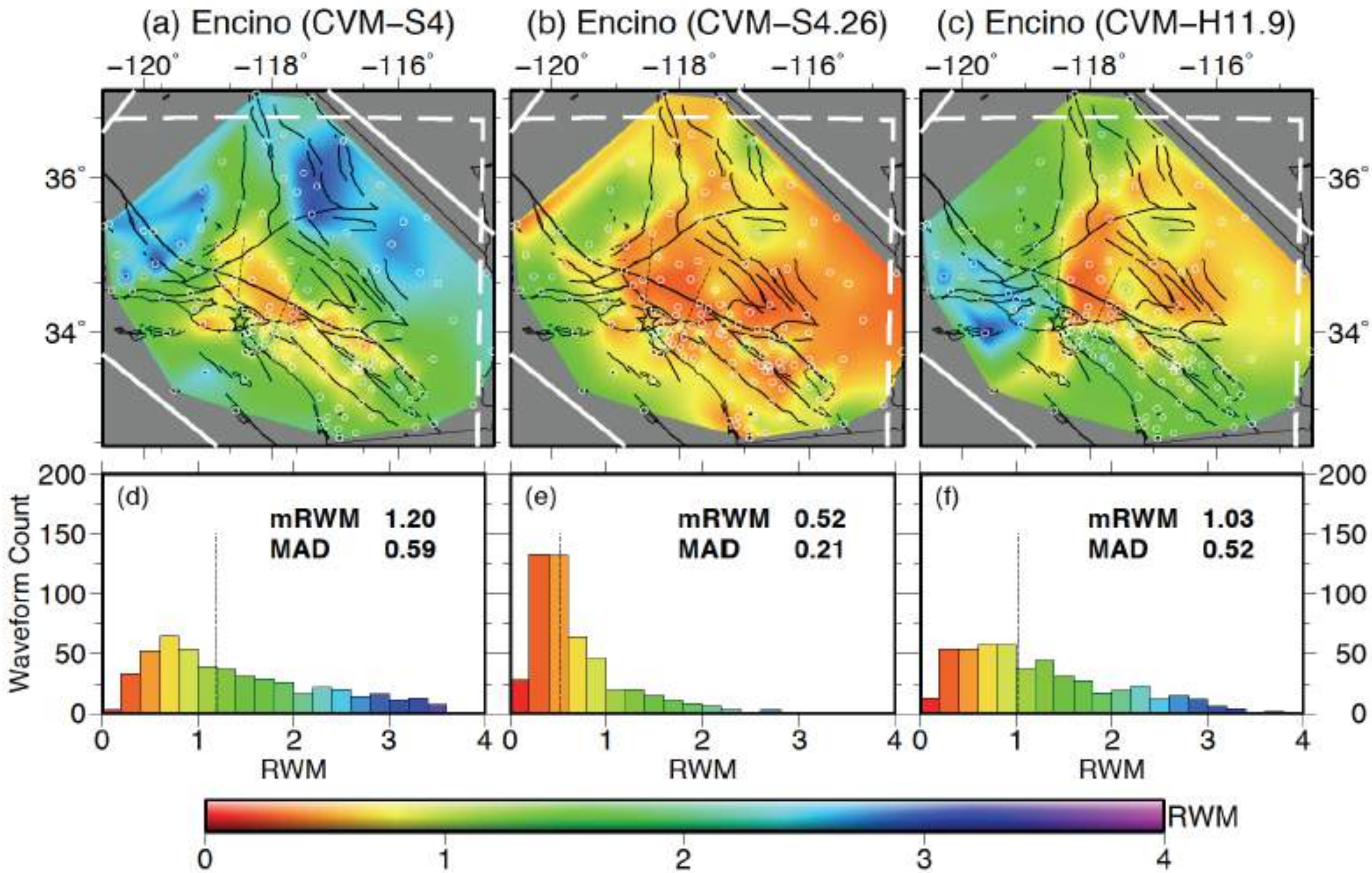




03/28/14 La Habra Earthquake (M5.1)



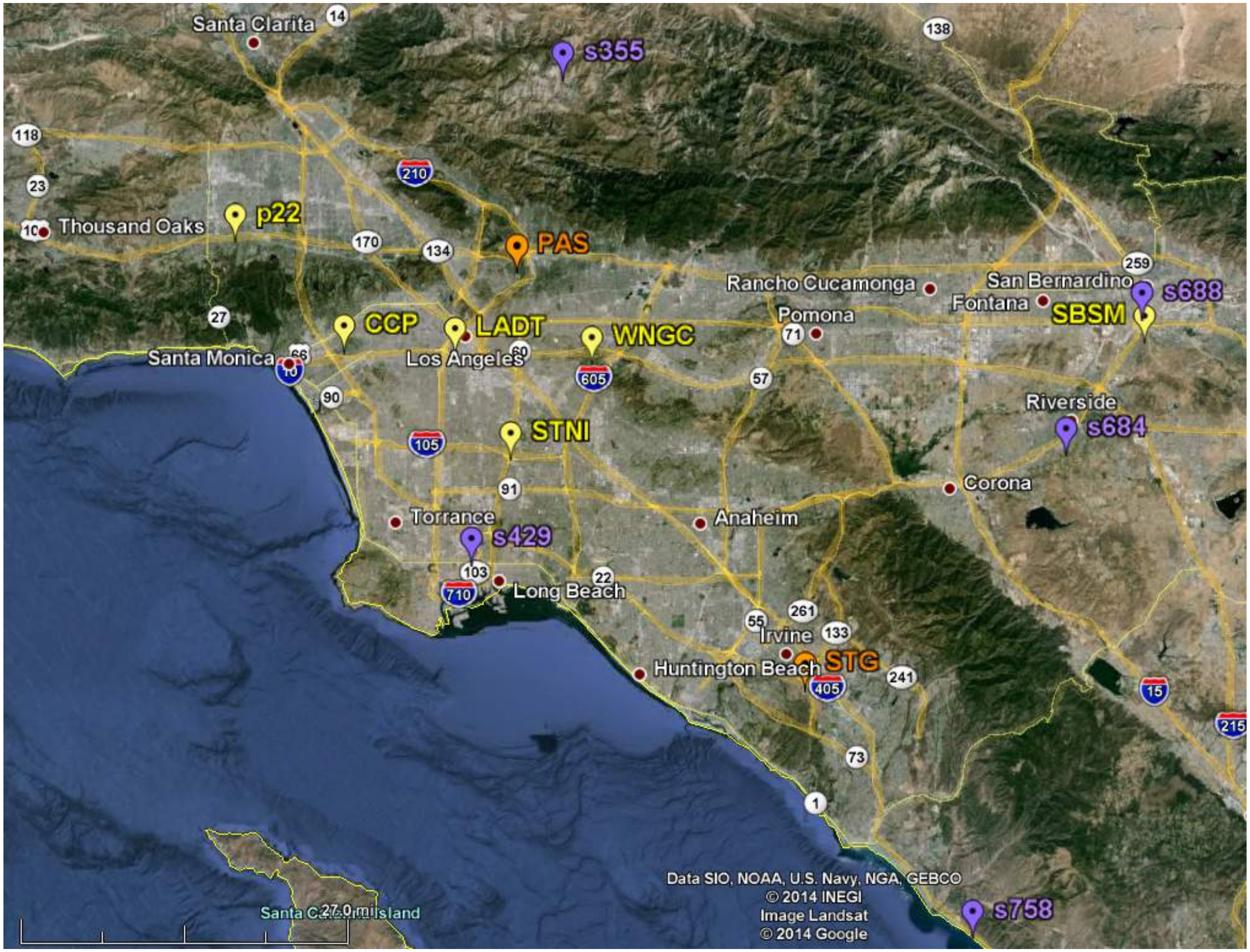
03/17/14 Encino Earthquake (M4.4)



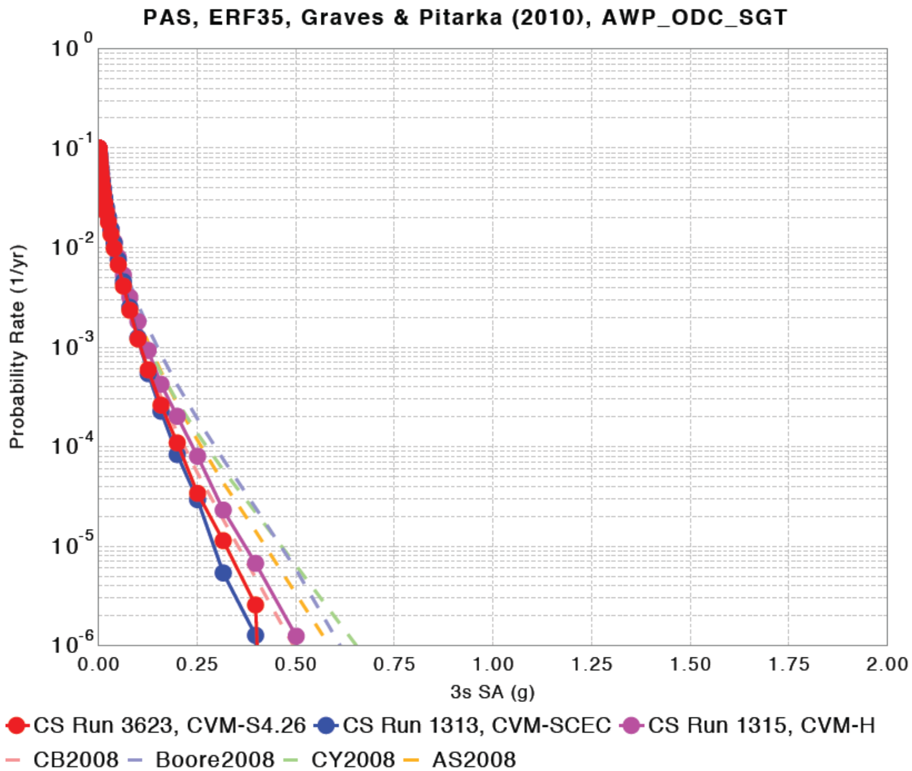
Comparisons Among CyberShake Models and NGA GMPEs

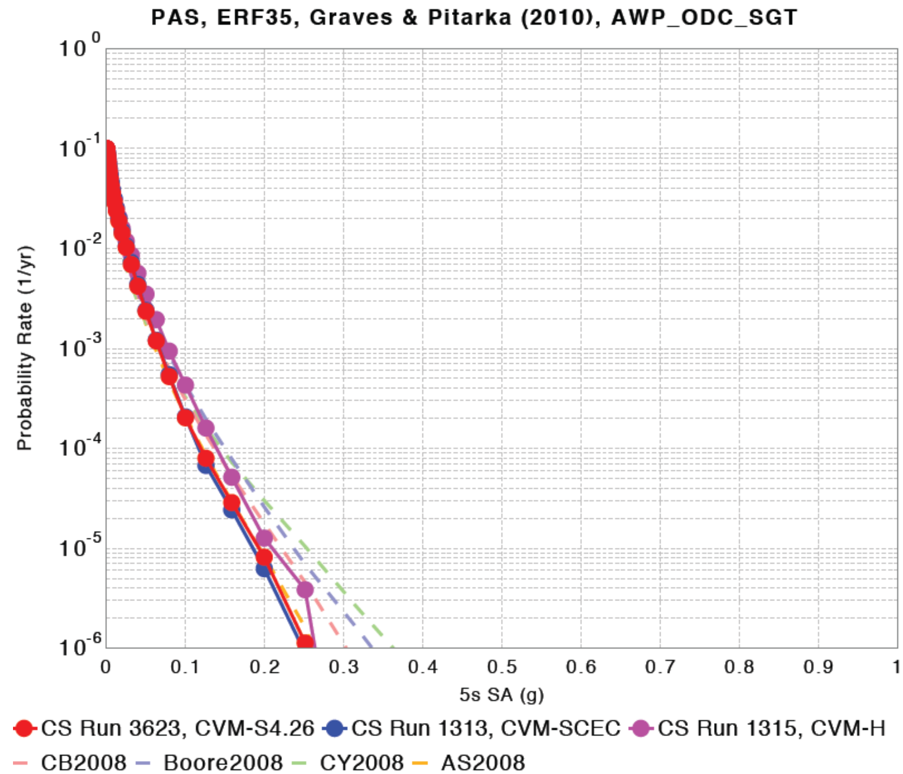
Sites for NGA-CyberShake Comparisons

http://scec.usc.edu/scecpedia/SCEC_UGMS_Committee_Meeting#CyberShake_Hazard_Maps



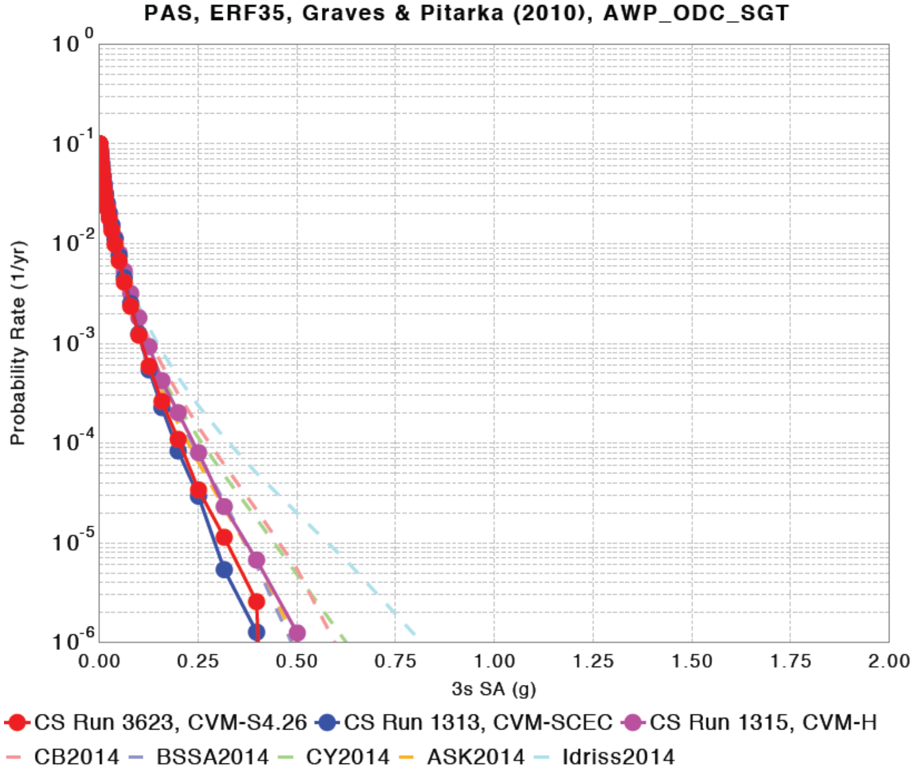
NGA08-CyberShake Comparisons

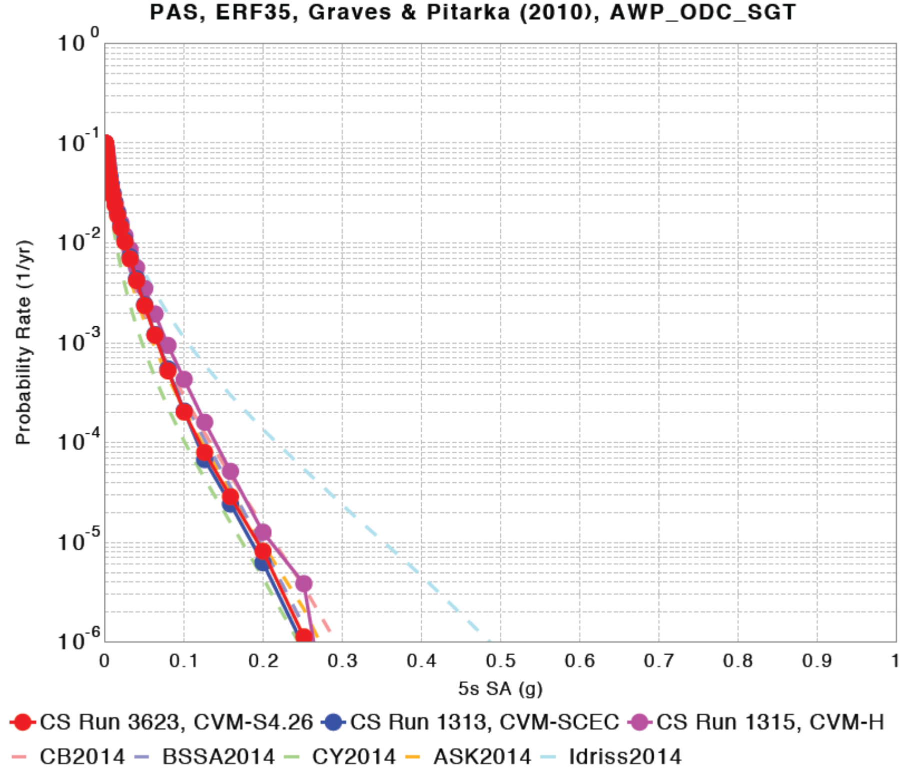




Site PAS

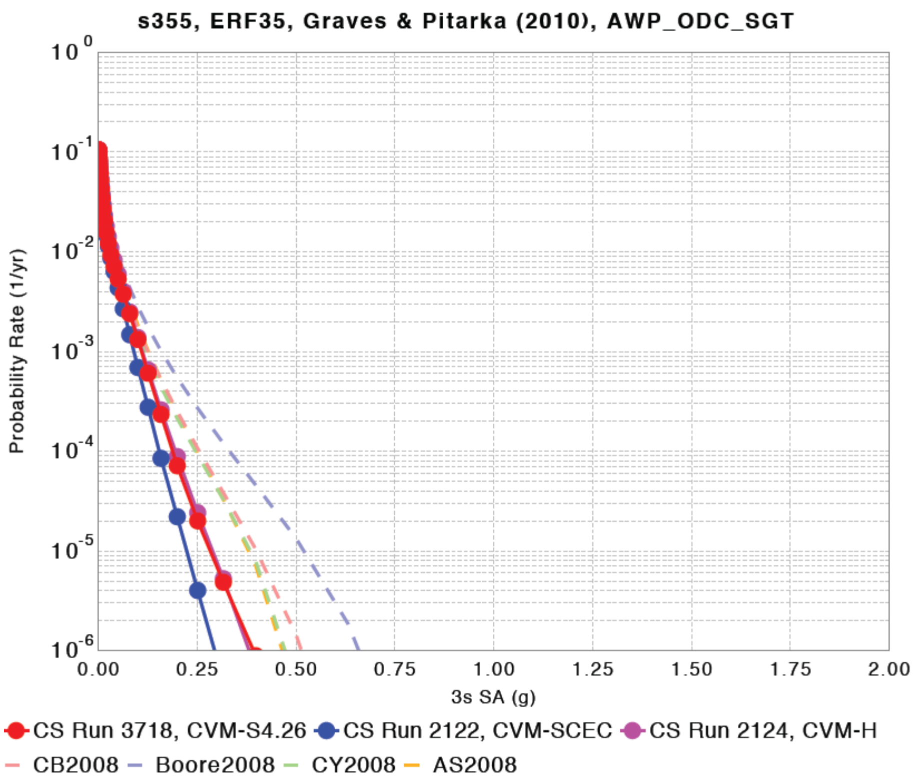
NGA14-CyberShake Comparisons

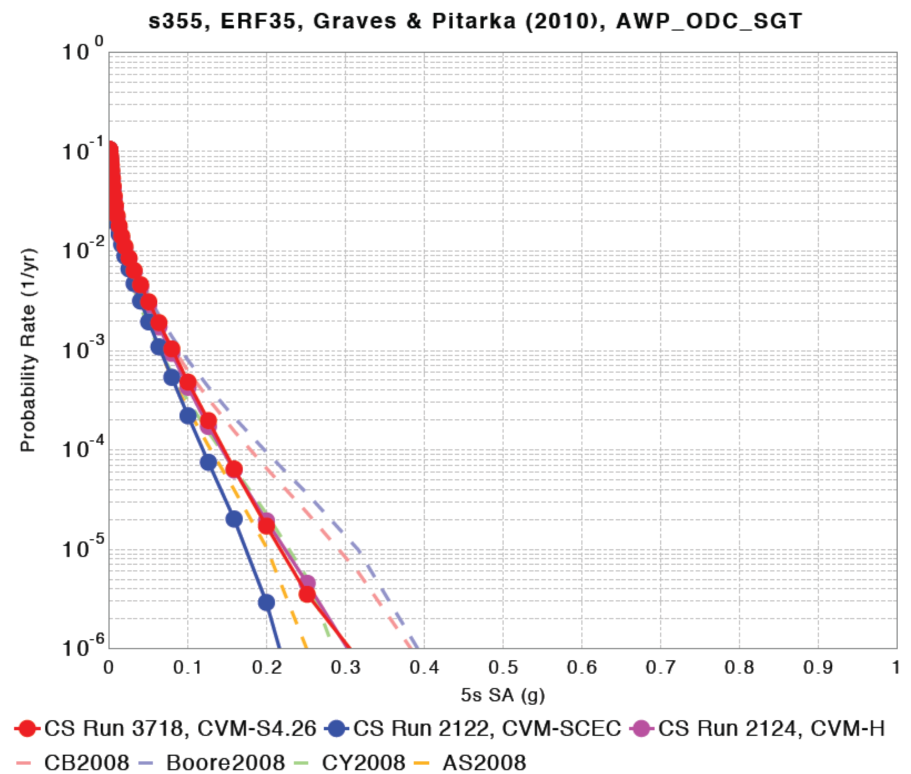




Site PAS

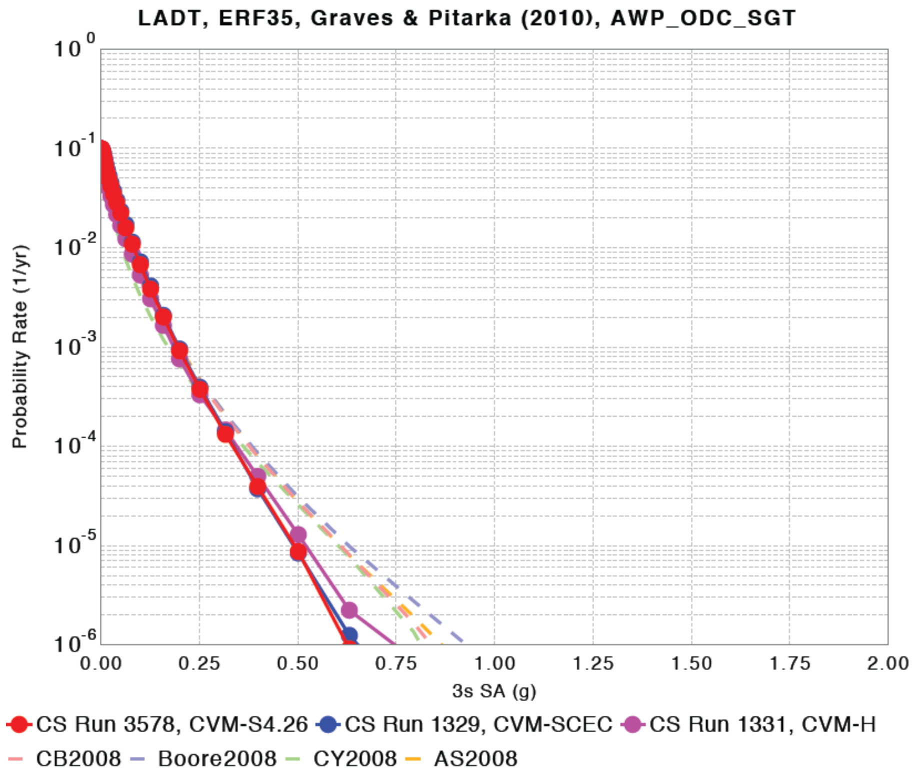
NGA08-CyberShake Comparisons

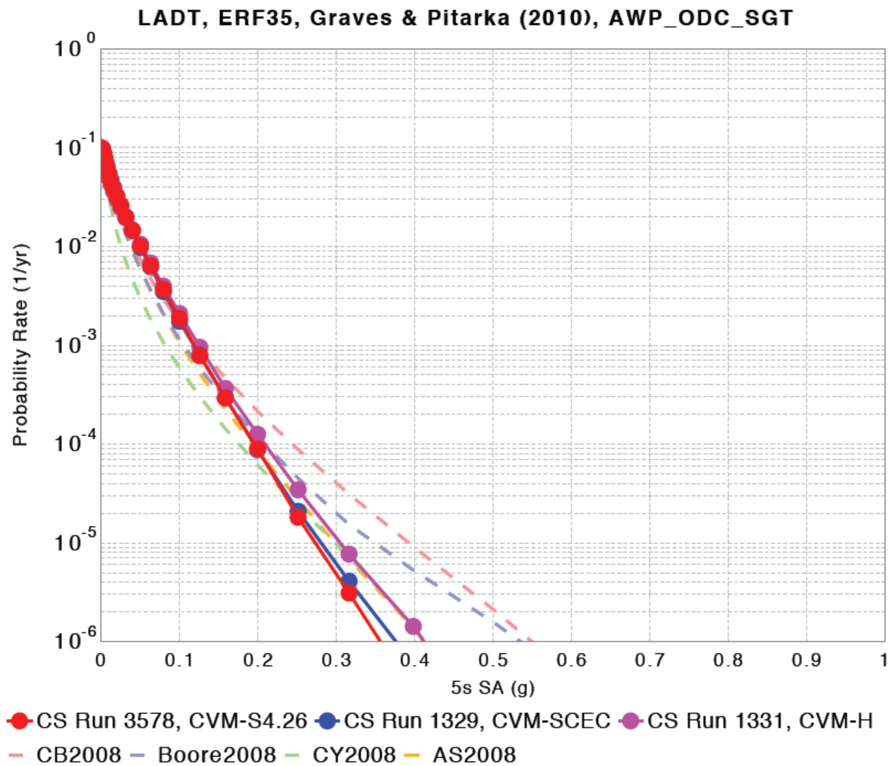




Site s355

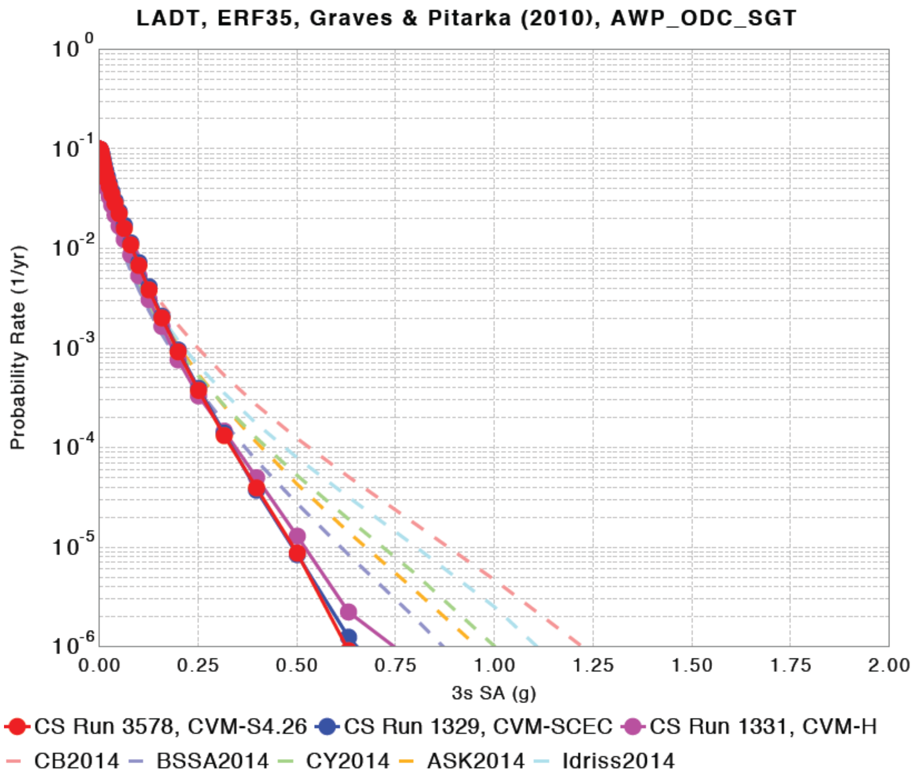
NGA08-CyberShake Comparisons

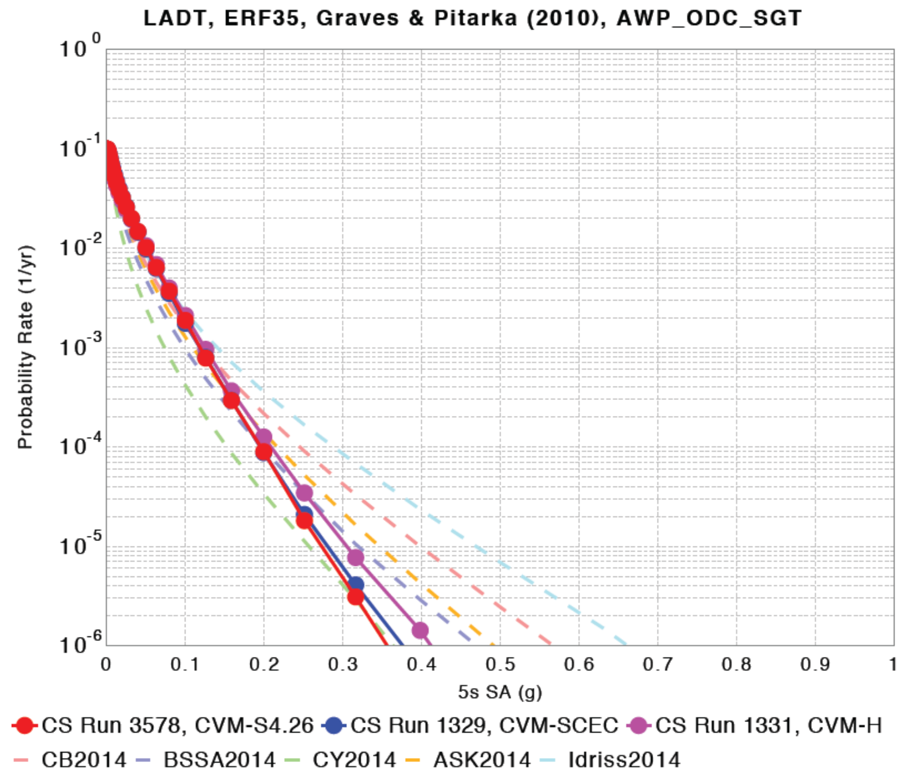




Site LADT

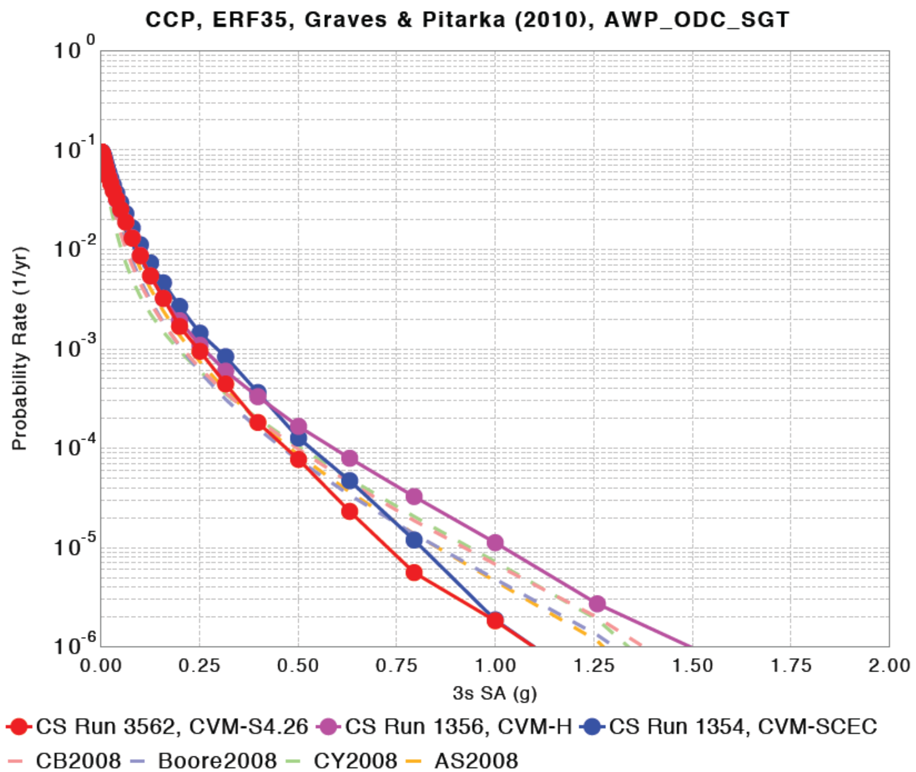
NGA14-CyberShake Comparisons

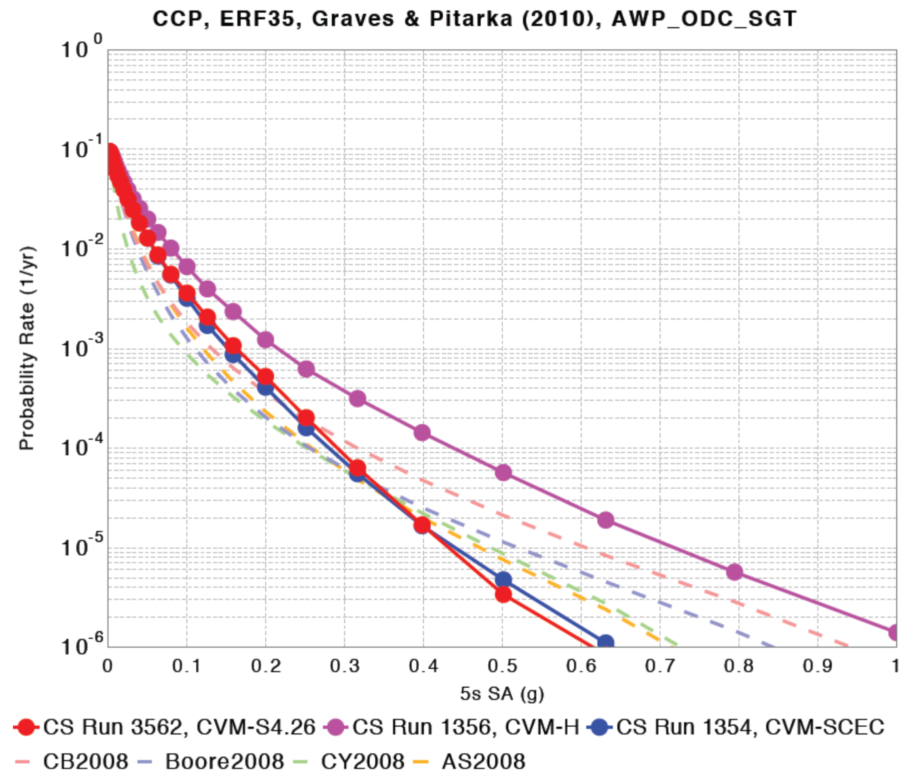




Site LADT

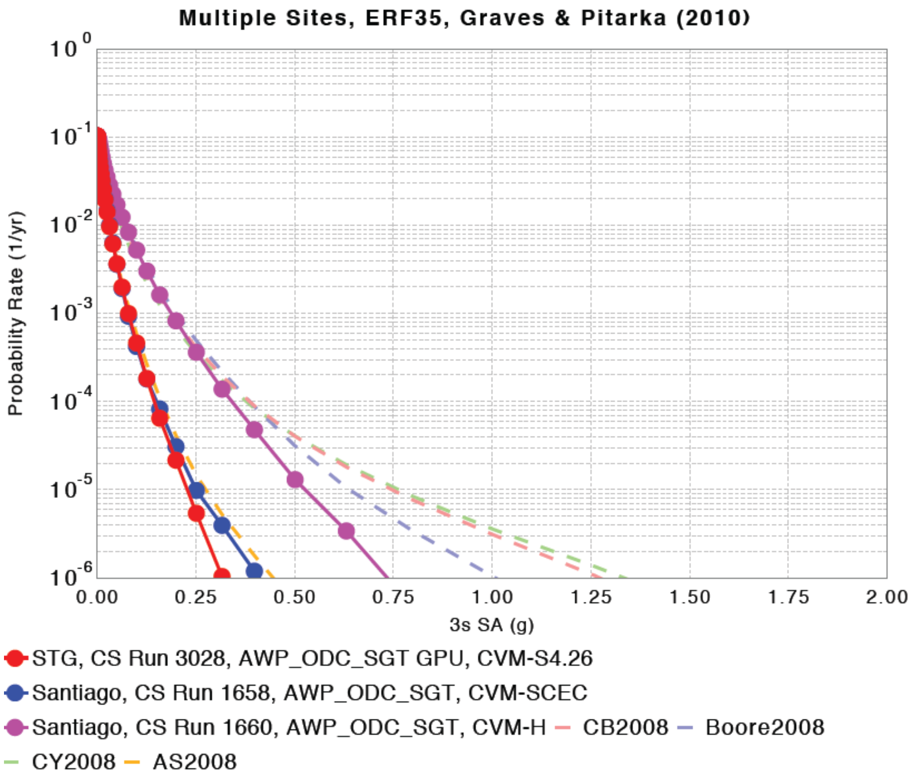
NGA08-CyberShake Comparisons

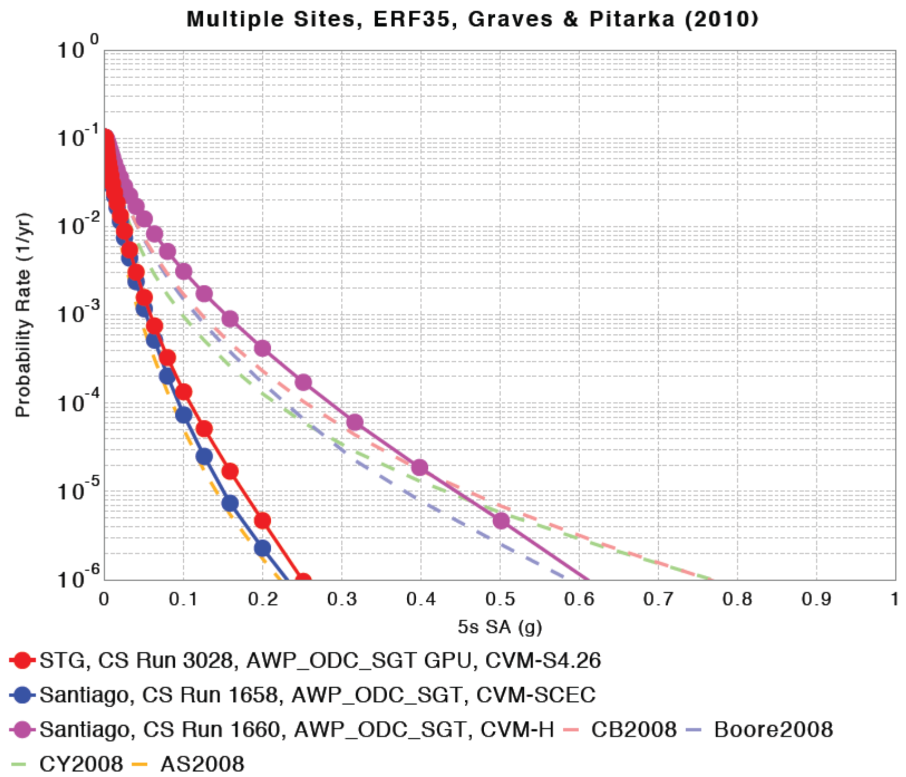




Site CCP

NGA08-CyberShake Comparisons



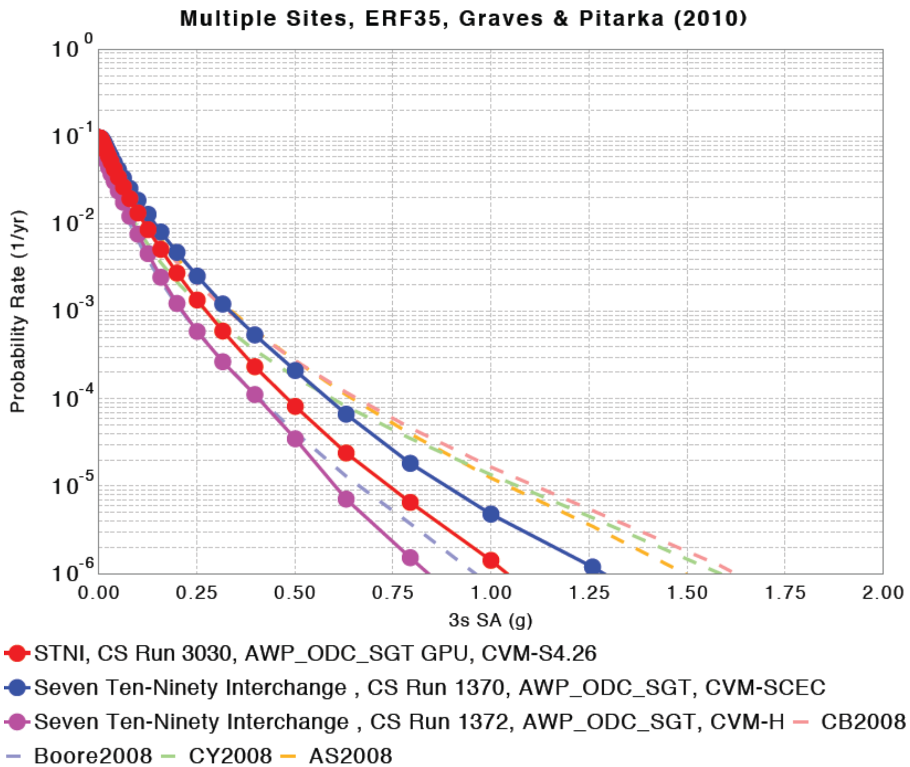


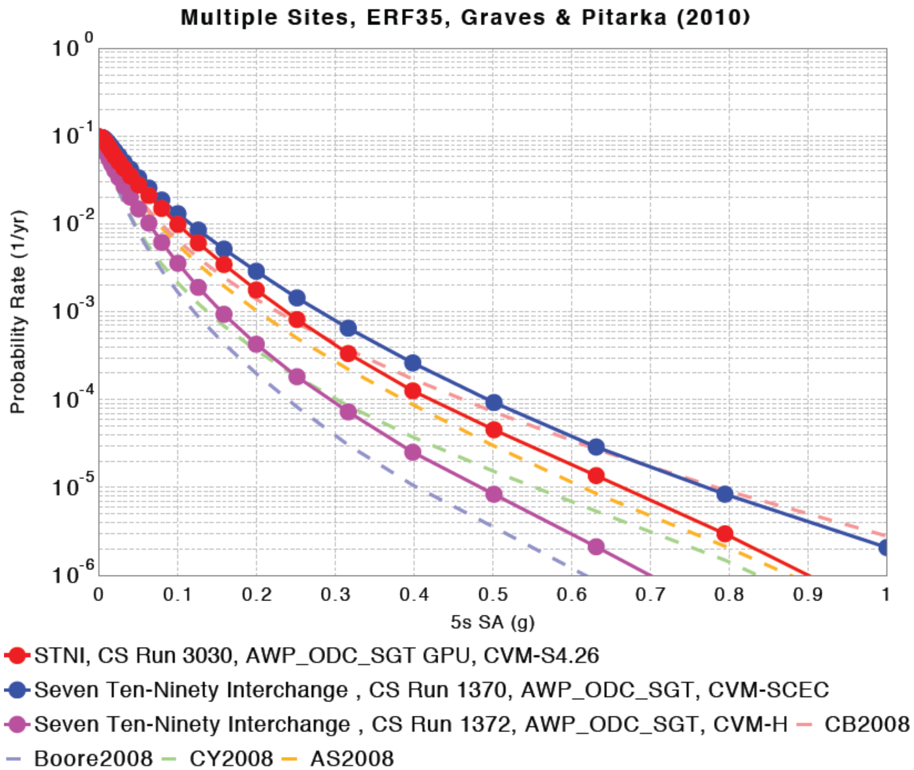
Site SGT

SCEC

Southern California
Earthquake Center

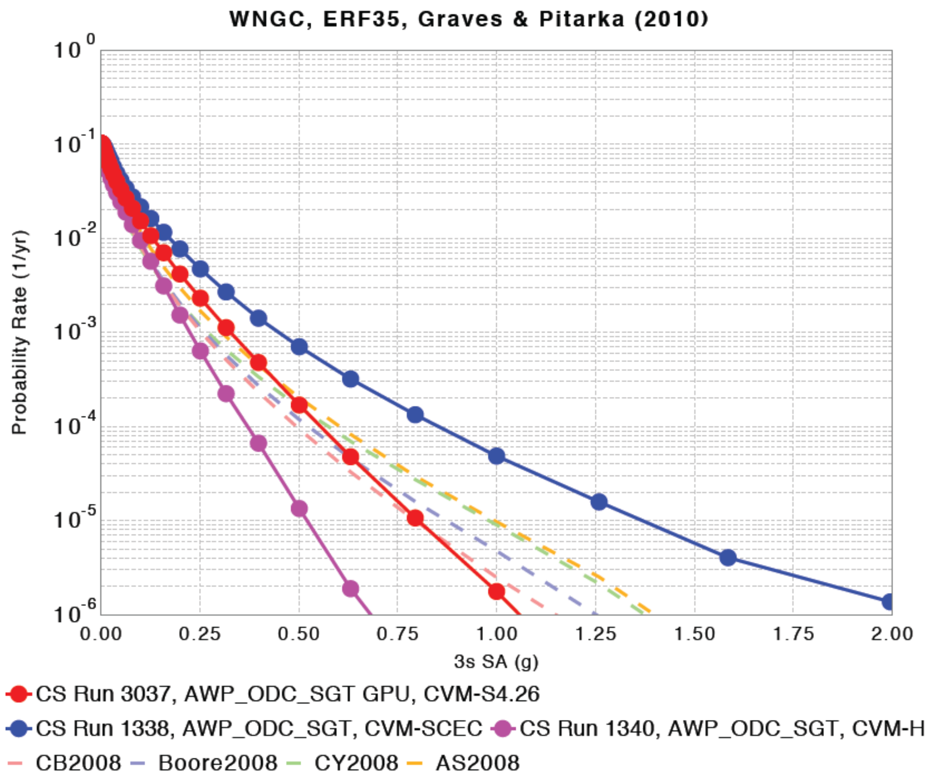
NGA08-CyberShake Comparisons

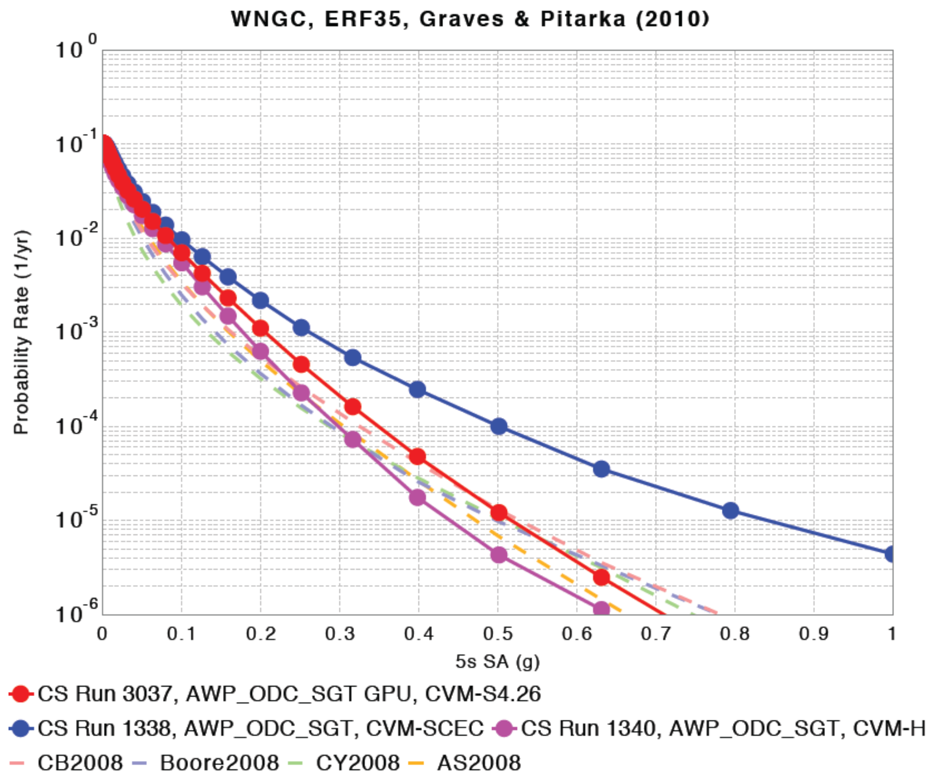




Site STNI

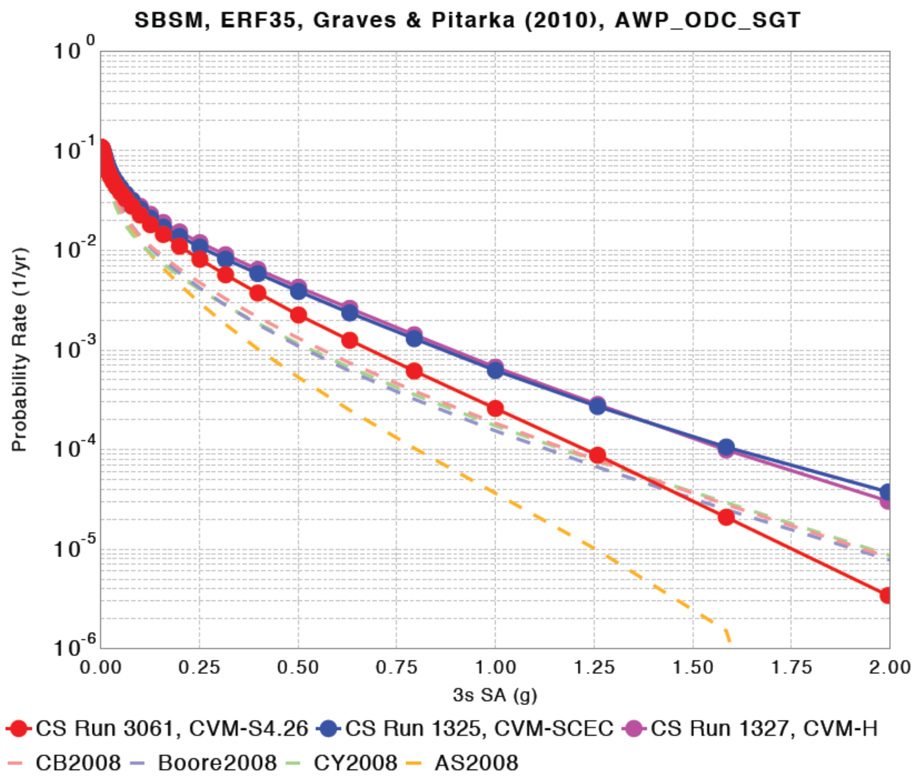
NGA08-CyberShake Comparisons

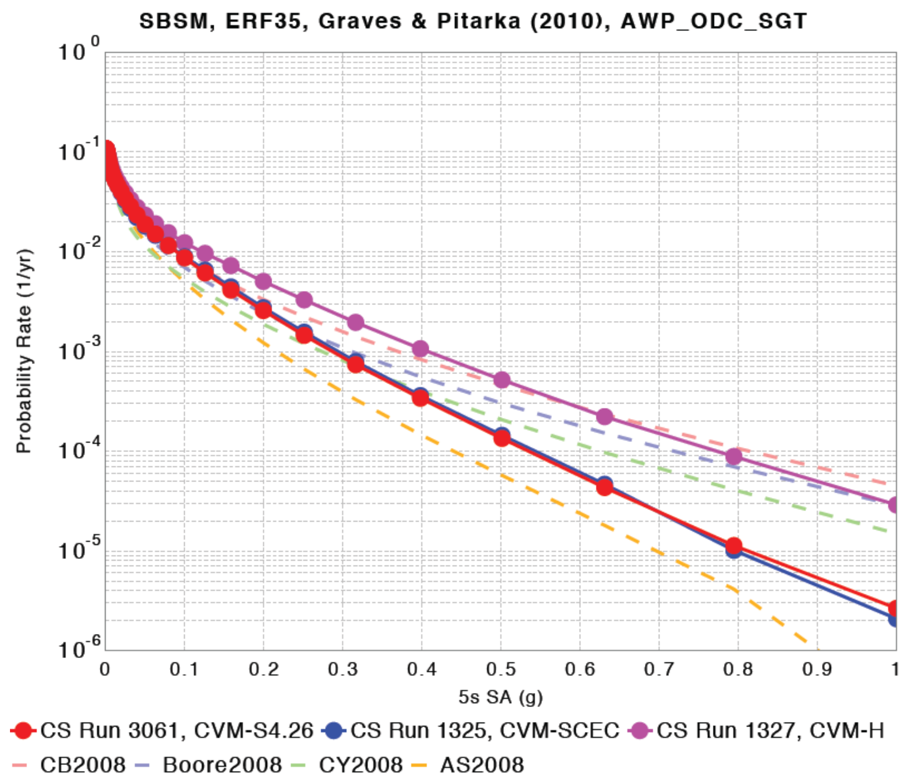




Site WNGC

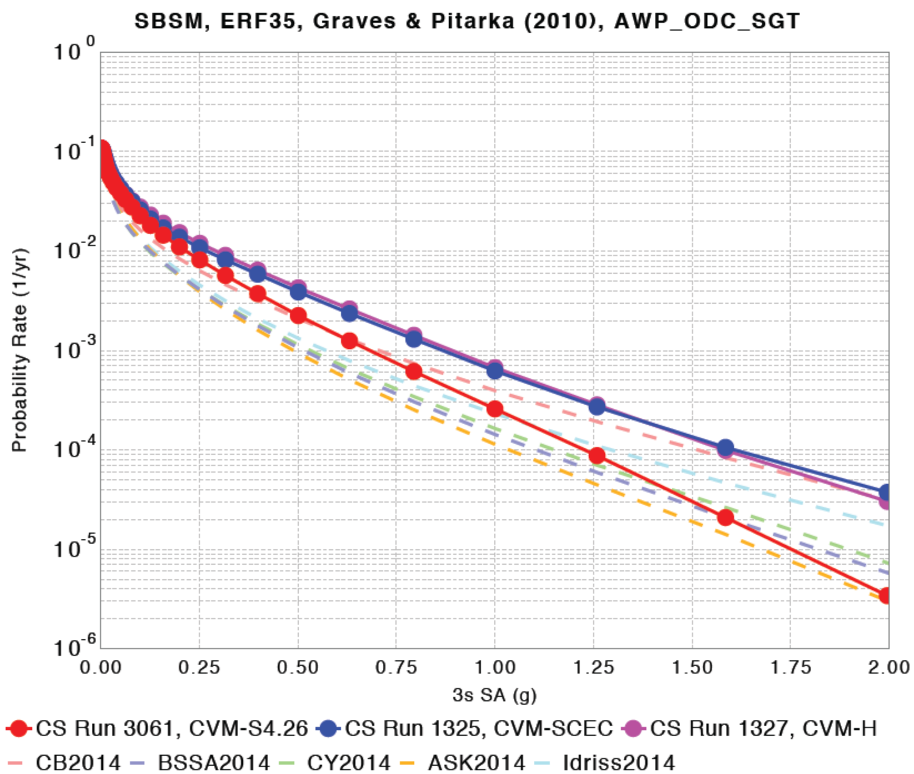
NGA08-CyberShake Comparisons

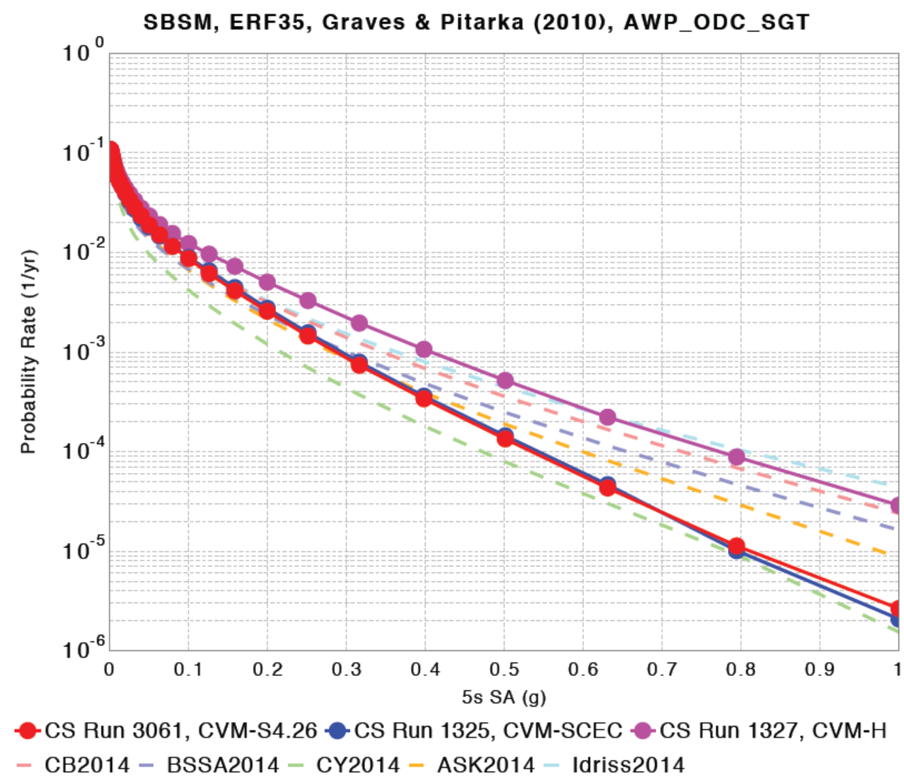




Site SBSM

NGA14-CyberShake Comparisons





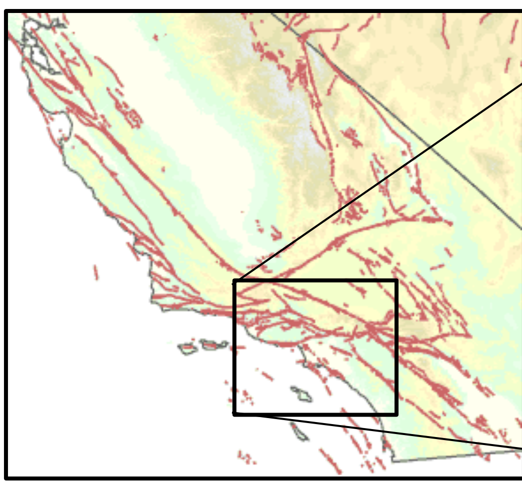
Site SBSM

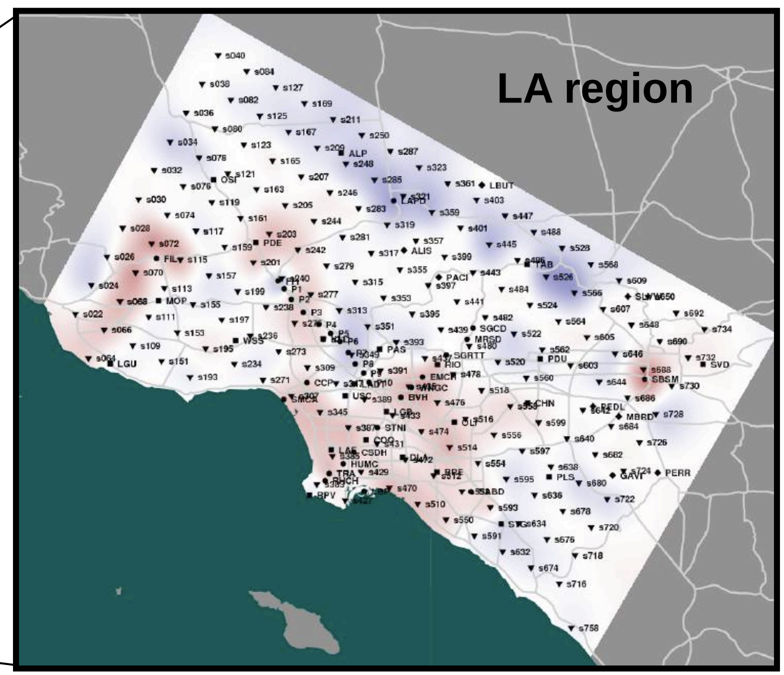
Seismological Hierarchy of CyberShake

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

- **Site set:** $r \in R$
  - 283 sites in the greater Los Angeles region
  - Regional structure specified by CVM-S4 or CVM-H11
- **Rupture set:** $k \in K(r)$
  - All UCERF2 ruptures within 200 km of site
- **Conditional hypocenter distribution:** $x \in X(r, k)$
  - Uniform distribution along fault strike with $\Delta x \approx 20$ km
- **Conditional slip distribution:** $s \in S(r, k, x)$
  - Pseudo-dynamic rupture models of Graves & Pitarka (2007, 2010)

Approximately 415,000 rupture variations per site, 235 million synthetic seismograms per model





Averaging-Based Factorization

- **GMPEs are the multiplication of factors representing attenuation, site effects, directivity effects, etc.**
  - **This model-based factorization is not available for CyberShake**
- **We can compare simulation-derived models with GMPEs using "averaging-based factorization" (Wang & Jordan, *BSSA*, 2014)**
  - **Expected shaking intensities are constructed from a hierarchy of averaging operations over slip variations (s), hypocenters (x), sources (k), and sites (r)**

$$G(r,k,x,s) = A + B(r) + C(r,k) + D(r,k,x) + E(r,k,x,s)$$

- $G(r,k,x,s)$: ln (Y)
- A: level
- $B(r)$: site effect
- $C(r,k)$: attenuation effect
- $D(r,k,x)$: directivity effect
- $E(r,k,x,s)$: slip complexity effect

Averaging-Based Factorization

Define expectation of excitation G with respect to distribution $p(q)$:

$$\langle G \rangle_q = \sum_{q \in Q} p(q)\, G(q) \quad \text{where} \quad \sum_{q \in Q} p(q) = 1$$

Average G over $p(q)$ and subtract:

- conditional slip distribution $p(s|k,x)$: $\quad G(r,k,x,s) - \langle G(r,k,x,s) \rangle_s$
- conditional hypocenter distribution $p(x|k)$: $\quad \langle G(r,k,x,s) \rangle_s - \langle G(r,k,x,s) \rangle_{s,x}$
- rupture distribution $p(k)$: $\quad \langle G(r,k,x,s) \rangle_{s,x} - \langle G(r,k,x,s) \rangle_{s,x,k}$
- site distribution $p(r)$: $\quad \langle G(r,k,x,s) \rangle_{s,x,k,r} \qquad \langle G(r,k,x,s) \rangle_{s,x,k} - \langle G(r,k,x,s) \rangle_{s,x,k,r}$

$$G(r,k,x,s) = A + B(r) + C(r,k) + D(r,k,x) + E(r,k,x,s)$$

- $G(r,k,x,s)$: ln (Y)
- A: level
- $B(r)$: site effect
- $C(r,k)$: attenuation effect
- $D(r,k,x)$: directivity effect
- $E(r,k,x,s)$: slip complexity effect

This averaging-based decomposition is unique and exact

Averaging-Based Factorization

The residuals between excitation functions of an arbitrary target model $G(r, k, x, s)$ and an arbitrary reference model $\tilde{G}(r,k,x,s)$ can be factorized in a similar manner,

$$G(r,k,x,s) - \tilde{G}(r,k,x,s) = g(r,k,x,s)$$
$$= a + b(r) + c(r,k) + d(r,k,x) + e(r,k,x,s)$$

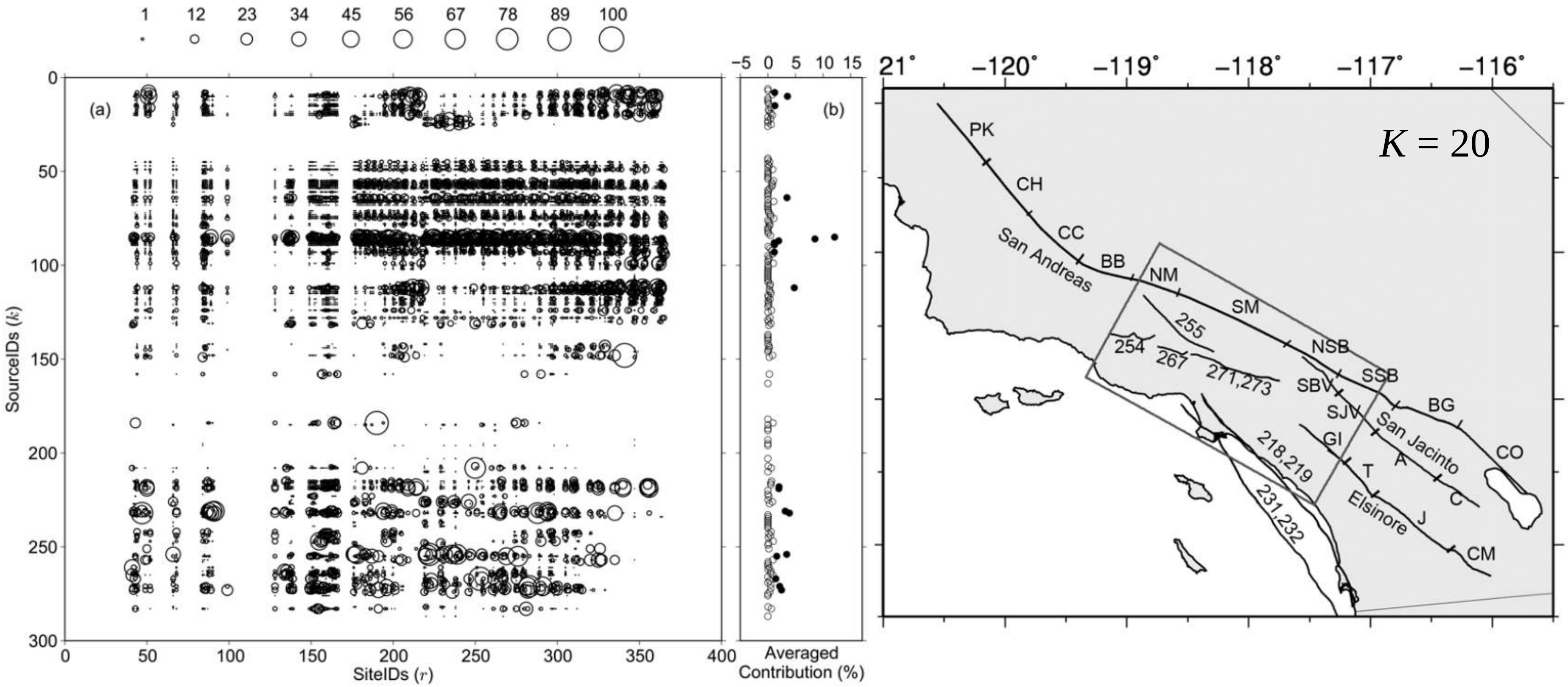
and the individual terms also average to zero:

$$\langle e(r,k,x,s)\rangle_s = \langle d(r,k,x)\rangle_x = \langle c(r,k)\rangle_k = \langle b(r)\rangle_r = 0$$

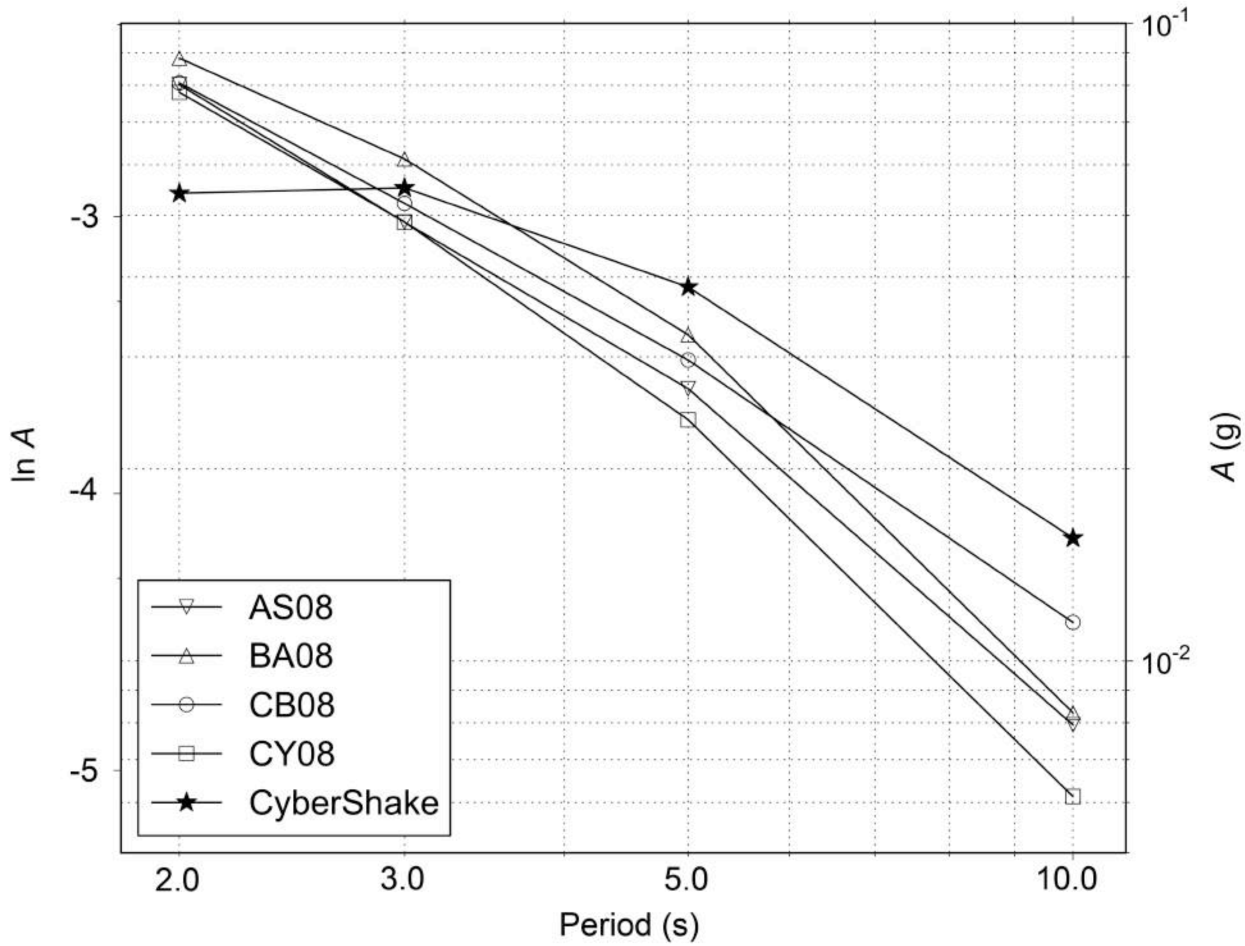
$$G(r,k,x,s) = A + B(r) + C(r,k) + D(r,k,x) + E(r,k,x,s)$$

- $G(r,k,x,s)$: ln (Y)
- A: level
- $B(r)$: site effect
- $C(r,k)$: attenuation effect
- $D(r,k,x)$: directivity effect
- $E(r,k,x,s)$: slip complexity effect

Source Set for CyberShake ABF

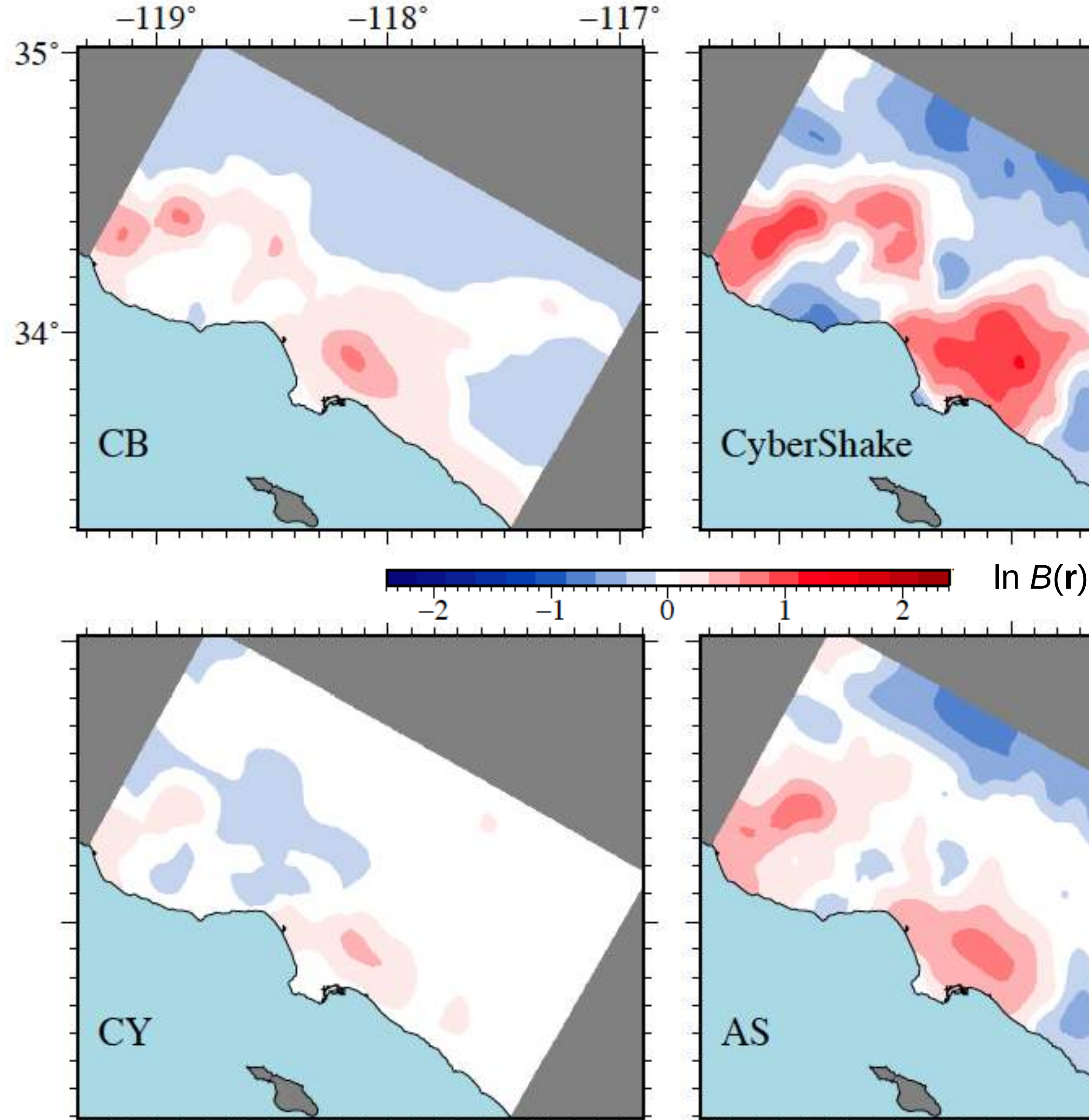


A-values of CyberShake models



ABF Basin Amplification Maps

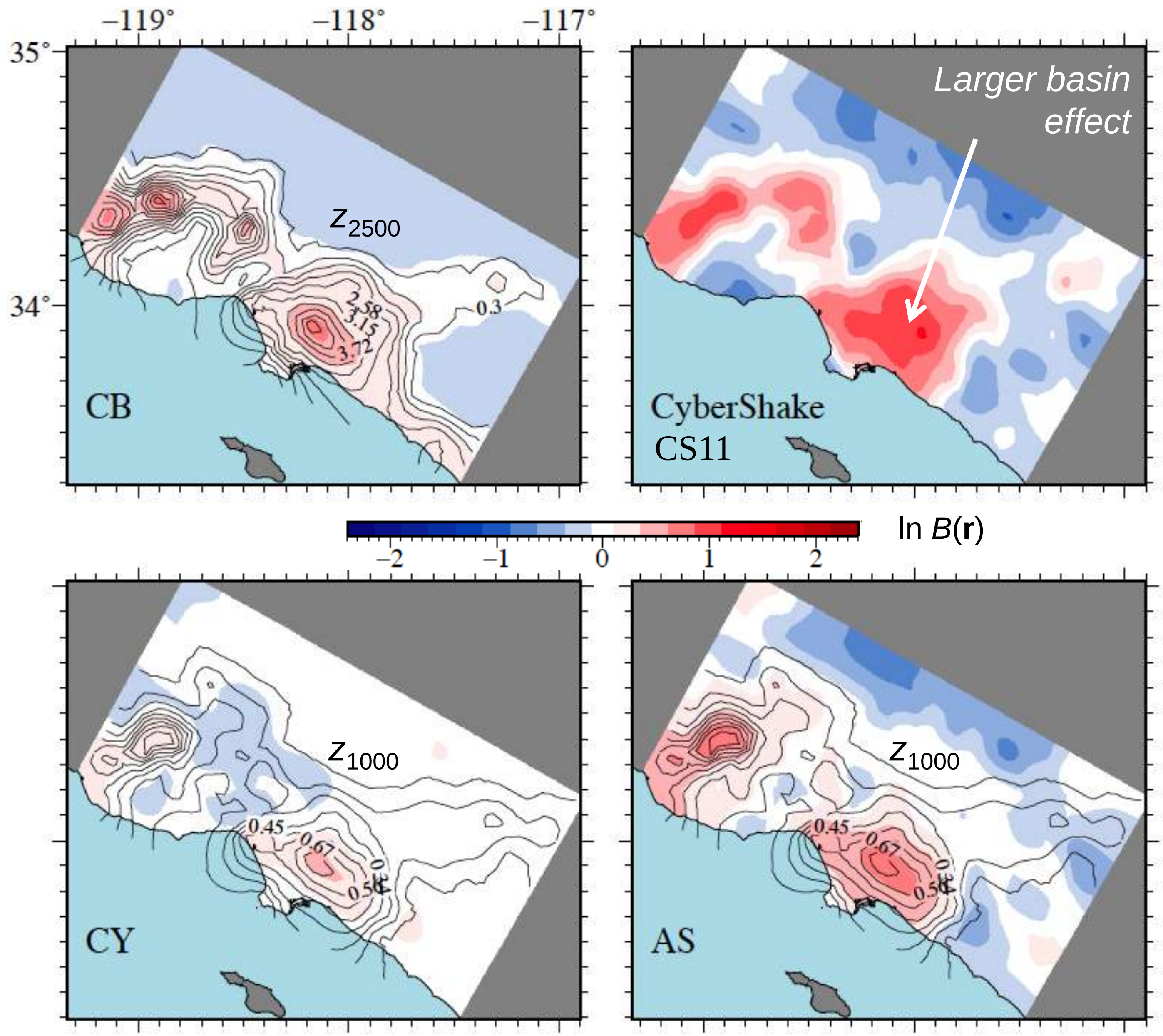
(SA-3s corrected for $V_S 30$ using BA08)



Wang & Jordan (2014)

ABF Basin Amplification Maps

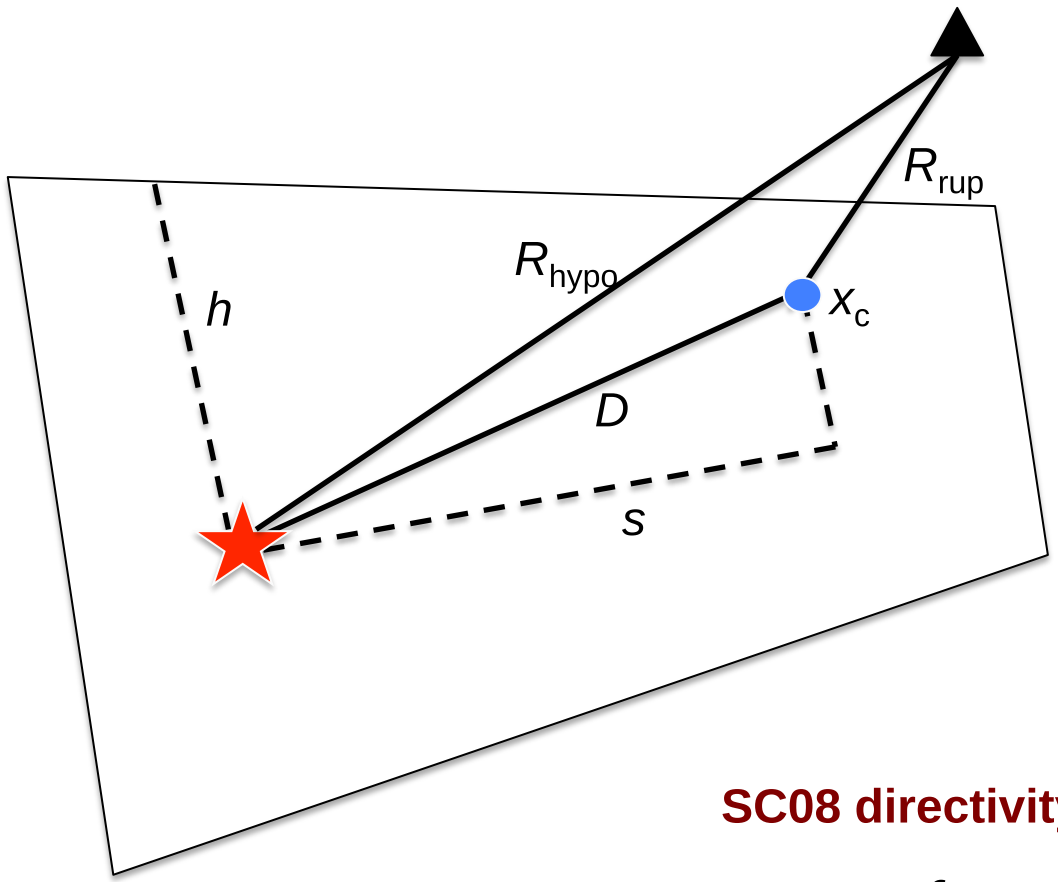
(SA-3s corrected for V_S30 using BA08)



Wang & Jordan (2014)

Empirical Directivity Modeling

Isochrone directivity predictor (IDP) of Spudich & Chiou (2008)



$$IDP = R_{ri}\frac{\tilde{c}}{\beta}\ln\left(\max\left(s,h\right)\right)$$

$$\tilde{c} = \frac{D}{T_c - T_{hypo}}$$ isochrone velocity

$$T_c = \left(\frac{D}{V_r} + \frac{R_{rup}}{\beta}\right); T_{hypo} = \frac{R_{hypo}}{\beta}$$

$$V_r = 0.8\beta$$

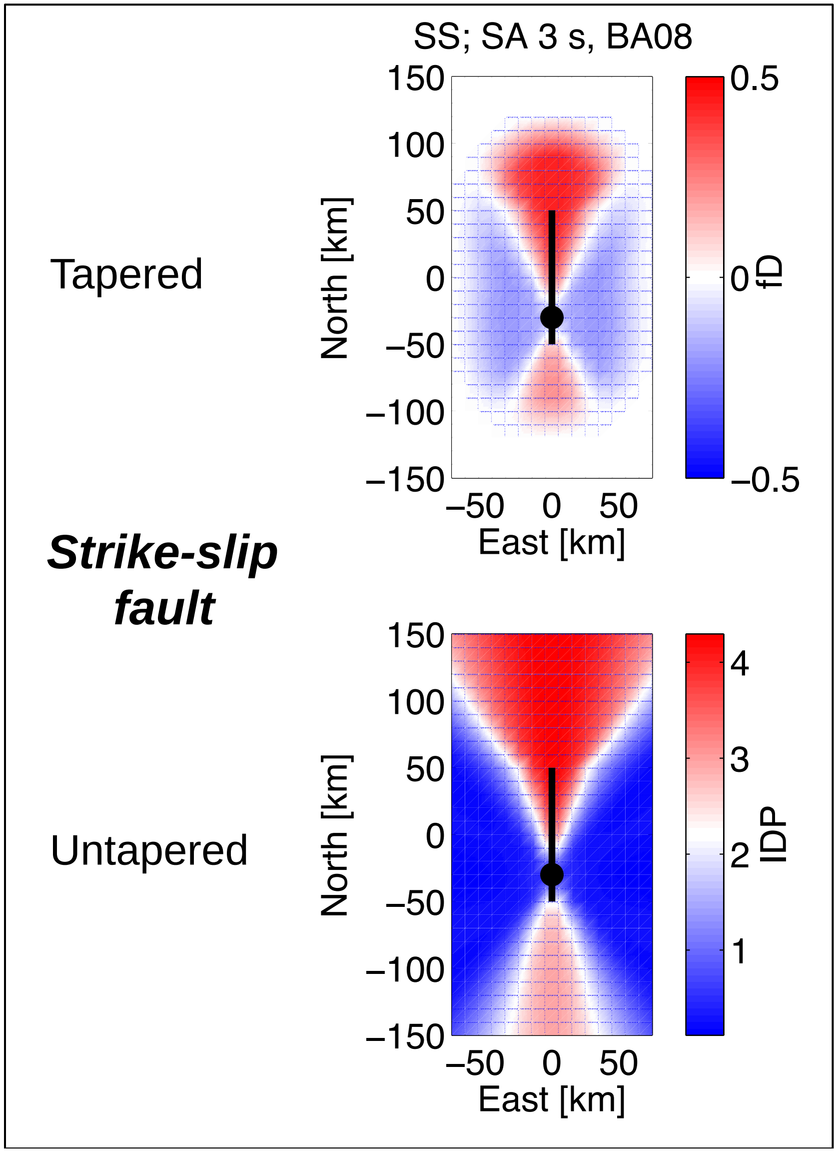
SC08 directivity correction to the NGA GMPEs:

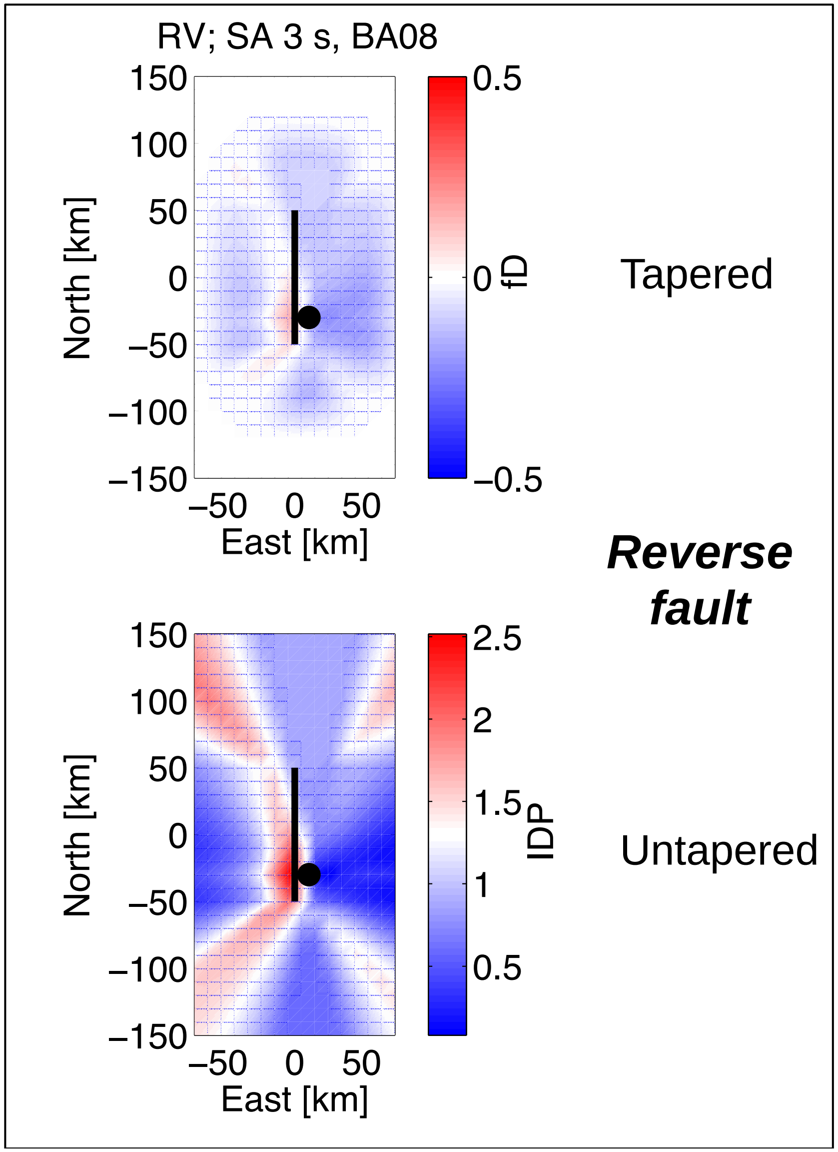
$$f_D = Taper\left(M, R_{rup}\right)\left(a + b * IDP\right)$$

Southern California Earthquake Center

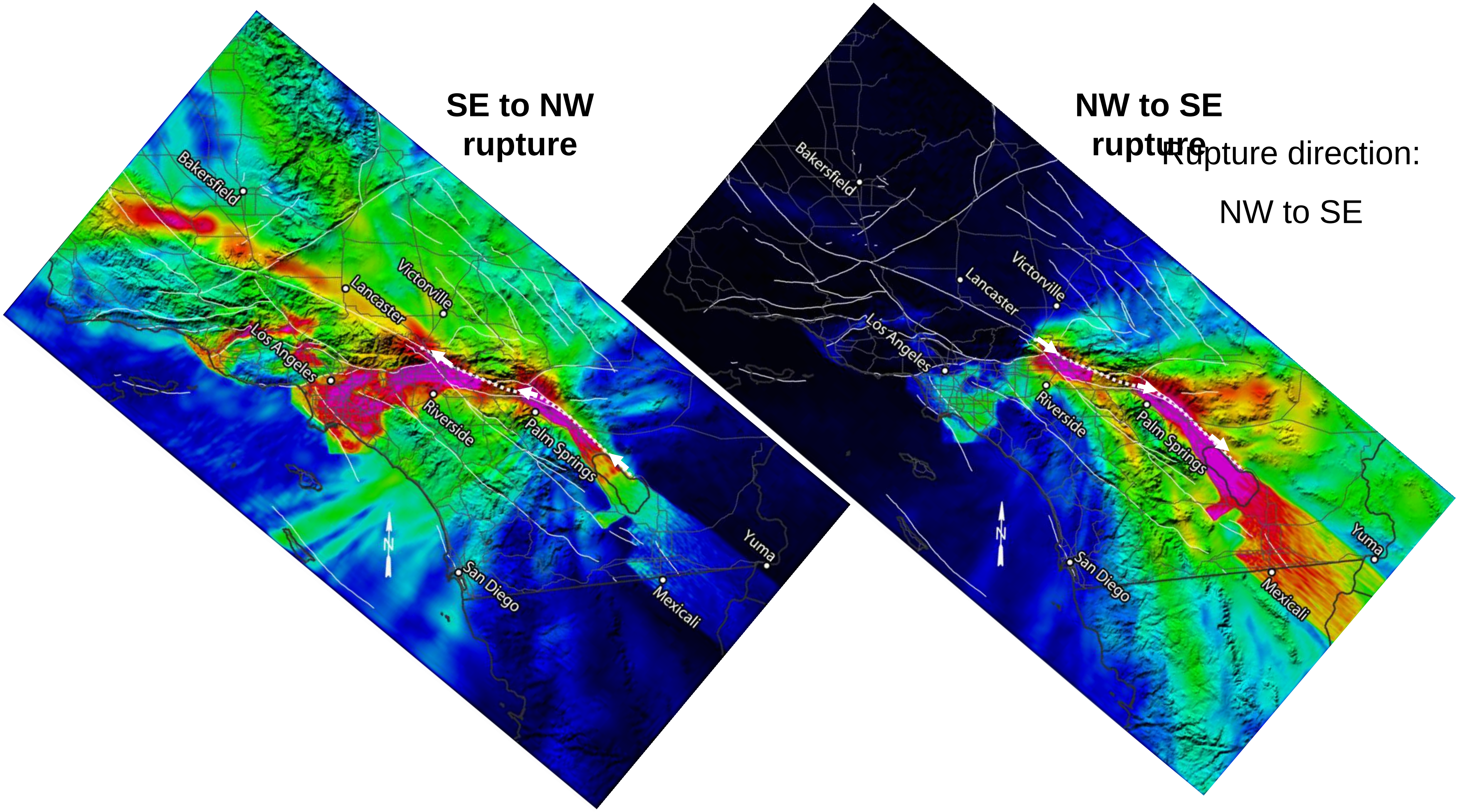
Empirical Directivity Modeling

Isochrone directivity predictor (IDP) of Spudich & Chiou (2008)





Coupling of Directivity and Basin Effects

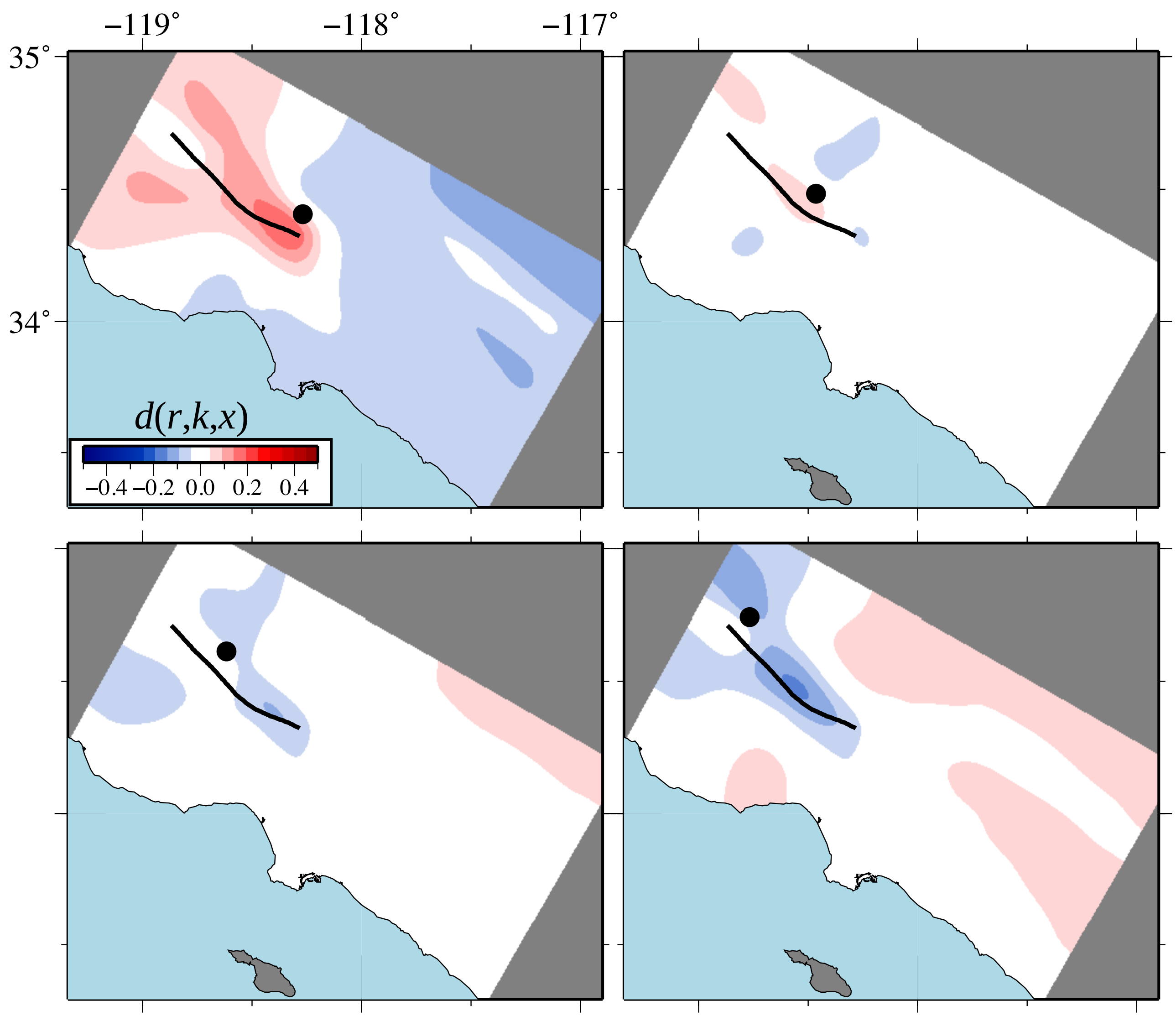


TeraShake simulations of M7.7 earthquake on Southernmost San Andreas (Olsen et al. 2006)

ABF Recovery of SC08 Directivity Correction

Target model: BA08 with SC08 directivity correction

Reference model: BA08 without SC08 directivity correction

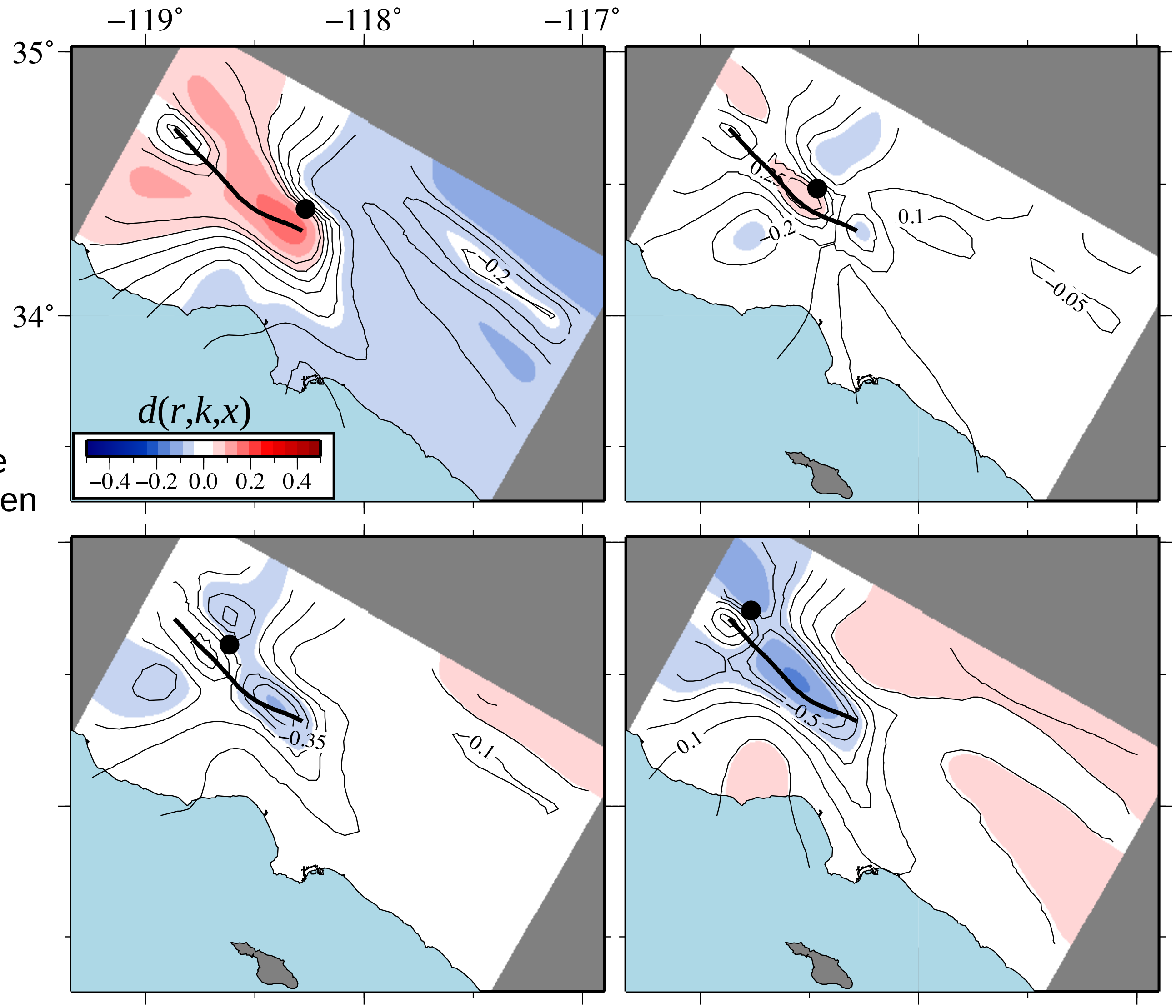


Wang & Jordan **(**2014**)**

ABF Recovery of SC08 Directivity Correction

Target model: BA08 with SC08 directivity correction

Reference model: BA08 without SC08 directivity correction



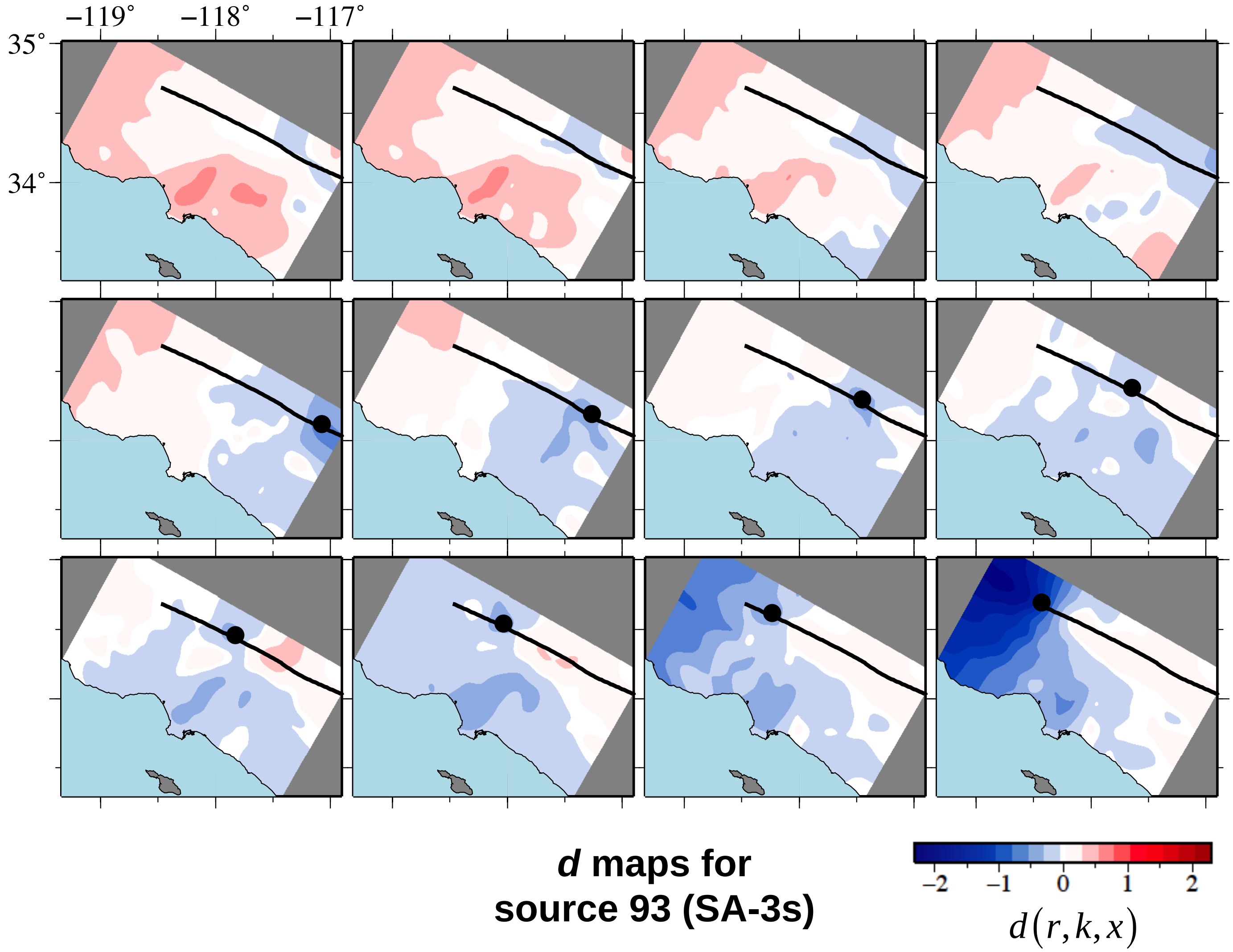
Contour lines are normalized f_D given by SC08 model

Wang & Jordan (2014)

ABF Directivity Comparison: CS11 vs. SC08

Target model: CS11 (GenSlip v2.1)

Reference model: BA08 with SC08 directivity correction



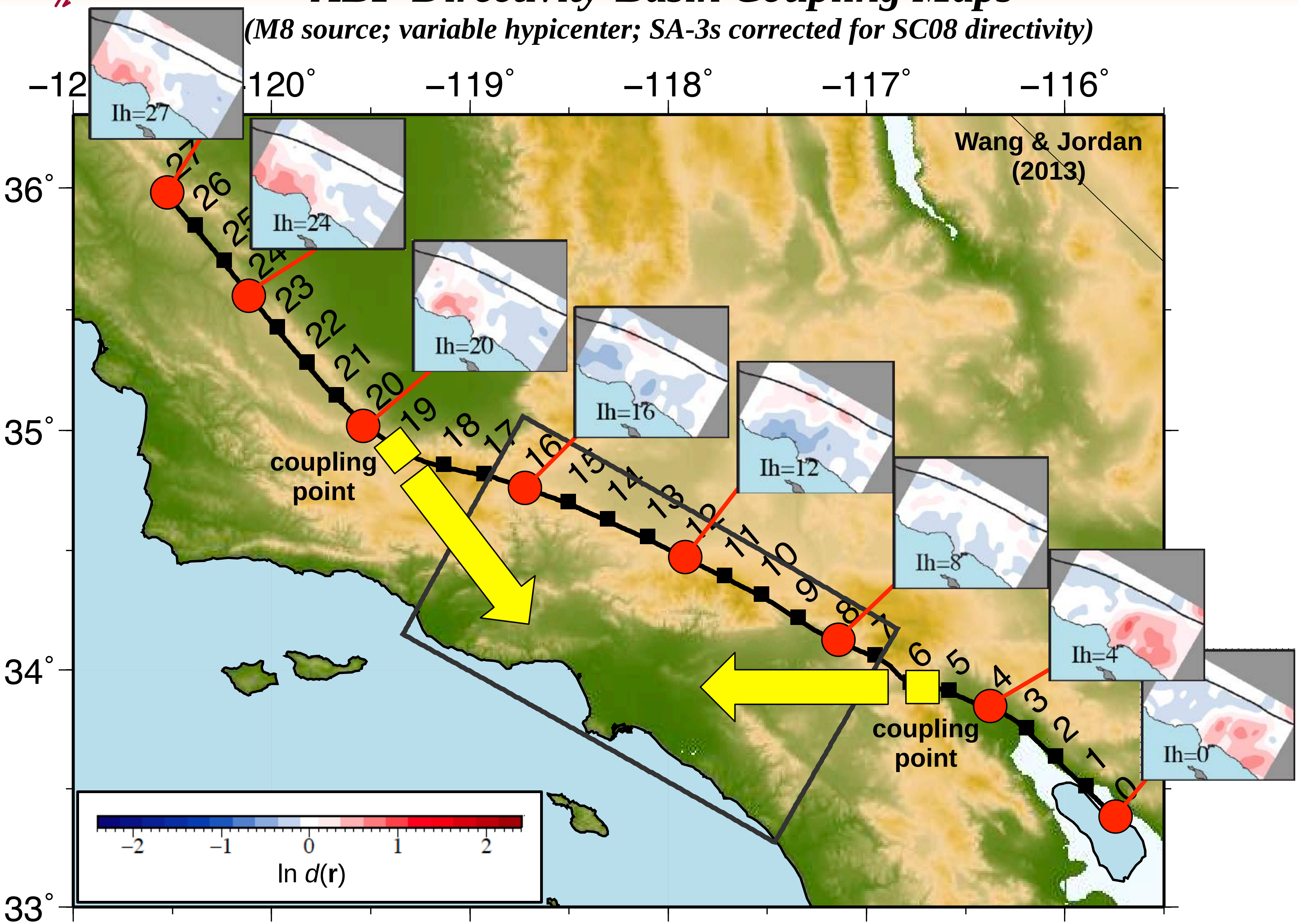
d maps for source 93 (SA-3s)

SCEC

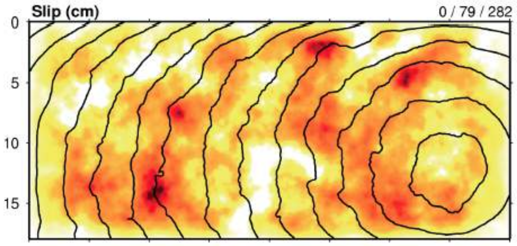
Southern California Earthquake Center

ABF Directivity-Basin Coupling Maps

(M8 source; variable hypicenter; SA-3s corrected for SC08 directivity)



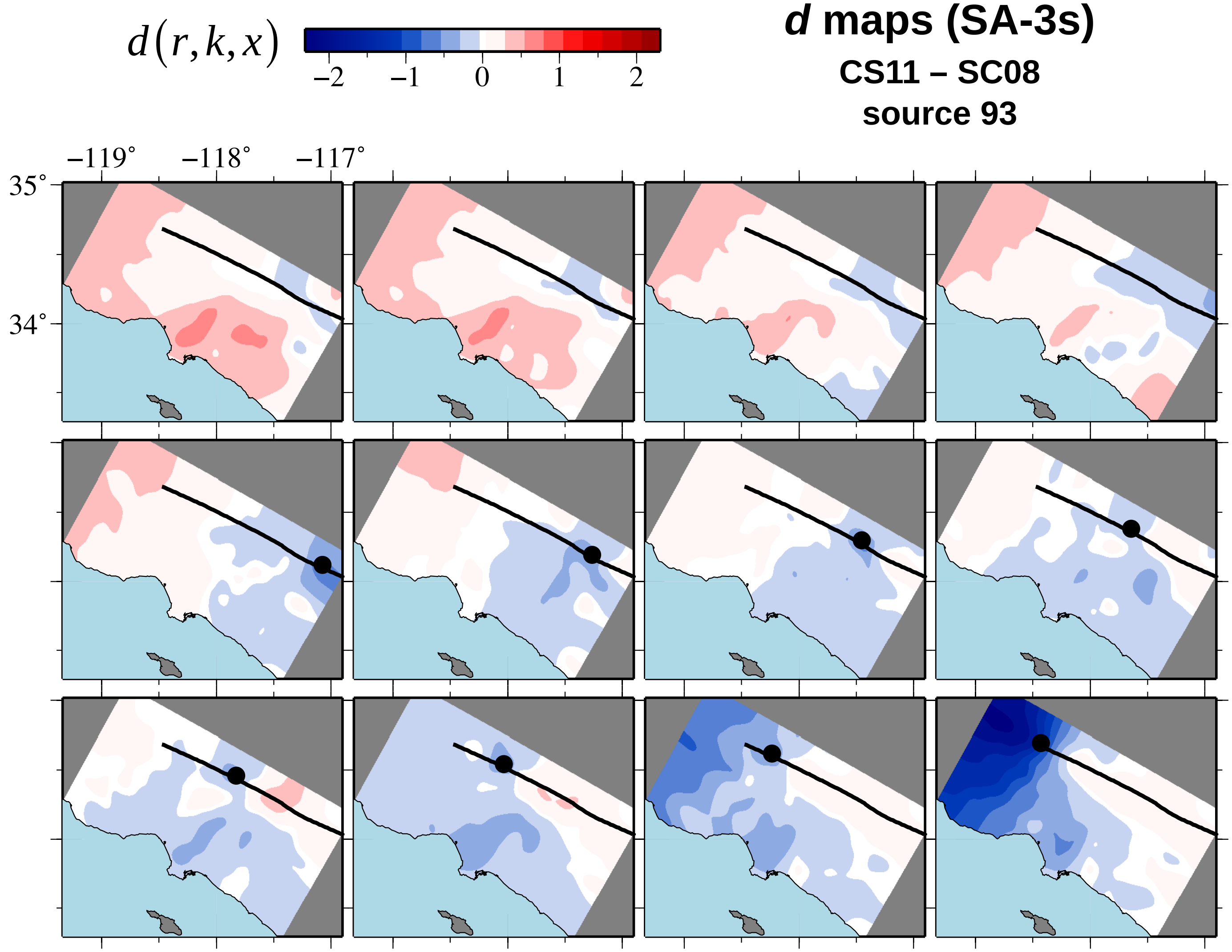
Dependence of Directivity Effects on Rupture Complexity



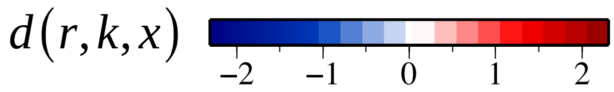
GP07
used in CS11

d maps (SA-3s)

CS11 – SC08
source 93



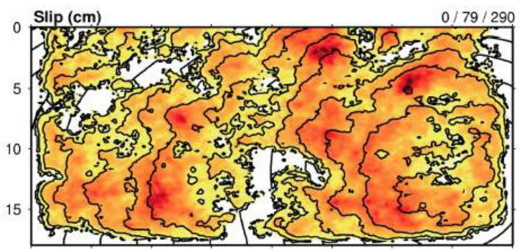
Dependence of Directivity Effects on Rupture Complexity



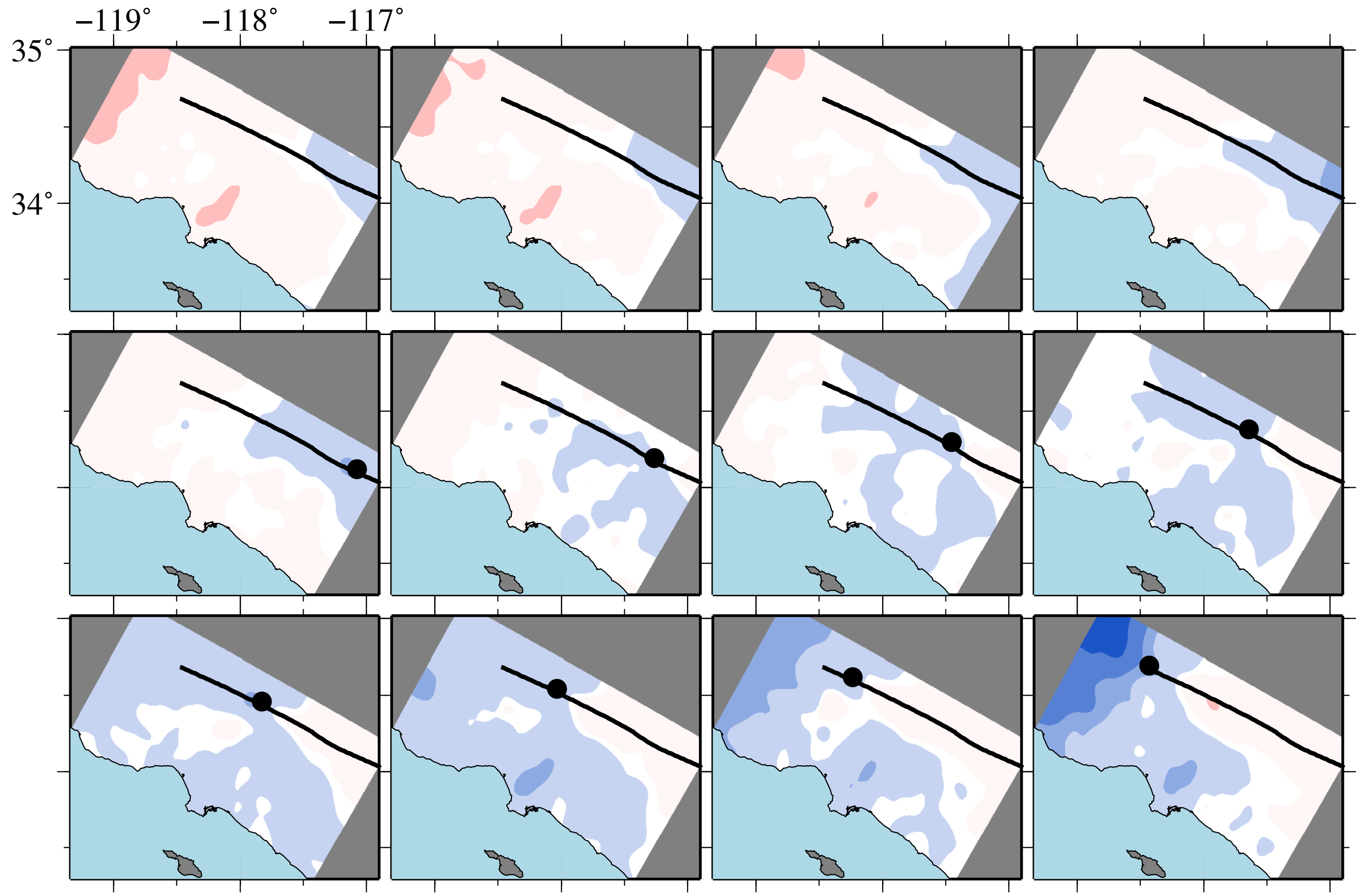
d maps (SA-3s)

CS13a – SC08

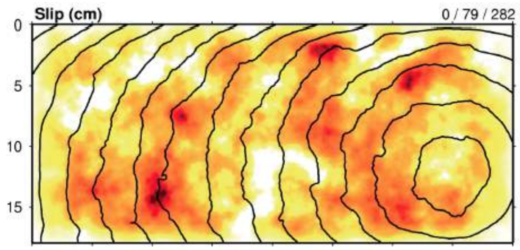
source 93



GP10
used in CS13a



Dependence of Directivity Effects on Rupture Complexity

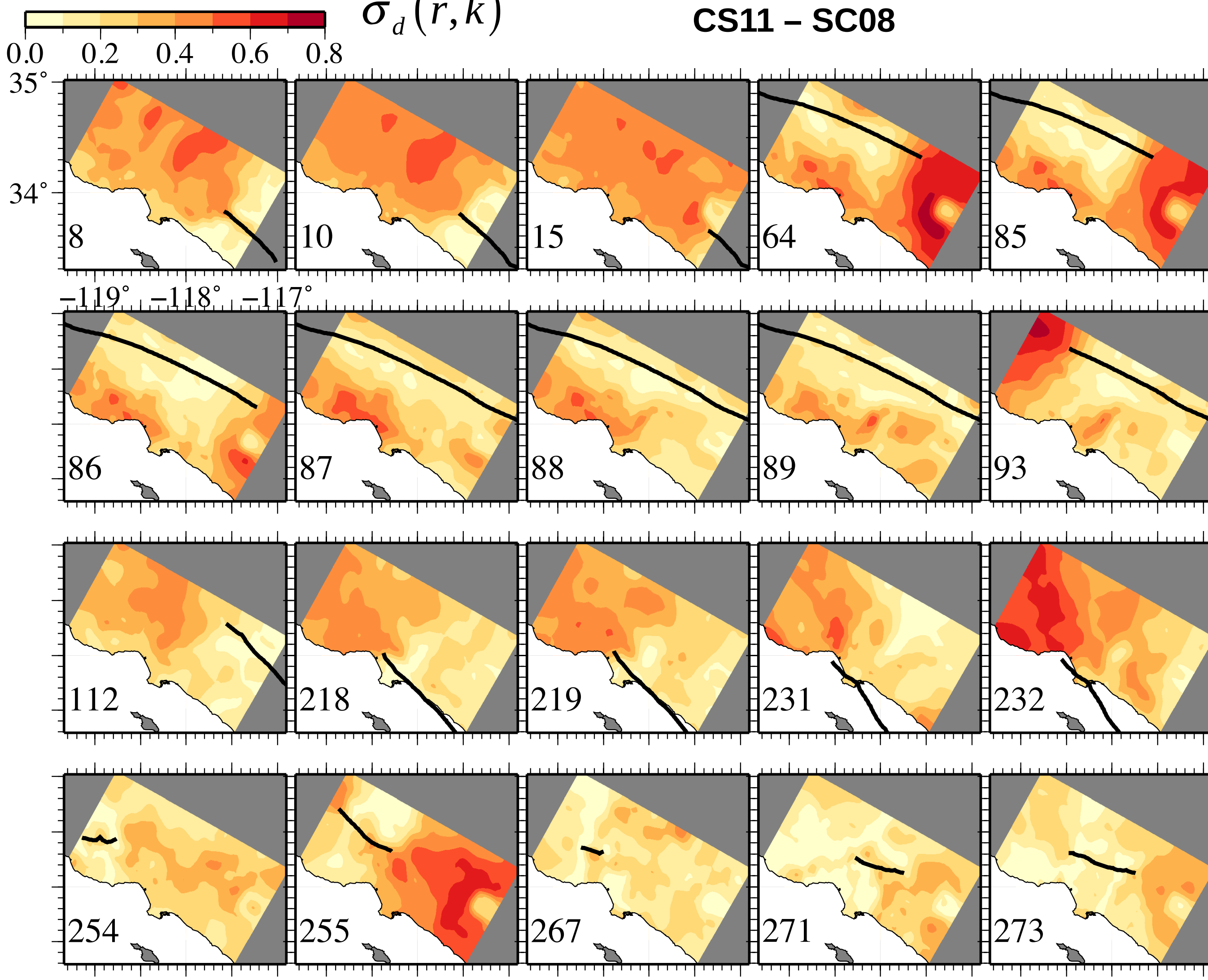


GP07
used in CS11

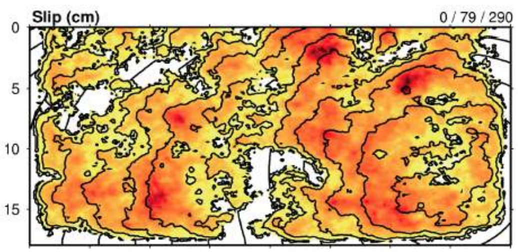
Model	$\overline{\sigma}_d$
GP07 raw	0.41
GP07-SC08	0.31

σ_d maps (SA-3s)

CS11 – SC08



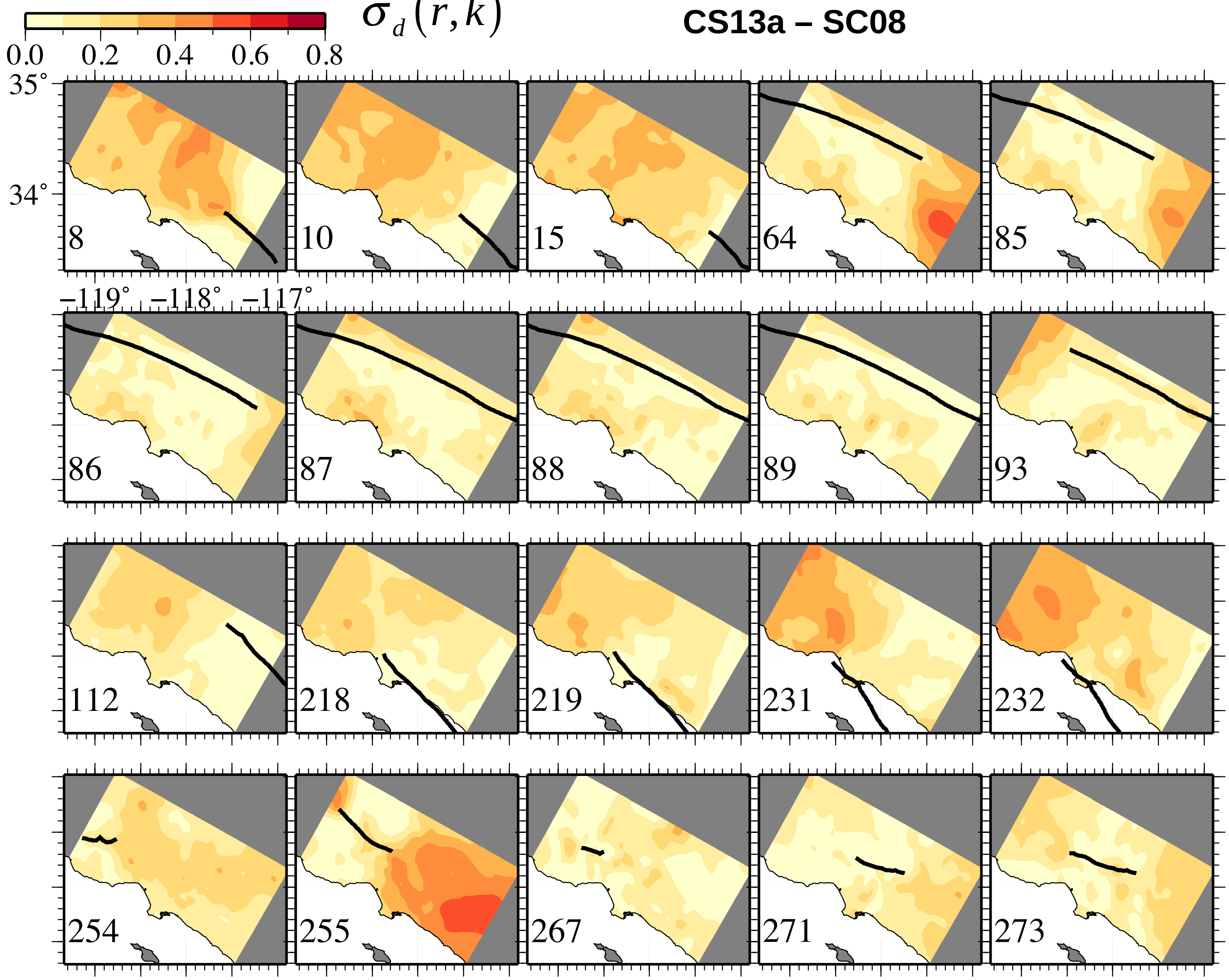
Dependence of Directivity Effects on Rupture Complexity



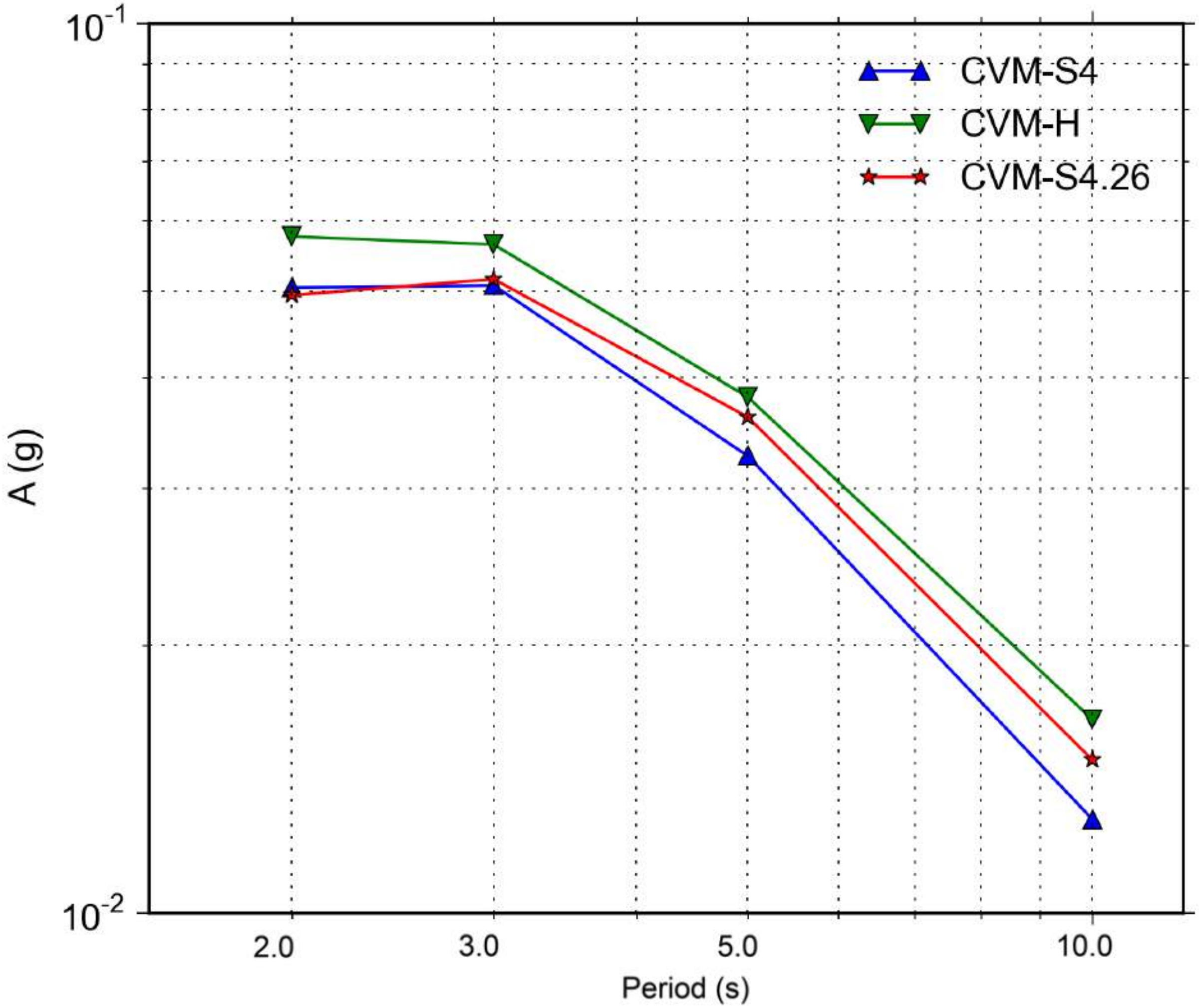
GP10
used in CS13a

Model	$\overline{\sigma}_d$
GP07 raw	0.41
GP07-SC08	0.31
GP10 raw	0.26
GP10-SC08	0.17

σ_d maps (SA-3s)
CS13a – SC08

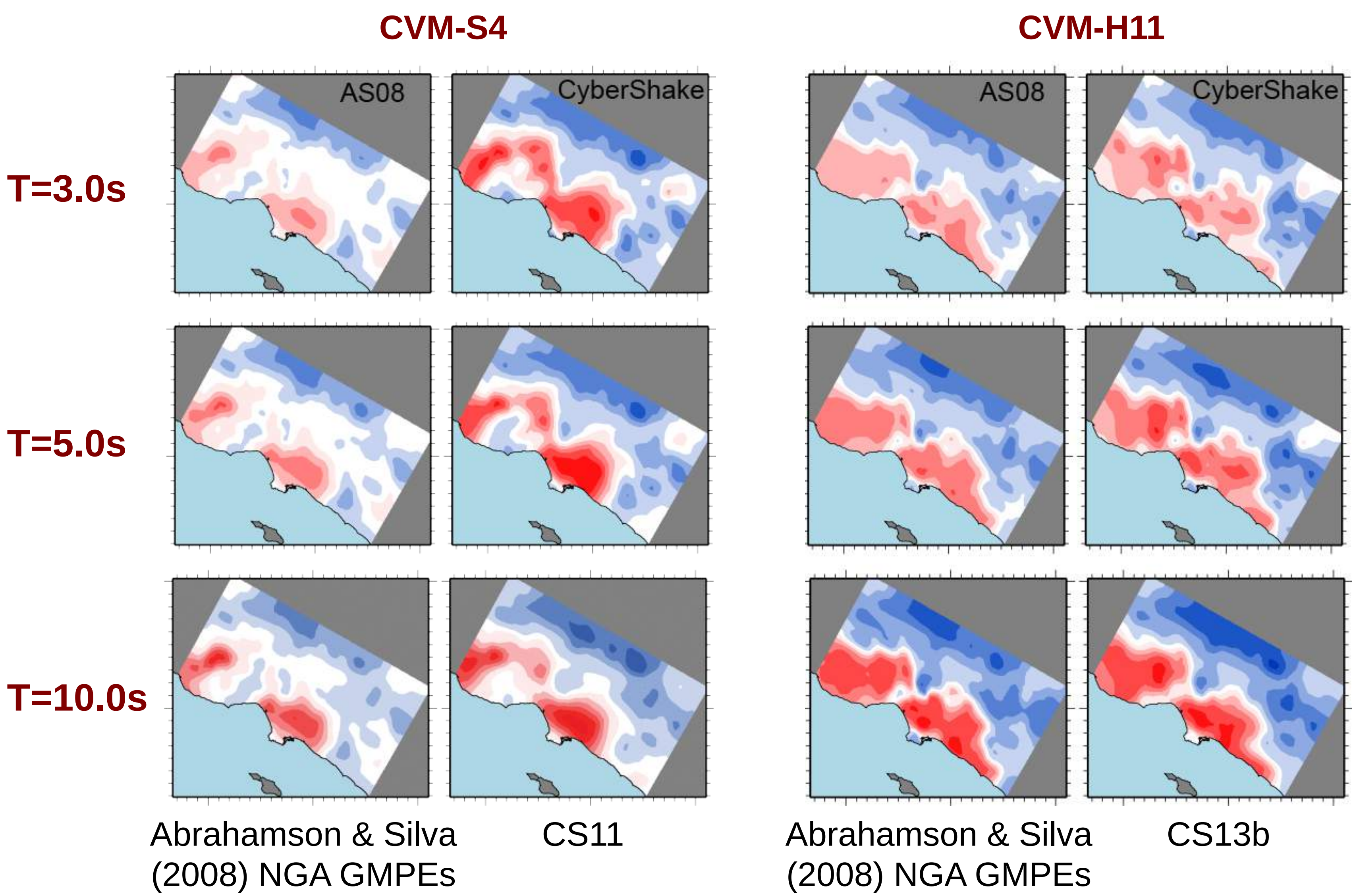


A-values of CyberShake models



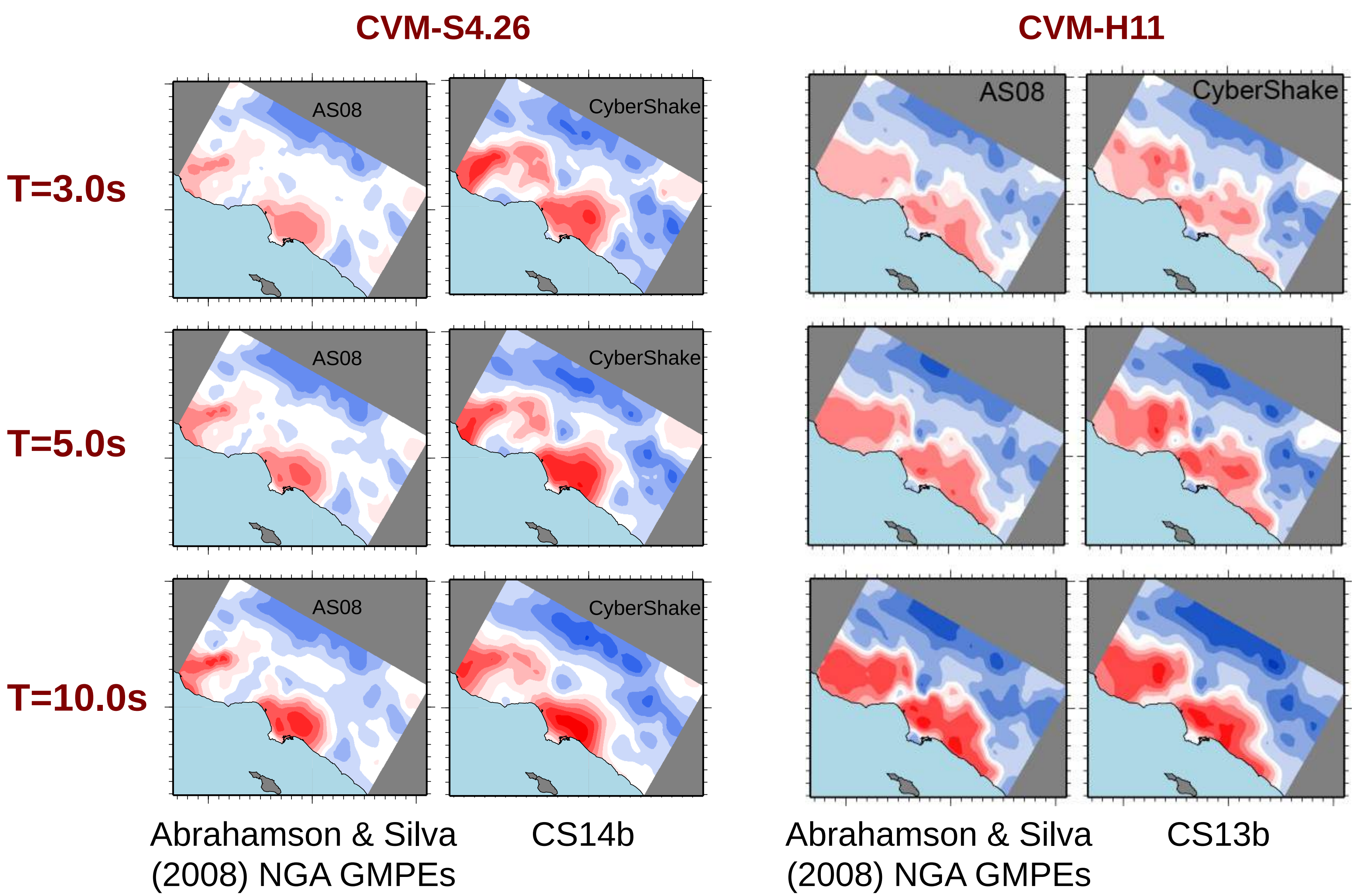
Dependence of Basin Effects on Velocity Structures

(SA corrected for V_{S30} using BA08)



Dependence of Basin Effects on Velocity Structures

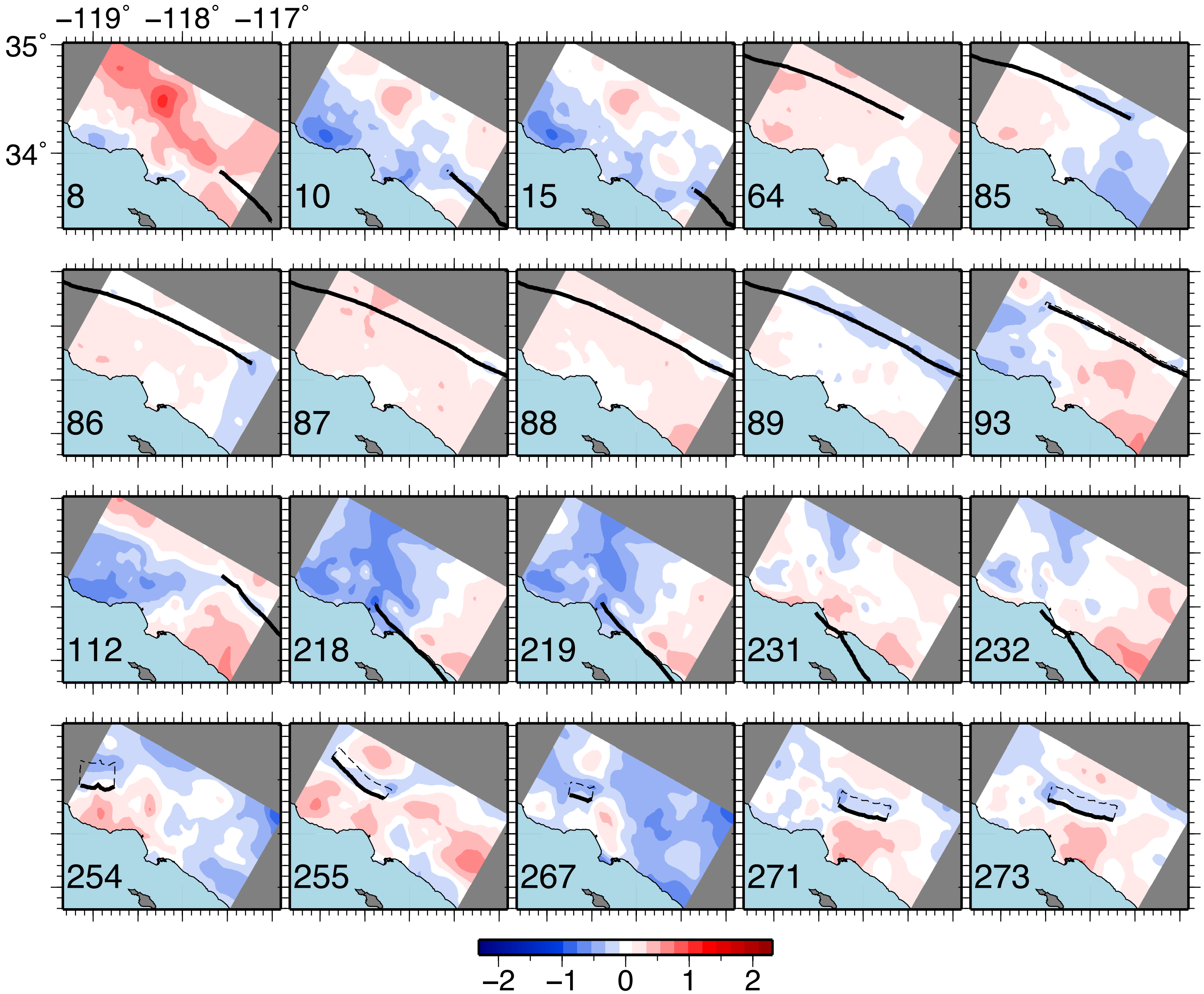
(SA corrected for V_{S30} using BA08)



Dependence of Path Effects on Velocity Structures

(SA-3s corrected for V_{S30} using BA08)

CS11
CVM-S4



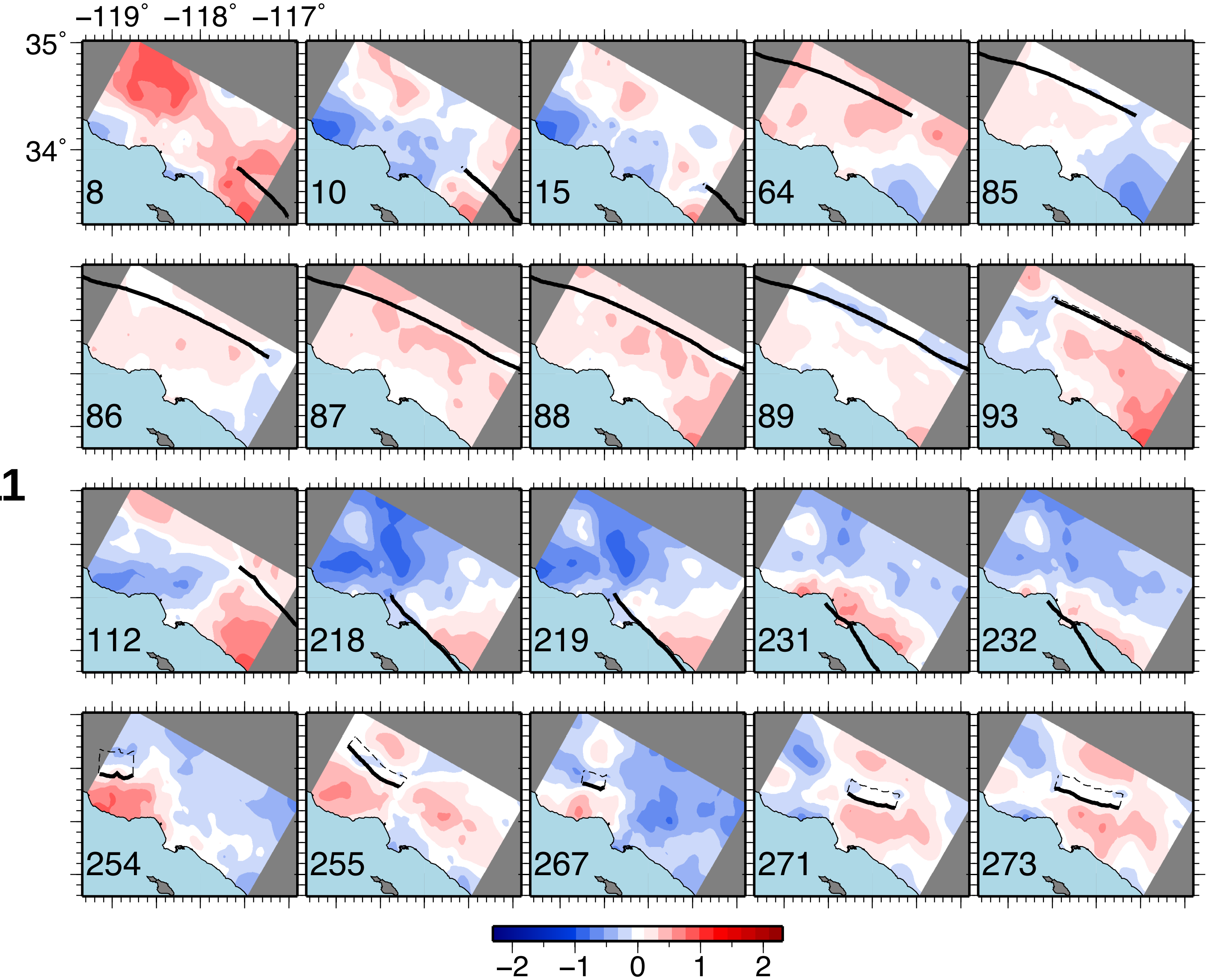
SCEC

Southern California
Earthquake Center

Dependence of Path Effects on Velocity Structures

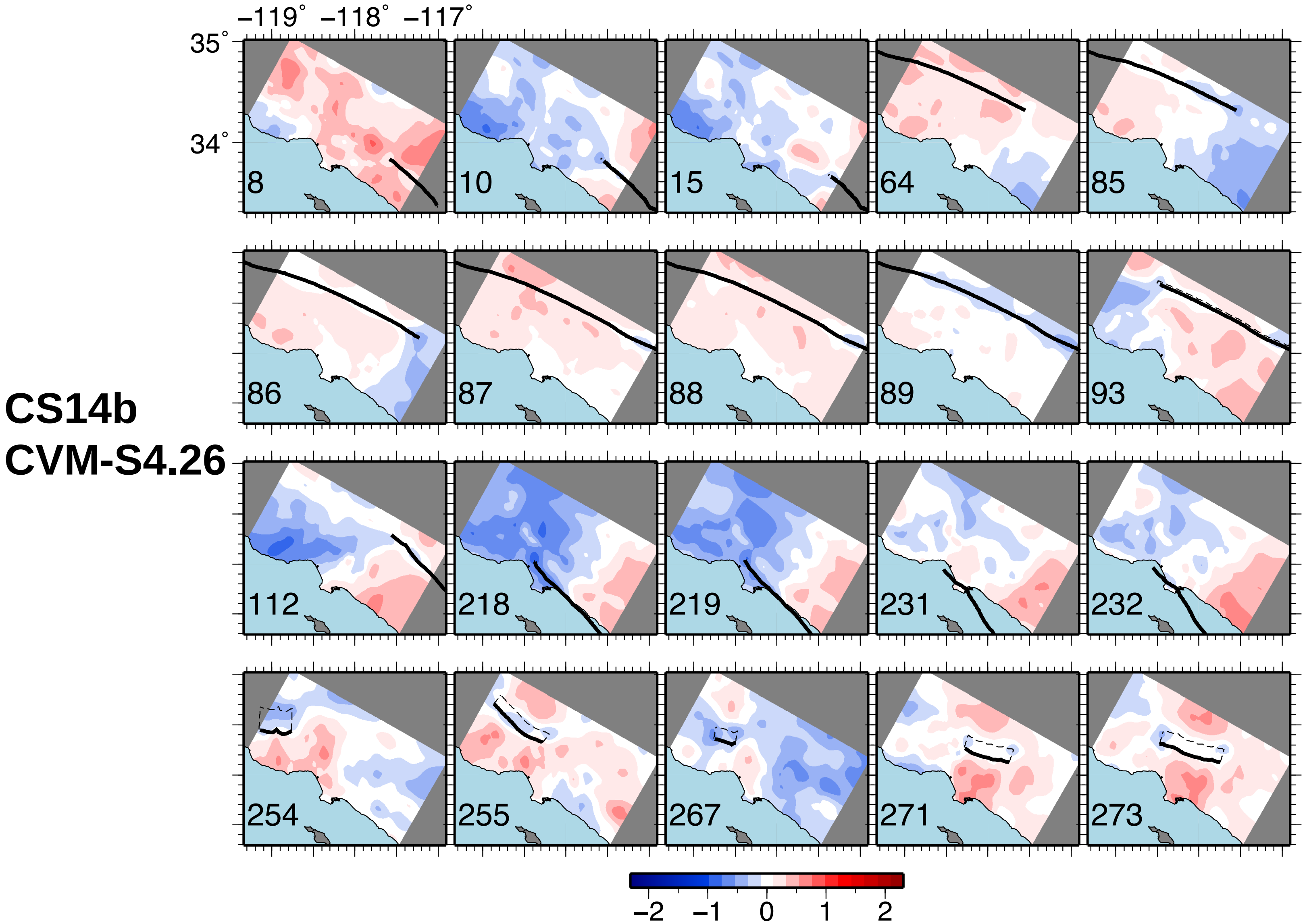
(SA-3s corrected for V_{S30} using BA08)

CS13a
CVM-H11



Dependence of Path Effects on Velocity Structures

(SA-3s corrected for V_{S30} using BA08)



Averaging-Based Factorization

- **ABF representation of excitation functionals**

$$G(r,k,x,s) = A + B(r) + C(r,k) + D(r,k,x) + E(r,k,x,s)$$

- **ABF representation of excitation variance**

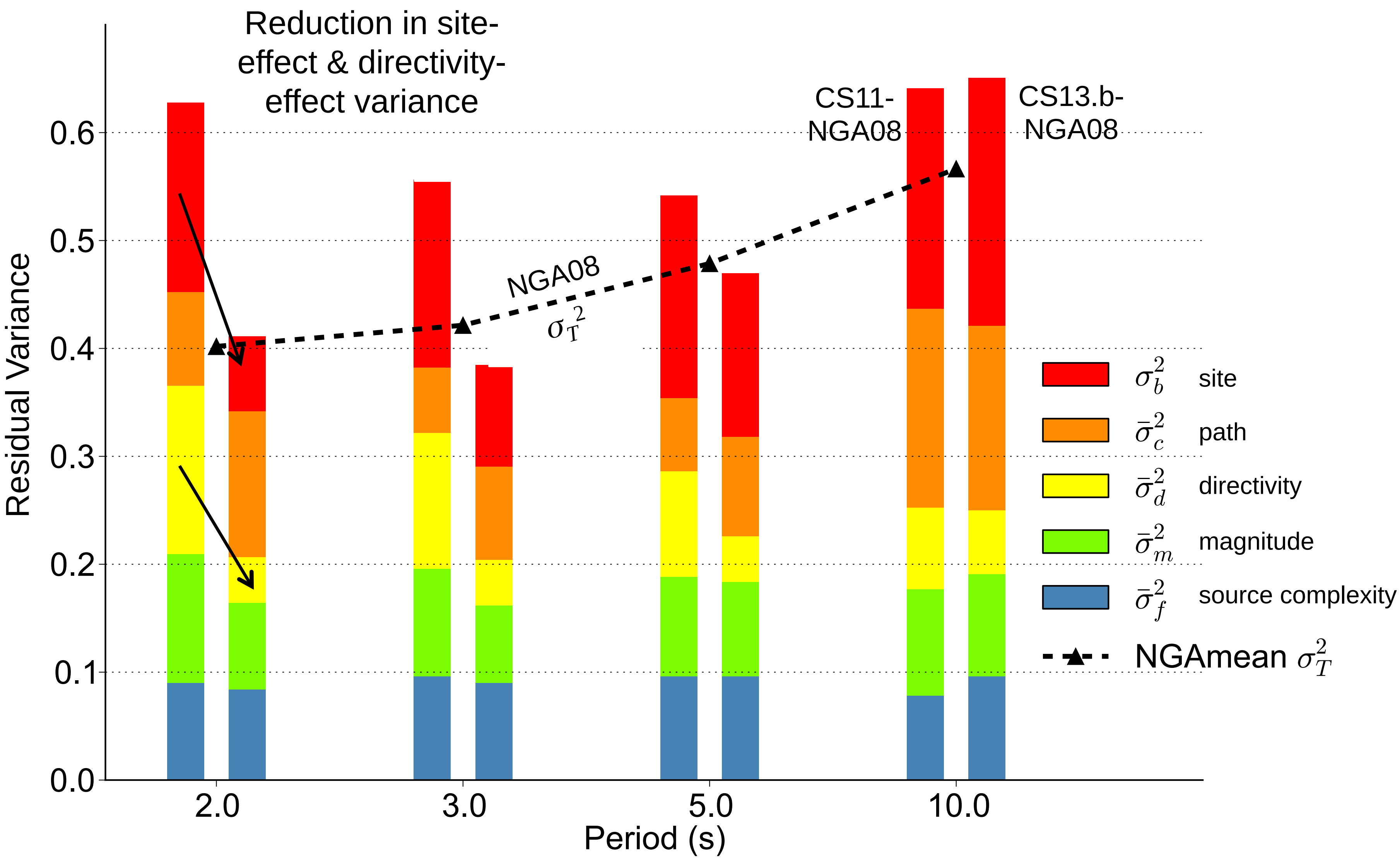
$$\mathrm{E}[G] = \langle G(r,k,x,s) \rangle_{S,X,K,R} = A$$

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

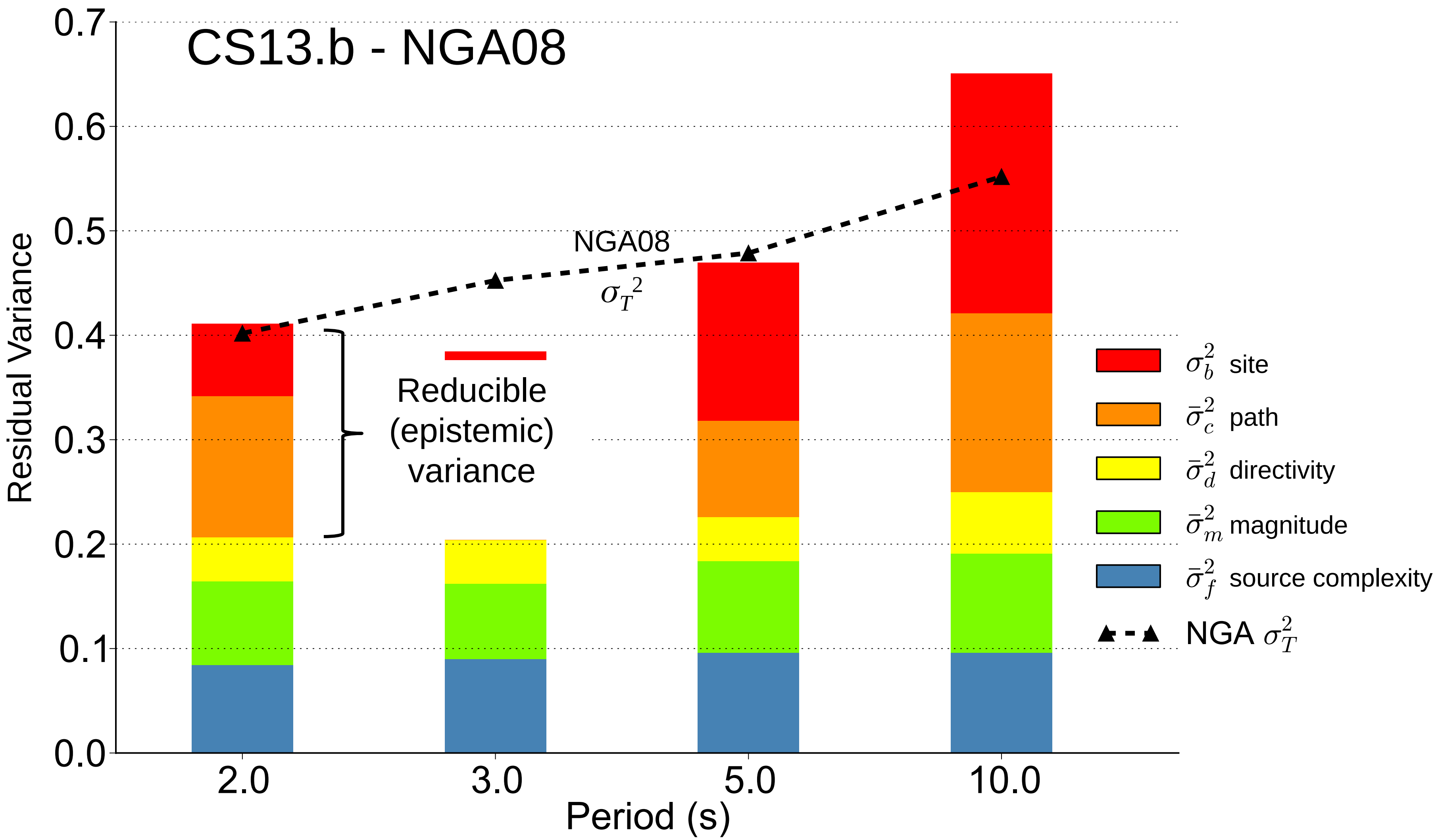
$$= \sigma_B^2 + \langle \sigma_C^2(r) \rangle_R + \langle \sigma_D^2(r,k) \rangle_{K,R} + \langle \sigma_E^2(r,k,x) \rangle_{X,K,R}$$

$$\equiv \sigma_B^2 + \bar{\sigma}_C^2 + \bar{\sigma}_D^2 + \bar{\sigma}_E^2$$

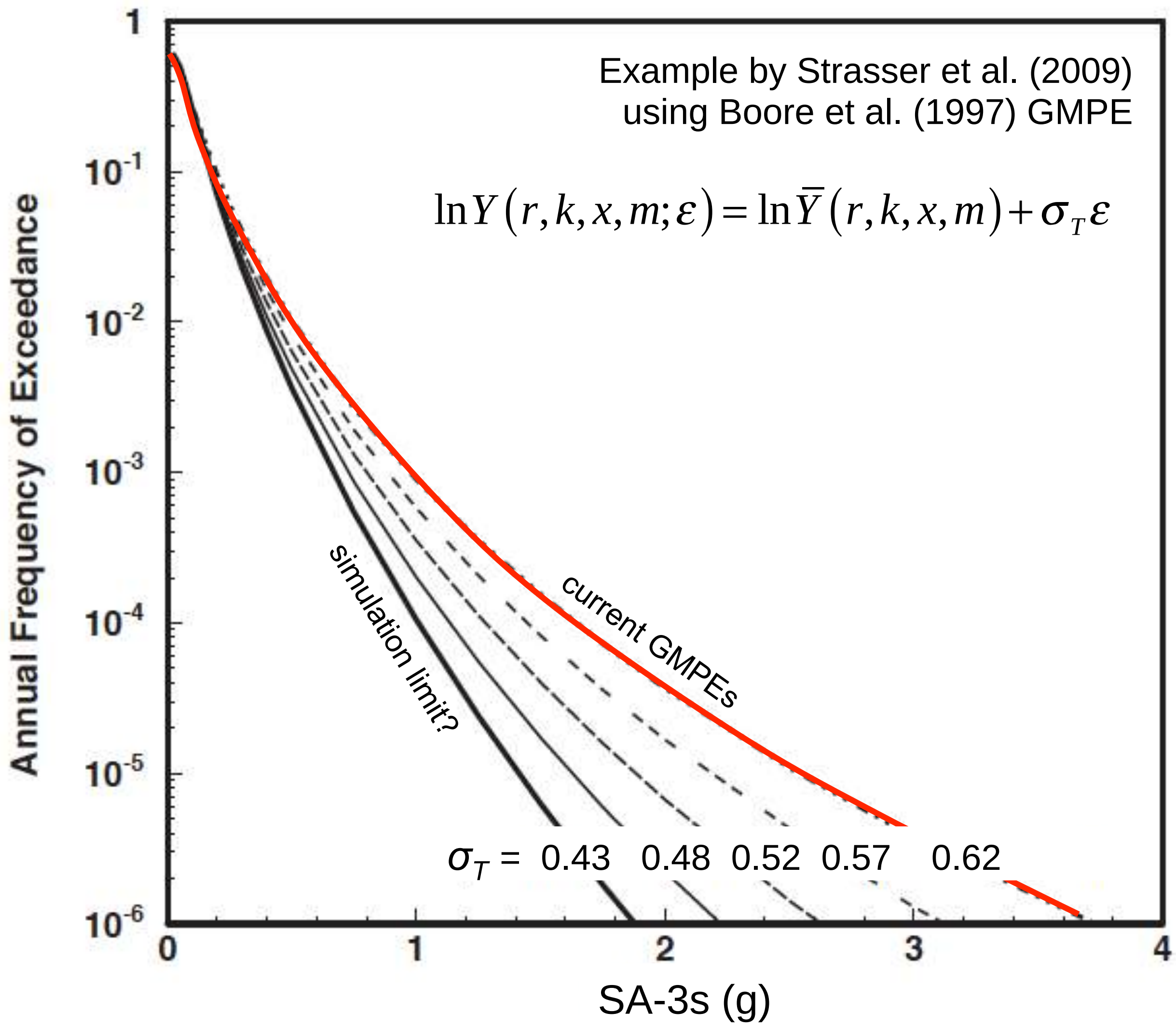
ABF Variance Analysis



ABF Variance Analysis



Importance of Reducing Aleatory Variability

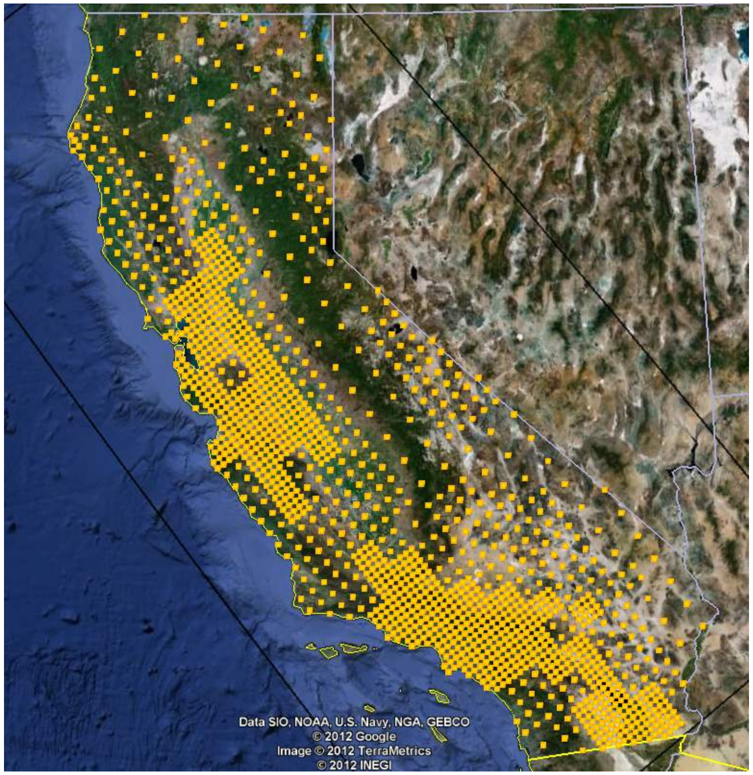


Plans for CyberShake Research

CyberShake: Initiative to Compute a Statewide Physics-Based Hazard Model

Southern California Earthquake Center

- **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

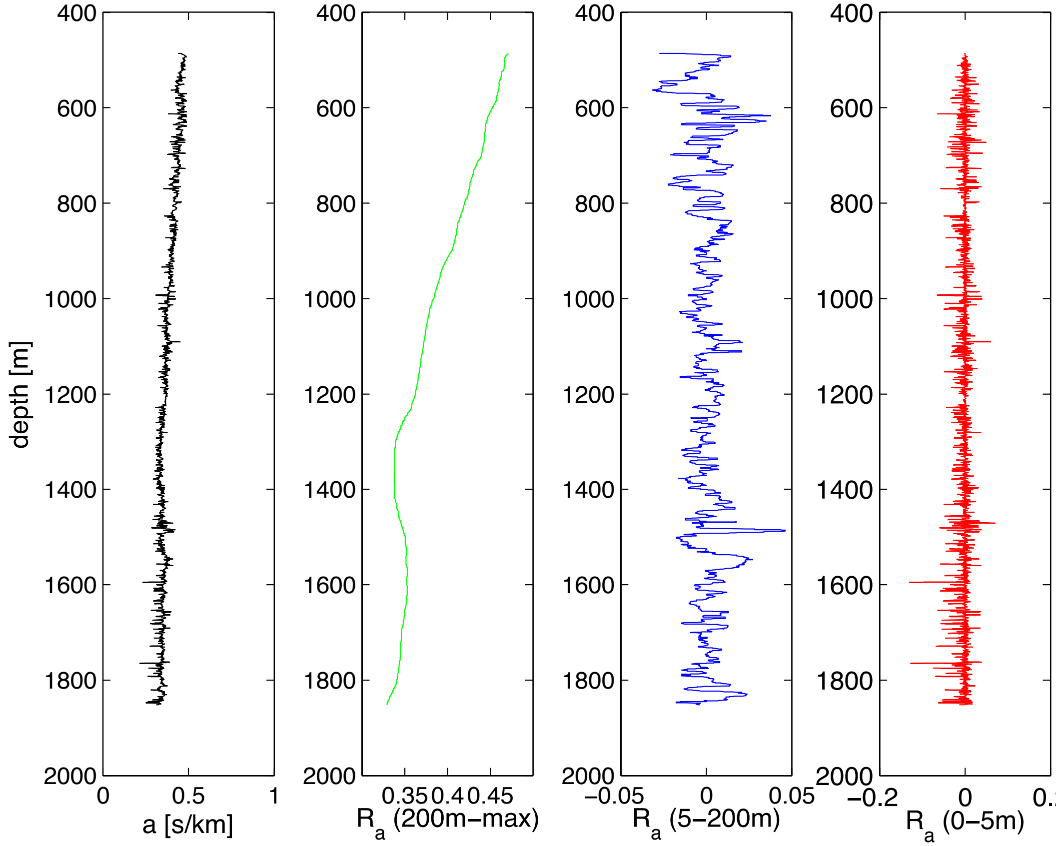
- Computational requirements for 1 Hz deterministic, 10 Hz stochastic:
  - Number of jobs: 23.2 billion
  - Storage: 2800 TB seismograms
  - Computer hours: 392 million

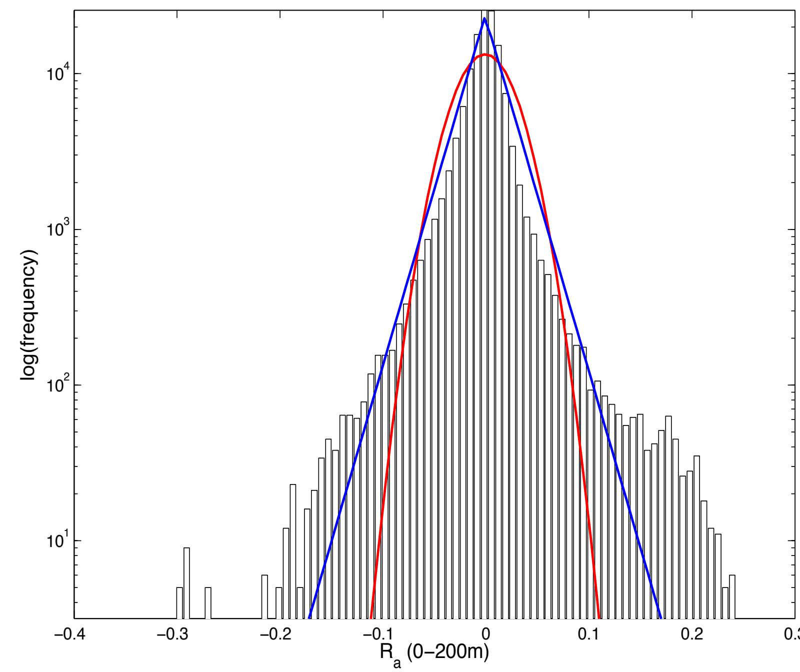
CyberShake Science Challenges

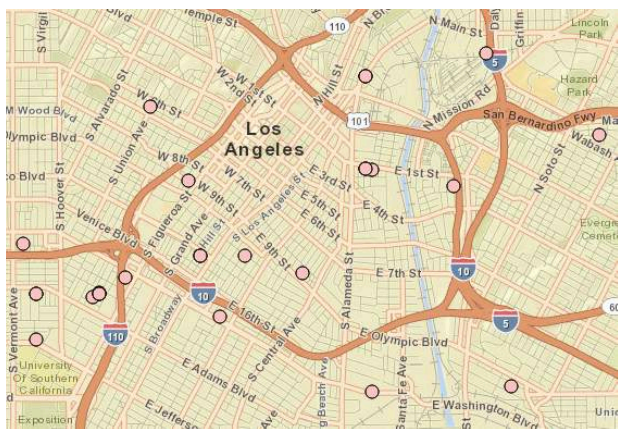
- **Move towards**
  - **higher frequencies (0.5 Hz → 5 Hz)**
  - **more ruptures (UCERF3)**
  - **more sites (1440 for statewide)**
- **This will require better physics…**
  - **Frequency-dependent attenuation**
  - **Fault roughness**
  - **Near-fault plasticity**
  - **Soil nonlinearities**
  - **Near-surface heterogeneities**

… and much more computation!

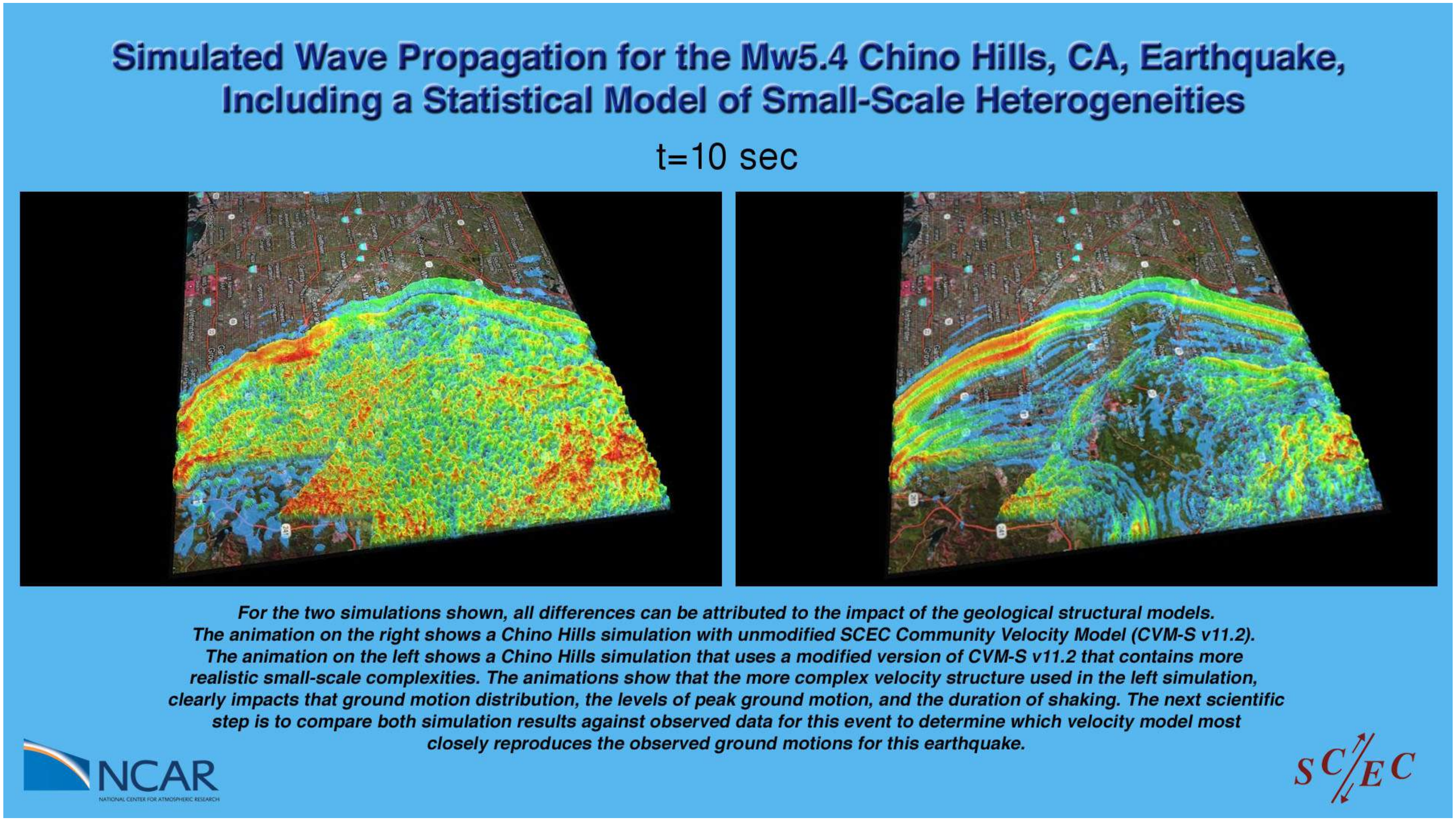
Near-Surface Heterogeneities





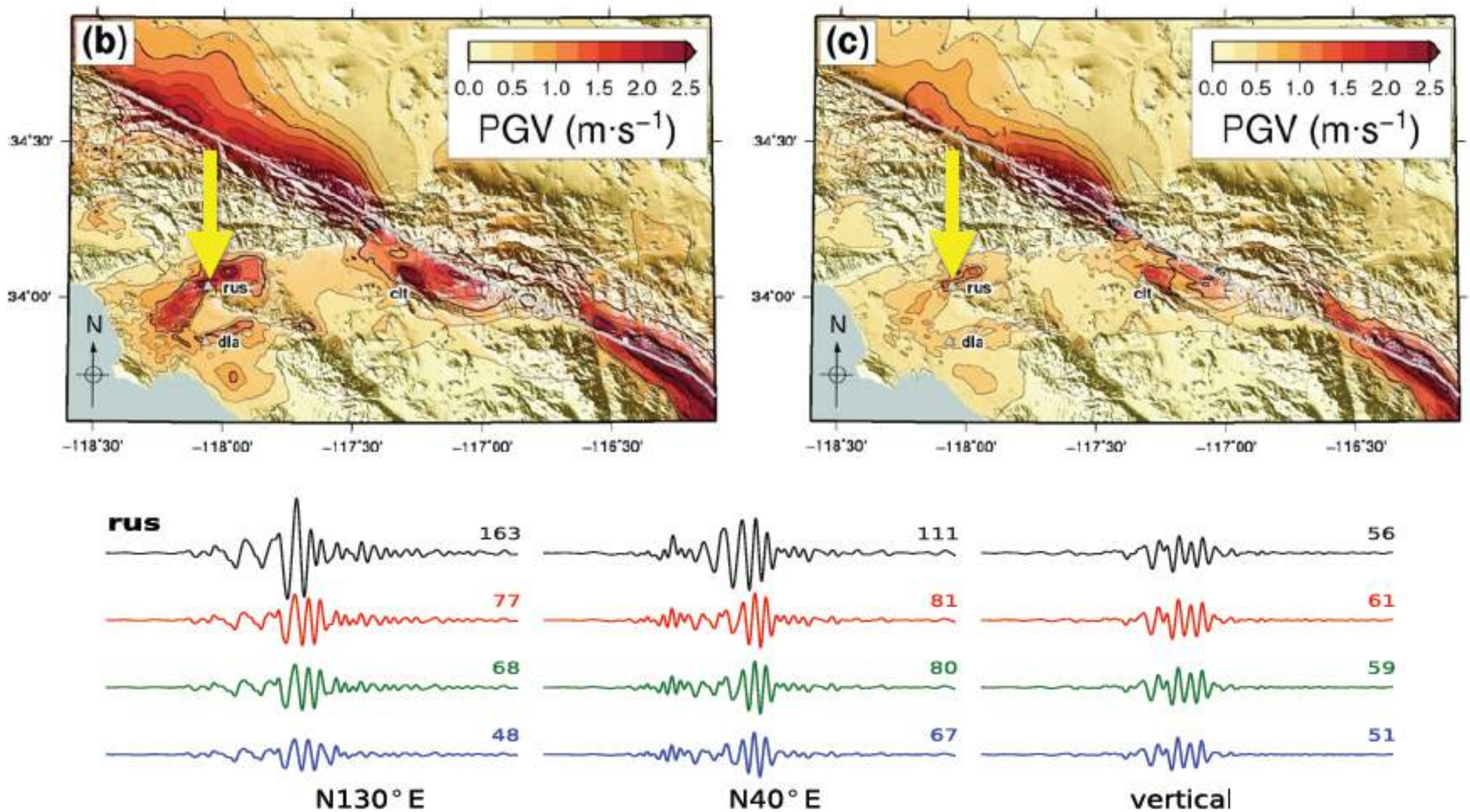


Effects of Near-Surface Heterogeneities



First results on SCEC's High-F project from Yellowstone

Nonlinear Simulations of the ShakeOut Scenario



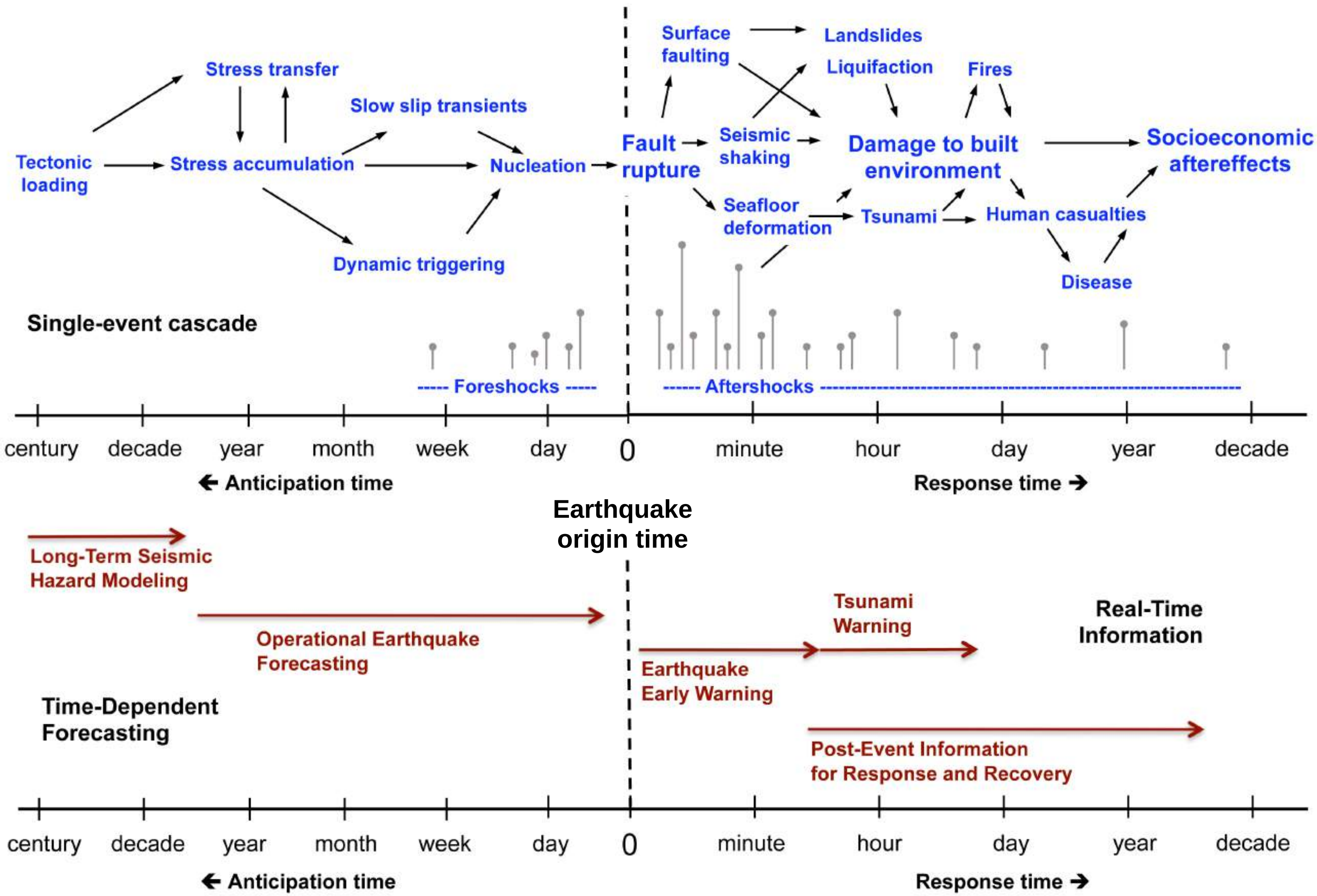
Roten et al. (2014)

Issues for the UGMS Committee

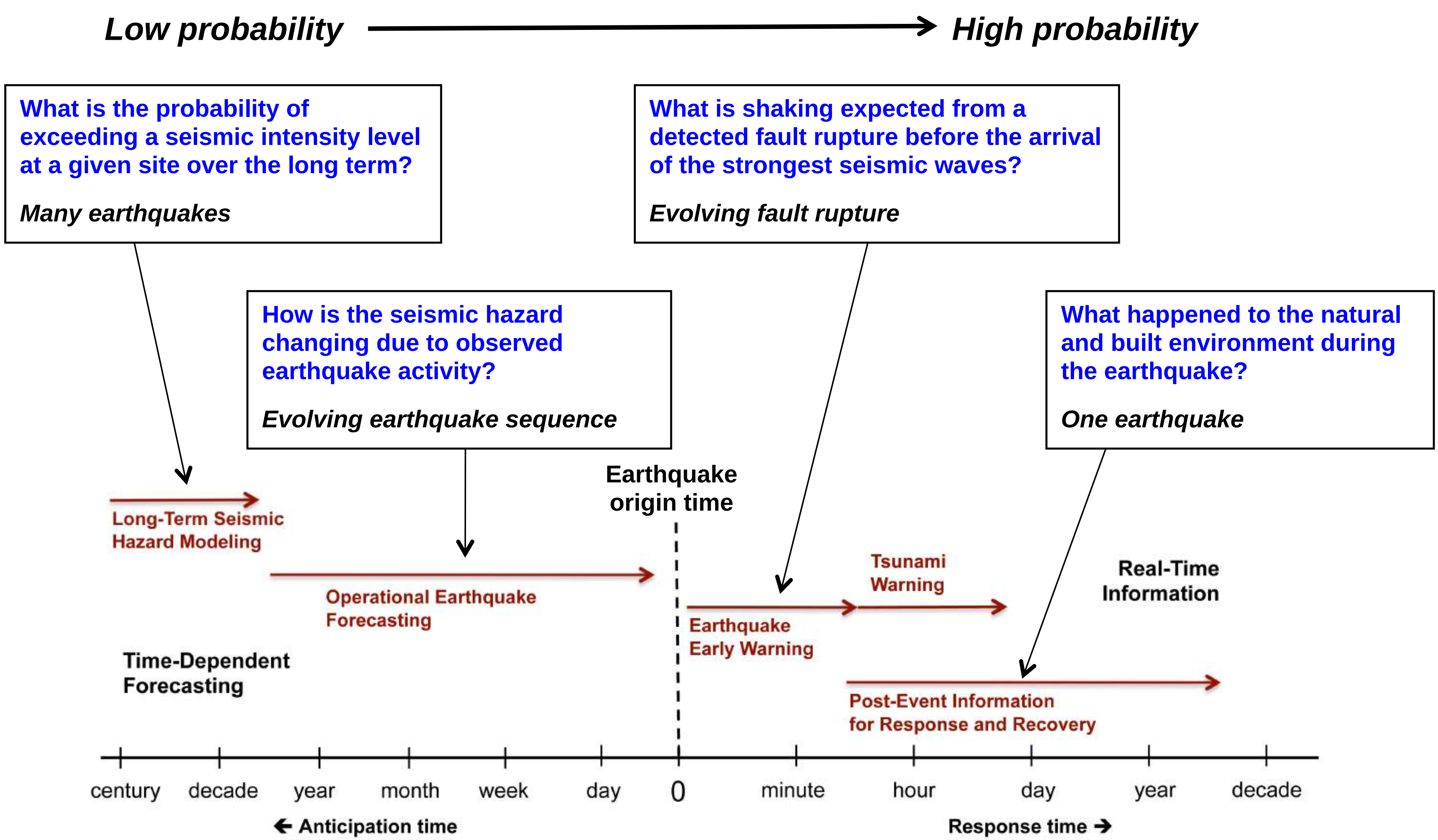
- What frequency range should CyberShake strive towards?
- How useful would be a hybrid (deterministic/stochastic) broadband CyberShake?
- What frequencies should be sampled in the CyberShake database?
- Should we compute vertical component seismograms for CyberShake?
- What performance measures should we use in validating our simulations?

Time-Dependent Ground Motion Forecasting using CyberShake

Tracking Earthquake Cascades

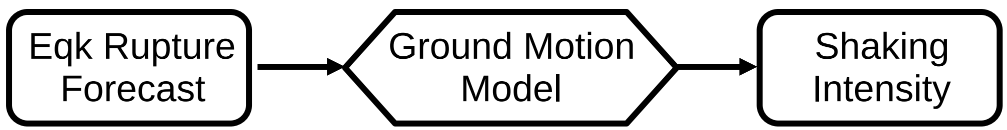


Tracking Earthquake Cascades



CyberShake: Application to Short-Term Earthquake Forecasting

- **Pre-computed CyberShake ground motion models are easily coupled to short-term forecasting models, such as STEP and UCERF3**
  - **Output is a time-dependent seismic hazard estimate**

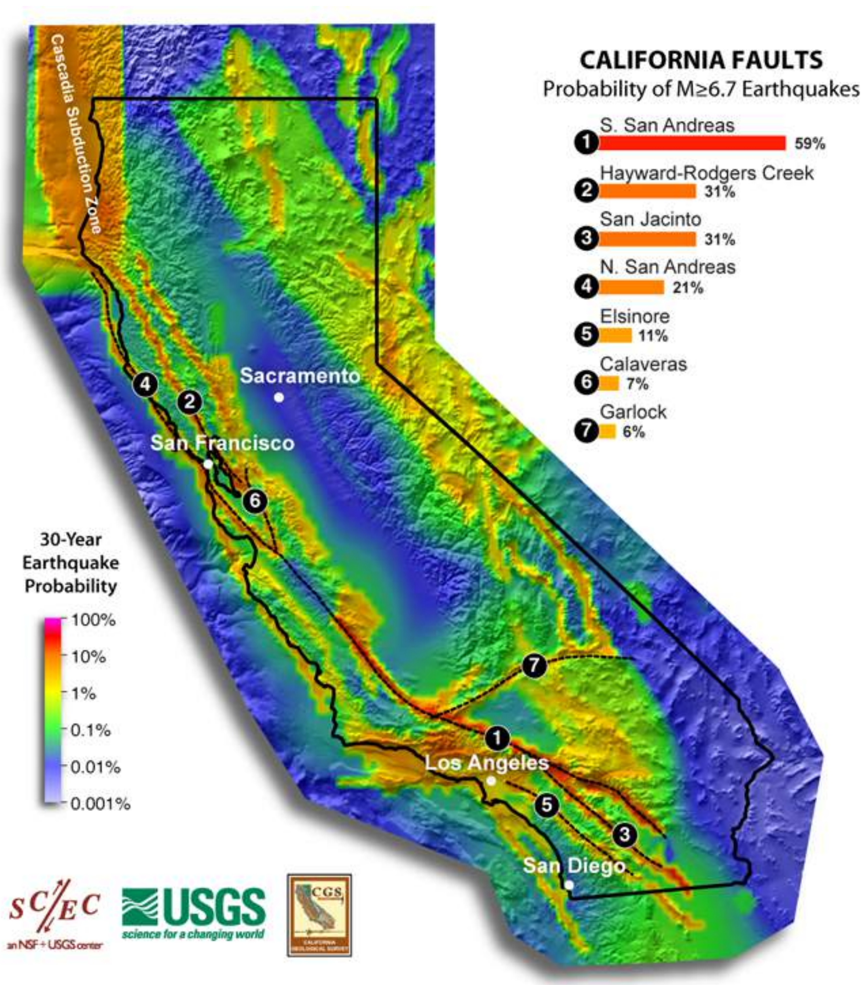


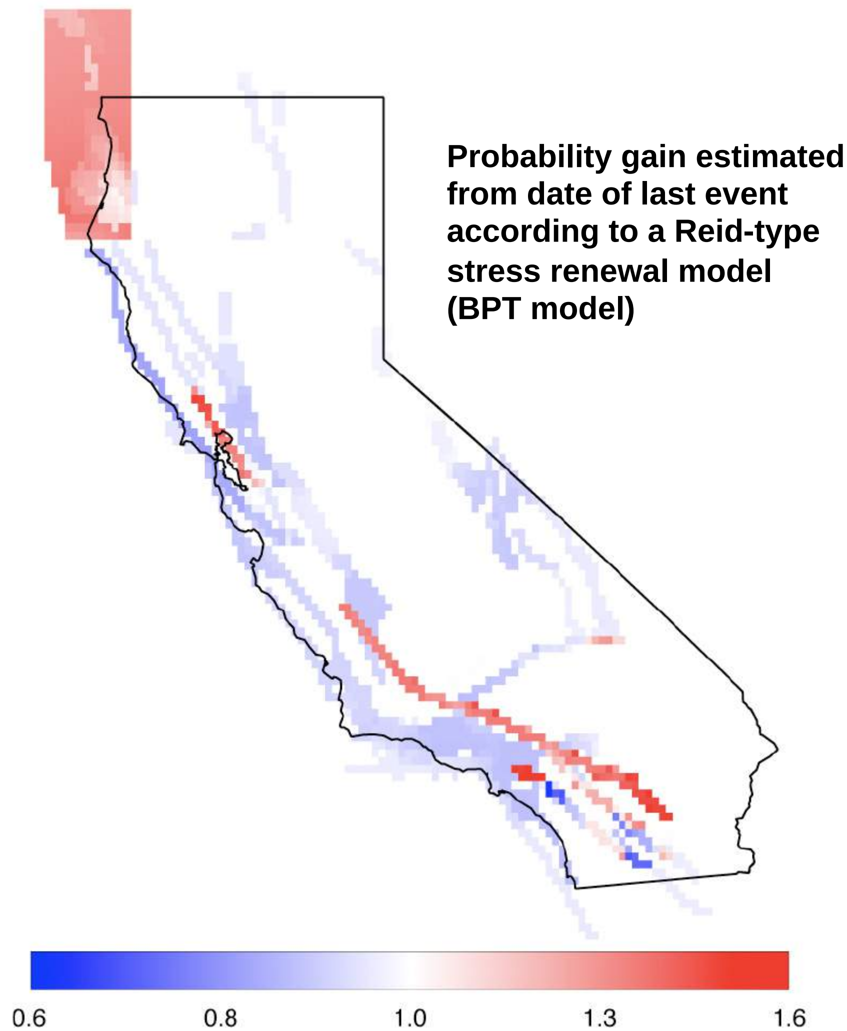
$P(S_k, T)$ $P(Y_n \mid S_k)$ $P(Y_k, T)$ **T = forecast time**

- **Short-term forecasting localizes epicenter probabilities**
  - **Coupled model achieves significant gains in ground motion probabilities through the forecasting of source directivity and directivity-basin coupling**

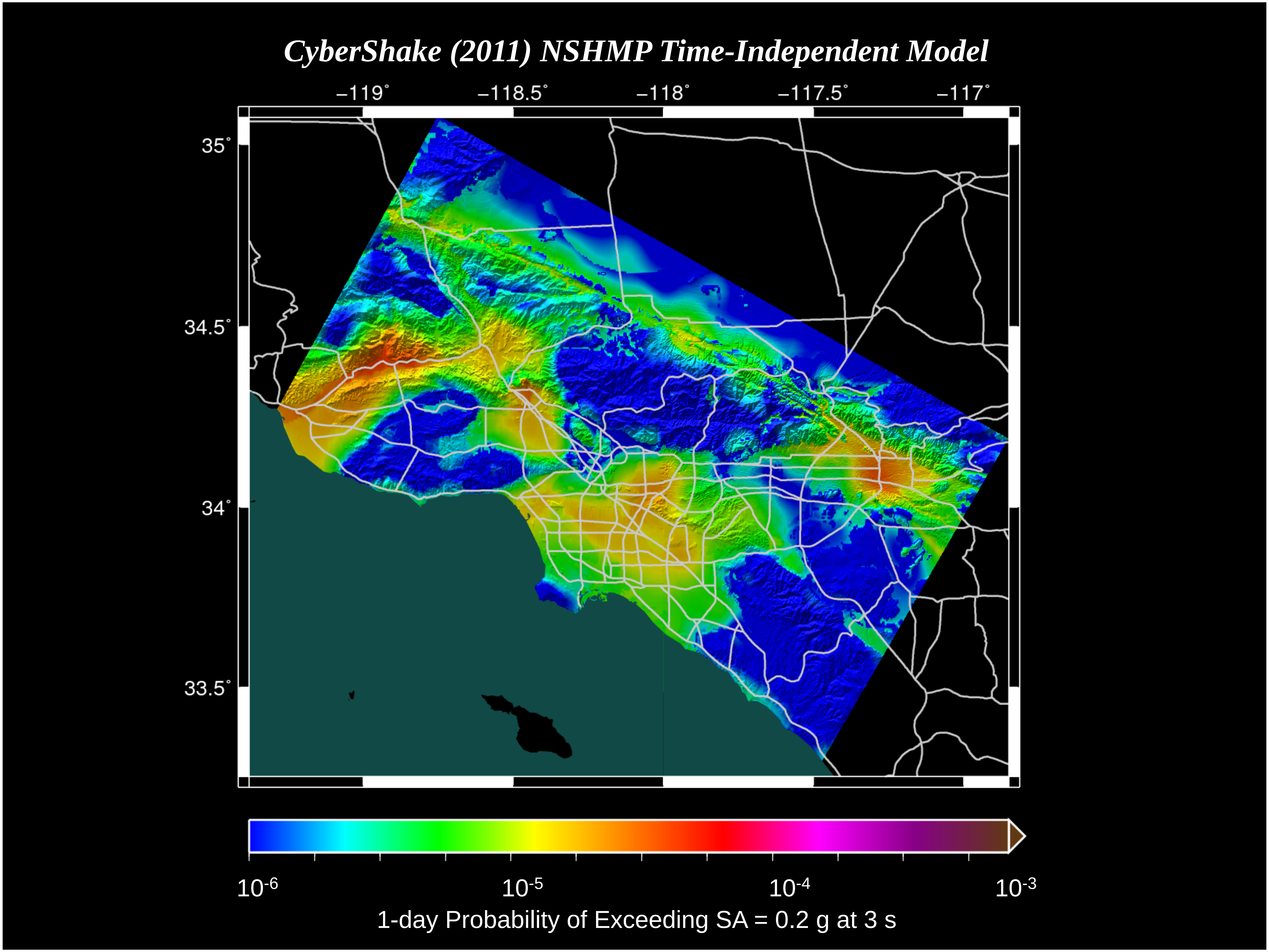
Working Group on California Earthquake Probabilities (2007)

Uniform California Earthquake Rupture Forecast (UCERF2)

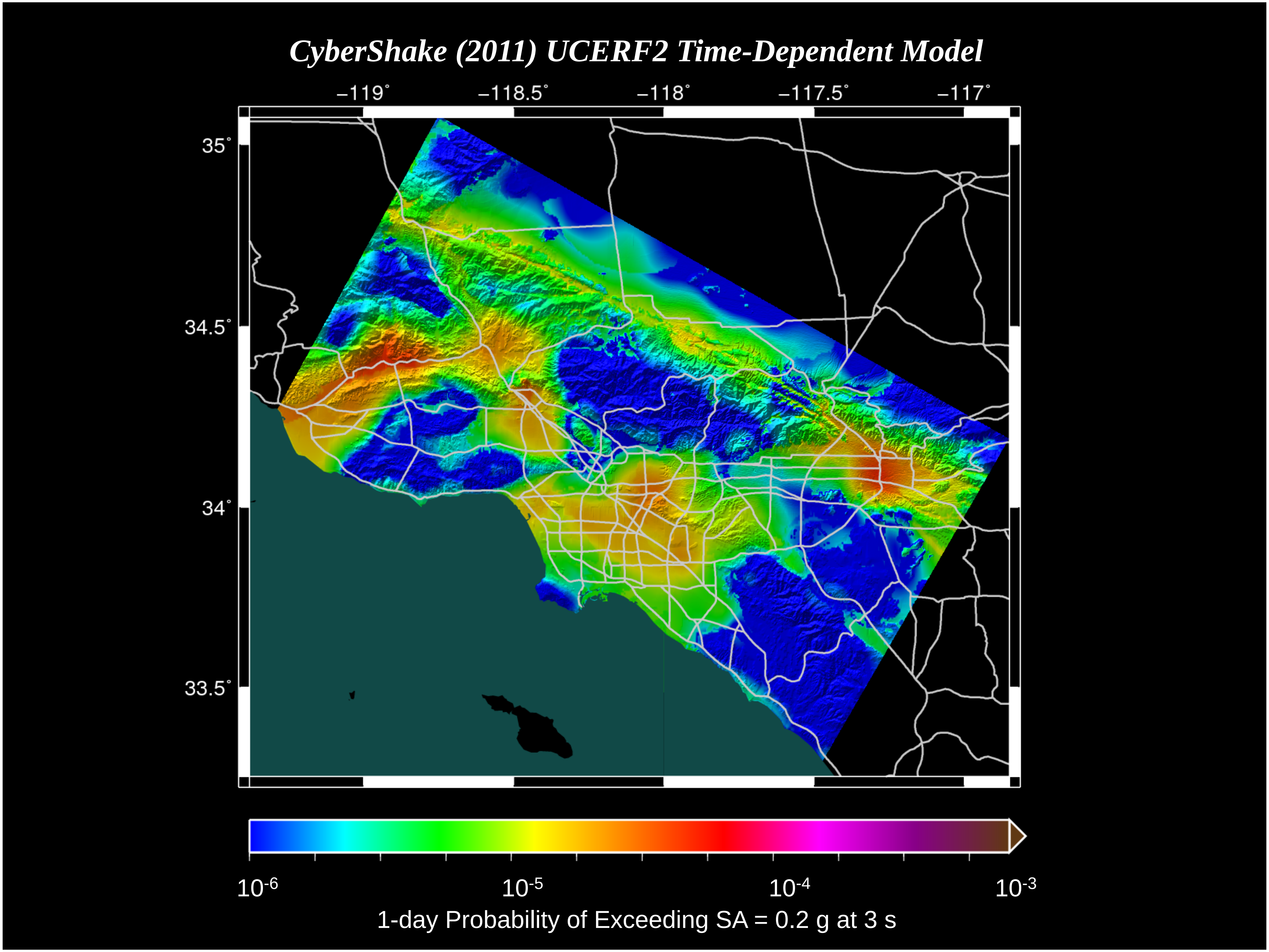




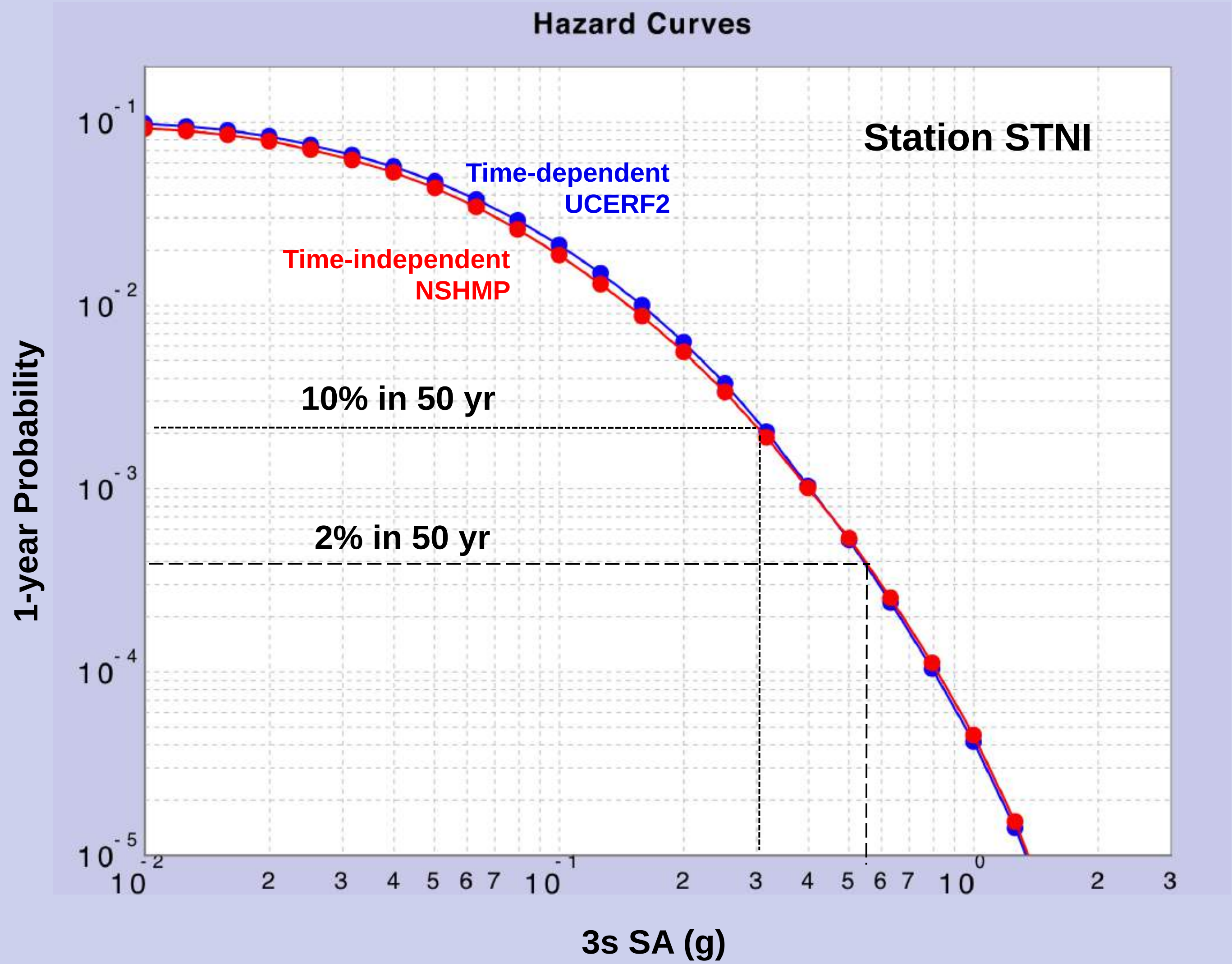
Ratio of time-dependent to time-independent participation probabilities for M ≥ 6.7



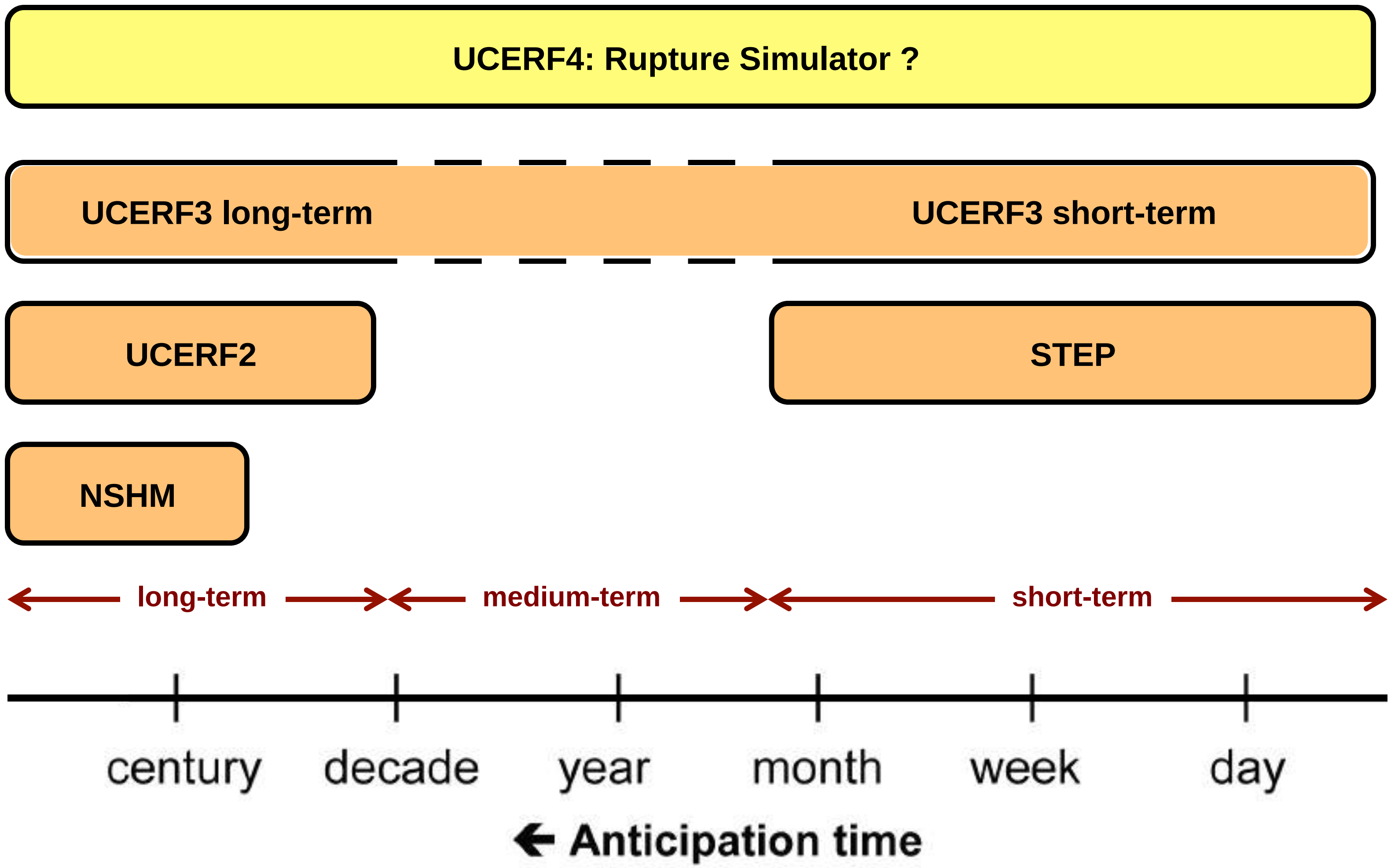
CyberShake (2011) NSHMP Time-Independent Model
−119˚
−118.5˚
−118˚
−117.5˚
−117˚
35˚
34.5˚
34˚
33.5˚
10-6
10-5
10-4
10-3
1-day Probability of Exceeding SA = 0.2 g at 3 s


CyberShake (2011) UCERF2 Time-Dependent Model
–119˚
–118.5˚
–118˚
–117.5˚
–117˚
35˚
34.5˚
34˚
33.5˚
10-6
10-5
10-4
10-3
1-day Probability of Exceeding SA = 0.2 g at 3 s

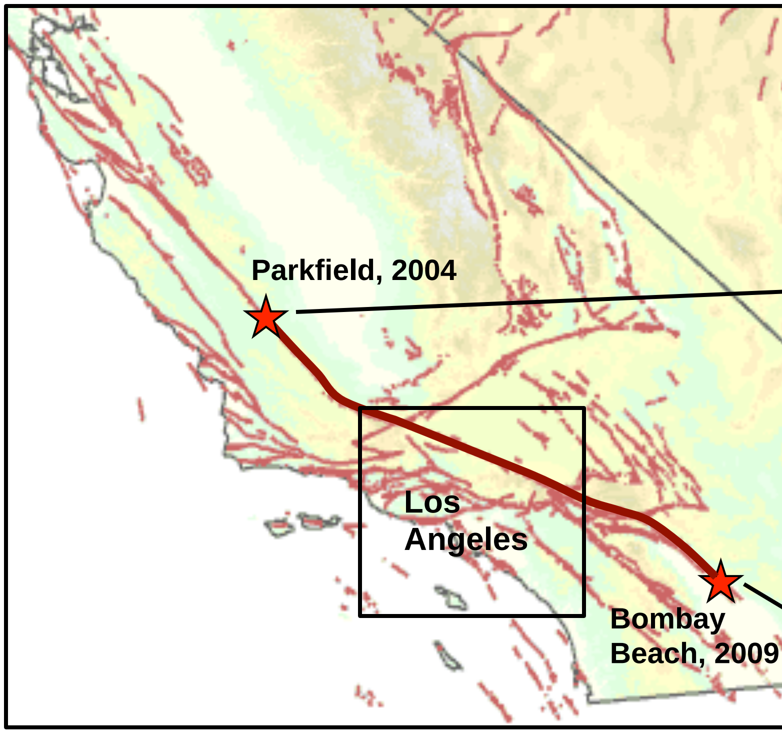
CyberShake Time-Independent Hazard Curves

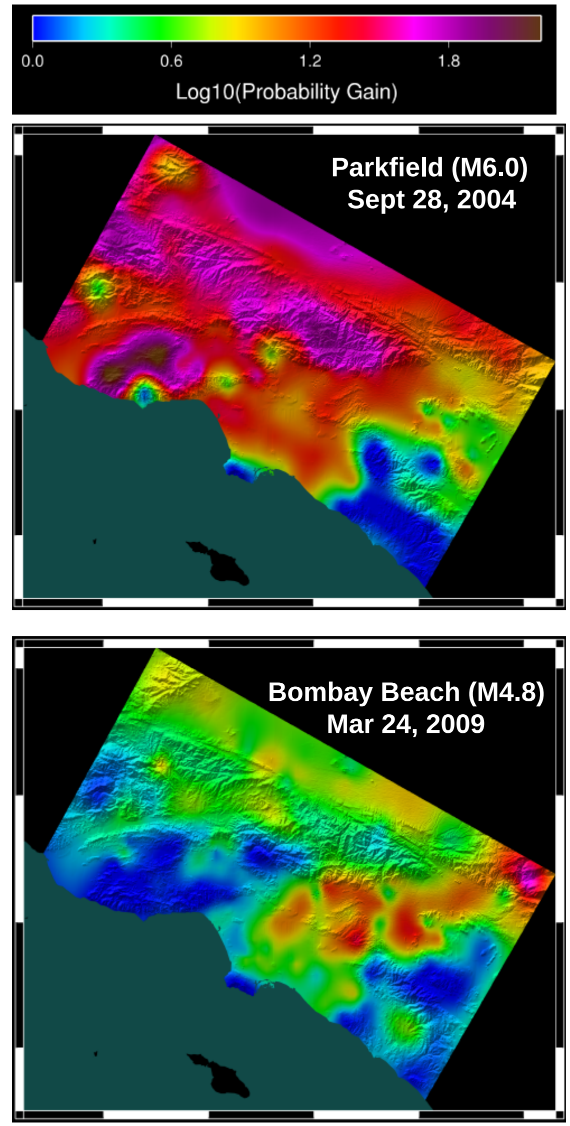


California Earthquake Forecasting Models



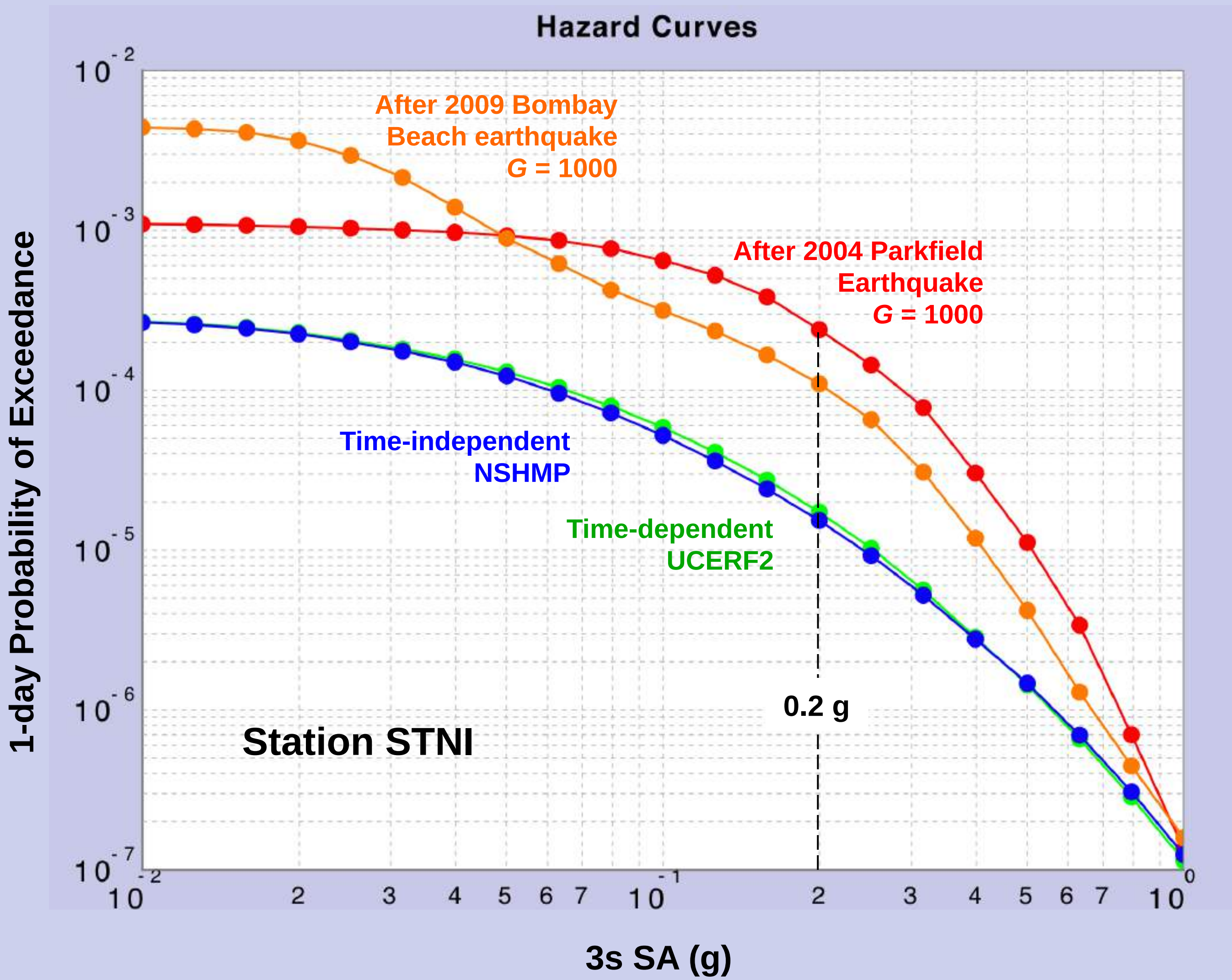
CyberShake: Application to Short-Term Earthquake Forecasting

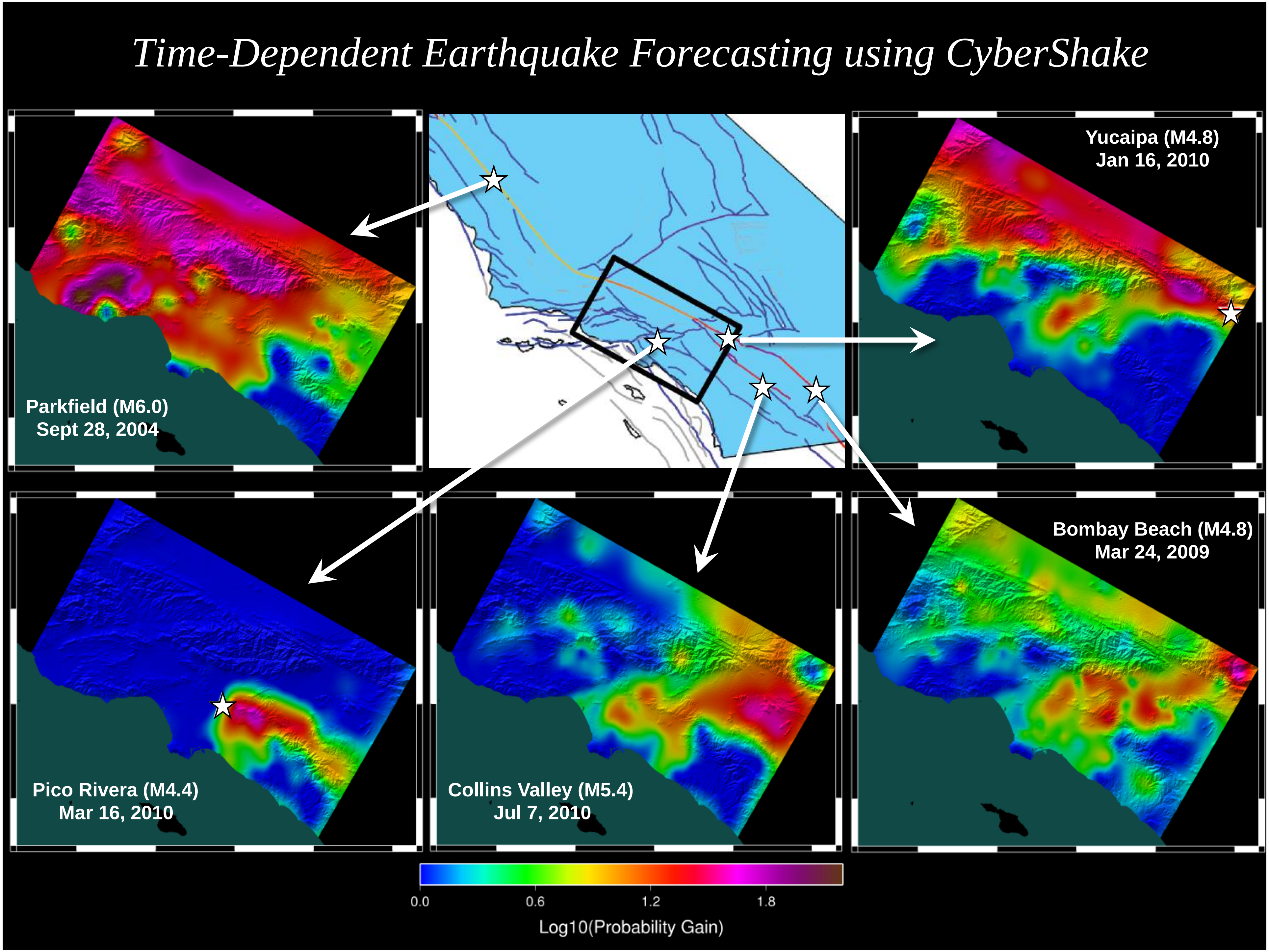




- **Compute probability gain from forecasting model.**
  Example: *G* = 1000 for *R* ≤ *10 km*
- **Apply probability gain to CyberShake ruptures and re-compute ground motion probabilities for short interval following events. Example: 1 day**

CyberShake Time-Dependent Hazard Curves





Time-Dependent Earthquake Forecasting using CyberShake
Parkfield (M6.0)
Sept 28, 2004
Yucaipa (M4.8)
Jan 16, 2010
Pico Rivera (M4.4)
Mar 16, 2010
Collins Valley (M5.4)
Jul 7, 2010
Bombay Beach (M4.8)
Mar 24, 2009
0.0
0.6
1.2
1.8
Log10(Probability Gain)

Thank you!