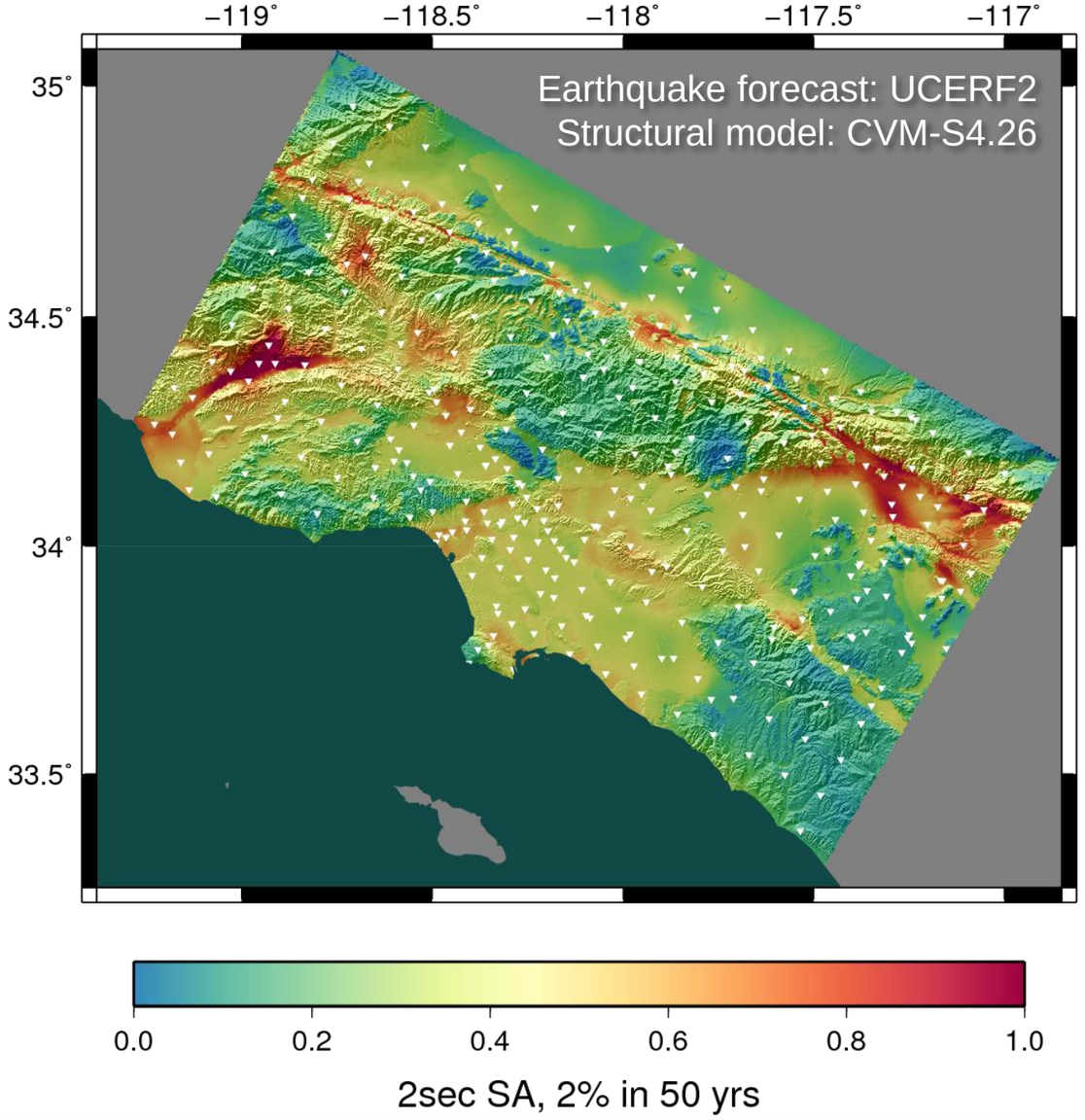
Current SCEC CyberShake Activities

Scott Callaghan

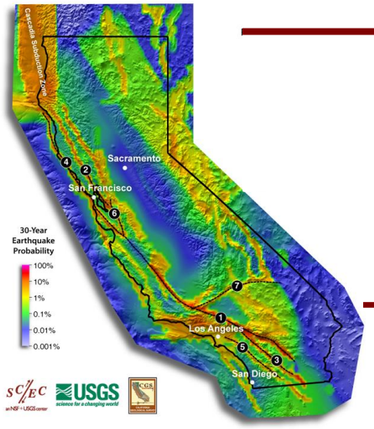
QuakeCoRE GMS&V meeting
May 24, 2018

CyberShake Overview

- Southern California Earthquake Center's 3D physics-based probabilistic seismic hazard analysis (PSHA) platform
- UCERF2 used as earthquake rupture forecast (M≥6.5, <200 km)
- Reciprocity-based approach to simulate seismograms (~500,000 per site)
- Intensity measures (geometric mean, RotD50, RotD100) & duration calculated
- Hazard results from individual sites interpolated with GMPEs to make map

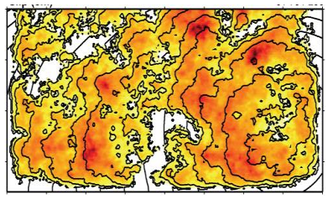


CyberShake Components



Uniform California Earthquake Rupture Forecast

Graves-Pitarka (2014) kinematic rupture generator



UCVM

Mesh generation
1 job per site
MPI, 1500-4000 cores

AWP-ODC

SGT computation
2 jobs per site
MPI, 200-800 GPUs

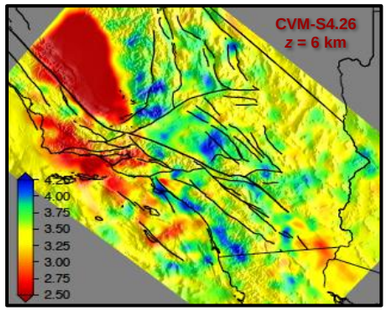
Seismogram Synthesis

Post-processing
500,000 events per site
MPI master/worker, ~4000 cores

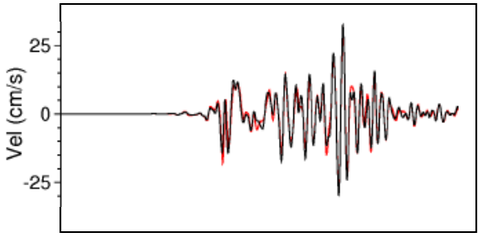
11 GB
data transfer

Data Product Generation

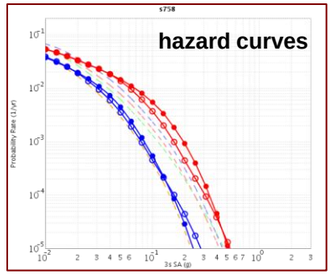
Populate DB, construct images
6 jobs per site



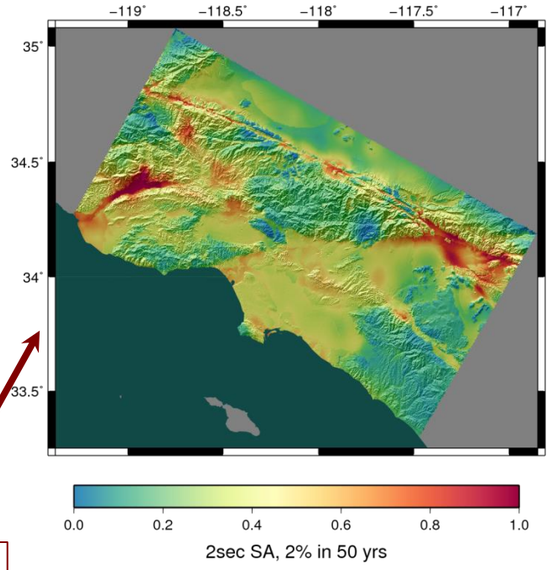
Community Velocity Model



500,000 Seismograms
75M intensity measures



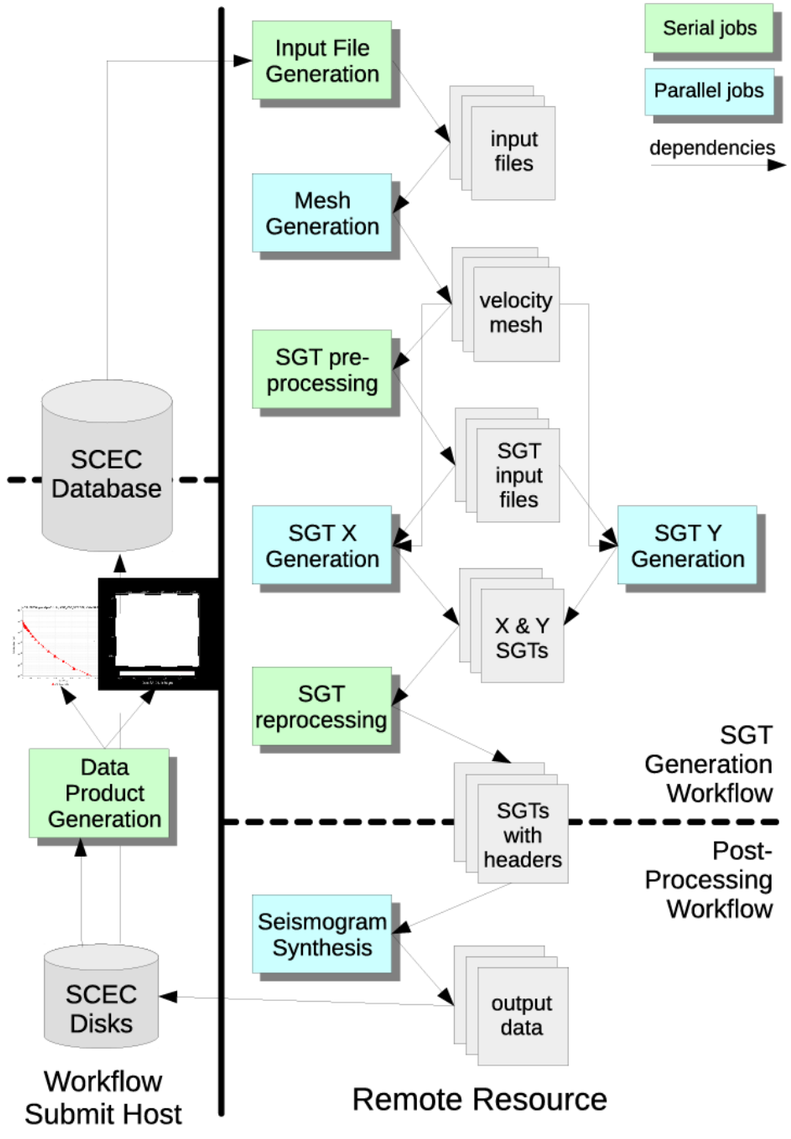
CyberShake Hazard Map



Most recent CyberShake study took 4 weeks of real time and used 21 million core-hours

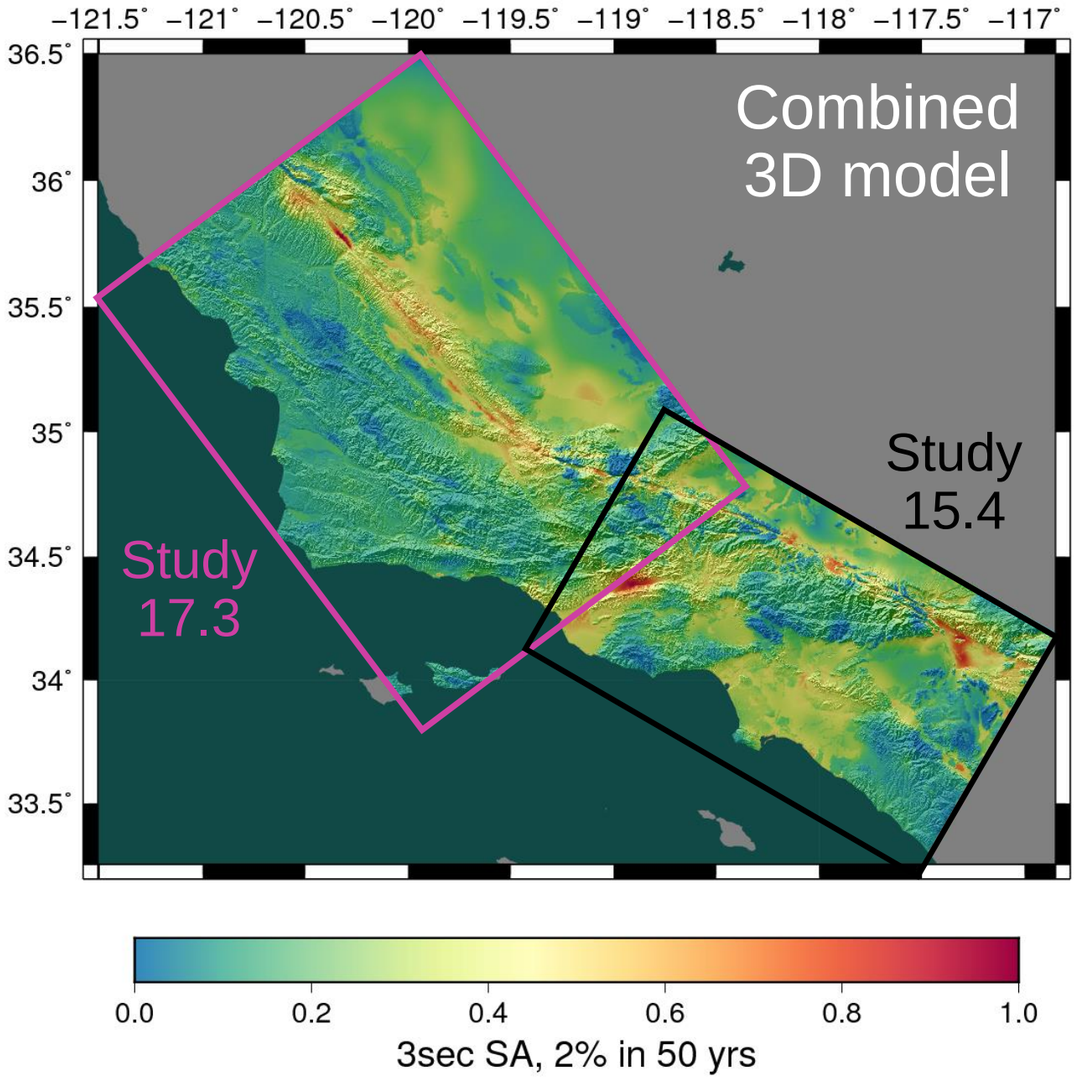
CyberShake Workflows

- Scientific workflow tools orchestrate CyberShake simulations
  - Pegasus-WMS, HTCondor, Globus Toolkit
  - Create description of workflow with files and dependencies
  - Tools then manage real-time execution of workflow
- Provide key benefits
  - Automation: supports running millions of jobs over weeks
  - Data management: files are automatically staged in and out
  - Resource provisioning: jobs submitted to multiple clusters
- Enabled SCEC to scale CyberShake since 2007
  - 9 different supercomputers
  - 100 million core-hours



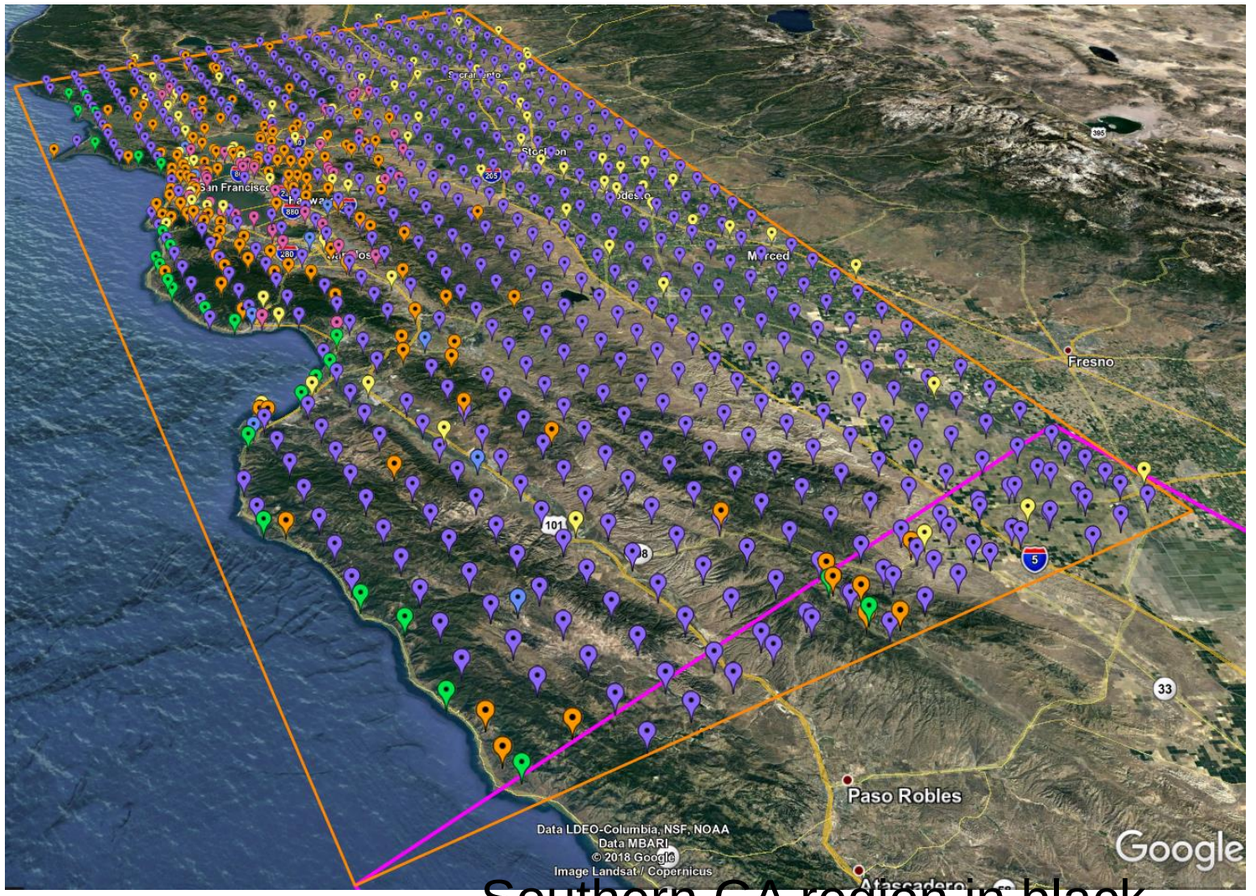
2017 Results

- Study 17.3: Central California
- 438 sites (+57) x 2 velocity models
  - Tomographically-derived 3D (CCA-06)
  - 1D average of 3D model
- Burned 21.6M core-hours on NCSA *Blue Waters* and OLCF *Titan*
- Ran 15,581 jobs using workflows
- Managed 777 TB data
  - 308 TB transferred between systems
  - 10.7 TB archived (285M seismograms)



CyberShake Study 18.5

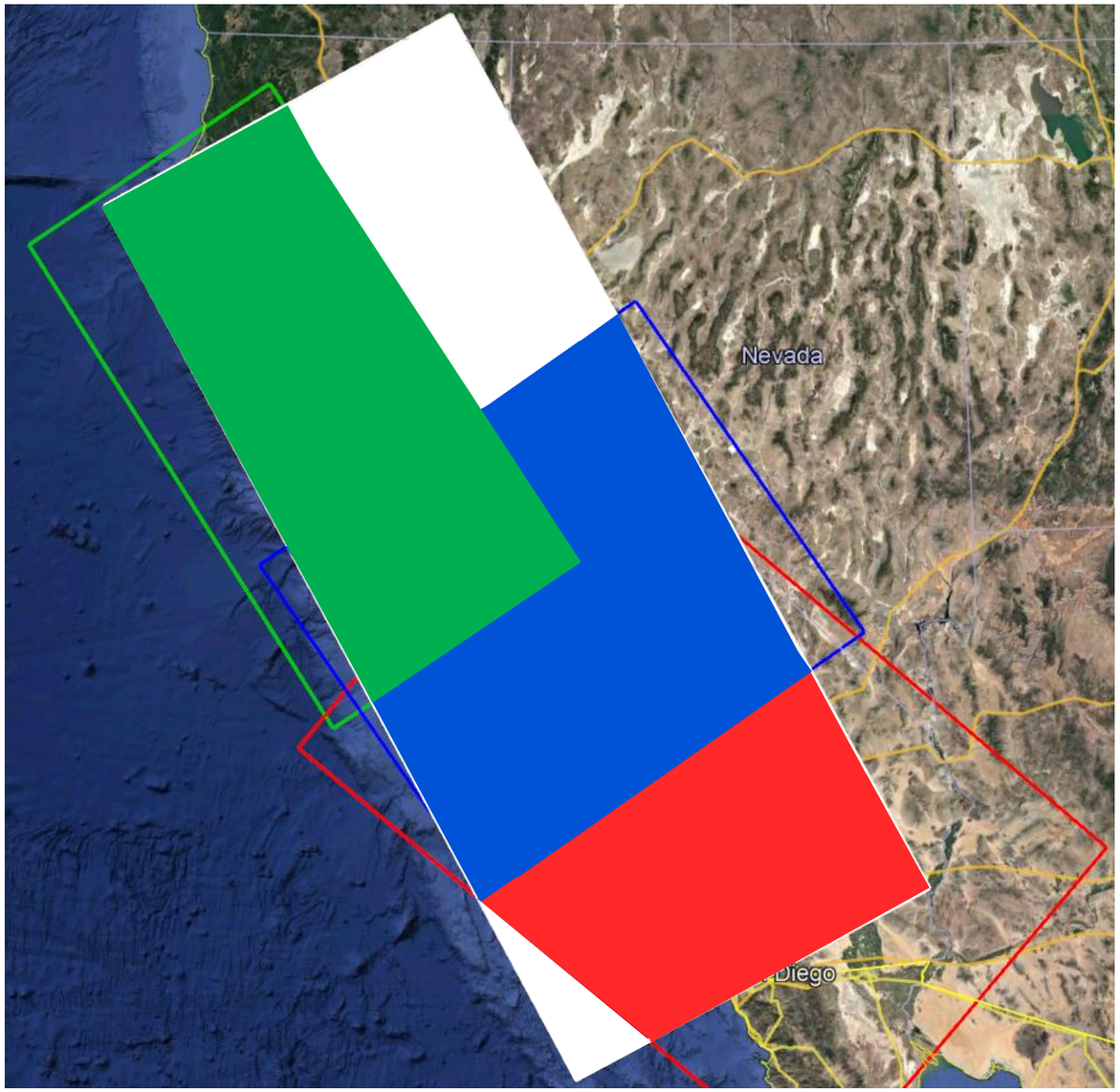
- Migrate CyberShake farther north, to greater San Francisco (“Bay Area”)
- 837 new sites + 32 for verification
- New combination of velocity models
- First physics-based PSHA results for this region
- Largest study computationally



Southern CA region in black
869 sites, densest near San Francisco
Central CA region in magenta
Bay Area region in orange

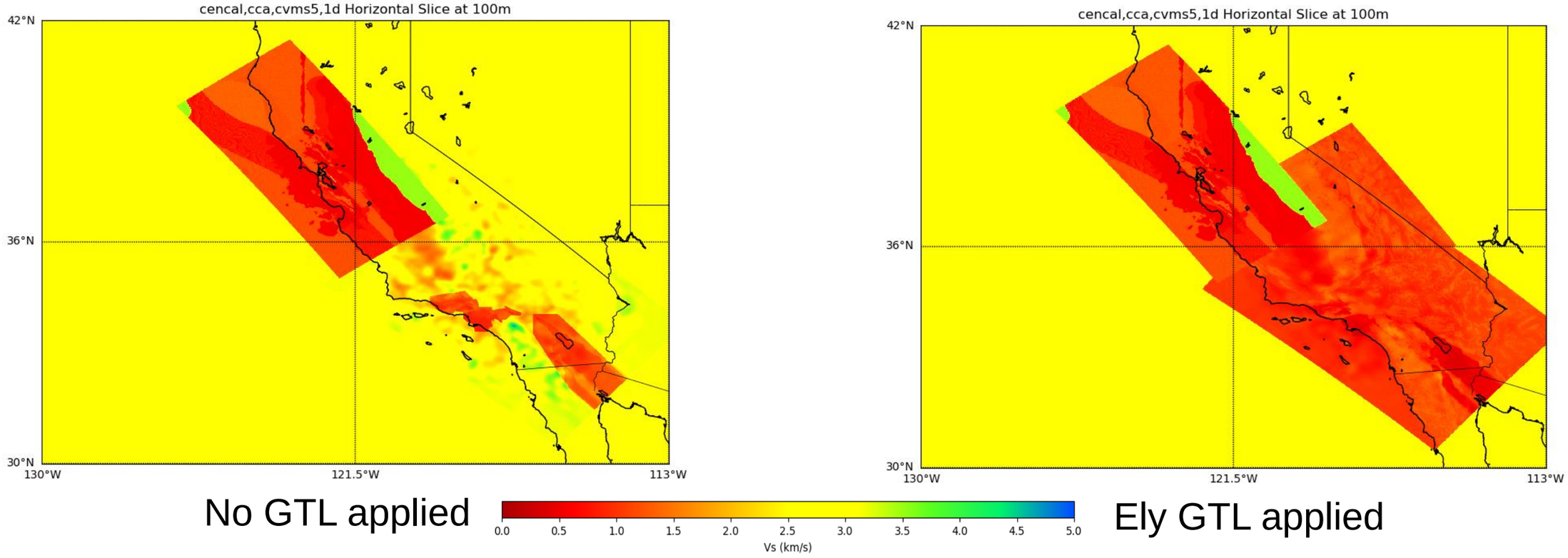
Velocity Model

- No single model large enough for statewide volume
- Must stitch together models
  - USGS Bay Area (green)
  - CCA-06 (blue)
  - CVM-S4.26 (red)
  - 1D background model (white)
- Apply smoothing along model interfaces



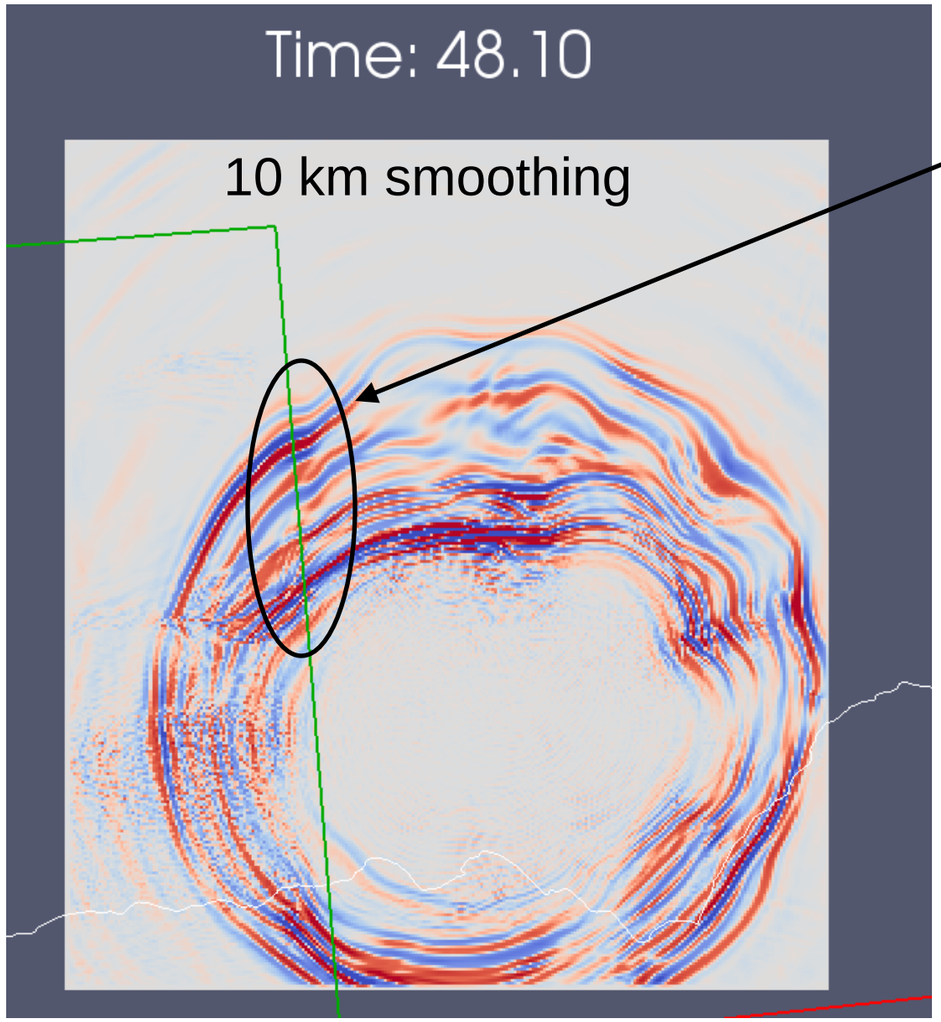
Geotechnical Layer

- Study 17.3 used Vs min=900 m/s due to tomography with CCA-06
- Plan to perform Study 18.5 at V_S min=500 m/s
- Added Vs30-derived GTL (from Wills (2015))



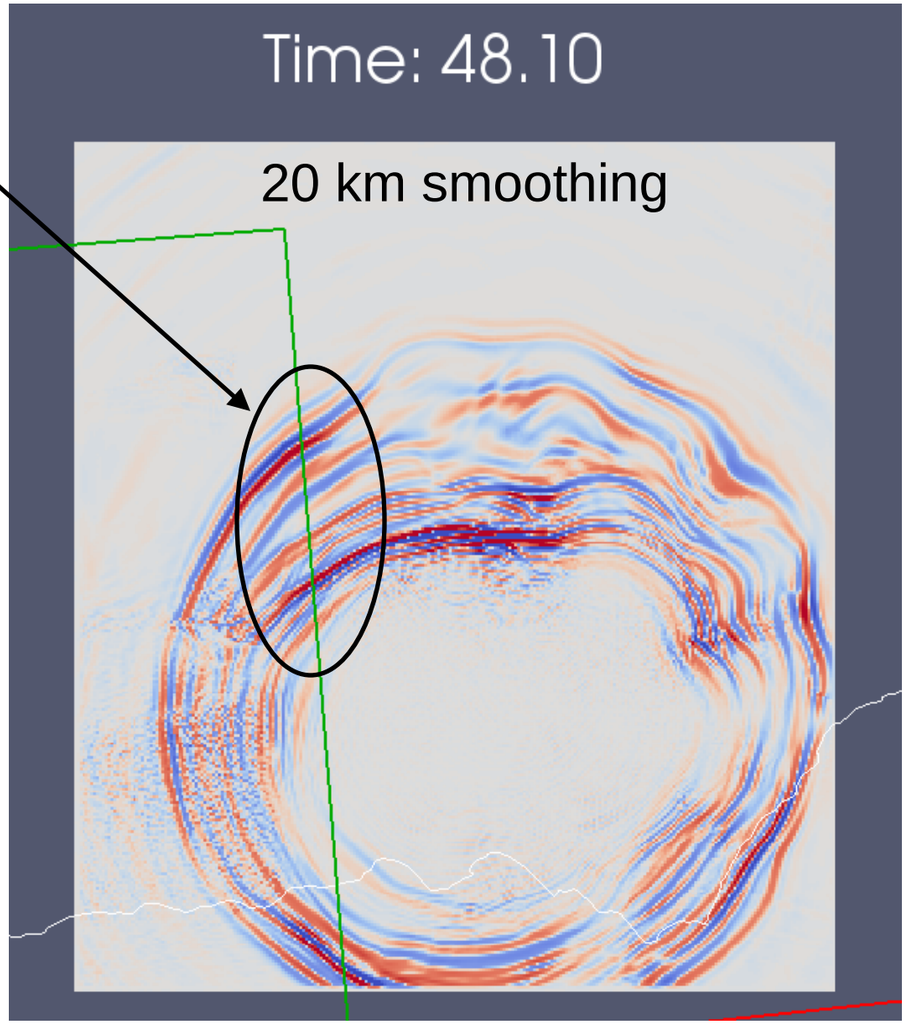
Velocity Model Verification

- Performed forward simulations near model interface to test smoothing



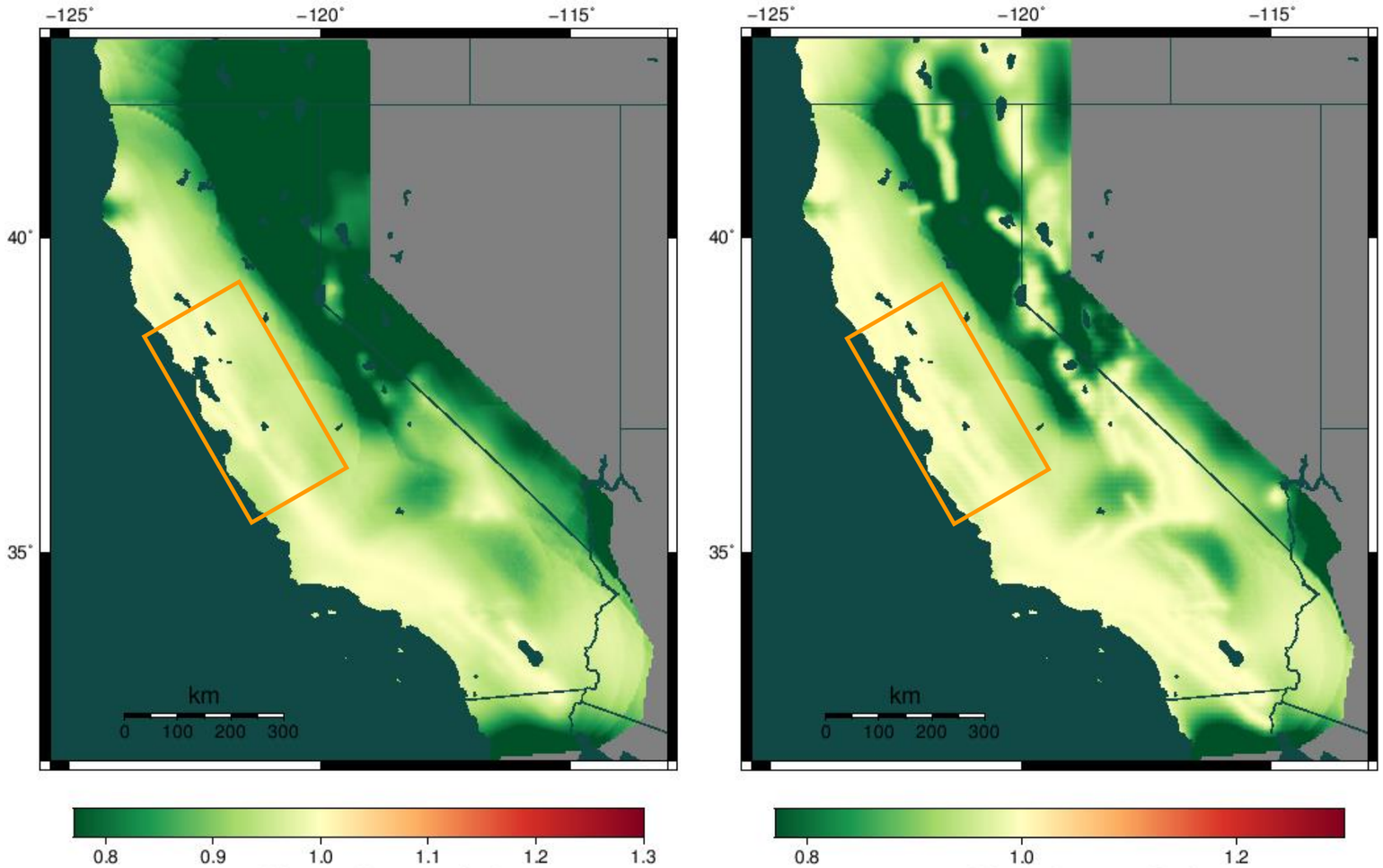
Wavefront encounters smoothed region

20 km smoothing shows reduced refraction effects



Background Seismicity

- UCERF models include off-fault background seismicity
- In previous CyberShake regions can be ignored
- UCERF2 GMPEs show impact of up to 10% on eastern edge of region
- Planning to add background seismicity to Study 18.5

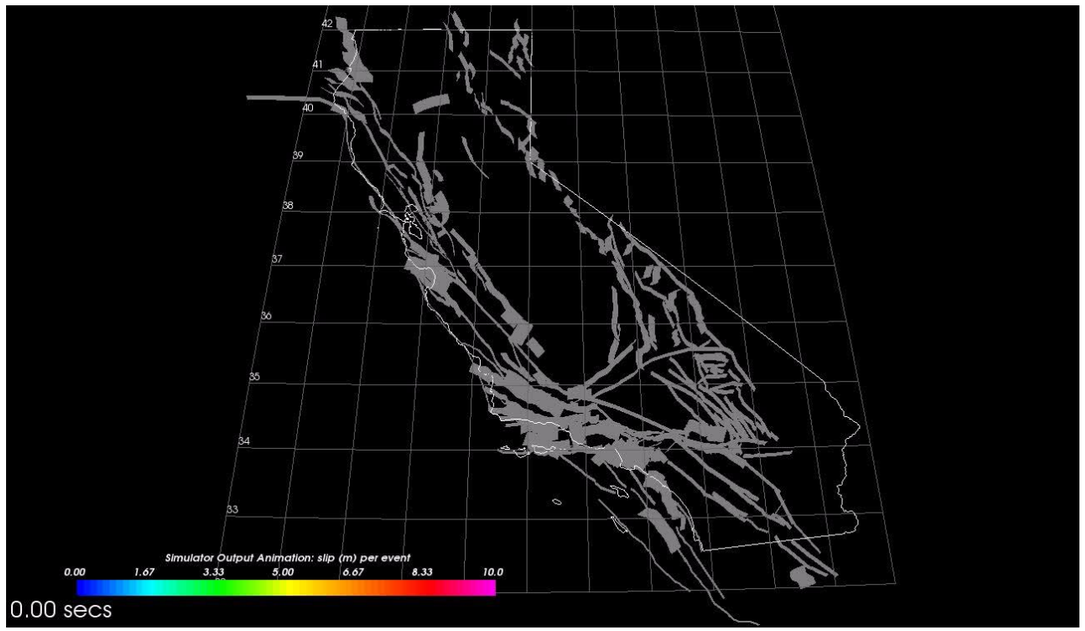


Study 18.5 Requirements

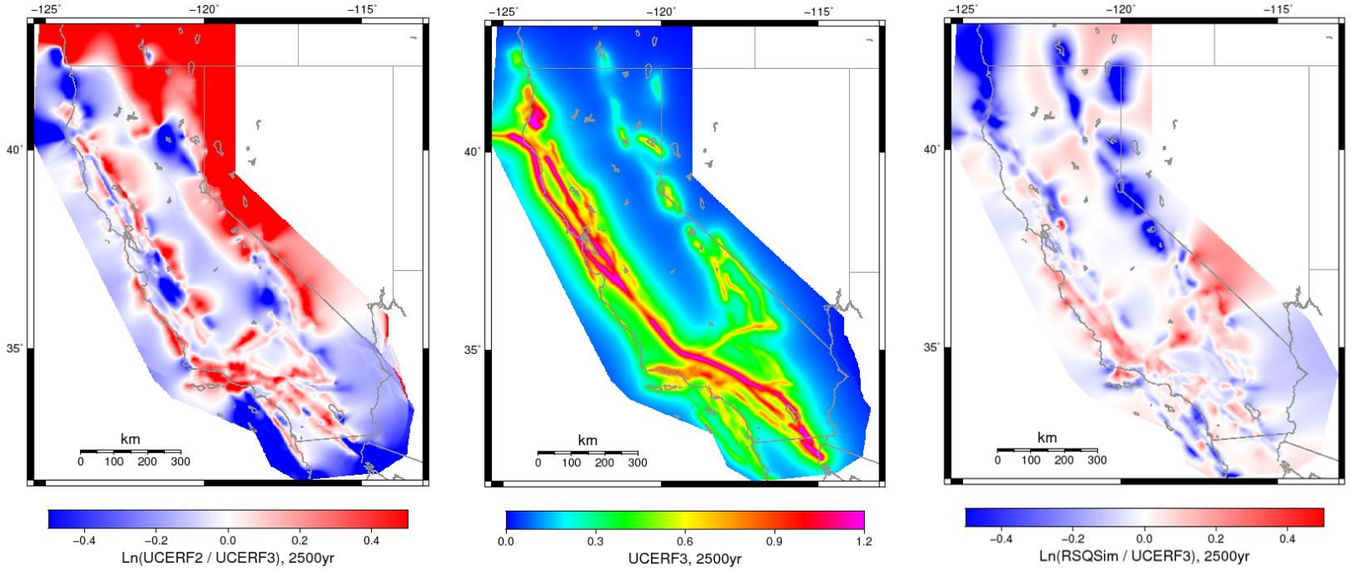
- 869 sites
- 1 Hz, V_S min=500 m/s, UCERF 2 ERF
- ~80M core-hours, divided between Titan and Blue Waters
  - GPUs for SGT generation
  - CPUs for mesh generation, reciprocity calculations
- 750 TB intermediate data
- 12.6 TB output data products
- Planning to start in 2 weeks

RSQSim

- Updating ERF from UCERF2
  - UCERF3 difficult for CyberShake
- Alternative: use RSQSim to generate long seismicity catalogs
  - Based on rate-and-state friction
  - Similar ERF to UCERF3



UCERF2 | UCERF3 | RSQSim

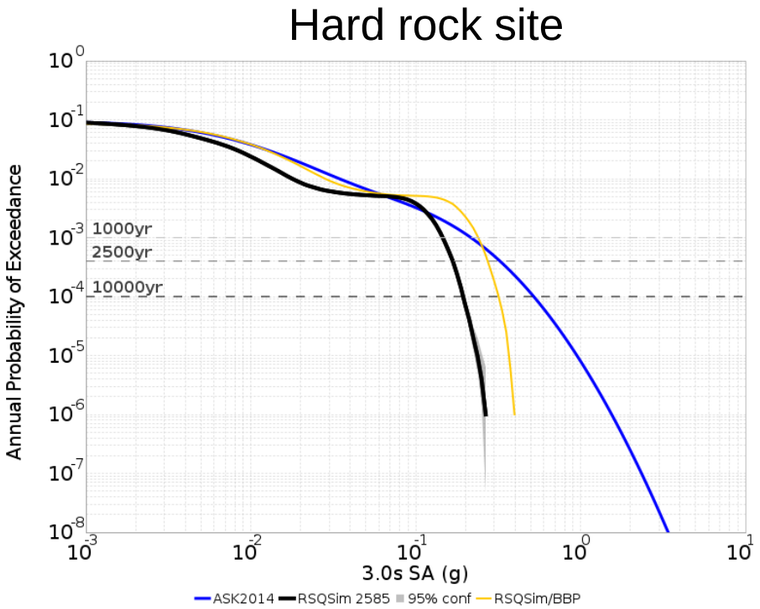


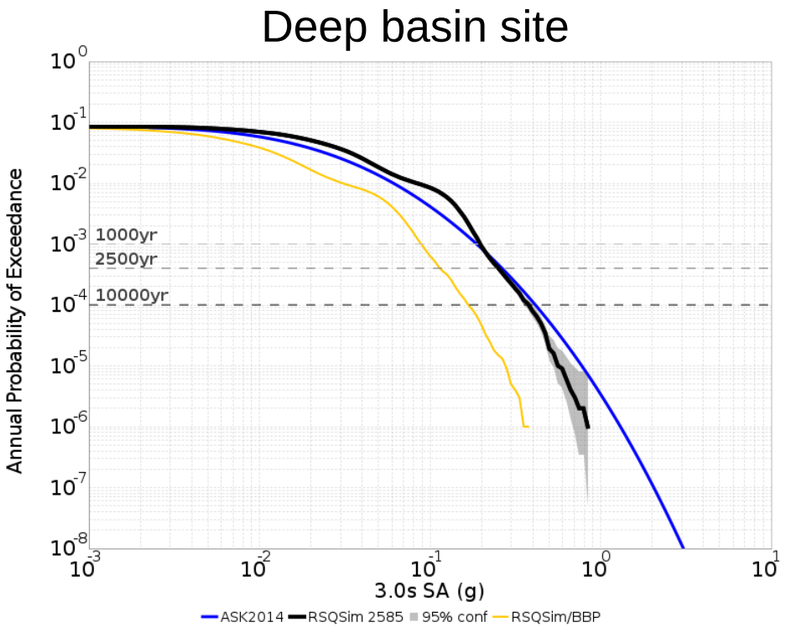
PGA (g), 2% in 50 yrs

- RSQSim events fed directly into CyberShake
  - Slip time functions produced, no need for generator
- Fully physics-based PSHA
- Undergoing extensive validation

RSQSim Results

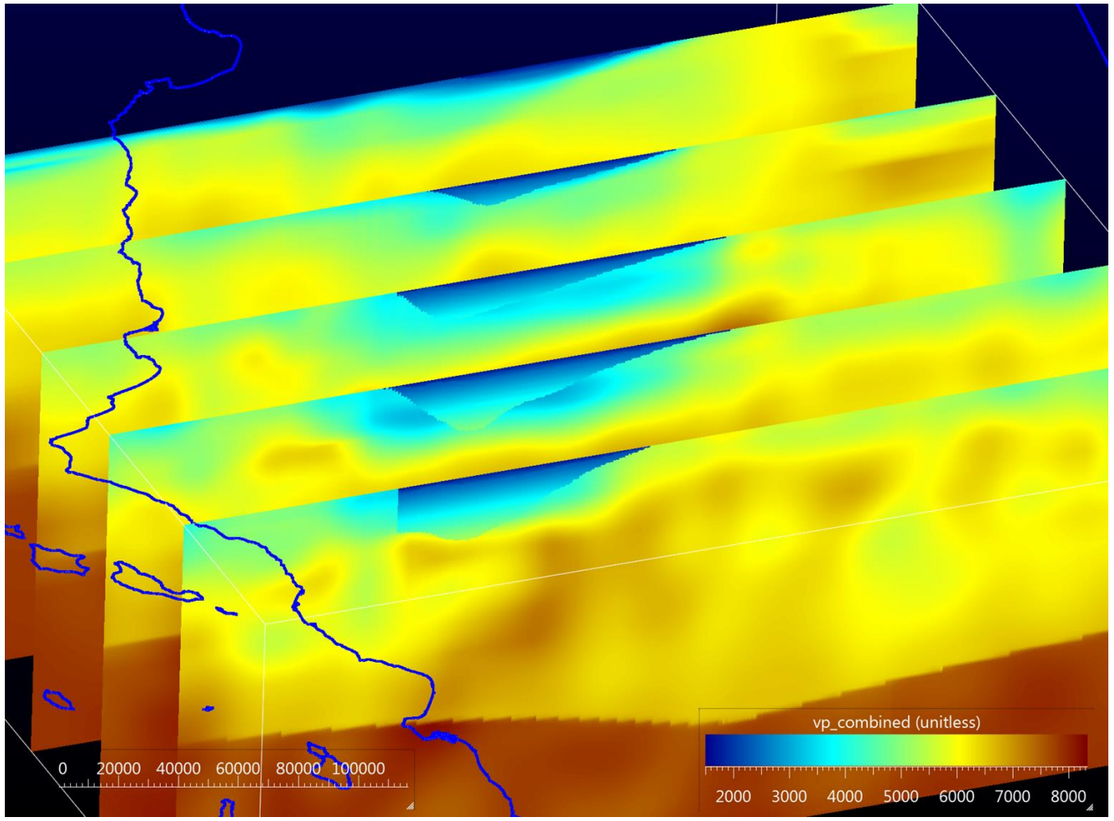
- Hazard curves
  - RSQSim/CyberShake (3D)
  - RSQSim/GMPE (ASK 2014)
  - RSQSim/SCEC Broadband Platform (1D)
- Curves less smooth due to less variability for individual sources
  - About 20% as many events
- Simulation curves are truncated
  - 1 million year catalog: minimum probability of 10^{-6}
- Working to improve match with observed rupture velocities using SCEC BBP



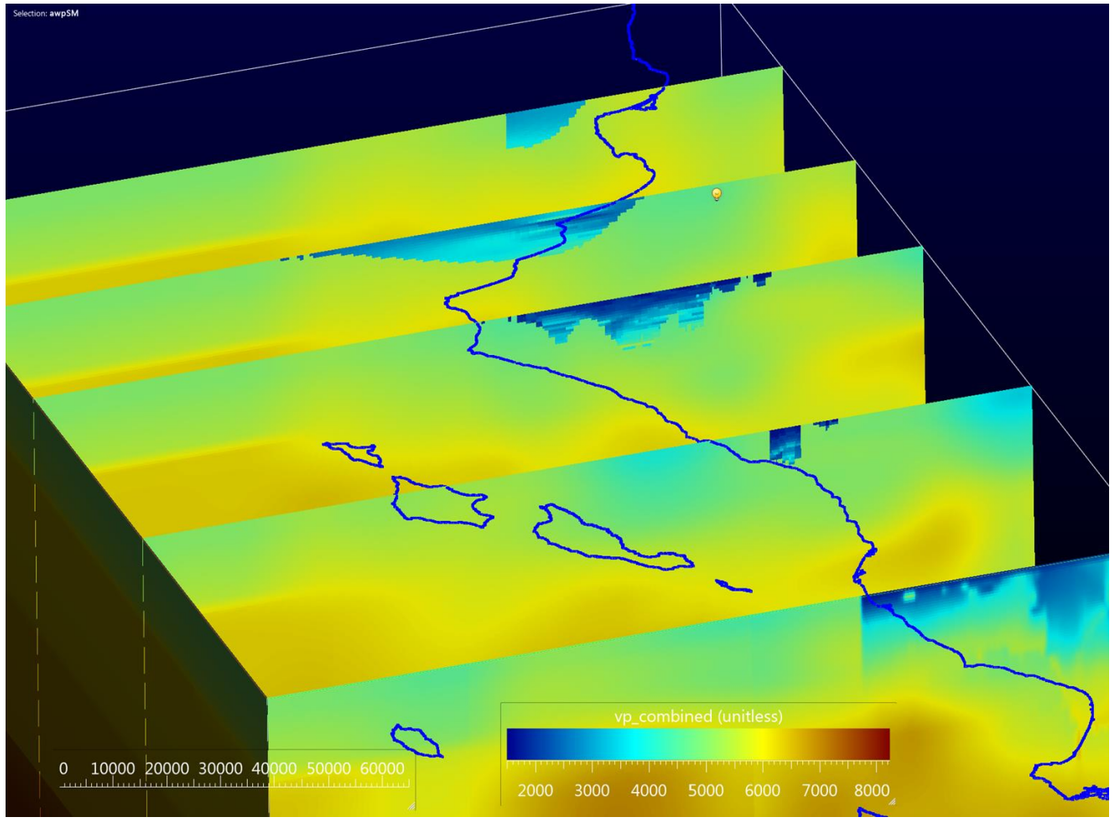


Improved Central California Basins

- Integrating improved basins from Shaw and Plesch into CCA-06
- Will rerun Study 17.3 with improved model



San Joaquin basin



Santa Maria basin

On the horizon

- Discontinuous mesh implementation of SGT code
  - Current version uses regular mesh
  - Performance improvement
  - Includes frequency-dependent Q, plasticity
- Nonlinearity
  - Breaks reciprocity: will require changes to CyberShake
  - Exploring multiple paths forward
    - Use nonlinear approximations in reciprocity
    - Perform forward simulation of small subset of events, use reciprocity for others
      - Use machine learning to identify which events should be forward-simulated
    - Create ‘equivalent kinematic sources’, which reproduce nonlinear effects

Questions?

SCEC
AN NSF+USGS CENTER

SDSC
SAN DIEGO SUPERCOMPUTER CENTER

U.S. DEPARTMENT OF ENERGY

USGS
science for a changing world

BLUE WATERS
SUSTAINED PETASCALE COMPUTING

NCSA

TACC

ISI
Information Sciences Institute

OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

U.S. DEPARTMENT OF ENERGY
INCITE
LEADERSHIP COMPUTING

National Science Foundation

USCGRID

Pegasus