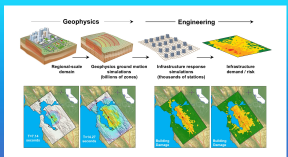


PEER International Pacific Rim Forum
June 16-17, 2021

Verification and Validation of Earthquake Ground Motion Simulation in the San Francisco Bay Area with SW4

Arthur Rodgers
Lawrence Livermore National Laboratory

June 17, 2021

Seismology
Lawrence Livermore National Laboratory

ECP
EXASCALE COMPUTING PROJECT

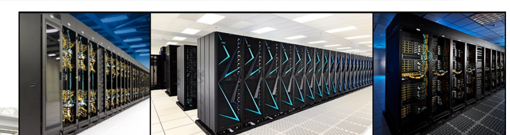
PEER

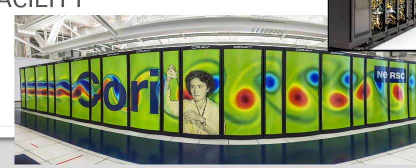
LLNL-PRES-823639

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Thanks to many coworkers over the years

- Anders Petersson (LLNL, CASC)
- Bjorn Sjogreen (LLNL, CASC)
- Arben Pitarka (LLNL, AEED)
- Ramesh Pankajakshan (LLNL, LC)
- David McCallen, Floriana Petrone, Maha Kenawy (UNR, LBNL)
- Houjun Tang (LBNL)
- Norm Abrahamson (UCB & UCD)
- Albert Kottke (PG&E)
- Rob Graves, Brad Aagaard (USGS)
- **Computing Facilities**
  - Livermore Computing (LLNL)
  - OLCF (Oak Ridge Lab)
  - NERSC (LBNL)





3D Seismic wave propagation simulations are complex, require verification and validation

- Ground motions depend on several factors
  - simulations must all of these get correct
- V&V are crucial to demonstrating accuracy and confidence
  - Key to acceptance of simulated ground motions

Verification

- Numerical method must be accurate
- Computer code must solve the algorithm correctly

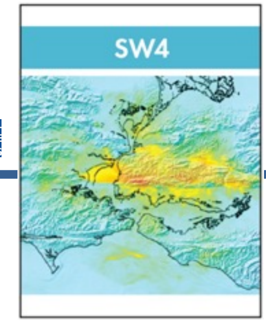
Related topic: Performance benchmarking

Validation

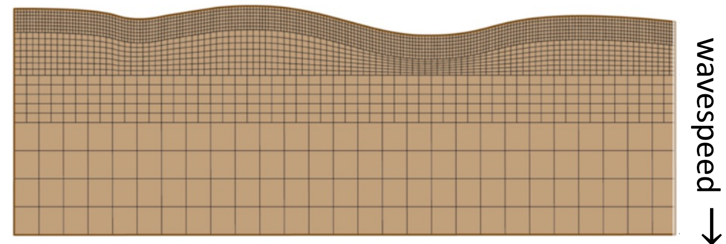
- Inputs must be accurate and physically meaningful
  - Source, earthquake rupture
  - Earth model must represent true 3D structure
    - Path propagation effects: crustal, basin, topographic structure
    - Site effects: minimum shear wavespeed, linear and non-linear response

SW4: seismic wave propagation code based on the summation-by-parts FD method

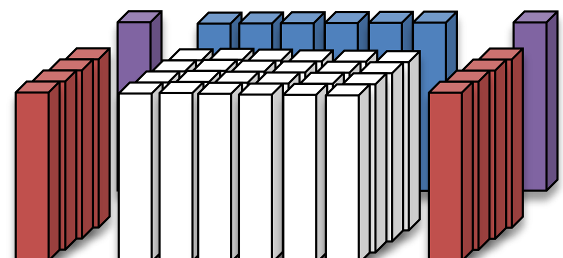
github.com/geodynamics/sw4



- Summation-by-parts FDTD
  - Node-centered, displacement formulation
    - Not velocity-stress staggered grid!
  - Accurate, provably stable & energy conserving
  - Super-grid boundary conditions

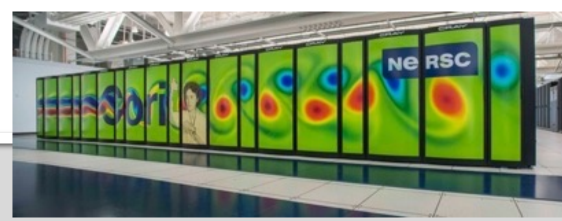


- SW4 is 4th order accurate (time & space)
  - Fully 3D material models (iso- and anisotropic)
  - Topography (curvilinear mesh)
  - Mesh refinement in Curv. & Cart. meshes
    - Accurate at boundaries & interfaces, w/ hanging nodes



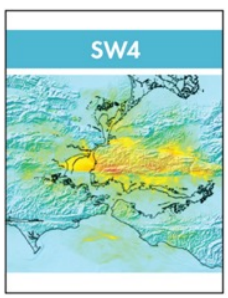
- Optimized for the hardware
  - Many core CPU's (e.g. NERSC's Cori-II)
    - Hybrid MPI/OpenMP communications
  - GPU's (e.g. Sierra & Summit)
    - RAJA directs work on GPU's

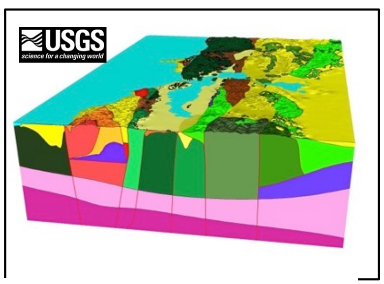
ECP EXASCALE COMPUTING PROJECT

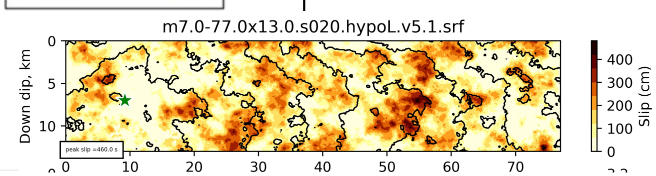


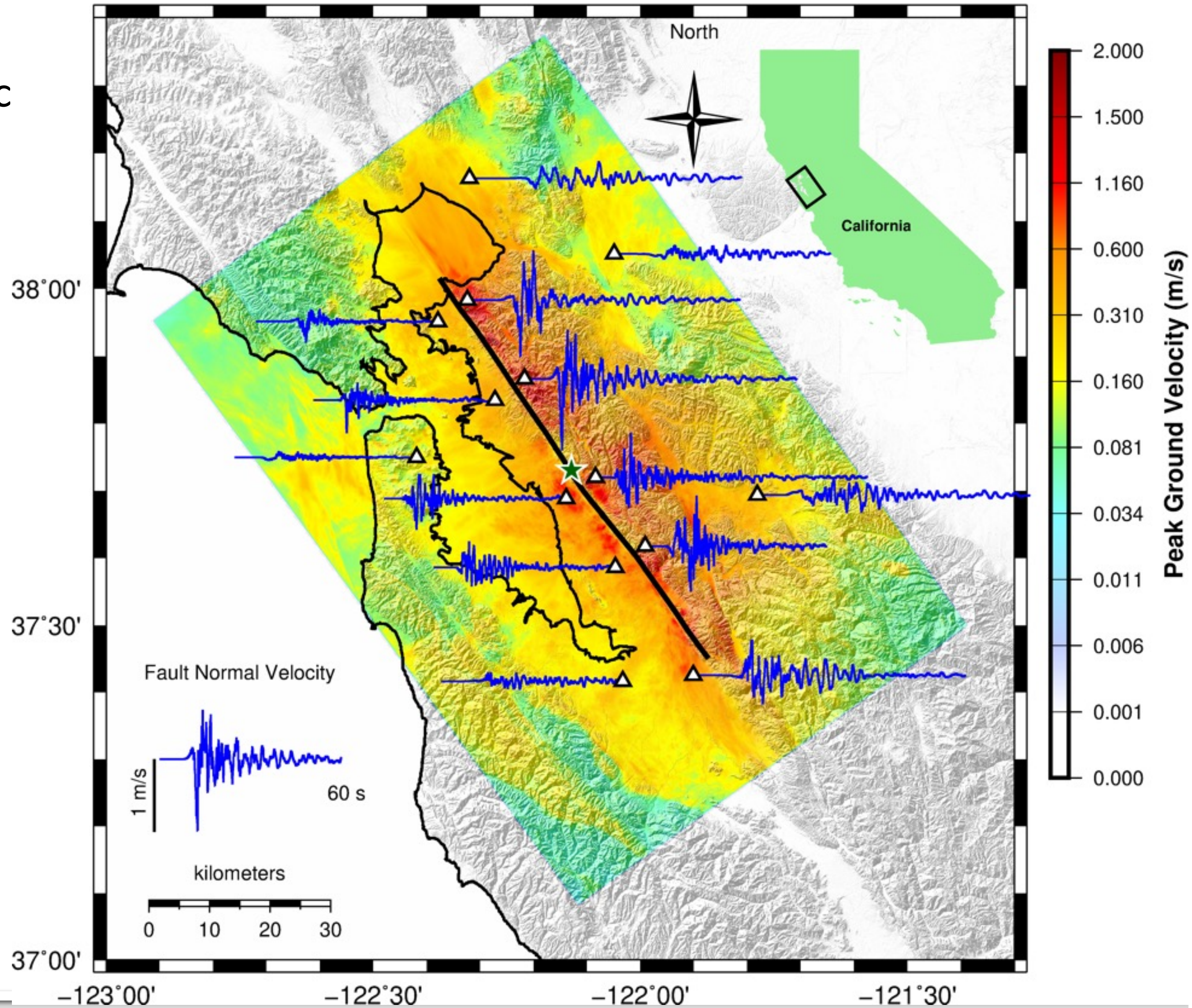
Our demonstration problem: Hayward Fault M_W 7.0 scenario earthquake

- Regional-scale
- Broadband, fully deterministic
  - f_{max} = 10 Hz @ 8 PPW
  - h_{min} = 6.25, 3.125 m
  - λ_{min} = 50, 25 m
  - v_{Smin} = 500, 250 m/s
- 3D USGS model (Jachens, Brocher, Aagaard & coworkers)
  - Surface topography
  - Anelasticity, Q
- Graves & Pitarka ruptures

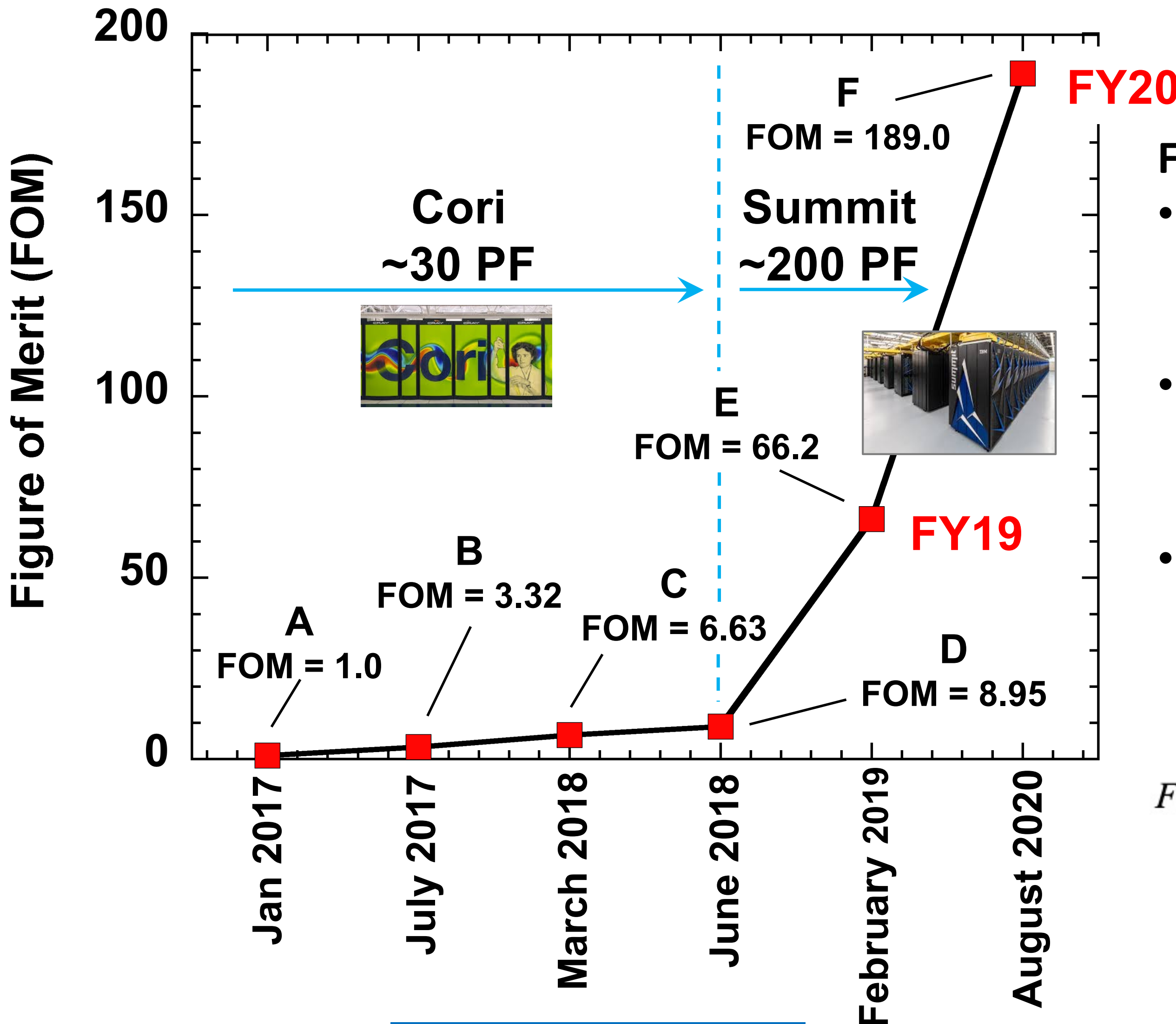








FY20 EQSIM performance evaluation for a M7 Hayward fault SFBA simulation



Vsmin = 500 m/s

Figure of Merit (FOM)

- Measure of computational efficiency for our demonstration problem
- Normalized to performance at the beginning of the project (2017)
- Maximum resolved frequency, *Freq*, increasing

$$FOM = \frac{(Freq)^4}{(Wall\ Clock\ Time \times 7.6)} \left(\frac{500}{V_{smin}}\right)^4$$

EXASCALE COMPUTING PROJECT

(McCallen et al., Eq Spectra 2020a)

Verification: the accuracy of mathematics and computation of numerical simulations

- Comparison of computed solutions (time-series) against analytic solutions or other computed solutions
- Goal: to obtain accurate and reliable simulations of seismic response and build confidence in solutions
- Criteria for comparing waveforms
  - Pointwise differences with analytic or 1D semi-analytic solution (e.g. reflectivity)
  - Anderson (2004) Goodness-of-fit score of waveform measurements
  - Kristekova et al. (2006) time-frequency phase and envelope misfit
    - wavelet-based decomposition
- A few notable examples:
  - S. M. Day, J. Bielak, D. Dreger, S. Larsen, R. Graves, A. Pitarka, and K. B. Olsen (2003). Test of 3D elastodynamic codes: Lifelines program task 1A02. Technical report, PEER & SCEC.
  - Moczo et al. (2006). Comparison of Numerical Methods for Seismic Wave Propagation and Source Dynamics - the SPICE Code Validation, ESG2006, Grenoble, France
  - Bielak et al. (GJI, 2010) Compares 3 SCEC ShakeOut simulations
  - Chaljub et al. (GJI, 2015) Mygdonian Basin, Greece: stringent methods

SW4 uses the method of manufactured solutions for accuracy & convergence: "twilight mode"

Elastodynamic equations of motion in 3D

$$\rho \mathbf{u}_{tt} = \nabla \cdot \mathcal{T} + \mathbf{F}(\mathbf{x}, t), \quad \mathbf{x} \text{ in } \Omega,\ 0 \le t \le T,$$
$$\mathbf{u}(\mathbf{x}, 0) = 0, \quad \mathbf{u}_t(\mathbf{x}, 0) = 0, \quad \mathbf{x} \text{ in } \Omega.$$

A.1 Method of manufactured solutions

The method of manufactured solutions provides a general way of testing the accuracy of numerical solutions of partial differential equations, including effects of heterogeneous material properties and various boundary conditions on complex geometries. The test scripts can be found in the directory

```
.../sw4/examples/twilight
```

In these tests, we take the material properties to be

$$\begin{aligned}
\rho(x,y,z) &= A_\rho \left(2 + \sin(\omega_m x + \theta_m)\cos(\omega_m y + \theta_m)\sin(\omega_m z + \theta_m)\right), \\
\mu(x,y,z) &= A_\mu \left(3 + \cos(\omega_m x + \theta_m)\sin(\omega_m y + \theta_m)\sin(\omega_m z + \theta_m)\right), \\
\lambda(x,y,z) &= A_\lambda \left(2 + \sin(\omega_m x + \theta_m)\sin(\omega_m y + \theta_m)\cos(\omega_m z + \theta_m)\right).
\end{aligned}$$

} 3D Earth model

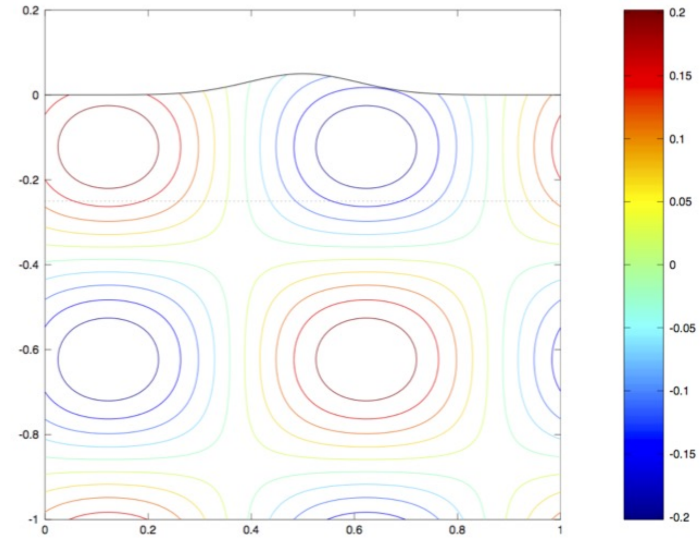
The internal forcing, boundary forcing and initial conditions are chosen such that the exact (manufactured) solution becomes[1]

Chosen response

$$\begin{aligned}
u_e(x,y,z,t) &= \sin(\omega(x - c_e t))\sin(\omega y + \theta)\sin(\omega z + \theta), \\
v_e(x,y,z,t) &= \sin(\omega x + \theta)\sin(\omega(y - c_e t))\sin(\omega z + \theta), \\
w_e(x,y,z,t) &= \sin(\omega x + \theta)\sin(\omega y + \theta)\sin(\omega(z - c_e t)).
\end{aligned}$$

The values of the material parameters (ω_m, θ_m, A_ρ, A_λ, A_μ) and the solution parameters (ω, θ, c_e), can be modified in the input script. Since the exact solution is know, it is possible to evaluate the error in the numerical solution. By repeating the same test on several grid sizes, it is possible to establish the convergence rate of the numerical method.

Computed response



- Compare computed and analytic solutions
- Compute norm for different grid spacing
- Measure convergence

From SW4 User Guide, Petersson & Sjogreen, 2021

Recent porting of SW4 to GPU/CPU platforms requires verification of code: 0-5 Hz HF M7

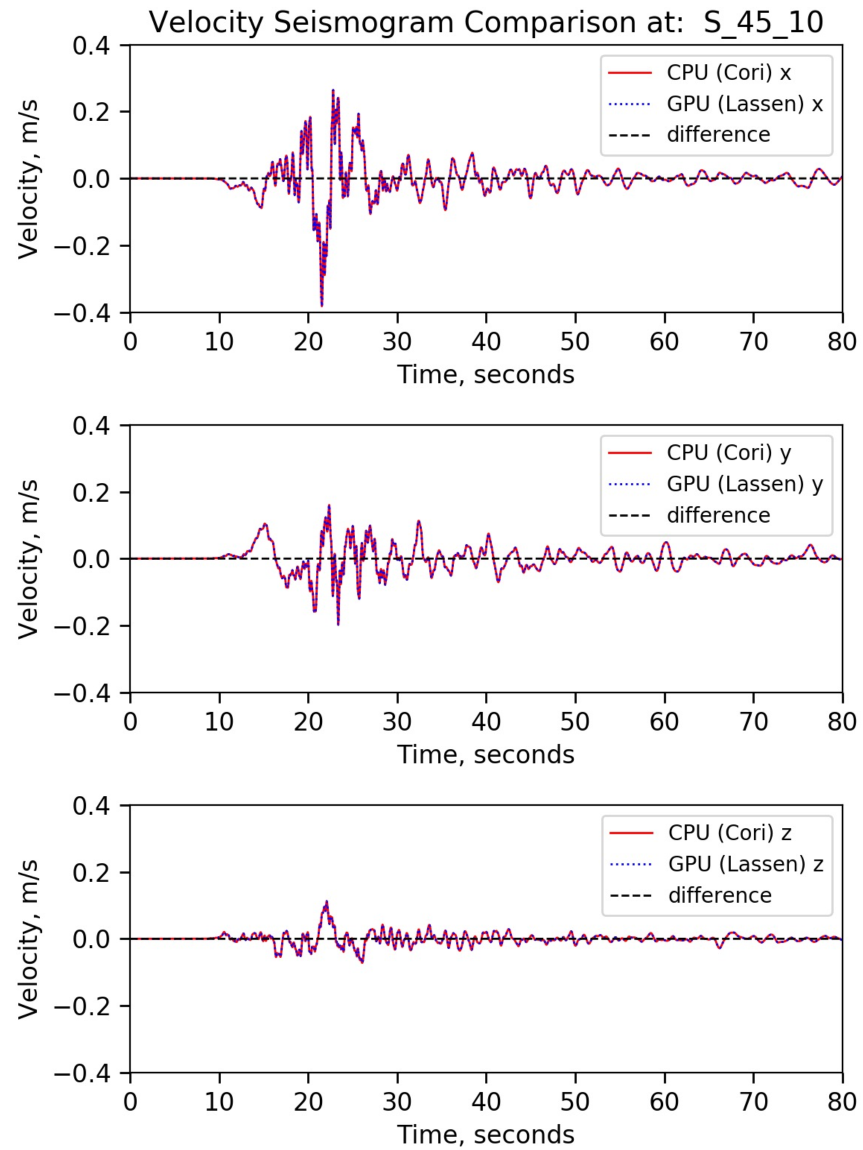
SW4 uses RAJA C++ package to manage work on GPU

SW4-RAJA uses the same source code as SW4 and a machine-specific profile to know how to offload compute intensive loops to GPU

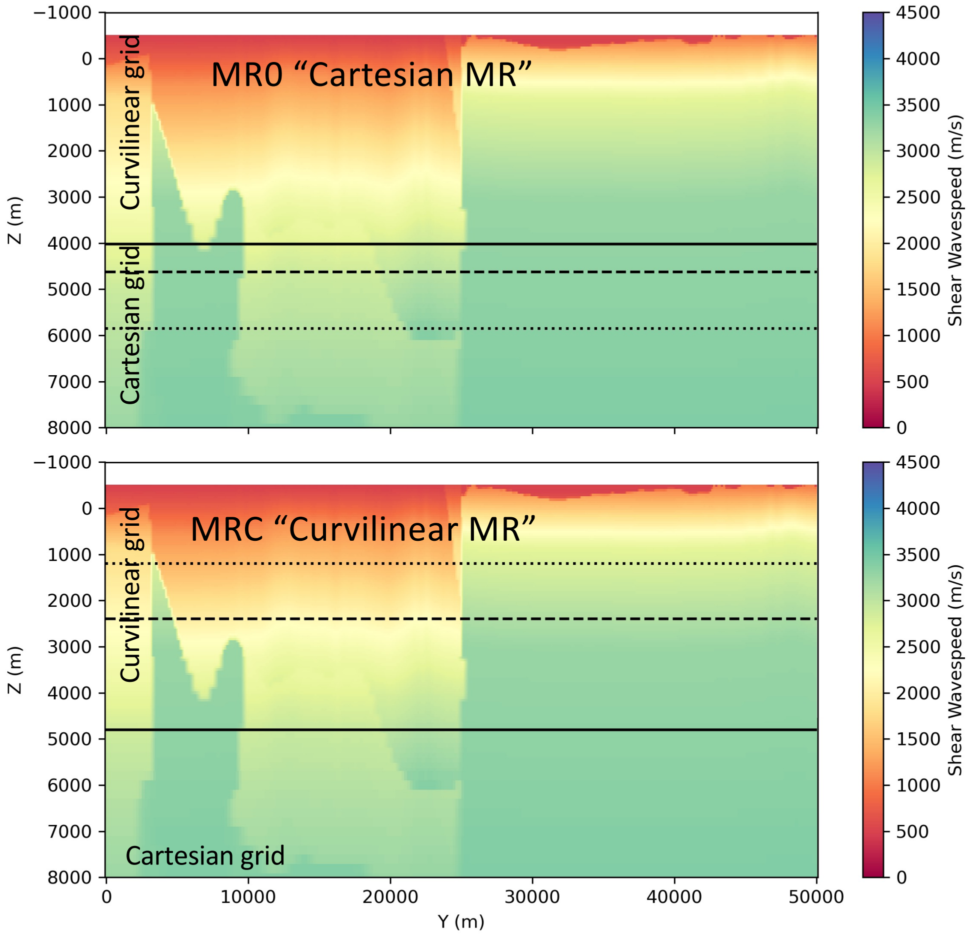
Hayward Fault, M_W 7.0, resolved 0-5 Hz

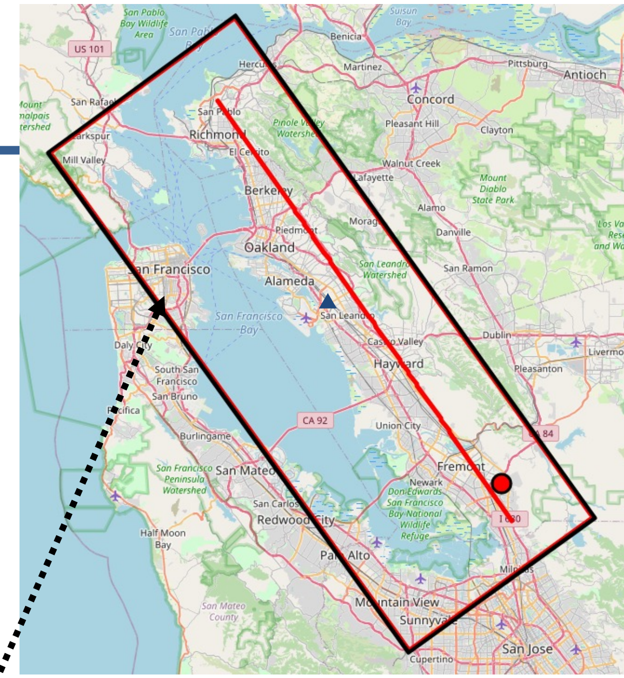
Use point-wise differences

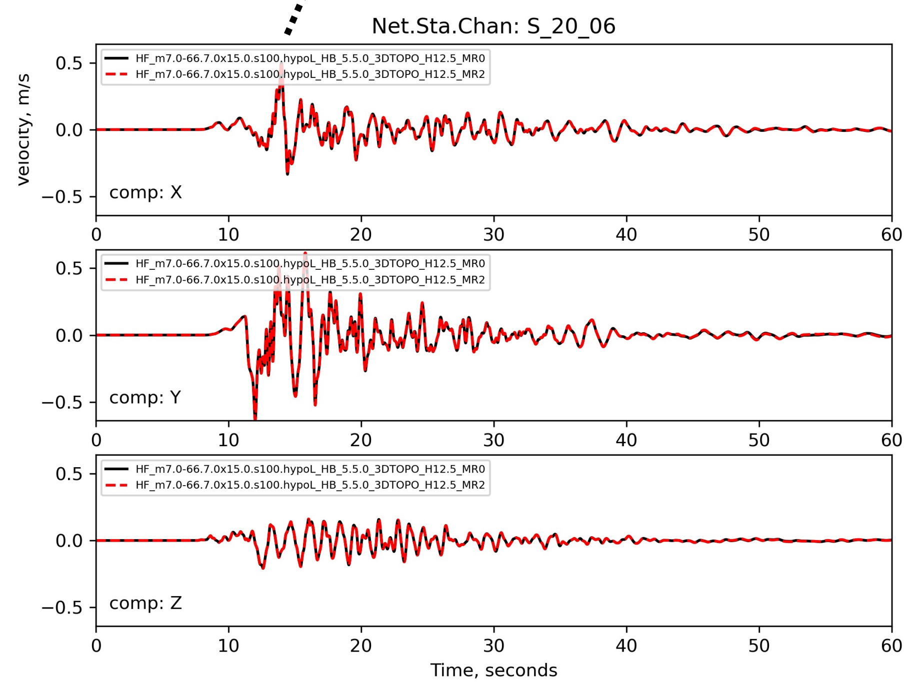
- We get excellent agreement
- Waveforms agree to 10^{-7}
- 3-component waveforms at 2301 sites



Verfication of mesh refinement cases







Run	MR Type	Npoints ratio	Run time ratio
MR0	Cart	1.0	1.0
MRC	Curv	0.386	0.44

Validation: the evaluation of the physical accuracy of numerical simulation

- Comparison of computed solutions against empirical data
  - Tests inputs: source and Earth model with a verified code
- Goal: to ensure that simulation predictions are realistic and consistent with empirical observations, to build confidence
- Moderate (M_W 3-5) earthquakes provide data sets for testing 3D Earth models in California
  - A few examples:
    - Rodgers et al., (BSSA, 2008) SFBA moderate events, waveforms
    - Kim et al. (BSSA, 2021) SFBA moderate events, intensities
    - Olsen & Mayhew (SRL, 2010) 2008 Chino Hills GOF
    - Taborda & Bielak (BSSA, 2013) 2008 Chino Hills
    - Hirakawa and Aagaard (SSA, 2021) update(s) of USGS SFBA model
- Large scenario events without empirical data
  - Compare with ground motion models (GMM's, GMPE's) or data from similar events

Earlier evaluation of the USGS 3D model using moderate events & long-periods (33-4 sec.)

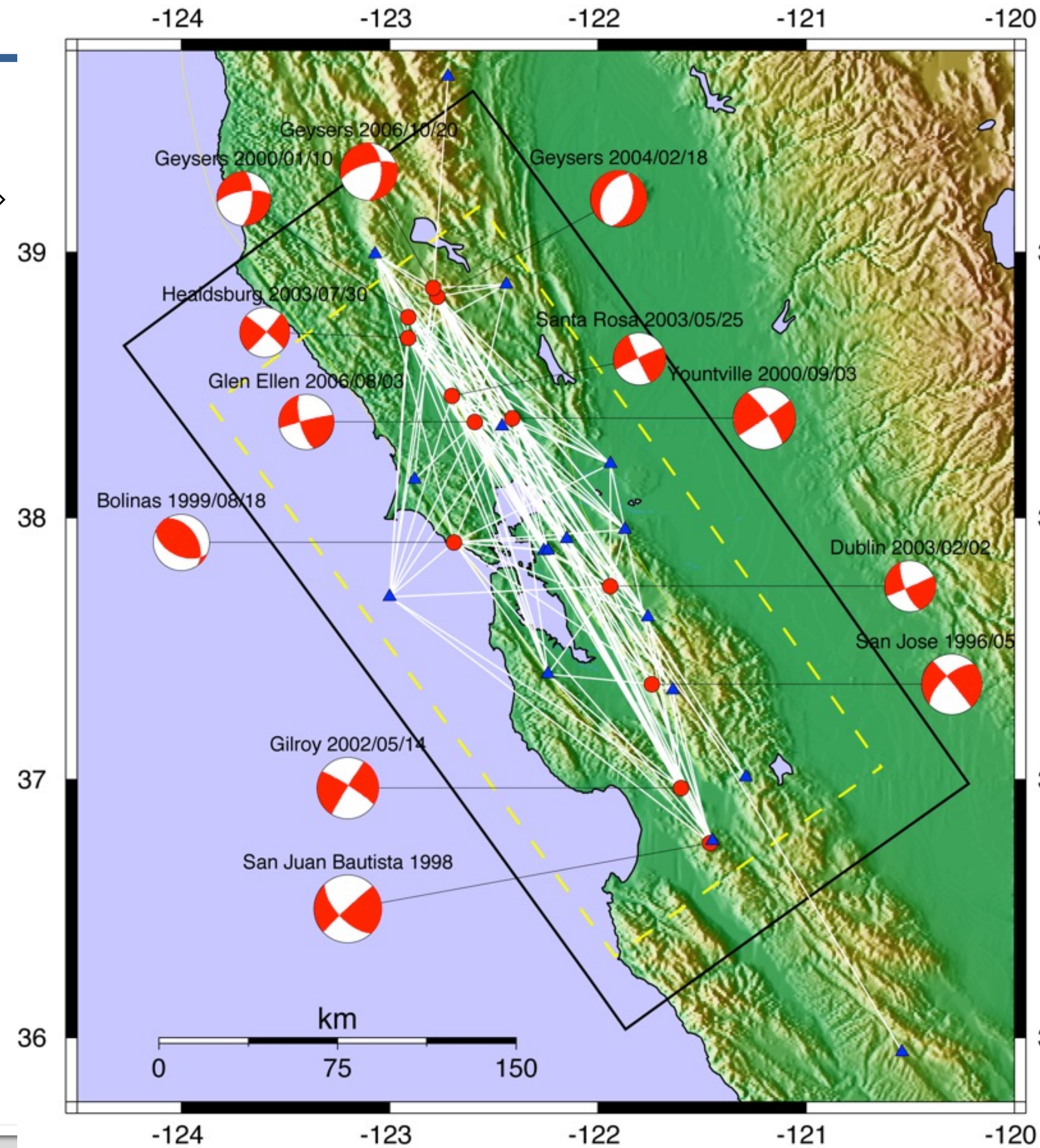
Events, BB (BK) Stations & Paths

Moderate (M_W 4-5)

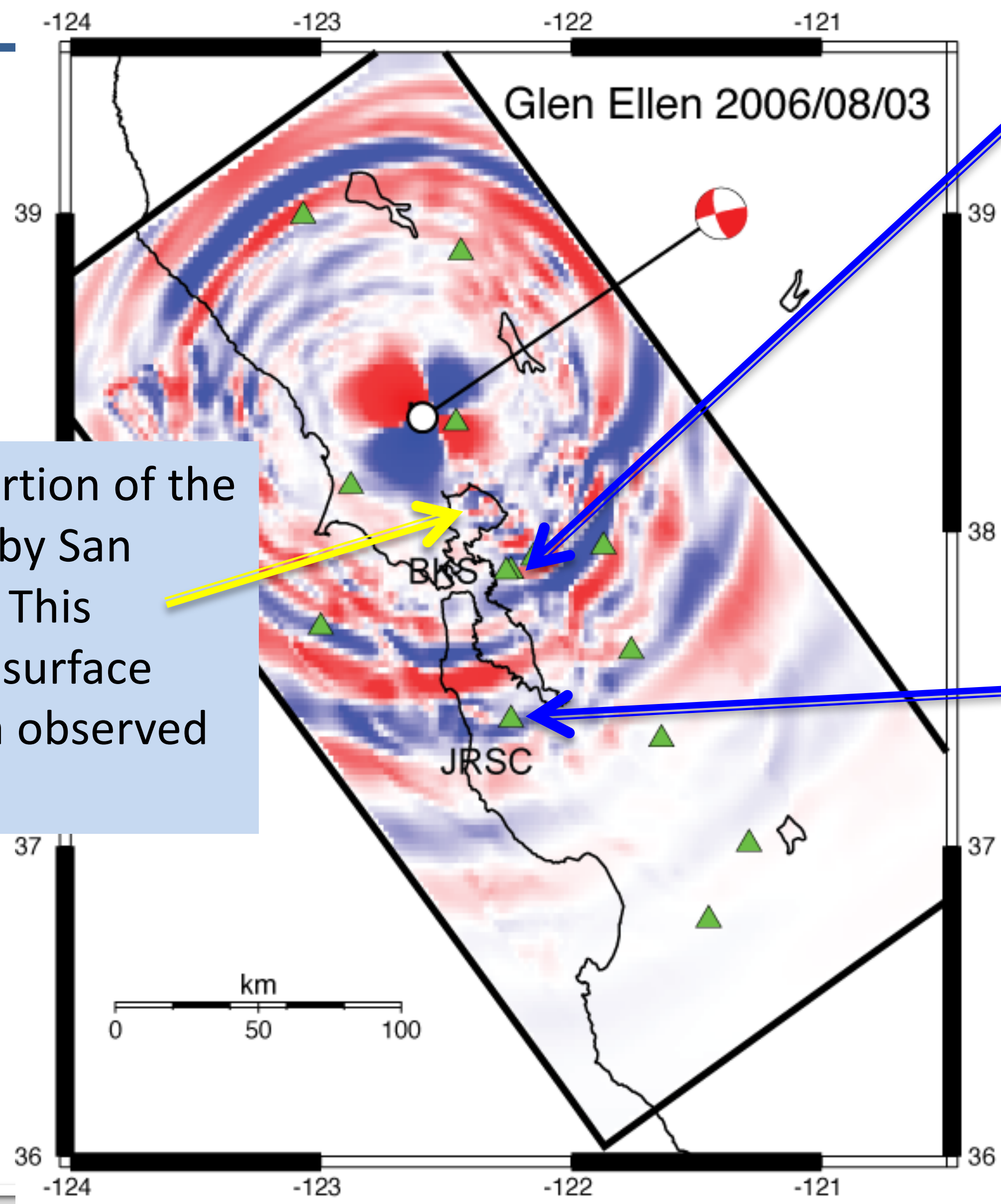
12 events (circles)
15 stations (triangles)

Coverage uneven, many paths along Hayward Fault

Rodgers et al. (BSSA, 2008)

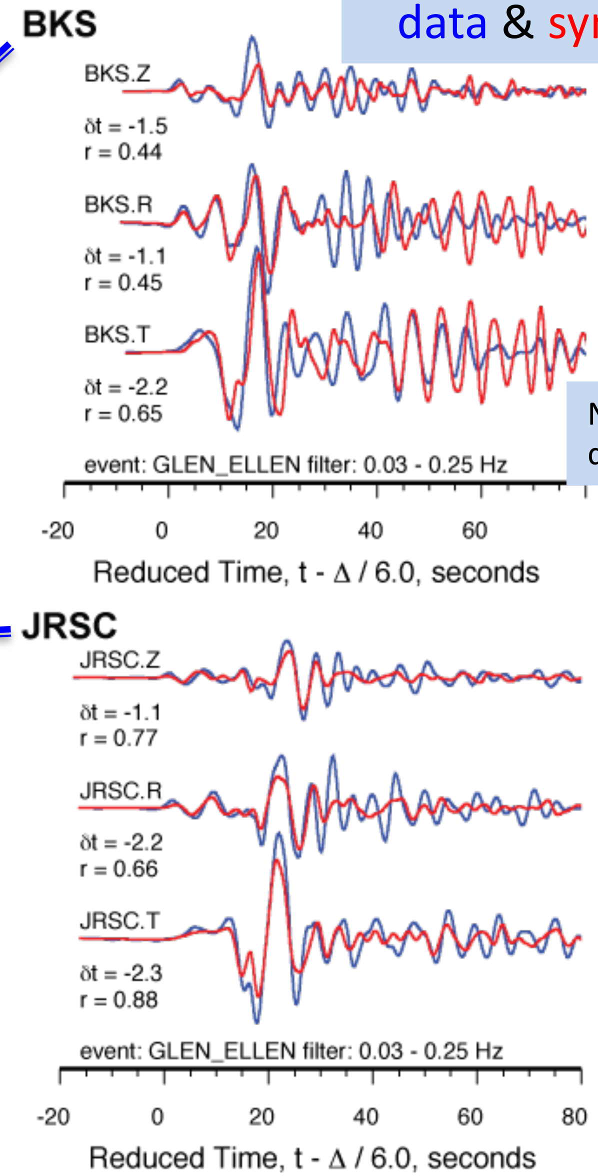


Long-period waveform comparisons for August 3 2006, Glen Ellen Earthquake



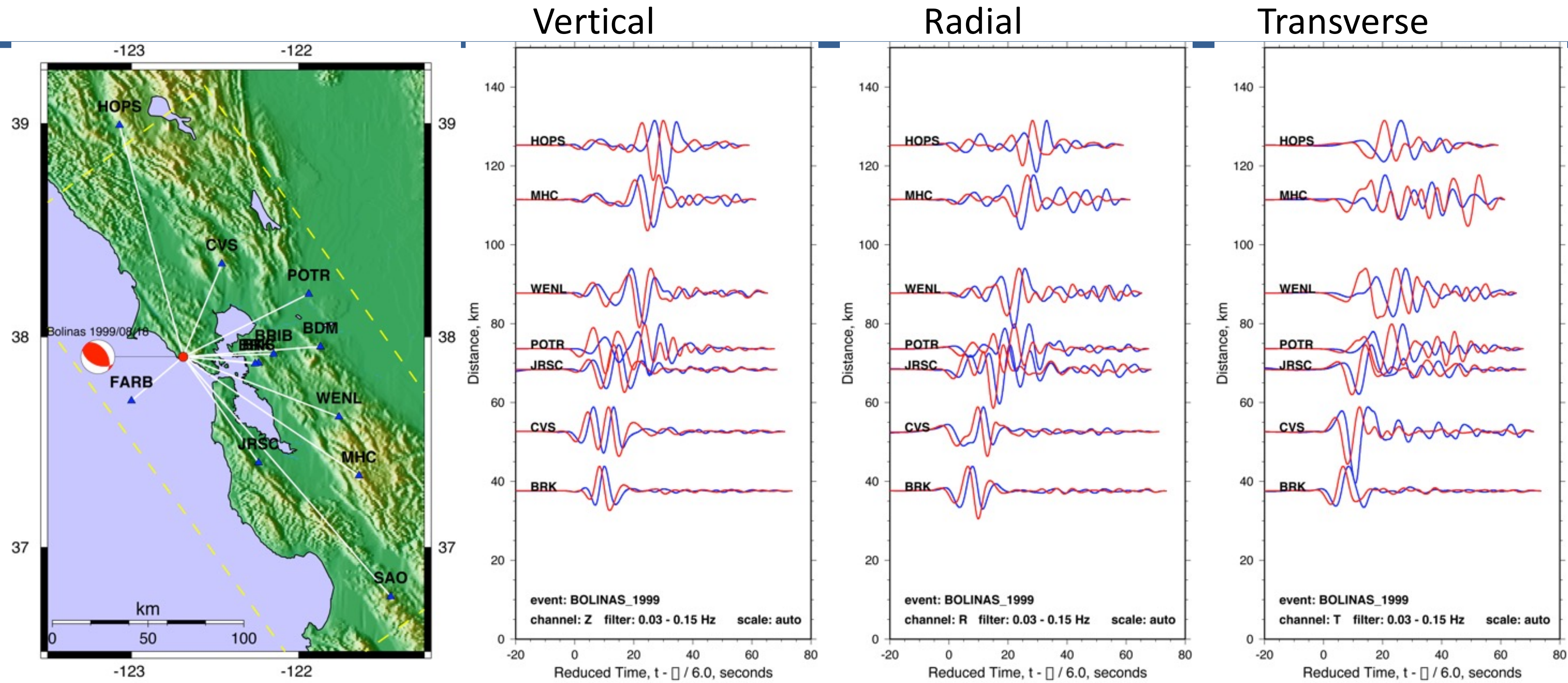
Note distortion of the wavefield by San Pablo Bay. This generates surface wave coda observed at BKS.

Shifted and aligned data & synthetic



Note long duration

Long-period waveform comparisons for Aug. 18 1999, Bolinas Earthquake



Frequencies = 0.03-0.15 Hz
Periods = 7-33 seconds

Delays increase with distance, suggests systematic bias

Note amplitudes are well matched

- see Kim, Dreger & Larsen, BSSA (2010)

Data such these are useful for waveform tomography

More recent effort looking at shorter periods: 1-32 sec.

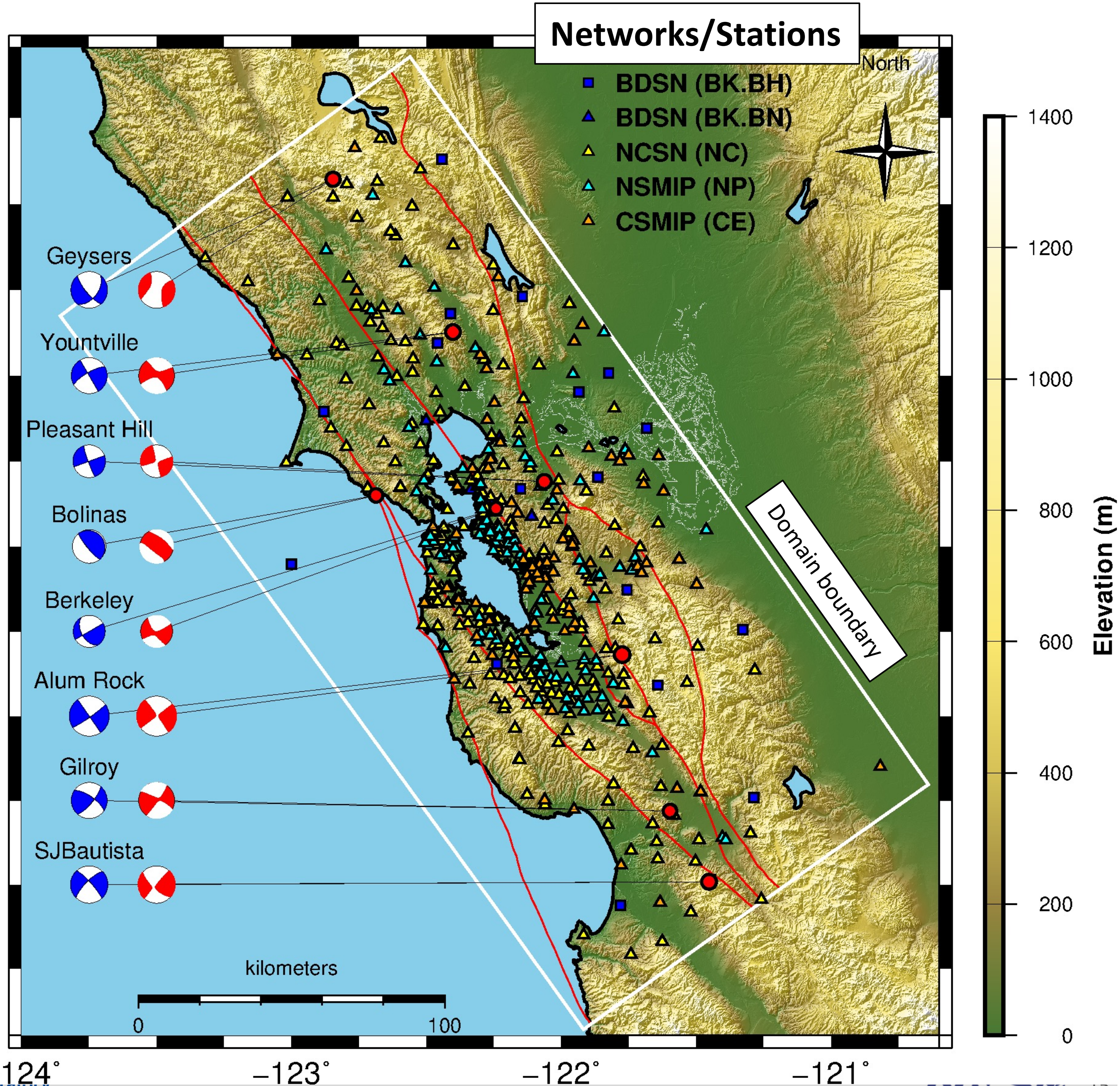
8 events
M_W 4.4-5.5
1998-present

With strong-motion sites

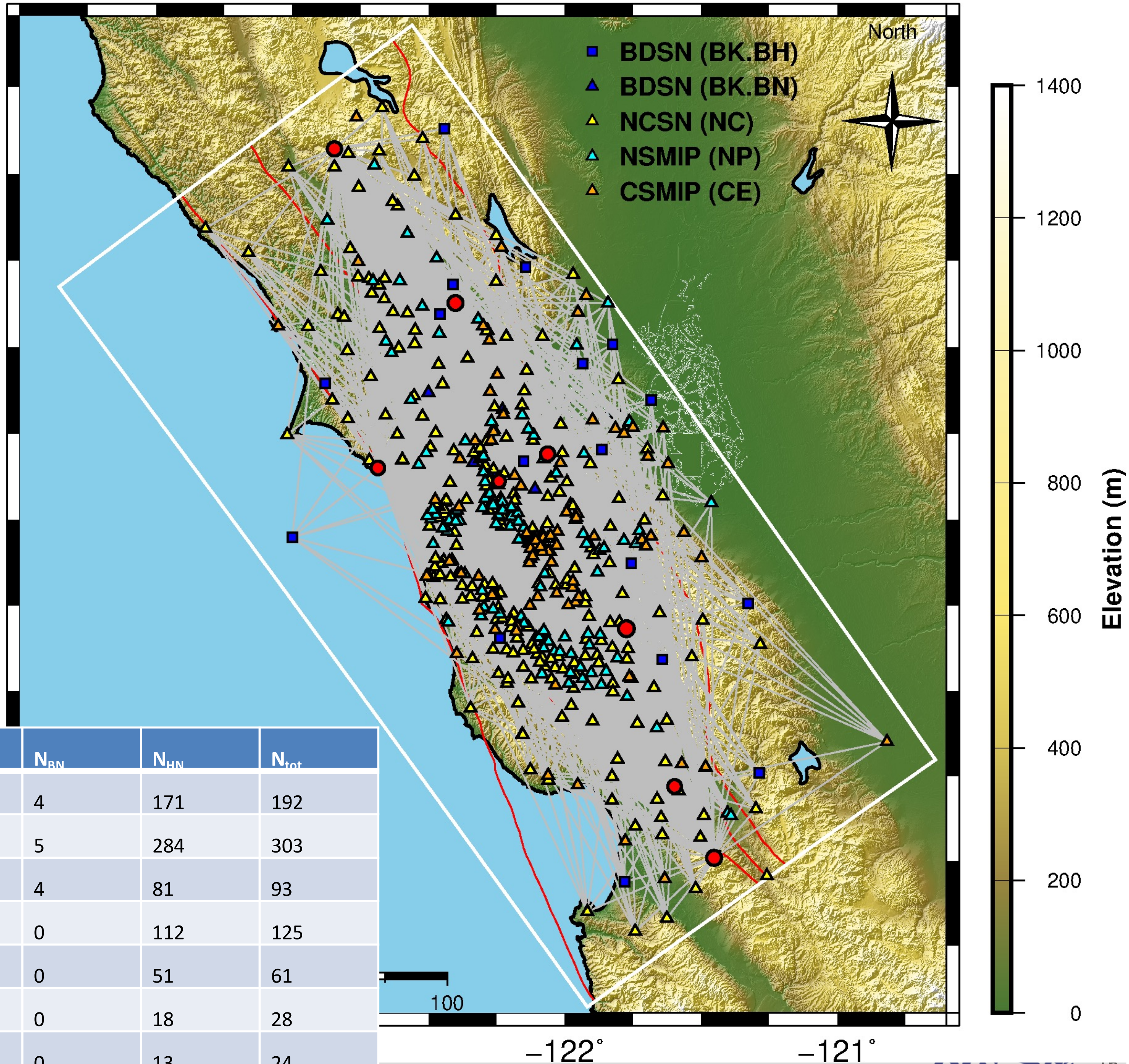
SW4 domain same as USGS Detailed model

Earth models:
- 3D + topography
- 1D flat free-surface

DC and MT focal mechanisms



Path Map

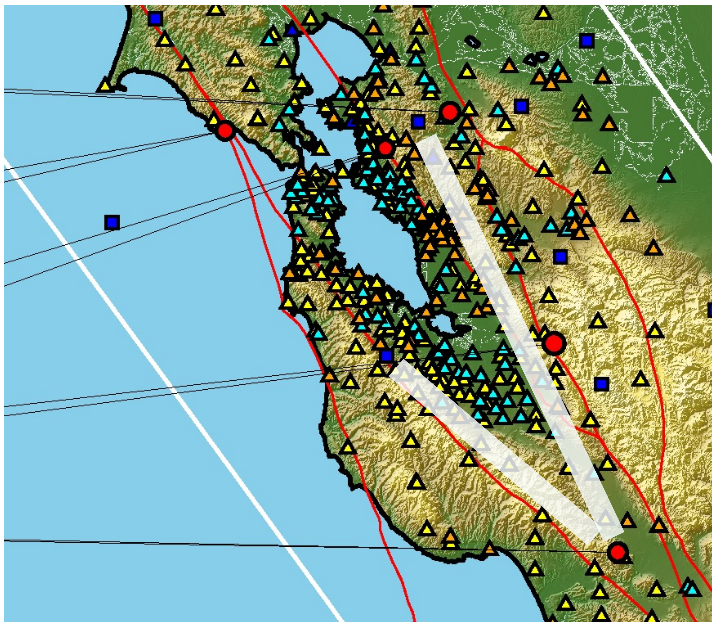


Number of paths/event
839 total for 8 events

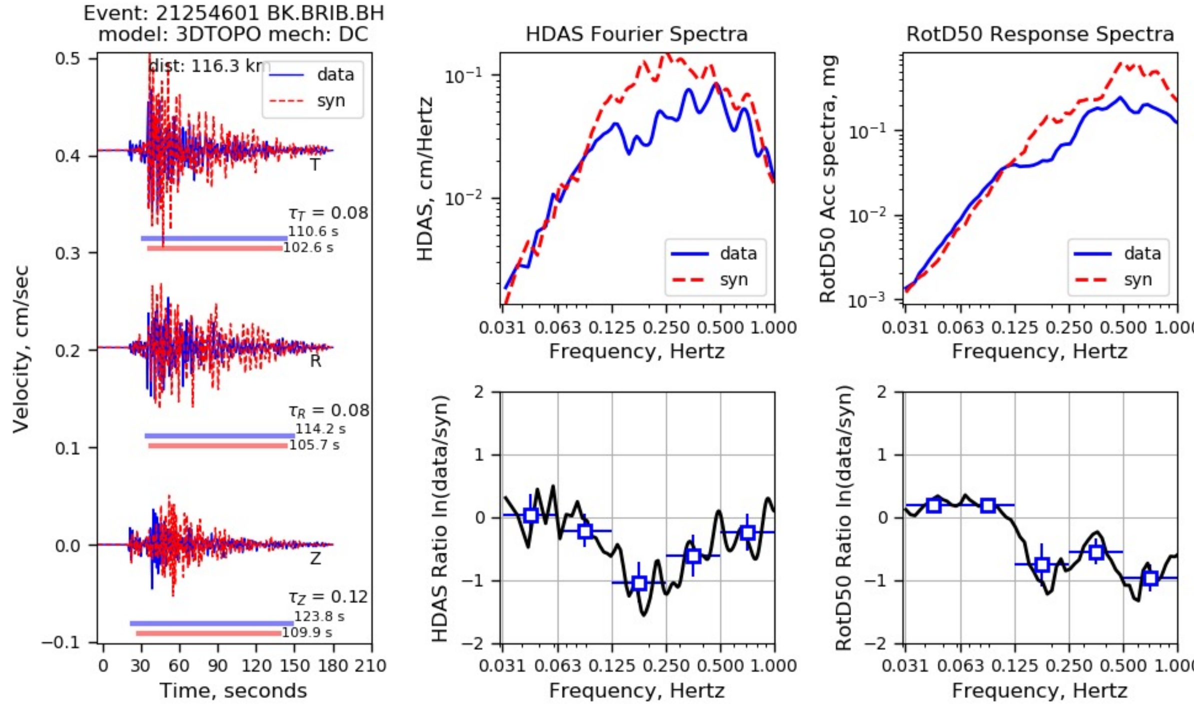
Event	Event ID	N_{BH}	N_{BN}	N_{HN}	N_{tot}
Pleasant Hill	73291880	17	4	171	192
Berkeley	72948801	14	5	284	303
Geysers	72737985	8	4	81	93
Alum Rock	40204628	13	0	112	125
Gilroy	21254601	10	0	51	61
Yountville	21123384	10	0	18	28
Bolinas	21044694	11	0	13	24
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

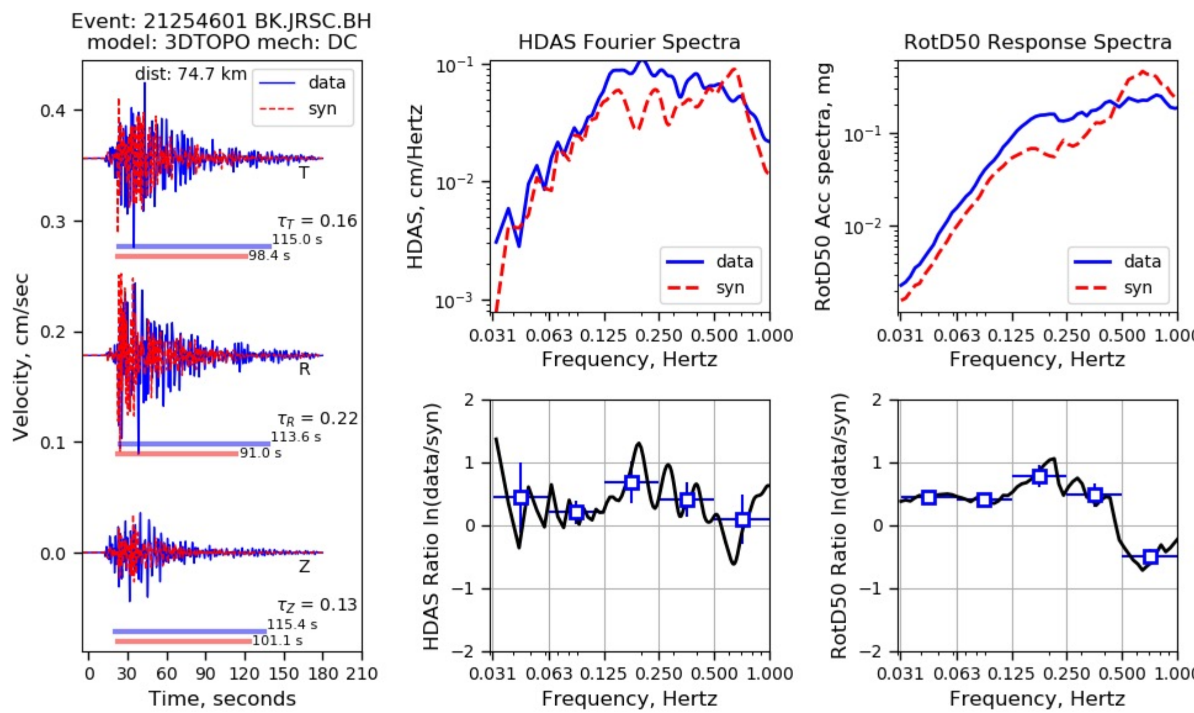
Gilroy 2002-05-14 Mw 4.9
to BDSN broadband stations

Path through East Bay Hills:
USGS 3D model produces higher amplitudes than observed



Path through Santa Cruz Mtns:
USGS 3D model produces relatively unbiased amplitudes



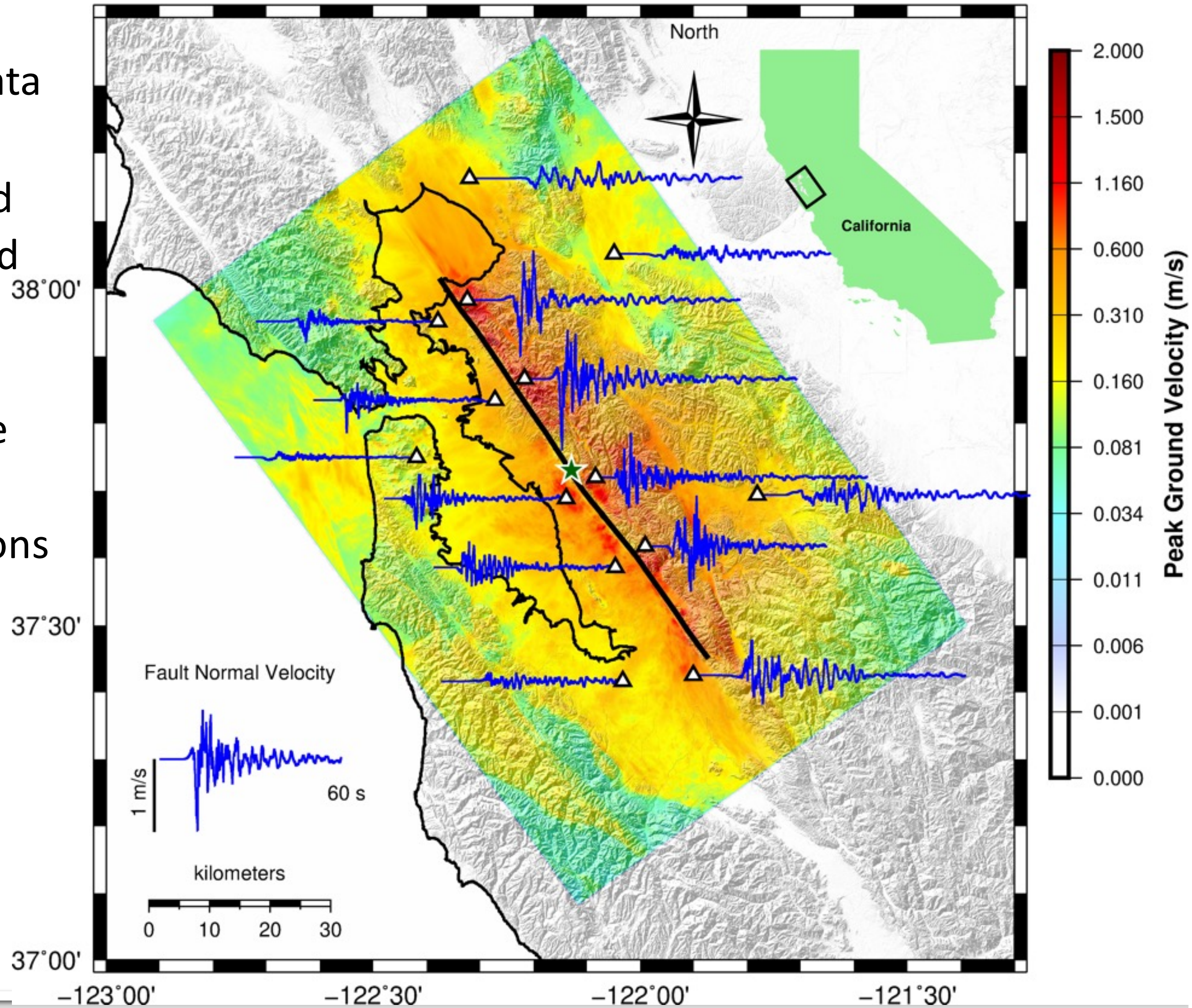


Our demonstration problem: Hayward Fault M_W 7.0 scenario earthquake

- In the absence of empirical data for scenario earthquakes,
- we compare simulated ground motion intensities with ground motion models

- Recall that we are pushing the limits of fully deterministic scenario earthquake simulations to f_{max} = 10 Hz

- This potentially exposes shortcomings of our:
  - source model
  - v_{Smin} = 500 m/s
  - linear wave propagation



Several recent papers focus on simulations of M7 Hayward Fault scenario earthquakes

Computing in Science & Engineering — IEEE

Toward Exascale Earthquake Ground Motion Simulations for Near-Fault Engineering Analysis

Hans Johansen | Lawrence Berkeley National Laboratory
Arthur Rodgers and N. Anders Petersson | Lawrence Livermore National Laboratory
David McCallen | Lawrence Berkeley National Laboratory
Bjorn Sjogreen | Lawrence Livermore National Laboratory
Mamun Miah | Lawrence Berkeley National Laboratory

Application modernization for massively parallel time-domain simulations of earthquake ground motion in 3D models is increasing application resolution and providing ground motion estimates for critical infrastructure risk evaluations. Improvements to the geophysics application code SW4 algorithms, developed while porting the code to systems at Lawrence Berkeley National Laboratory, revealed that reorganizing operation order can improve performance for massive problems.

AGU PUBLICATIONS — Geophysical Research Letters

RESEARCH LETTER

Broadband (0–4 Hz) Ground Motions for a Magnitude 7.0 Hayward Fault Earthquake With Three-Dimensional Structure and Topography

Arthur J. Rodgers[1,2,3], Arben Pitarka[1], N. Anders Petersson[4], Björn Sjögreen[4], and David B. McCallen[2,5]

Research Paper — EARTHQUAKE SPECTRA

EQSIM—A multidisciplinary framework for fault-to-structure earthquake simulations on exascale computers part I: Computational models and workflow

Earthquake Spectra
1–29
DOI: 10.1177/8755293020970982
journals.sagepub.com/home/eqs

David McCallen, M. EERI[1,2], Anders Petersson[3], Arthur Rodgers[3], Arben Pitarka[3], Mamun Miah[2], Floriana Petrone, M. EERI[1,2], Bjorn Sjogreen[3], Norman Abrahamson[4], and Houjun Tang[2]

SRL Seismological Research Letters

Broadband (0–5 Hz) Fully Deterministic 3D Ground-Motion Simulations of a Magnitude 7.0 Hayward Fault Earthquake: Comparison with Empirical Ground-Motion Models and 3D Path and Site Effects from Source Normalized Intensities

by Arthur J. Rodgers, N. Anders Petersson, Arben Pitarka, David B. McCallen, Bjorn Sjogreen, and Norman Abrahamson

1268 Seismological Research Letters Volume 90, Number 3 May/June 2019

BSSA Bulletin of the Seismological Society of America

Regional-Scale 3D Ground-Motion Simulations of M_w 7 Earthquakes on the Hayward Fault, Northern California Resolving Frequencies 0–10 Hz and Including Site-Response Corrections

Arthur J. Rodgers[*1,2], Arben Pitarka[1], Ramesh Pankajakshan[1], Bjorn Sjögreen[4], and N. Anders Petersson[1]

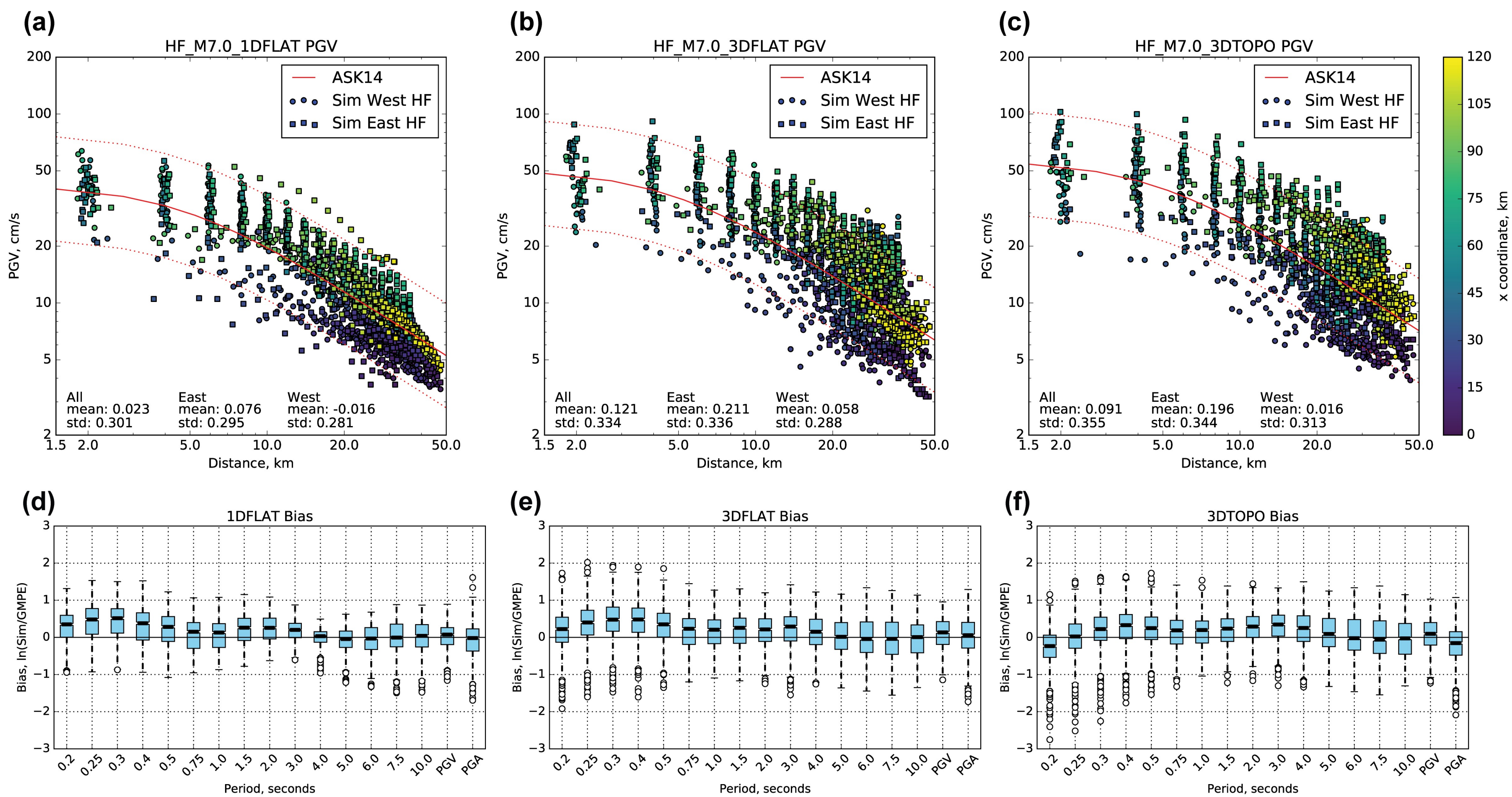
Research Paper — EARTHQUAKE SPECTRA

EQSIM—A multidisciplinary framework for fault-to-structure earthquake simulations on exascale computers, part II: Regional simulations of building response

Earthquake Spectra
1–26
DOI: 10.1177/8755293020970980
journals.sagepub.com/home/eqs

David McCallen, M.EERI[1,2], Floriana Petrone, M.EERI[1,2], Mamun Miah[2], Arben Pitarka[3], Arthur Rodgers[3], and Norman Abrahamson[4]

HF M7 0-4 Hz compared to ASK (2014) GMM



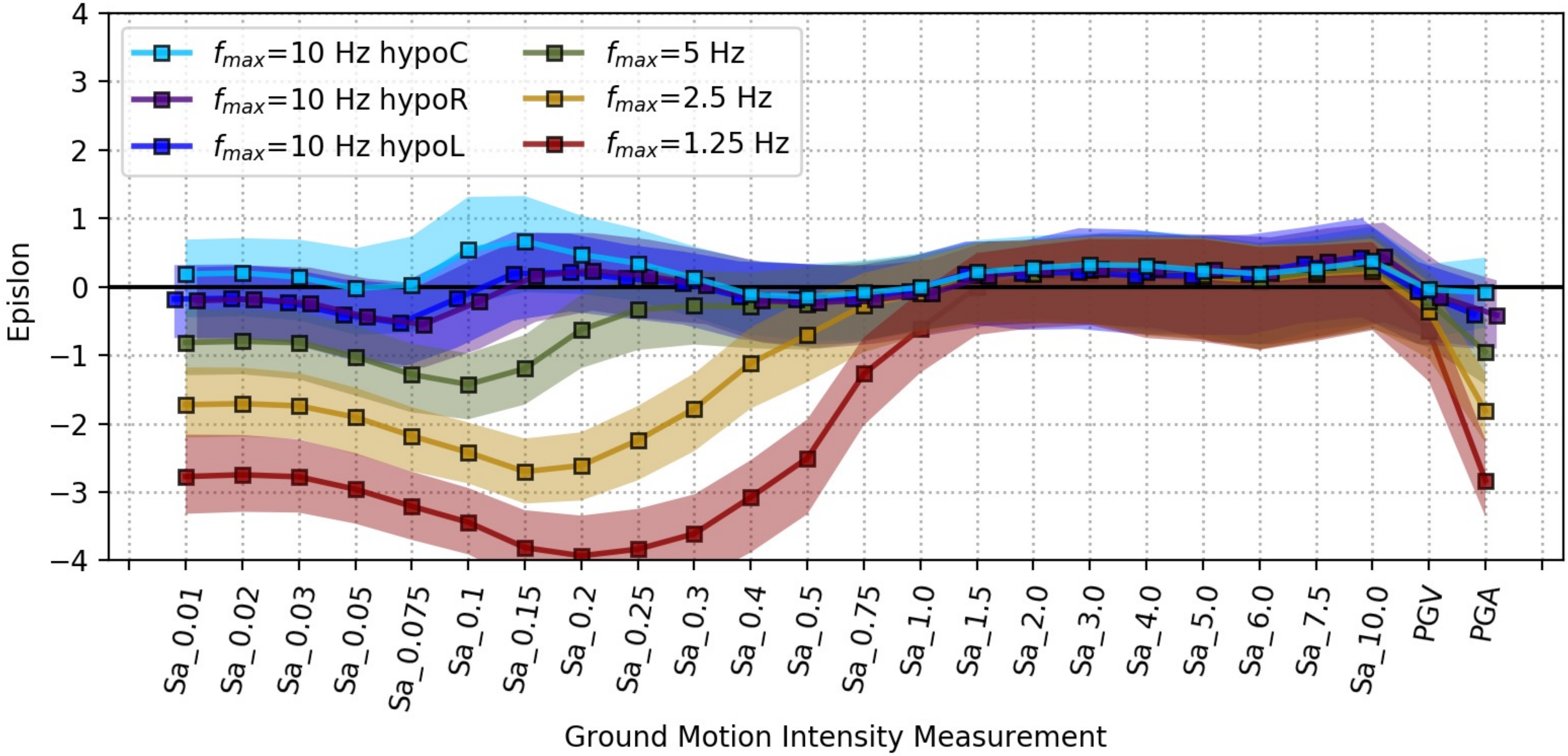
Median bias are near zero, within one σ of GMM

Variations $\sigma_{1D} < \sigma_{3DFLAT} < \sigma_{3DTOPO} < \sigma_{GMM}$

Rodgers et al. (GRL, 2018)

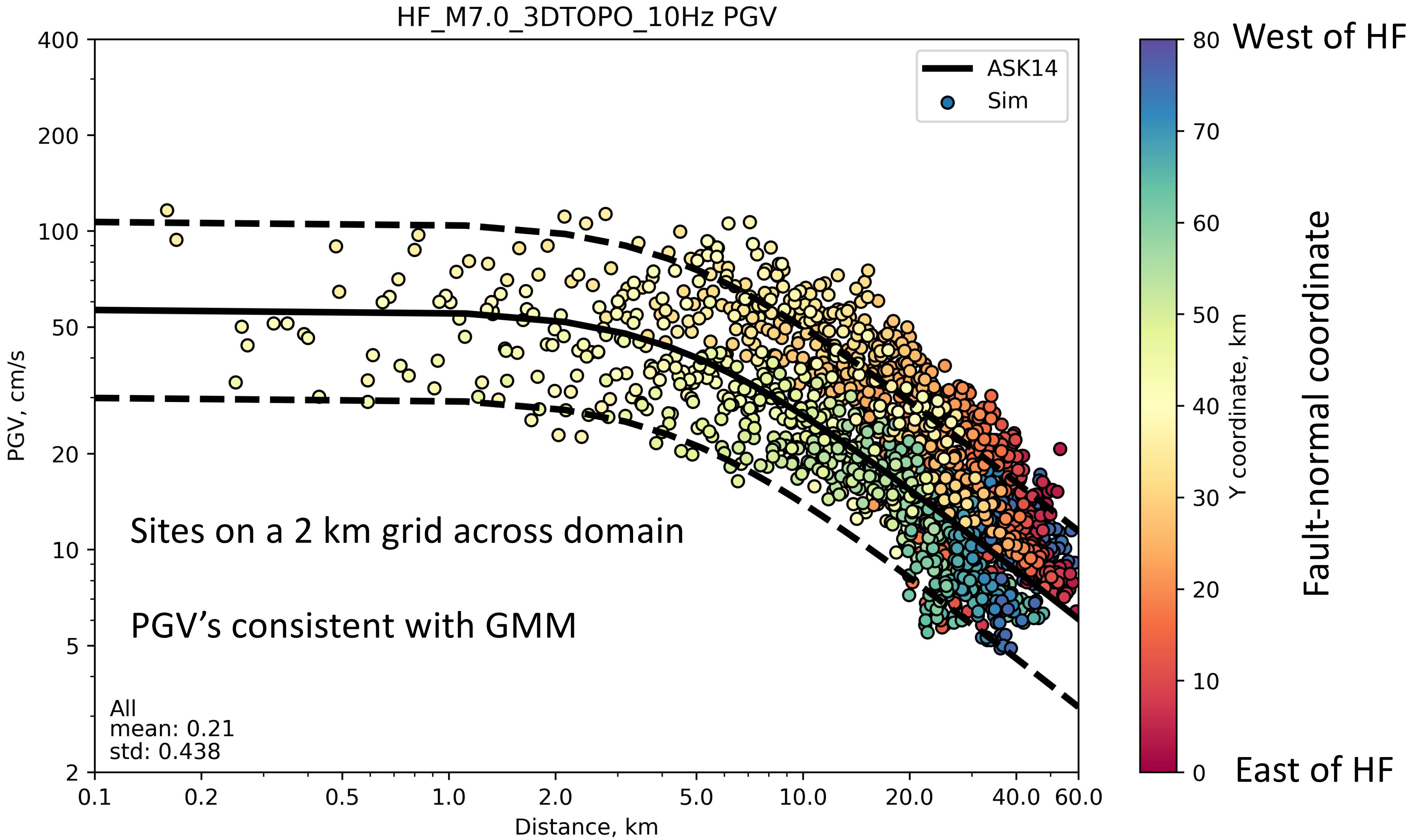
Epsilon for all sites, 3D model, f_{max} = 1.25 - 10 Hz

Colored bands show 50% of data, interquartile range



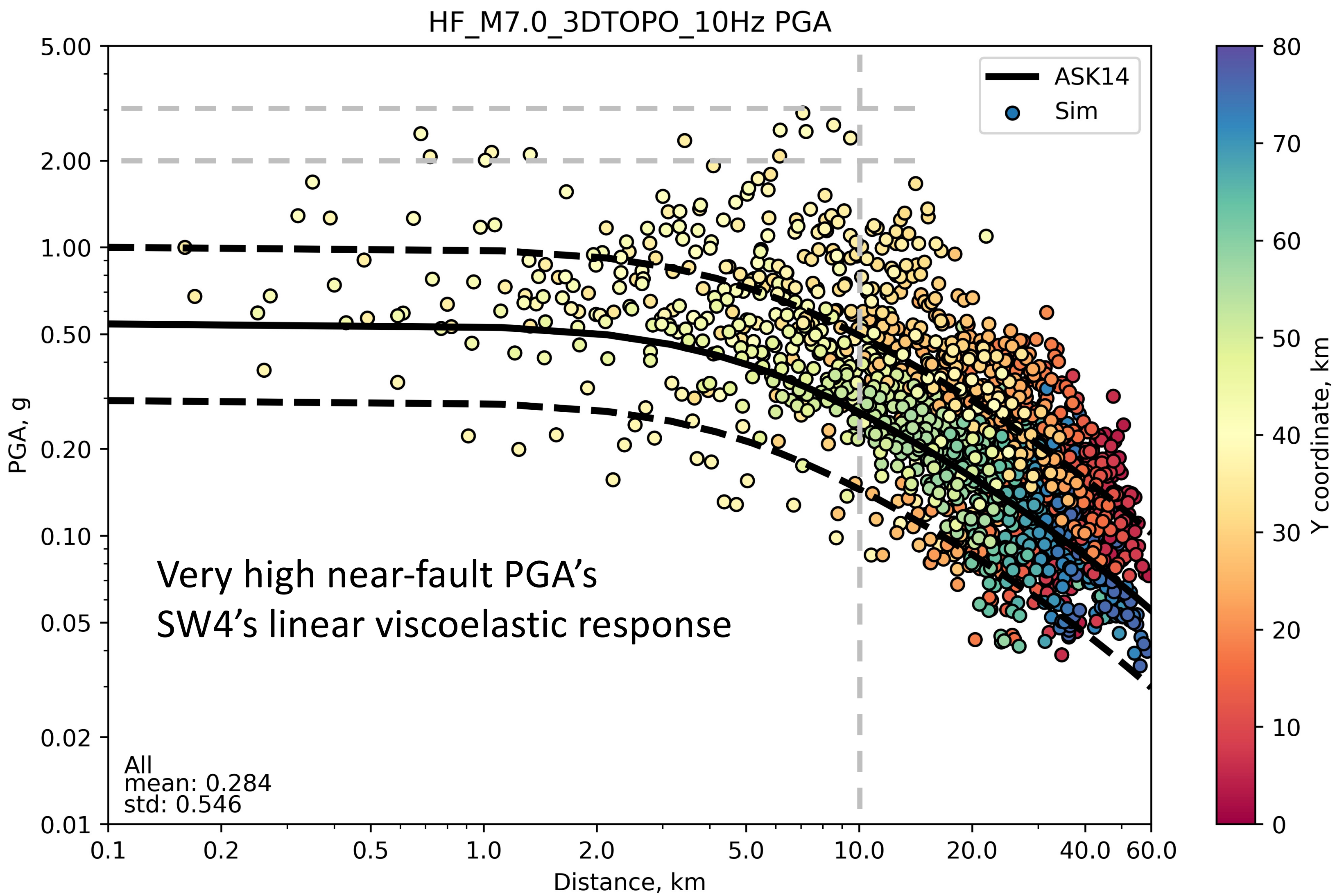
SW4 (viscoelastic) PGV versus distance compared with ASK (2014) GMM

HF M_W 7.0 10.0 Hz - PGV



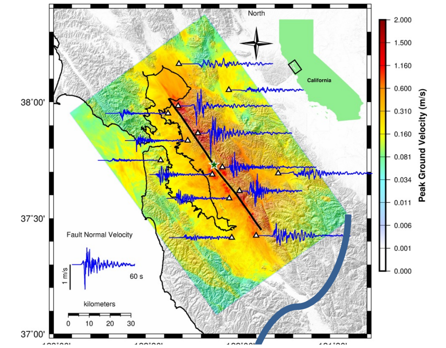
SW4 (viscoelastic) PGA versus distance compared with ASK (2014) GMM

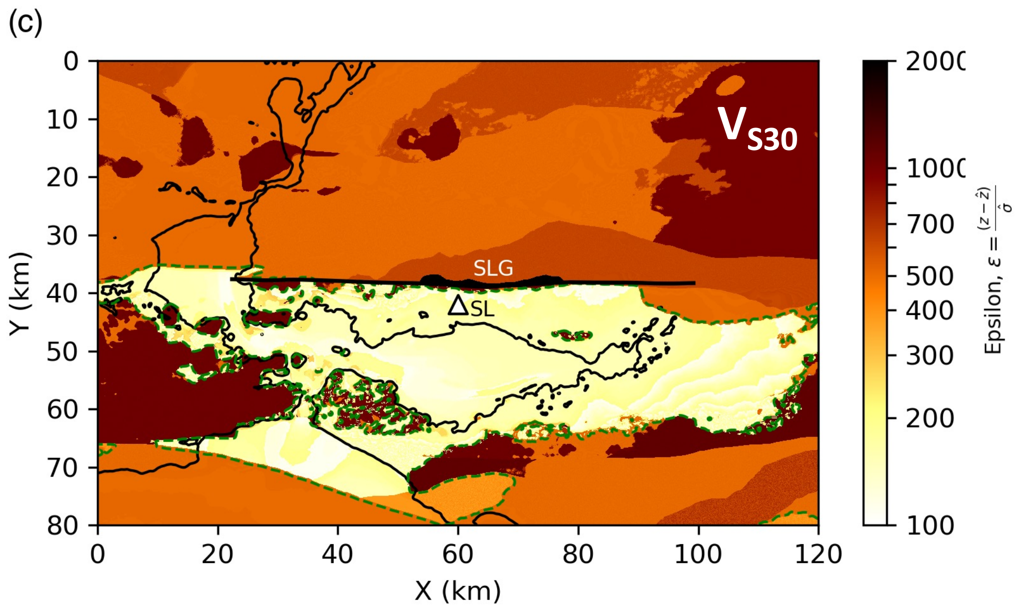
HF M_W 7.0 10.0 Hz - PGA

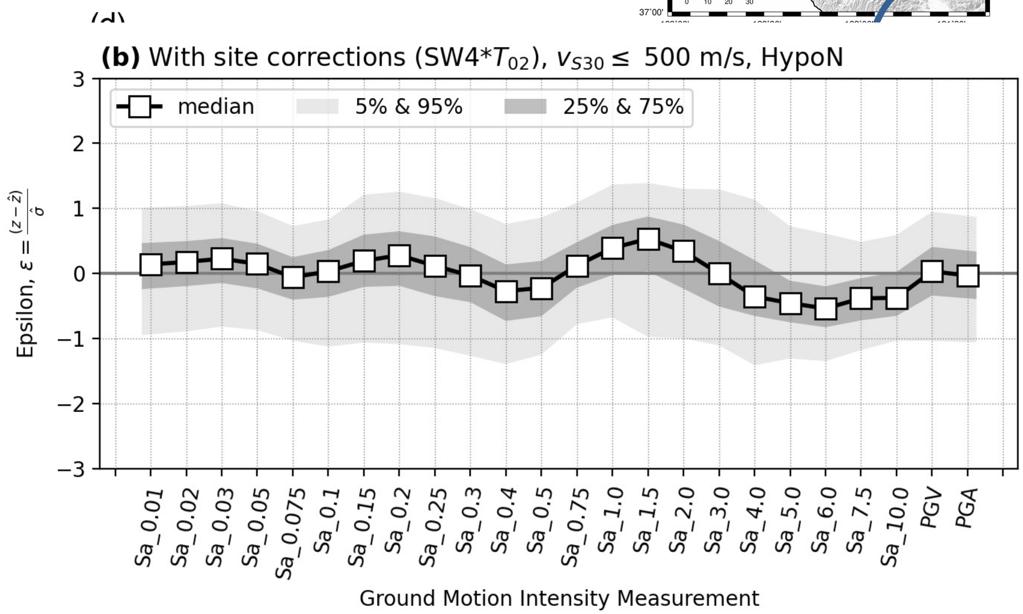


High PGA values indicate shortcomings in our simulations: linear viscoelasticity & assumed v_{Smin}

- Assumed v_{Smin} = 500 m/s does not honor weak near-surface soils
- These can respond with competing effects:
  - Amplify long-period weak motion
  - Dampen short-period strong motion

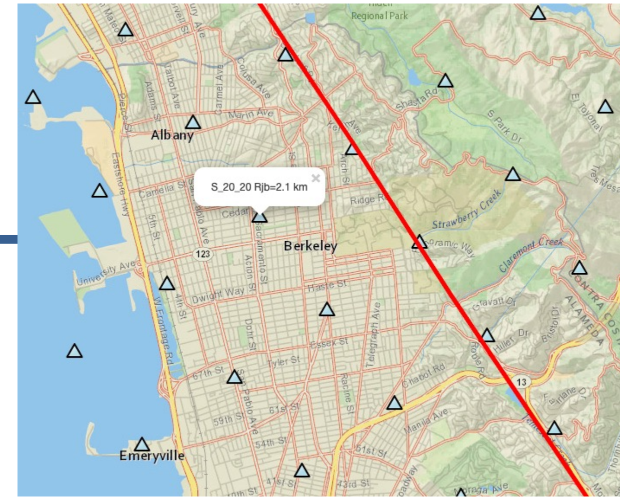




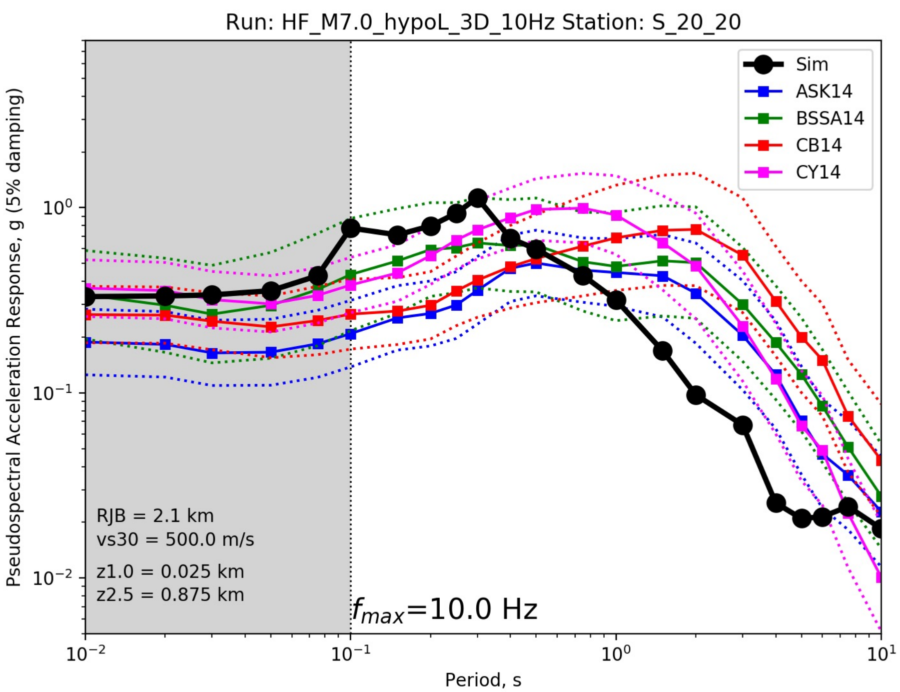


Site correction method reduces bias
At low V_{S30} sites Rodgers et al. (BSSA 2020)

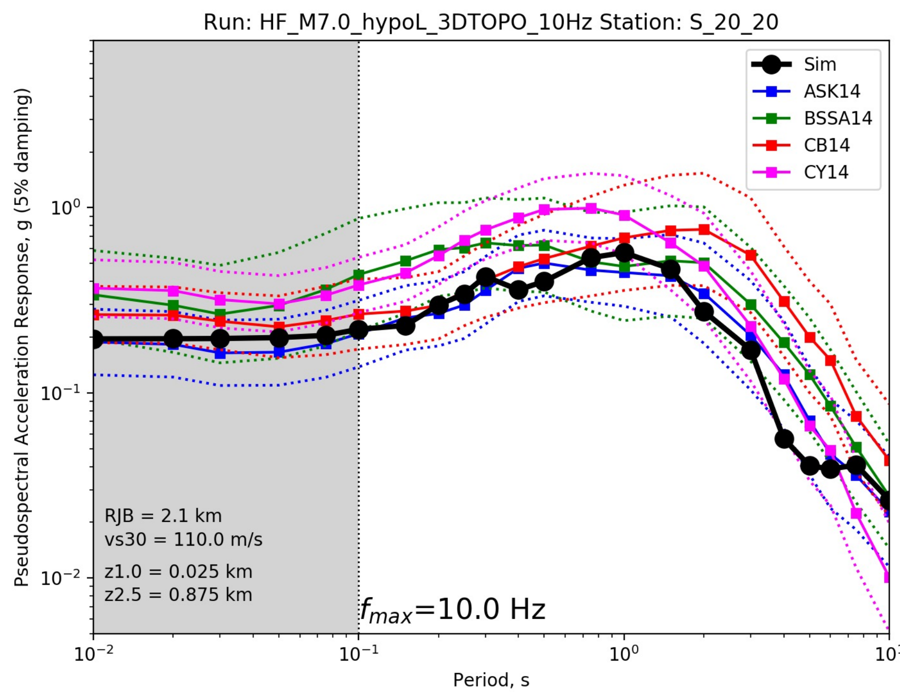
Ground motions without and with site response corrections: Berkeley



SW4 output, linear
v_{Smin} = 500 m/s
v_{S30} = 500 m/s



SW4 with site response correction
v_{S30} = 110 m/s



Conclusions

- Verification must be an essential and ongoing task for seismic simulation codes undergoing continuous development
- Validation is needed for both source and 3D Earth structure models
- Validation of path propagation in 3D Earth models with moderate events is important
  - We must learn as much as possible from smaller events
    - Basin effects, crustal structure, waveform tomography
  - We are awaiting update of the SFBA model from USGS (Hirakawa, Aagaard)
- Validation of large event ruptures is more complex due to:
  - Lack available empirical data
  - Dependence of intensities on source, path and site effects
  - *Simulated data may be consistent with GMM's, but is the Earth model correct?*
  - Additional criteria must be considered such as
    - Median epsilon
    - Within-event and between-event variability
    - Spectra correlation
    - Duration
    - Building response, engineering demand

Recommendations for community-based V&V

- Encourage FAIR (Findable, Accessible, Interoperable & Reusable) best-practices
  - Version control on data sets, synthetics and 3D models
- Standardization of waveform and event parameter data used in simulations for validation
  - Assembly with Python, ObsPy, Jupyter notebooks
  - Storage as ASDF (single HDF5 file per event)
- Standardization of simulated event data and metadata
  - Simulation metadata, e.g. input file(s) so others can reproduce
  - Storage as ASDF (single file per event)
  - Source and site parameters used in GMM's
- Standardize metrics for comparison
  - Waveform and intensity measurements

https://seismic-data.org/