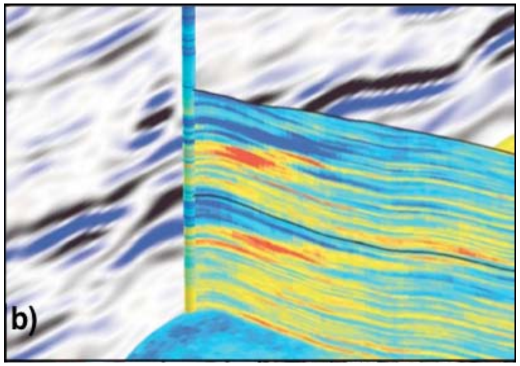
Validation of Dynamic North America Models: Benefit from finite frequency?

Carene Larmat & Monica Maceira, EES17-LANL
David M. Higdon (CCS, LANL)
Rob Porritt & Richard Allen (UC Berkeley)

4th Quest Meeting, May 19-25, 2013 LA-UR-13-23665

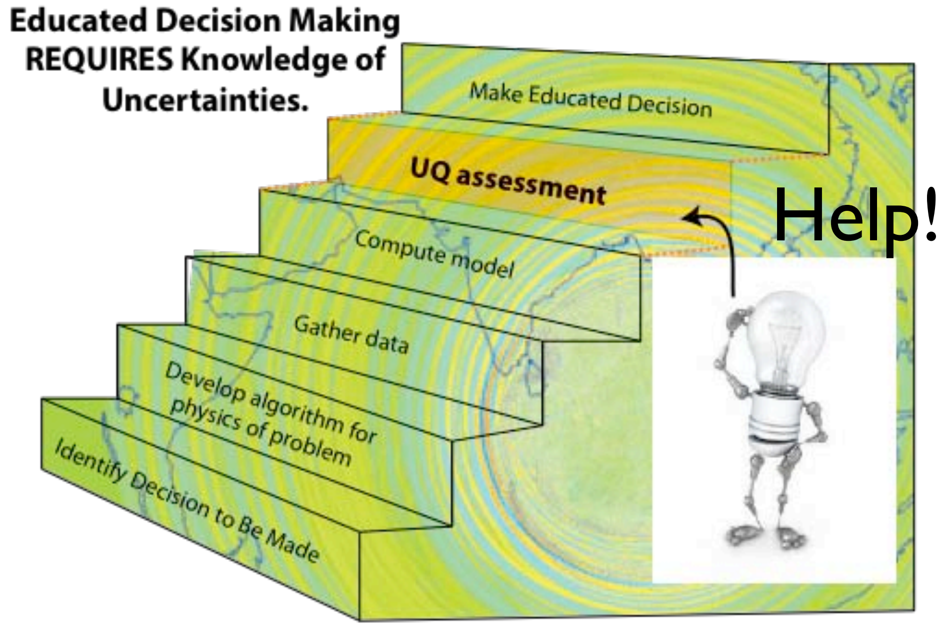
Motivations(1)

High-resolution 3D geophysical models are at the base of earthquake studies but are also pivotal for national security problems: monitoring, Fossil Energy between others.

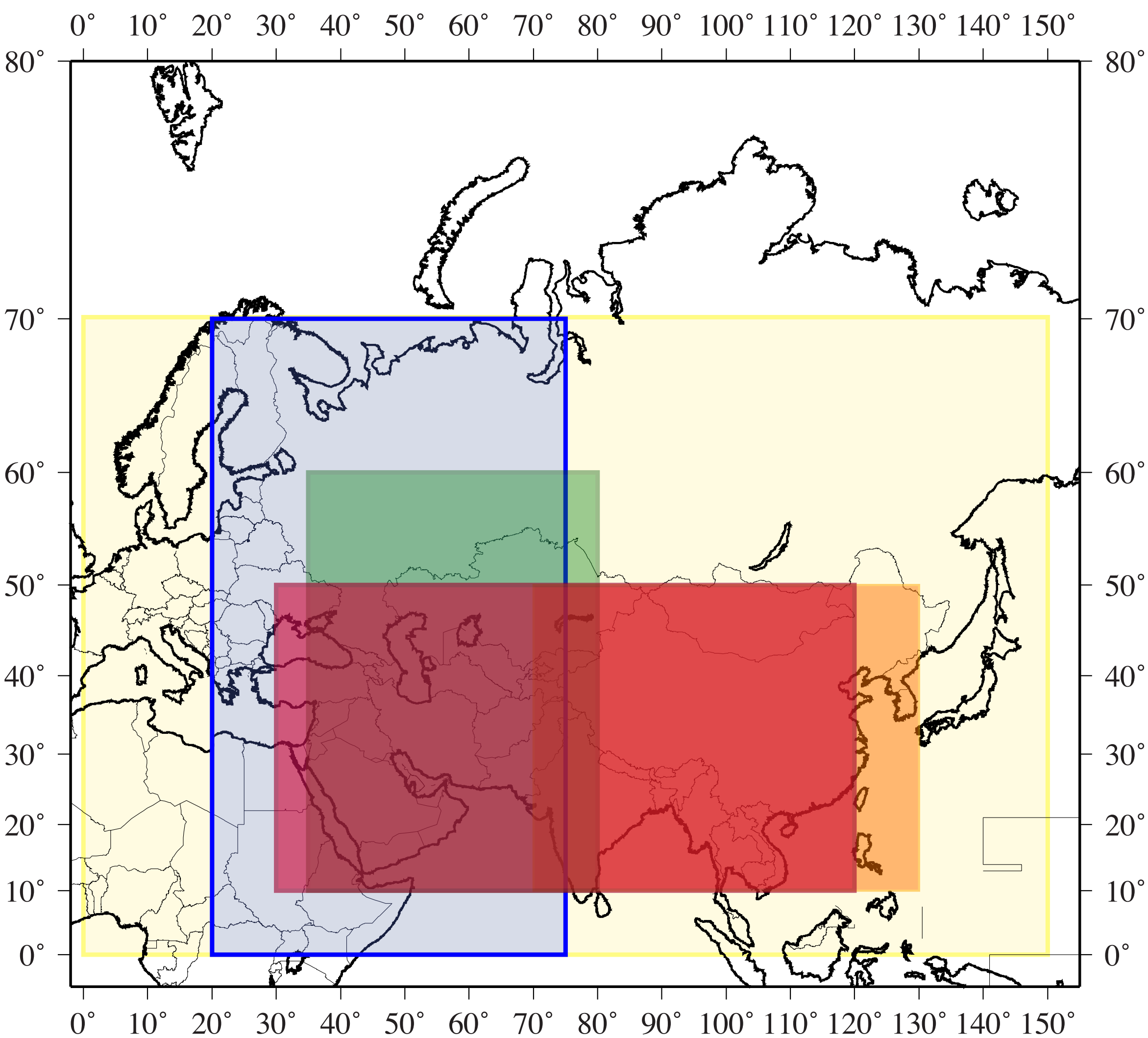


<u>Uncertainty Quantification (UQ)</u>: the discipline that develops theory, methods and tools to assess the reliability of scientific inferences.

- Range of possible solutions permitted by the set of observables.
- Key element for resource managers to properly plan for the future as understanding the limits of the models allows educated decision making



Motivations(2)



JWM (Weston/MIT)
Pn + Rayleigh waves

DOE Unified (LANL/LLNL)
A priori model

EVA09 (Northwestern/LLNL)
Waveforms and telseismic S

Li and van der Hilst (MIT)
Body waves travel times

RSTT (LANL/LLNL)
Body waves travel times

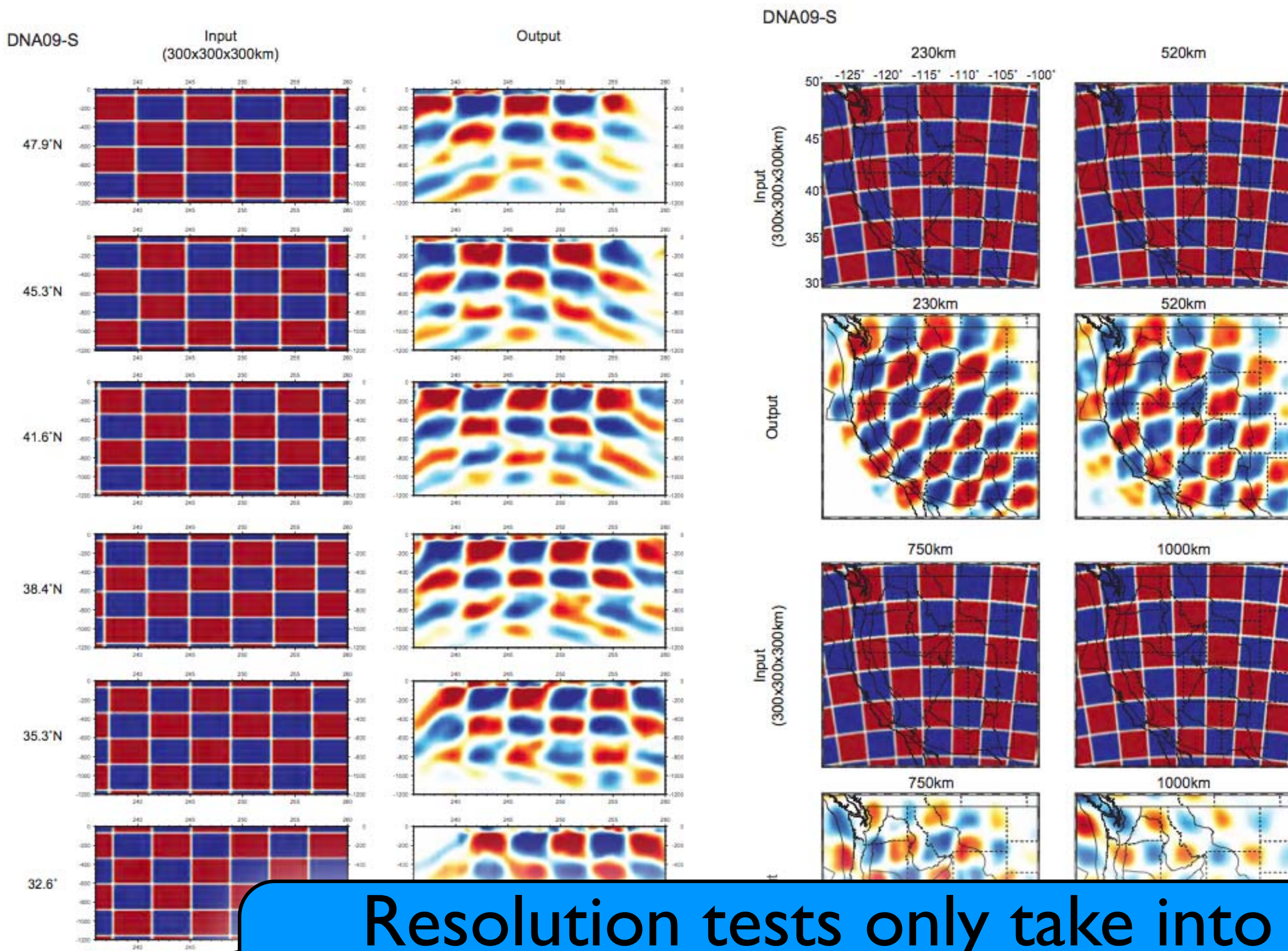
SALSA3D(LANL/SNL)
Body waves travel times

CUB20_J362D28 (CUB/Harvard)
Surface waves group velocity

LP2008(SAIC)
Surface waves

?

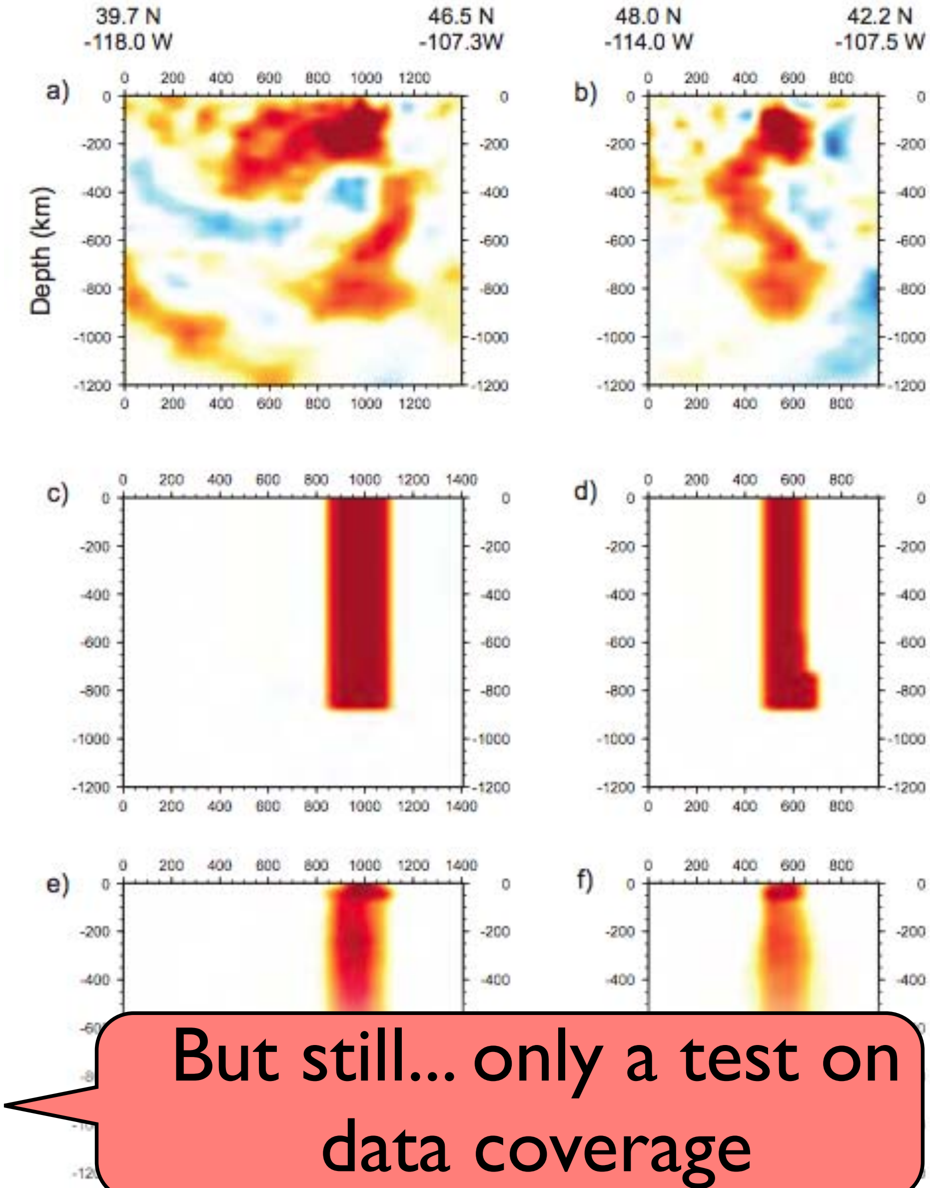
Checkerboard tests



Resolution tests only take into account the data coverage

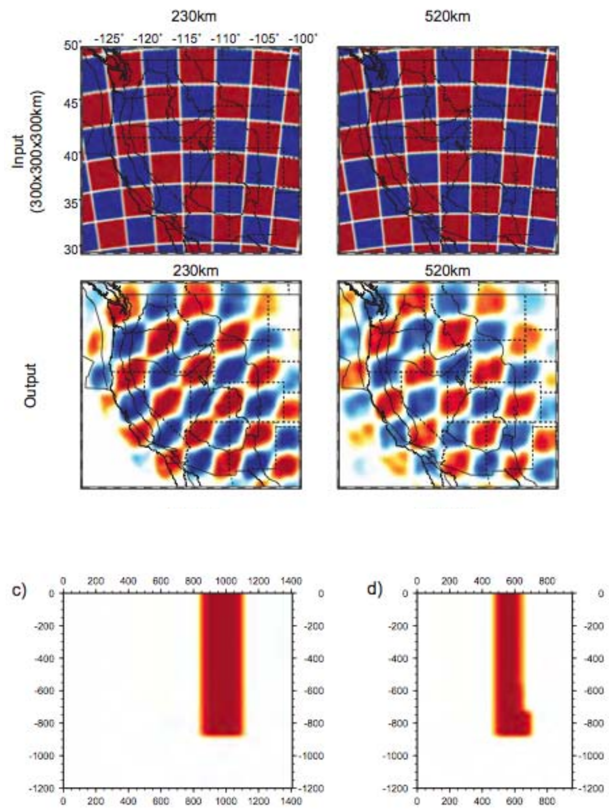
More dedicated tests

Resolution tests customized to a plume conduit



But still... only a test on data coverage

Approaches

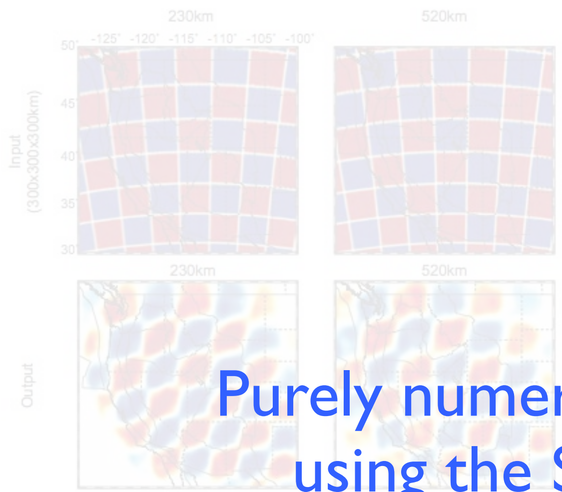


- Only indicates which part of the model can be potentially resolved by the data coverage. Does NOT give any indication of the range of parameter values that is consistent with the data.

- Very efficient psychological tendency to mistake the great detail seen in the models as an indicator of comparatively high resolution (e.g. Fichtner 2011).

- May not always provide meaningful quantifications of nonuniqueness (Menke, 1989, Lévêque et al., 1993).

Approaches

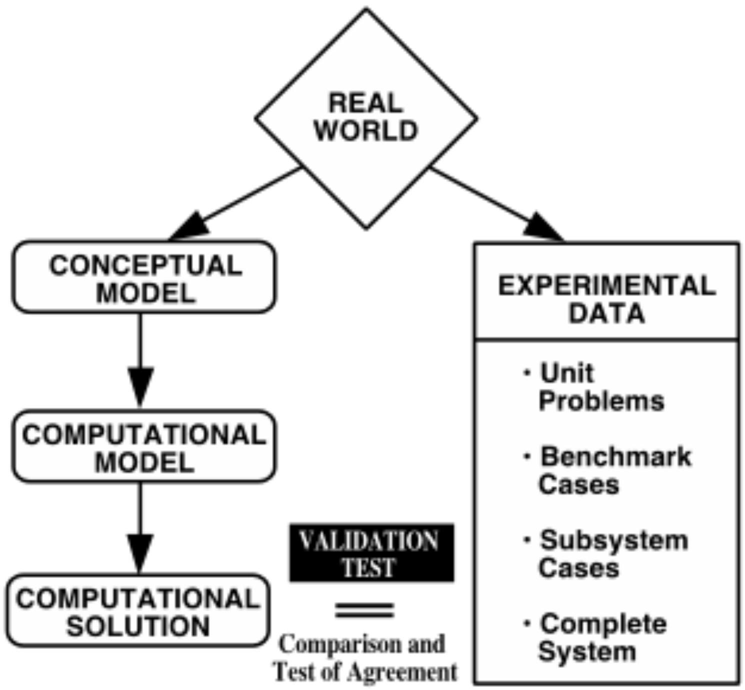


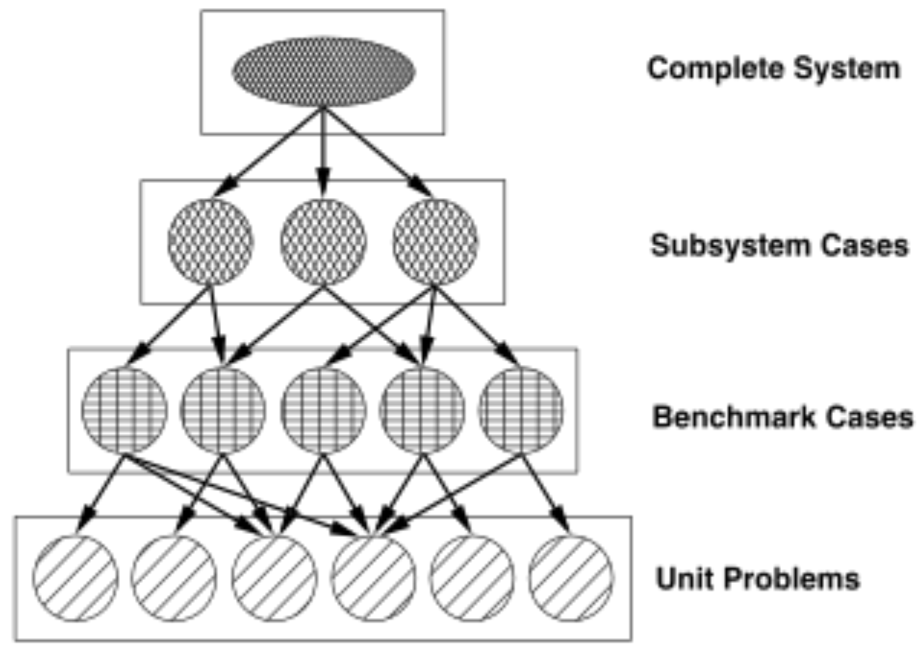
- Only indicates which part of the model can be potentially resolved by the data coverage. Does NOT give any indication of the range of parameter values that is consistent with the data.
- Very efficient psychological tendency to mistake the great detail seen in the models as an indicator of comparatively high resolution (e.g. Fichtner 2011).
- May not always provide meaningful quantifications of nonuniqueness (Menke, 1989, Lévêque et al., 1993).

Our Approach:

Purely numerical validation of 3D models using the Spectral Element Method.

Validation

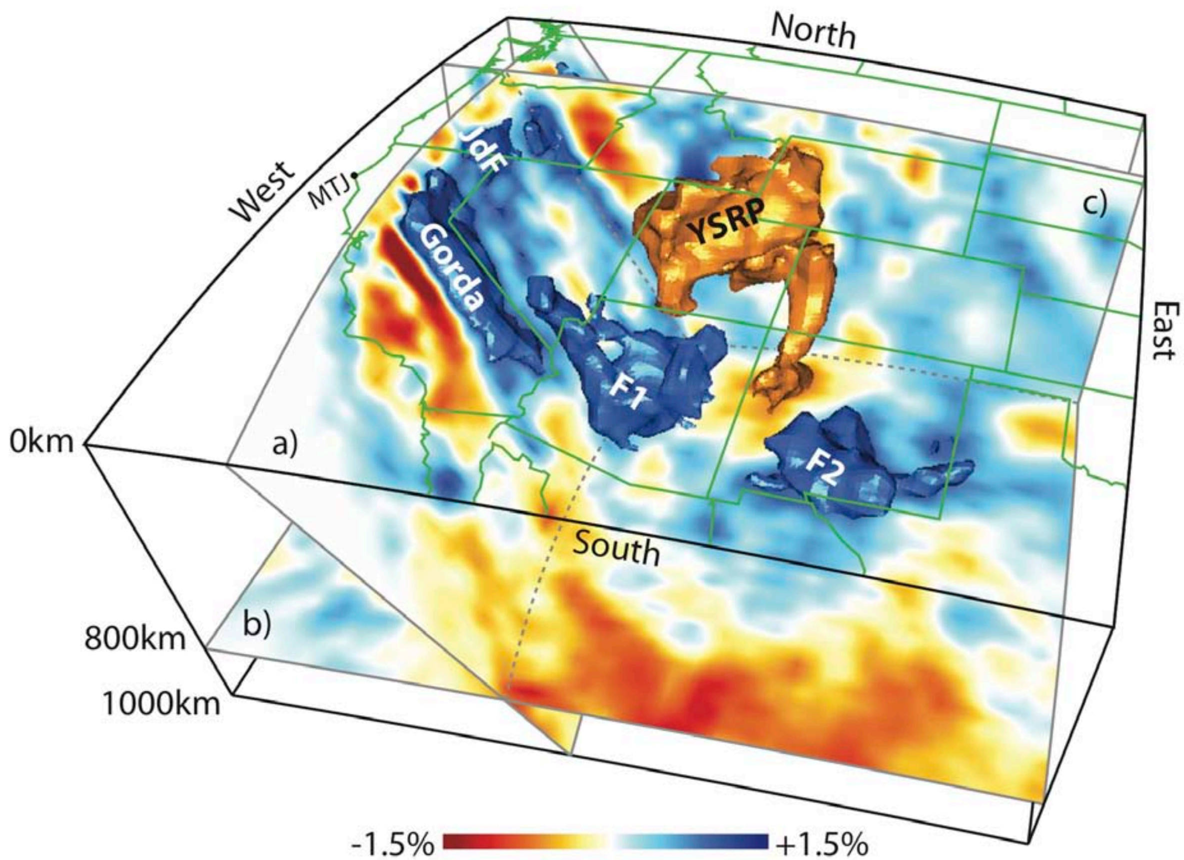




Validation: the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.

Breaking complex system into a hierarchy of successively less complicated subsystems, and eventually separated effects.
Physical experiments at different levels of this hierarchy for calibration, estimating prediction accuracy, and assessing quality of model-based predictions.

DNA**09-P -S** velocity models



Obrebski, Allen, Xue & Hung, 2010

Mathias Obrebski

Robert Porritt

Caroline Eakin

Richard Allen

UC Berkeley

Fred Pollitz

US Geological Survey

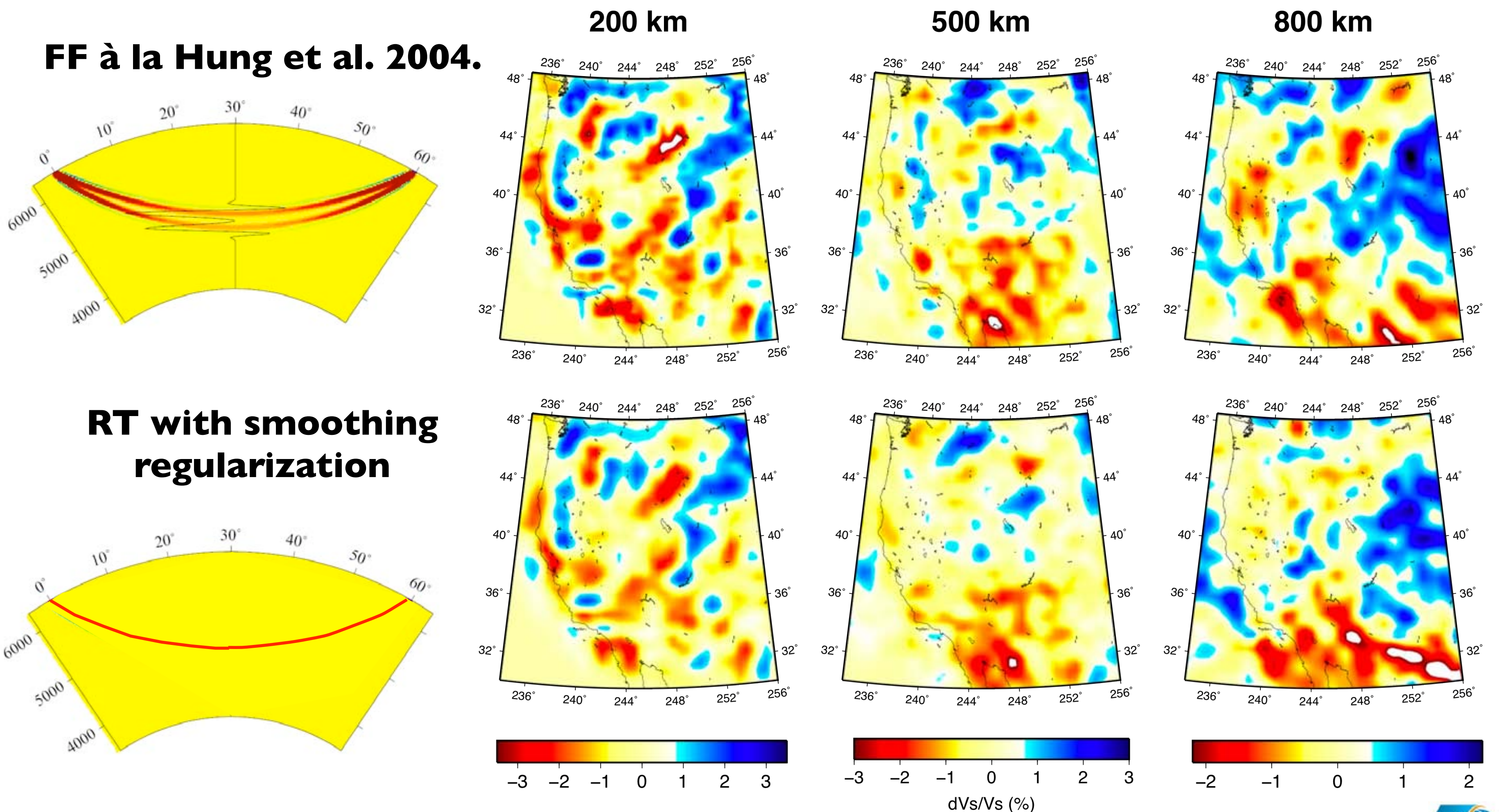
Shu-Huei Hung

National Taiwan University

DNA models available to download and slice at **http://dna.berkeley.edu**

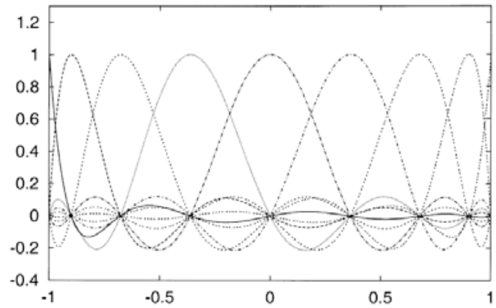
Testing the imaging method

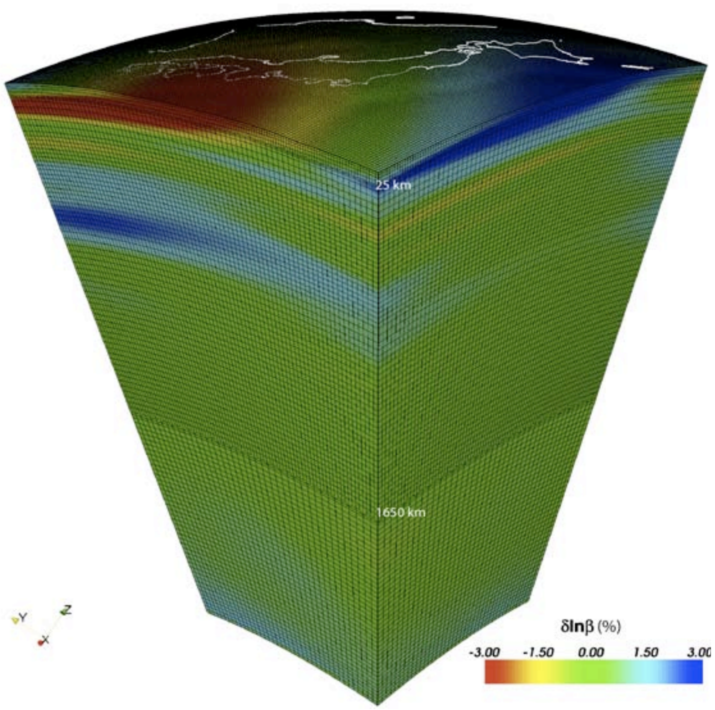
Comparing Finite Frequency and Ray Theory Methods



Larmat & Maceira, Validation of DNA models

Independent Wave Propagator





SPECFEM3D high-accuracy modeling in 3D Earth model

- finite element method based on GLL quadrature and Lagrange interpolation combines the **exponential convergence** of a high order Finite Element method with the geometric flexibility of Finite Element methods (Komatitsch & Vilotte, 1998; *Komatitsch & Tromp, 1999*).
- Current versions incorporate 3D topography of seismic interfaces, anisotropy and attenuation, mesh refinement to handle low velocity and complexity in the crust.

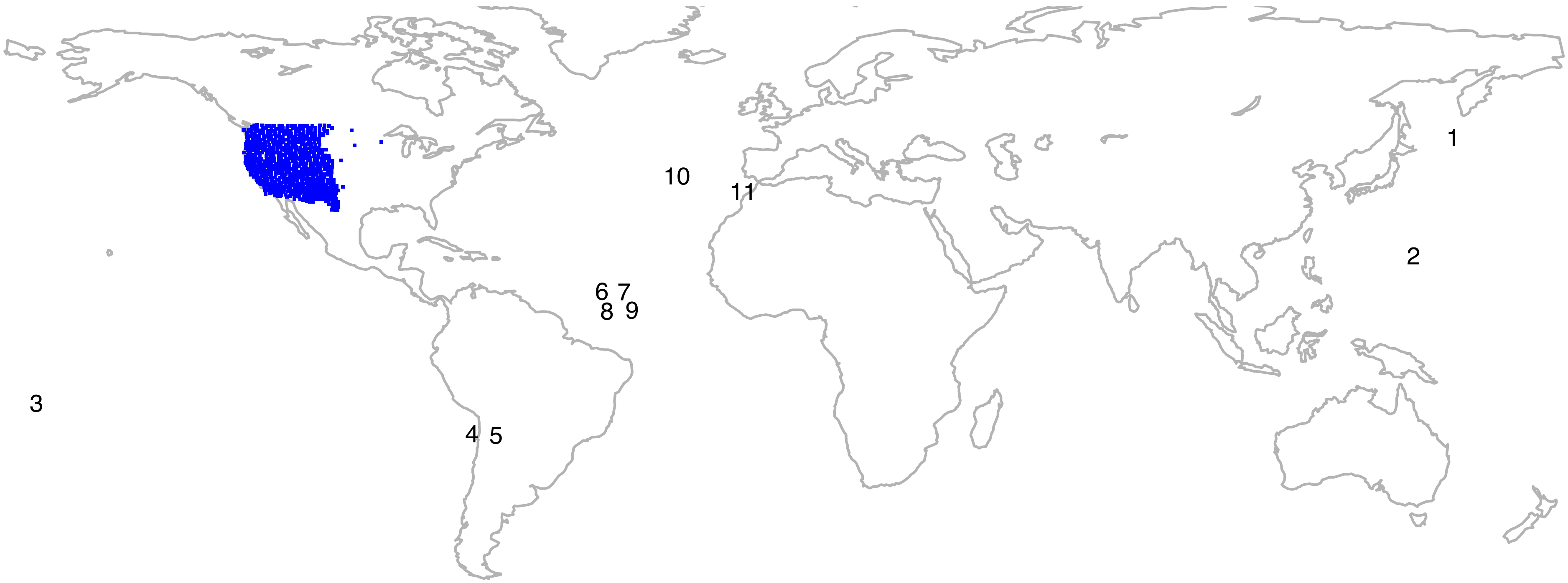
HPC resources

- Number of CPU-hrs: 600k CPU-hrs.

	Resolution		Nb CPU-hrs
actual	**10s**	486 cores of Conejo, each run with 1.3 billion grid points and 32,900 time steps	150x4,000 = 600k
goal	**5s**	runs on 4056 cores of Mustang, each modeling with a total of 8 billion grid points and 54,200 time steps	40,000 each

- Proposals of million(s) of CPU-hrs: 600k CPU-hrs.
- External access to LANL through an Open Network.

Validation set



✓ 11 events with different characteristics (depth, magnitude).

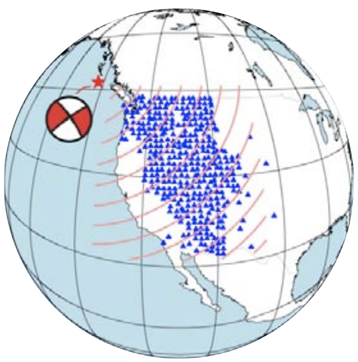
✓ Good azimuthal coverage.

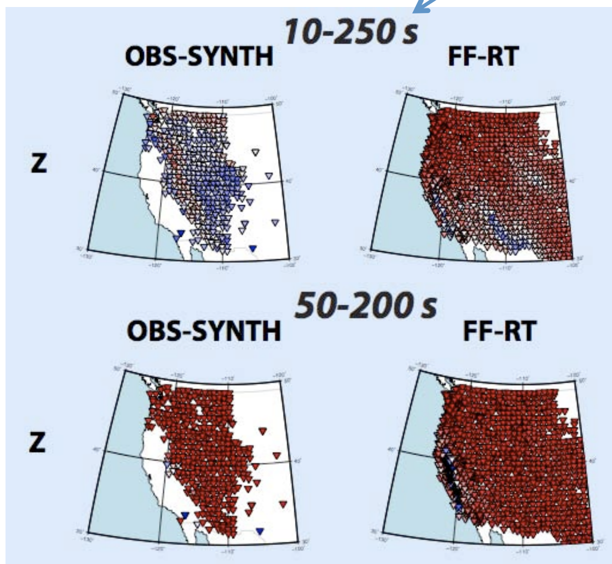
“full” waveform(1)

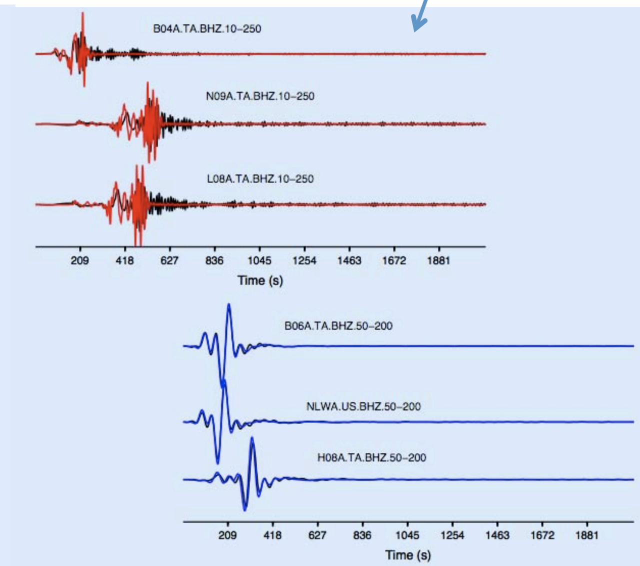
Cross-correlation coefficients computed for two frequency-bands

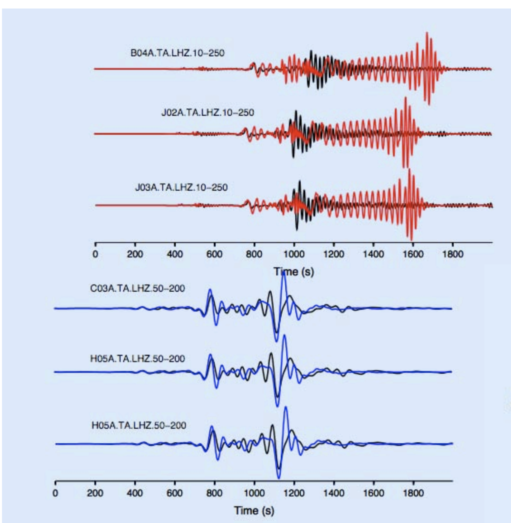
Comparison synthetic seismograms and data

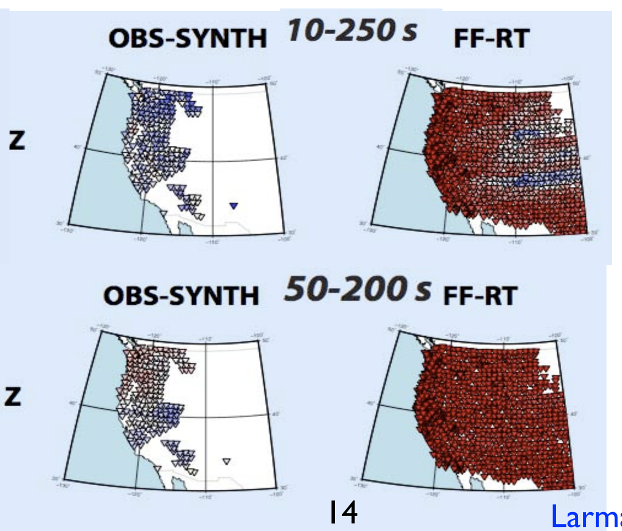
QUEEN CHARLOTTE

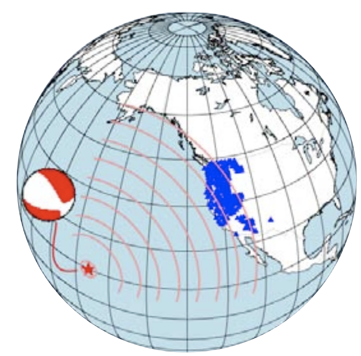






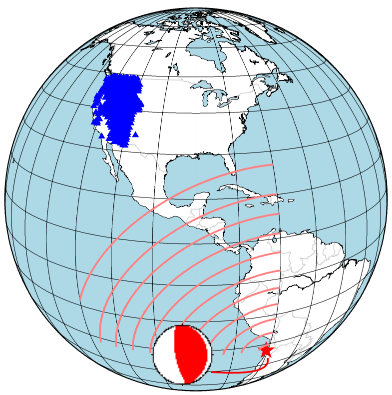




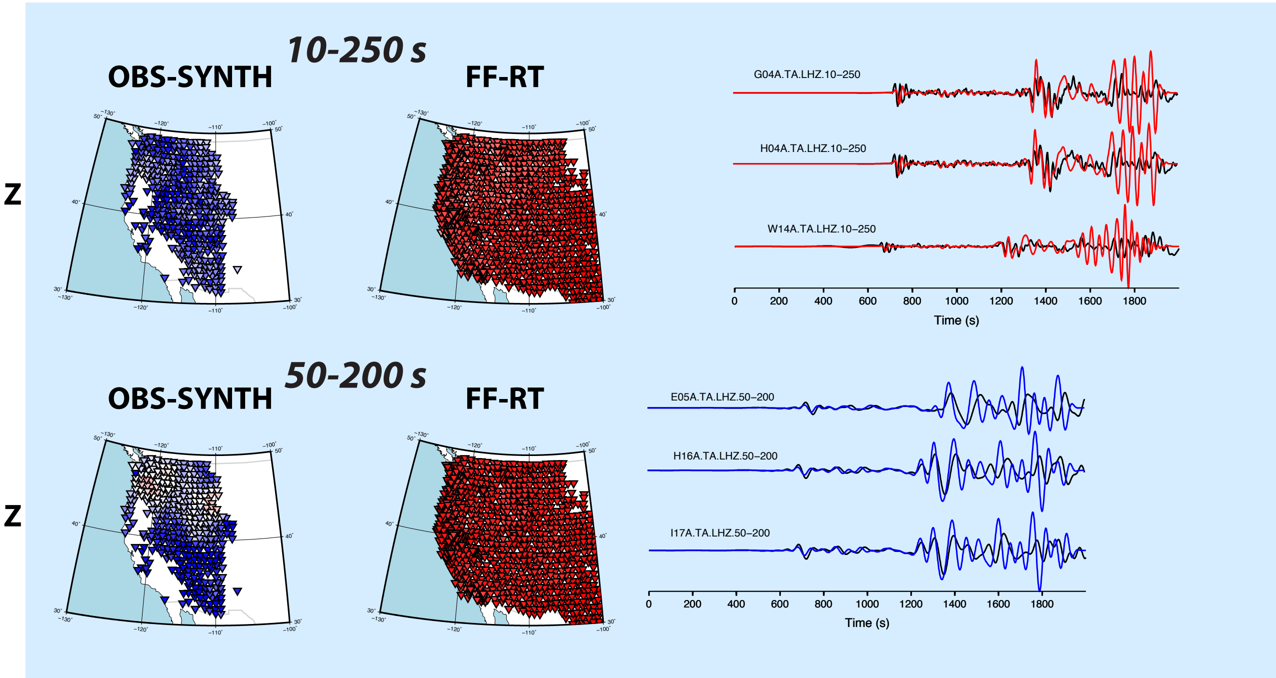


HAWAII

“full” waveform(2)



NORTHERN CHILE



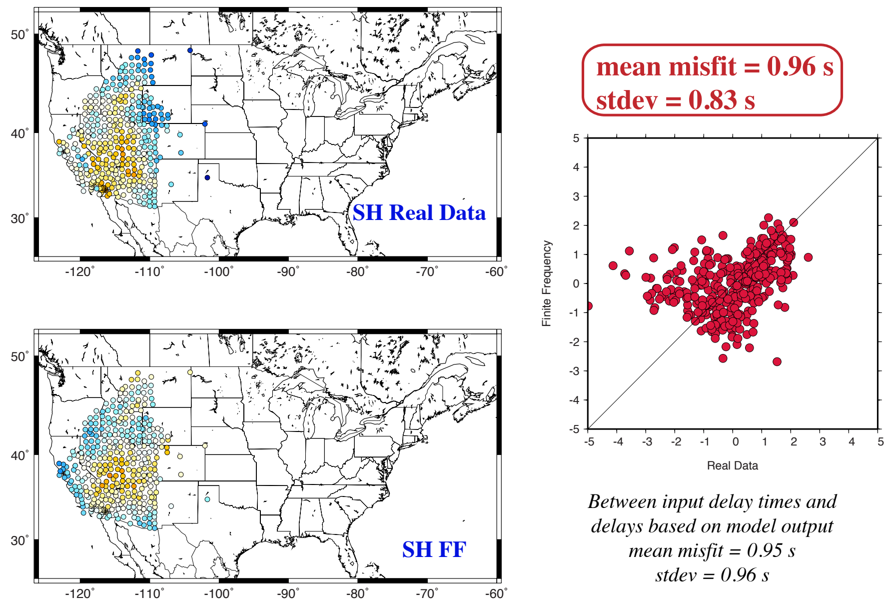
Seismic Velocity Models:

- fit the observations well for the period band 50-200s.
- Fit degrades with larger event-receiver distances.

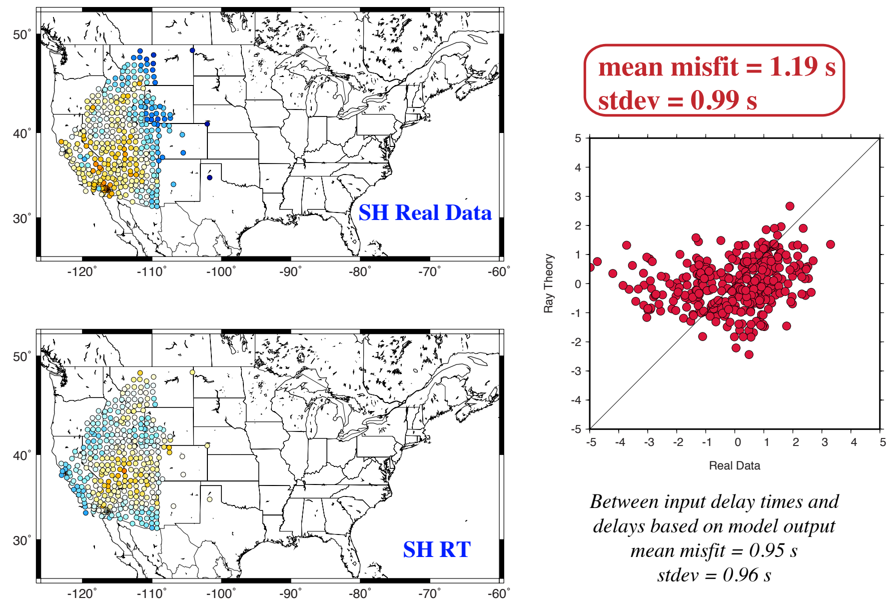
Seismic Imaging Methods:

- no significant differences between RT and FF at intermediate periods (50-200s)
- Differences start to appear for periods < 15s.

Arrival Time of the S wave (1)



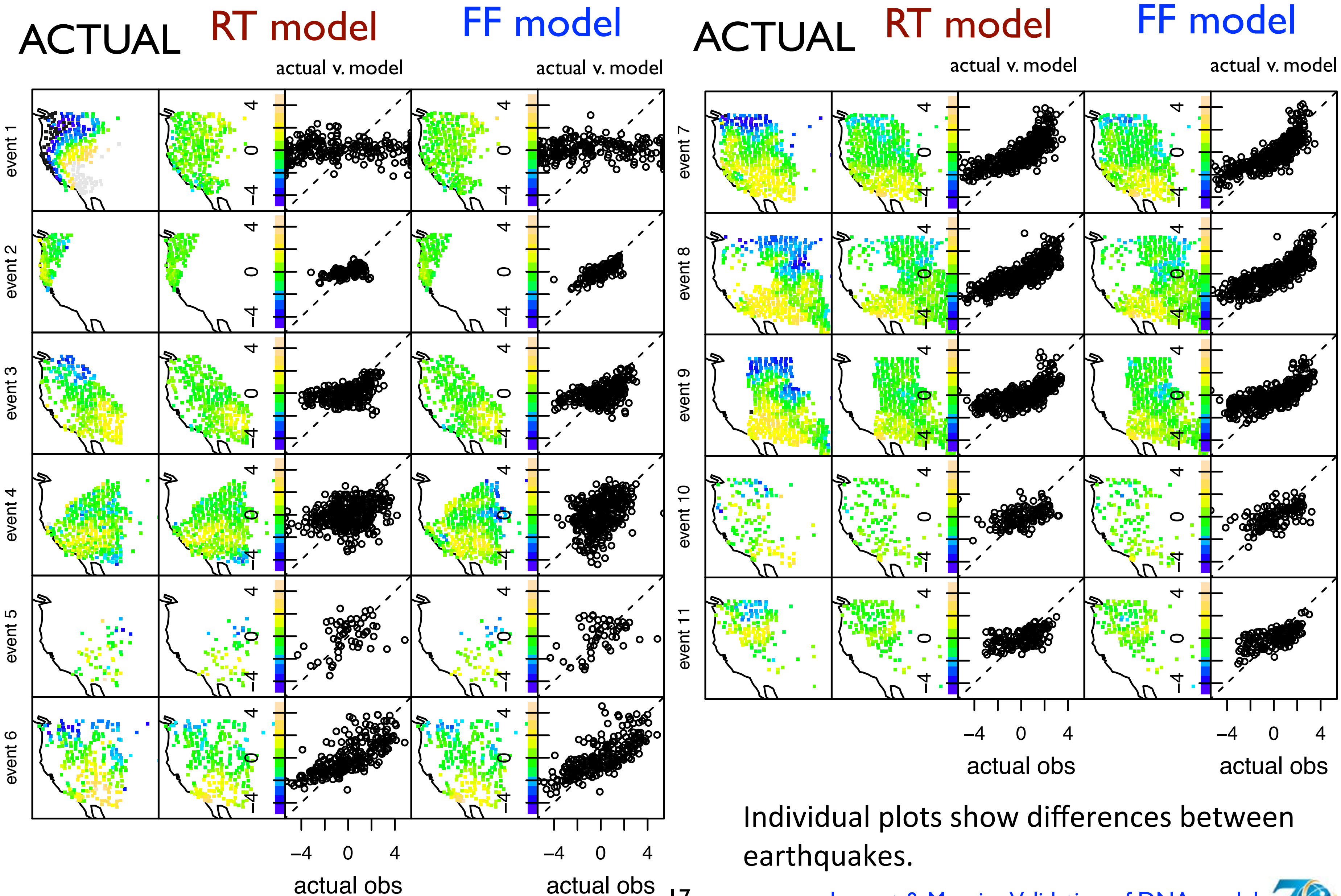
Comparing S-wave delay times between FF synthetics and earthquake data for the Northern Chile event



Comparing S-wave delay times between RT synthetics and earthquake data for the Northern Chile event

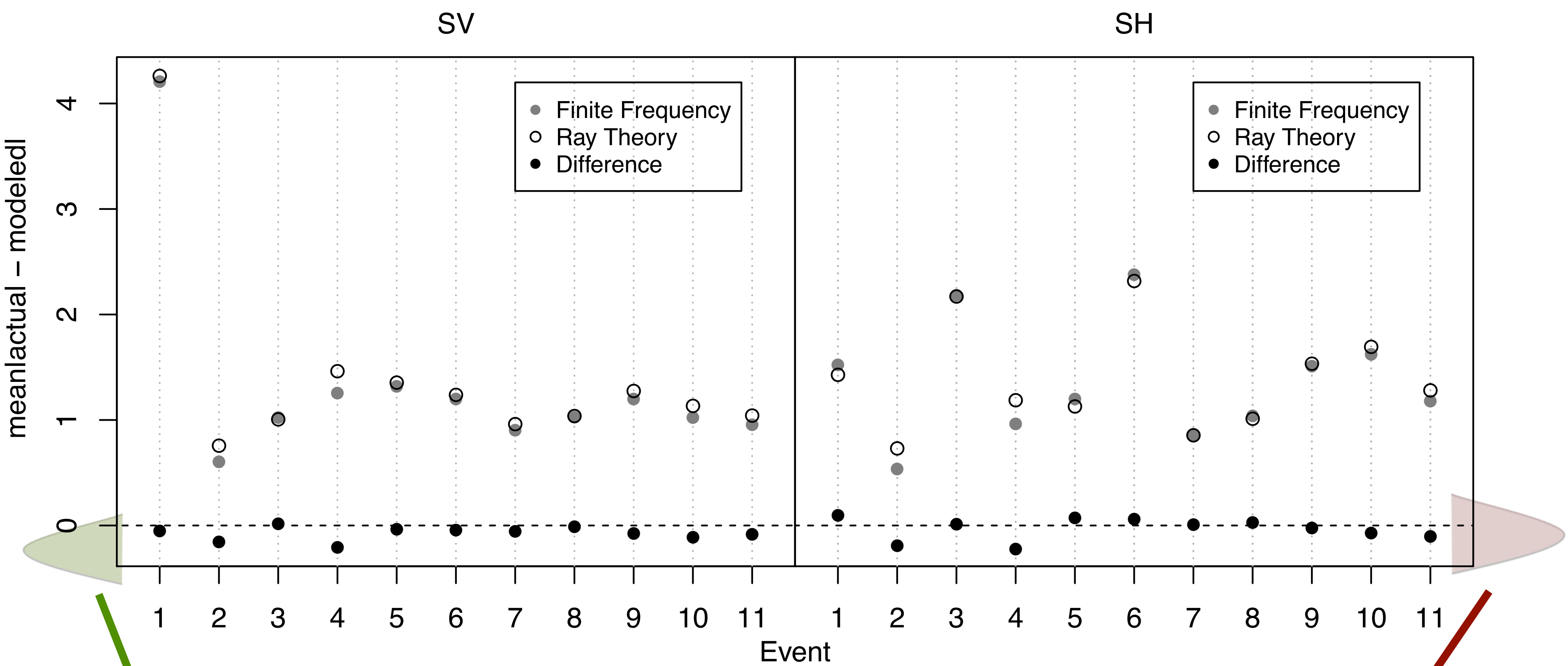
Similar clustering for RT and FF, maybe more correlation for FF.

Arrival Time of the S wave (2)



Individual plots show differences between earthquakes.

“Matched pair” two-sided t-test



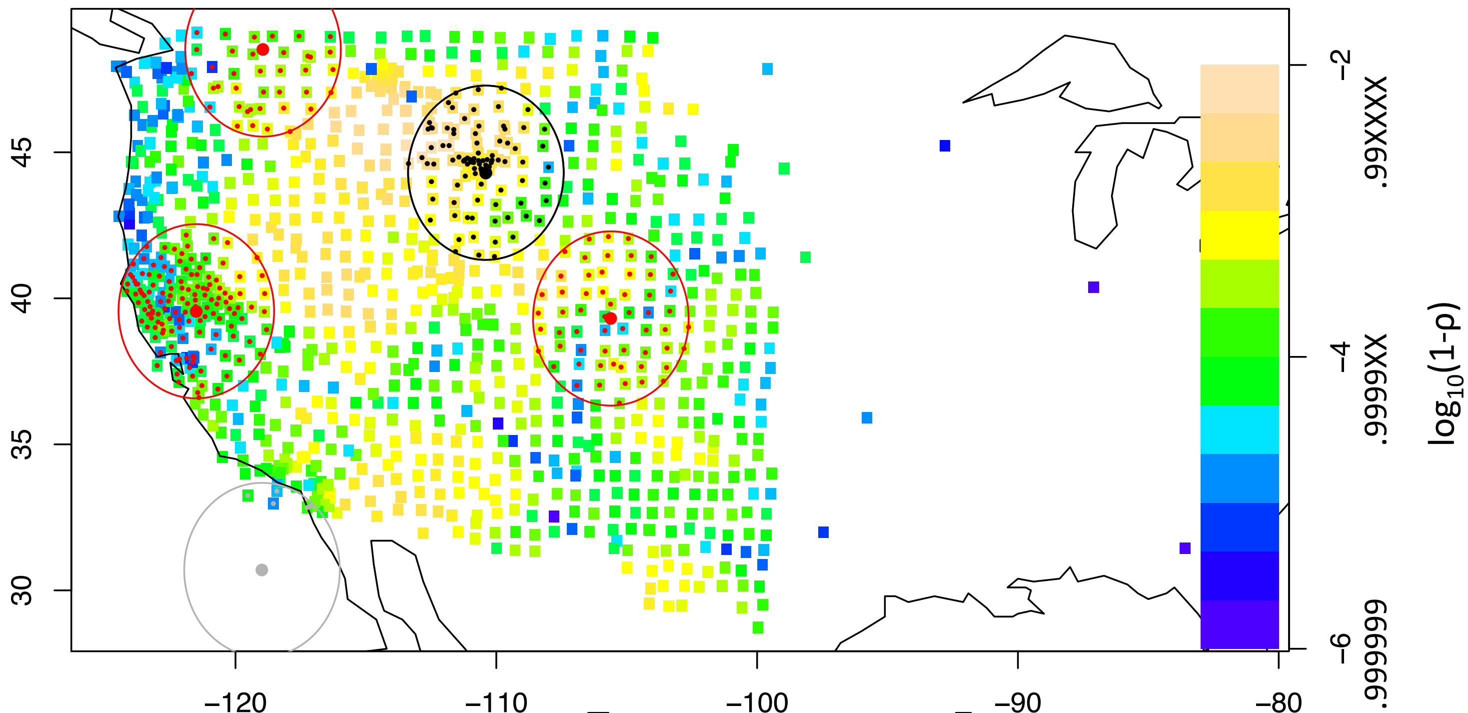
p=0.003

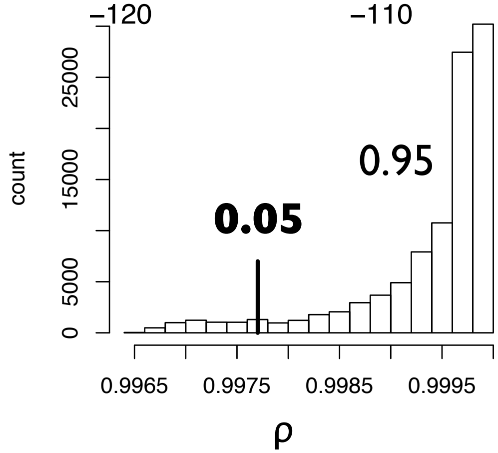
Significant! Delay times from FF model are closer to the observed delay times

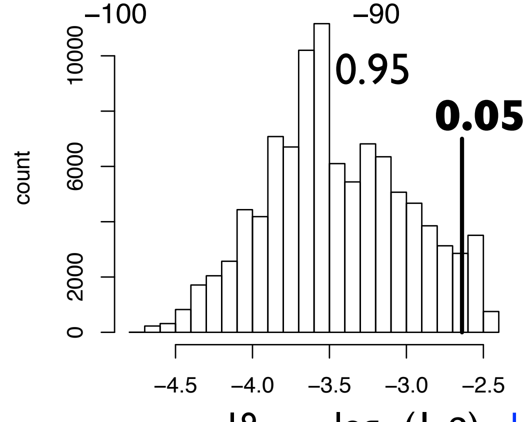
p=0.4

Cross-correlation S-wave (1)

focussing on Yellowstone hot-spot





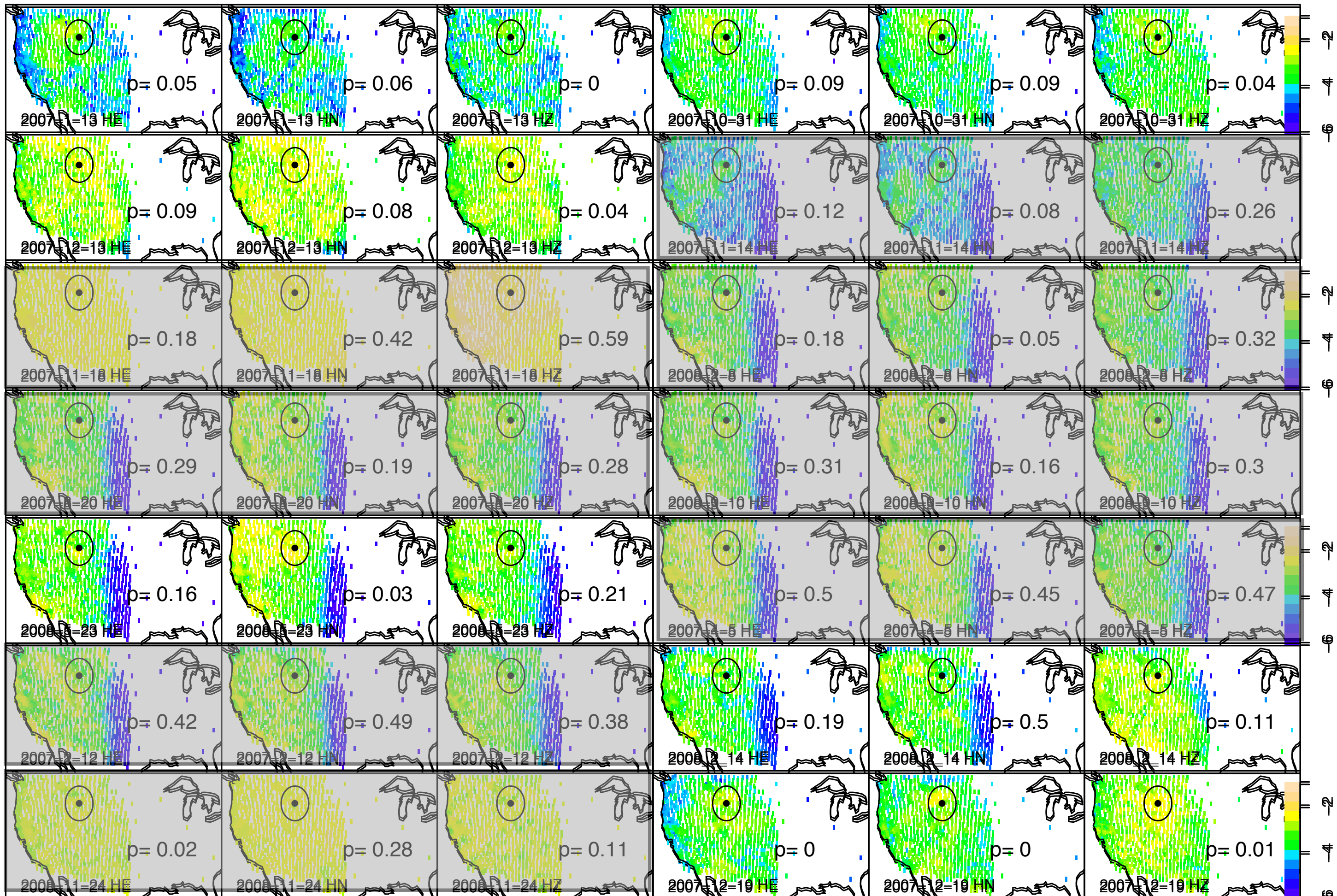


1,000,000 random locations

- event 1: Yellowstone is in the first 5th percentile.

Cross-correlation S-wave (2)

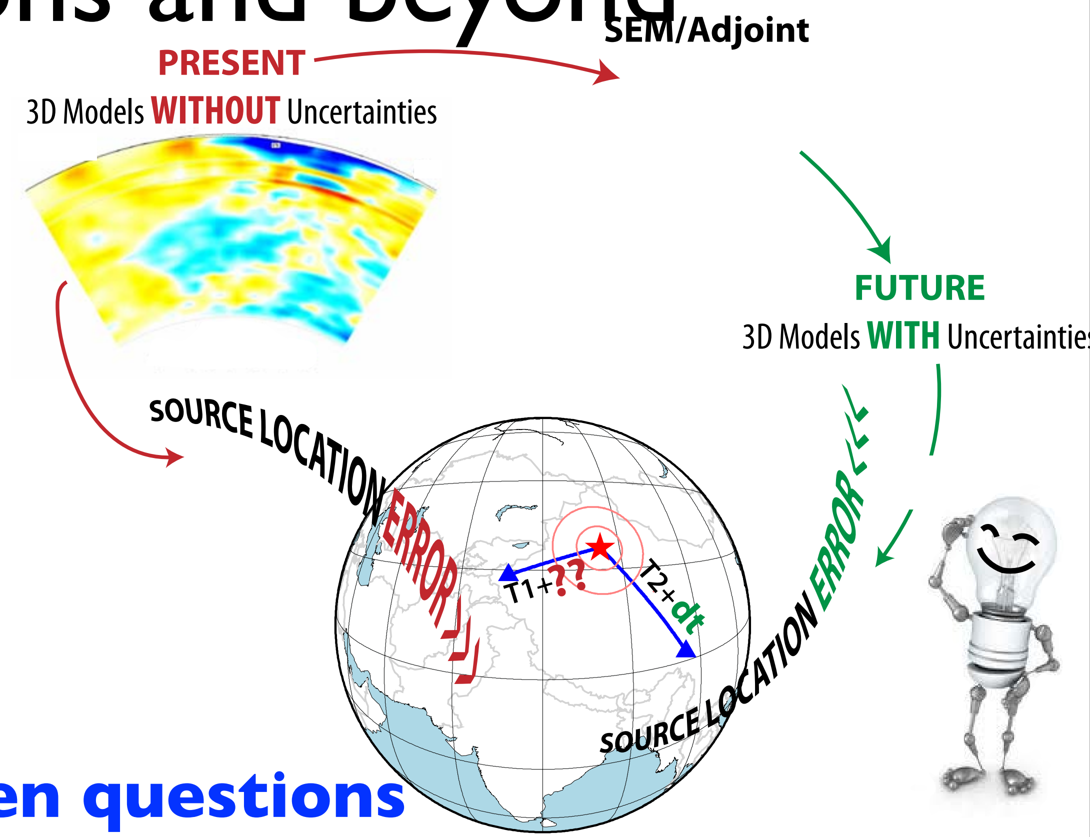
focussing on Yellowstone hot-spot



−6: cross cor = .999999, −4: cross cor = .9999XX,−2: cross cor = .99XXXX

1. Kuril 2007-1-13
2. Mariana 2007--10-31
3. Samoa 2007-12-13
9. Central-Mid-Atlantic 2008-5-23
12.Greece 2008-2-14
14.Aleutians 2007-12-19
4. Chile 2007-11-14
5. Jujuy 2007-11-18
6. NorthernMid-Atlantic 2008-2-8
7. Central-Mid-Atlantic 2007-8-20
8. Central-Mid-Atlantic 2008-9-10
10.Azores 2007-4-5
11.Azroes-Cape 2007-2-12
13. Sea-of-Ohkost 2008-11-24

Conclusions and beyond



Validation of Dynamic North America Models: Benefit from finite frequency***?***

- No conclusive statistical result.
- What is a fair comparison between two different methods?

Open questions

I. Accuracy: How dependent is a model on its numerical methods?

II. Sensivity&resolution: How many data are needed for a model to be considered "good enough"?

III. Parameterization&Uncertainty: How well can we uniquely identify sources of error?

IV. How good are 3D geophysical models at representing true physics of the Earth? How closely do models represent the true world?