Full-3D Waveform Tomography

Po Chen

LDEO, Columbia University

&

James Gaherty

LDEO, Columbia University

&

SCEC CME Collaboration

University of Southern California

Outline

- Problem statement and background
- Nonlinearity, data conditioning, time & frequency bootstrapping
- Born & Rytov linearization, generalized data functionals and their Fréchet kernels
- Computational considerations, CPU-hour, disk storage, memory…
- Examples:
  - LA basin full-3D tomography
  - Canadian Cordillera
  - Eastern Eurasia
  - Western US
- Discussion & summary

Linearization

$$\chi^2(\mathbf{m}) = \left\|\Delta\mathbf{d}(\mathbf{m})\right\|_D + \lambda\left\|\mathbf{m}\right\|_M$$

d: Waveform data in time or frequency domain

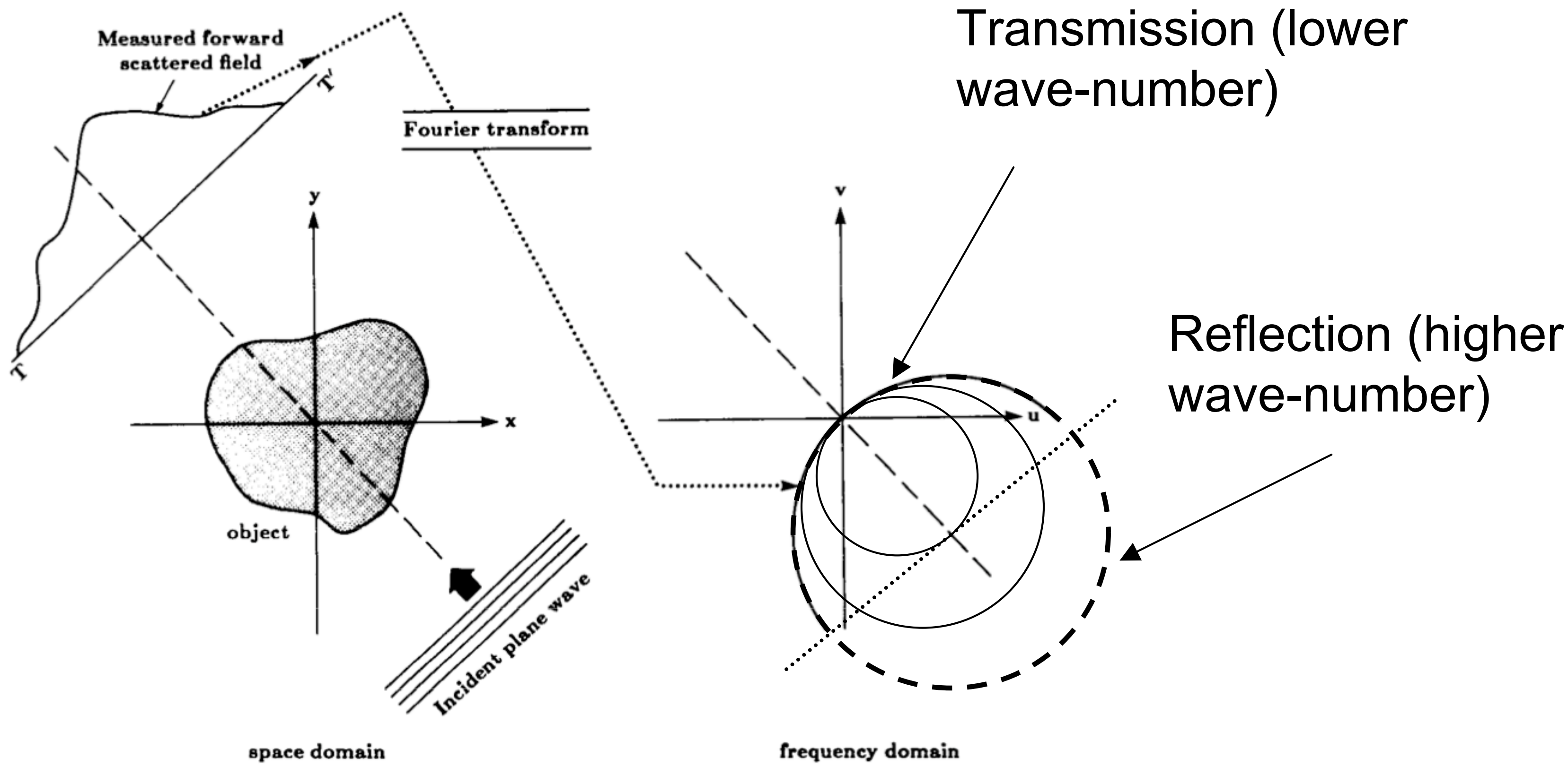
$F[\mathbf{m}]$: Two-way 3D elastic/anelastic wave equation

$\Delta\mathbf{d}$:

$$u_s = \overline{u} - u$$

$$\phi_s = \overline{\phi} - \phi \quad (\overline{u} = e^{\overline{\phi}}, u = e^{\phi})$$

m: 3D in space, smooth or include reflectors



Fourier Diffraction Theorem:

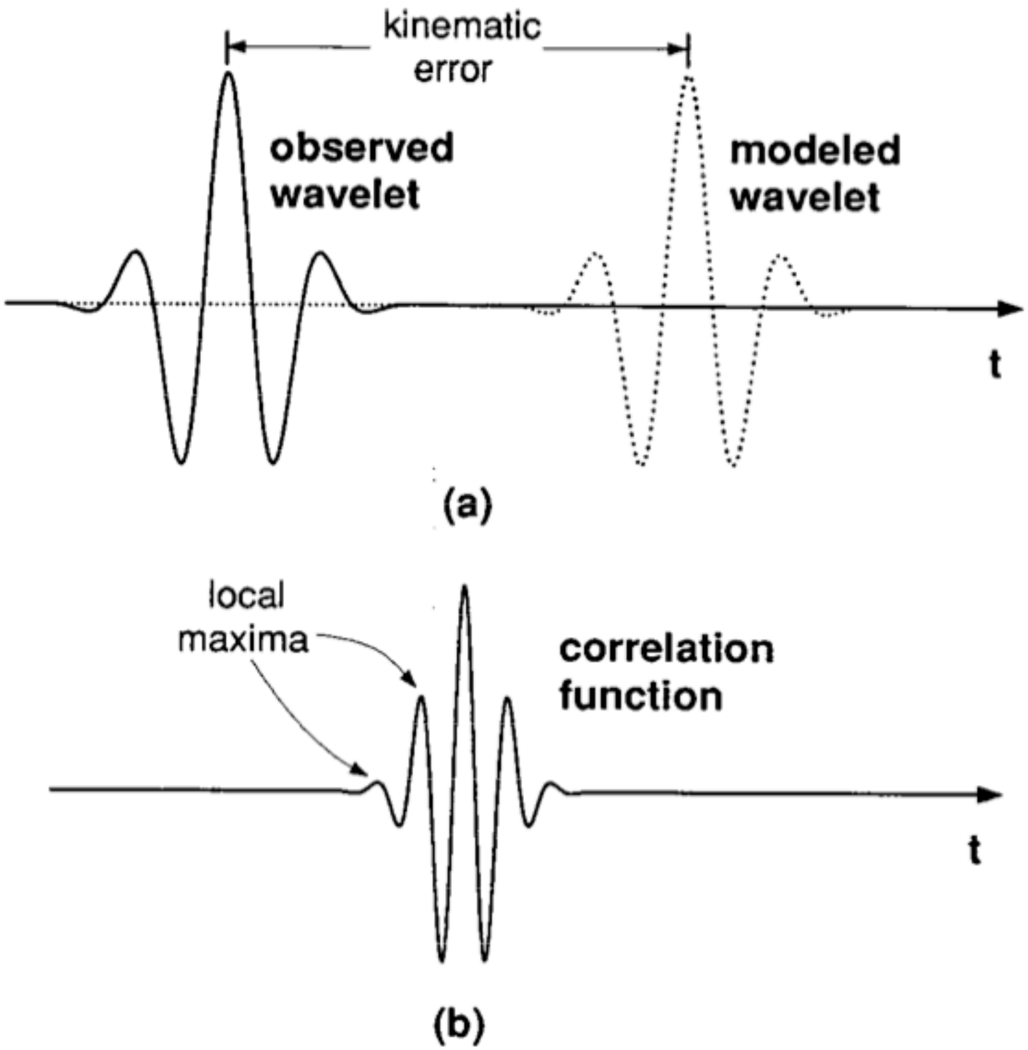
Born $U_s(\Lambda) \approx G(\Lambda)\{O(\Lambda) * U_0(\Lambda)\}$

Rytov $F_{\mathbf{x}\to\Lambda}[\phi_s(\mathbf{x})u_0(\mathbf{x})] \approx G(\Lambda)\{O(\Lambda) * U_0(\Lambda)\}$

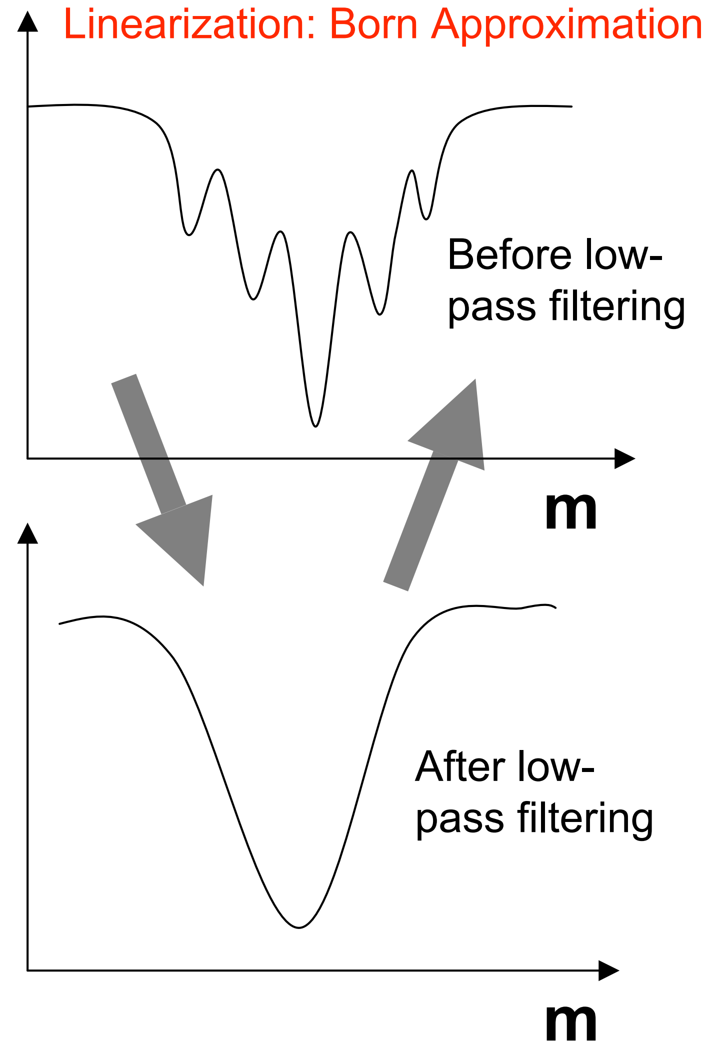
To Cover the Spectrum of the Object

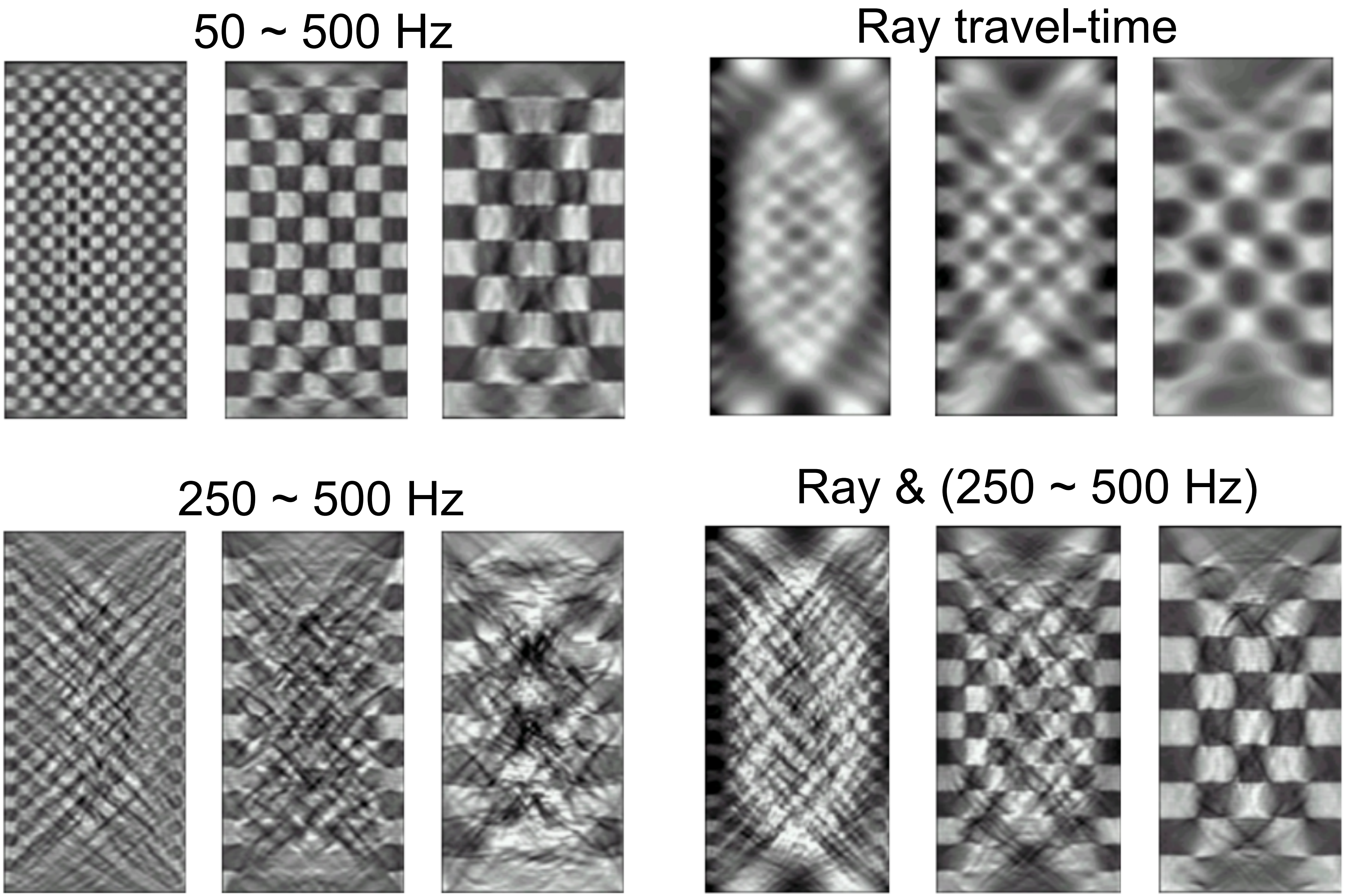
- Transmission + Reflection
  - early arrivals → late arrivals
- Angular diversity
  - large offsets → small offsets
- Frequency diversity
  - low frequency → high frequency

$$\chi^2 = ||u^{data} - u^{model}||_{L2}$$



Bunks *et al*. 1995

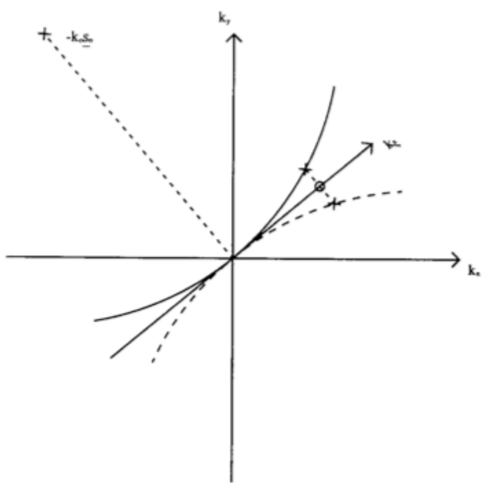




Cross-well transmission waveform tomography (Pratt *et al.* 2002)

Resolution Limit of Ray Travel-time Tomography

- error in physics: infinite-frequency assumption for finite-frequency data
- In-complete information: only travel-time, not complete waveform

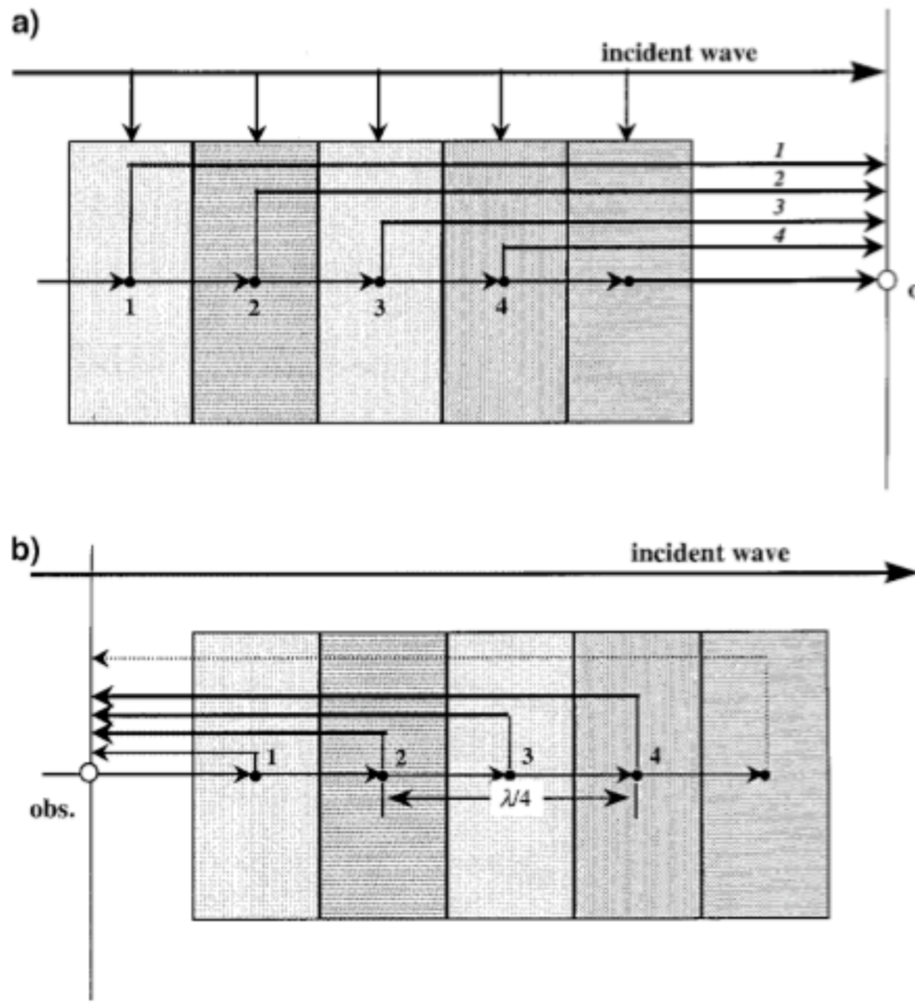


Resolution limit: ~ size of the first Fresnel zone

$$\sqrt{L\lambda}$$

Williamson, 1991

Forward-scattering and Back-scattering



Coherent interference among incident and scattered waves

Summation of scattered waves

Wu 1989

Validity Conditions of Born and Rytov Approximations

Born:

$$u(\mathbf{x}) = u_0(\mathbf{x}) + u_s(\mathbf{x})$$

$$a \cdot n_\delta < \frac{\lambda}{4}$$

Good for back-scattering

Rytov:

$$u(\mathbf{x}) = u_0(\mathbf{x}) \exp\left[\phi_s(\mathbf{x})\right]$$

$$n_\delta >> \left[\frac{\nabla \phi_s \lambda}{2\pi}\right]^2$$

Good for forward-scattering

Chernov 1960; Tatarskii, 1971; Ishimaru, 1978;
Wu 1989; Kak & Slaney 1999;

Generalized Data Functionals and Exact Fréchet Kernels

Data functional:

$$d_{in}^{sr} = D_n\left[u_i^s(\mathbf{x}_r, t),\ \tilde{u}_i^s(\mathbf{x}_r, t)\right]$$

Seismogram perturbation kernel:

$$\delta d_{in}^{sr} = \int dt\, \boxed{J_{in}^{sr}(t)}\, \delta u_i^s(\mathbf{x}_r, t)$$

Fréchet kernel:

$$K_{d_{in}^{sr}}^{c_{jklm}}(\mathbf{x}) = -\int dt \int d\tau\, J_{in}^{sr}(t)\, \partial_k \boxed{G_{ji}(\mathbf{x}, t-\tau; \mathbf{x}_r)}\, \partial_l u_m^s(\mathbf{x}, \tau)$$

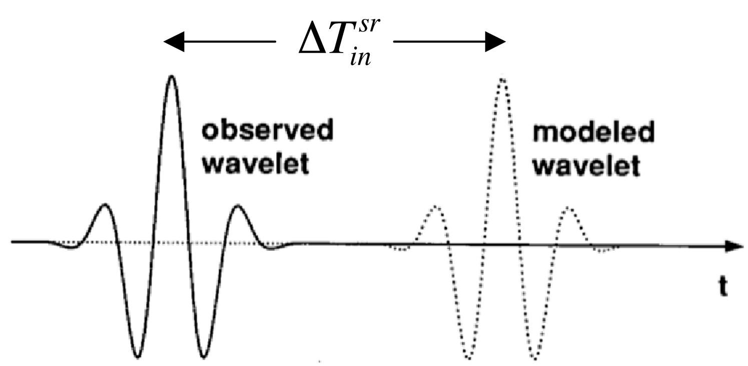
Receiver Green Tensor

- Differential Waveform:

$$d_{in}^{sr} = u_i^s(\mathbf{x}_r, t_n) - \tilde{u}_i^s(\mathbf{x}_r, t_n)$$

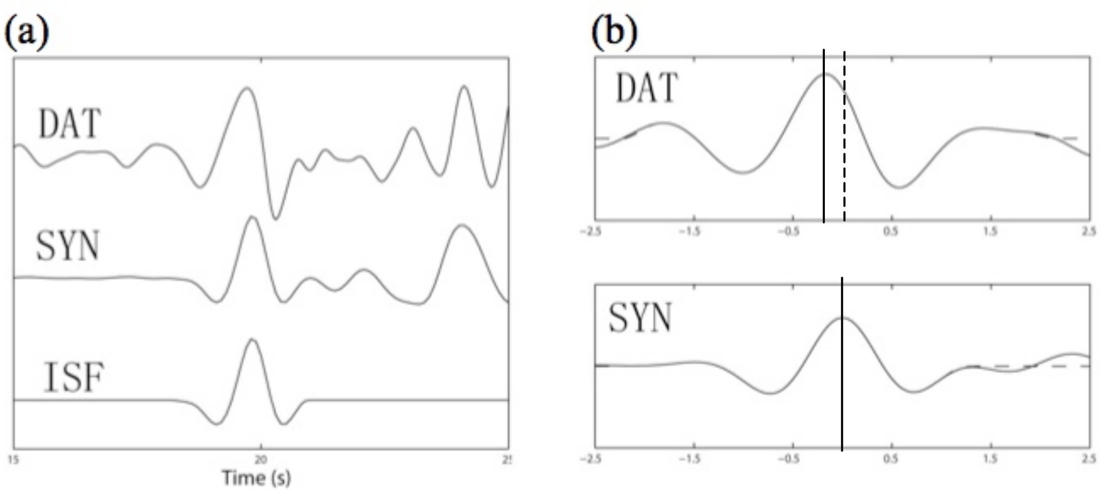
$$J_{in}^{sr}(t) = \delta(t - t_n)$$

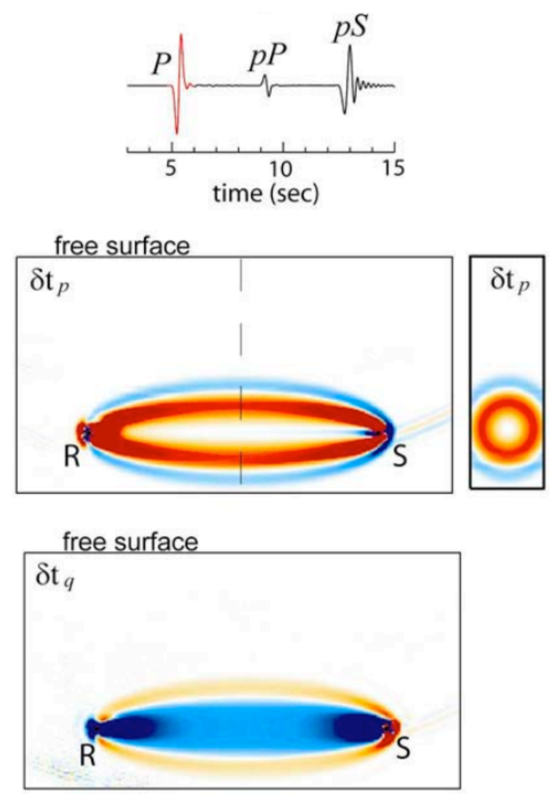
Born validity condition: $\omega_0 \, \Delta T_{in}^{sr} \ll 1$



- Broadband cross-correlation travel-time:

$$d_{in}^{sr} = \Delta T_{in}^{sr} \qquad J_{in}^{sr}(t) = -\frac{\partial_t u_i^s(\mathbf{x}_r, t)[H(t - t_n) - H(t - t_n')]}{\int_{t_n}^{t_n'} \left|\partial_t u_i^s(\mathbf{x}_r, t)\right|^2 dt}$$



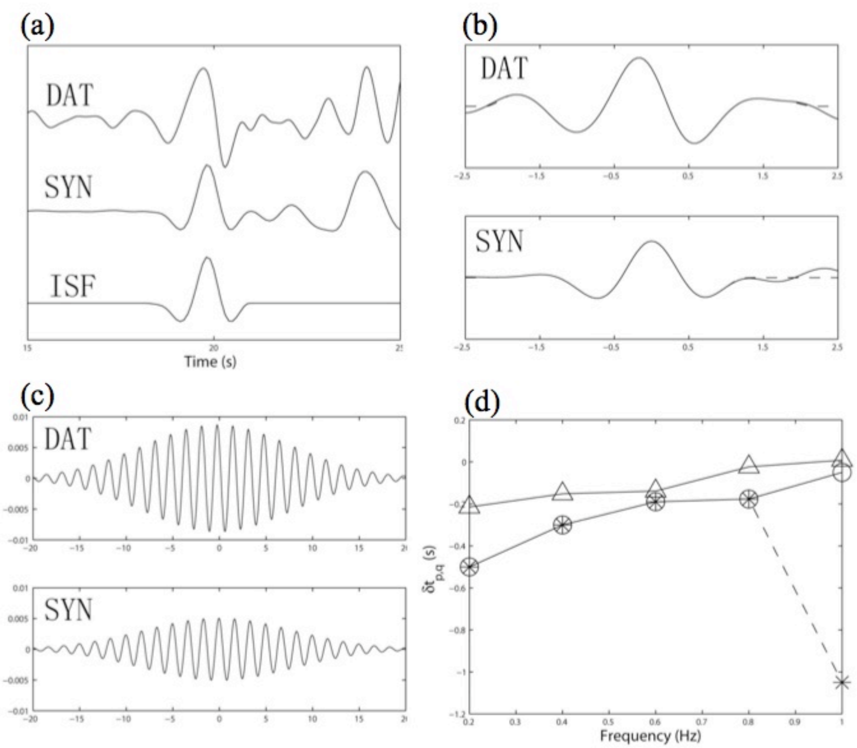


Luo & Schuster 1991; Dahlen *et al.* 2000

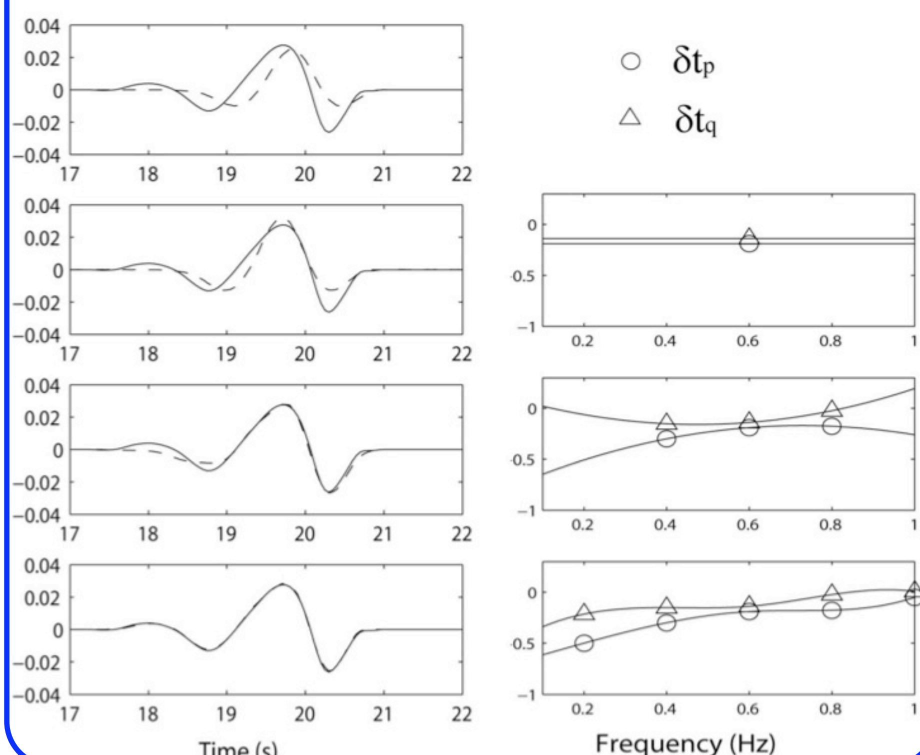
- Generalized seismological data functionals (GSDF):

$$\bar{u}(\omega) = u(\omega)\exp\left[i\omega\delta\tau_{\mathrm{p}}(\omega) - \omega\delta\tau_{\mathrm{q}}(\omega)\right]$$

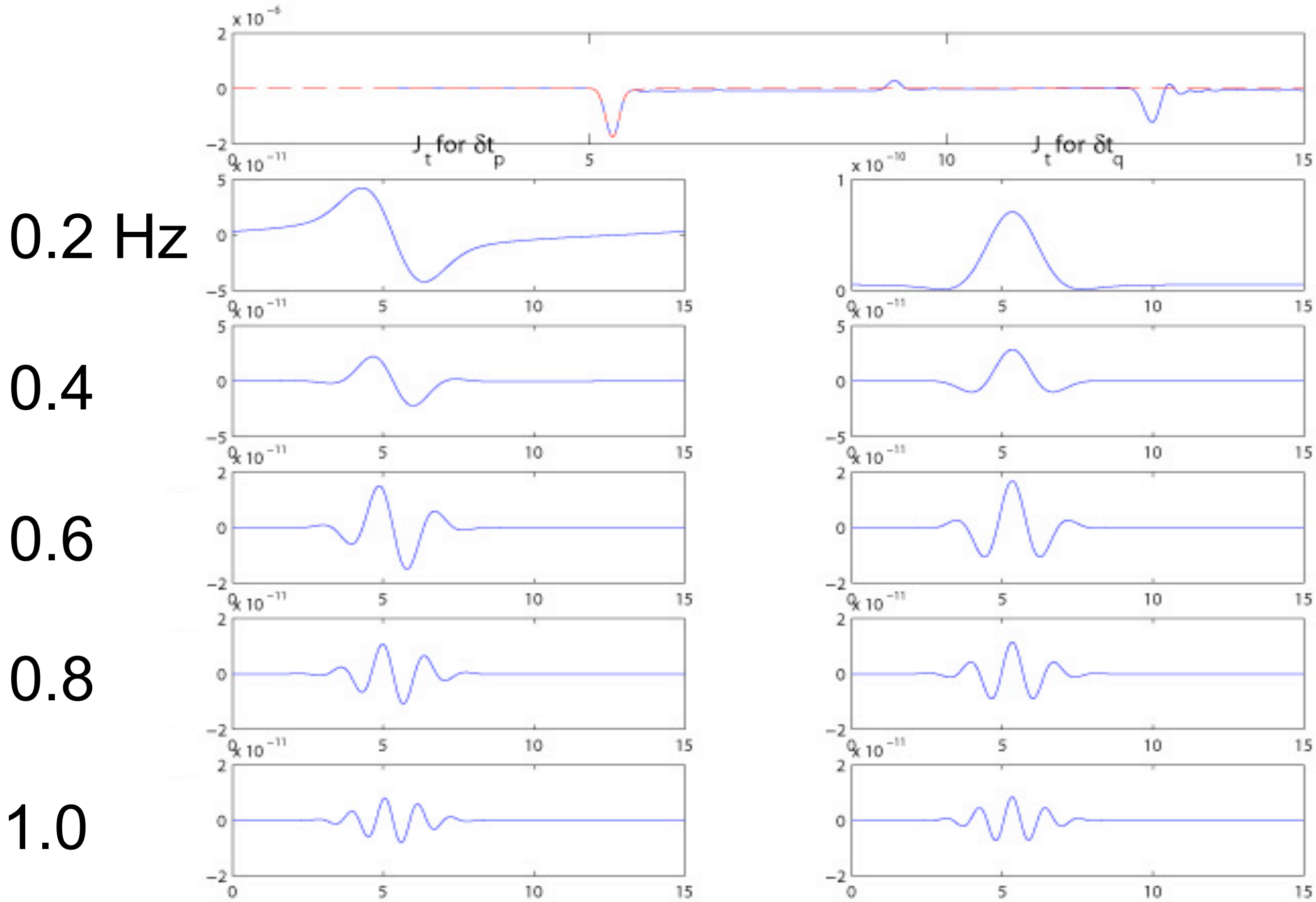
Windowed-Fourier analysis



10 numbers capture waveform difference

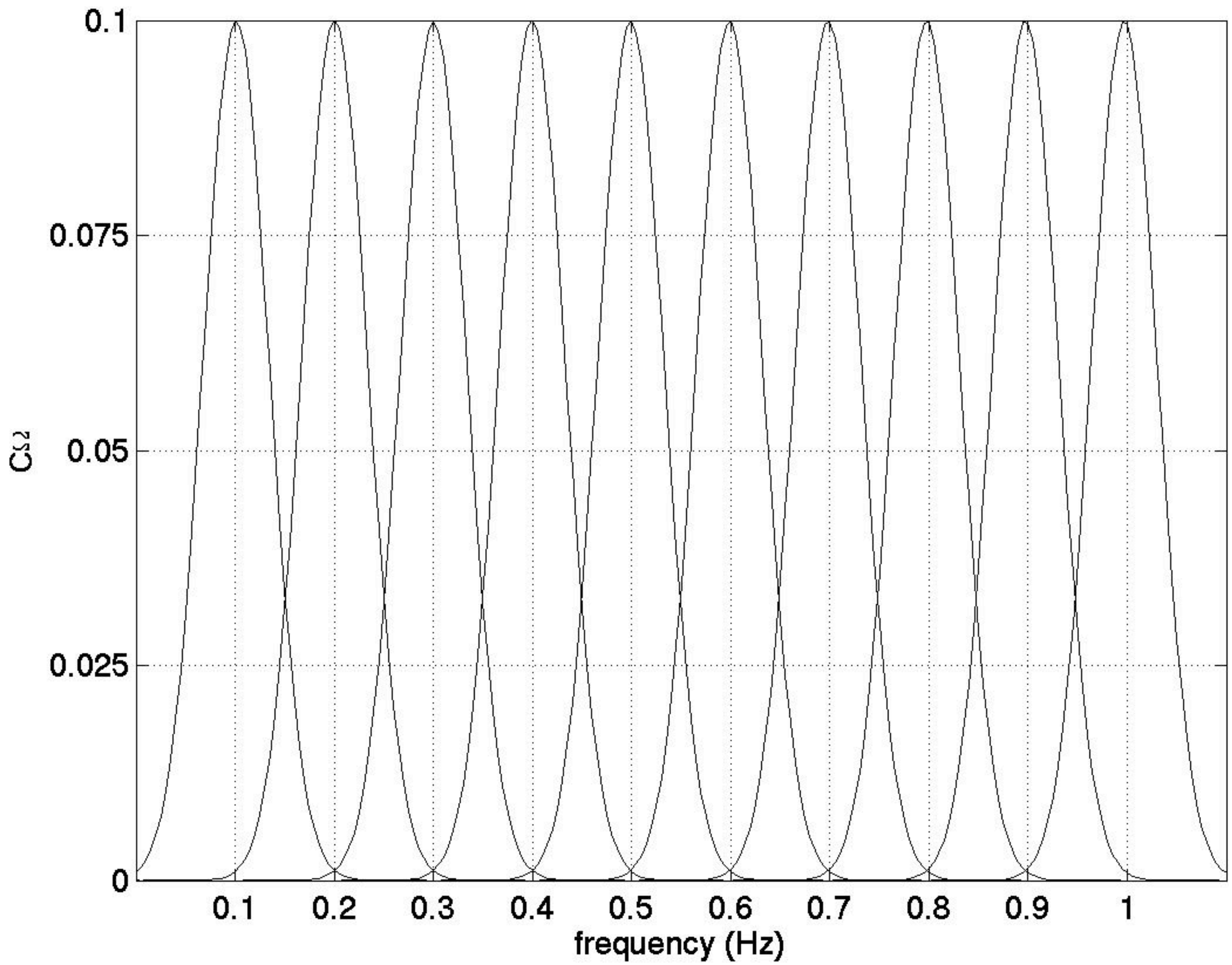


Exact Seismogram Perturbation Kernels for Generalized Data Functionals

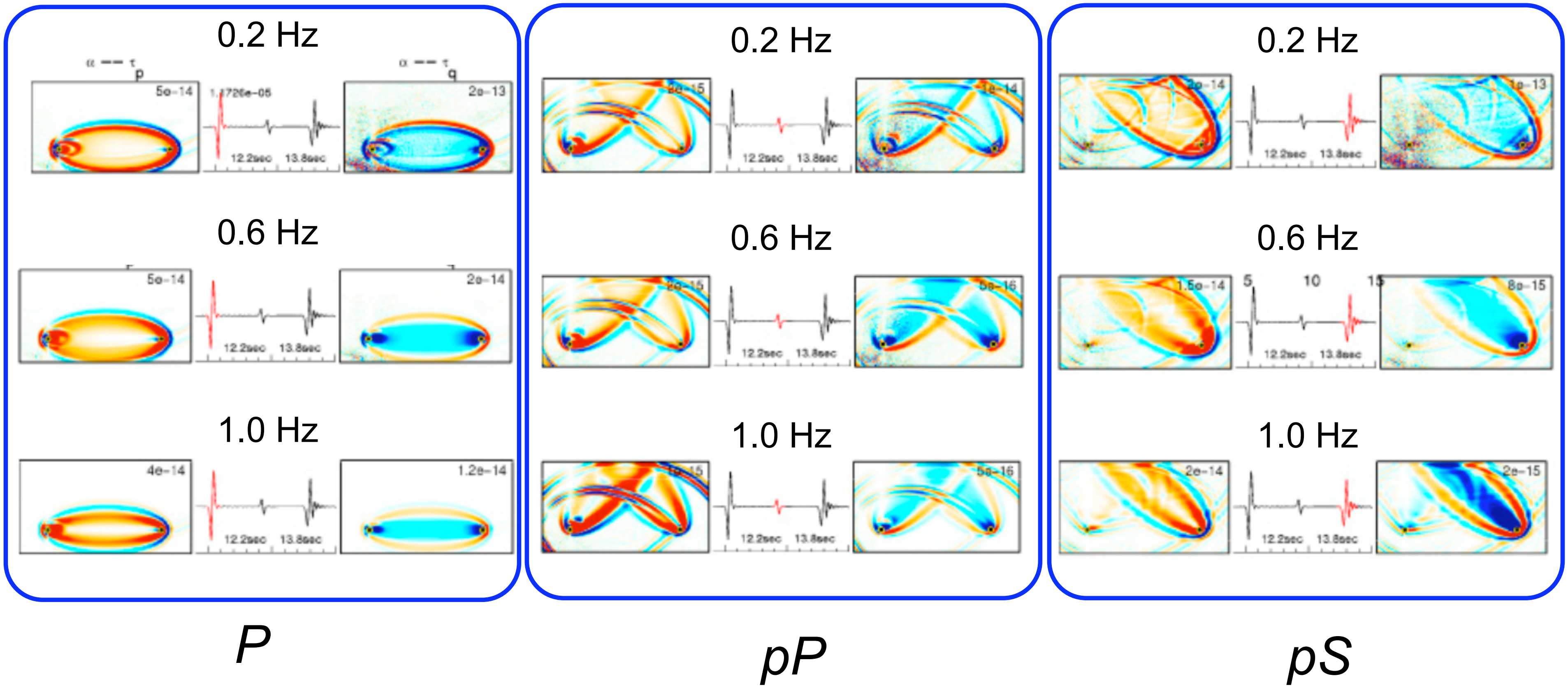


$J(t)$ for $\delta t_p(\omega_n)$ $\quad$ $J(t)$ for $\delta t_q(\omega_n)$

J is also compact in frequency domain



Fréchet Kernels for Generalized Data Functionals (K_α in half-space)

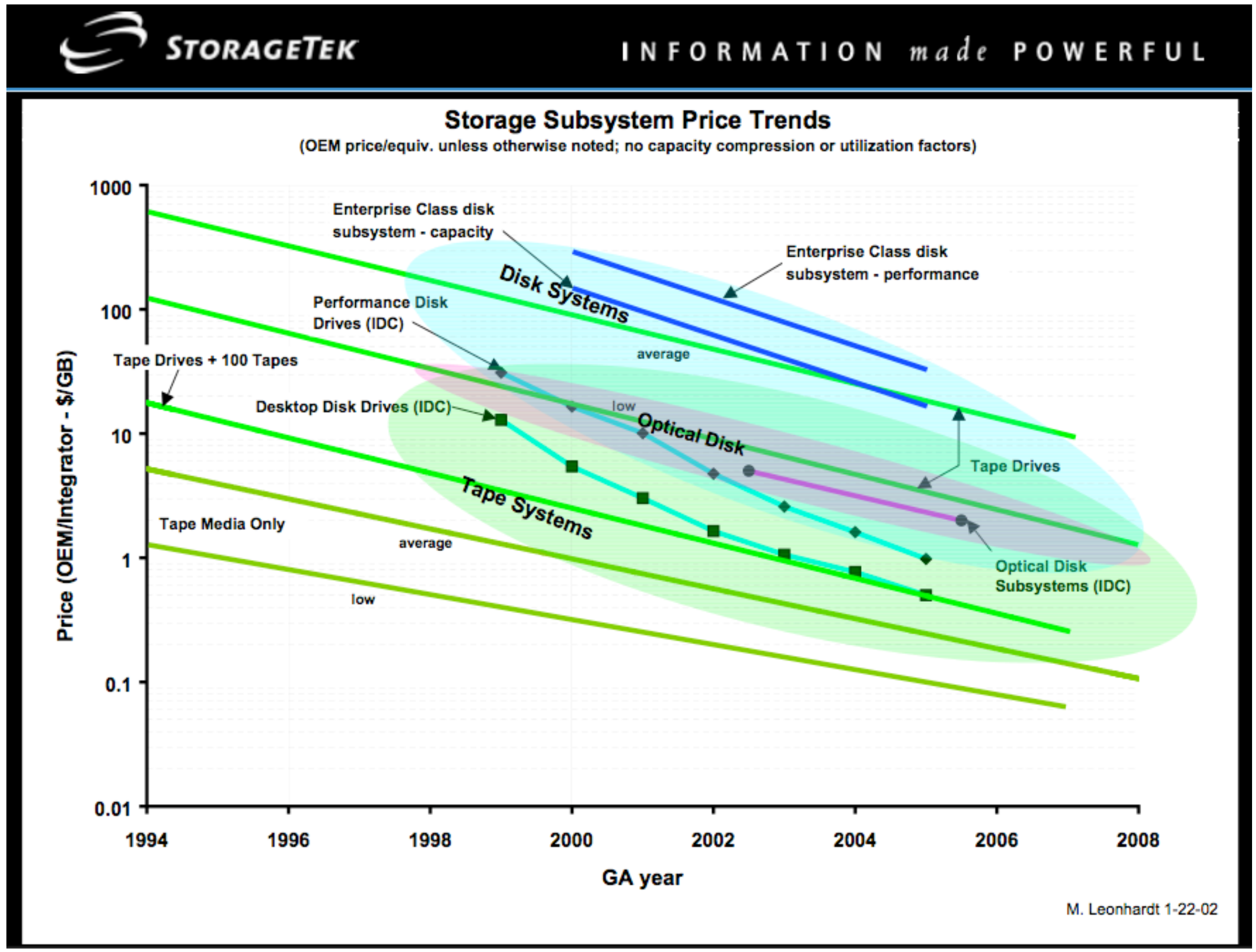


Computation

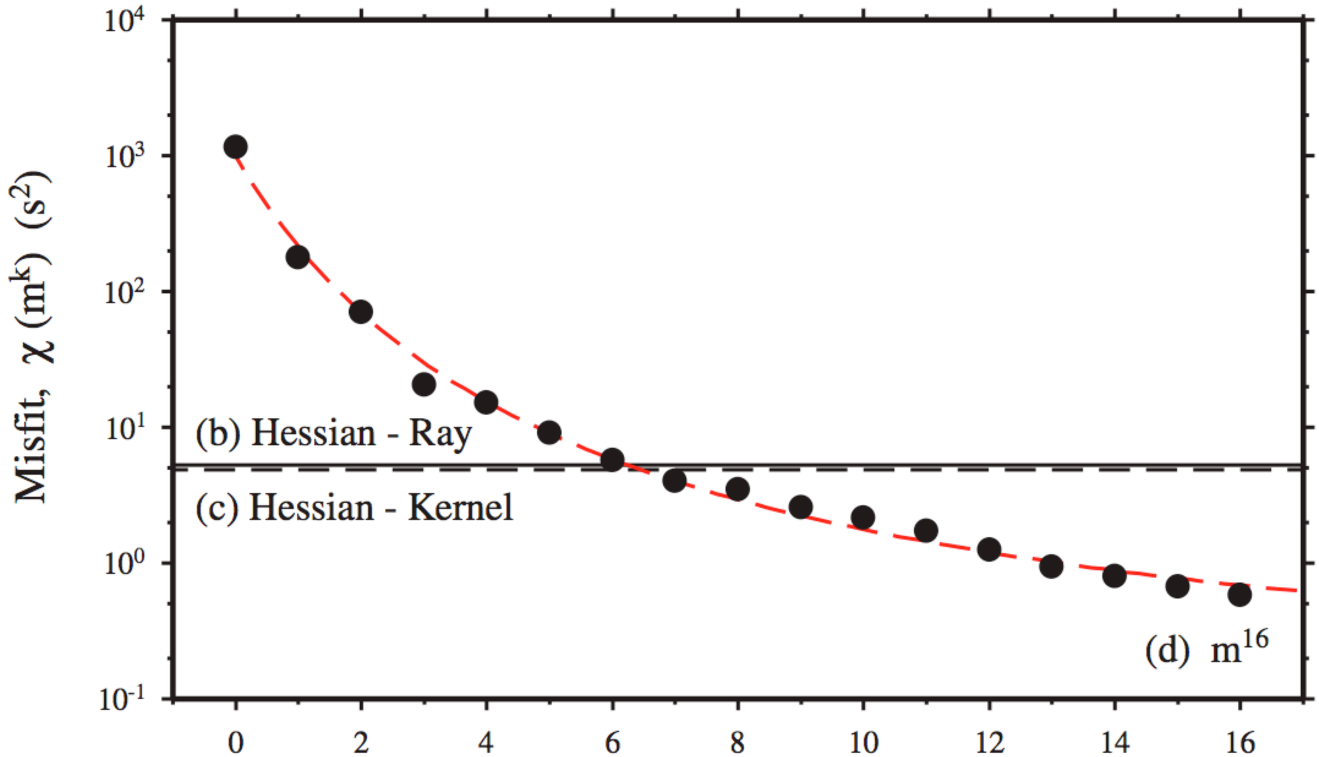
Disk Storage ⇔ CPU-Hour

Storage density is doubling every 12 months since 1994.

Transistor density is doubling every 18 months since 1997.



The Hessian



2D Numerical Experiment (Tape *et al*. 2007)

For 3D problems, 1 GN ~ 15-30 CG

- Misfit functional (least-squares):

$$\chi^2(\mathbf{m}, \tilde{\mathbf{m}}) = \frac{1}{2} \sum_{s=1}^{N_s} \sum_{r=1}^{N_r} \sum_{i,n} \left| d_{in}^{sr}(\mathbf{m}, \tilde{\mathbf{m}}) \right|^2$$

- Misfit functional gradient:

$$K_{\chi^2}^{c_{jklm}}(\mathbf{x}) = -\sum_{s} \int d\tau \, \partial_l u_m^s(\mathbf{x}, \tau) \partial_k \left[u_j^s \right]^+ (\mathbf{x}, -\tau)$$

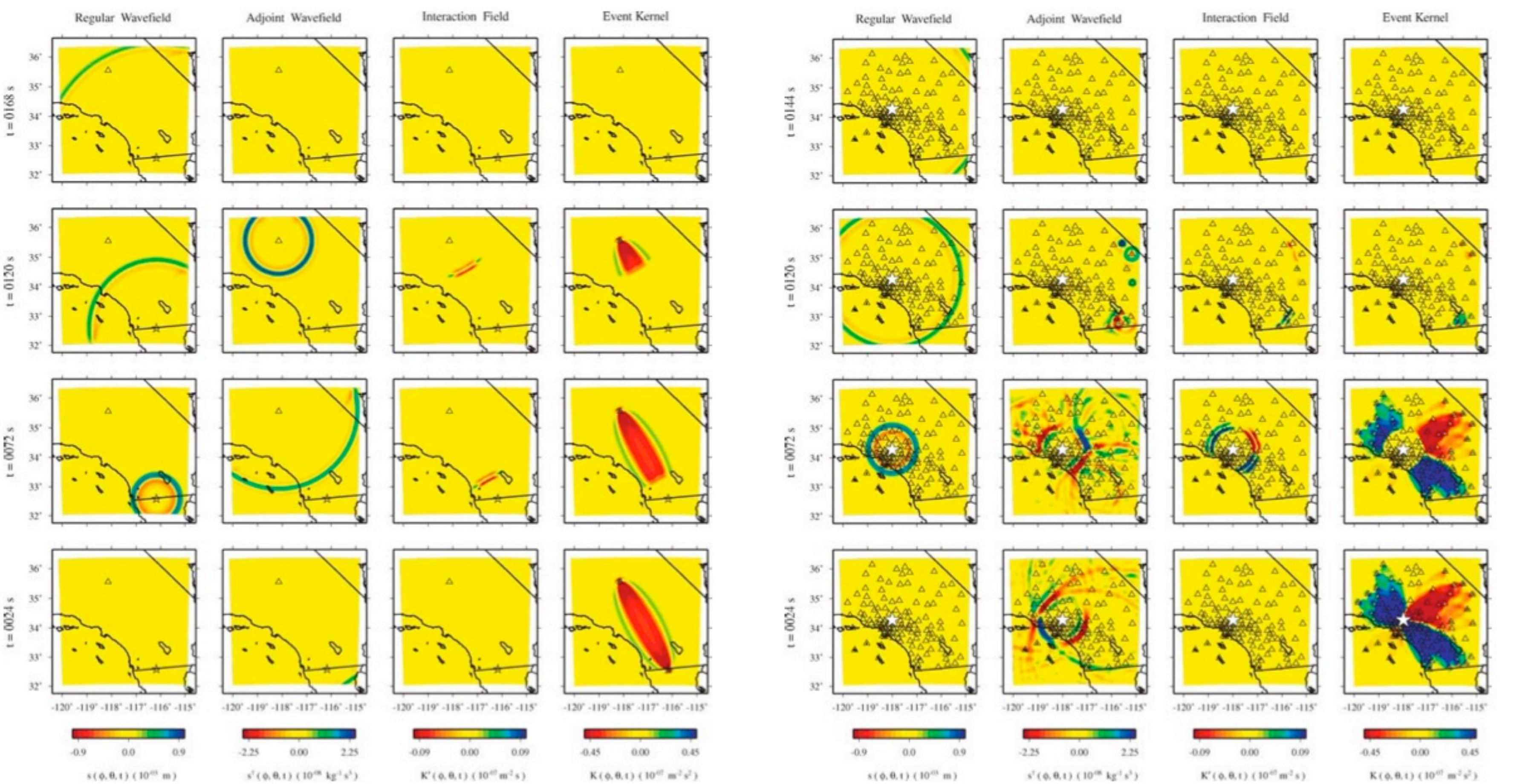
- Adjoint source field:

$$\left[f_i^s \right]^+ (\mathbf{x}, t) = \sum_{rn} J_{in}^{sr}(-t) d_{in}^{sr} \delta(\mathbf{x} - \mathbf{x}_r)$$

- Adjoint wave field:

$$\left[u_j^s \right]^+ (\mathbf{x}, \tau) = \int dt \sum_{rin} \boxed{G_{ji}(\mathbf{x}, \tau - t; \mathbf{x}_r)} J_{in}^{sr}(-t) \, d_{in}^{sr}$$

• 2D example (wave-equation travel-time tomography using adjoint method):



Tape *et al*. 2007

- Data functional Fréchet kernel:

$$K_{d_{in}^{sr}}^{c_{jklm}}(\mathbf{x}) = -\int dt \int d\tau \, J_{in}^{sr}(t)\, \partial_k G_{ji}(\mathbf{x}, t-\tau; \mathbf{x}_r)\, \partial_l u_m^s(\mathbf{x}, \tau)$$

- Relation between misfit functional gradient and data functional Fréchet kernel :

$$K_{\chi^2}^{\mathbf{m}}(\mathbf{x}) = \sum_{srin} d_{in}^{sr} K_{d_{in}^{sr}}^{\mathbf{m}}(\mathbf{x})$$

- Hessian:

$$\mathbf{H} = \nabla_{\tilde{\mathbf{m}}} \nabla_{\tilde{\mathbf{m}}} \chi^2 = \mathbf{A}^{\mathrm{T}} \mathbf{C}_{\mathrm{d}}^{-1} \mathbf{A} + \mathbf{C}_{\mathrm{m}}^{-1} + \boxed{(\nabla_{\tilde{\mathbf{m}}} \mathbf{A})^{\mathrm{T}} \mathbf{C}_{\mathrm{d}}^{-1} \mathbf{d}}$$

- Normal equation:

$$\begin{bmatrix} \mathbf{C}_{\mathrm{d}}^{-1/2} \mathbf{A} \\ \mathbf{C}_{\mathrm{m}}^{-1/2} \end{bmatrix} \delta \mathbf{m} = \begin{bmatrix} \mathbf{C}_{\mathrm{d}}^{-1/2} \mathbf{d} \\ 0 \end{bmatrix}$$

Computational Cost Comparison

Cost	SI method	AW–CG[a]
Storage requirement	$3N_r N_V N_T$	N_V
Number of simulations	$3N_r + N_s$	$6N_s$
Number of time integrations	$2N_t\ N_V\ N_u$	$2\ N_V\ N_s$
I/O cost	$N_u N_T N_V$	$2N_s N_V$
Requires solving a linear system?	Yes	No
Optimization algorithm	Gauss–Newton	Conjugate-Gradient
Number of iterations needed to match one Gauss–Newton step	1	15-30

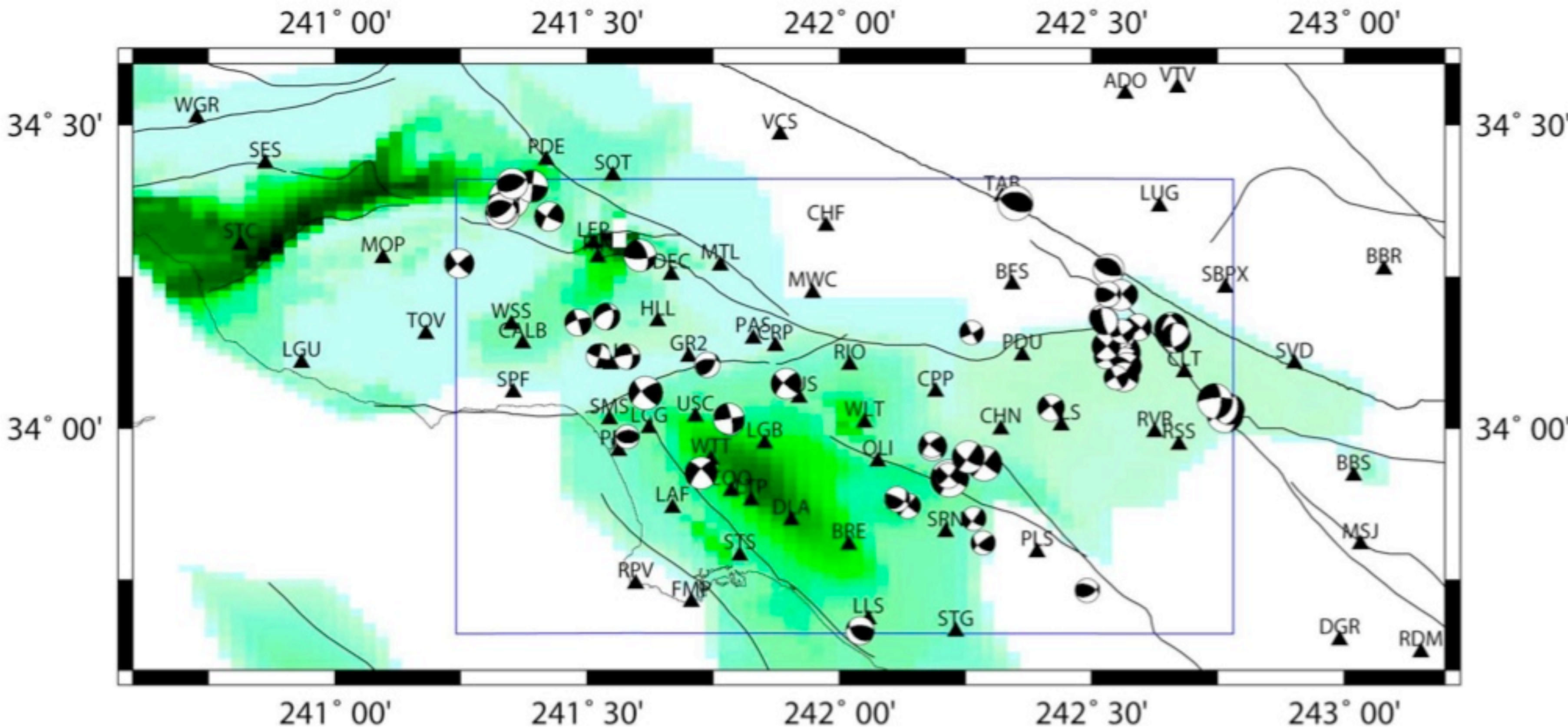
For pre-stack reverse-time migration
$3N_S$ simulations, N_V storage, or
$2N_S$ simulations, $N_V N_T$ storage

Difficulties of the Frequency-Domain Approach

- Time-domain data conditioning (computational cost proportional to number of frequencies)
- Memory expensive for large 3D problems (L, U matrices not very sparse)
- Low parallel efficiency of LU factorization algorithms

Examples

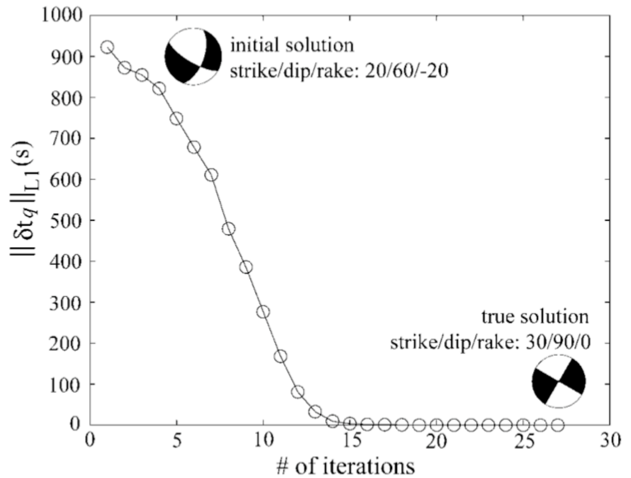
Application to LA Basin



LAB Inversion Computational Cost For One GN Iteration

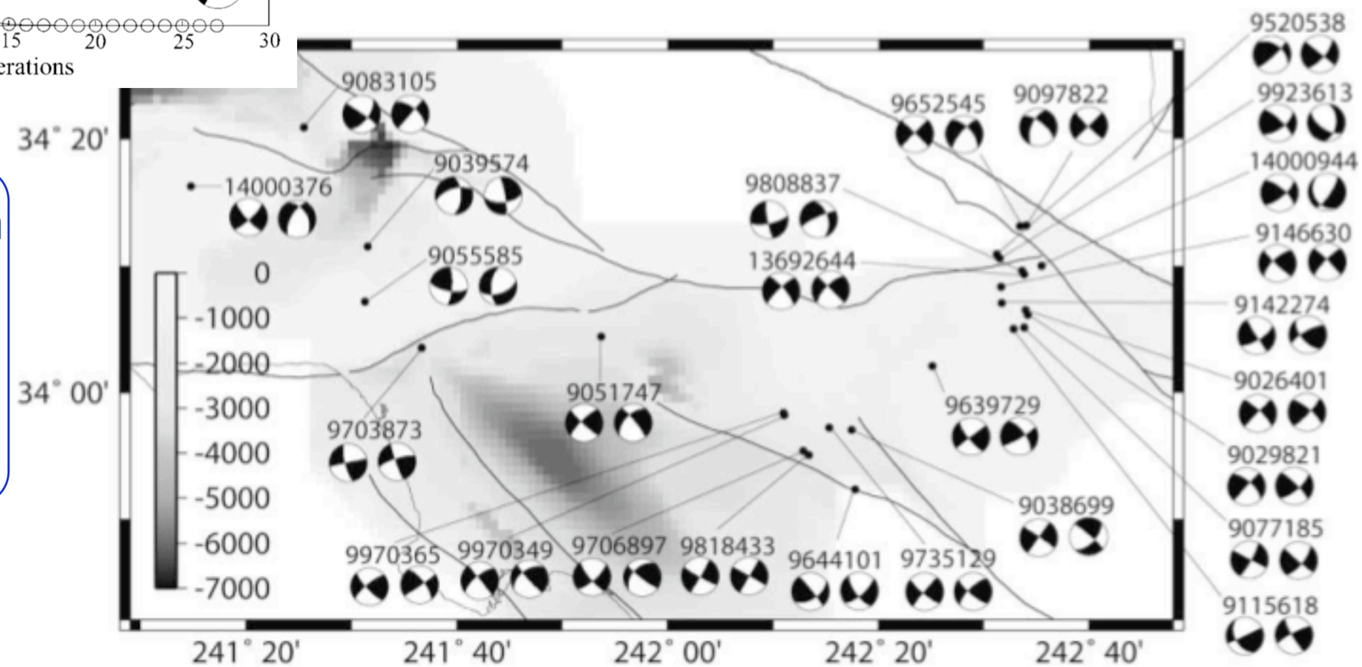
Number of stations N_r	48
Number of earthquakes N_s	67
Number of seismograms N_u	2000
Number of FD simulations $3N_r + N_s$	211
Simulation grid spacing, time interval	200 m, 0.01 s
Simulation grid points N_V, time steps N_T	36 140 440, 6000
Number of CPUs	128
Total CPU time per iteration	62 000 CPU-hours
Total disk space $3N_rN_VN_T$	24 TB

Rapid CMT Inversion Using Waveforms computed in a 3D Earth Structural Model

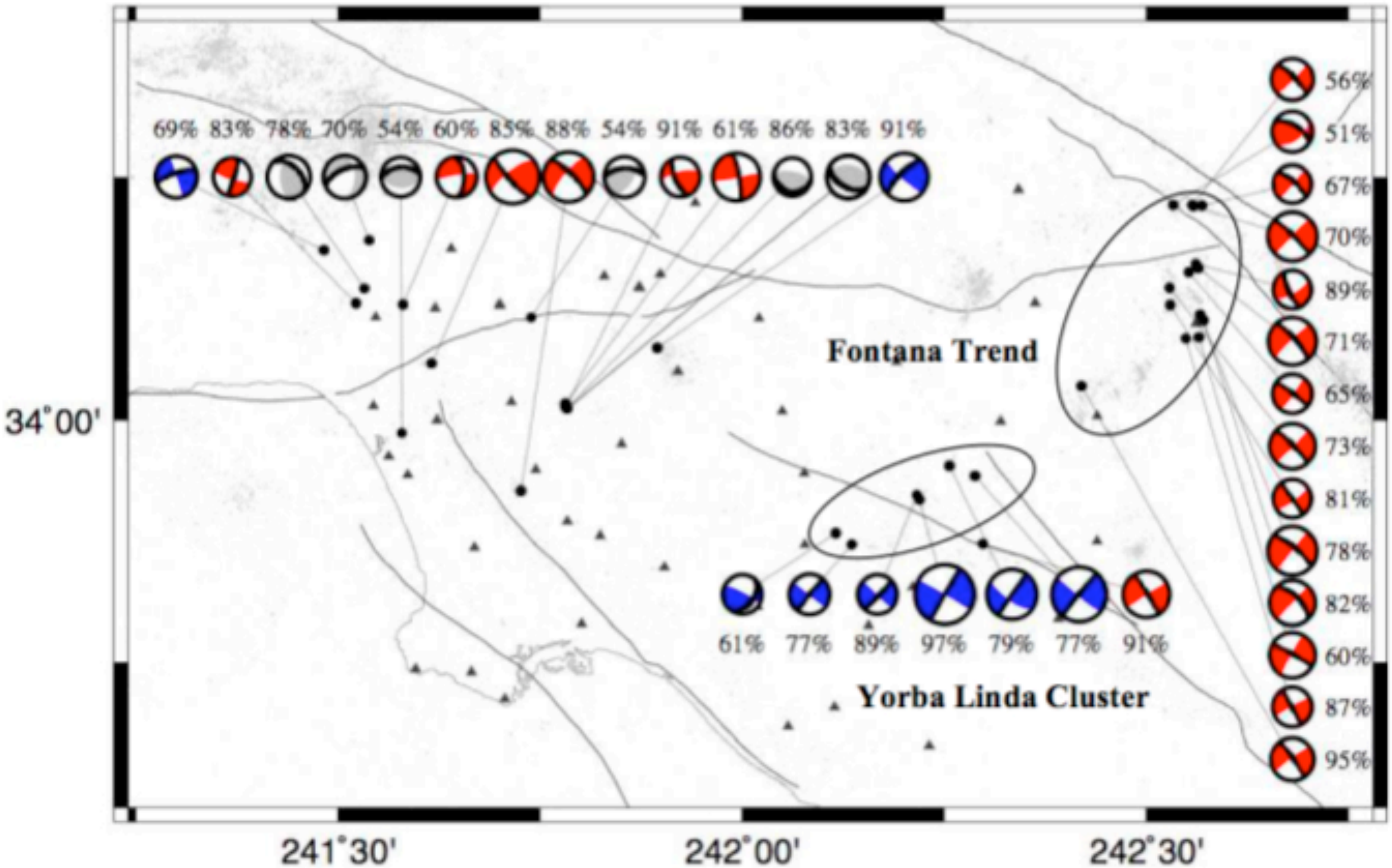


Numerical tests to verify inversion algorithm

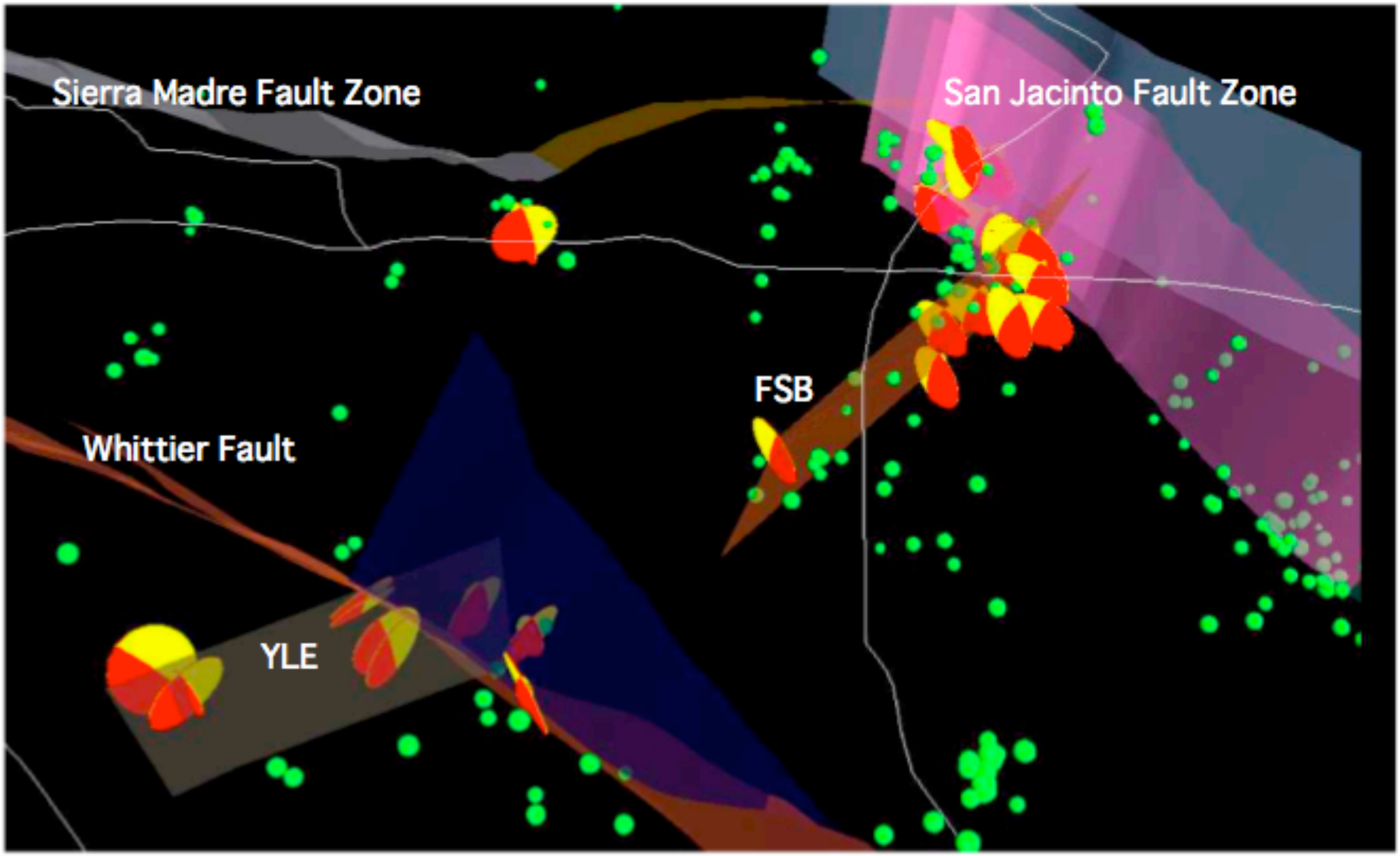
Waveform inversion using 3D RGT synthetics .vs. first-motion focal mechanisms



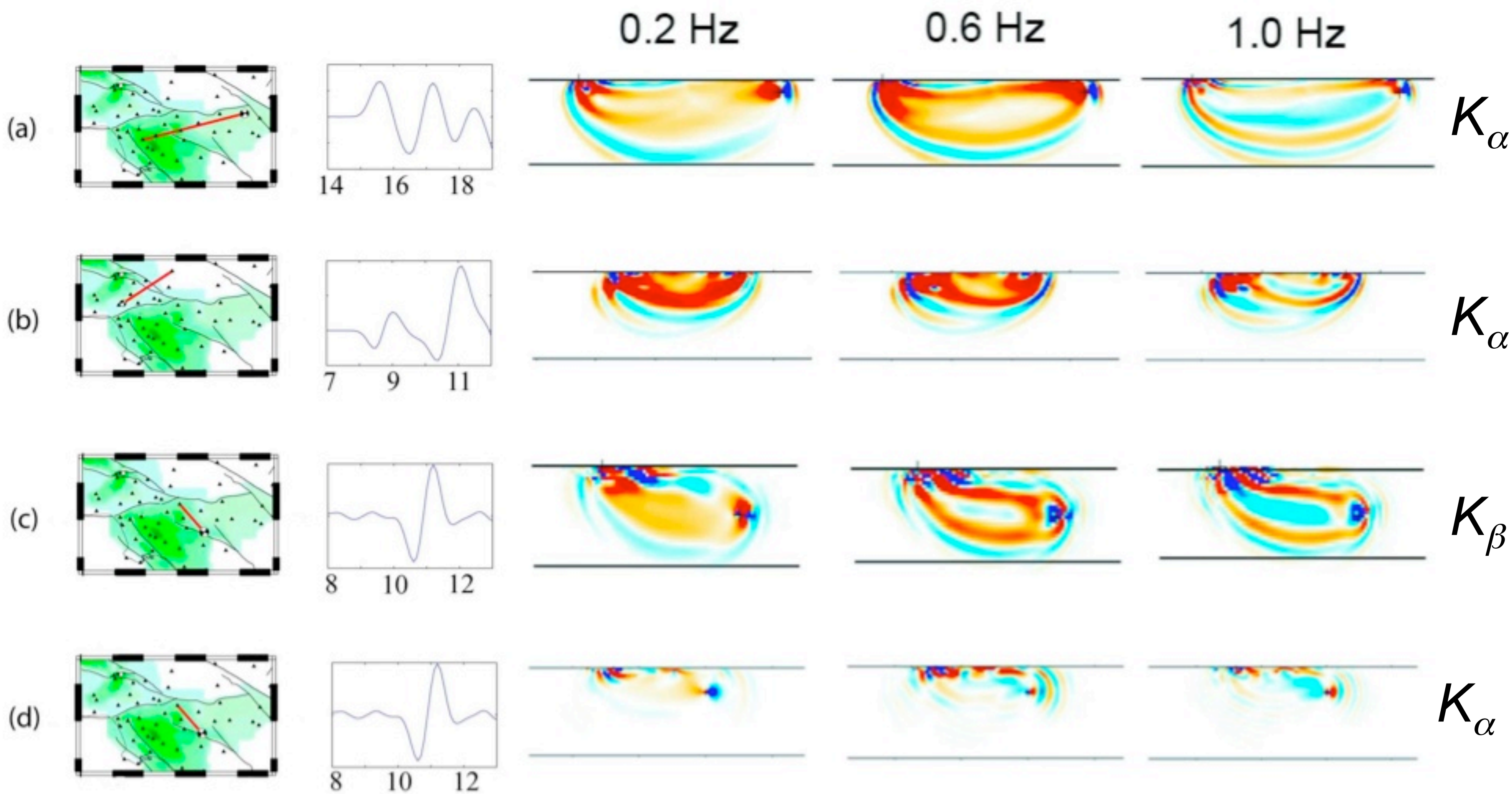
Resolving Fault-plane-ambiguity for Small Earthquakes

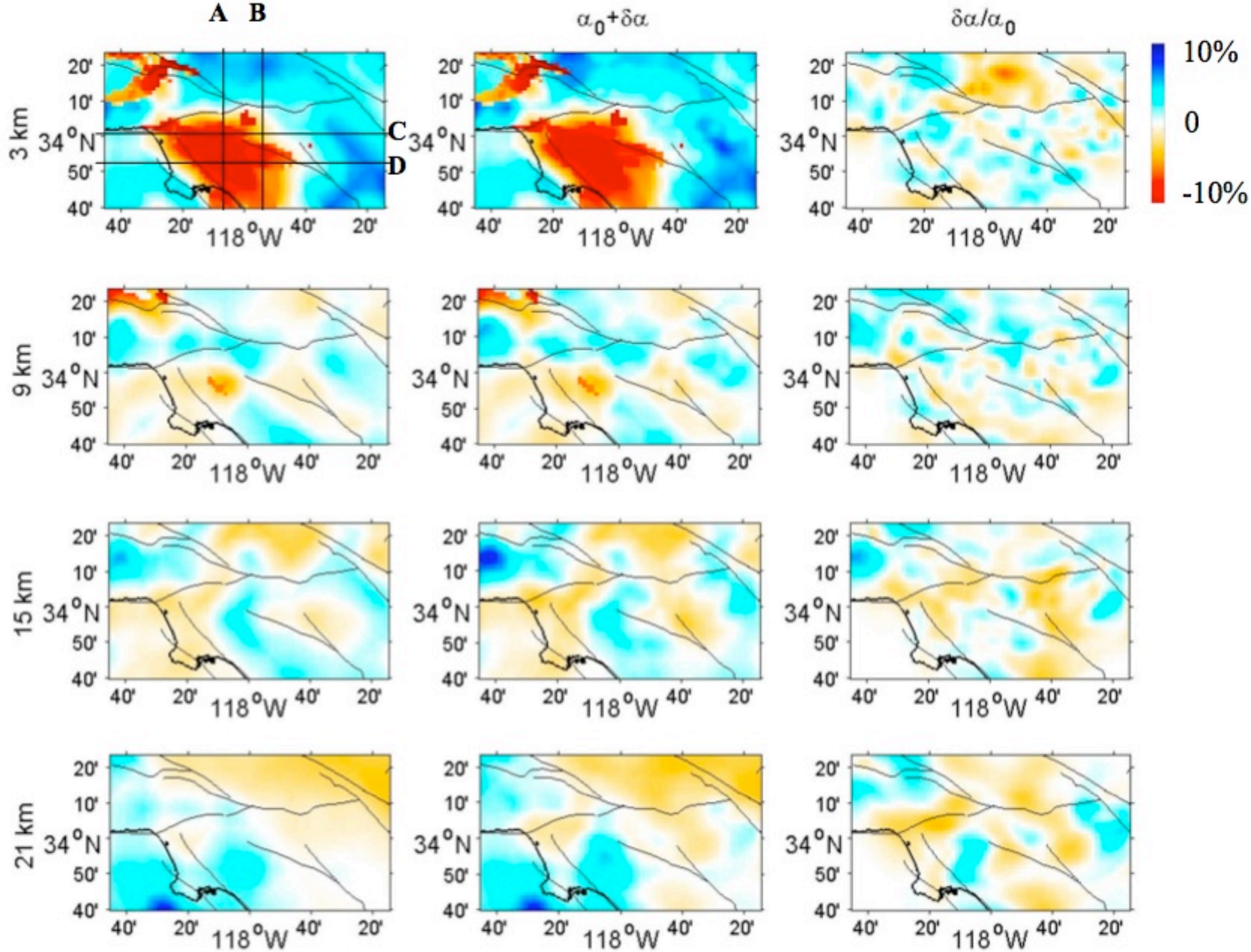


A new representation of the two nodal planes with different probabilities

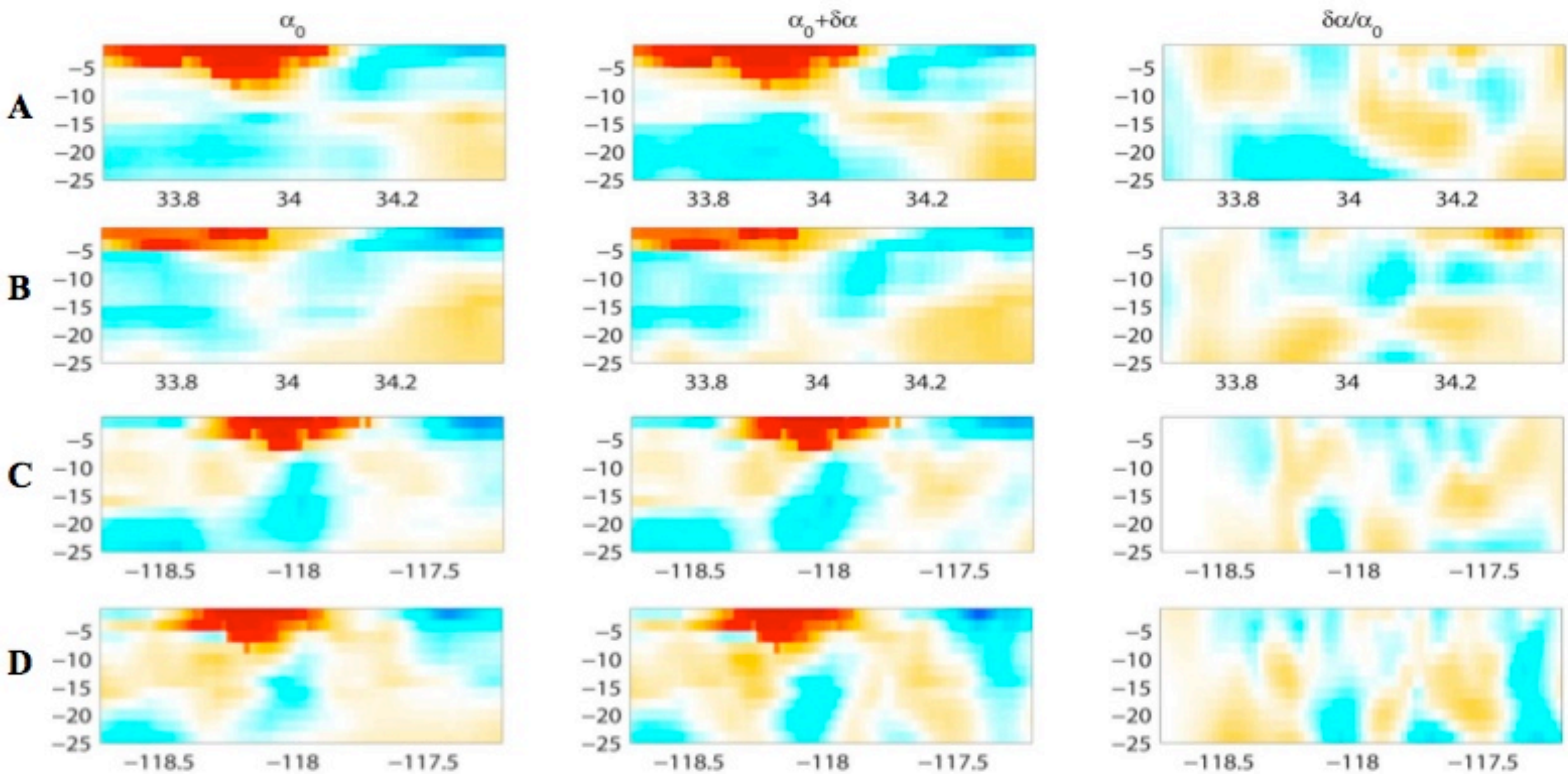


Full-3D Fréchet Kernel Examples

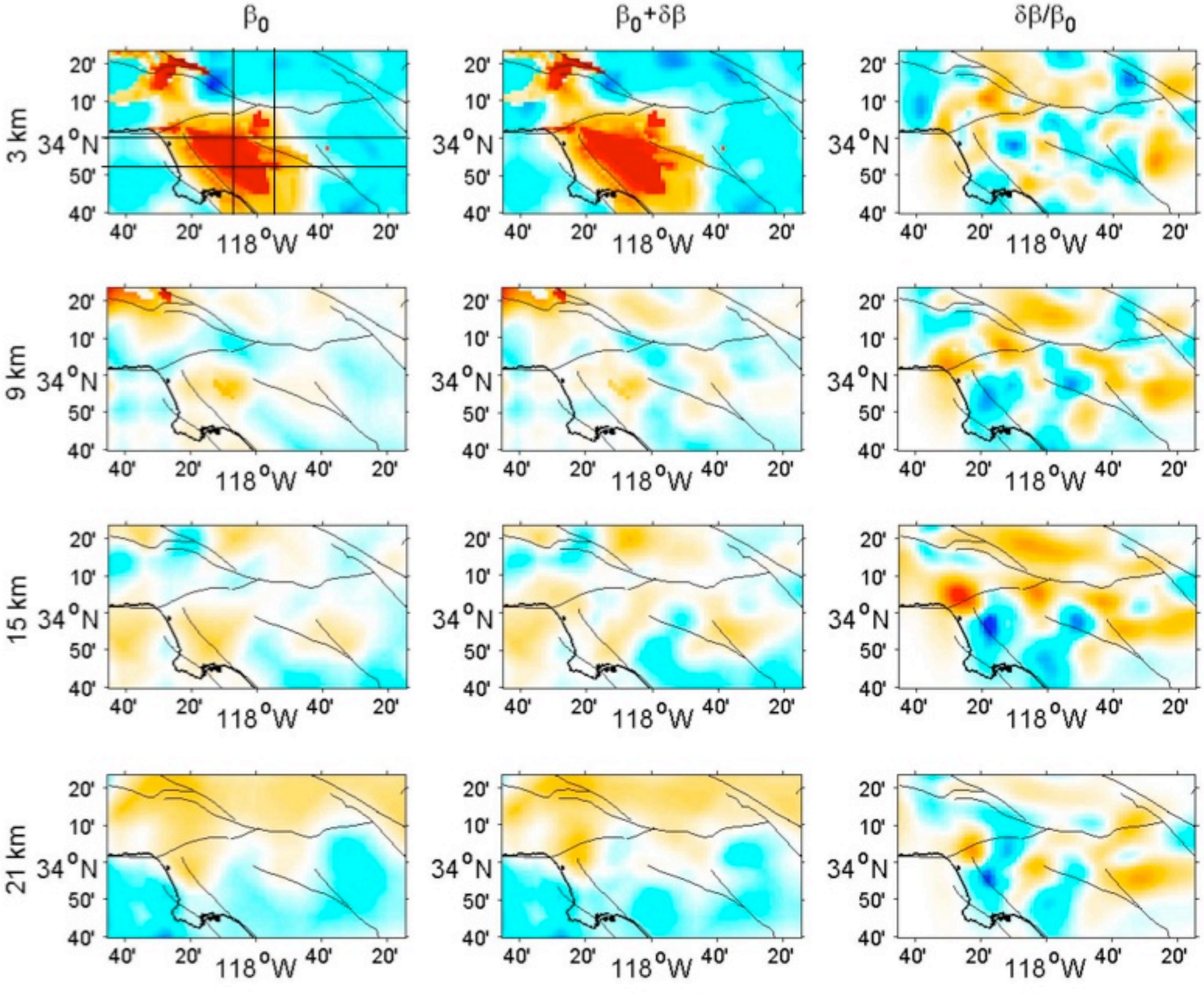


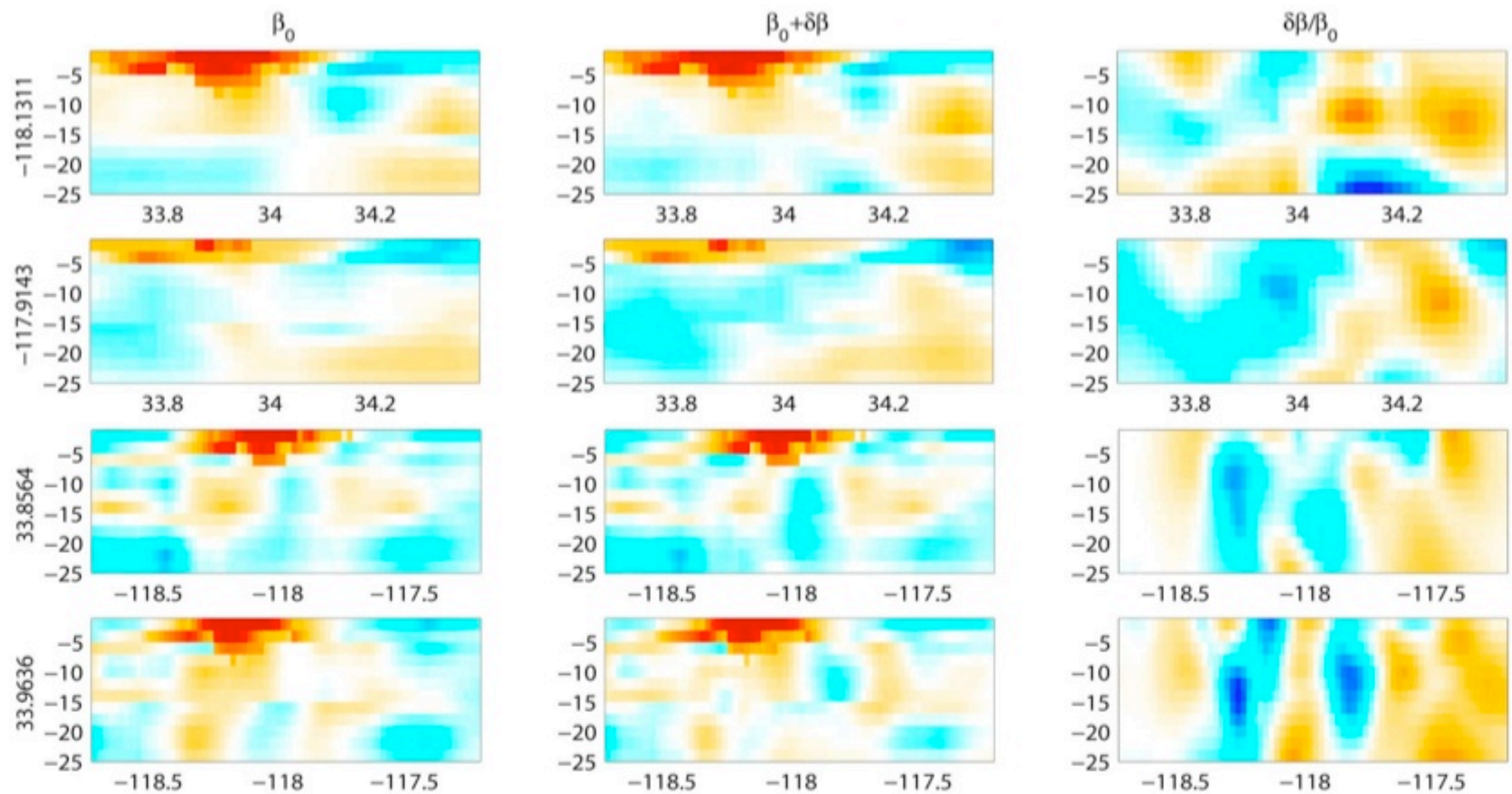


A
B
C
D
α0+δα
δα/α0
10%
0
-10%
3 km
9 km
15 km
21 km
20'
10'
34°N
50'
40'
40'
20'
118°W
40'
20'

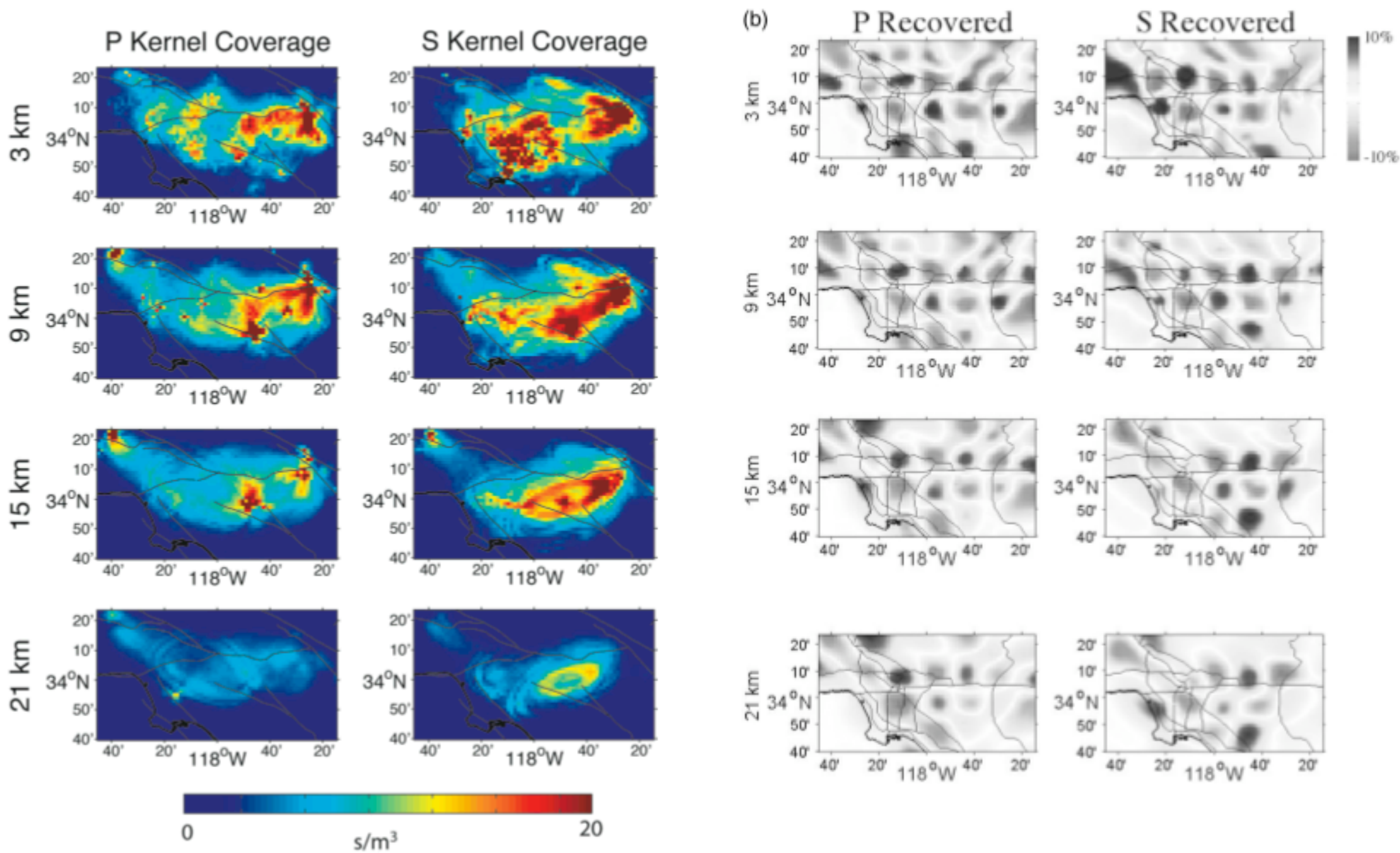


α0
α0+δα
δα/α0
A
B
C
D
−5
−10
−15
−20
−25
33.8
34
34.2
−118.5
−118
−117.5

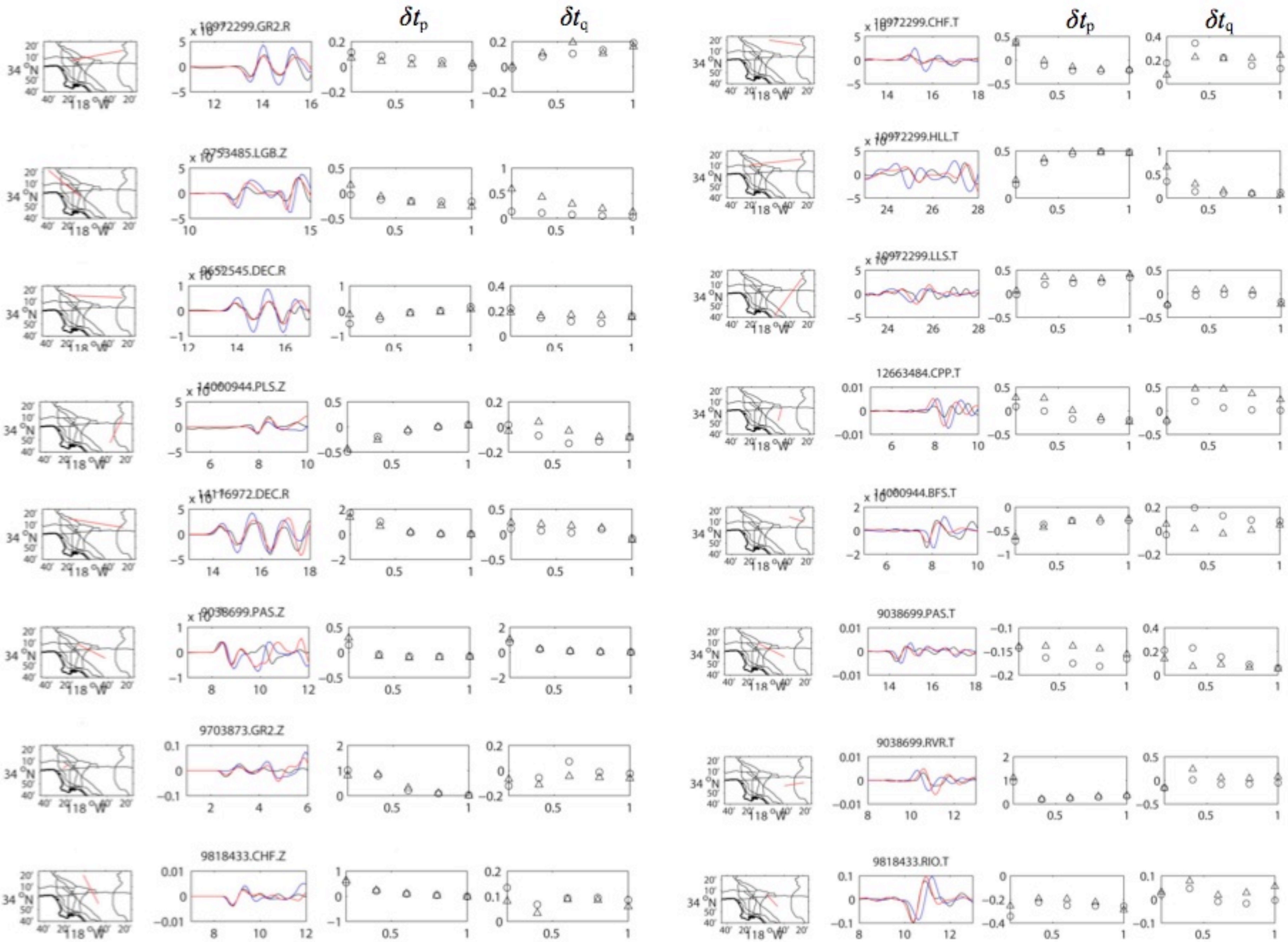

β0
β0+δβ
δβ/β0
3 km
9 km
15 km
21 km
20'
10'
34°N
50'
40'
40'
20'
118°W
40'
20'



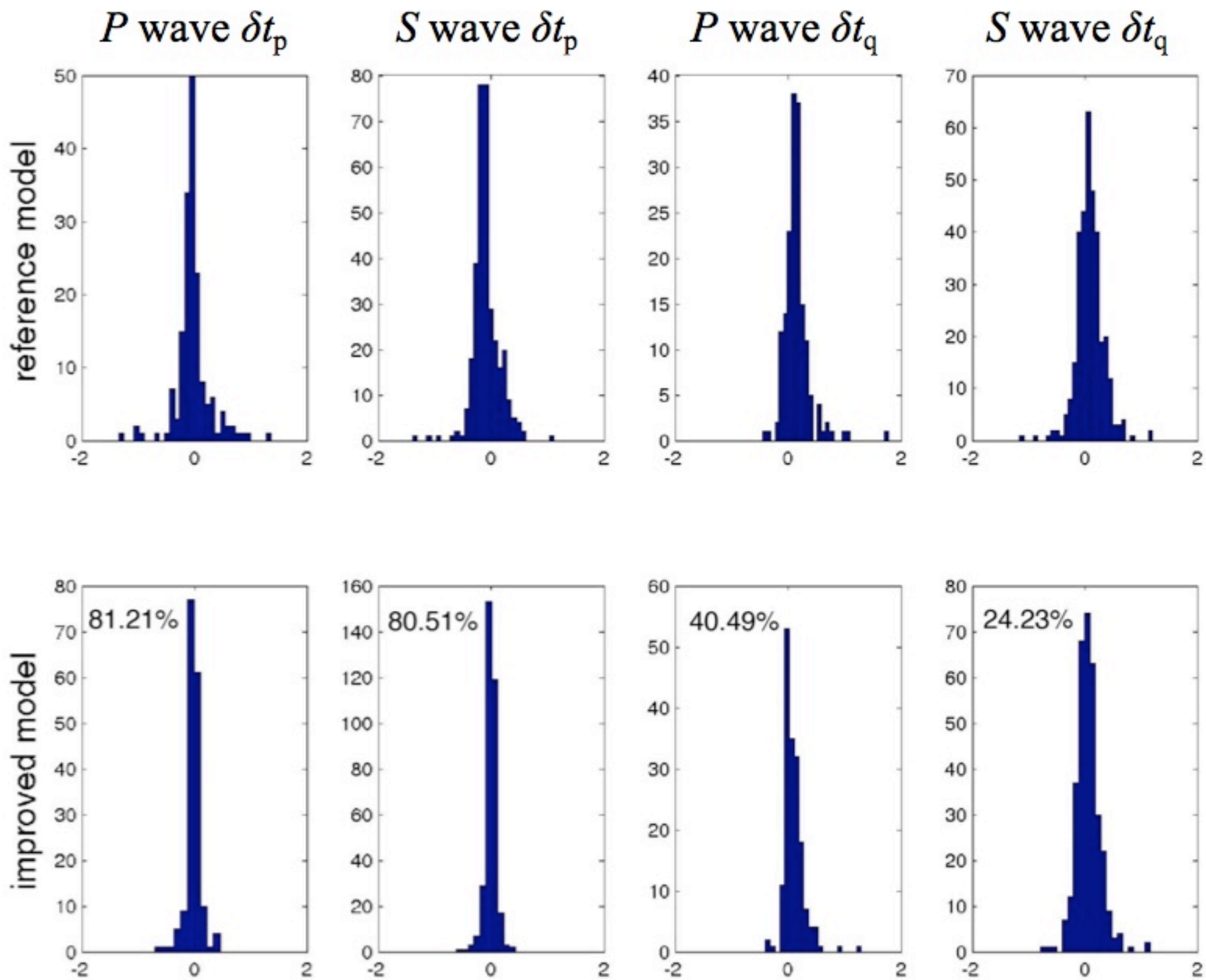
β0
β0+δβ
δβ/β0
−118.1311
−117.9143
33.8564
33.9636
−5
−10
−15
−20
−25
33.8
34
34.2
−118.5
−118
−117.5



P Kernel Coverage
S Kernel Coverage
3 km
9 km
15 km
21 km
20'
10'
34°N
50'
40'
40'
20'
118°W
0
20
s/m³
(b)
P Recovered
S Recovered
10%
-10%

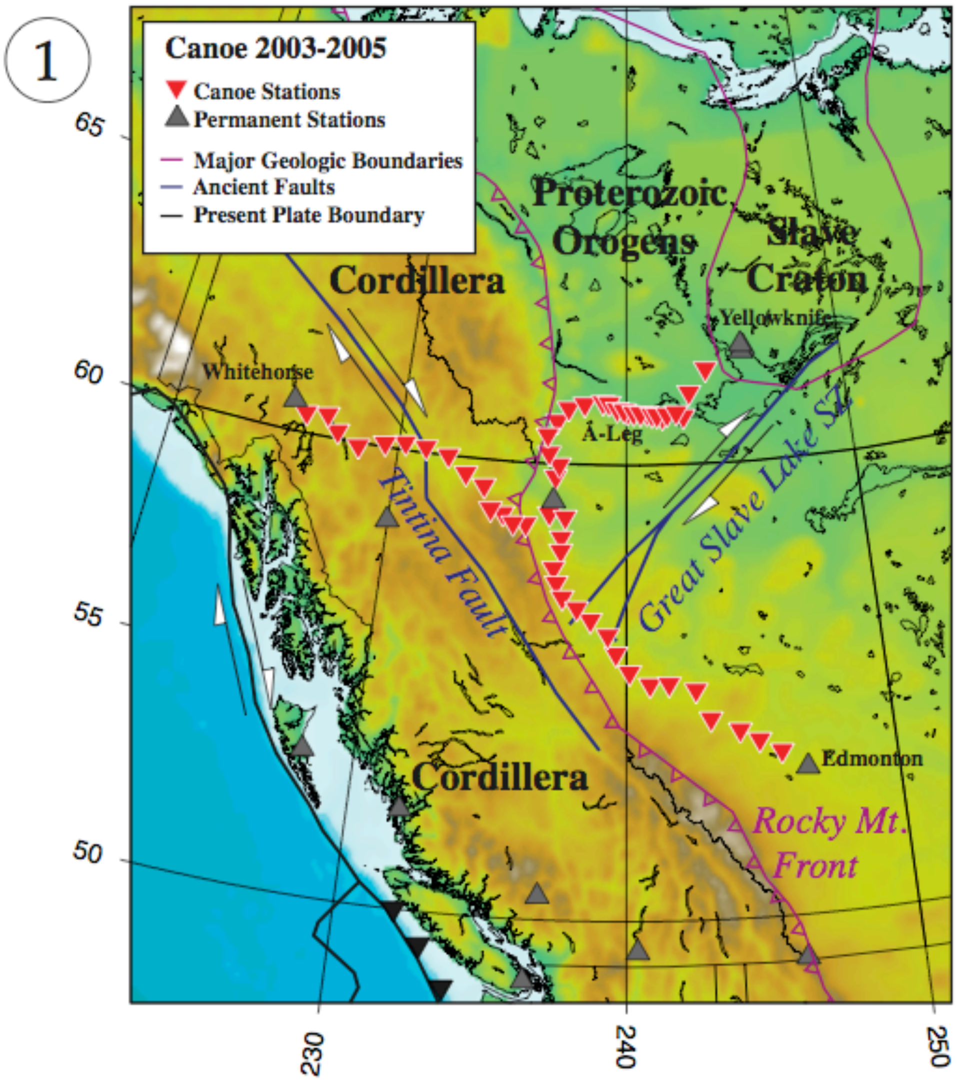


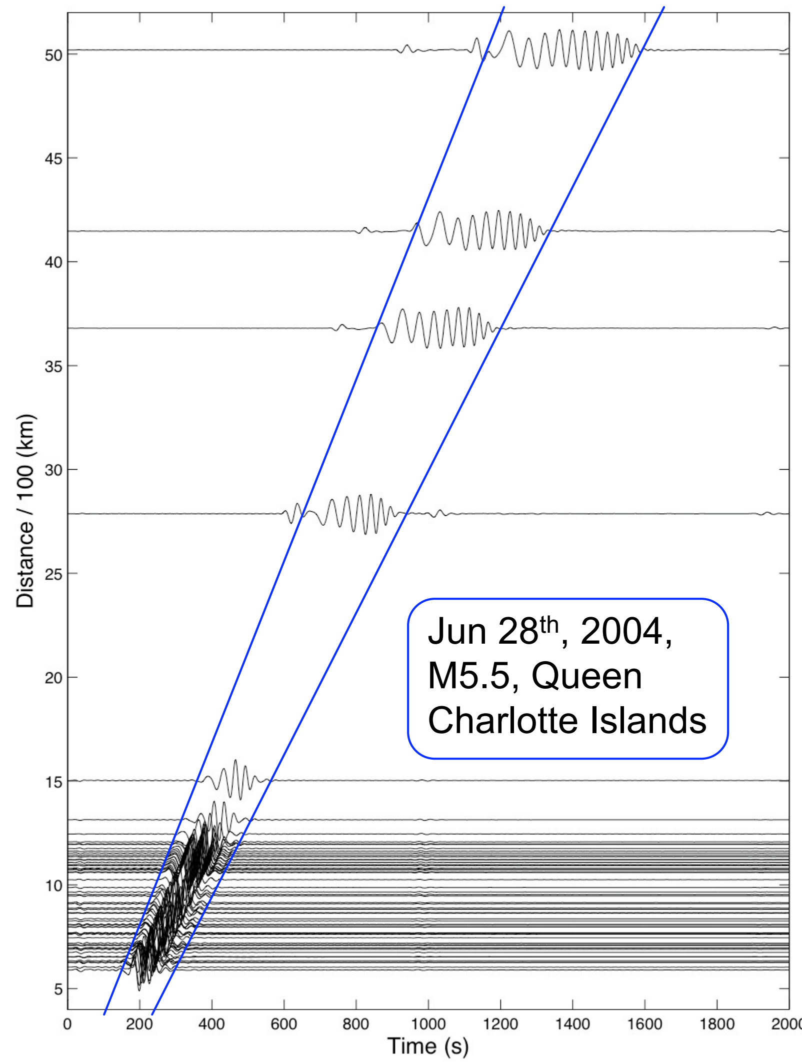
δt_p
δt_q
10972299.GR2.R
9753485.LGB.Z
9652545.DEC.R
14000944.PLS.Z
14116972.DEC.R
9038699.PAS.Z
9703873.GR2.Z
9818433.CHF.Z
10972299.CHF.T
10972299.HLL.T
10972299.LLS.T
12663484.CPP.T
14000944.BFS.T
9038699.PAS.T
9038699.RVR.T
9818433.RIO.T
34 °N
118 °W



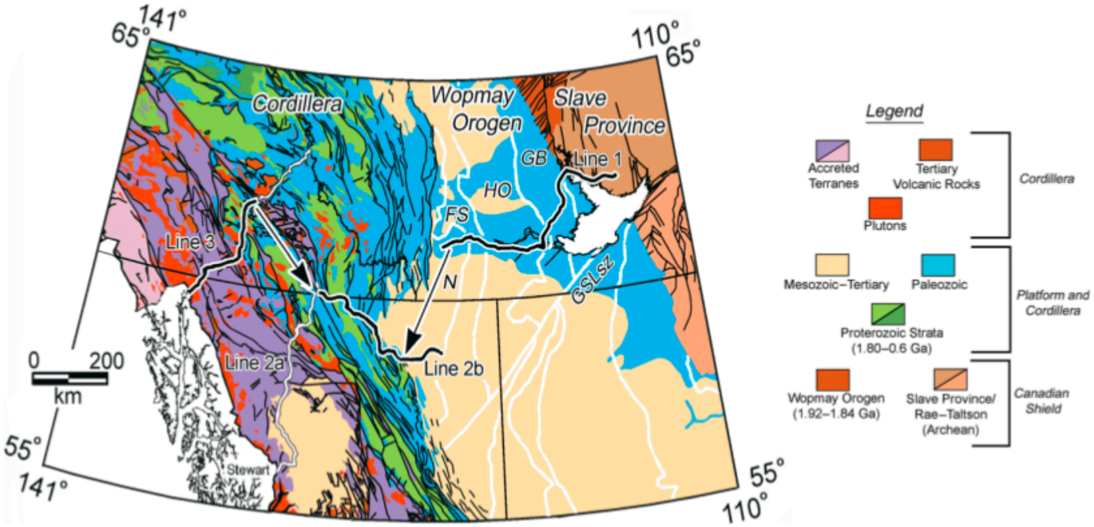
P wave δt_p
S wave δt_p
P wave δt_q
S wave δt_q
reference model
improved model
81.21%
80.51%
40.49%
24.23%

The Canadian Cordillera



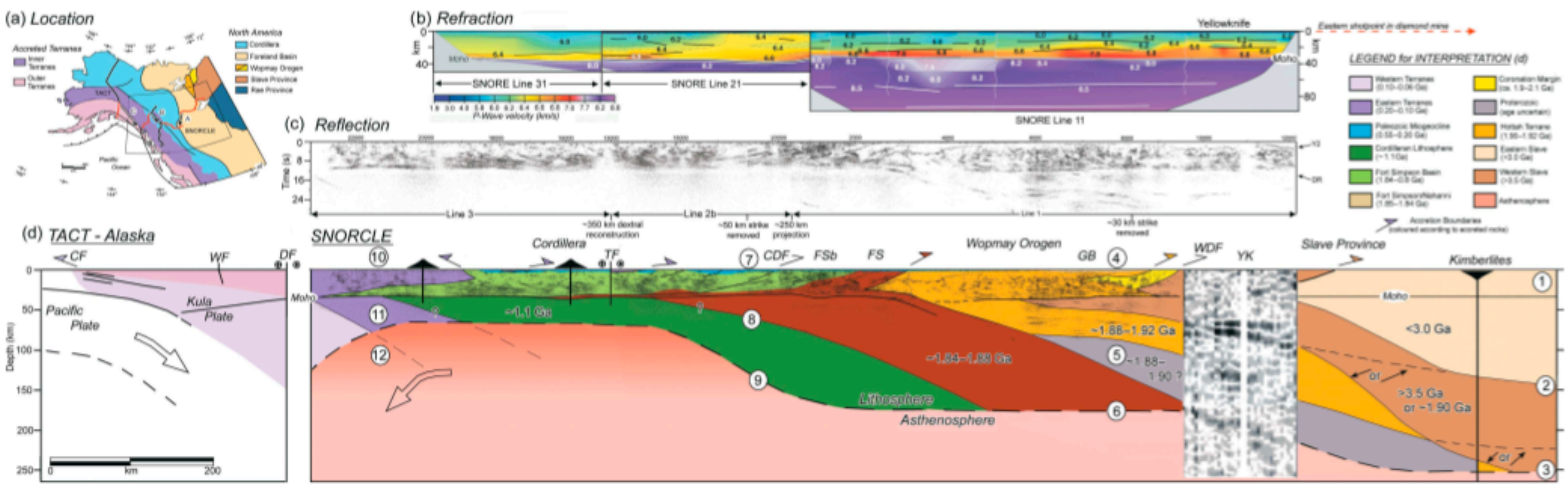


Lithoprobe

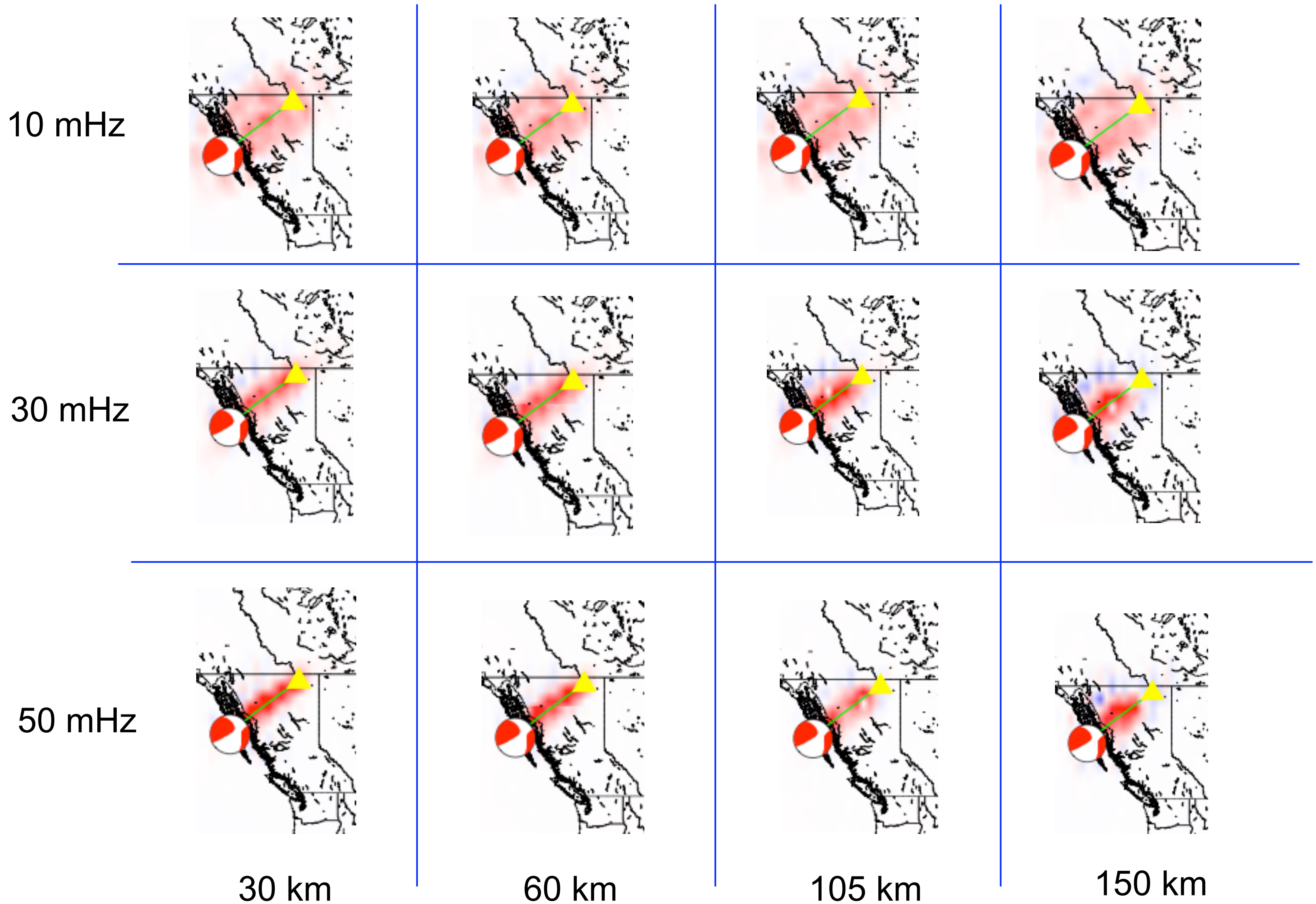


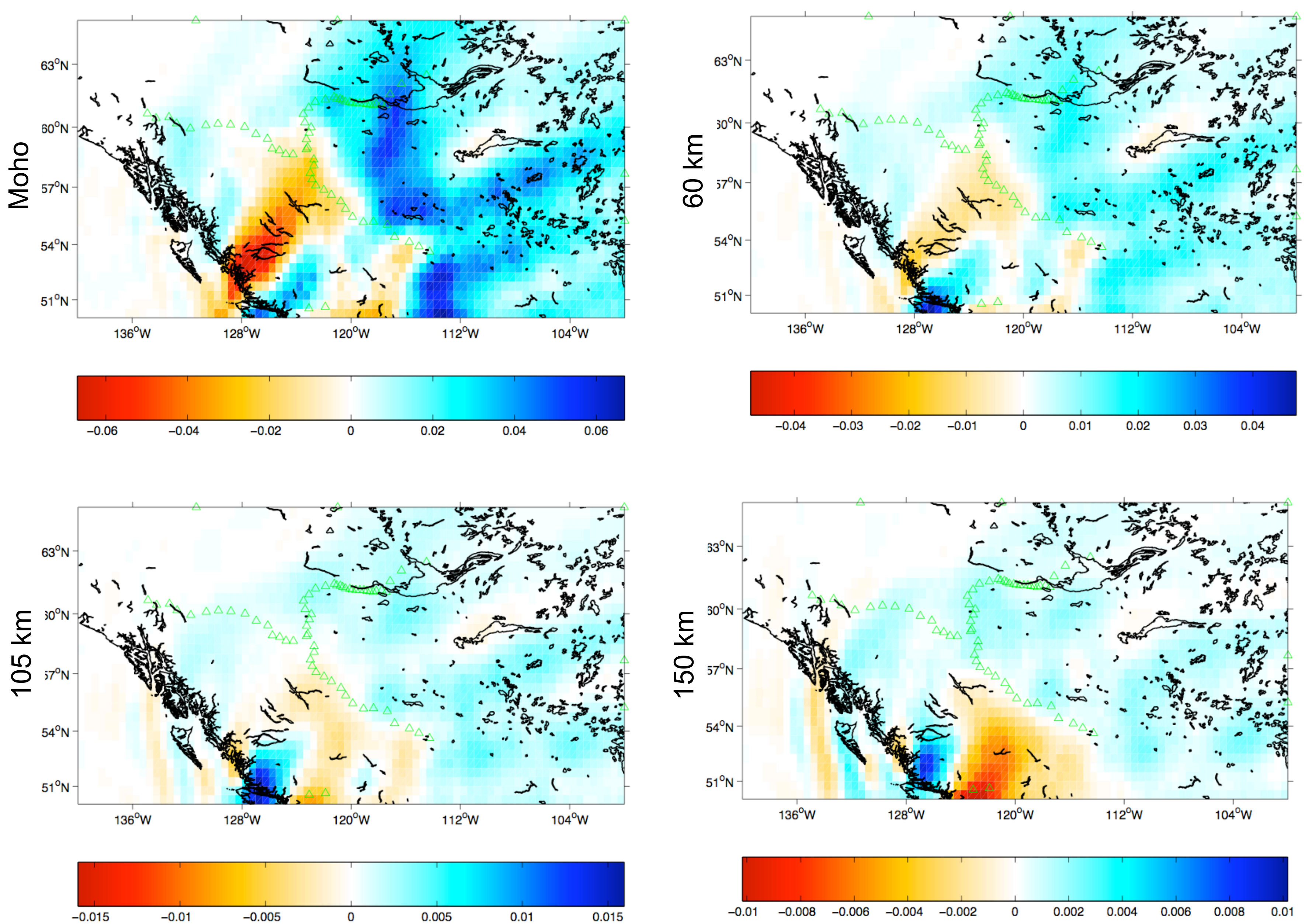
Refraction/reflection (Cook *et al.* 2005)

Basement below the cordillera is old.



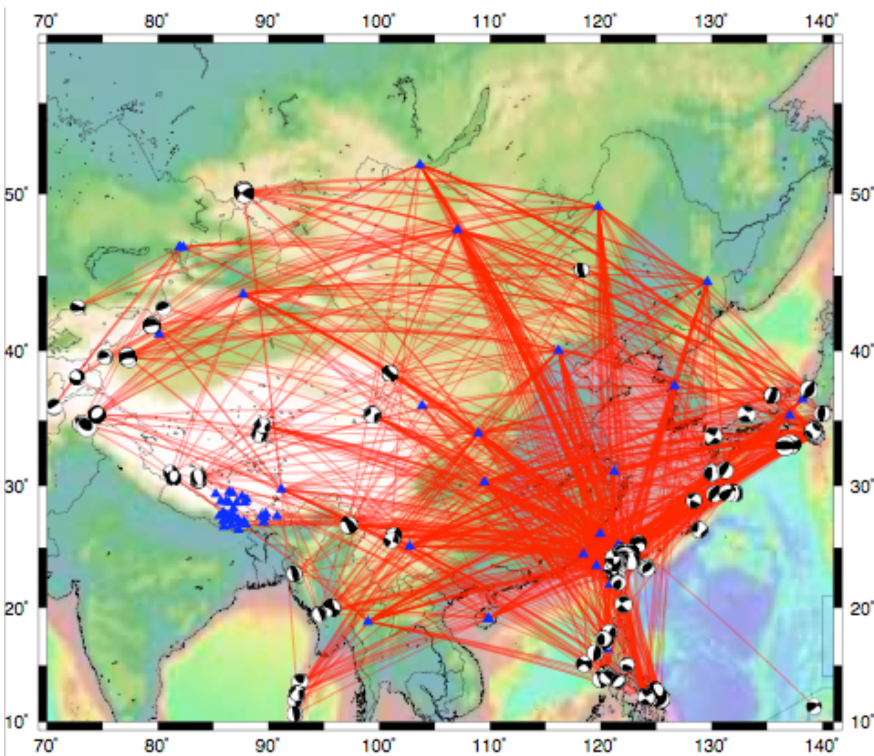
Frequency-dependent Full-wave Kernels

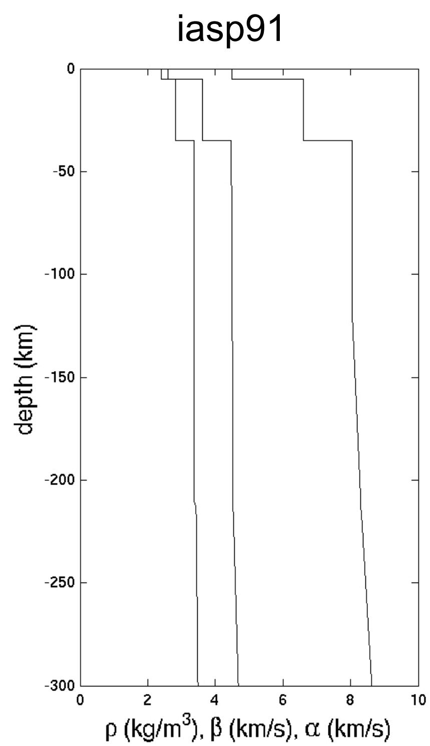


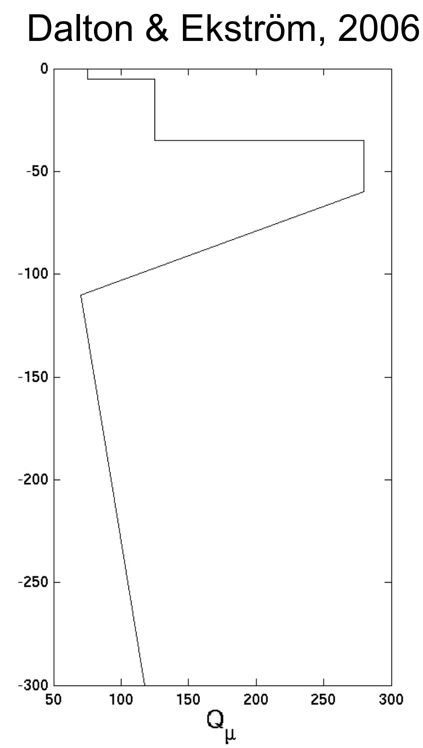


Moho
60 km
105 km
150 km
63°N
60°N
57°N
54°N
51°N
136°W
128°W
120°W
112°W
104°W
−0.06 −0.04 −0.02 0 0.02 0.04 0.06
−0.04 −0.03 −0.02 −0.01 0 0.01 0.02 0.03 0.04
−0.015 −0.01 −0.005 0 0.005 0.01 0.015
−0.01 −0.008 −0.006 −0.004 −0.002 0 0.002 0.004 0.006 0.008 0.01

Eastern Eurasia

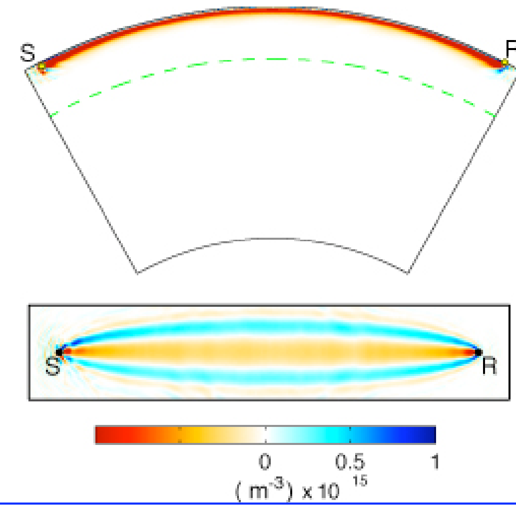






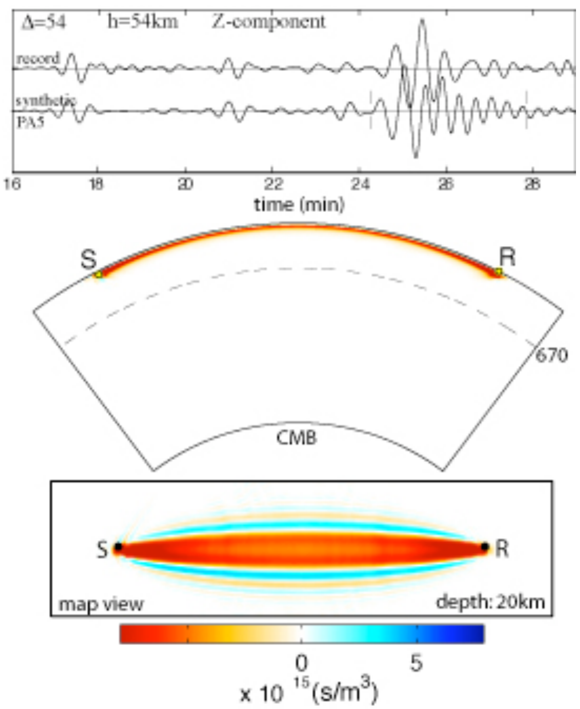
β and Q_μ Joint Inversion Using Surface Wave Phase and Amplitude Anomalies

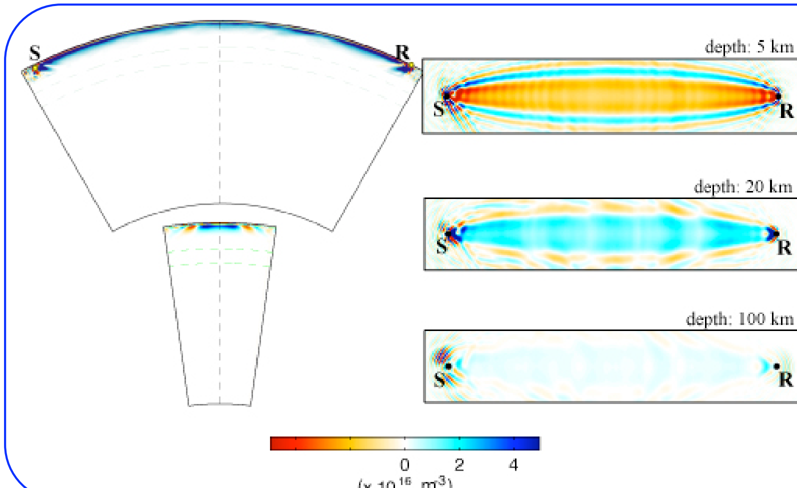
$$C_{jklm} = \left\{ \kappa(\omega)[1 + iQ_\kappa^{-1}(\omega)] - \frac{2}{3}\mu(\omega)[1 + iQ_\mu^{-1}(\omega)] \right\} \delta_{jk}\delta_{lm} + \mu(\omega)[1 + iQ_\mu^{-1}(\omega)](\delta_{jl}\delta_{km} + \delta_{jm}\delta_{kl}) + \gamma_{jklm},$$



K_β^q

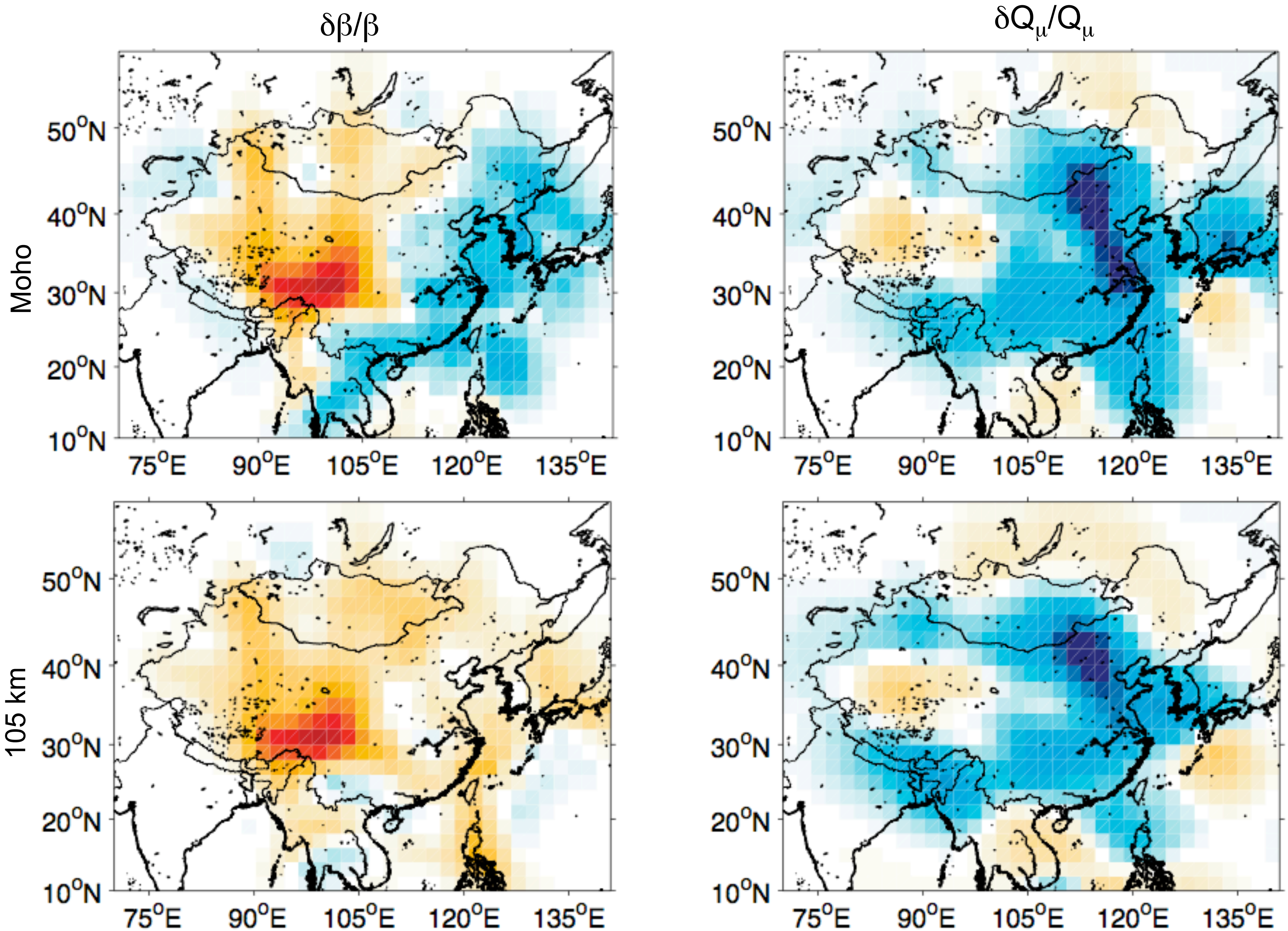
K_β^p



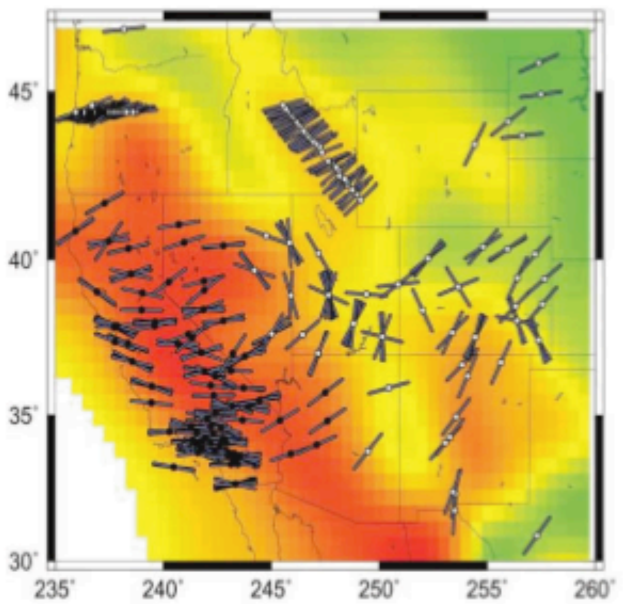


$K_{Q_\mu}^q$

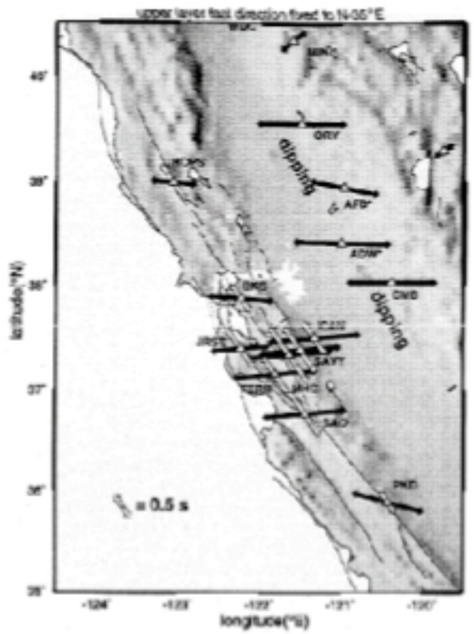

SCEC


δβ/β
δQμ/Qμ
Moho
105 km
50°N
40°N
30°N
20°N
10°N
75°E
90°E
105°E
120°E
135°E

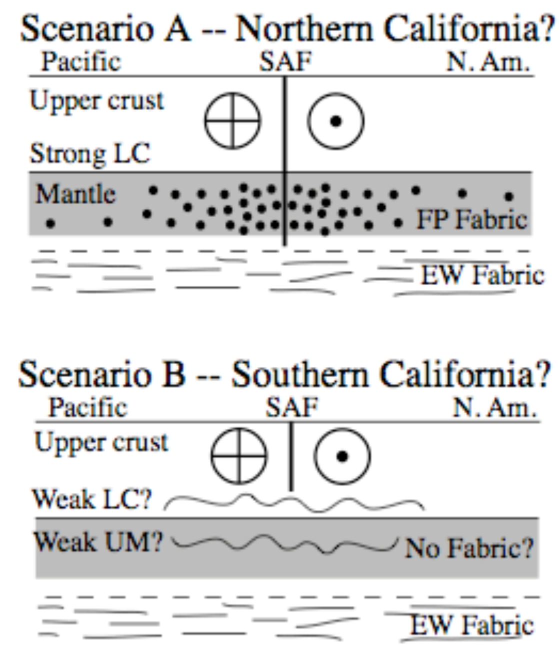
Western US



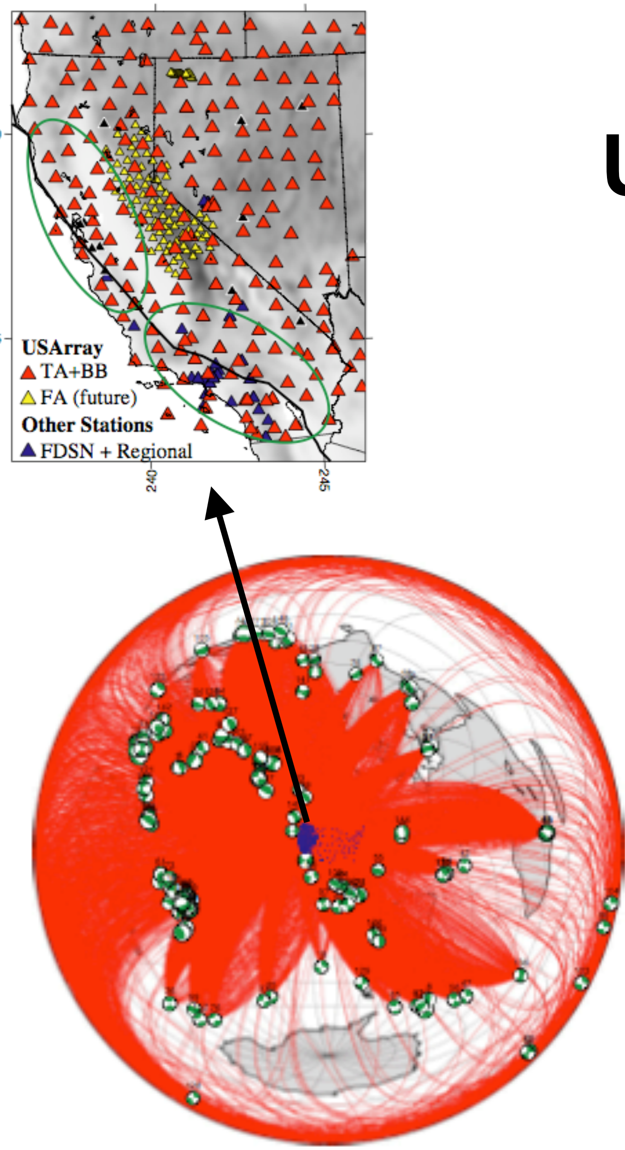
Silver & Holt, 2002



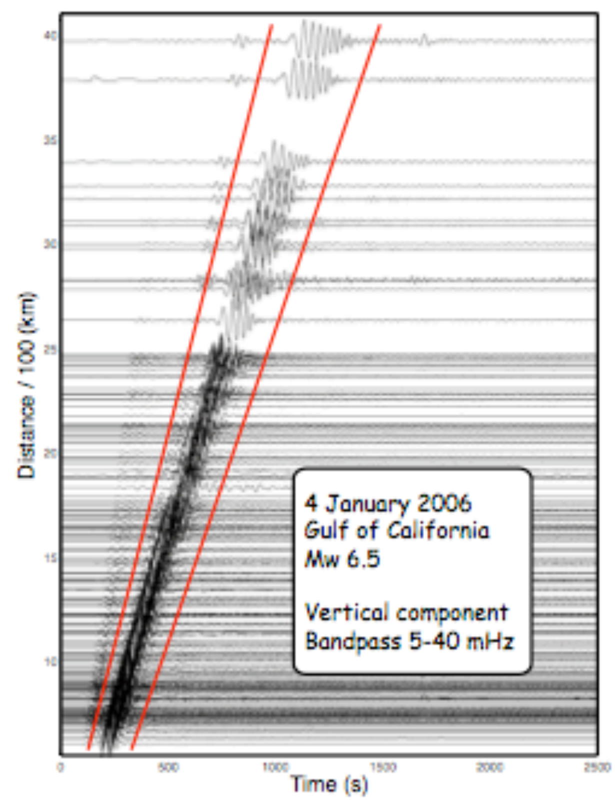
Hartog & Schwartz, 2001



USArray

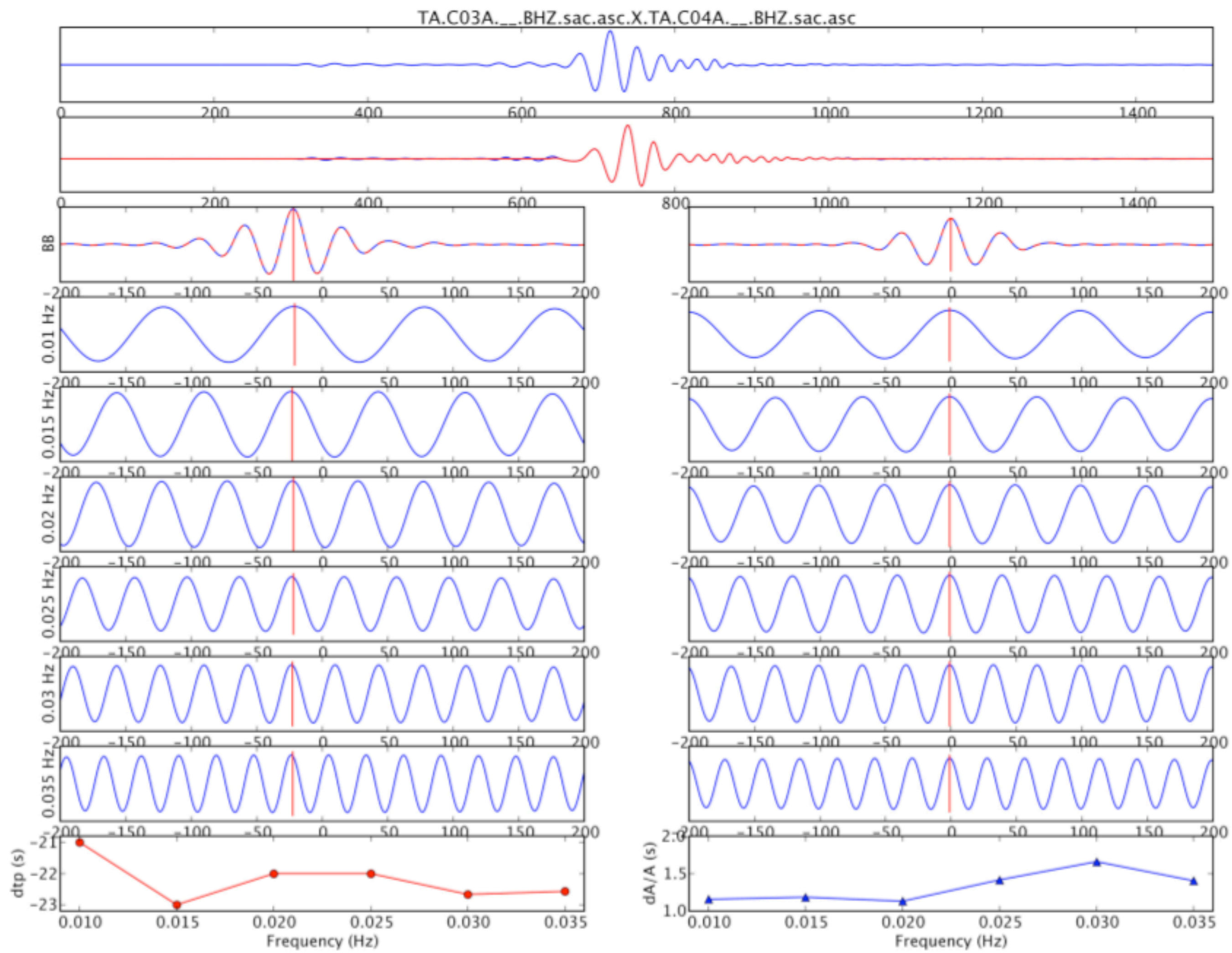


40
35
240
245
USArray
TA+BB
FA (future)
Other Stations
FDSN + Regional

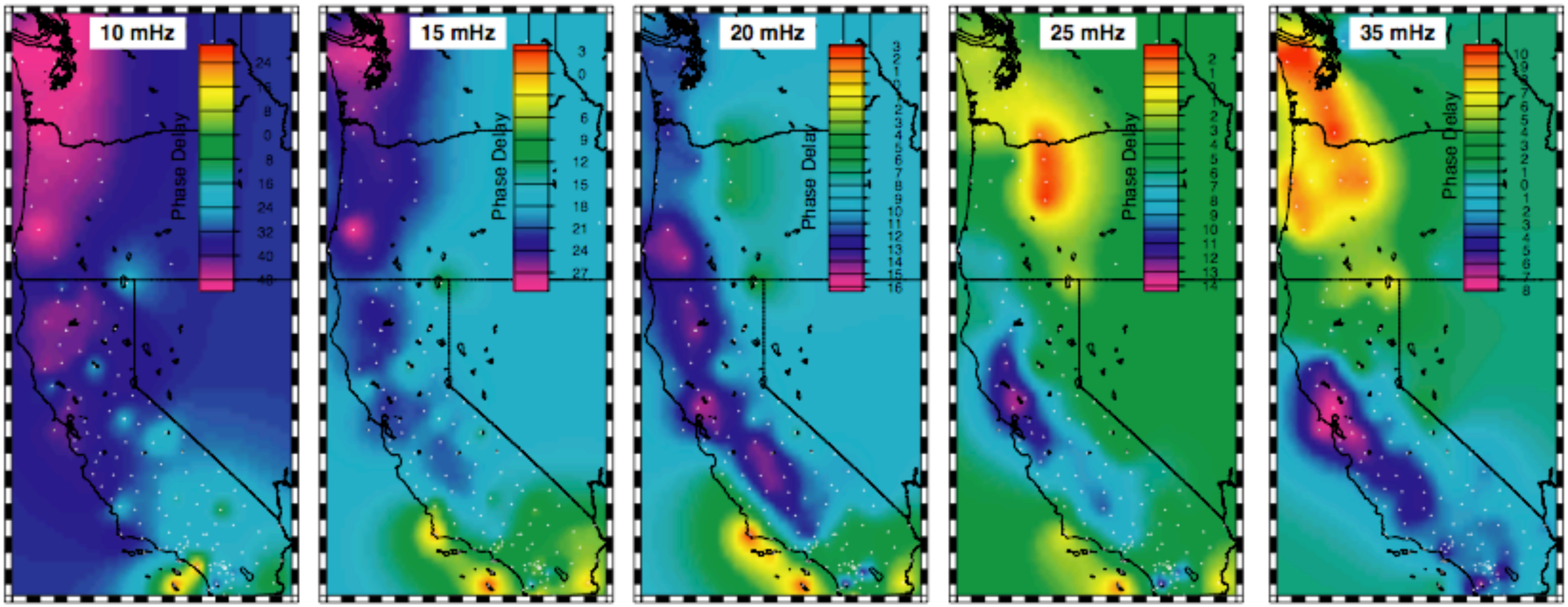


Distance / 100 (km)
Time (s)
0
500
1000
1500
2000
2500
4 January 2006
Gulf of California
Mw 6.5
Vertical component
Bandpass 5-40 mHz

Double-Difference GSDF



Frequency-dependent Phase-delay Maps



Discussion & Summary

Wave-equation Migration Velocity Analysis

$$\chi^2(\mathbf{m}) = \left\|\Delta I(\mathbf{m})\right\|_I + \lambda\left\|\mathbf{m}\right\|_m$$

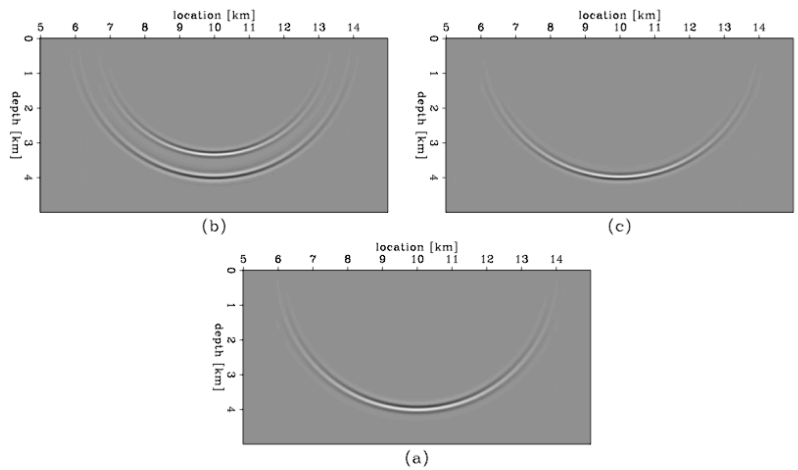
$$I(\mathbf{m}) = \sum_s u^s \otimes [u^s]^+$$

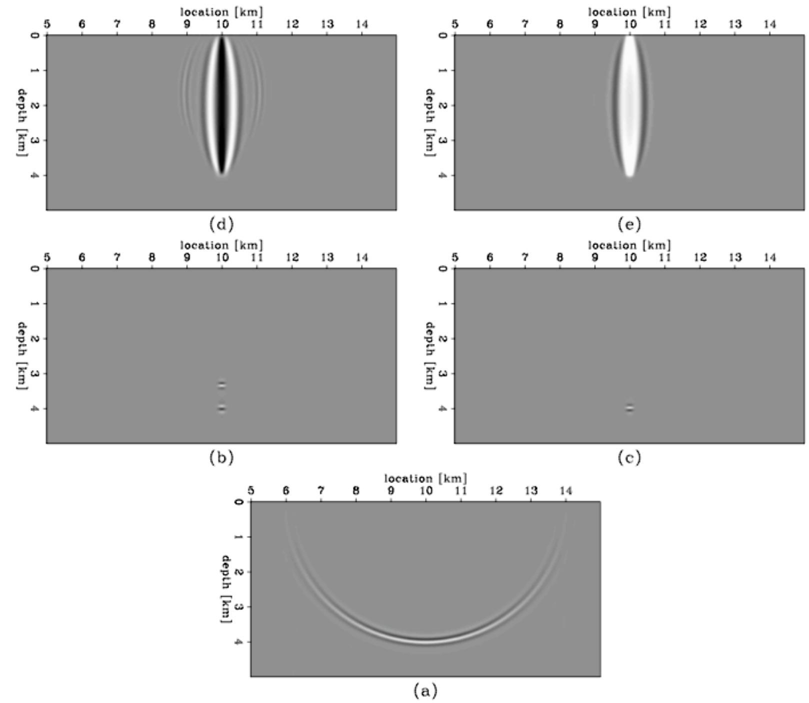
$$\delta I(\mathbf{m}) = \sum_s \delta u^s \otimes [u^s]^+ + \sum_s u^s \otimes \left\{\delta \mathbf{G}^{\mathbf{T}} * [f^s]^+\right\} + \sum_s u^s \otimes \left\{\mathbf{G}^{\mathbf{T}} * \delta [f^s]^+\right\}$$

$$\delta I(\mathbf{m}) = \mathbf{K}_{\mathbf{m}}^I \cdot \delta \mathbf{m}$$

$$\Delta I : \begin{cases} I - I_0 \\ \Psi - \Psi_0 \; (I = e^{\Psi}, I_0 = e^{\Psi_0}) \qquad ? \end{cases}$$

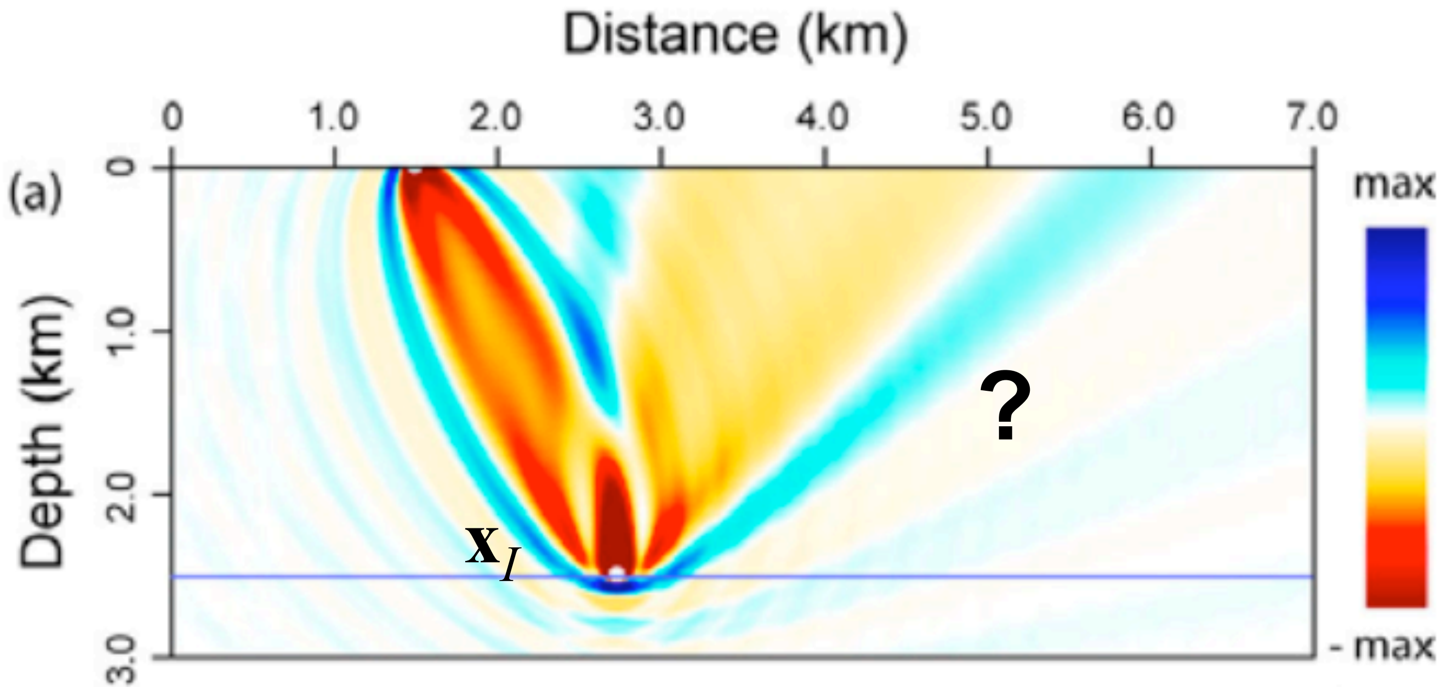
Sava & Biondi 2004





$I = K(\rho)[I_0]$	Stolt residual migration
$\Delta I = I - I_0$	Difference image perturbation
$\Delta I \approx K'\big\vert_{\rho=1}[I_0]\Delta\rho$	Linearized image perturbation
$I_0 \Delta\Psi \approx K'\big\vert_{\rho=1}[I_0]\Delta\rho$	Image phase perturbation

Broadband Phase-shift Kernel



Xie & Yang (2007),
Modeling and Imaging
Laboratory (MILAB),
UC Santa Cruz

$$\delta I^s(\mathbf{x}_I) = \delta u^s \otimes [u^s]^+ + u^s \otimes \left\{\delta \mathbf{G}^{\mathbf{T}} * [f^s]^+\right\}$$
$$+ u^s \otimes \left\{\mathbf{G}^{\mathbf{T}} * \delta[f^s]^+\right\}$$

Summary

- Combine time-bootstrapping with frequency-bootstrapping.
- Use Rytov linearization for velocity tomography, Born linearization for impedance mapping.
- Construct the Hessian when source number is not much less than receiver number.
- Full-3D two-way wave equation accounts for complete physics.