High-resolution seismic array imaging based on spectral element simulations

Piero Basini, Ping Tong, Qinya Liu

Department of Physics
University of Toronto

July 14-17, 2013

Collaborators:

Yingjie Yang (Macquarie U), Carl Tape (U of Alaska)

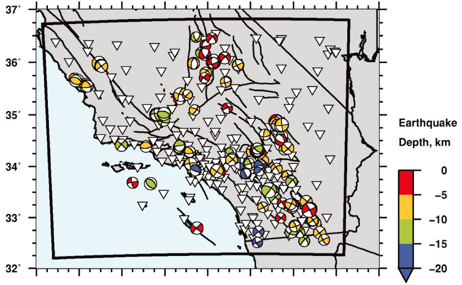
Chinwu Chen (National Taiwan U), Dimitri Komatitsch (Aix-Marseille U)

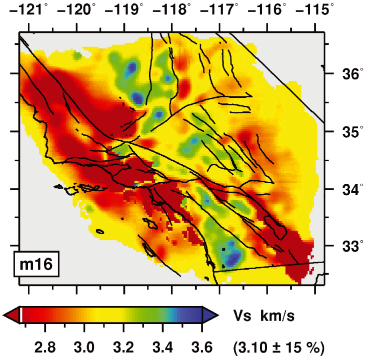
Adjoint tomography of southern California crust

Previous adjoint tomography of SoCal crust (Tape et al 2009, 2010):

- 143 crustal earthquakes + 243 stations
- traveltime measurements over 2-30, 3-30, and 6-30 sec bands.

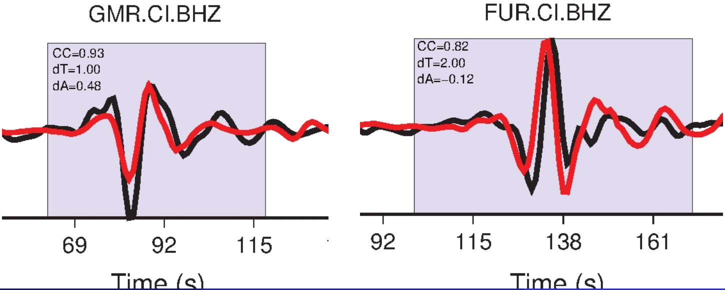
t



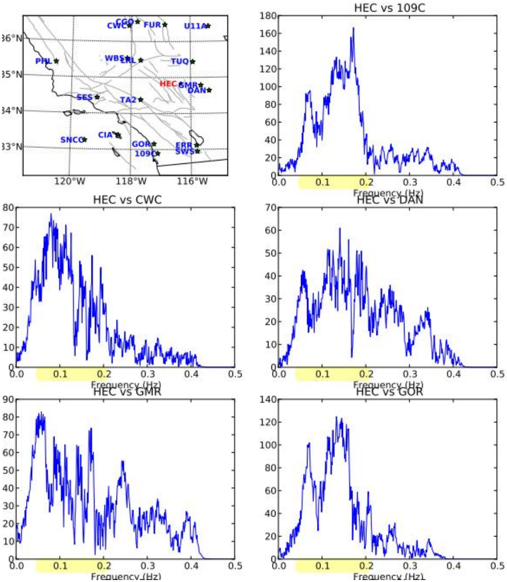


NCFs for socal stations

- NCF: cross-correlation of ambient noise recording between two seismic stations stacked over long period of time
- Data processing: remove IR, filter between 3-50 sec, spectral whitening, stack cross-correlation of daily data over three years (Bensen et al. 2007)
- $\sim$ 13,000 V-V NCFs between 147 stations; only phase is used
- We assume $\frac{\partial}{\partial t}$ (NCF) $\sim$ Greens functions between master and receiver station



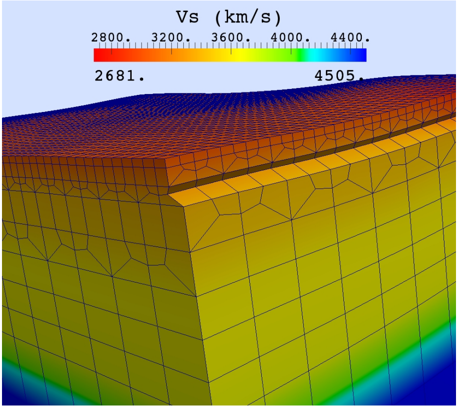
NCFs for socal stations: spectra



- primary ($\sim$ 14 s) and secondary ($\sim$ 7 s) microseismic peaks; dominant energy between 5-20 sec.
- measurements by FLEXWIN (Maggi et al, 2008)
  5 – 50 ($\sim$ 8600)
  10 – 50 ($\sim$ 16700)
  20 – 50 ($\sim$ 2000)
- start iteration for *Vs* at 10 – 50 sec, and go down to 5 – 50 sec subsequently

Computation

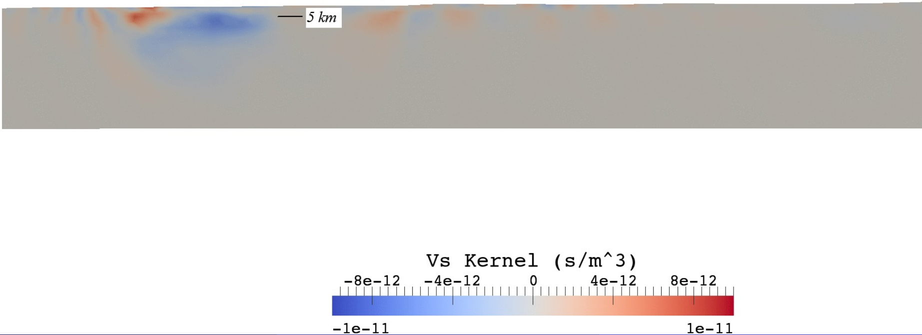
- Forward and adjoint simulations (i.e., synthetics and kernels) are computed by the SPECFEM3D package (an early version)



- $\sim$ 30 min for one forward and $\sim$ 1 hour for one adjoint simulation

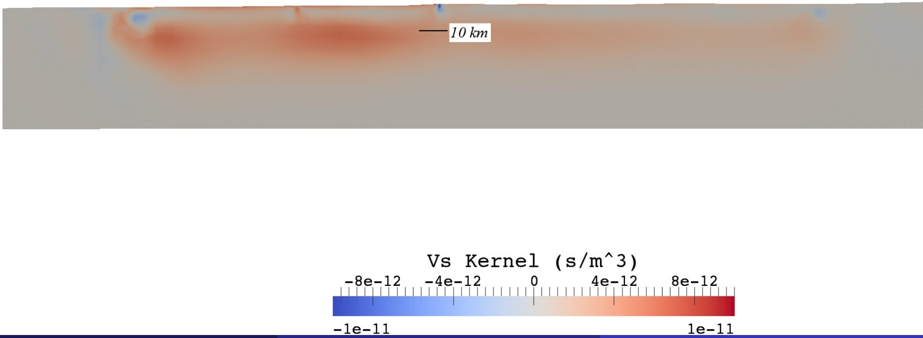
NCF depth sensitivities

Shear-wave speed (Vs) depth sensitivity of a single traveltime measurement peaks at $\sim 1/3$ of wavelength.
Cross-sections of event kernels between stations: 5 sec $\rightarrow$ 5 km $\Rightarrow$ help resolve mid-to-lower crustal structure



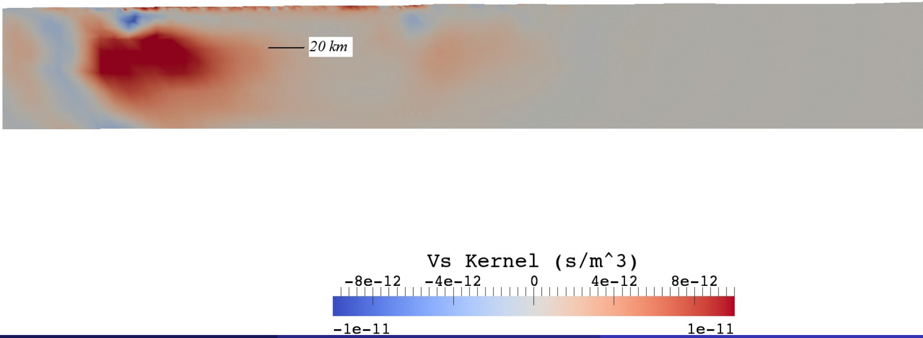
NCF depth sensitivities

Shear-wave speed (Vs) depth sensitivity of a single traveltime measurement peaks at $\sim 1/3$ of wavelength.
Cross-sections of event kernels between stations: 10 sec $\rightarrow$ 10 km $\Rightarrow$ help resolve mid-to-lower crustal structure



NCF depth sensitivities

Shear-wave speed (Vs) depth sensitivity of a single traveltime measurement peaks at $\sim 1/3$ of wavelength.
Cross-sections of event kernels between stations: 20 sec $\rightarrow$ 20 km $\Rightarrow$ help resolve mid-to-lower crustal structure



Model update by subspace method at each iteration

$$\text{Total misfit} \quad \Phi = \frac{1}{2}\frac{1}{N_{meas}}\sum_{e,s,p}\left(\frac{T_s - T_d}{\sigma_T}\right)^2$$

Event kernels (for one master station):

$$\delta\phi_e = \int_V K^e_{m;s,p}(\mathbf{x})\frac{\delta m}{m}\, d^3\mathbf{x}$$

Assume model update is a linear combination of event kernels at every iter (Tape et al 2009):

$$\frac{\delta m}{m}(\mathbf{x}) = \sum_e K^e(\mathbf{x})\mathbf{C},$$

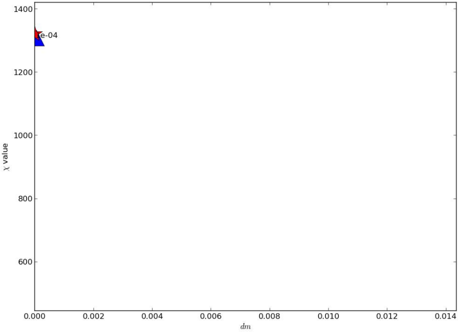
where

$$C = (N + \lambda I)^{-1}\Phi, \quad \text{kernel Hessian} \;\; N_{ij} = \int K^{e_i} K^{e_j} d^3\mathbf{x}$$

Selection of λ

Choose a set of λ values, for $\lambda = 10^{-4}$

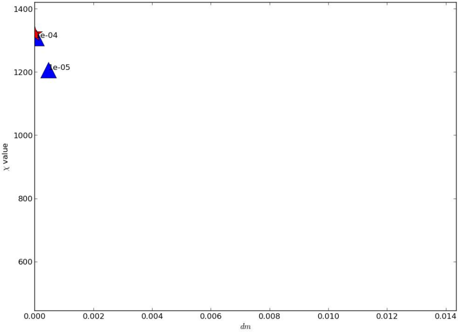
- compute model update $\frac{\delta m}{m}$ based on the coefficients $\mathbf{C} = (N + \lambda I)^{-1}\Phi$
- compute sum of event misfits ϕ for a subset of master stations (~ 20), and plot ϕ against δm (L-curve):



Selection of λ

Choose a set of λ values, for $\lambda = 10^{-5}$

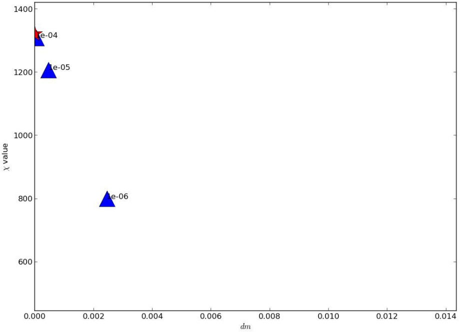
- compute model update $\frac{\delta m}{m}$ based on the coefficients $\mathbf{C} = (N + \lambda I)^{-1}\Phi$
- compute sum of event misfits ϕ for a subset of master stations ($\sim$ 20), and plot ϕ against δm (L-curve):



Selection of λ

Choose a set of λ values, for $\lambda = 10^{-6}$

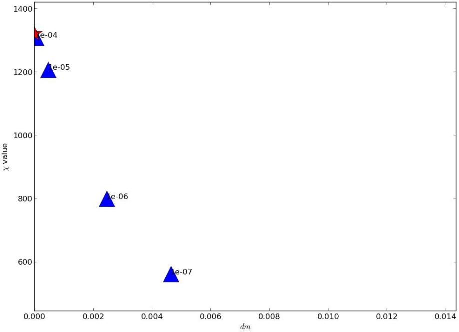
- compute model update $\frac{\delta m}{m}$ based on the coefficients $\mathbf{C} = (N + \lambda I)^{-1}\Phi$
- compute sum of event misfits ϕ for a subset of master stations (~ 20), and plot ϕ against δm (L-curve):



Selection of λ

Choose a set of λ values, for $\lambda = 10^{-7}$

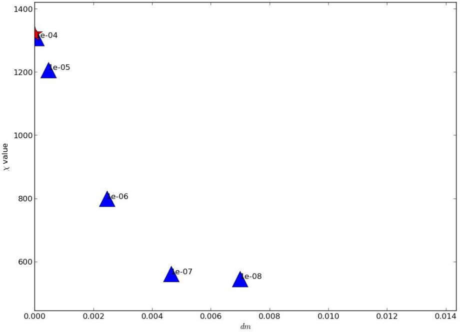
- compute model update $\frac{\delta m}{m}$ based on the coefficients $\mathbf{C} = (N + \lambda I)^{-1}\Phi$
- compute sum of event misfits ϕ for a subset of master stations (~ 20), and plot ϕ against δm (L-curve):



Selection of λ

Choose a set of λ values, for $\lambda = 10^{-8}$

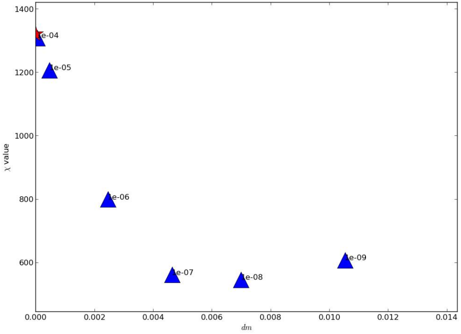
- compute model update $\frac{\delta m}{m}$ based on the coefficients $\mathbf{C} = (N + \lambda I)^{-1}\Phi$
- compute sum of event misfits ϕ for a subset of master stations (~ 20), and plot ϕ against δm (L-curve):



Selection of λ

Choose a set of λ values, for $\lambda = 10^{-9}$

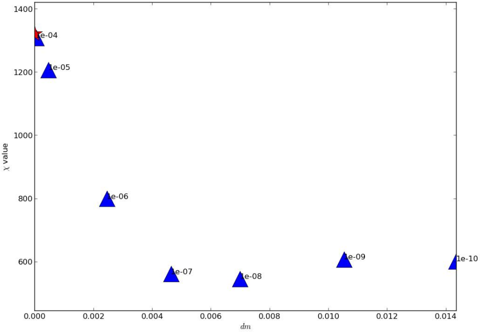
- compute model update $\frac{\delta m}{m}$ based on the coefficients $\mathbf{C} = (N + \lambda I)^{-1}\Phi$
- compute sum of event misfits ϕ for a subset of master stations (~ 20), and plot ϕ against δm (L-curve):



Selection of λ

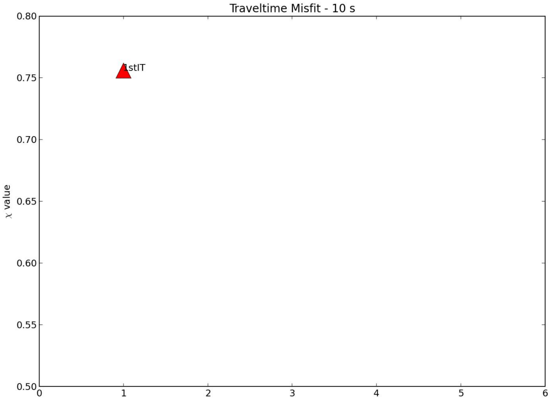
Choose a set of λ values, for $\lambda = 10^{-10}$

- compute model update $\frac{\delta m}{m}$ based on the coefficients $\mathbf{C} = (N + \lambda I)^{-1}\Phi$
- compute sum of event misfits ϕ for a subset of master stations (~ 20), and plot ϕ against δm (L-curve): $\lambda = 10^{-7}$



Reduction of total misfit over iterations

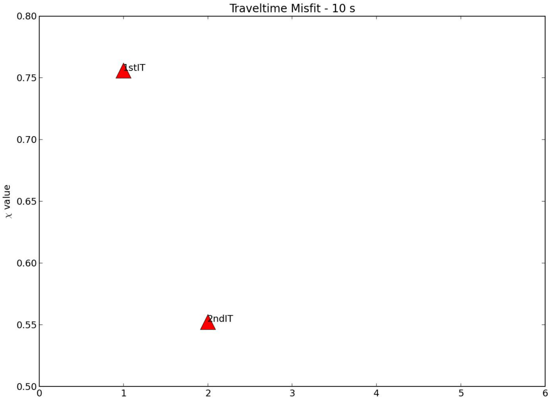
$$\chi = \frac{1}{2}\frac{1}{N_{meas}}\sum_{e,s,p}\left(\frac{T_s - T_d}{\sigma_T}\right)^2$$



- Best value is 0.5 when all time shifts are within STD
- Convergence at $4-5$th iter for $10-50$ sec

Reduction of total misfit over iterations

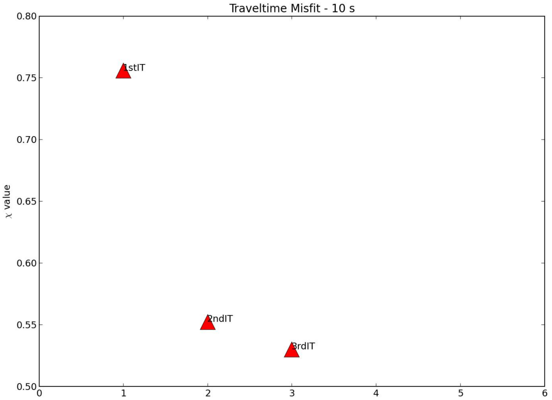
$$\chi = \frac{1}{2}\frac{1}{N_{meas}}\sum_{e,s,p}\left(\frac{T_s - T_d}{\sigma_T}\right)^2$$



- Best value is 0.5 when all time shifts are within STD
- Convergence at $4-5$th iter for $10-50$ sec

Reduction of total misfit over iterations

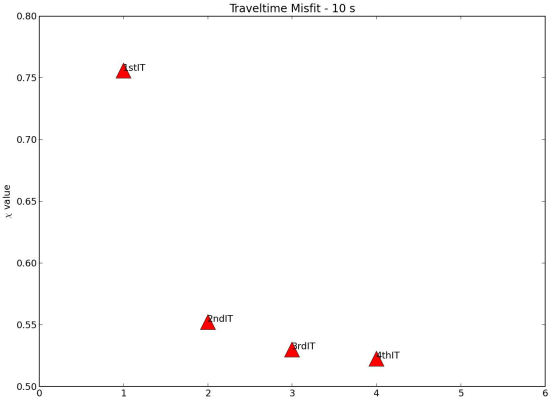
$$\chi = \frac{1}{2}\frac{1}{N_{meas}}\sum_{e,s,p}\left(\frac{T_s - T_d}{\sigma_T}\right)^2$$



- Best value is 0.5 when all time shifts are within STD
- Convergence at $4-5$th iter for $10-50$ sec

Reduction of total misfit over iterations

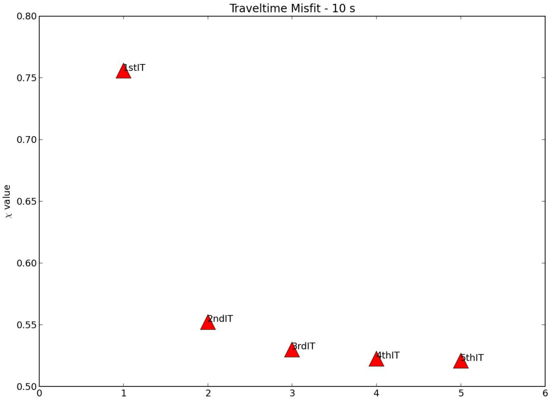
$$\chi = \frac{1}{2}\frac{1}{N_{meas}}\sum_{e,s,p}\left(\frac{T_s - T_d}{\sigma_T}\right)^2$$



- Best value is 0.5 when all time shifts are within STD
- Convergence at $4 - 5$th iter for $10 - 50$ sec

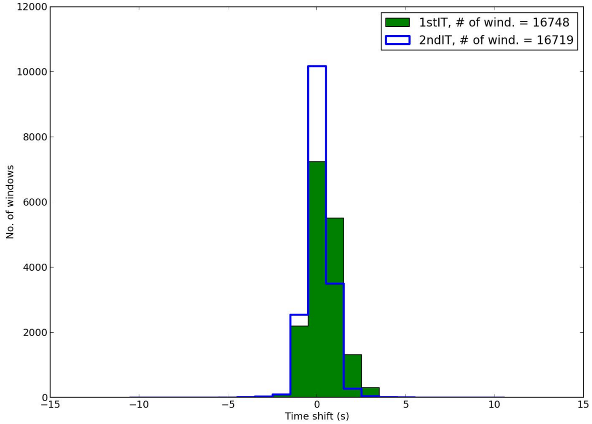
Reduction of total misfit over iterations

$$\chi = \frac{1}{2}\frac{1}{N_{meas}}\sum_{e,s,p}\left(\frac{T_s - T_d}{\sigma_T}\right)^2$$



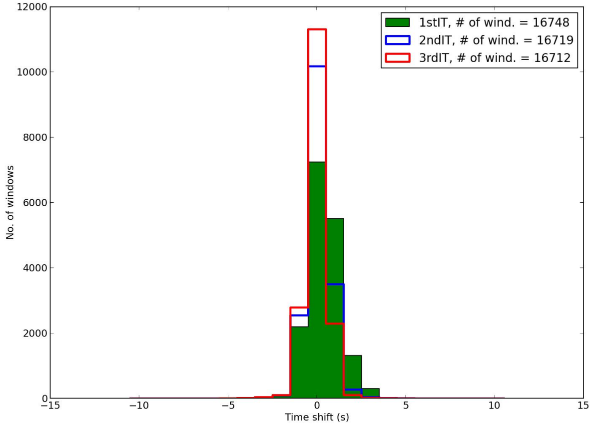
- Best value is 0.5 when all time shifts are within STD
- Convergence at $4-5$th iter for $10-50$ sec

Selected time windows over iterations



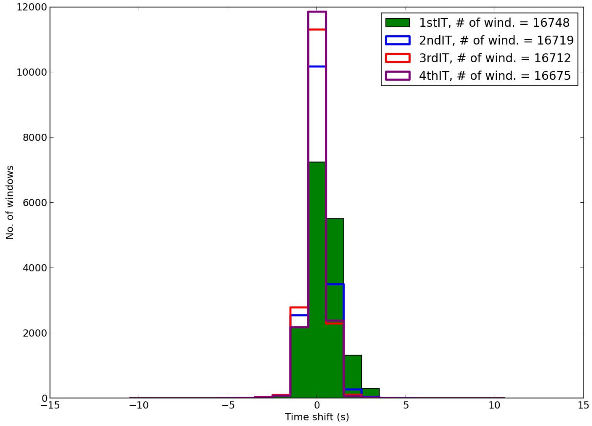
- 10 – 50 sec windows
- Reduction of traveltime anomalies (esp. the slow arrivals)
- More symmetric distribution of traveltime anomaly

Selected time windows over iterations



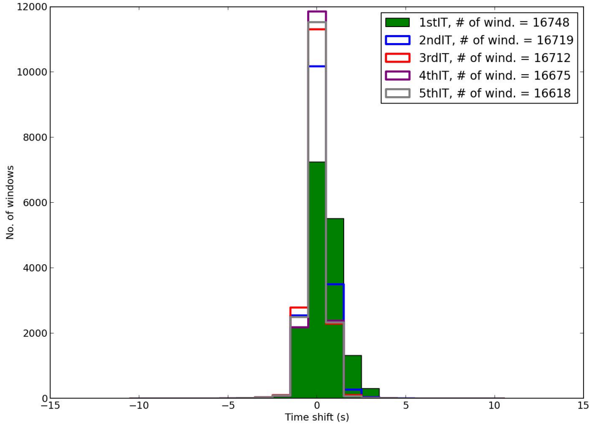
- 10 – 50 sec windows
- Reduction of traveltime anomalies (esp. the slow arrivals)
- More symmetric distribution of traveltime anomaly

Selected time windows over iterations



- 10 – 50 sec windows
- Reduction of traveltime anomalies (esp. the slow arrivals)
- More symmetric distribution of traveltime anomaly

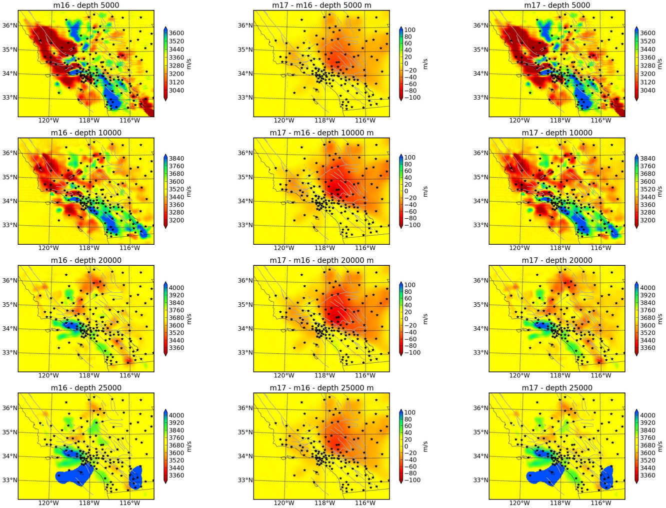
Selected time windows over iterations



- 10 – 50 sec windows
- Reduction of traveltime anomalies (esp. the slow arrivals)
- More symmetric distribution of traveltime anomaly

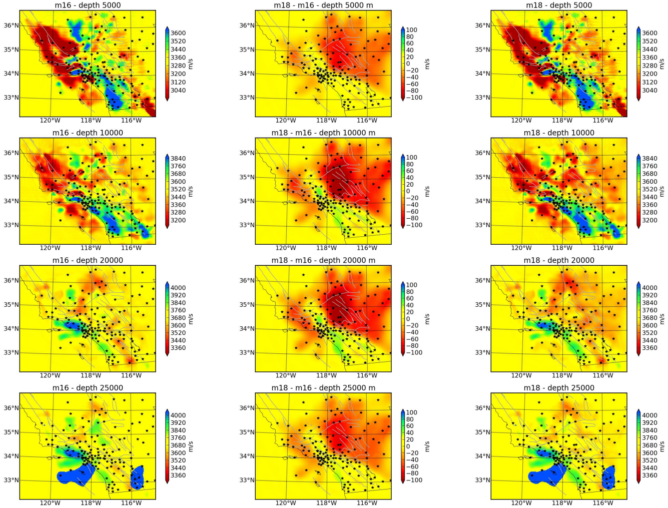
Vs Model updates: from m16 to m20

Smoothed by 15km/5km. Up to 3 − 4% model update. 1st iter: m17



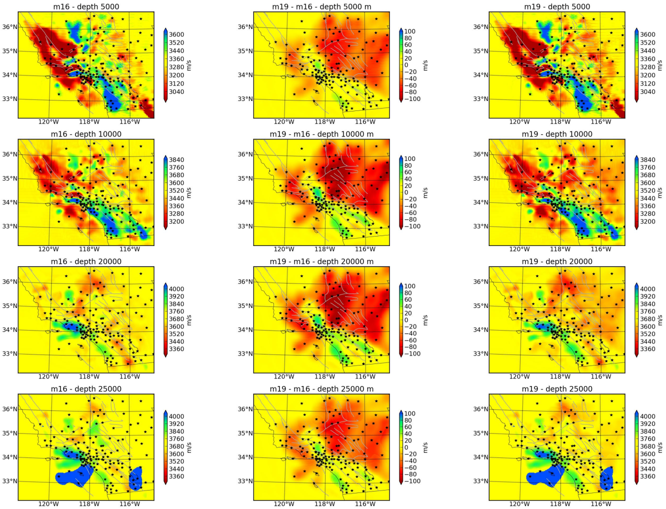
Vs Model updates: from m16 to m20

Smoothed by 15km/5km. Up to $3-4\%$ model update. 2nd iter: m18



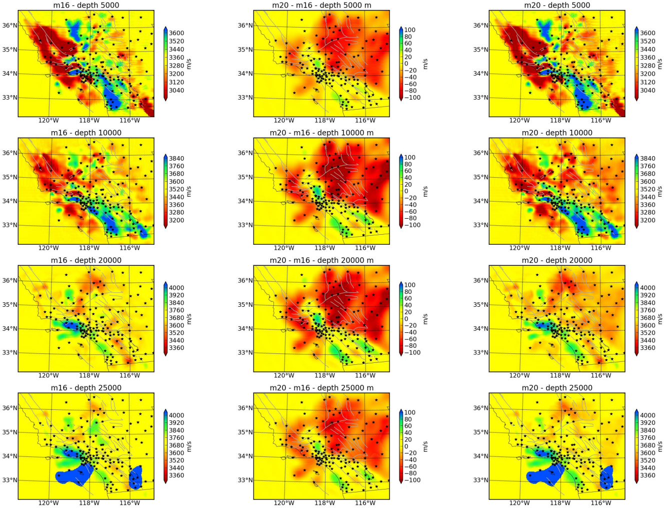
Vs Model updates: from m16 to m20

Smoothed by 15km/5km. Up to 3 – 4% model update. 3rd iter: m19



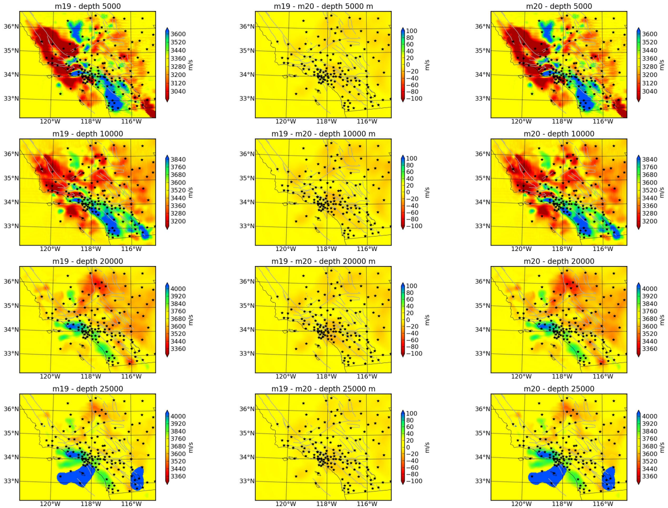
Vs Model updates: from m16 to m20

Smoothed by 15km/5km. Up to 3 – 4% model update. 4th iter: m20



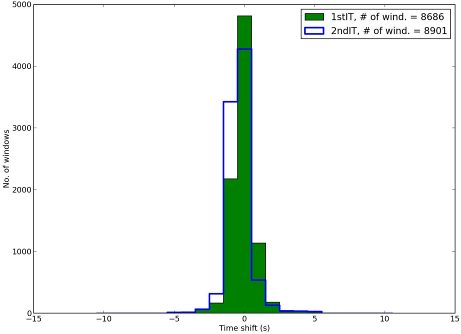
Vs Model updates: from m16 to m20

Smoothed by 15km/5km. Up to 3 − 4% model update. Last model update: m20-m19



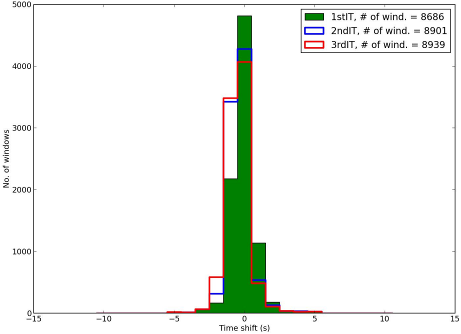
Future work

- Validation by earthquake data (annually $\sim$ 40 eqks with $M_w >= 3.5$ in socal)
- How do $5 - 50$ sec data do over the iterations?



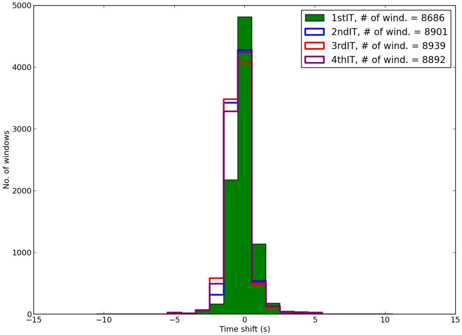
Future work

- Validation by earthquake data (annually $\sim$ 40 eqks with $M_w >= 3.5$ in socal)
- How do $5 - 50$ sec data do over the iterations?



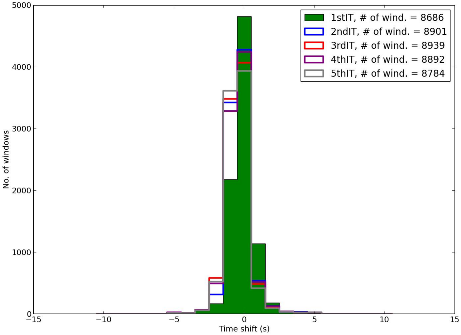
Future work

- Validation by earthquake data (annually $\sim$ 40 eqks with $M_w >= 3.5$ in socal)
- How do $5 - 50$ sec data do over the iterations?



Future work

- Validation by earthquake data (annually $\sim$ 40 eqks with $M_w >= 3.5$ in socal)
- How do $5-50$ sec data do over the iterations?

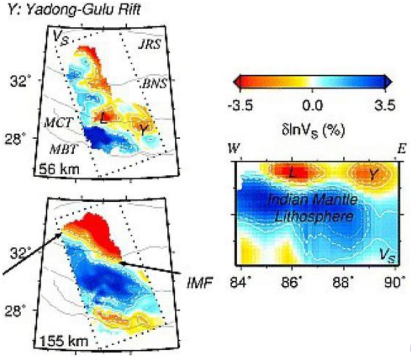


$\Rightarrow$ include $5-50$ sec NCF data.

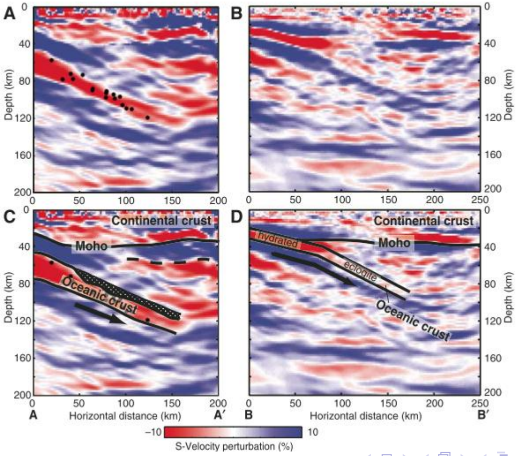
- include other components (e.g., T-T etc) of NCFs.

Seismic array imaging

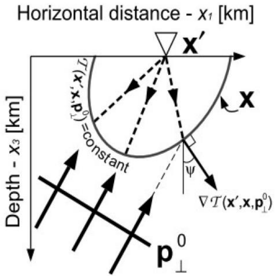
- body- and surface-wave tomography adapted to regional arrays. Traveltime inversions → resolve structure of the size of Fresnel zone.
  e.g., data-adaptive, multiscale approach of finite-frequency, traveltime tomography of central Tibet (*HI-CLIMB*, *Hung et al, 2011*)

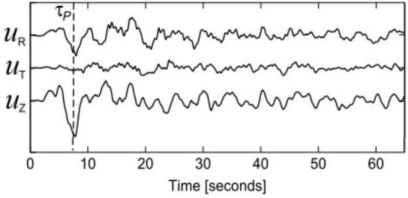


- converted/scattered wave imaging → resolve structure of the size of wavelength of dominant waves (receiver function, migration, generalized Radon transform)
  e.g., migrated images of subducted slabs based on data from *BEAAR* array in Alaska and *CASC93* array in Oregon (*Rondenay et al, 2008*)



Scattered-wave imaging





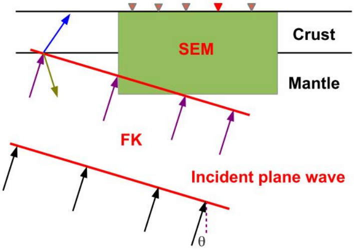
RF: 1-3 sec for P waves and 3-6 sec for S waves; Full global seismic wave simulation at $T < 8.0$ sec numerically costly.
Adaptation of adjoint tomography to scattered-wave imaging:

- forward simulation: response of local media to tele-seismic waves (for upper mantle imaging, assume to be plane waves)
- interface 1D FK solution with SEM simulations (*Tong et al, submitted, 2013*)

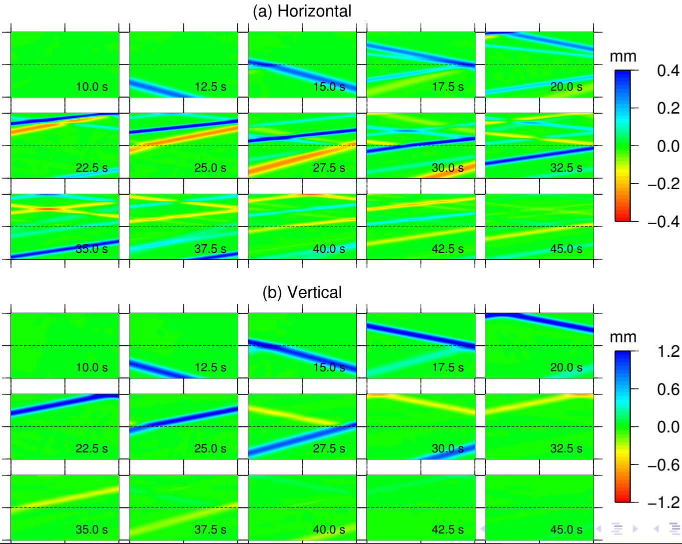
SEM-FK hybrid method

- Response of 1D media to plane-wave incidence computed by frequency-wavenumber (*FK*) method (*Zhu & Rivera, 2002*) and saved on the boundary as (x, t)
- Spectral-element method (*SPECFEM2D*) is applied to the computation domain that includes all local heterogeneities
- interfacing occurs through absorbing boundary condition:

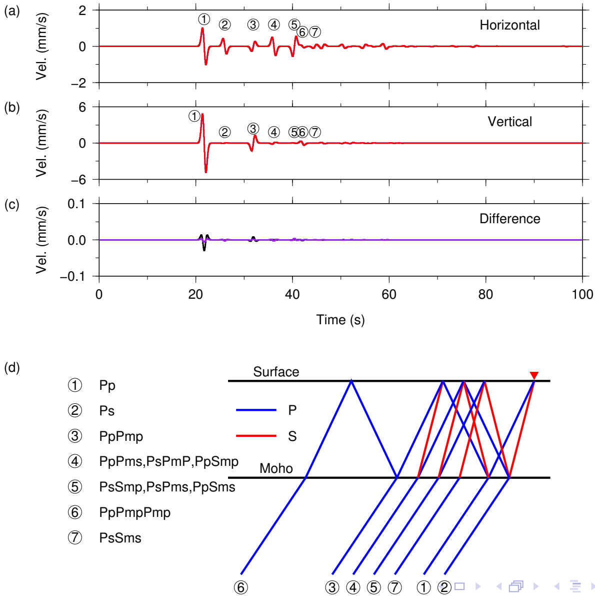
$$\mathbf{T}_{scatt} \cdot \hat{n} = \rho\alpha[\hat{n} \cdot \partial_t \mathbf{u}_{scatt}]\hat{n} + \rho\beta[\hat{t} \cdot \partial_t \mathbf{u}_{scatt}]\hat{t}$$



Validation of SEM-FK hybrid method: two-layer model



Validation of SEM-FK hybrid method



Adjoint tomography for scattered waves

Minimizing

$$\phi(\mathbf{m}) = \frac{1}{2}\sum_{i,r,e}\int ||\mathbf{w}_i(t)[\mathbf{u}(\mathbf{x}_r^i, t, \theta_e; \mathbf{m}) - \mathbf{d}(\mathbf{x}_r^i, t, \theta_e)]||^2 dt, \qquad (1)$$

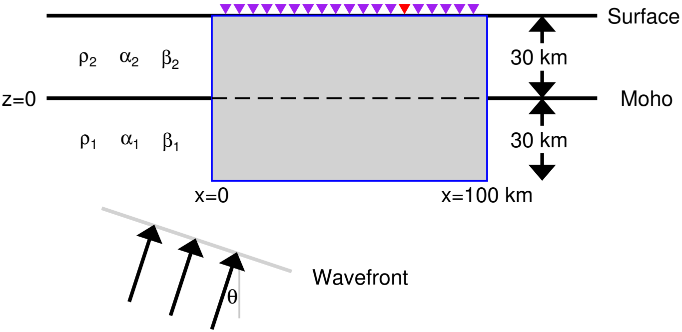
and its variation

$$\delta\phi = \int_S (K_{\rho'}\delta \ln\rho + K_\alpha \delta \ln\alpha + K_\beta \delta \ln\beta) d^2\mathbf{x} + \int_\Gamma K_d \delta \ln d \, d\mathbf{x} \qquad (2)$$

$K_{\rho'}$, K_α, K_β and K_d: Fréchet kernels of density, P- and S-wave speed, and *Moho topography* (e.g. Tromp et al. 2005)

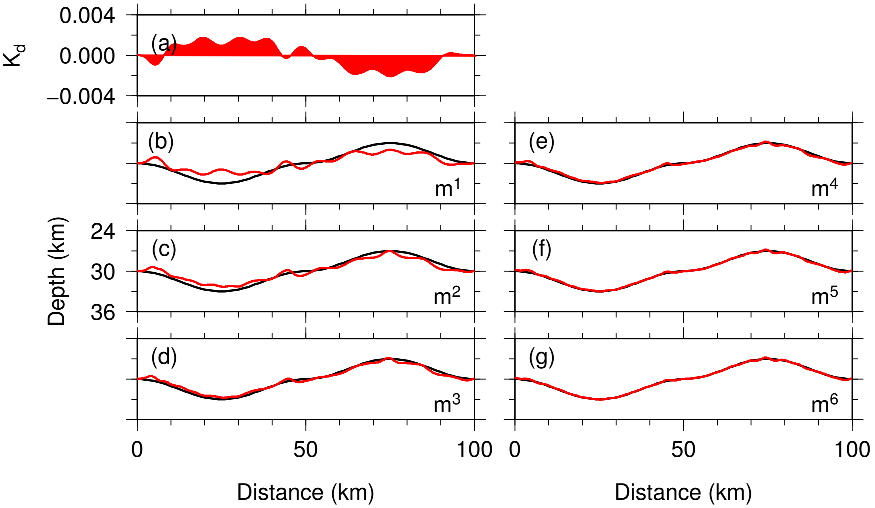
Then a nonlinear conjugate gradient method is used to minimize ϕ iteratively.

Synthetic tests: Moho-only inversion

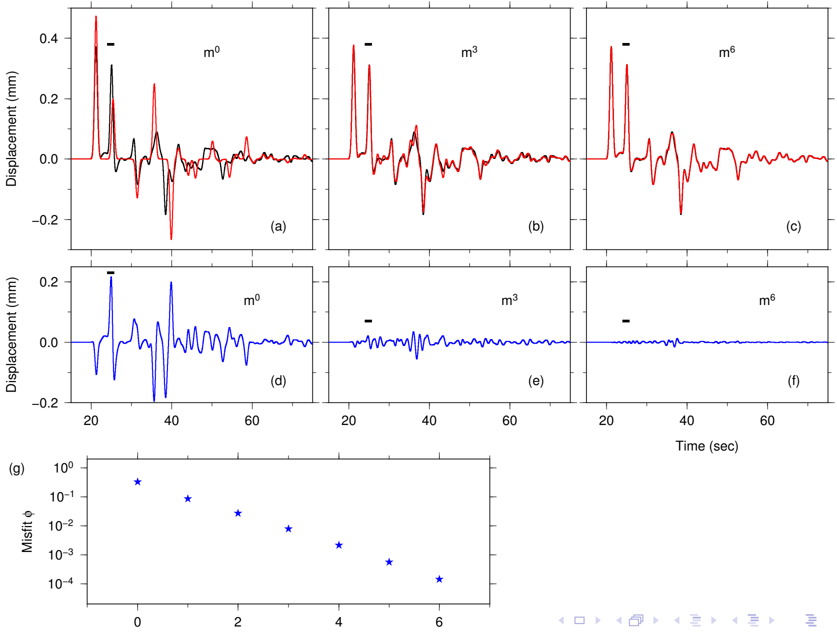


- Target model: sinusoidal Moho undulation with maximum of 3 km.
- Event: P plane-wave incidence angles of 4°, 12°, 20° and 28° from both left and right with a cut-off frequency of 2 Hz.
- Station: *Ps* waveforms from 10 receivers with equal spacing of 10 km.

Moho inversion with Ps phase

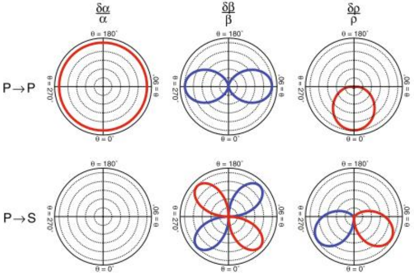


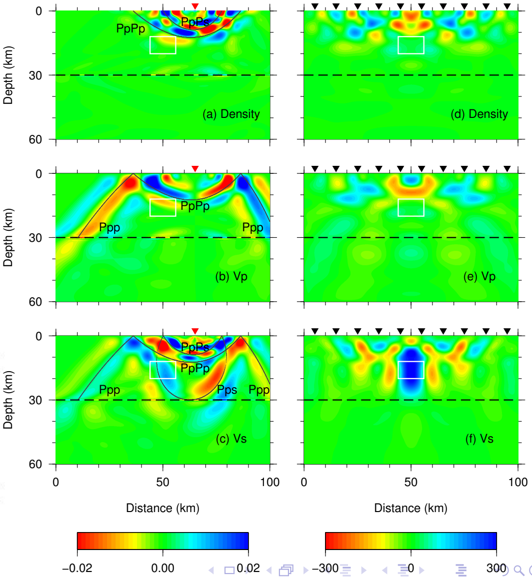
Moho inversion with Ps phase



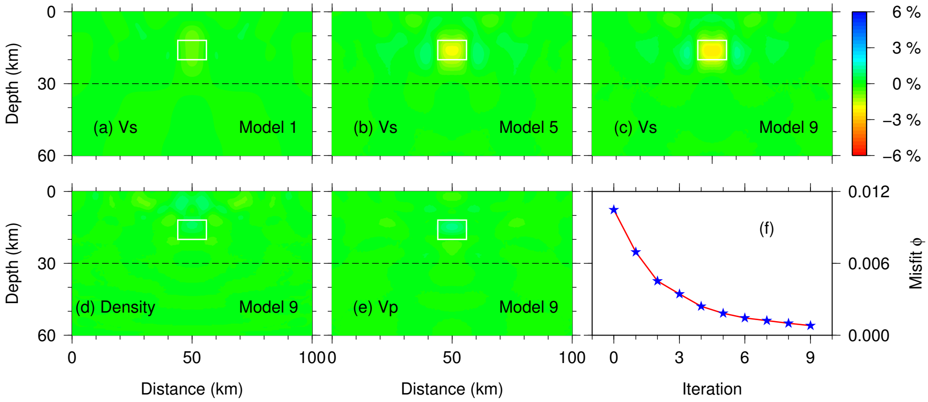
Volumetric structural inversion with Ps waveform

- 6% slow Vs anomaly (12x8 km) in the mid-crust
- 8 events, 10 receivers, cut-off period 1 sec
- individual kernels → isochrons (Rondenay 2008)



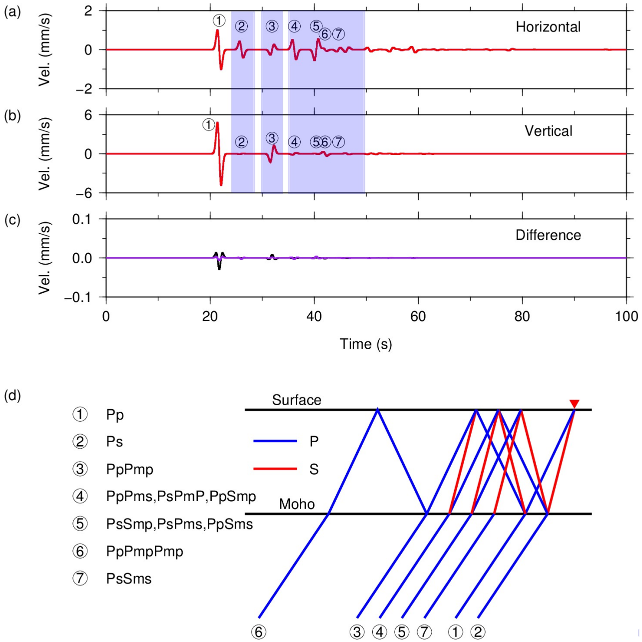


Volumetric structural inversion with Ps waveform

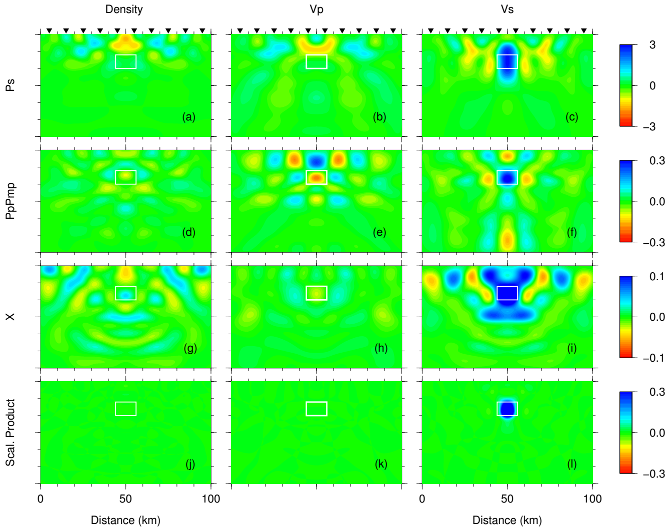


However, we only recovered on average 1.8% slow Vs anomaly.

Ps waveform inversion with preconditioner



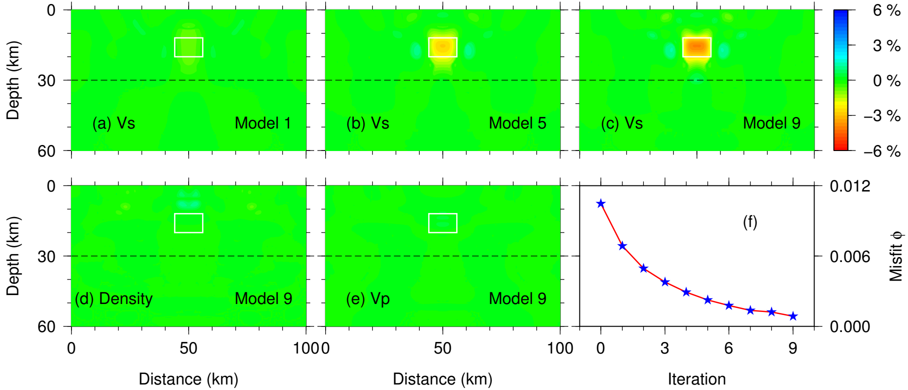
Ps waveform inversion with preconditioner



Scaled product of *Ps*, *PpPmp* and the remaining coda waves *X* kernels as preconditioner

$$g_\rho(\mathbf{x}) = \frac{1}{\eta}\Pi_{i=1}^{3} g_{\rho,i}(\mathbf{x}), \qquad (3)$$

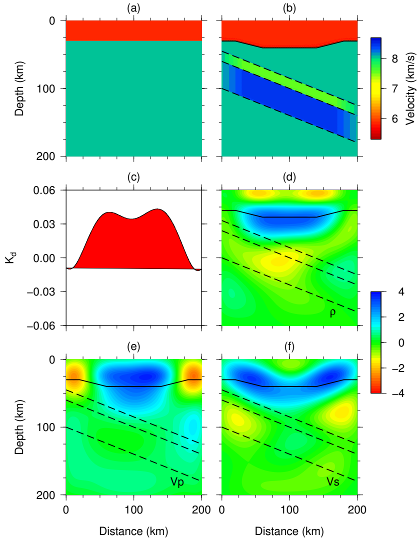
Ps waveform inversion with preconditioner



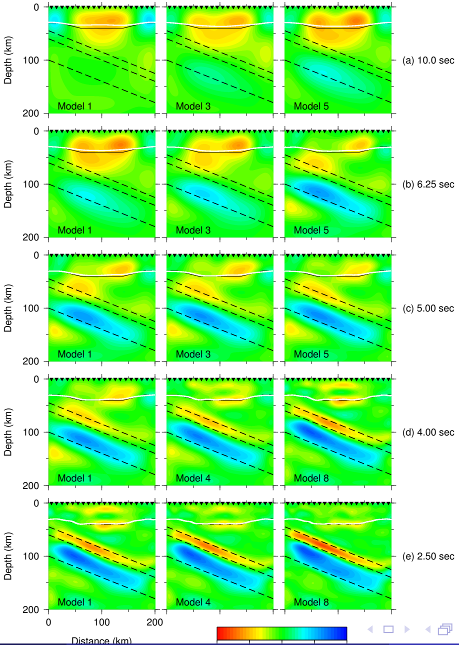
With scaled product of kernels as preconditioner, we recover almost the full amplitude of Vs anomaly.

Full waveform inversion of a subducted slab model

- slab model: -6% slow oceanic crust (14 km) atop +4% fast oceanic mantle (37 km)
- continental Moho depression 10 km
- 16 events, 20 receivers with spacing of 10 km
- hierarchical inversion over 10, 6.25, 5, 4 and 2.5-sec waveform of both P and its coda waves.

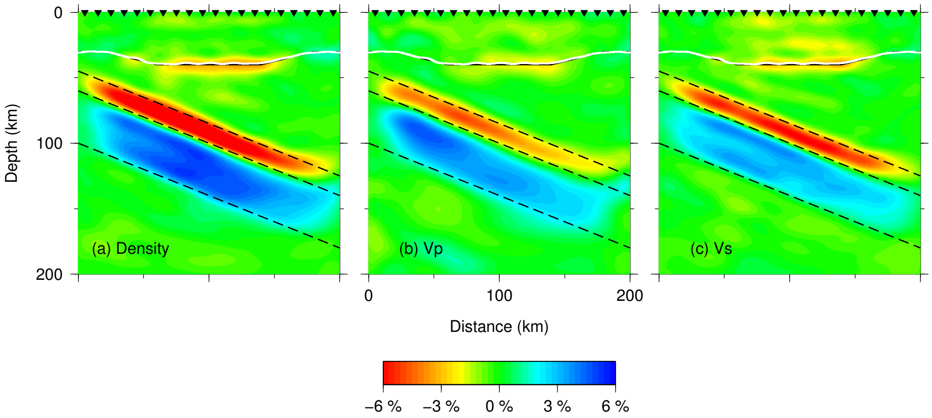


Full waveform inversion: 2-layer subducted slab model



Full waveform inversion of 2-layer subducted slab: Final model

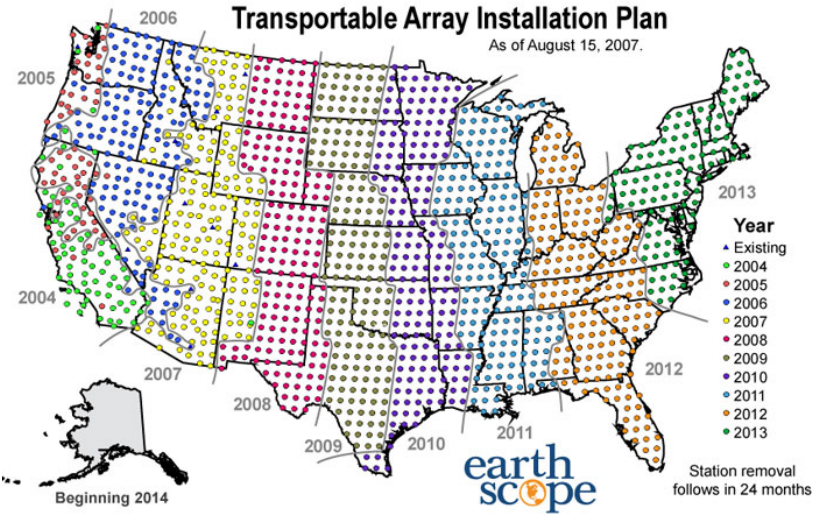
- recovery of Moho variation
- sharp subducted oceanic Moho
- almost full recovery of anomaly amplitude



Conclusions

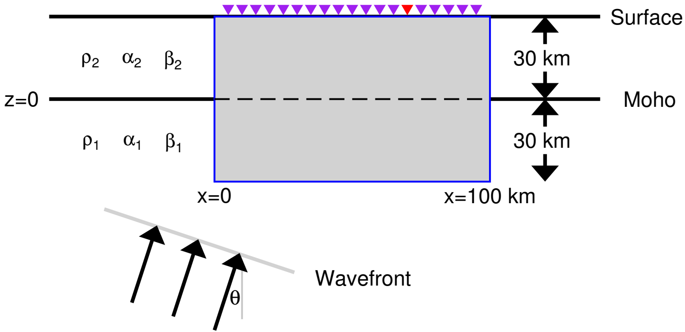
- Preliminary adjoint tomography based on NCFs of southern California stations, which provide improved resolution for the mid- and lower crust, complementary to earthquake data.
- Implement an SEM-FK hybrid method to simulate the response of local heterogeneities to plane-wave incidence, and allow full waveform inversions of teleseismic converted/scattered waves for the imaging of crust and upper-mantle structures beneath seismic arrays.
- Scaled product of sensitivity kernels for different phases as preconditioner, the combination of traveltime and waveform inversion, and hierarchical inversions from long- to short-period waveforms proved to be beneficial in high-resolution seismic array imaging based on SEM-FK hybrid method.

USArray TA deployment plan



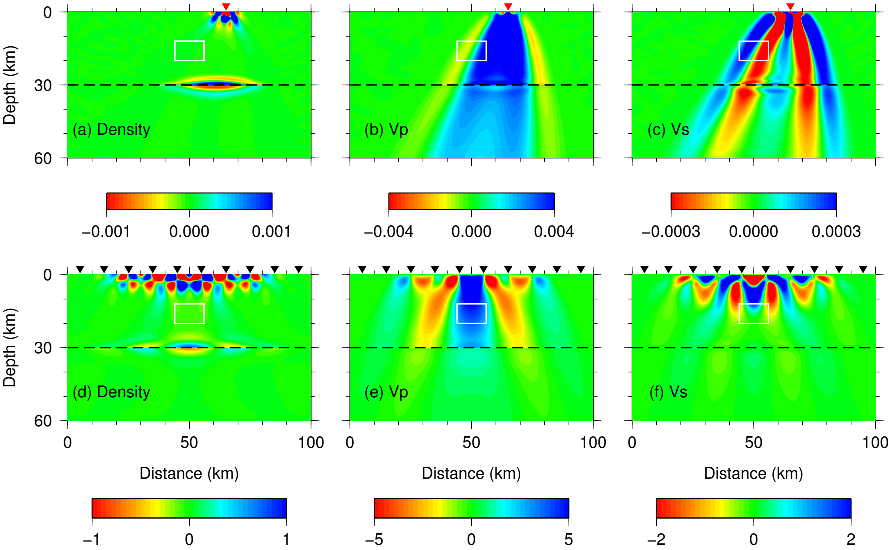

Thank You!

Combined traveltime and waveform inversion



- Event: P plane-wave incidence angles of 4°, 12°, 20° and 28° from both left and right with a cut-off frequency of 1 Hz.
- Station: 10 receivers with equal spacing of 10 km.
- Anomaly: a 6.0% slower Vp anomaly (12 km by 8 km) in the middle crust.

Combined traveltime and waveform inversion



Combined traveltime and waveform inversion