Supplementary Materials

SURFACE WAVE TOMOGRAPHY METHODS

To measure the Rayleigh wave phase velocities, we employ the two-plane-wave method described by Forsyth and Li (2005) and Yang and Forsyth (2006), with some modifications. There is a basic description of the procedures in the submitted text and for most of the details we refer the reader to these original references. In this supplement, we briefly describe the important modifications. Our modifications include:

1. Instead of using the average phase velocity for the region as the starting model for inverting for lateral variations in phase velocity, we use an *a priori* crustal model to predict the lateral variations in phase velocity, then solve for deviations from that starting prediction. This is probably the most important modification, so we describe the construction of the starting crustal model below under a separate heading.

2. Instead of using a Gaussian interpolation or weighting function for the velocities, we interpolate linearly between each grid point so that the coefficient for velocity at a grid point is equal to the velocity at that point, rather than the velocity being a weighted function of the coefficients of the grid points in the neighborhood.

3. In the two-plane-wave method, the structure of the incoming wavefield is approximated as the sum of two plane waves with unknown phase, amplitude and direction of approach, thus amounting to six wavefield parameters for each earthquake source. Because the area of this study is large, to allow for variations in the wavefield we subdivide the area into four, slightly overlapping sub-regions, applying the two-plane-wave approximation separately in each sub-region for a total of up to 24 wave parameters, depending on the distribution of stations available for each event. This is the same approach employed by Yang and Ritzwoller (2008) and Wang et al. (2009). Our tests show that perturbing the boundaries of the sub-regions have no significant effects.

4. In each iteration of the inversion there are two steps. We first use a grid search technique to solve for the wave parameters keeping the current velocity model fixed, then invert for perturbations to both the velocity model and wave parameters using a linearized inversion, but with the wave parameters relatively heavily damped. In the original approach, we used a simulated annealing method rather than grid search to solve for the non-linear wave parameters. The grid search technique is somewhat more stable.

5. We use a sequence of inversions for phase velocity on refined grids. We begin with a grid spacing of 1.2°, interpolate onto a fine grid, then invert for the shear velocity structure at each point beginning with the *a priori* crustal and mantle structure as a starting model. We then use the predicted phase velocities for that new structure as the starting model for the next phase velocity inversion on a finer grid, in this case with spacing 0.8°. We repeat the process, using the previous shear velocity model as the starting model for the shear velocity inversion, and successively refining the grid to 0.6° and 0.4°. At each step, in inverting for shear velocity structure we use only those periods with rank for phase velocity parameters (as measured from the resolution matrix) greater than 80% that of the rank of the period with the best resolution. This restriction prevents distortion of the vertical velocity structure from having different effective damping of the lateral variations in velocity at different periods. It also largely avoids underestimates of the amplitude of lateral velocity variations due to damping, because each successive step uses the laterally varying previous model as a starting model rather than a laterally uniform mantle. In practice, we use all periods from 18 to 143 s in the first step, and periods ≤ 100, 80, and 40 s in the subsequent steps. Because the depth range of sensitivity to shear velocity structure is roughly proportional to period, this means that in our final shear velocity model, we have higher lateral resolution at shallow levels in the mantle than deeper in the asthenosphere.

6. We double the nominal parameter standard deviations derived in the inversion for phase velocities to conservatively represent the effects of tradeoffs with the wave parameters that are restricted by the damping of the wave parameters in the linearized step. There is no damping of the wave parameters in the grid search part of each iteration. This doubling gives a more accurate indication of the scatter in phase velocity from period to period or from grid point to grid point in tectonically similar areas.

7. We solve for laterally varying azimuthal anisotropy as well as the azimuthally averaged velocities used to invert for the shear velocity structures presented here. In this study, we use a constant 1.2° grid for the anisotropy terms, which we restrict to 2θ variations, where θ is azimuth of propagation, because those are the dominant terms for Rayleigh waves (Smith and Dahlen, 1973).

8. In addition to the damping of the phase velocity and shear velocity inversions provided by assigning an *a priori* error to each parameter in the starting models (we use 0.25 km/s standard deviation), we employ a strong minimum curvature constraint on changes to the model. Rather than arbitrarily varying the weight given to this smoothness constraint, we choose the curvature constraint that effectively minimizes the collective covariance between adjacent model parameters. Both constraints are applied through the *a priori* model covariance matrix, and, since in each phase velocity inversion we iterate to a solution, we penalize changes to the starting model not just the changes in each iteration using the formulation of Tarantola and

Valette (1982). In the shear velocity inversions, we do not apply the smoothness constraint across the Moho.

9. Although Rayleigh waves are most sensitive to Sv velocity structure, they also depend on density and P-wave velocity. P-wave sensitivity is limited primarily to the crust. Rather than having density and P velocity be very poorly resolved free parameters, we couple the changes in density and P velocity to changes in S velocity, using fixed ratios of 0.3 (Mg/m^3)/(km/s) and 1.2 respectively.

CRUSTAL MODEL

The most nonlinear part of the inversion of phase velocities for S velocity structure is associated with changes in crustal structure, because a low-velocity surface layer affects the depth of penetration and sensitivity of Rayleigh waves to changes in shear velocity. Thus, a good crustal starting model is important. We begin to construct the crustal model with a default velocity model in which P/S wave velocity varies from 5.45/3.06 km/s at the surface to 6.58/3.72 km/s at 22 km. In the lower crust below 22 km, P/S velocity is 6.8/3.89 km/s. Although largely irrelevant to this paper, in the oceans we use a laterally varying water layer and a 6 km-thick crust. On the continental margins, where water depth is less than 3 km, we vary the crustal thickness from 6 km to 28 km at the coast with thickness changing inversely with water depth. In southern California, we replace the default upper crustal model with the Three-Dimensional Community Velocity Model from SCEC (Kohler et al., 2003). Next we modify the upper 10 km of the crust to be consistent with phase velocity data for 8 s period Rayleigh waves derived from ambient noise analysis (Lin et al., 2008). In this modification, we adjust the shear velocity in proportion 4:2:1 in 3 layers starting from the surface of 3.0, 3.0 and 4.0 km thickness, with the implicit assumption that most of the pronounced, short-period, Rayleigh wave velocity variations are associated with sedimentary basins that may have very slow shear wave velocities near the surface.

We constrain total crustal thickness with P to S receiver functions from the EarthScope Automated Receiver Study (EARS, Crotwell and Owens, 2005). This compilation of receiver functions in the form of H-k stacks (stacked amplitudes as a function of thickness and Vp/Vs ratio) includes all stations, not just EarthScope's USArray. Rather than automatically accepting the maximum stacking amplitude as the solution, we in effect assume that crustal rocks have Vp/Vs ratios of 1.80 with a standard deviation of about 0.075. Maxima with ratios greater than 1.95 or less than 1.65 are rejected. If there is another strong, local maxima within the acceptable range, that value is chosen. Otherwise, the station is discarded, as are those with ambiguous solutions in conflict with a priori estimates of crustal thickness in the region or with solutions for the crustal thickness at surrounding stations. We use a linear weighting function declining to zero at 1.0° distance to interpolate and smooth the results from individual stations. Our final input crustal thickness model with the stations used for receiver function analysis is shown in Figure DR1. We then solve for the predicted phase velocities for this crustal structure with an initially laterally uniform mantle velocity structure and use it as the starting model for the phase velocity inversion.

When we invert the phase velocities to obtain the shear velocity structure, the crustal velocities are allowed to vary, but the crustal thickness is fixed. With the period range we employ, typically there are 1 to 2 resolvable pieces of information about crustal structure. Usually the inversions make little or no change in the shallow crust, but the velocity of the lower crust is often significantly altered from the default starting model. For example, the shear velocity of the lowermost crust beneath the western Sierra Nevada exceeds 4.0 km/s, in agreement with the interpretation that there is eclogitic lower crust in that region. If the crustal thickness is incorrect, then there will be a tradeoff between Moho depth and the velocities in the lowermost crust and uppermost mantle, as described in the next section. However, we believe the thicknesses based on the receiver function analysis are more reliable than any adjustments we would make with the surface waves, so we choose to leave the thicknesses fixed and interpret the velocities near the Moho with caution.

SHEAR VELOCITY RESOLUTION

Due to the underdetermined nature of the inverse problem stemming from the limited vertical resolution of surface waves given the errors in the data, we are able to uniquely resolve only averages of the velocity over depth ranges rather than the absolute velocity at any given depth. There is some tradeoff between resolving length or depth range and the uncertainty in the average. Details of the model smaller than the resolving length are highly dependent on the starting model. Examples of the vertical resolution length beneath the amagmatic zone are shown in Figure DR2. The black line indicates the particular velocity model found in the final shear wave velocity inversion. The boxes represent resolvable depth ranges and the uncertainty ($\pm 2\sigma$ assuming linearity) in the velocity we can resolve within that depth range. Here, the average velocity beneath the amagmatic zone within the depth range of 50-100 km is 4.06 ± 0.03 km/s, which is within the range and below the average value for regions that have experienced recent (<1 Ma) volcanism. Nonlinearities can increase the uncertainty, with the primary nonlinearity in this problem being the crustal thickness.

To demonstrate the effects of variations in crustal thickness, we introduce starting models with crustal thicknesses spanning the range we consider to be a reasonable range of possibility for this area (thinned by 5 km and thickened by 10 km from the 35 km *a priori* value). Results from these inversions are shown in Figure DR2A, where it can be seen that there is little effect deeper than 60 km and that the averages for these models lie within the confidence limits for the original model. Thus, we can say with confidence that these low velocities beneath the amagmatic zone are independent of the starting crustal model. We also demonstrate that the velocity averages are largely independent of the starting mantle structure by using three different input mantle structures: our original model that iterated with progressively refined grids; one with laterally (but not vertically) uniform velocities based on the average velocity structure beneath southern California (Mantle 1, Yang and Forsyth, 2006), and one with laterally and vertically uniform velocities at 4.2 km/s from the Moho down to 170 km (and increasing shear velocity below 170 km, Mantle 2). The latter two models are a little faster in the 50 to 100 km depth range because the shear velocity inversion was performed in a single step, they had higher starting velocities in this depth range, and damping limited the perturbations from the starting model. Nevertheless, the models are similar (Figure DR2B), indicating that slow shear wave velocities beneath the amagmatic zone are a robust feature.

REFERENCES CITED

Crotwell, H. P., and Owens, T. J., Automated receiver function processing, Seismological Research Letters, v. 76, p. 702-708

Forsyth, D.W. and Li, A., 2005, Array analysis of two-dimensional variations in surface wave phase velocity and azimuthal anisotropy in the presence of multipathing interference, in SeismicEarth: Array Analysis of Broadband Seismograms, Geophysical Monograph Series, Levander, A. and G. Nolet, eds, v. 157, p. 81-97

Kohler, M., Magistrale, H., and Clayton, R., 2003, Mantle heterogeneities and the SCEC three-dimensional seismic velocity model version 3, Bulletin Seismological Society of America v. 93, p. 757-774

Lin, F., M.P. Moschetti, and M.H. Ritzwoller, 2008, Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.*, doi:10.1111/j1365-246X.2008.03720.x

Smith, M. L., and Dahlen, F.A., 1973, The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium, Journal of Geophysical Research, v. 78, p. 3321-3333

Tarantola, T. and B. Valette, 1982, Generalized non-linear problems solved using the least-squares criterion, Reviews in Geophysics, v. 20, p. 219-232

Yang, Y. and Forsyth, D.W., 2006, Regional tomographic inversion of the amplitude and phase of Rayleigh waves with 2-D sensitivity kernels, Geophysical Journal International, v. 166, p. 1148-1160

Yang, Y. and Ritzwoller, M.H., 2008, Teleseismic surface wave tomography of western US using the Transportable Array component of USArray, Geophysical Research Letters, L04308, doi:10.1029/2007GL032278.

FIGURE CAPTIONS

Figure DR1. Crustal thickness model used as input for the phase velocity inversion. Crustal thickness estimates were made by weighting measurements from P to S receiver functions observed at the shown stations (black triangles) according to distance from the station.

Figure DR2. Resolvable averages (gray boxes, where width of the box represents ± 2σ) of the shear wave velocity beneath the amagmatic zone from the model in this study (black line). Experimental results to determine the effect of variations in the input (A) crustal thickness and (B) underlying mantle on the retrieved shear wave velocities. In both cases, the retrieved shear wave velocities fall within the bounds of the resolvable average shear wave velocity of the original model, indicating that the slow shear wave velocities beneath the amagmatic zone are a robust feature of the model.

Figure DR1

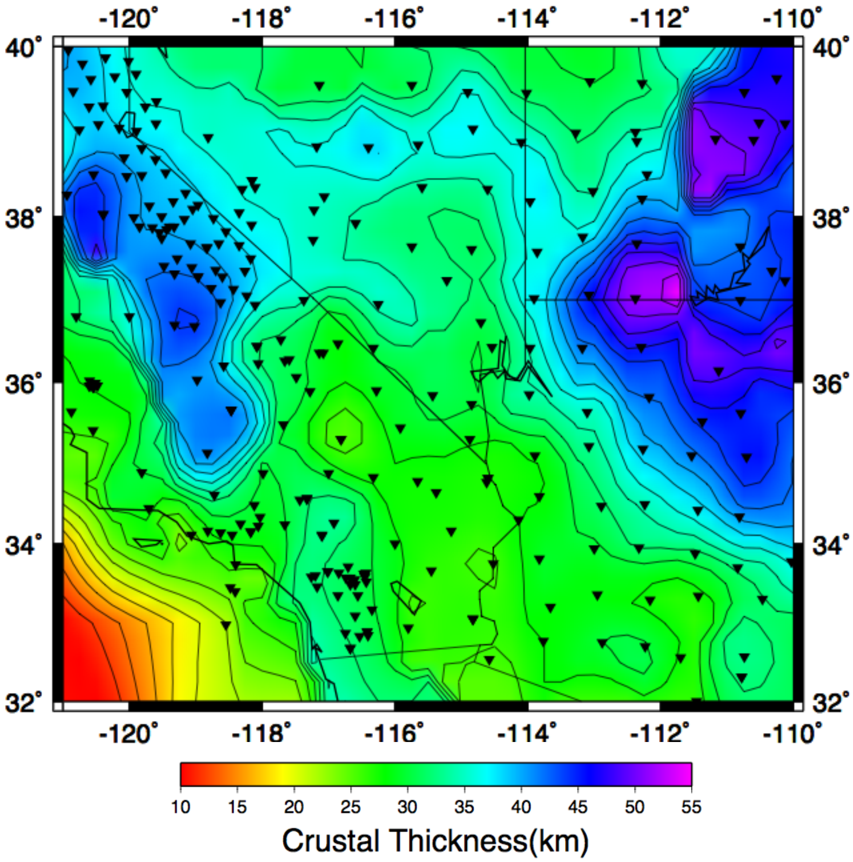
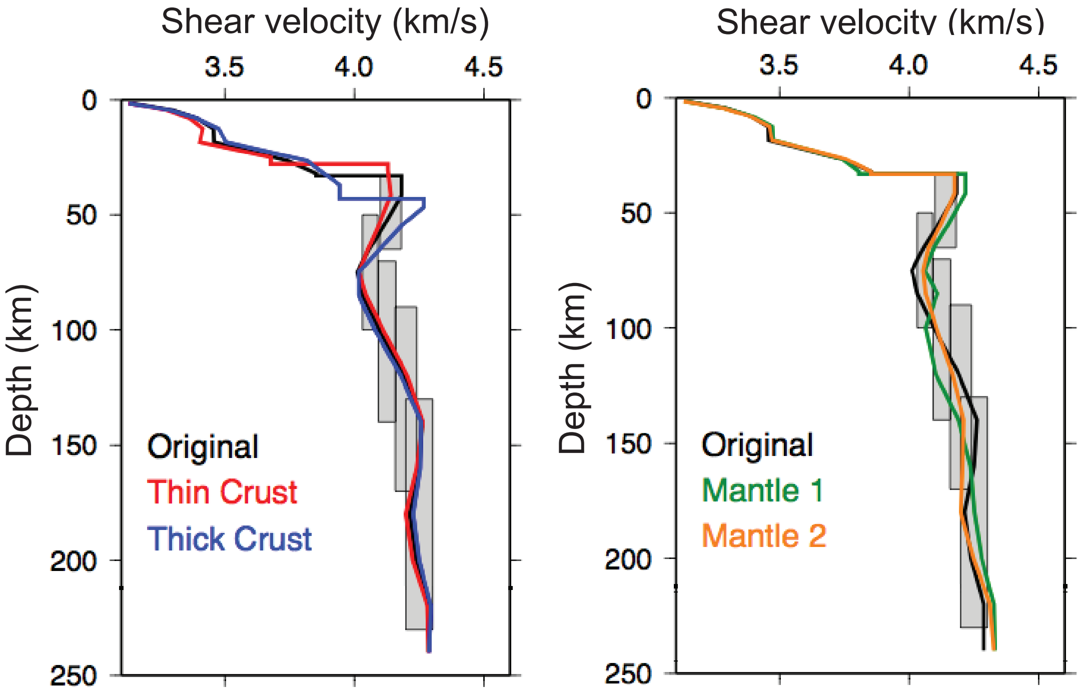

-120°
-118°
-116°
-114°
-112°
-110°
40°
38°
36°
34°
32°
10
15
20
25
30
35
40
45
50
55
Crustal Thickness(km)

Figure DR2


Shear velocity (km/s)
3.5
4.0
4.5
0
50
100
150
200
250
Depth (km)
Original
Thin Crust
Thick Crust
Shear velocity (km/s)
3.5
4.0
4.5
0
50
100
150
200
250
Depth (km)
Original
Mantle 1
Mantle 2