

Effects of shallow density structure on the inversion for crustal shear wavespeeds in surface wave tomography

Guangchi Xing¹, Fenglin Niu^{1,2,*}, Min Chen¹, Yingjie Yang³

1. Department of Earth Science, Rice University, Houston, Texas, USA
2. State Key Laboratory of Petroleum Resource and Prospecting, and Unconventional Natural Gas Institute, China University of Petroleum, Beijing, China
3. Australian Research Council Centre of Excellence for Core to Crust Fluid Systems/GEMOC, Department of Earth and Planetary Sciences, Macquarie University, Sydney, New South Wales, Australia

(*Corresponding author, Tel.: +1-713-348-4122; email address: niu@rice.edu)

SUMMARY. Surface wave tomography routinely uses empirically scaled density model in the inversion of dispersion curves for shear wavespeeds of the crust and uppermost mantle. An improperly selected empirical scaling relationship between density and shear wavespeed can lead to unrealistic density models beneath certain tectonic formations such as sedimentary basins. Taking the Sichuan basin east to the Tibetan plateau as an example, we investigate the differences between density profiles calculated from four scaling methods and their effects on Rayleigh wave phase velocities. Analytical equations for 1-D layered models and adjoint tomography for 3-D models are used to examine the tradeoff between density and S-wave velocity structures at different depth ranges. We demonstrate that shallow density structure can significantly influence phase velocities at short periods, and thereby affect the shear wavespeed inversion from phase velocity data. In particular, a deviation of 25% in the initial density model can introduce an error up to 5% in the inverted shear velocity at middle and lower crustal depths. Therefore one must pay enough attention in choosing a proper velocity-density scaling relationship in constructing initial density model in Rayleigh wave inversion for crustal shear velocity structure.

KEYWORDS: Surface wave tomography; density; shear wavespeed; adjoint tomography; Monte-Carlo inversion.

1. INTRODUCTION

The dispersion of surface wave phase velocities has long been utilized to image the crust and upper mantle shear velocity structure (e.g. Shapiro & Ritzwoller, 2002; Shapiro et al., 2005; Yao et al., 2006; Lin et al., 2007; Bensen et al., 2009; Yang et al., 2012). Moreover, development in surface wave tomography has been accelerated by wave dispersion datasets derived from teleseismic earthquakes and seismic interferometry, i.e., empirical Green's functions (EGFs) of surface waves extracted from cross correlations of ambient noise (Shapiro et al., 2005) or diffuse earthquake coda (Campillo & Paul, 2003). In particular, EGFs extracted from seismic interferometry, which serve as complementary data to earthquake recordings, have two major advantages: 1) EGFs compensate for limited data coverage due to uneven earthquake distribution; 2) EGFs avoid the source property uncertainties in earthquake studies (Yang et al., 2007).

Traditional surface wave tomography based on seismic interferometry generally consists of the following steps. First, Love or Rayleigh wave phase or group velocity dispersion curves are measured from cross-correlation of each station pair. Subsequently, 2-D tomography based on ray theory is adopted to generate phase or group velocity maps in a study area. Finally, local dispersion curves at individual grids are extracted from these maps and then inverted for individual 1-D shear velocity profiles to compose 3-D velocity model. This multistep analysis has been widely employed in surface wave studies at various local or regional scales (Lin et al., 2007; Bensen et al., 2009; Yang et al., 2012).

Surface wave phase or group velocities are more sensitive to S-wave velocity (β , V_s) than P-wave velocity (α , V_p) and density (ρ). The sensitivity to density seems to decrease with increasing periods, and becomes nearly negligible for long-period earthquake data. Also, α and ρ are proportional to β for crustal and mantle rocks. Therefore most of the surface wave studies only invert for β and compute α and ρ by using an assumed V_p/V_s ratio and an empirical velocity-density scaling relationship derived

from experiment data of crustal and mantle rocks.

Four empirical scaling relationships have been widely used, which are: (1) the constant velocity-density scaling relationship (referred to as CM in this paper), $\rho=0.81\beta$ (Christensen & Mooney 1995; examples of studies: Bensen et al., 2009; Yang et al., 2012); (2) the Gardner's rule, $\rho=1.74\alpha^{0.25}$ (Gardner et al., 1974; examples of studies: Giancarlo, 2010; Cauchie & Saccorotti, 2013); (3) the Birch's law, $\alpha=3.125\rho-2.40$ (Birch, 1961; examples of studies: Li et al., 2012; Guo et al., 2015); and (4) the Nafe-Drake relation (Ludwig et al., 1970) represented by 5th order polynomial (Brocher, 2005; examples of studies: Yao et al., 2008; Huang et al., 2014):

$$r = 1.6612a - 0.4721a^2 + 0.0671a^3 - 0.0043a^4 + 0.000106a^5 \quad (1)$$

In the remainder of this paper, we will refer to these four scale relations as the CM, Gardner, Birch and Nafe-Drake velocity-density scaling models.

The velocity-density empirical relationship has been used in dispersion data inversion in two ways. It is first used in constructing the initial density model, i.e., using an absolute velocity model to compute the absolute values of the density profile, which is referred to absolute velocity-density scaling here. Some studies also use the S-wave velocity updates to scale density changes during the inversion. Here since scaling is performed in terms of perturbations, i.e., $\delta \ln \rho = k \delta \ln \beta$, where k is a constant, we refer to it as the perturbation or relative scaling relationship. Some studies fix the density profile during the inversion (e.g., Yang et al., 2012), which means that k=0 in the relative density-velocity scaling. The V_p/V_s scaling is also used similarly in both constructing the initial P-wave velocity model and in updating the velocity models during the inversion.

As mentioned above, traditional surface wave tomography requires intermediate steps of measuring dispersion curves, projects them onto a set of grids by using a 1-D raypath approximation, and then inverts grid-based dispersion data by assuming a 1-D layered model. Both the approximation and assumption could affect the inverted 3-D velocity structure. Recently, Chen et al. (2014) utilized the adjoint tomography technique that formulates the inversion by a direct matching of the surface

waveform data with synthetics of 3-D models, without performing the intermediate steps. The initial density model is also computed by using the Nafe-Drake scaling, which is iteratively updated with a relative scaling of $\delta \ln \rho / \delta \ln \beta = 0.33$. Thus the issue of density-velocity and V_p/V_s scaling also exists in the adjoint surface waveform tomography here.

Recently, the advent of the interferometry method brings the short-period data into the surface wave community and these kind of data are sensitive to densities and velocities with comparable magnitude at shallow depths (Yang et al., 2012). Thus it is necessary to investigate whether and how different scalings could affect the final S-wave velocity model in the conventional and adjoint tomographic inversions. To do so, we first calculate the sensitivity kernels of phase velocities to the S-, P-wave velocities and density using an analytical equation derived from the variation principle (Aki & Richard, 2002). We find short period phase velocities are indeed very sensitive to density structures at shallow depth. We further generate a set of synthetic dispersion and waveform data, and conduct 1-D and 3-D inversions by using different initial models with a wide range of density and P-wave velocities. We also investigate the effect of different perturbation scalings ($\delta \ln \rho / \delta \ln \beta$, $\delta \ln \alpha / \delta \ln \beta$) on the final S-wave velocity model.

We find significant tradeoffs exist between shallow-depth density and mid-to-lower crust shear wavespeeds in ambient noise tomography. More specially, the absolute density-velocity scaling used in constructing the initial model is the most important scaling relationship that can significantly affect the final velocity model, while the effects of the perturbation scaling, $\delta \ln \rho / \delta \ln \beta$, and the assumed V_p/V_s ratios used in the inversion seem to be very limited.

2. FORWARD MODELING

Our modeling region includes the eastern margin of the Tibetan plateau and the adjacent Sichuan basin (Figure 1a). We choose this region because it contains two different types of tectonic units, plateau and basin, which have very distinct crustal and uppermost mantle structures (Yao et al., 2008; Yang et al.,

Fig. 1

2012; Chen et al., 2014). We set up a total of 97 virtual stations/sources in the region, which have the same geographic distribution as the broadband stations of the CEArray and IRIS/PASSCAL arrays used in a previous tomographic study (Yang et al. 2012) (Figure 1b). The dense ray coverage allows to conduct both 1-D and 3-D inversions to investigate the scaling effects.

2.1. Rayleigh wave phase velocities of different 1-D density models

To probe how different density models affect Rayleigh wave phase velocity dispersion, we select the grid point A (105°E, 31°N) (red star in Figure 1a) in the Sichuan basin with a thick sediment layer (~8 km) (Laske et al., 2013) as an example. Based on the shear velocity profile produced by Y12 (abbreviation for Yang et al., 2012), we generate four density profiles with four empirical scaling models mentioned above (Figure 2a). The four empirical scaling relationships were built based on rock samples with different original depths, each scaling is thus expected to apply to certain depth range better than the others. For example, the Birch scaling model is more appropriate for deeper and denser rocks while the Nafe-Drake model is more reliable for sedimentary rocks (Maceira & Ammon, 2009). The two models exhibit a ~18% difference at shallow depths, but are consistent below the sediment layer. The Gardner's rule is established for studying shallow structure (eg: Giancarlo, 2010; Cauchie & Saccorotti, 2013), therefore it matches the Nafe-Drake model at shallow depths but a ~12% lower density in the upper mantle than the Birch predicts. Compared to the Nafe-Drake density model, the CM model has a much lower density (up to ~25%) at shallow depths and a higher density (~13%) in the middle to lower crust and the upper mantle. Thus, it seems that the Nafe-Drake model gives better density values over the entire depth range for the point A inside the Sichua basin.

Fig. 2

We adopt the analytical equations of the Thomson-Haskell method (Aki & Richards, 2002) to calculate Rayleigh wave dispersion curves in the period range of 10-50 s using the above four density models (Figure 2b). The four models have the same 1-D shear and compressional wavespeed ($\alpha=1.732\beta$) profiles extracted from Y12. As mentioned above, the Nafe-Drake density model seems to fit the density structure beneath the Sichuan basin in the entire crustal range better than the other three models, we thus

use its dispersion curve as the reference to make comparisons. Overall, the calculated phase velocities from the Birch and CM models are higher than those of the Nafe-Drake model over the entire period range of 10-50 s, while the Gardner phase velocities are generally smaller than the Nafe-Drake predictions (Figure 2b). The corresponding percentage difference is $\sim +1.5$ - 5.5% , $\sim +0.5$ - 3.0% , and ~ -0.5 - 2.1% for the CM, Birch, and Gardner models, respectively. Interestingly, although the Birch and Nafe-Drake models have nearly the same density structure at depths >5 km below the surface, the corresponding phase velocities of the two models are pretty large, especially at the short-period range (Figure 2c). On the other hand, the short period phase velocities between the Gardner and Nafe-Drake models differ by $<1\%$ despite the large density difference between the two models in depth greater than 5 km (Figure 2c). Based on Figure 2, we conclude that there are substantial differences in the phase velocities calculated from the four density models, and shallow density structure appears to have significant influence on the short period phase velocities.

2.2. The depth sensitivity kernels of phase velocity to Vs, Vp, and density

To better understand how shallow density structure affects phase velocities, we represent the “partial derivatives” equation of Rayleigh wave phase velocity, which is shown as a function of the two Lamé constants in Aki and Richards (2002), in terms of P-, S-wave velocities and density:

$$\left(\frac{dc}{c}\right)_w = \frac{1}{2k^2 U c I_1} \left\{ \int_0^\infty r a \left(k r_1 + \frac{dr_2}{dz} \right)^2 da dz + \int_0^\infty r b \left[\left(k r_2 - \frac{dr_1}{dz} \right)^2 - 4 k r_1 \frac{dr_2}{dz} \right] db dz + \int_0^\infty \left[\frac{1}{2} a^2 \left(k r_1 + \frac{dr_2}{dz} \right)^2 + \frac{1}{2} b^2 \left(k r_2 - \frac{dr_1}{dz} \right)^2 - \frac{1}{2} w^2 (r_1^2 + r_2^2)^2 - 2 k r_1 \frac{dr_2}{dz} b^2 \right] dr dz \right\}, \quad (2)$$

where ω denotes the angular frequency, k the corresponding wavenumber, c the Rayleigh wave phase velocity, U the group velocity, (r_1, r_2, r_3, r_4) the motion-stress vector, and I_1 the integral term that is defined as:

$$I_1 = \frac{1}{2} \int_0^\infty r (r_1^2 + r_2^2) dz. \quad (3)$$

We use the Nafe-Drake density model and the velocity profile of Y12 at grid point A to compute the

“partial derivatives”, which are the sensitivity kernels of phase velocity to V_s , V_p , and density (Figure 3). At shallow crustal depths (< 10 km for 10 s Rayleigh wave and < 20 km for 20 s) an increase in density leads to a decrease in phase velocity (Figures 3a and 3b). On the contrary, at deeper depths (~10-30 km for the 10 s and ~20-60 km for 20 s Rayleigh waves) an increase in density results in an increase in phase velocity (Figures 3a and 3b). Meanwhile, the phase velocity of 10-s Rayleigh waves appears to be more sensitive to density perturbations than that of the 20-s wave. The former seems to be more sensitive to shallow density while the latter has a broader depth sensitivity range of 0 - 60 km. Roughly speaking, the 10-s Rayleigh wave is about twice as sensitive as the 20-s Rayleigh wave to shallow density perturbations, which indicates that shallow density structures have more profound effects on short period Rayleigh wave phase velocities. We further vary $\delta\rho/\rho$ from 0 to -30% and use equation (2) to compute the corresponding $\delta c/c$ for the periods of 10 and 20 s (Figure 4), which are found to be consistent with the phase velocity differences between the CM and Nafe-Drake models (Figure 2b). Furthermore, the S-wave sensitivity kernels appear to peak at greater depths than the corresponding density kernels (Figure 3). This difference in peak depth gives rise to the tradeoff between the shallow density structure and the shear velocity in the mid-to-lower crustal depths.

3. 1-D AND 3-D INVERSIONS WITH DIFFERENT INITIAL DENSITY MODELS

In the previous section we illustrated that shallow density structure has significant effects on Rayleigh wave phase velocities. In this section we further quantify its effects in 1-D and 3-D shear velocity inversion used in the traditional and adjoint surface wave tomography.

In the 1-D shear wavespeed inversion, we use the Monte-Carlo method (Shapiro & Ritzwoller, 2002) to search for suitable models that fit the phase velocity dispersion curve of a target model. In the 3-D shear wavespeed inversion, we apply an iterative adjoint tomography method developed by Chen et al. (2014) to minimize frequency-dependent traveltimes between the “data” (Green’s functions produced from a target 3-D model computed with the spectral-element method) (Komatitsch et al., 2004; Chen et al., 2014) and the synthetics (synthetic Green’s functions produced from an initial model and

subsequent updated 3-D models).

3.1. 1-D Shear wavespeed inversions

We use the 1-D profile at the grid point A to conduct 1-D inversions of synthetic data to investigate the effect of the absolute density-velocity scaling. To set up the target profile, we take the 1-D Vs profile at grid point A from Y12 and use the Poisson solid assumption and the Nafe-Drake density model to calculate Vp and density, respectively. We then calculate the target phase velocity dispersion curves for this target model, which are the dispersion data for the Monte-Carlo inversion. The synthetic model is hence referred to as the “target model” in this section.

For each test, we start with an initial model that has exactly the same P- and S-wave velocities as the target model, but with a different density structure. In particular, we employ three initial density models, the CM, Gardner and Birch models, and perform a Monte-Carlo 1-D inversion (Shapiro & Ritzwoller, 2002). We first set the model depth range to be 0-150 km and divide it into three segments, the shallow depth (0.0 - 5.0 km), the mid-to-lower crustal depth (5.0 - 40.8 km), and the upper mantle depth (40.8 – 150.0 km). Then in each segment we parameterize the shear wavespeed as a function of depth using five cubic B-splines. During the Monte-Carlo inversion we perturb the five coefficients of each B-spline to produce trial models and employ the simulated annealing algorithm (Shapiro & Ritzwoller, 2002) to explore the model space. During the inversion, we also update the P-wave velocity and density by using the relative scalings of $\delta \ln \alpha / \delta \ln \beta = 0.5$ and $\delta \ln \rho / \delta \ln \beta = 0.33$ (Panning & Romanowicz, 2006). We will further discuss the influences of different relative scalings in the following discussion section. Vertical model smoothness is taken into account in the objective function, and the recovered model is obtained by averaging all accepted models whose objective functions are below certain threshold.

For the inversion with the CM initial model (Figure 5a), the recovered S-velocity model is lower by ~4.5% in the mid-to-lower crustal depths (10 - 40 km) when compared to the target model. This large difference between the recovered and target shear-wave velocity are largely due to the difference in density at shallow depths between the CM and Nafe-Drake model (Figure 2a). The inversion using the

Fig. 5

Birch model shows a similar but a less significant shear velocity reduction in the mid-lower crust (Figure 5c), which corresponds to the smaller discrepancy between the Nafe-Drake and Birch density models in the shallow depths (Figure 2a). Meanwhile, the recovered shear-wave velocity with the Gardner is slightly higher only in the lower crust level (Figure 5b), which is likely caused by its lower density in the lower crust (Figure 2a) based on the positive density sensitivity kernel in this depth range (Figure 3a). For validation, we conduct another inversion which uses the Nafe-Drake density model, the same as the target model, plus an initial V_s model (orange line in Figure 5d) that is 5% lower than Y12 above the Moho. The resulting shear velocity profile (black line in Figure 5d) shows a good match to the target model with a very subtle discrepancy, probably caused by parameterization and/or regularization in the inversion.

Combined with the results shown in the previous sections, it is clear that density structures in the upper five kilometers have significant tradeoff with shear-wave velocity structure in the mid-to-lower crust in the above 1-D Monte-Carlo inversion. In particular, if the initial model has a density $\sim 25\%$ lower than the target model in the top five kilometers, the recovered shear-wave velocity can be $\sim 4.5\%$ higher than the true model in the mid-to-lower crustal depths. Since shear wavespeed anomalies at this level are significant in tomography, cautions have to be taken in choosing a proper density model at shallow depths when implementing the inversion. We will further confirm this argument with a 3-D inversion example based on adjoint tomography in the following section.

3.2. 3-D Shear wavespeed inversions

In this application of the ambient noise adjoint tomography (Chen et al., 2014), we first set up a target 3D model for the Sichuan Basin and surrounding regions by taking the shear wavespeeds from Y12 and using the Poisson solid assumption and the Nafe-Drake density model to calculate compressional wavespeeds and density, respectively. We then generate synthetic Green's functions (Chen et al., 2014) between all the station pairs (Figure 1b) for this target model using the spectral-element method (Komatitsch et al., 2004). These synthetic Green's functions are treated as observed

empirical Green's functions (EGFs) derived from ambient noise, and the target model is assumed as a known ground truth model. In the 3D adjoint tomography, we use the CM density model to calculate the initial density model. We only invert for shear wavespeeds β and scale $\delta\rho$ and $\delta\alpha$ to $\delta\beta$ based on the same scaling relations described in the 1-D inversion.

The synthetic Green's functions (SGFs) are also calculated for the starting model and the subsequent updated models. The frequency-dependent traveltime misfits between the EGFs and the SGFs are measured at four period bands: 10 - 20 s, 15 - 30 s, 20 - 40 s, and 25 - 50 s. The overall misfit is iteratively minimized based on a preconditioned conjugate gradient method while the 3-D model gets refined using 3-D finite-frequency kernels (Chen et al., 2014). The inverted 3-D model is obtained after 3 iterations (Figures 6a–c).

Fig. 6

Density differences between the initial model and the target model averaged over the upper five kilometers (Figure 6d) shows an amplitude of up to 25% inside the Sichuan basin. We also compute phase velocities at 10 s and 20 s periods using the *geopsy* software package for both the initial 3-D model and the target model. The largest phase-velocity differences (up to ~5%) between these two models (Figures 6e and 6f) are located within the Sichuan basin, which are correlated with the shallow density deficit in the initial model (Figure 6d). In order to match the observed phase velocities, the shear wavespeeds in the mid-to-lower crust have to be reduced by ~4.5% (Figure 6b). The 1-D shear wavespeed profile extracted from the inverted model at grid point A is also consistent with the 1-D inversion result (Figure 5a). Therefore, the 3-D inversion results further confirm that large model errors in density at shallow depths can result in significant artificial shear wavespeed anomalies at mid-to-lower crustal depths.

4. DISCUSSION AND CONCLUSIONS

In the above 1-D and 3-D inversions, we have assumed a Poisson solid ($V_p=1.732V_s$) and a fixed perturbation scaling ($\delta\ln\alpha/\delta\ln\beta=0.5$ and $\delta\ln\rho/\delta\ln\beta=0.33$) in constructing the initial P-wave model and in model updating, respectively. It is well known that sediments could have very high V_p/V_s ratio, very

different from that of a Poisson solid. To investigate the absolute scaling between V_p and V_s , we create an S-wave velocity model by adding a 300 m low velocity layer with $V_s=0.5$ km/s on the top of the Y12 grid A profile. We then generate two sediment P-wave velocity models by using a V_p/V_s ratio of 1.732 and 4.0, respectively. Everything below the sediment is exactly the same (Figure 7a). We compute the phase velocities of these two models, which differs by less than 0.1% at periods shorter than ~20s (Figures 7b and 7c).

Fig. 7

In order to further evaluate how the assumed V_p/V_s scaling in the whole crust affects the 1-D inversion results, we generate another crustal P-wave model using the Nafe-Drake V_p-V_s scaling relationship (Brocher, 2005) (red line in Figure 7d). We notice that the Nafe-Drake P-wave velocity model is actually not so different from the Poisson solid model. Therefore their corresponding phase velocities are very similar, with a discrepancy less than 1% over the entire periods of 10-50 s (Figures 7e and 7f). Therefore we think the Poisson's solid assumption used in building the initial model is valid, and the absolute V_p-V_s scaling seems to have very small influence on the final models.

In the above 1-D and 3-D inversions, we employ a relative scaling relationship of $\delta \ln \alpha / \delta \ln \beta = 0.5$ and $\delta \ln \rho / \delta \ln \beta = 0.33$ (Panning & Romanowicz, 2006) in updating the density and P-wave velocity for both crustal and mantle layers. We further conduct the 1-D inversions with other three relative scalings: (1) $\delta \ln \alpha / \delta \ln \beta = 0.0$, $\delta \ln \rho / \delta \ln \beta = 0.0$; (2) $\delta \ln \alpha / \delta \ln \beta = 1.0$, $\delta \ln \rho / \delta \ln \beta = 0.0$; (3) $\delta \ln \alpha / \delta \ln \beta = 1.0$, $\delta \ln \rho / \delta \ln \beta = 1.0$. (1) means that both V_p and density are fixed, and (2) implies for a fixed V_p/V_s ratio and a fixed density (Yang et al., 2012), while (3) indicates for a fixed V_p/V_s ratio and a fixed ρ/V_s ratio, such as the CM model, during the inversions. The S-wave velocity and density are kept the same in all the inversions. The final S-wave velocity models with the three perturbation scalings are shown in Figures 7h, 7i and 7j, respectively. For comparison we also show the S-wave velocity model obtained in the previous 1-D inversion in Figure 7g. It seems that the final V_s model does depend slightly on how V_p is updated during the inversion, but the model difference is generally less than 1%. Therefore, the potential bias on the inversion caused by improper perturbation scalings is almost negligible.

By conducting both 1-D and 3-D shear wave speed inversions, we demonstrate that density structures

at shallow depths have great effects on the short-period surface wave dispersion curves and there is a significant tradeoff between density structures at shallow depths and shear wavespeed at mid/lower crust in these inversions. Improper scaling of density from shear velocity in the shallow depths can lead to spurious shear velocity anomalies in the mid/lower crust in surface wave inversion. Specifically, for sedimentary basins, such as the the Sichuan basin, where sediment deposits can reach several kilometers, ~25% deviation in density in the top sedimentary layer can introduce a 5% error in shear velocity in the mid/lower crust. The artificial anomalies induced by density errors in the top layers can mislead interpretations of tomographic results. Therefore, it is very important to adopt a reasonable density model during the 1-D and 3-D inversions to obtain a robust shear wavespeed model in surface wave tomography.

Recent studies (e.g. Lin et al., 2012) on the joint inversion of H/V ratio and phase velocity suggest that addition of ellipticity data can help constrain the density in the shallow depth of the crust. Meanwhile gravity data are always helpful in providing additional constraints on density structure. Maceira & Ammon (2009) proposed a joint inversion algorithm of surface wave and gravity data, which integrates the data with empirical relations between velocity and density. Furthermore, although suffering from the intrinsic nonuniqueness, the inversion of gravity data itself can yield smooth and geologically reasonable density model (Bear et al., 1994; Li & Oldenburg, 1998), which can be used as the initial density model to improve the reliability of the surface wave tomography.

ACKNOWLEDGEMENTS

This research was supported by NSF grants 1345096 and ACI-1053575 through XSEDE research allocation. We thank the operators for the various networks that contribute the seismic data (Chinese Provincial networks, FDSN, PASSCAL experiments), as well as the Rice Research Computing Support Group. We also thank the editor, Dr. Jörg Renner, and two anonymous reviewers for their constructive comments and suggestions, which significantly improved the quality of this paper. The open source spectral-element software package SPECFEM3D is freely available for download via the Computational

Infrastructure for Geodynamics (CIG; geodynamics.org). The *geopsy* software package (geopsy.org) is used to calculate the dispersion curve. This is contribution xx from the ARC Centre of Excellence for Core to Crust Fluid Systems (<http://www.ccfs.mq.edu.au>) and xxx in the GEMOC Key Centre (<http://www.gemoc.mq.edu.au>).

REFERENCES

- Aki, K., & Richards, P. G., 2002. *Quantitative seismology*, 2nd edn, University Science Books, Sausalito, CA, USA.
- Bear, G., Al-Shukri, H. & Rudman, A., 1995. Linear inversion of gravity data for 3-D density distributions, *Geophysics*, 60, 1354–1364.
- Bensen, G.D., Ritzwoller, M.H. & Yang, Y., 2009, A 3-D shear velocity model of the crust and uppermost mantle beneath the United States from ambient seismic noise, *Geophys. J. Int.*, 177, 1177–1196.
- Birch, F., 1961. The velocity of compressional waves in rocks to 10 kilobars, part 2, *J. geophys. Res.*, 66, 2199–2224.
- Brocher, T.M., 2005. Empirical relations between elastic wavespeeds and density in the Earth's crust, *Bull. seism. Soc. Am.*, 95(6), 2081– 2092.
- Campillo, M. & Paul, A., 2003. Long-range correlations in the diffuse seismic coda. *Science*, 299, 547–549, doi:10.1126/science.1078551.
- Cauchie, L., & Saccorotti, G., 2013. Probabilistic inversion of Rayleigh-wave dispersion data: an application to Mt. Etna, Italy, *J. seism.*, 17(2), 335–346.
- Chen, M., Huang, H., Yao, H., van der Hilst, R. & Niu, F., 2014. Low wave speed zones in the crust beneath SE Tibet revealed by ambient noise adjoint tomography, *Geophys. Res. Lett.*, 41(2), 334–340.
- Christensen, N.I. & Mooney, W.D., 1995. Seismic velocity structure and composition of the continental crust: a global view, *J. geophys. Res.*, 100, 9761–9788.

- Gardner, G.H.F., Gardner, L.W. & Gregory, A.R., 1974. Formation velocity and density – the diagnostic basis for stratigraphic traps, *Geophysics*, 39, 770–780.
- Giancarlo, D. M., 2010. Insights on surface wave dispersion and HVSR: Joint analysis via Pareto optimality, *J. Appl. Geophys.*, 72(2), 129–140.
- Guo, Z., Chen, Y. J., Ning, J., Feng, Y., Grand, S. P., Niu, F., Kawakatsu, H., Tanaka, S., Obayashi, M., & Ni, J., 2015. High resolution 3-D crustal structure beneath NE China from joint inversion of ambient noise and receiver functions using NECESSArray data, *Earth planet. Sci. Lett.*, 416, 1–11.
- Huang, Y. C., Yao, H., Wu, F. T., Liang, W. T., Huang, B. S., Lin, C. H., & Wen, K. L., 2014. Crustal and upper mantle S-wave velocity structures across the Taiwan Strait from ambient seismic noise and teleseismic Rayleigh wave analyses, *J. Asian Earth Sci.*, 81, 38–52.
- Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.*, 149, 390–412.
- Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation—II. 3-D models, oceans, rotation, and self-gravitation, *Geophys. J. Int.*, 150, 303–318.
- Komatitsch, D., Liu, Q., Tromp, J., Suess, P., Stidham, C. & Shaw, J.H., 2004. Simulations of strong ground motion in the Los Angeles Basin based upon the spectral-element method, *Bull. seism. Soc. Am.*, 94, 187–206.
- Laske, G., Masters, G., Ma, Z. & Pasayanos, M., 2013. Update on CRUST1.0: a 1-degree global model of Earth's crust, *Geophys. Res. Abstr.*, 15, Abstract EGU2013-2658.
- Li, H., Li, S., Song, X.D., Gong, M., Li, X. & Jia, J., 2012. Crustal and uppermost mantle velocity structure beneath northwestern China from seismic ambient noise tomography, *Geophys. J. Int.*, 188, 131–143.
- Li, Y. & Oldenburg, D.W., 1998. 3-D inversion of gravity data, *Geophysics*, 63(1), 109–119.
- Lin, F.-C., Ritzwoller, M.H., Townend, J., Bannister, S. & Savage, M.K., 2007. Ambient noise Rayleigh wave tomography of New Zealand, *Geophys. J. Int.*, 170(2), 649–666. doi:10.1111/j.1365–246X.2007.03414.x

- Lin, F.C., Schmandt, B. & Tsai, V.C., 2012a. Joint inversion of Rayleigh wave phase velocity and ellipticity using USArray: constraining velocity and density structure in the upper crust, *Geophys. Res. Lett.*, 39, L12303, doi:10.1029/2012GL052196.
- Ludwig, W. J., Nafe, J. E., & Drake, C. L., 1970. Seismic refraction, in *The Sea*, A. E. Maxwell (Editor), Vol. 4, Wiley-Interscience, New York, 53-84.
- Maceira, M. & Ammon, C.J., 2009. Joint inversion of surface wave velocity and gravity observations and its application to central Asian basins shear velocity structure, *J. geophys. Res.*, 114, B02314, doi:10.1029/2007JB005157.
- Panning, M. & Romanowicz, B., 2006. A three dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.*, 167, doi:10.1111/j.1365-246X.2006.03100.x.
- Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005. High resolution surface wave tomography from ambient seismic noise, *Science*, 307, 1615–1618.
- Shapiro, N.M. & Ritzwoller, M.H., 2002. Monte-Carlo inversion for a global shear velocity model of the crust and upper mantle, *Geophys. J. Int.*, 151, 88–105.
- Tromp, J., Komatitsch, D. & Liu, Q., 2008. Spectral-element and adjoint methods in seismology, *Commun. Comput. Phys.*, 3(1), 1–32.
- Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, 160, 195–216.
- Yang, Y., Ritzwoller, M.H., Levshin, A.L. & Shapiro, N.M., 2007. Ambient noise Rayleigh wave tomography across Europe, *Geophys. J. Int.*, 168, 259–274.
- Yang, Y., Ritzwoller, M.H., Zheng, Y., Shen, W., Levshin, A. & Xie, Z., 2012. A synoptic view of the distribution and connectivity of the mid-crustal low velocity zone beneath Tibet, *J. geophys. Res.*, 117, B04303, doi:10.1029/2011JB008810.
- Yao, H., van der Hilst, R.D. & de Hoop, M.V., 2006. Surface-wave tomography in SE Tibet from ambient seismic noise and two-station analysis—I. Phase velocity maps, *Geophys. J. Int.*, 166, 732–744, doi: 10.1111/j.1365-246X.2006.03028.x.

Yao, H., Beghein, C. & van der Hilst, R.D., 2008, Surface wave array tomography in SE Tibet from ambient seismic noise and two-station analysis—II. crustal and upper-mantle structure, *Geophys. J. Int.*, 173, 205–219.

Figure Captions

Figure 1. (a) The topography of the study area. Gray lines delineate the boundaries of tectonic blocks. Red star marks the grid point A at 105° E and 31° N, which is used to illustrate the density effect on Rayleigh wave phase velocity and shear wavespeed inversion in 1-D earth models. (b) Station and raypath coverage for the 3D wavespeed inversion. Yellow triangles denote the 97 stations used in 3-D shear wavespeed inversion. Black lines represent interstation great-circle raypaths.

Figure 2. (a) The magenta dashed line denotes the shear wavespeed profile at grid point A in the Sichuan Basin obtained from Yang et al. (2012). The solid lines of different colors indicate the 1-D density depth profiles calculated from shear wavespeeds using the CM, Gardner, Birch and Nafe-Drake relations. (b) Rayleigh wave phase velocities in the period range from 10 s to 50 s computed with the four velocity-density models shown in (a). (c) Differences in percentage among the phase velocities shown in (b), all plotted relative to the one using the Nafe-Drake density models.

Figure 3. (a) Depth profiles of 1-D phase velocity sensitivity kernels at grid point A, calculated based on eq. (2) for the 10-s Rayleigh wave, with respect to $\delta\alpha$ (cyan), $\delta\beta$ (green), and $\delta\rho$ (red), respectively. Gray dashed line indicates zero sensitivity at all depths is used for reference. (b) Same with (a) but for the 20-s Rayleigh wave.

Figure 4. Phase velocity perturbations ($\delta C/C$) are shown as a function of density perturbations ($\delta\rho/\rho$) averaged over the upper five kilometers at grid point A, for 10-s (black line) and 20-s (red line) Rayleigh waves, respectively.

Figure 5. (a)-(c) Inverted shear-wave velocity depth profiles of the 1-D Monte-Carlo inversion at grid point A with different initial density models: (a) CM, (b) Gardner; (c) Birch. Blue line represents the target shear velocity model (Y12). Gray lines indicate all acceptable models of the 1-D Monte-Carlo inversion, and the black line the finally inverted model obtained by averaging all the acceptable models. The dashed red line in (a) shows the final S velocity model at grid point A derived from the 3-D adjoint tomography. (d) Same as (a)-(c) but for a validation test with a Nafe-Drake initial density model and an initial Vs model (orange line) which is 5% slower than Y12 at the depths above the Moho.

Figure 6. Maps showing the difference in percentage between the recovered and target models at 5 km (a), 20 km (b), and 50 km depths (c), obtained by the 3-D iterative adjoint tomography. The contour interval is 1%. (d) Color contour map showing the average density difference of the upper five kilometers between the initial CM and the target Nafe-Drake models. The contour interval is 8%. (e) and (f) show the phase velocity differences between the initial CM and target Nafe-Drake model for 10-s and 20-s Rayleigh waves, respectively. The contour interval here is 1%.

Figure 7. Influence of the absolute Vp-Vs scaling and the perturbation scalings ($\delta \ln \alpha / \delta \ln \beta$, $\delta \ln \rho / \delta \ln \beta$) on the inversion. (a) The sediment P-wave velocity model with a Vp/Vs ratio of 4.0 (blue solid line) is shown together with the Poisson solid model with a Vp/Vs ratio of 1.732 (red dashed line). The two model share with similar S-wave velocity (black dotted line) and density structures. (b) and (c) show the calculated phase velocities of the two models in the period range of 10 to 50s, and their difference in percentage. (d) The Nafe-Drake P-wave velocity model (blue solid line) is shown with the Poisson solid model (red dashed line) for comparison. The phase velocities of the two models and their difference are shown in (e) and (f), respectively. The S-wave velocity depth profiles of the 1-D Monte-Carlo inversions at grid point A with different perturbation scalings: (g) $\delta \ln \alpha / \delta \ln \beta = 0.5$, $\delta \ln \rho / \delta \ln \beta = 0.33$; (h) $\delta \ln \alpha / \delta \ln \beta = 0.0$, $\delta \ln \rho / \delta \ln \beta = 0.0$; (i) $\delta \ln \alpha / \delta \ln \beta = 1.0$, $\delta \ln \rho / \delta \ln \beta = 0.0$; (j) $\delta \ln \alpha / \delta \ln \beta = 1.0$, $\delta \ln \rho / \delta \ln \beta = 1.0$. The

models are plotted similarly as Figure 5.

Figures

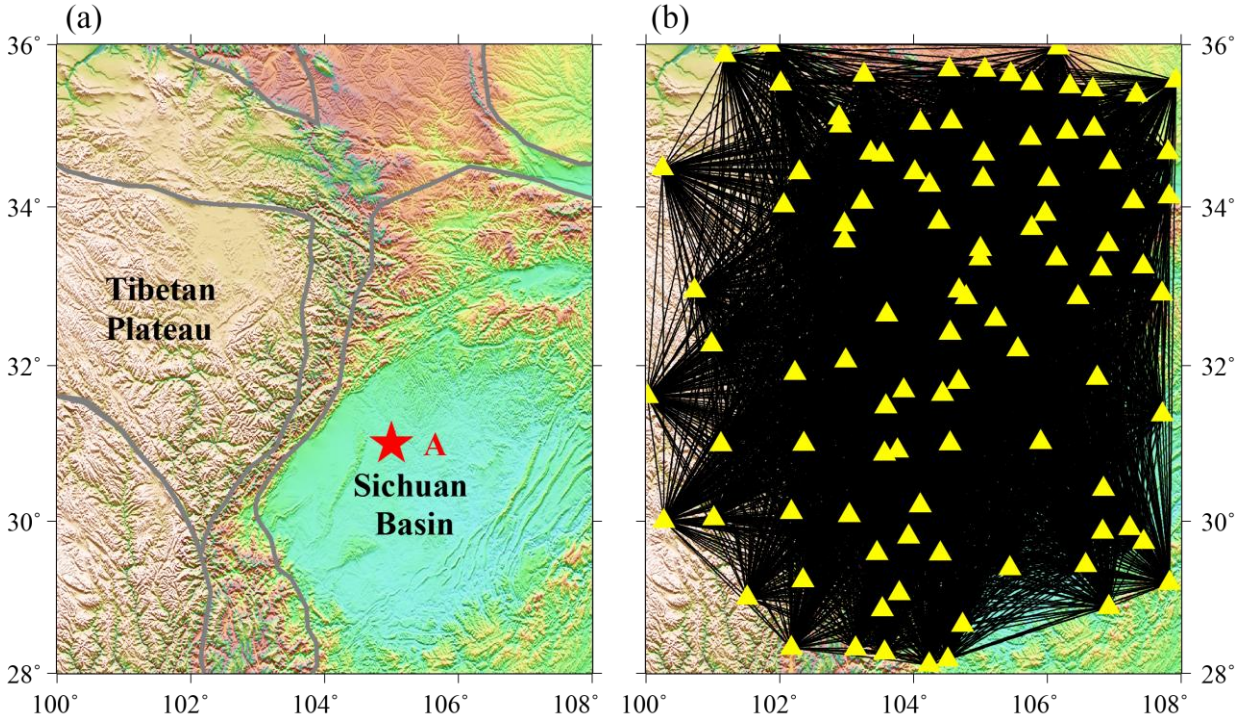


Figure 1. (a) The topography of the study area. Gray lines delineate the boundaries of tectonic blocks. Red star marks the grid point A at 105° E and 31° N, which is used to illustrate the density effect on Rayleigh wave phase velocity and shear wavespeed inversion in 1-D earth models. (b) Station and raypath coverage for the 3D wavespeed inversion. Yellow triangles denote the 97 stations used in 3-D shear wavespeed inversion. Black lines represent interstation great-circle raypaths.

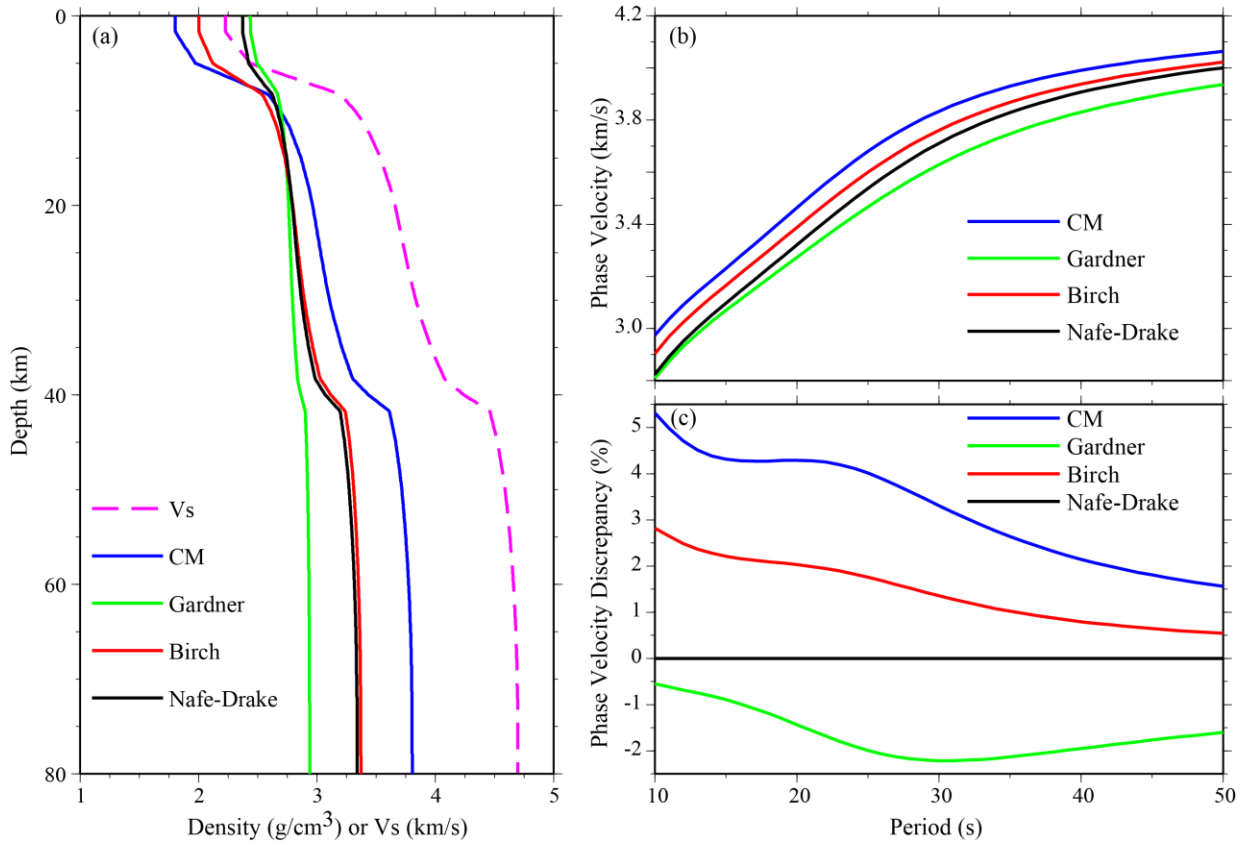


Figure 2. (a) The magenta dashed line denotes the shear wavespeed profile at grid point A in the Sichuan Basin obtained from Yang et al. (2012). The solid lines of different colors indicate the 1-D density depth profiles calculated from shear wavepeeds using the CM, Gardner, Birch and Nafe-Drake relations. (b) Rayleigh wave phase velocities in the period range from 10 s to 50 s computed with the four velocity-density models shown in (a). (c) Differences in percentage among the phase velocities shown in (a), all plotted relative to the one using the Nafe-Drake density models.

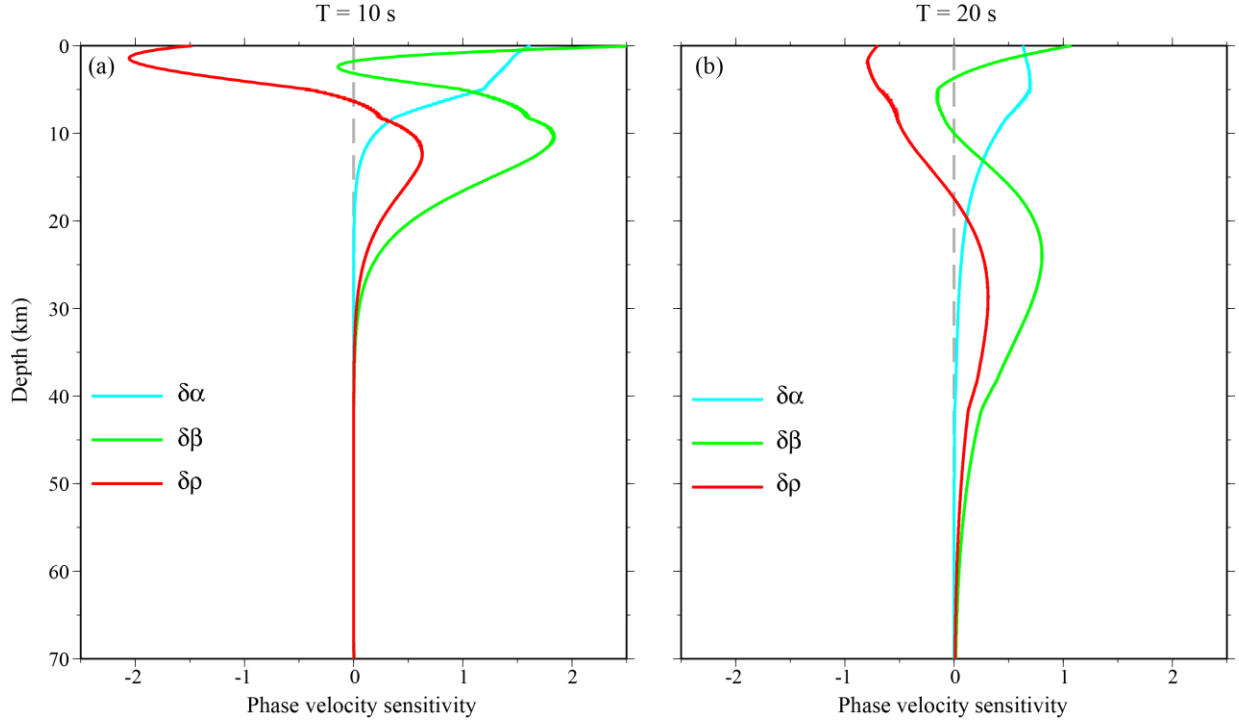


Figure 3. (a) Depth profiles of 1-D phase velocity sensitivity kernels at grid point A, calculated based on eq. (2) for the 10-s Rayleigh wave, with respect to $\delta\alpha$ (cyan), $\delta\beta$ (green), and $\delta\rho$ (red), respectively. Gray dashed line indicates zero sensitivity at all depths is used for reference. (b) Same with (a) but for the 20-s Rayleigh wave.

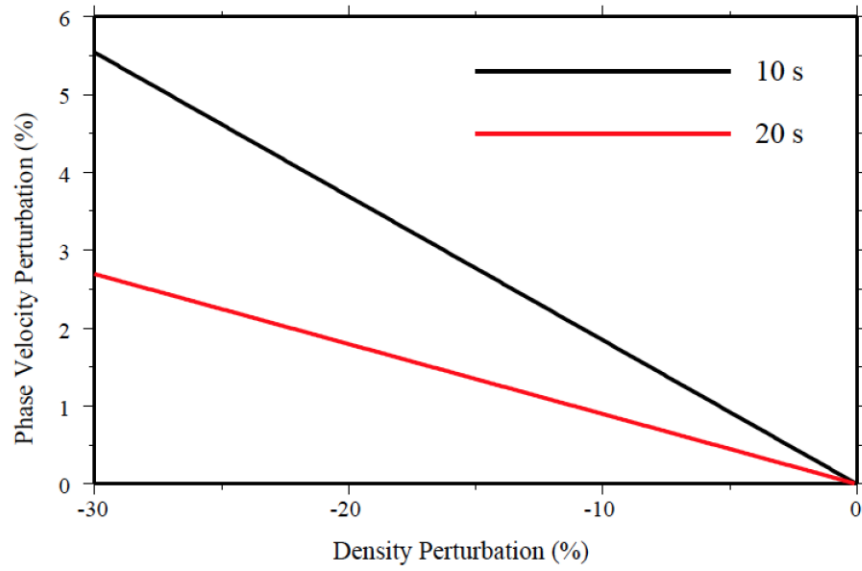


Figure 4. Phase velocity perturbations ($\delta C/C$) are shown as a function of density perturbations ($\delta\rho/\rho$) averaged over the upper five kilometers at grid point A, for 10-s (black line) and 20-s (red line) Rayleigh waves, respectively.

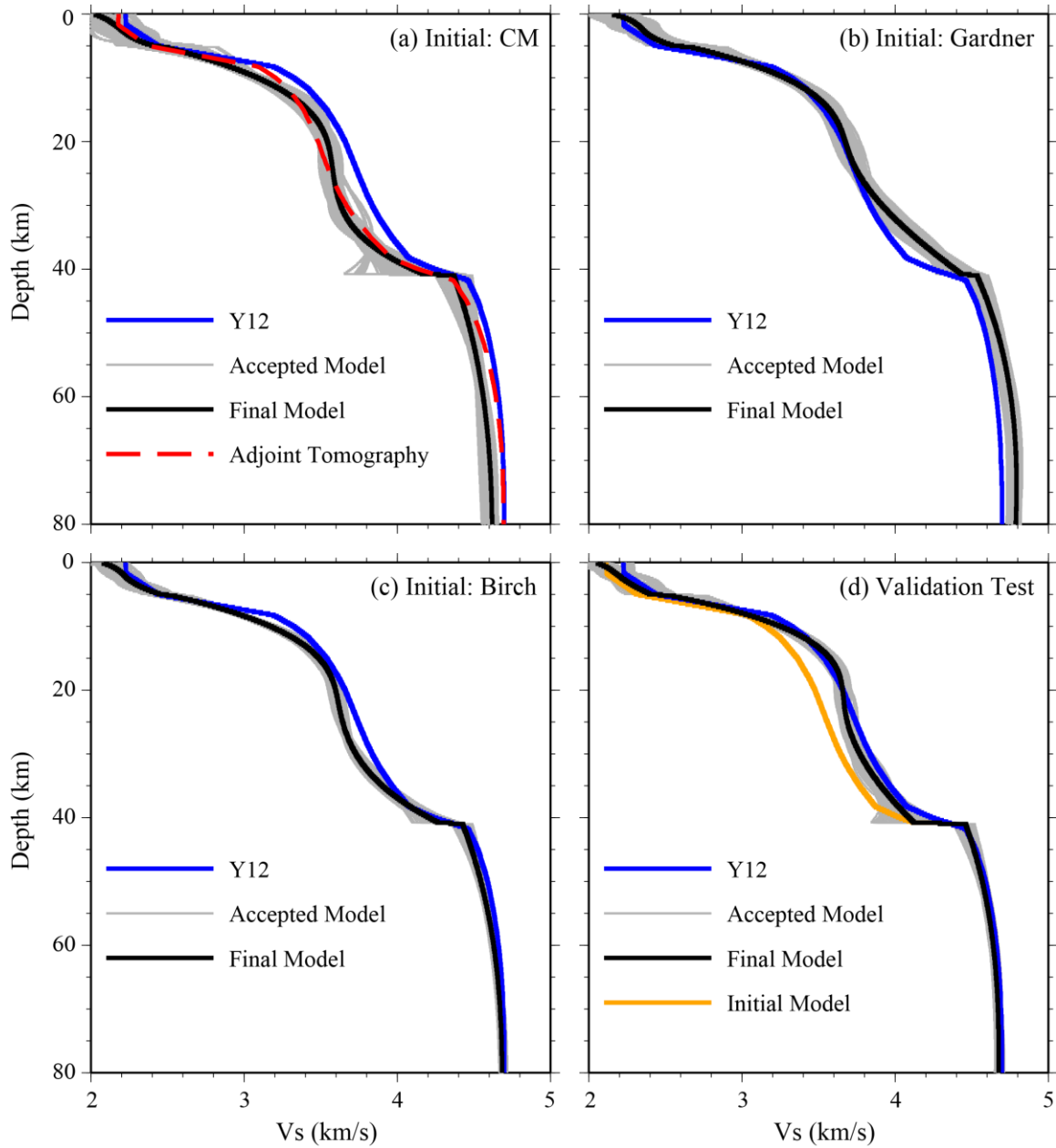


Figure 5. (a)-(c) Inverted shear-wave velocity depth profiles of the 1-D Monte-Carlo inversion at grid point A with different initial density models: (a) CM, (b) Gardner; (c) Birch. Blue line represents the target shear velocity model (Y12). Gray lines indicate all acceptable models of the 1-D Monte-Carlo inversion, and the black line the finally inverted model obtained by averaging all the acceptable models. The dashed red line in (a) shows the final S velocity model at grid point A derived from the 3-D adjoint tomography. (d) Same as (a)-(c) but for a validation test with a Nafe-Drake initial density model and an initial Vs model (orange line) which is 5% slower than Y12 at the depths above the Moho.

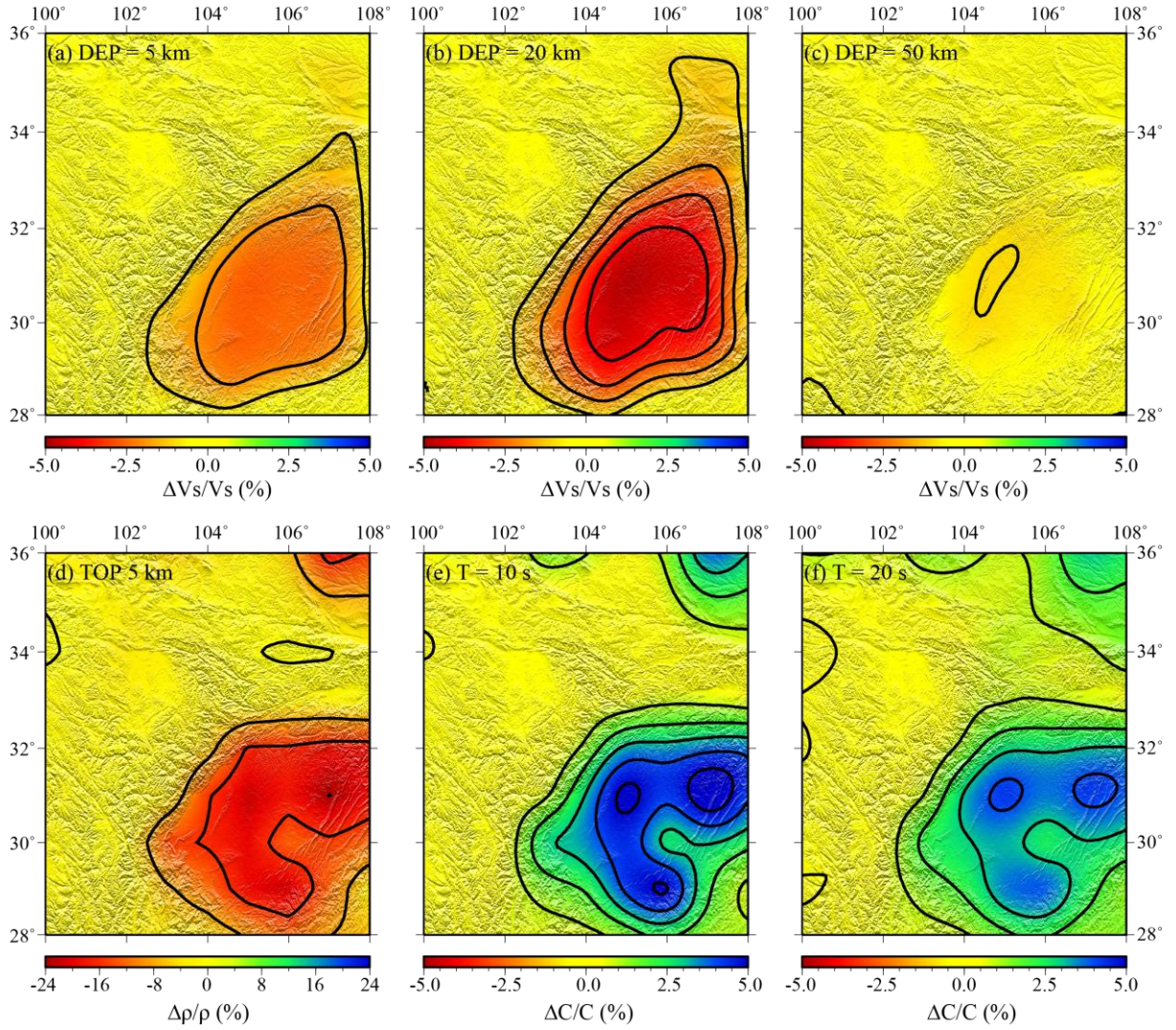


Figure 6. Maps showing the difference in percentage between the recovered and target models at 5 km (a), 20 km (b), and 50 km depths (c), obtained by the 3-D iterative adjoint tomography. The contour interval is 1%. (d) Color contour map showing the average density difference of the upper five kilometers between the initial CM and the target Nafe-Drake models. The contour interval is 8%. (e) and (f) show the phase velocity differences between the initial CM and target Nafe-Drake model for 10-s and 20-s Rayleigh waves, respectively. The contour interval here is 1%.

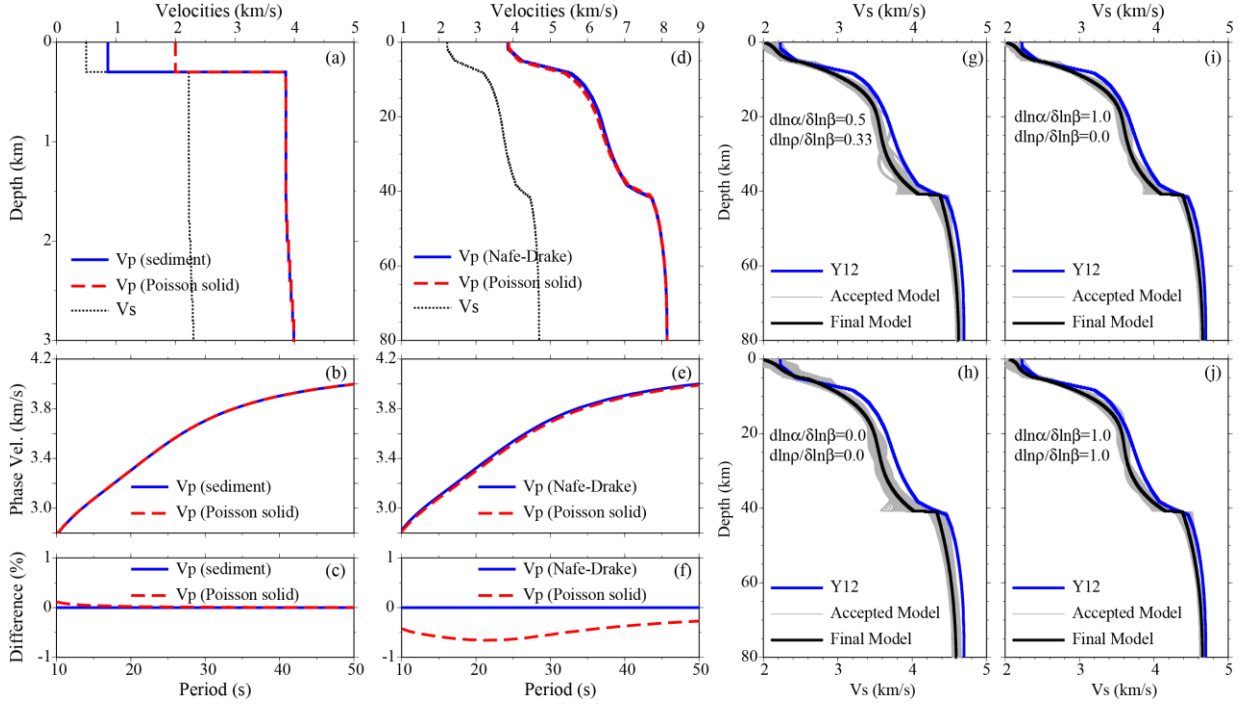


Figure 7. Influence of the absolute Vp-Vs scaling and the perturbation scalings ($\delta \ln \alpha / \delta \ln \beta$, $\delta \ln \rho / \delta \ln \beta$) on the inversion. (a) The sediment P-wave velocity model with a Vp/Vs ratio of 4.0 (blue solid line) is shown together with the Poisson solid model with a Vp/Vs ratio of 1.732 (red dashed line). The two models share the same S-wave velocity (black dotted line) and density structures. (b) and (c) show the calculated phase velocities of the two models in the period range of 10 to 50s, and their difference in percentage. (d) The Nafe-Drake P-wave velocity model (blue solid line) is shown with the Poisson solid model (red dashed line) for comparison. The phase velocities of the two models and their difference are shown in (e) and (f), respectively. The S-wave velocity depth profiles of the 1-D Monte-Carlo inversions at grid point A with different perturbation scalings: (g) $\delta \ln \alpha / \delta \ln \beta = 0.5$, $\delta \ln \rho / \delta \ln \beta = 0.33$; (h) $\delta \ln \alpha / \delta \ln \beta = 0.0$, $\delta \ln \rho / \delta \ln \beta = 0.0$; (i) $\delta \ln \alpha / \delta \ln \beta = 1.0$, $\delta \ln \rho / \delta \ln \beta = 0.0$; (j) $\delta \ln \alpha / \delta \ln \beta = 1.0$, $\delta \ln \rho / \delta \ln \beta = 1.0$. The models are plotted similarly as Figure 5.