

Comparison of 3-D Raytracing and Finite Frequency Tomography

S. K. Suhardja^{1*}, H. L. Gaol¹, A. Abdullah¹, A. D. Nugraha², Z. Zulfakriza²

¹Department of Geophysical Engineering, University of Pertamina,
Jl. Teuku Nyak Arief, Simprug, Kebayoran Lama, Jakarta 12220, Indonesia

²Global Geophysics Research Group, Faculty of Mining and Petroleum Engineering Institute of Technology Bandung,
Jalan Ganesa No. 10, Bandung 40132, Indonesia

*Email: s4ndy104@gmail.com

Submit: 3 July 2019; Revised: 3 July 2019; Accepted: 1 September 2019

Abstract: We performed 3-D seismic tomography using teleseismic arrival time at Southwest Mexico. The Mexican subduction zone results from successive fragmentation events that affected the ancient Farallon plate as various segments of the East Pacific rise approached the paleo-trench off western North America. The complexity in this region is related to two subducting oceanic plates, the Rivera and Cocos plates, that have different ages, compositions, convergence velocities and subduction dip angles. In this study, we compared the 3-D raytracing tomography model with finite frequency tomography model. Final models show the differences in amplitude and pattern between the raytracing and finite frequency. 3D raytracing models produced sharper images of fast velocity structures in the mantle. The deeper slabs are more coherent and show less broadening with depth than using 1D finite frequency kernels. However, although the finite frequency and 3-D ray tracing models show some differences in amplitude and pattern, the overall agreement of the models supports the interpretation of previous study that slab rollback is occurring in South Western Mexico. One possible different interpretation between the raytracing and finite frequency theory results concerns the deep structure of the Rivera slab. The finite frequency models show that the Rivera slab is clearly observable at a depth of about 300km but fades away at greater depths. However, the 3-D ray tracing model shows a clear fast velocity band down to a depth of 400 km and thus our model does not support a slab tear of the Rivera plate above 400 km depth.

Keywords: 3-D Raytracing, Finite Frequency, Teleseismic Tomography, Southwest Mexico, Subduction Zone

1 INTRODUCTION

Seismic tomography is a technique to image 3D seismic velocities in the Earth's subsurface by employing seismic waves generated by earthquakes and/or explosions. Some variations of tomography include: waveform tomography that works by determining a velocity model through synthetic waveform matching, reflection travel time tomography that is popular in industry and works by optimizing velocity and reflector depth to minimize travel time of reflected waves, finite frequency tomography and ray based travel time to-

mography. The last two methods have been used extensively in global seismology.

Travel time tomography assumes the travel time of a wave depends on the rock velocities along the geometric raypath corresponding to the observed wave. Technically, this is true only at infinite frequency. The finite frequency tomography technique works on the hypothesis that the observed seismic waves are finite frequency signals. For finite frequency waves, travel times are sensitive to the velocities in a three-dimensional volume around the geometric raypath. Seismic waves passing through velocity heterogeneities with dimensions smaller than the width of the Fresnel zone undergo significant wave-front healing, which results in reduced travel-time shifts compared to the predictions of ray theory (Hung et al., 2001). The 3-D Frechet travel-time sensitivity kernels (Marquering et al., 1999; Dahlen et al., 2000; Zhao et al., 2000) provide a way to account for wave-front healing and off-ray scattering. Therefore, finite frequency seismic tomography based on 3-D Frechet sensitivity kernels is theoretically more accurate in imaging velocity anomalies (Hung et al., 2001; Baig et al., 2003). This method has been applied to regional (e.g., Hung et al., 2004; Yang et al., 2006) and global tomographic studies (Montelli et al., 2004) with great success. However, the finite frequency Frechet kernels are generally only computed once in a given tomography experiment due to the computation cost and effort to do so. The kernels are usually computed for a velocity model that only varies with depth because, again, the kernel calculations are easier in this case than for a fully three dimensionally varying structure. Thus, although using finite frequency kernels is theoretically more accurate than using geometric rays to calculate the sensitivity of travel time anomalies to velocity perturbations in the subsurface, the kernels themselves are approximations to the true kernels due to the fact that they are sensitive to the velocity model itself.

Although ray theory is an approximation, it does have the advantage that one can calculate ray paths through fully three-dimensional structures. Thus, when structures are complex with large seismic velocity variations, calculating the actual raypaths through a complicated 3D model may provide better tomography images than using finite frequency kernels computed using a simple 1D model. Subduction zones may be regions where the distortion of raypaths

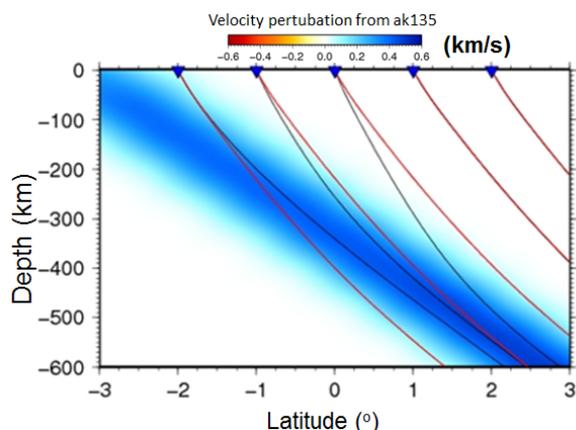


Figure 1. A 2-D synthetic subduction model that has dipping fast velocity 6% faster than the reference ak135 (Kennet et al., 1995) velocity model. The red line are raypaths associated with reference model, the black line are raypaths corresponds to 2-D dipping model that bend into the fast velocity nodes to travel on the shortest time.

is severe due to the high amplitude velocity anomaly within slabs compared to surrounding mantle. Figure 1 shows an example of this. Here I show the raypaths from a distant earthquake to a line of stations computed for a velocity model that varies in depth only as well as the raypaths for a model that has a dipping 6% fast velocity anomaly that could represent a subducting slab. Note how the raypaths are focused into the slab in the 3D model. It should also be noted that if one uses regional waves that interact with the Moho and upper mantle discontinuities, 3D ray tracing is likely necessary to properly locate the paths through the Earth that such waves take.

Yang et al. (2009) presented a P wave finite frequency study of Southwestern Mexico using MARS and CODEX data. We use the same data measured by Yang et al. (2009) and perform teleseismic seismic tomography using 3-D ray tracing to investigate the difference in results between the finite frequency method and the 3-D ray theory method. Montelli et al (2004) performed a comparison between 3-D ray theoretical and finite-frequency travel time on global tomography. Their results show that depending on the depth and size of the anomaly, the amplitudes of the velocity perturbations in the finite-frequency tomography images can be up to 50 percent larger than in the corresponding ray-theory images. We will investigate this for a regional tomography experiment. We are so motivated to do this study because this study recorded many regional earthquakes that can be incorporated into the ray tomography method but not the finite frequency method.

1.1 Tectonic Setting

The current tectonic configuration of the Mexican subduction zone results from successive fragmentation events that affected the ancient Farallon plate as various segments of the East Pacific rise approached the paleo-trench off western North America. When a spreading center encounters a subduction zone, a major change in stress and plate bound-

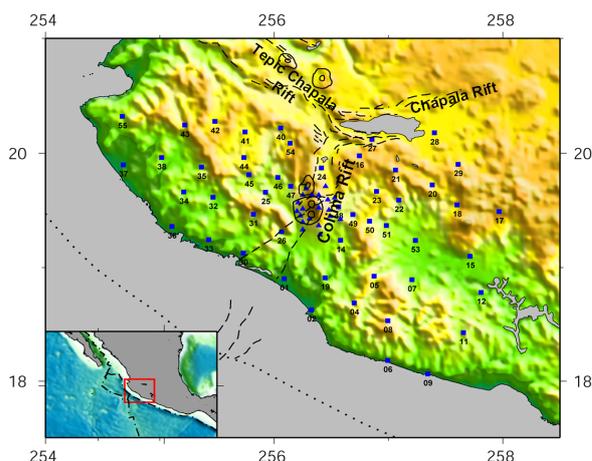


Figure 2. Seismic stations deployed in the MARS experiment. Rectangles show MARS broadband stations and triangles show short period seismometers installed by collaborators from the University of Alaska (CODEX) at the same time as the MARS deployment.

aries occurs. The change in stress affects the tectonics of the overriding plate which may include unusual volcanism, fragmentation of the overriding plate, and micro-plate capture (Stock and Lee, 1994).

The complexity and uniqueness of this region is related to the presence of two independent subducting oceanic plates, the Rivera and Cocos plates, with different ages, compositions, convergence velocities and subduction dip angles. Each sector of the arc is also controlled by a complex stress and deformation history that includes extensional and strike-slip faulting concurrent with magmatism. In addition, the convergent margin and the magmatic arc have not remained static throughout their histories, but instead have undergone notable modifications in terms of geometry and composition.

2 DATA AND METHOD

This study use data recorded on Mapping the Rivera Subduction Zone (MARS) array project. The MARS array consisted of 50 broadband sensors that were deployed throughout Southwestern Mexico for 18 months beginning in January 2006. Stations were spaced from 35 to 50 km apart with a combination of Strekeisen STS-2 and Guralp 3T and ESP sensors. The MARS experiment involved two U.S. institutions, the University of Texas at Austin and New Mexico State University, in collaboration with two Mexican institutions, Centro de Geociencias, Universidad Nacional Autónoma de México, and the Volcanic Observatory at the Universidad de Colima. We also use data collected by the Colima Volcano Deep Seismic Experiment (CODEX). CODEX consisted of a deployment of 20 short-period instruments around Colima Volcano and overlapped the MARS experiment for 5 months in time (Gardine et al., 2007). Finally, we also used two broadband seismic stations from the Mexican National Seismic network located within Southwestern Mexico.

To perform 3-D tomography model using 3-D ray trac-

ing, we used a recently developed tomography inversion package called tomo3d (Rawlinson, 2005) to generate a 3-D velocity model of the crust and upper mantle beneath Southwestern Mexico. Parameterization is represented by 3-D velocity nodes defined in spherical coordinates that have continuous smooth velocity fields between the nodes that are interpolated with cubic B spline functions. The tomo3d method utilizes 3-D raytracing and a subspace inversion technique [Kennett et al, 1988] to solve the delay time tomography problem. The 3-D ray tracing is a grid-based eikonal-solver that uses a modified finite difference method that is known as the fast-marching method. This method has been tested and shown to be computationally efficient and stable (Rawlinson and Sambridge, 2003).

The last step in seismic tomography is to solve the inverse problem that can be seen as a way to adjust the value of model parameters (m) to satisfy the data ($dobs$). Seismic tomography inverts the differences between measured and predicted travel times (travel time residuals) of seismic waves to obtain deviations in seismic velocity from the starting model used to make the travel time predictions. Two common approaches to solve the inversion problem are using the back projection and gradient based method. In the back-projection method, the travel time can be seen as an integral quantity along a raypath that is estimated using ray tracing technique. Then, the relationship between travel time perturbation (d) and slowness perturbations (m) can be linearized as,

$$d = Gm \tag{1}$$

where G is a matrix of ray lengths corresponding to the distance traversed by each ray in each block. The new updated (m) is then solved by either using the algebraic reconstruction technique (ART) or simultaneous iterative reconstruction technique (SIRT). Once the new updated model (m) is calculated, new ray paths can be re-calculated and the back projection is repeated until satisfying some convergence criteria.

The second method used in this study is the gradient based method that uses the derivative of model predictions in order to produce a solution. Let the data d be a vector dependent on a model m that can be written as $d=g(m)$. The difference between observed travel time and calculated data from a model [$dobs - g(m)$] gives an indication of the accuracy of the model and can be quantified by constructing an objective function $S(m)$ consisting of a weighted sum of data misfit and regularization terms. A typical objective $S(m)$ function can be written as,

$$S(m) = \underbrace{(g(m) - d_{obs})^T C_d^{-1} (g(m) - d_{obs})}_{1} + \underbrace{\epsilon(m - m_0)^T C_m^{-1} (m - m_0)}_{2} + \underbrace{\eta m^T D^T D m}_{3} \tag{2}$$

where m are the model parameters, $dobs$ the observed data residuals, $g(m)$ the predicted residuals, C_d the a priori data covariance matrix, C_m the a priori model covariance matrix, ϵ is damping parameter, η is smoothing parameter and D is a second derivative smoothing operator. The first part of the objective function equation 2 that contains data covariance matrix C_d , is an operator that calculate the difference between observed and predicted data. The second part of the equation that contains damping parameter is a regulariza-

tion term that provides additional constraints on solution and reduces the non-uniqueness of the solution. The third term of the equation that contains the smoothing parameter is additional regularization that attempts to find an acceptable trade-off between satisfying the data and finding a model with a minimum amount of structural variation (Constable et al., 1987).

In this study, we use an iterative nonlinear approach to minimize equation 2 as the solution to the inverse problem. Starting with an initial model [m], the objective function is then minimized by updating the initial model using a subspace inversion scheme (Kennett et al., 1998). Note that we use ak135 velocity model (Kennett et al, 1995) as the initial model. Once the new model is obtained ($m_{i+1}=m_i+\Delta m$), new 3-D ray tracing is then reconstructed to update ray path and travel time information. The looping continues until the misfit between observed data and calculated data becomes small and stagnant. The subspace inversion solution of equation 1 is derived as:

$$\delta m = -A[A^T(G^T C_d^{-1} G + \epsilon C_m^{-1} + \eta D^T D)A]^{-1} A^T \hat{\gamma} \tag{3}$$

where $A=[a^j]$ is the $M \times N$ projection matrix (M is the number of unknowns and N is the subspace matrix dimension), G is the Frechet derivatives matrix and γ is the gradient vector. A full derivation of this equation can be found in Rawlinson and Sambridge [2003].

3 INVERSION

In this study, we used the same data measured by Yang et al. (2009) in their tomographic study of southwestern Mexico. In the Yang et al. (2009) study, travel times were measured in three different frequency bands as finite frequency kernels are different depending on the frequency band used. Ray theory is an infinite frequency approximation so we only used the high frequency band measurements (.5-2 Hz) from their study. The travel times came from 269 earthquakes with body-wave magnitudes greater than 5.3 at epicentral distances from 300 to 900 for direct P phases. Additional PKPpdf phases at epicentral distances greater than 1580 were also used. The relative travel-time delays of waves recorded by MARS and CODEX stations were measured by a multi-channel cross-correlation (MCCC) method [VanDecar and Crosson, 1990] for each event with respect to the IASP91 seismic model (Kennett and Engdahl, 1991). The resulting data set of relative travel times is highly accurate as shown by the value of the MCCC standard deviations of 12ms. In total, 7890 high-quality P wave arrivals, including 774 PKPpdf phase arrivals were collected. The azimuthal distribution of P earthquake sources is good as can be seen in Figure 2.

The tomo3D tomography technique uses absolute residual travel times, recorded across an array, as data. The residuals measured at a seismic network are due to seismic velocity variations throughout the Earth and are also susceptible to errors in the location and origin time of the source. Figure 3 shows the average residuals for each station. Negative residuals mean the observed data arrives faster than the reference model predicted and positive residuals indicate slower velocity rock beneath the station. The residuals

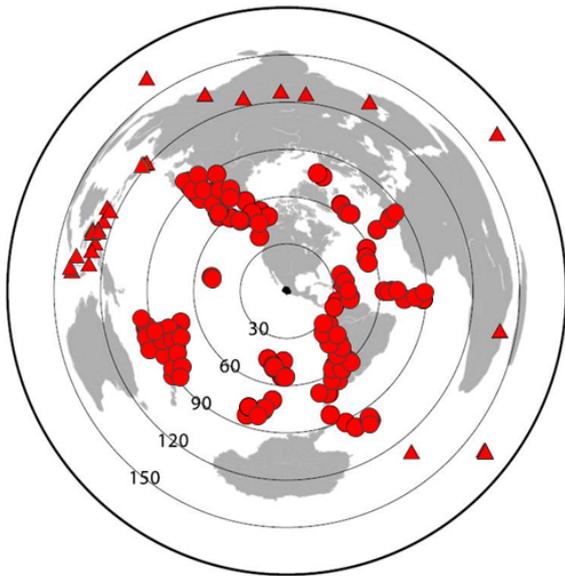


Figure 3. Source distribution map. The map is centered on Southwestern Mexico with the MARS seismic array shown as a black circle. Red circles are earthquakes for which P waves measured and red triangles are events that produced PKP arrival times.

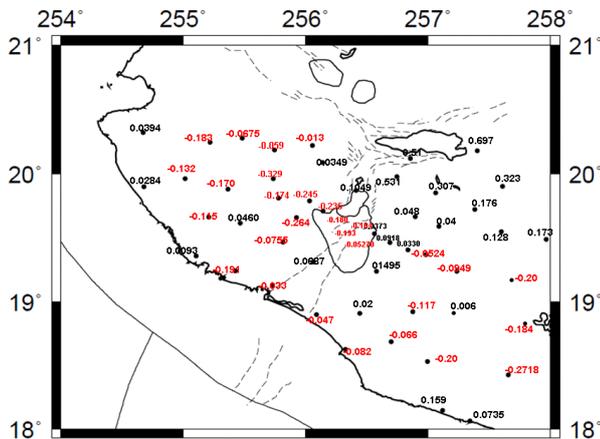


Figure 4. Map of average station residuals. The red numbers represent negative time residuals and black numbers are positive time residuals. All numbers are plotted in seconds.

shown in Figure 3 are relatively small ranging from -0.6 to 0.6 sec with faster stations through the middle of the array.

Figure 4 shows station residuals from four events with different back azimuths (northwest, northeast, southeast and southwest). In this figure large changes in residuals can be seen depending on the incoming direction of the waves. For events northeast of the array fast residuals are seen in the eastern part of the array at a latitude of about 19°N and from the northwest fast residuals are measured for stations near Colima Volcano. These waves are likely passing through deeper Rivera and Cocos slabs. From the south, waves pass perpendicular through the slabs and do not result in strong variations in times across the array.

The model volume was comprised of 61,380 velocity

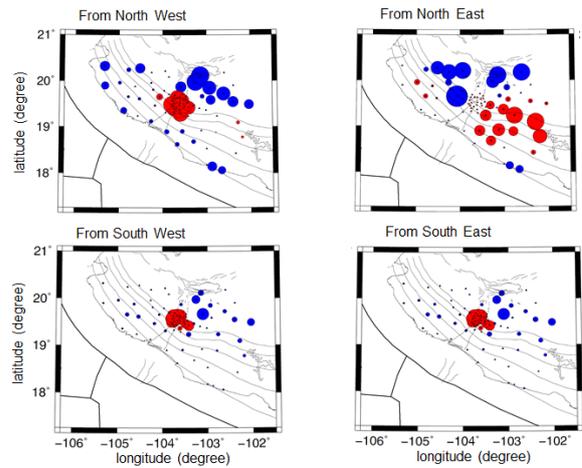


Figure 5. Circles show arrival time residuals (in seconds) from four different events from different back azimuths. Blue color represents positive relative time residuals (signal arrives slower than the reference model) and red color represents negative time residuals (signal arrives faster than the reference model).

nodes spaced at approximately 20 km intervals in all three dimensions (latitude, longitude and depth) to represent velocity structure in the crust and upper mantle beneath Southwestern Mexico. The model spans 400 km in depth, 3.4° in latitude and 5° in longitude. Crustal thickness and velocity, obtained from the receiver function study, were incorporated into the starting model. For the starting model below the crust I used the ak135 global reference model (Kennett et al, 1995). The travel time residuals were also corrected for elevation differences of the stations.

The inversion procedure was carried out in six iterations of a 20-D subspace inversion routine. The forward problem is solved at every iteration to obtain new travel times, ray paths, and Fréchet derivatives. The difference between observed data and calculated data decreased significantly in the first 3 iterations but did not change much after the fifth iteration. For this reason the inversion was terminated after six iterations. A trade-off analysis of the model norm versus variance reduction is performed to determine the damping factor and the smoothing weight. Figure 5 shows the steps to determine the number for damping and smoothing parameters. we initially start with damping parameter (ϵ) equal to 1 and vary the smoothing parameter from 300 to zero. The optimum smoothing parameter is chosen as the model with maximum smoothness that still fits the data to a high degree. Once the optimum number is selected, we repeated the same process by varying the damping parameter with the chosen smoothing parameter. The optimum number for the damping parameter is 5 and the optimum smoothing parameter is 10. These values are used for the final inversion.

The final model reduces the data variance by 84% from 0.1423 s² to 0.0392 s², which corresponds to an RMS reduction from 377 ms to 199 ms. Histograms showing the distribution of time residuals using the initial and final models are shown in Figure 6. Before the inversion, most of the residual arrival times fall in a range -0.75 s to 0.75s with a relatively normal distribution and wide variance. After inversion most data now fall in a range of -0.3s to 0.3s.

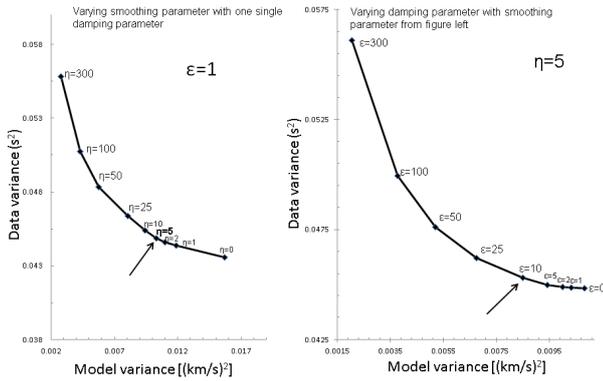


Figure 6. Steps to determine the damping and smoothing parameter for the inversion. Left, the model variance is plotted against the data variance while varying the smoothing parameter between 300 and 0 with the damping parameter $\epsilon=1$. Right, varying the damping parameter between 300 and 0 with a $\eta=5$. The final optimum number for the inversion (shown with the arrow symbol) are: smoothing parameter $\eta=5$ and damping parameter $\epsilon=10$.

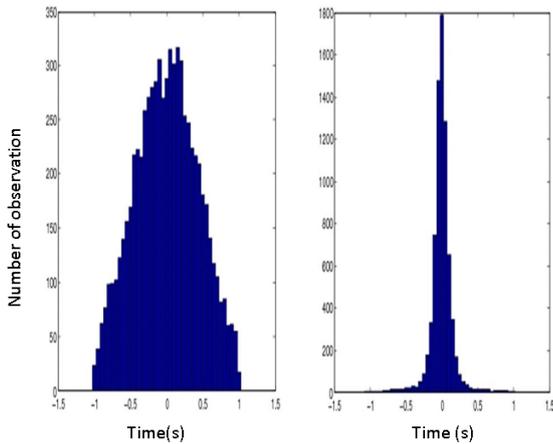


Figure 7. Residual histograms before (left) and after inversion (right).

4 RESOLUTION TEST

Resolution in most teleseismic tomography studies suffers to some extent due to most problems being under determined which happens when there is an un-balanced condition between unknown parameters and the available recorded data set that can lead to a non-unique solution in inversion. We performed a “checkerboard” synthetic experiment to test the robustness of the model. In the checkerboard test, the model is divided into alternating regions of high and low velocities. We divided the model into squares that have a maximum velocity of 0.1 km/s and a minimum velocity of -0.1 km/s relative to the ak135 velocity model. This corresponds to about 1.2-1.5% velocity differences (Figure 7). The highest amplitude of the anomaly is located at the center of the cube and decreases outward. The size of the cubes is about 160 km horizontally and vertically and spaces are located between cubes. A set of travel time residuals are

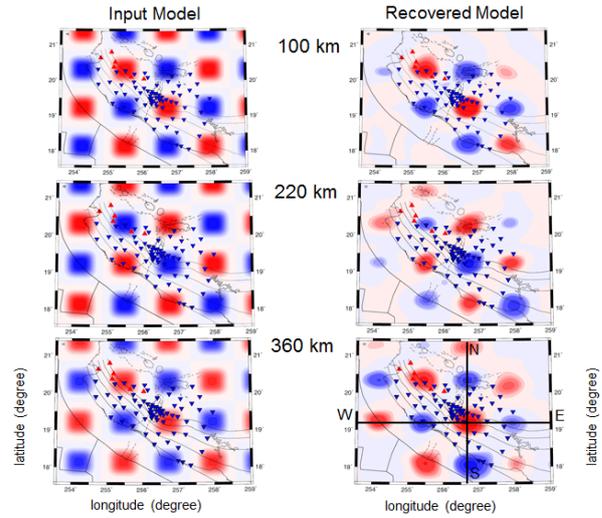


Figure 8. Inverted checker board model at different depths. Red color shows slow velocity anomalies and blue color shows fast velocity anomalies. The A-B and C-D lines are cross sections shown in Figure 8.

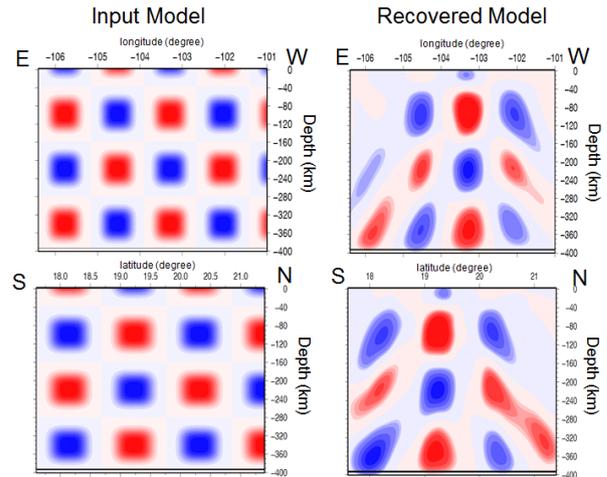


Figure 9. The right figures are the cross section recovered models and the left figures are the input models. Smearing is seen on the edges of the model especially for slow velocity anomalies. A-B and C-D map view can be seen on Figure 7.

calculated using identical sources, receivers and phase types as the observational data. The data are then inverted using the subspace inversion technique outlined above. Gaussian noise with a standard deviation of 10 ms is also added to the synthetic data sets to simulate the noise content of the observed data. The difference between the true model and the recovered structures indicate which regions of the model are well constrained by the data. Horizontal and vertical slices through the recovered models are shown in Figure 7 and Figure 8.

The inverted synthetic models show good recovery in terms of the shape of the input anomalies down to 400-km-depth although the amplitude recovered in some cubes is lower than the input model. The quality of recovered images decreases on the sides of the model as seen in the Figure 8.

Vertical streaking is visible through most of the model but especially at the edges where there are no crossing paths. Still, the streaking does not extend more than about 50 km.

In classical tomography, where raypaths are fixed by the starting model, resolution tests such as discussed above do not depend on the amplitude of the anomaly i.e., one can scale the results above to larger or smaller amplitude anomalies. In this work, however, rays are traced through 3D structures. Even with a similar pattern, the raypaths will be different depending on the amplitude of lateral heterogeneities. To investigate this effect on resolution, we performed checkerboard tests with different amplitude anomalies keeping the spatial extent of anomalies the same. We ran inversions with the maximum input anomaly ± 0.2 km/s and ± 0.9 km/s that corresponds to about 2.6-3% and 11-15% velocity differences respectively. Note that we use the same data input, noise level, and inversion parameters in the inversion process as before.

In general, the higher amplitude tests show worse results than the low amplitude study (Figure 9 and 10). The model with a 3% amplitude input shows similar results in pattern compared to the 1.5% amplitude model but has less amplitude recovery especially at deeper depths. The inversion result for an input model with 13% amplitude variations is far worse than the lower amplitude model results. The checkerboard pattern is not seen well and fast velocity anomalies dominate the slow velocity anomalies especially at shallow depths. This model also shows that slow anomalies in general are less well recovered compared to the lower amplitude results. The reason for this is that according to Fermat's principle, raypaths follow the path of least time. Thus in 3D structure the rays tend to follow the high velocity zones and avoid the slower regions and thus the density of raypaths ends up higher in the fast anomalies. This effect, combined with the smoothing and damping in the inversion results in the strong fast anomalies showing up in the inversion and the relatively weak recovery of slow anomalies.

5 TOMOGRAPHIC MODEL

The final tomography model is shown and compared to the model of Yang et al. (2009) in Figure 11 and 12. It should be noted that the data used in both studies was similar but different approximations were used in the two studies. The two models show similar results with a few possibly important differences. Both models show a northwest-southeast band of high velocities that dips to the southeast. Yang et al. (2009) interpret this feature to be the subducting Rivera and Cocos slabs. A gap in the high velocity band has been interpreted as the boundary between the two subducting plates that diverge at depth thus allowing upwelling asthenosphere that feeds Colima Volcano. We find the gap to be a little deeper (230 km) than the Yang et al. (2009) study but agree on the interpretation.

The main difference between the two studies is in the sharpness of the high velocity structures (slabs) and the amplitude of the slow velocity anomalies. These differences are seen more clearly in Figure 13 that shows a comparison of cross sections through the two models. The finite frequency results showed high velocity anomalies that broaden with depth and were interpreted as slabs. Due to the broadening

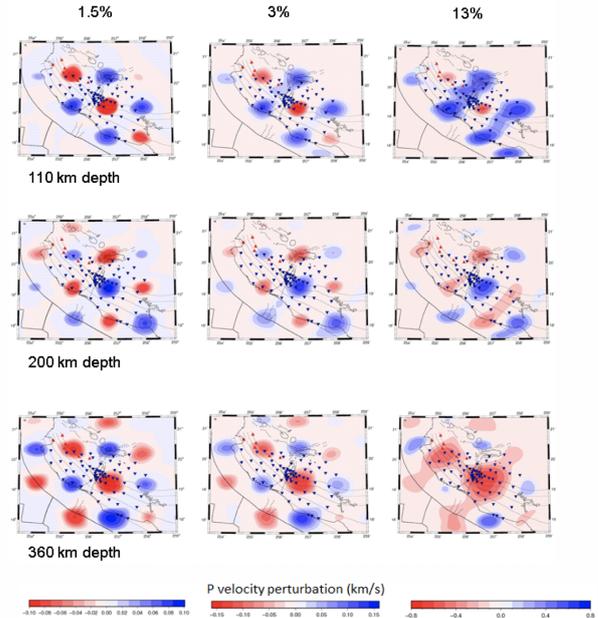


Figure 10. Recovered checker board tomography models at different anomaly input. Left is recovered model using a 1.5% anomaly input, the center is the recovered model using a 3% anomaly input. On the right is the recovered model using a 13% anomaly input with respect to the ak135 velocity model.

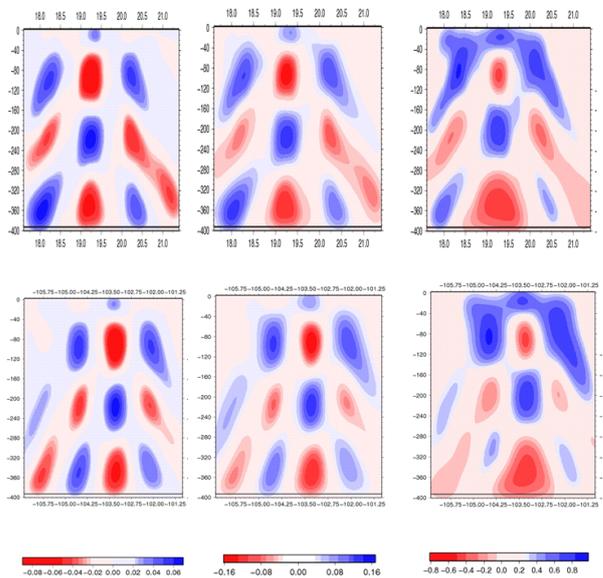


Figure 11. Cross sections along line A-B (top) and C-D (bottom) for different anomaly inputs. Left is using a 1.5% anomaly input, center is using a 3% and right is using a 13% anomaly input.

it appeared that the slabs changed dip at depth. In contrast, the 3D raytracing results show slabs that maintain their thin width and dip with increasing depth (Fig. 15. sections A4-B4 and A1-B1). I feel the 3D raytracing result is more reliable because rays will tend to travel within the high velocity anomalies to obey Fermat's principle and can thus be concentrated into thin fast structures. The finite

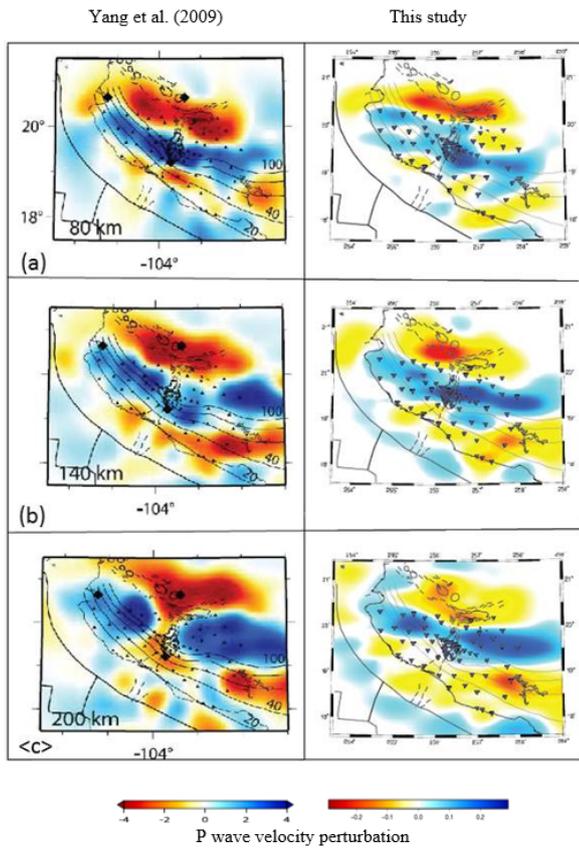


Figure 12. The final inverted model at depths of 80 km, 140 km and 200 km. The left side shows the results of a finite frequency inversion by Yang et al. (2009) and the right side is the final inversion from this study.

frequency results do not account for the perturbation of the raypaths and thus fast arrivals are back-projected over a broad region as predicted by 1D raytracing when in fact the raypaths have been perturbed into the narrow slab. Our result shows imaging and interpretation of subducting slabs require 3D raytracing to properly account for the width of the slab and its trajectory.

The second difference between the two approaches is in the amplitude of the slow anomalies. In the finite frequency results, Figure 11 and 12 show very slow velocities in the north from shallow depths to 300 km depth. It should be noted that these slow anomalies are slower than the mantle to the south of the slab near the coast. Since the mantle beneath the slab to the south is presumably asthenosphere the implication of the slow velocities to the north is that this mantle is anomalous, perhaps with higher temperatures or with higher water content than normal asthenosphere. The 3D raytracing results also show anomalously slow mantle to the north down to a depth of 150 km or so but below 150 km depth, the velocity of the mantle in the north is similar to the mantle velocities to the south beneath the slab. As discussed above, ray theory tomography does a poor job of imaging slow anomalies whereas finite frequency tomography includes wavefront healing and thus does a better job of imaging slow anomalies. On the other hand, if the fast anomalies (slabs) are artificially widened in the fi-

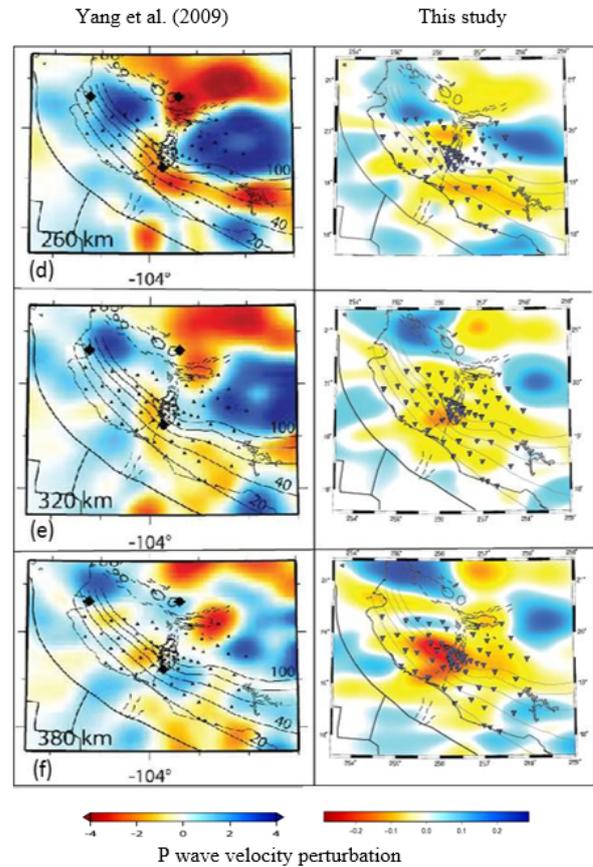


Figure 13. Tomography model at a depth of 260 km, 320 km and 380 km. The left side shows the result of previous finite frequency inversion from Yang et al. (2009) and the right side is the final inversion of this study.

nite frequency tomography, it may be that the extreme slow anomalies are just compensation for the misplaced deep fast anomalies and thus are not real features of the mantle beneath Southwestern Mexico. We note that the discrepancy in slow structure between the two inversion techniques occurs at the edge of the seismic array and thus is a region with relatively poor coverage. To determine which model is correct would likely require increasing data coverage to the north or perhaps attempting to model 3D regional or teleseismic waveforms that sample the northern part of our study area. With our current results, it is difficult to determine which inversion results are closer to the true structure.

6 DISCUSSION

In this chapter we have compared two approaches to regional tomography applied to Southwestern Mexico. The two methods both involve approximations and show differences in the images produced. Using 3D raytracing but neglecting the finite sampling of structure by seismic waves produced sharper images of fast velocity structures in the mantle. In our example, the deeper slabs are more coherent and show less broadening with depth than inversions using 1D finite frequency kernels. This has implications for the amount of deformation in slab as they descend through the

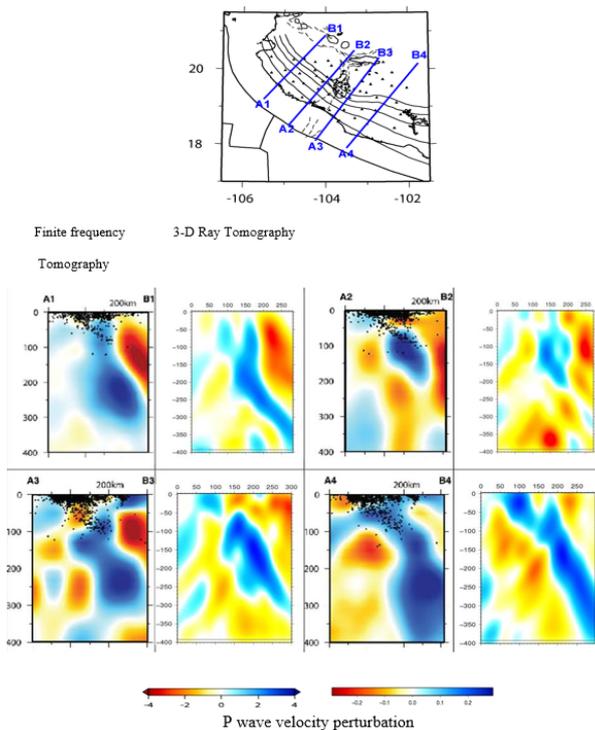


Figure 14. Comparison between finite frequency tomography results (left) and 3-D ray tracing tomography results (right) along cross sections located in the top.

upper mantle. We also find that deep slow anomalies that were found using finite frequency kernels are more subdued using the ray theory approach. This is to be expected as the finite frequency kernels take into account wavefront healing. Although the finite frequency and 3-D ray tracing models show some differences in amplitude and pattern, the overall agreement of the models supports the interpretation of Yang et al. (2009) that slab rollback is occurring in South Western Mexico leading to a coastward migration of volcanism. Our models support the idea of Ferrari et al (2001) who proposed that the dip of the Rivera slab increased when the convergence rate slowed between 8.5 Ma and 4.6 Ma resulting in slab roll back. This model can explain the trenchward migration of the volcanic front and the mixed geochemical signatures observed on the western TMBV. Both models also show a gap between the Cocos and Rivera slabs near 200 km depth and close to the location of Colima volcano.

One difference between the finite frequency and ray theory results that does contradict an interpretation from Yang et al. (2009) concerns the deep structure of the Rivera slab. Ferrari (2004) analyzed the age of mafic volcanism and proposed a tear in the subducting slab that propagated from the Gulf of California to the Gulf of Mexico during the late Miocene. The location of the west to east tear is projected to be roughly beneath the northern part of the TMBV. The finite frequency models show that the Rivera slab is clearly observable at a depth of about 300km but fades away at greater depths. Yang et al (2009) suggest that the Rivera slab tore at depth and the shallow part then steepened and rolled back. They also suggest that the Cocos slab may have also torn but the detachment is at a deeper depth. As can

be seen in Fig. 13 cross section A1-B1, the finite frequency inversion shows an end to the subducting Rivera plate at about 300 km depth. Yang et al. (2009) interpreted this as the location of the end of the slab that tore and then rolled back. As can be seen in the same figure, the 3-D ray tracing model shows a clear fast velocity band down to a depth of 400km along A1-B1 and thus our model does not support a slab tear of the Rivera plate above 400 km depth. If a tear in the slab did occur, the slab end is deeper. Unfortunately, the resolving power of both our tomography inversions becomes weak at deeper depths, especially to the north.

References

- Baig, A., F. A. Dahlen, and S.-H. Hung, (2003), Traveltimes of waves in random media, *Geophysical Journal International*, 153, 467–482.
- Buske, S., and Kästner, U., (2004), Efficient and accurate computation of seismic travel times and amplitudes. *Geophysical Prospecting*, 52, 313–322.
- Constable, S.C., Parker, R.L., and Constable, C.G., (1987), Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, 52, 289–300.
- Dahlen, F. A., Hung, H. S. and Nolet, G., (2000), Fréchet kernels for finite frequency traveltimes, I: Theory, *Geophysical Journal International*, 141, 157–174.
- Ferrari, L., Petrone, C. and Francalanci, L., (2001), Generation of oceanic-island basalt-type volcanism in the western Trans-Mexican volcanic belt by slab rollback, asthenosphere infiltration, and variable flux melting, *Geology*, 29, 507–510.
- Ferrari, L., (2004). Slab detachment control on mafic volcanic pulse and mantle heterogeneity in central Mexico, *Geology*, 32(1): 77 – 80.
- Gardine, M., Dominguez, T., West, M., Grand, S., Suhardja, S., (2007), The Deep Seismic Structure of Volcan de Colima, Mexico *Eos Trans. AGU*, 88(23), Fall Meeting Supplementary, Abstract T51A-02.
- Hole, J. A., and Zelt, B. C., (1995), 3-D finite-difference reflection travel times, *Geophysical Journal International*, 121, 427–434.
- Hung, S.H., Dahlen, A. F. and Nolet G., (2001), Wavefront healing: A banana-doughnut perspective, *Geophysical Journal International*, 146, 289–312.
- Hung, S.H., Shen, Y. and Chiao, Y. L., (2004), Imaging seismic velocity structure beneath the Iceland hot spot: A finite frequency approach, *Journal of Geophysical Research*, 109, B08305, doi: 10.1029/2003JB002889.
- Julian, B. R., and Gubbins, D., (1977), Three-dimensional seismic ray tracing: *Journal of Geophysics*, 43, 95–113.
- Kennett, B. L. N., and Engdahl, E. R., (1991), Traveltimes for global earthquake location and phase identification, *Geophysical Journal International*, 122, 429–465.
- Kennett, B. L. N., Engdahl, E. R. and Buland, R., (1995), Constraints on seismic velocities in the Earth from travel times, *Geophysical Journal International*, 122, 108 – 124.
- Kennett, B. L. N., Sambridge, M. S. and Williamson, P. R., (1988), Subspace methods for large scale inverse problems involving multiple parameter classes, *Geophys. J.*, 94, 237 – 247.

- Marquering, H., F. A. Dahlen, and G. Nolet., (1999), Three-dimensional sensitivity kernels for finite-frequency travel times: The banana-doughnut paradox, *Geophysical Journal International*, 137, 805– 815.
- Montelli, R., Nolet, G., Masters, G., Dahlen, F. A. and Hung, S.-H., (2004), Global P and PP travel time tomography: rays versus waves. *Geophysical Journal International*, 158, 637–654.
- Rawlinson, N. and Sambridge, M., (2003), Seismic travel-time tomography of the crust and lithosphere, *Advances in Geophysics*, 46, 81-198.
- Rawlinson, N. and Sambridge M., (2005), The fast marching method: An effective tool for tomographic imaging and tracking multiple phases in complex layered media, *Explor. Geophys.*, 36, 341-350.
- Sambridge, M. S., and Kennett, B. L. N., (1990), Boundary value ray tracing in a heterogeneous medium: A simple and versatile algorithm: *Geophysical Journal International*, 101, 157–168.
- Stock J. M. and Lee, J.,(1994). Do microplates in subduction zones leave a geological record? *Tectonics*, 13, 6, 1472-1487.
- Urn, J. and Thurber, H. C., (1987), A fast algorithm for two-point seismic ray tracing, *Bull. Seismic Society of America*, 77, 972-986.
- Van Decar, J. C. and Crosson, S. R., (1990), Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. Seismol. Soc. Am.*, 80, 150– 169.
- Van Trier, J., and Symes, W. W., (1991), Upwind finite-difference calculation of travel times, *Geophysics*, 56, 812–821.
- Vidale, J. E., (1988), Finite difference calculations of traveltimes: *Bulletin of the Seismological Society of America*, 78, 2062–2076.
- Yang, T., Shen, Y., van der Lee S., Solomon, C. S. and Hung, H. S., (2006), Upper mantle structure beneath the Azores hotspot from finite-frequency seismic tomography, *Earth and Planetary Science Letter*, 250, 11 –26.
- Yang, T., Grand, P. S., Wilson, D., Guzman-Speziale, M., Gomez-Gonzalez, M. J., Dominguez-Reyes, T. and Ni. J., (2009), Seismic structure beneath the Rivera subduction zone from finite-frequency seismic tomography, *Journal of Geophysical Research*, 114, B01302.
- Zhao, L., Jordan, H. T. and Chapman, H. C., (2000), Three-dimensional Frechet differential kernels for seismic delay times, *Geophysical Journal International*, 141, 558– 576.