91

Seismic Tomography

Robert L. Nowack and Cuiping Li

Department of Earth and Atmospheric Science, Purdue University, West Lafayette, IN 47907, USA

1 • Introduction

The inversion of seismic travel-time data for radially varying media was initially investigated by Herglotz, Wiechert, and Bateman (the HWB method) in the early part of the 20th century [1]. Tomographic inversions for laterally varying media began in seismology starting in the 1970's. This included early work by Aki, Christoffersson, and Husebye who developed an inversion technique for estimating lithospheric structure beneath a seismic array from distant earthquakes (the ACH method) [2]. Also, Alekseev and others in Russia performed early inversions of refraction data for laterally varying upper mantle structure [3]. Aki and Lee [4] developed an inversion technique using travel-time data from local earthquakes. The first global tomographic images were then obtained using both body waves and surface waves [5,6]. In this chapter, we provide an overview of seismic tomography mostly applied to seismic body waves. We first describe travel-time tomography using different approaches including variable damping, iterative methods, irregular gridding, autoregressive extrapolation, nonlinear inversions, and multi-grid methods. We next provide an overview of applications of seismic tomography to cross-hole, refraction and reflection data, local earthquake data, and teleseismic data. An example is given for the inversion of data from the TomoVes experiment to image beneath the Mt. Vesuvius volcano. Finally,

a brief discussion of seismic attribute, waveform, and finite-frequency tomography for seismic body waves is given.

2 • Seismic Travel-Time Tomography

The goal of seismic travel-time tomography is to find a velocity model for a subsurface volume consistent with measured travel-times along ray paths that pass through the volume. The travel-time for high-frequency seismic waves is given by

$$T_i = \int_{\Gamma_i} \frac{1}{v(x,y,z)}\, d\vec{x} = \int_{\Gamma_i} s(x,y,z)\, d\vec{x} \tag{1}$$

where $v(x,y,z)$ is the velocity distribution, T_i is the observed travel-time, Γ_i is the ray path, and $s(x,y,z)$ is the slowness or inverse of velocity. The velocity or slowness fields can be represented by a limited set of parameters after discretization with a set of basis functions. The problem can then be written as

$$d_i = G_i(x_j) \tag{2}$$

where d_i are the observed travel-times and x_j are the model parameters. Ray tracing algorithms and model parameterization are essential to seismic tomography and a detailed overview is given in [7]. Here we mostly describe the tomographic inversion aspects.

The tomography problem can be linearized about an initial model x_0 with corresponding predicted travel-times T_0 as

$$A\,\delta x = \delta d \tag{3}$$

where $\delta d = T - T_0$ is the vector of travel-time residuals (observed minus predicted travel-times), $\delta x = x - x_0$ is the vector of model deviations from the initial model, and $A_{ij} = \partial G_i / \partial x_j$ are the components of the sensitivity matrix A evaluated for the model x_0. Often, this set of equations is operationally under-determined where $(A^T A)$ is rank deficient. In this case the observed data alone cannot completely resolve the model and the inverse solution requires some kind of regularization. For under-determined problems, one approach is to use a combined cost function which minimizes both the observed minus predicted data residuals and some norm of the model variations. For simple damped least squares (DLS), the regularization can be written as

$$\begin{bmatrix} A \\ \varepsilon I \end{bmatrix} \delta x = \begin{bmatrix} \delta d \\ 0 \end{bmatrix} \tag{4}$$

where ε is a small, positive damping parameter. Solving this combined system using least squares results in

$$\delta x = (A^T A + \varepsilon^2 I)^{-1} A^T \delta d \tag{5}$$

where T indicates the transpose. In general, this results in a compromise in finding a model that fits the data and has model deviations that are not too large.

In Eq. (4), the model parameters all have the same weight. But for most seismic tomographic experiments some weighting of the travel-times and model parameters is required. To overcome the shortcomings of the DLS method, weighting matrices W_1 for the data and W_2 for the model which are symmetric and positive definite can be used. An objective function can then be written as

$$S = (\delta d - A\,\delta x)^T W_1 (\delta d - A\,\delta x) + \varepsilon^2 \delta x^T W_2\, \delta x \tag{6a}$$

where the diagonal elements of W_1 and W_2 allow for nonuniform weighting of the data and model parameters, and the off-diagonal elements account for data correlations and model smoothing. These also result in modified inner products of the data and model spaces. Hansen [8] proposed an L-curve approach to obtain an appropriate value for ε^2. In this approach the log of the normed data misfit is plotted against the log of the normed model deviation. The trade-off parameter ε^2 at the 'knee' of this curve provides a trade-off between tightly fitting the observed data (as ε^2 goes to zero) and minimizing a norm of the model deviations (as ε^2 goes to infinity) (Figure 1). With this value of ε^2, minimizing Eq. (6a) gives

$$\delta x = (A^T W_1 A + \varepsilon^2 W_2)^{-1} A^T W_1\, \delta d \tag{6b}$$

For the above linear tomographic problem, a number of different forms of the weighting matrices W_1 and W_2 have been suggested. Using a Bayesian approach, these can be related to the data and prior model covariances as $C_D = E[\delta d\, \delta d^T] = W_1^{-1}$ and $C_{x_0} = E[\delta x\, \delta x^T] = \varepsilon^{-2} W_2^{-1}$ [9]. In this approach, the trade-off parameter is absorbed into the prior model covariance. But the model covariance must be strictly based on prior information from a Bayesian standpoint [9]. Nonetheless, Eq. (6b) is often used with weighting matrices applied operationally.

For nonuniform ray coverage and variable model cell sizes, the weighting matrices can be chosen to approximately compensate for these effects [10–12]. However, additional damping may still

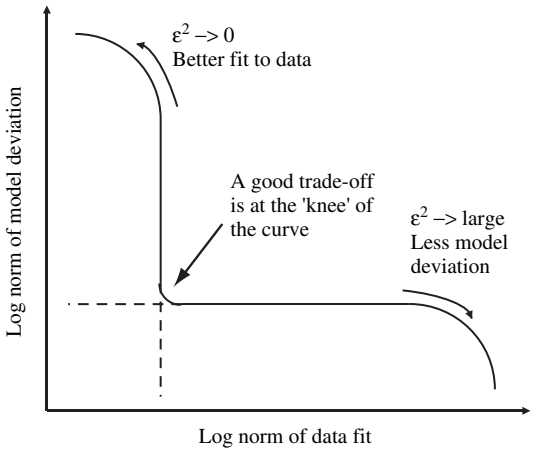

FIGURE 1 The generic form of the L-curve. The L-curve displays the compromise between minimizing the model deviation and minimizing the data misfit. The trade-off parameter can be estimated from the "knee" of the L-curve [8].

be required to avoid instabilities for cells with few rays. An example is shown in Figure 2 in which a cross-hole geometry is used to illustrate the effects of variable damping based on cell size and ray length in a given cell. Figure 2A shows the true model which is a linear gradient from 5 to 5.8 km/s. Figure 2B shows the homogeneous starting model of 5 km/s. Also shown in Figure 2B are the locations of the spline interpolated velocity nodes of the model displayed as solid circles. The rays are shown by white lines with the sources displayed as open circles and the receivers by open triangles on the left and right boundaries. This results in variable ray coverage with dense ray coverage in the upper part of the model but only two rays in the lower half of the model. Also, the cell sizes for the nodes along the boundaries of the model are effectively smaller than those for the interior nodes. The partial derivatives are computed from integrations of the perturbed travel-times along the initial rays for small perturbations of the spline interpolated nodes.

A linearized inversion for the model using simple damping is shown in Figure 2C where both W_1 and W_2 are identity matrices and ε^2 is chosen to give the best overall magnitude of model perturbations. As can be seen, the unequal ray density and different cell sizes result in differences of the inverted model from the true model. These effects can be minimized by adding variable damping terms to W_2. First a term related to the square-root of the diagonal elements of $(A^T A)$ is used to account for the

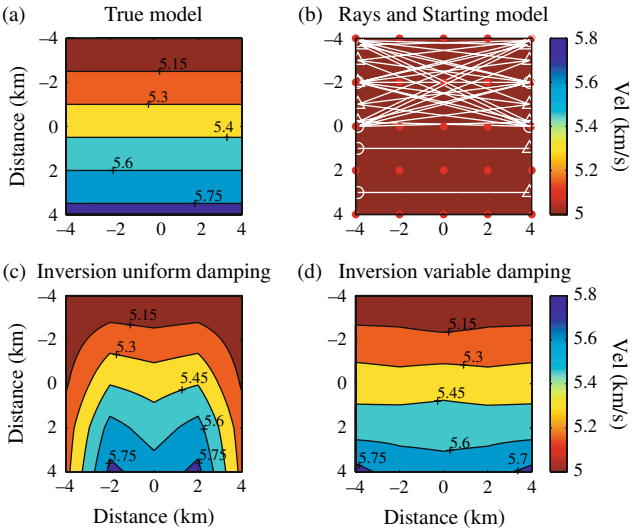

FIGURE 2 Tomographic example using unequal cell sizes and variable ray coverage. (**a**) The true model. (**b**) The starting model with ray coverage and nodes positions with the boundary nodes having smaller cells. (**c**) The inversion with uniform damping. (**d**) The inversion with variable damping.

ray lengths within different cells of the model [11]. Also, a small amount of simple damping is added to this for stability. This is then multiplied by a term which equalizes for the small and large block sizes of the model parameterization [10]. Finally, ε^2 is chosen to give the best overall magnitude of the model perturbations. Figure 2D shows the result of using this variable damping to balance the variability in ray coverage and cell size. Although this is a simple example, it illustrates the utility of a variable damping approach.

As an alternative approach, Deal and Nolet [13] suggested a nullspace shuttle technique for constraining poorly resolved parts of the model which simultaneously conserves the data fit and incorporates an a priori model into the solution. The null space component of the prior model is preserved when the data is insufficient.

2.1 Iterative Inversion for Linearized Problems

The solution of Eq. (6b) requires solving a matrix equation which can be prohibitively time consuming for large numbers of model parameters. Large-scale linear problems are usually solved using iterative methods and there are two popular iterative methods used for tomography, SIRT (simultaneous iterative reconstruction technique) and LSQR, a modified conjugate

gradient method. Dines and Lytle [14] introduced the SIRT method from medical imaging to geotomography. However, Van der Sluis and Van der Vorst [15] showed that the SIRT method does not solve the original inverse problem, Eq. (3), but instead solves a weighted least squares problem in modified data and model norms. Trampert and Leveque [16] also analyzed these properties of SIRT. In a similar fashion to using variable damping, different least squares solutions can be obtained if the metrics in the original and transformed model and data spaces are changed.

The LSQR method belongs to the family of conjugate gradient methods and was introduced by Nolet [17] to seismic tomography problems. The LSQR approach seeks the solution δx that minimizes the $||A\,\delta x - \delta d||_2$ subject to also minimizing $||\,\delta x\,||_2$. Van der Sluis and Van der Vorst [14] showed that LSQR starts the iterative process by ignoring those components associated with the smallest eigenvalues of $(A^T A)$. After more iterations, the contributions of the smaller eigenvalues eventually enter the solution. This property contributes to the intrinsic damping of the algorithm.

Another iterative approach is to use subspace methods in which the iterative process is carried out over higher dimensional subspaces of the model. An early description of this method can be found in Kennett et al. [18] (for more details see [7]).

2.2 Irregular Grid Inversion

Seismic tomography using a regular grid and uniform damping may have undesirable properties in the presence of inhomogeneous sampling of the model by the ray paths, and may result in an unnecessary over-parameterization of the model space [19,20]. The use of irregular grids may be less sensitive to regularization issues than inversions based on regular grids. Curtis and Snieder [21] suggested a genetic algorithm to do an irregular parameterization of the model space by searching a configuration of Delaunay triangles that result in the best conditioned inverse problem. Sambridge and Gudmundsson [22] proposed an irregular cell approach based on the tetrahedral and Voronoi diagrams and gave several options to do the irregular cell design. Spakman and Bijwaard [19] optimized the irregular cell parameterizations by equalizing the ray hit-counts in each cell. Sambridge and Faletic [23] recently proposed an adaptive tetrahedral cell scheme based on the maximum spatial gradients in the seismic velocity across each tetrahedral face. For their adaptive method, the study region was initially divided into a uniform tetrahedral mesh. The scheme then refined the parameterization during the inversion in response to the structural features detected.

2.3 Autoregressive Extrapolation

Autoregressive (AR) extrapolation methods have been used in signal processing to predict time series, as well as perform prediction-error filtering. Claerbout [24,25] suggested a two-stage scheme to apply autoregressive techniques to the linear least-squares problem. The first stage is to find the optimal prediction-error filter (PEF). For the second stage, the PEF is assumed to be known, and values for the missing data are found so that the power out from the prediction-error filter is minimized. This is similar to maximizing an entropy criterion [26].

In terms of an unknown model, one approach to apply the AR technique to the seismic inversion problem is to separate the problem into two steps. For model extrapolation, this requires some information about the model in order to compute the PEF. First nodes with few or no rays nearby are removed from the inversion, and only nodes with good ray coverage are inverted. Then, making use of the initial inversion result, the AR interpolation technique can be used to obtain the values for the unconstrained model nodes.

An alternate approach is to apply AR extrapolation directly to the travel-time data instead of the model. In this approach, the known data is first extrapolated to estimate the unknown data. Then, the combined known and extrapolated data are inverted for the model parameters. Using this approach, the missing data and the prediction-error filter are jointly estimated using the known data. The combined observed and extrapolated data are then used to invert for the model. This can be written as

$$F \begin{bmatrix} \delta d_a \\ \delta d_b \end{bmatrix} = r_2 \approx 0 \tag{7}$$

and

$$A\,\delta x - \begin{bmatrix} \delta d_a \\ \delta d_b \end{bmatrix} = r_1 \approx 0 \tag{8}$$

where in Eq. (7) the missing data δd_b and the PEF, F, are found from the known data δd_a by minimizing the residual r_2. In Eq. (8) the model deviations δx are then found to minimize r_1 using the known data and δd_b the extrapolated data. The extrapolated data play the role of regularization terms in Eq. (8). Equation (7) is nonlinear for the unknown F and δd_b, so this can be solved in an iterative fashion [25], and then Eq. (8) is solved for δx. Li and Nowack [27] used this approach to apply autoregressive extrapolation to the seismic tomography problem.

2.4 Nonlinear Inversion

The tomography problem can become nonlinear under large velocity perturbations if a velocity (rather than slowness) representation is used, or if the ray paths are substantially perturbed

by the velocity changes. In nonlinear inversion, the objective function to be minimized can be written as

$$S = (d - G(x))^T W_1 (d - G(x)) + \varepsilon^2 (x - x_0)^T W_2 (x - x_0) \quad (9)$$

where d is the data, $G(x)$ is the nonlinear forward problem, x is the model, x_0 is the prior model, and W_1 and W_2 are weighting matrices similar to those given earlier for the linear case. A linearization approach can be applied iteratively to solve the nonlinear problem [9]. However, the regularization is now with respect to the prior model x_0 instead of the solution at the previous iteration. For each linearization step, the linearized problem can be solved either by direct matrix methods or by iterative methods. For the error analysis, the linearized covariance and resolution operators are often computed for the converged final solution.

For strongly nonlinear problems, linearized gradient methods may fail because of a multi-minimum objective function and therefore a good starting solution is required. Alternatively, global optimization methods can be applied including MCMC (Markov Chain, Monte Carlo), simulated annealing, and genetic algorithms. A recent survey of Monte Carlo methods applied to geophysical inverse problems was given by Sambridge and Mosegaard [28]. However, Rawlinson and Sambridge [7] noted that computational expense may still limit the use of global optimization methods for seismic tomography and other inverse problems when the number of model parameters is large.

2.5 Multigrid Inversion

Traditional iterative inversion methods cannot always invert seismic data obtained from complicated earth models because of variable convergence of the high and low wavenumber components of the solution. Multigrid approaches to seismic tomography apply preliminary iterations on a coarse grid and then use this solution as an initial model on a finer grid. The use of an initial coarse grid can reduce the low wavenumber components of the model more efficiently and also allow for a variable ray coverage.

Nemeth et al. [29] summarized the following multigrid strategy for the travel-time tomography problem. First an initial iteration is performed on a smoothed model with few unknowns. After several iterations the convergence rate is likely to stall, and then the gridded model is refined by a factor of 2. The model is smoothed after refinement and new iterations are performed until convergence again stalls. This is repeated until final convergence occurs. Simple applications of a multi-grid approach were applied by Lutter and Nowack [30] for travel-times and

Nowack and Matheney [31] for mixed seismic attributes. Zhou [32] applied multi-scale tomography to cope with poor raypath coverage and solve for velocity interfaces. For global travel-time data, Chiao and Kao [33] used a wavelet approach.

3 • Examples of Travel-Time Tomography

In this section, we provide examples of travel-time tomography inversions applied to cross-borehole, refraction and reflection data, local earthquake data, and global teleseismic data. A particular example is given from the TomoVes experiment to image the shallow structure beneath the Mt. Vesuvius volcano.

3.1 Cross-Borehole, Refraction, and Reflection Tomography

In cross-borehole tomographic imaging, sources are placed in one borehole and possibly along the surface and the receivers are placed in a nearby borehole. Several examples of seismic ray tomography applied to cross-borehole data include McMechan [34], Peterson et al. [35], and Bregman et al. [36]. Menke [37] showed that cross-borehole geometries have less resolution in the crosshole direction depending on the source and receiver apertures. Lo and Inderseisen [38] provide case studies of cross-borehole tomography including monitoring steam flood injections in oil fields and imaging fault systems. Pratt and Chapman [39] and Pratt et al. [40] describe travel-time tomographic applications for anisotropic media.

Examples of 2D refraction tomography include that of White [41], Lutter et al. [42], and Zelt and Smith [43]. Lutter and Nowack [44] inverted for velocities and interface depths in a layer stripping fashion. Examples of 3D refraction tomography for crustal velocities are given by Hole [45] and Zelt and Barton [46]. Extensive reviews of refraction and wide-angle tomography are given by Nowack and Braile [47] and Rawlinson and Sambridge [7].

Examples of travel-time tomography used in reflection seismology include Bishop et al. [48] and Farra and Madariaga [49]. However, there can be an ambiguity between velocities and reflector depths as described by Williamson [50] and Stork [51] unless further a priori information is provided. The resolution is best near the surface and decreases with depth depending on the source–receiver offset range. An application of tomography to reflection seismology called stereotomography was given by Billette and Lambare [52] and uses travel-time and travel-time slope information with offset to estimate smooth velocity models. A recent assessment of this method was given by Lambare [53].

3.2 Inversion of TomoVes Experiment Data

In this example, seismic refraction data from the TomoVes experiment [54] is used to invert for shallow structure beneath the Mt. Vesuvius volcano. The imaging of the shallow structure beneath an active volcano was one of the scientific motivations of this experiment. The source and receiver geometry for the experiment is displayed in Figure 3A and can be seen to be very heterogeneous. Because of the large topographic relief in the area, a 3-D

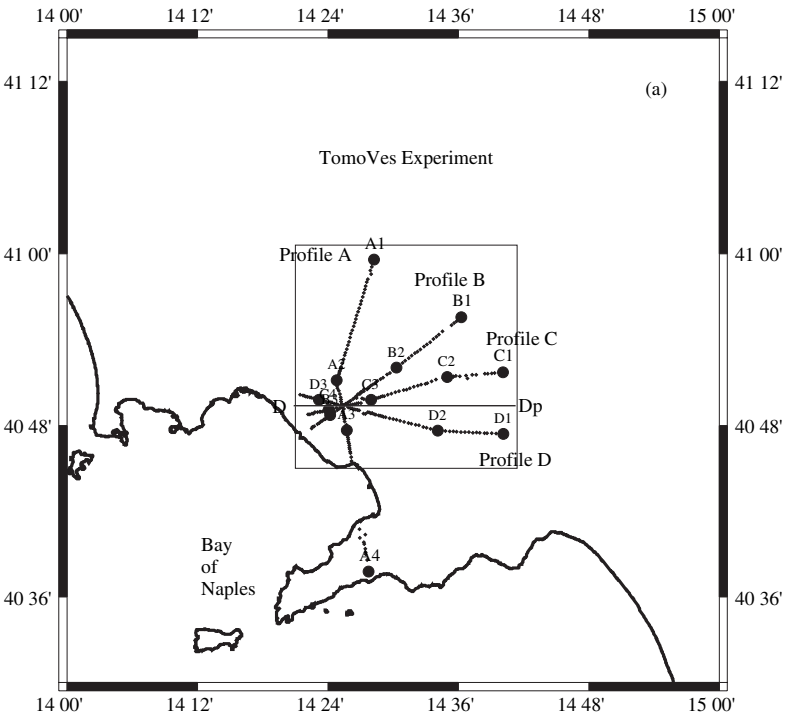

FIGURE 3 (**a**) Map of the geometry of the refraction lines from the TomoVes tomography experiment [54] going across the top of the Mt. Vesuvius volcano. The *solid dots* are the shot points and the different intersecting profiles of receivers are noted by *small dots*. The *line* D-Dp shows the location of the vertical cross-section through the inverted 3D velocity model from Nowack et al. [55,56]. (**b**) The topography and 3D node model for the Mt. Vesuvius region corresponding to the inset square in (a). The volcano is on the left and the background hills to the upper right. The *solid dots* show the tomographic nodes position used for the 3-D inversion. Rays from one shot point are also shown diving beneath the volcano. (**c**) This shows a cross-section of the tomographic velocity model from [55,56] along line D-Dp in (a).

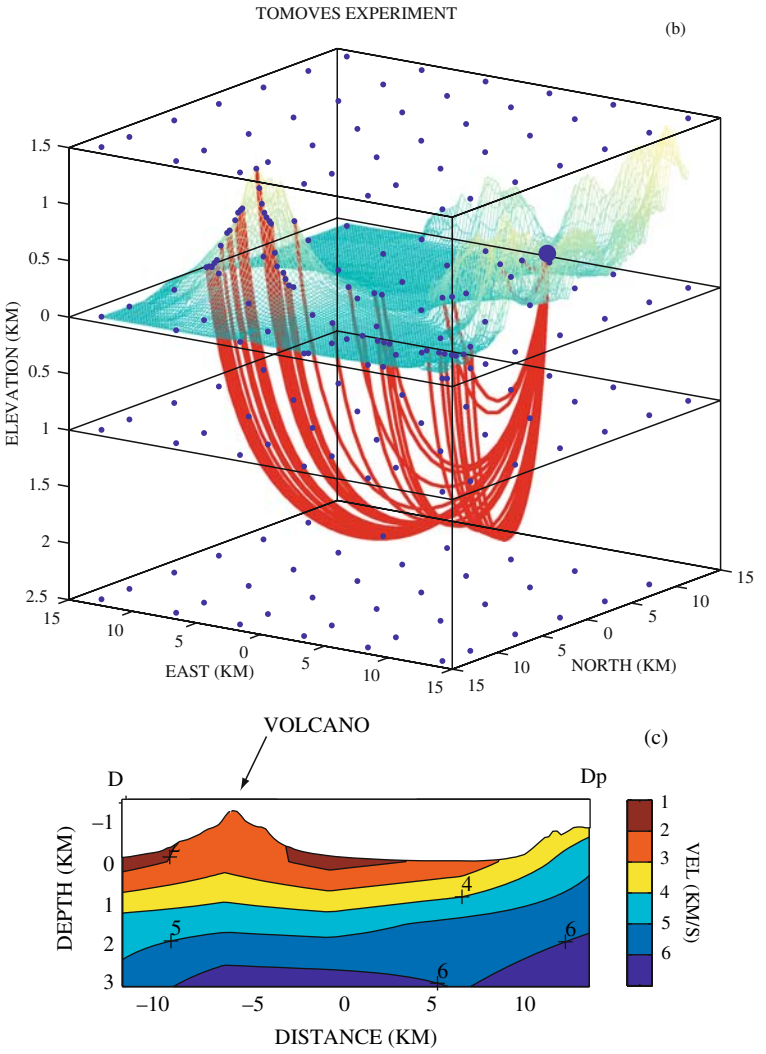

FIGURE 3 (continued)

uniform velocity grid model shown by the dots in Figure 3B was used. The velocities have been spline interpolated. The rays for one shot point are shown diving beneath the volcano. A variable damping inversion technique was applied to the first arrival travel-times and other seismic attributes of the TomoVes data by Nowack et al. [55,56]. Figure 3C illustrates the 3-D inversion result after 3 iterations shown along profile D-Dp in Figure 3A. Low surface velocities can be seen surrounding the volcano, as

well as lower velocities in the edifice of the volcano. At depths of 1.0–2.0 km, the 4 and 5 km/s velocity contour lines are seen to be shallow toward the background hills to the east and the corresponding formations are fairly continuous beneath the volcanic plain. Also, the 3 and 4 km/s contours tend to shallow directly beneath the volcano and this has been suggested to be a paleo-volcanic or sub-volcanic structure [57]. Although the derived 3D velocity model is smoothly varying, the results are similar to the more detailed velocity models along the individual 2D profiles obtained by Zollo et al. [57].

3.3 Local Earthquake Tomography

Local earthquake tomography developed as an extension of earthquake location estimation, and early examples include Aki and Lee [3] and Thurber [58]. Kissling [59] provided a detailed review of the method and Eberhart-Phillips [60] extended the method to S-waves. Hirahara [61] described using local earthquakes and teleseisms for both velocity and anisotropy estimation.

A new approach for local earthquake tomography has been the use of double-difference techniques. Zhang and Thurber [62] described an approach in which both differential arrival times as well as absolute travel-times are used for local earthquake tomography. The use of differential travel-times can improve the velocity imaging and earthquake locations by canceling some of the path effects outside the source region.

3.4 Global Tomography

Early inversions for the entire mantle were performed by Clayton and Comer [63] and Dziewonski and Anderson [64], using body waves, and by Woodhouse and Dziewonski [65], using surface waves. Since then numerous P-velocity and S-velocity global tomographic studies using both body waves and surface waves have been conducted. Recent overviews include Iyer and Hirahara [6], Kennett and Van der Hilst [66], Romanowicz [67] and Curtis and Snieder [68]. Cara [69] gives an overview of seismic anisotropy and discusses the trade-offs between anisotropy and velocity heterogeneity as well as possible anisotropy-induced artifacts on isotropic tomographic inversions.

4 • Extensions of Travel-Time Tomography

In addition to travel-times, other seismic attributes including amplitudes and pulse frequencies have been used to perform seismic tomography. Thomson [70] used body wave amplitudes to invert for velocity structure beneath the Norsar seismic array

in Norway. Neele et al. [71] performed a joint inversion of travel-times and amplitudes for upper mantle velocity structure. Using a cross-borehole geometry, Bregman et al. [72] performed an initial travel-time inversion and then used focusing-corrected amplitudes to invert for seismic attenuation. For wide-angle refraction data, Nowack and Matheney [73] inverted for both velocity and attenuation using a combined data set of travel-times, amplitudes, and instantaneous pulse frequencies. Body wave spectral ratios could also be used for the estimation of attenuation. Mitchell and Romanowicz [74] include a number of studies on the estimation and inversion of seismic attenuation. Wang [75] applied seismic amplitude inversion to reflection seismology.

The Fresnel zone depends on the seismic wavelength and the path length, and if the heterogeneity scales are on this order then wave diffraction will occur. If the heterogeneities are weak then a single scattering model, the Born approximation, can be used to model seismic wave diffraction. An early application of diffraction tomography to seismic waves was presented by Devaney [76]. Other examples include and Wu and Toksoz [77] and Pratt and Worthington [78] who used waveform tomography in the frequency-domain for cross-hole data. An overview of diffraction tomography was given by Lo and Inderwiesen [38]. Pratt et al. [79] applied waveform tomography in the frequency-domain to wide-angle seismic data, and Brenders and Pratt [80] used this approach to successfully invert for small-scale features from a synthetic, wide-angle data set. However, a starting travel-time inversion was required to provide a smooth initial velocity model for the waveform inversion.

Recently, diffraction kernels for travel-times have been investigated which allow for wavefront healing of finite frequency wavefronts. This can occur for low-frequency seismic waves having long travel paths, such as for waves traveling through the mantle. Early examples of computing Fresnel zones for finite-frequency effects include Woodward [81] and Yomogida [82]. Dahlen et al. [83] derived a fast way to compute kernels relating travel-time perturbations measured by cross-correlation of waveforms to velocity anomalies using the paraxial approximation. This was used by Montelli et al. [84] to perform finite-frequency tomography of P and PP correlation times for mantle velocity structure, and a number of hotspots fed by plumes from the lower mantle were found. They compared their results with ray tomography and concluded that finite-frequency tomography provided clearer images when using long-period seismic waves. However, this was questioned by van der Hilst and de Hoop [85] who suggested that the benefits of the more

complete theory may still not yet be realized in global tomography because of practical considerations such as damping, weighting of different data, choice of the data fit, and limited ray coverage. Although discussion of these issues has continued (Montelli et al. [86] and van der Hilst and de Hoop [87]), it is clear that further developments to improve ray coverage, incorporate finite frequency effects and include other seismic wavefield attributes will result in improved tomographic images of the Earth's interior.

References

1. K. Aki and P. G. Richards: Quantitative Seismology, (Freeman, San Francisco 1980)
2. K. Aki, A. Christoffersson and E. S. Husebye: Determination of three-dimensional seismic structure of the lithosphere, J. Geophys. Res. 82 (1977) 277–296
3. A. S. Alekseev, M. M. Lavrentiev, V. G. Romanov and M. E. Romanov: Theoretical and computational aspects in seismic tomography, Surv. Geophys. 11 (1990) 395–409
4. K. Aki and W. H. K. Lee: Determination of three-dimensional velocity anomalies under a seismic array using first P-arrival time from local earthquakes 1. A homogeneous initial model, J. Geophys. Res. 81 (1976) 4381–4399
5. G. Nolet (ed.): Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics, (Reidel, Hingham MA 1987)
6. H. M. Iyer and K. Hirahara (eds.): Seismic Tomography Theory and Practice, (Chapman and Hall, London 1993)
7. N. Rawlinson and M. Sambridge: Seismic traveltime tomography of the crust and lithosphere, In: Advances in Geophysics Vol. 46 ed. by R. Dmowska, (Elsevier, Amsterdam 2003) 81–198
8. P. C. Hansen: Rank-deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion, (SIAM, Philadelphia 1998)
9. A. Tarantola: Inverse Problem Theory and Methods for Model Parameter Estimation, (SIAM, Philadelphia 2005)
10. G. Nolet: Seismic wave propagation and seismic tomography, In: Seismic Tomography ed. by G. Nolet, (Reidel, Hingham MA 1987) 1–23
11. B. Wang: Improvement of Seismic Travel-time Inversion Methods and Application to Observed Data, (Ph.D. thesis, Purdue University, West Lafayette, IN 1993)
12. J. G. Berryman: Weighted least-squares criteria for seismic travel-time tomography, IEEE Trans Geosci Remote Sens 27 (1989) 302–309
13. M. M. Deal and G. Nolet: Nullspace shuttles, Geophys. J. Int. 124 (1996) 372–380

14. K. A. Dines and R. J. Lytle: Computerized geophysical tomography, Proc. IEEE 67 (1979) 1065–1073
15. A. van der Sluis and H. A. van der Vorst: Numerical solution of large, sparse linear systems arising from tomographic problems, In: Seismic Tomography, ed. by G. Nolet, (Reidel, Hingham MA 1987) 49–83
16. J. Trampert and J. Leveque: Simultaneous iterative reconstruction technique: physical interpretation based on generalized least squares solution, J. Geophys. Res. 95 (1990) 12553–12559
17. G. Nolet: Solving or resolving inadequate and noisy tomographic systems, J. Comp. Phys. 61 (1985) 463–482
18. B L. N. Kennett, M. S. Sambridge and P. R. Williamson: Subspace methods for large scale inverse problems involving multiple parameter classes, Geophys. J. 94 (1988) 237–247
19. W. Spakman and H. Bijwaard: Optimization of cell parameterizations for tomographic, inverse problems, Pure Appl. Geophys. 158 (2001) 1401–1423
20. N. Rawlinson and M. Sambridge: Irregular interface parameterization in 3-D wide-angle seismic traveltime tomography, Geophys. J. Int. 155 (2003) 79–92
21. A. Curtis and R. Snieder: Reconditioning inverse problems using the genetic algorithm and revised parameterization, Geophysics 62 (1997)1524–1532
22. M. Sambridge and O. Gudmundsson: Tomographic systems of equations with irregular cells, J. Geophys. Res., 103 (1998) 773–781
23. M. Sambridge and R. Faletic: Adaptive whole Earth tomography, Geochemistry Geophysics Geosystems 4 (2003)
24. J. F. Claerbout: Earth Soundings Analysis: Processing Versus Inversion, (Blackwell Scientific, Oxford 1992)
25. J. F. Claerbout: Image Estimation by Example: Geophysical Soundings Image Construction: Multidimensional Autoregression, (web book, http://sepwww.stanford.edu/sep/prof/gee/toc_html 2004)
26. J. P. Burg: Maximum Entropy Spectral Analysis, (Ph.D. thesis, Stanford University, http://sepwww.stanford.edu/theses/sep06 1975)
27. C. Li and R. L. Nowack: Application of autoregressive extrapolation to seismic tomography, Bull. Seism. Soc. Am. 94 (2004) 1456–1466
28. M. Sambridge and K. Mosegaard: Monte Carlo methods in geophysical inverse problems, Rev. of Geophysics 40 (2002) doi10:1029/2000RG000089
29. T. Nemeth, E. Normark and F. H. Qin: Dynamic Smoothing in Crosswell Traveltime Tomography, Geophysics 62 (1997) 168–176
30. W. J. Lutter and R. L. Nowack: Inversion for crustal structure using reflections from the PASSCAL Ouachita experiment, J. Geophys. Res. 96 (1990) 4633–4646
31. R. L. Nowack and M. P. Matheney: Inversion of seismic attributes for velocity and attenuation structure, Geophys. J. Int. 128 (1997) 689–700

32. H. Zhou: Multi-scale traveltime tomography, Geophysics. 68 (2003) 1639–1649
33. L. Chiao and B. Kuo: Multiscale seismic tomography, Geophys. J. Int. 145 (2001) 517–527
34. G. McMechan: Seismic tomography in boreholes, Geophys. J. R. Astr. Soc. 74 (1983) 298–308
35. J. Peterson, B. Paulsson and T. McEvilly: Application of algebraic reconstruction techniques to crosshole seismic data, Geophysics 50 (1985) 1566–1580
36. N. D. Bregman, R. C. Bailey and C. H. Chapman: Crosshole seismic tomography, Geophysics 54 (1989) 200–215
37. Menke: The resolving power of cross-borehole tomography, Geophys. Res. Lett. 11 (1984) 105–108
38. T.W. Lo and P. Inderwiesen: Fundamentals of Seismic Tomography, (Society of Exploration Geophys., Tulsa OK 1994)
39. R. G. Pratt and C. H . Chapman: Traveltime tomography in anisotropic media II. Application, Geophys. J. Int. 109 (1992) 20–37
40. R. G. Pratt, W. J. McGaughey and C. H. Chapman: Anisotropic velocity tomography: A case study in a near-surface rock mass, Geophysics 58 (1993) 1748–1763
41. D. J. White: Two-dimensional seismic refraction tomography, Geophys. J. 97 (1989) 223–245
42. W. J. Lutter, R. L. Nowack and L. W. Braile: Seismic imaging of the upper crustal velocity structure using travel times from the PASSCAL Ouachita experiment, J. Geophys. Res. 95 (1990) 4621–4631
43. C. A. Zelt and R. B. Smith: Seismic travel time inversion for 2D crustal velocity structure, Geophys. J. Int. 108 (1992) 16–34
44. W. J. Lutter and R. L. Nowack: Inversion for crustal structure using reflections from the PASSCAL Ouachita experiment, J. Geophys. Res. 95 (1990) 4633–4646
45. J. A. Hole: Nonlinear high-resolution three-dimensional travel-time tomography, J. Geophys. Res. 97 (1992) 6553–6562
46. C. A. Zelt and P. J. Barton: Three-dimensional seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin, J. Geophys. Res. 103 (1998) 7187–7210
47. R. L. Nowack and L. W. Braile: Refraction and wide-angle reflection tomography: theory and results, In: Seismic Tomography: Theory and Practice ed. by H. M. Iyer and K. Hirahara, (Chapman and Hall, London 1993)
48. T. N. Bishop, K. P. Bube, R. T. Cutler, R. T. Langan, P. L. Love, J. R. Resnick, R. T. Shuey, D. A. Spindler and H. W. Wyld: Tomographic determination of velocity and depth in laterally varying media, Geophysics 50 (1985) 903–923
49. V. Farra and R. Madariaga: Non-linear reflection tomography, Geophys. J. 95 (1988) 135–147
50. P. R. Williamson: Tomographic inversion in reflection seismology, Geophys. J. Int. 100 (1990) 255–274

51. C. Stork: Singular value decomposition of the velocity-reflector depth tradeoff, part 2: High-resolution analysis of a generic model, Geophysics 57 (1992) 933–943
52. F. Billette and G. Lambare: Velocity macro-model estimation by stereotomography, Geophys. J. Int. 135 (1998) 671–680
53. G. Lambare: Stereotomography: where we are, and where should we go, SEG extended abstracts for the 74th annual meeting (2004) 2367–2370
54. P. Gasparini and the TomoVes Working Group: Looking inside Mt Vesuvius, Trans. Am. Geophys. Un. (EAS) 79 (1998) 229–231
55. R. L. Nowack, C. Li and J. Virieux: Inversion of seismic attributes from the 1996 3-D tomography experiment at Mt. Vesuvius, Italy, Trans. Am. Geophys. Un. (EOS) 81 (2000) 903
56. R. L. Nowack, C. Li and J. Virieux: Inversion of seismic attributes from the 1996 3-D tomography experiment at Mt. Vesuvius, Italy, In: The TomoVes Project, Looking Inside Mt. Vesuvius eds. A. Zollo, A. Bobbio, P. Gasparini, R. Casale and M. Yeroyanni, (CD Book, CUEN, Naples, Italy 2003)
57. A. Zollo, L. D'Auria, R. De Matteis, A. Herrero, J. Virieux and P. Gasparini: Bayesian estimation of 2-D P-velocity models from the active seismic arrival time data: imaging of the shallow structure of Mt. Vesuvius (Southern Italy), Geophys. J. Int. 151 (2002) 566–582
58. C. H. Thurber: Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, J. Geophys. Res. 88 (1983) 8225–8236.
59. E. Kissling: Geotomography with local earthquake data, Rev. of Geophys. 26 (1988) 659–698
60. D. Eberhart-Phillips: Three-dimensional P- and S-velocity structure in the Coalinga region, California, J. Geophys Res. 95 (1990) 15, 343–14363
61. K. Hirahara: Tomography using both local earthquakes and teleseisms: velocity and anisotropy – theory, In Seismic Tomography: Theory and Practice Ed. by H. M. Iyer and K. Hirahara (Chapman and Hall, London 1993)
62. H. Zhang and C. H. Thurber: Double-difference tomography: The method and its application to the Hayward Fault, California, Bull. Seism. Soc. Am. 93 (2003) 1875–1889
63. R. W. Clayton and R. P. Comer: A tomographic analysis of mantle heterogeneities from body wave travel time, Trans. Am. Geophys. Un. (EOS) 64 (1983) 776
64. A. M. Dziewonski and D. L. Anderson: Mapping the lower mantle: determination of lateral heterogeneity in P velocity up to degree and order 6, J. Geophys. Res. 89 (1984) 5929–5952
65. J. H. Woodhouse and A. M. Dziewonski: Mapping of the upper mantle: three-dimensional modeling of earth structure by inversion of seismic waveforms, J. Geophys. Res. 89 (1984) 5953–5986
66. B. L. N. Kennett and R. Van der Hilst: Seismic structure of the mantle: from subduction zone to craton, In The Earth's Mantle Ed. by I. Jackson (Cambridge Univ. Press, 1998)

67. B. Romanowicz: Global mantle tomography: progress status in the past 10 years, Ann. Rev. Earth Planet. Sci. 31 (2003) 303–328
68. A. Curtis and R. Snieder: Probing the Earth's interior with seismic tomography, In: International Handbook of Earthquake and Engineering Seismology (Academic Press, San Diego 2002)
69. M. Cara: Seismic Anisotropy, In: International Handbook of Earthquake and Engineering Seismology (Academic Press, San Diego 2002) 875–885
70. C. J. Thomson: Ray-theoretical amplitude inversion for laterally varying velocity structure below NORSAR, Geophys. J. R. Astr. Soc. 74 (1983) 525–558
71. F. Neele, J. VanDecar and R. Snieder: The use of P wave amplitude data in a joint inversion with travel times for upper mantle velocity structure, J. Geophys. Res. 98 (1993) 12,033–12,482
72. N Bregman, C. H. Chapman and R. C. Bailey: Travel time and amplitude analysis in seismic tomography, J. Geophys. Res. 94 (1989) 7577–7587
73. R. L. Nowack and M. P. Matheney: Inversion of seismic attributes for velocity and attenuation structure, Geophys. J. Int. 128 (1997) 689–700
74. B. J. Mitchell and B. Romanowicz (Eds.): Q of the Earth: Global, Regional and Laboratory Studies, (Birhhauser Boston 1999)
75. Y. Wang: Seismic Amplitude Inversion in Reflection Tomography, (Pergamon, Amsterdam 2003)
76. A. J. Devaney: Geophysical diffraction tomography, IEEE Trans. Geos. GE-22 (1984) 3–13
77. R. S. Wu and M. N. Toksoz: Diffraction tomography and multi-source holography applied to seismic imaging, Geophysics 52 (1987) 11–25
78. R. G. Pratt and M. H. Worthington: Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave-equation method, Geophys. Prosp. 38 (1990) 287–310
79. R. G. Pratt, Z. M. Song, P. Williamson and M. Warner: Two-dimensional velocity models from wide-angle seismic data by wavefield inversion, Geophys. J. Int. 124 (1996) 323–340
80. A. Brenders and R. G. Pratt: Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model. Geophys. J. Int. 168 (2007) 133–151.
81. M. J. Woodward: Wave equation tomography, Geophysics 57 (1992) 15–26
82. K. Yomogida: Fresnel zone inversion for lateral heterogeneities in the Earth, Pure Appl. Geophys. 138 (1992) 391–406
83. F. A. Dahlen, S. H. Hung and G. Nolet: Frechet kernels for finite frequency traveltimes – I. Theory, Geophys. J. Int. 141 (2000) 157–174
84. Montelli, G. Nolet, G. Masters, F. A. Dahlen and S. H. Hung: Global P and PP traveltime tomography: rays versus rays, Geophys. J. Int. 158 (2004) 637–654

85. van der Hilst, R. D. and M. V. de Hoop: Banana-doughnut kernels and mantle tomography, Geophys. J. Int. 163 (2005) 956–961
86. Montelli, R. G. Nolet and F. A. Dahlen: Comment on 'Banana-doughnut kernels and mantle tomography' by van der Hilst and de Hoop, Geophys. J. Int. 167 (2006) 1204–1210.
87. R. van der Hilst, and M. de Hoop: Reply to comment by R. Montelli, G. Nolet and F. A. Dahlen on 'Banana-doughnut kernels and mantle tomography', Geophys. J, Int. 167 (2006) 1211–1214.