JGR Solid Earth

RESEARCH ARTICLE
10.1029/2021JB021818

Key Points:

- Seismic traveltime tomography is reformulated and solved by the adjoint-state method
- One eikonal equation-based earthquake location method is derived to solve the coupled hypocenter-velocity problem
- Crustal fluids may have played critical roles in the genesis of the 2020 Mw 4.9 and other moderate-sized Anza earthquakes

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

P. Tong,
tongping@ntu.edu.sg

Citation:

Tong, P. (2021). Adjoint-state traveltime tomography: Eikonal equation-based methods and application to the Anza area in southern California. *Journal of Geophysical Research: Solid Earth*, *126*, e2021JB021818. https://doi.org/10.1029/2021JB021818

Received 29 JAN 2021
Accepted 25 APR 2021

Adjoint-State Traveltime Tomography: Eikonal Equation-Based Methods and Application to the Anza Area in Southern California

Ping Tong[1,2,3]

[1]Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, Singapore, [2]Asian School of the Environment, Nanyang Technological University, Singapore, Singapore, [3]Earth Observatory of Singapore, Nanyang Technological University, Singapore, Singapore

Abstract Ray tracing is avoidable in seismic traveltime tomography. In this study, seismic traveltime tomography is reformulated as an eikonal equation-constrained optimization problem solved by the ray-free adjoint-state method. The resultant approach is called adjoint-state traveltime tomography. For completeness, an eikonal equation-based earthquake location method is developed to locate the hypocenters of earthquakes when necessary. The multiple-grid model parameterization is adopted to discretize the relative slowness perturbation $\Delta s(\mathbf{x}) / s(\mathbf{x})$. The step-size-controlled gradient descent method, which has an effect of damping regularization, is used to find optimal earthquake hypocenters and velocity models. The performances of the earthquake location and adjoint-state traveltime tomography methods are tested in the source area of the 2020 M_w 4.9 Anza earthquake, a seismologically active place in southern California. The obtained high-resolution *P*-wave velocity model demonstrates that the 2020 M_w 4.9 Anza earthquake and other historic moderate-sized Anza earthquakes occurred near or at the boundaries of low V_P rocks, a typical seismogenic environment for moderate to large crustal earthquakes. Meanwhile, the source zone of the Anza earthquakes is characterized by high V_P/V_S, indicating the possible existence of crustal fluids as well as the critical role of fluids in the occurrence of those moderated-sized Anza earthquakes. On the whole, the well-performed eikonal equation-based earthquake location method and adjoint-state traveltime tomography method provide competent and attractive tools for hypocenter and tomographic inversions.

1. Introduction

Seismic traveltime tomography is one of the primary and routine techniques to estimate subsurface material properties. It has been contributing significantly to the understanding of heterogeneous structures and dynamic processes in the Earth's interior ever since the pioneer work of Aki and Lee (1976). Various seismic traveltime tomography methods are rooted in different mathematical physical models. In principle, we can categorize them into three major types: Ray-based traveltime tomography, wave equation-based traveltime tomography, and eikonal equation-based traveltime tomography.

Ray tracing is the key ingredient for seismic traveltime tomography methods based on ray theory (e.g., Aki & Lee, 1976; Thurber, 1983; Zhao et al., 1992). During the past few decades, plenty of seismic ray tracing techniques have been developed, including the widely used two-point ray tracing methods (e.g., Gou et al., 2018; Koketsu & Sekine, 1998; Um & Thurber, 1987) and the grid-based approaches (e.g., Liu et al., 2018; Rawlinson & Sambridge, 2004; Tong et al., 2017). As pointed out by Rawlinson et al. (2008), those two-point ray tracing methods may fail to converge to a true two-point path even in mildly heterogeneous media. The grid-based approaches have become popular in recent seismic tomographic inversions (e.g., Li et al., 2020; Liu et al., 2019; Zenonos et al., 2020). They first solve the eikonal equation numerically to find the traveltime from the source to every grid point, and then trace the raypath from the receiver to source along the negative traveltime gradient. The prominent advantage of the grid-based approaches is that accurate traveltime fields can be obtained even in complex heterogeneous media (Rawlinson et al., 2010). However, ray paths to be predicted do not directly rely on traveltime but traveltime gradient, which is estimated from the numerical solution of traveltime.

Wave equation-based traveltime tomography usually inverts cross-correlation traveltime measurements to infer subsurface velocity structure (e.g., Tong et al., 2014a; Tromp et al., 2005). Finite-frequency efforts of seismic wave propagation and off-ray structure influence on traveltime observations can be properly modeled. In contrast to simplified ray approximation, wave equation-based seismic tomography explores not only traveltime information but also amplitude, attenuation, generalized seismological data functionals, and even full waveform content to image subsurface structures (e.g., Chen et al., 2007; Fichtner et al., 2006; Tromp et al., 2005; Virieux & Operto, 2009). High-resolution images at half a wavelength scale are expected (Virieux & Operto, 2009). Although significant progress has been made, this type of seismic tomography remains to be computationally expensive and in most cases only uses more stable low-frequency data. Efforts must continue to release the full potential of the promising wave equation-based seismic tomography.

Under the high-frequency assumption, traveltime fields in smoothly varying velocity models can be described by the eikonal equation (Rawlinson et al., 2008). There are basically three different kinds of traveltime tomography methods related to the eikonal equation. The first kind has been mentioned earlier, which solves the eikonal equation to obtain the traveltime field and then finds the raypath based on the traveltime gradient. The second kind uses the Taylor series expansion to approximate a linear relation between the traveltime functional and slowness model (Ammon & Vidale, 1993). By slightly perturbing the slowness of one cell and recalculating the traveltime functional, the derivative of the traveltime functional with respect to the slowness of that cell is computed using a finite-difference scheme (Ammon & Vidale, 1993). But the brute-force way of evaluating the traveltime functional gradient is computationally prohibitive when a large number of cells are required to parameterize the model (Lelievre et al., 2011). The eikonal equation-based traveltime tomography of the third kind uses the adjoint-state method to compute the derivative of the traveltime functional with respect to slowness (or velocity) (e.g., Sei & Symes, 1994; Leung & Qian, 2006; Taillandier et al., 2009; Tavakoli et al., 2017). Once the traveltime field is computed, the functional derivative can be obtained by additionally solving an adjoint eikonal equation. The adjoint-state method is computationally favorable and has gained increasing popularity in recent years. Moreover, both the second and third kinds of methods advocate the avoidance of ray tracing in seismic tomographic inversions, which is attractive.

Seismic traveltime tomography has been widely used in earthquake seismology. To date, it is mainly ray-based or wave equation-based. To take advantage of the simple but precise mathematical physical model described by the eikonal equation and the efficacy of the adjoint-state method, we reformulate seismic traveltime tomography as an eikonal equation-constrained optimization problem solved by the adjoint-state method. The derivation of sensitivity kernel, multiple-grid model parameterization and optimization algorithm are discussed in details in Section 2 and Appendix A. The reformulated seismic traveltime tomography method is named adjoint-state traveltime tomography.

As is known, earthquake hypocenters and velocity structure are coupling factors that affect seismic-wave arrival times (Thurber, 1992). To construct an accurate velocity model, earthquake locations should be precisely determined in advance or simultaneously. As one of the oldest and most fundamental problems in seismology, a lot of methods have been developed to locate earthquakes since the beginning of the instrumental seismic recording in the late of the nineteenth century (Tong et al., 2016). More detailed reviews of earthquake location methods can be found in Tong et al. (2016) and references therein. In particular, Geiger's method is one of the most widely used earthquake location methods (e.g., Geiger, 1910; Thurber, 1983, 1992; Tong et al., 2016). To use this technique, the ray direction at the source location needs to be correctly measured. Again, the accuracy of raypath affects the performance of the Geiger's earthquake location method. To avoid ray tracing and to gain high accuracy in earthquake locations, an eikonal equation-based earthquake location method is developed in Section 3.

In Section 4, the eikonal equation-based earthquake location method and the adjoint-state traveltime tomography method are applied to the April 2020 M_w 4.9 Anza earthquake area in southern California. The obtained high-resolution V_P and V_P/V_S models allow us to investigate the relationship between crustal velocity heterogeneity and earthquake occurrence in the source zone of Anza earthquakes and provide some hints on such a relationship in other seismologically active areas.

2. Adjoint-State Traveltime Tomography

Earthquakes occur inside an Earth volume Ω at the locations $\mathbf{x}_{s,n}$ ($n = 1, 2, ..., N$). The traveltime of the wavefront arriving at any position $\mathbf{x}$ corresponding to the n-th earthquake is governed by the eikonal equation with a boundary condition as

$$\nabla T_n(\mathbf{x}) \cdot \nabla T_n(\mathbf{x}) = s^2(\mathbf{x}), \quad T_n(\mathbf{x}_{s,n}) = 0. \tag{1}$$

$s(\mathbf{x})$ is the P-wave or S-wave slowness (the reciprocal of wave speed). For the eikonal problem 1, the boundary $\partial\Omega$ includes the Earth surface and earthquake location $\mathbf{x}_{s,n}$. M seismic stations $\mathbf{x}_{r,m}$ ($m = 1, 2, ... , M$) deployed near the Earth surface record the first P-wave (S-wave) arrival times of these earthquakes. The traveltime observations $T_n^o(\mathbf{x}_{r,m})$ allow us to find an optimal slowness model by minimizing the objective function

$$\chi\Big(s(\mathbf{x})\Big) = \sum_{n=1}^{N}\sum_{m=1}^{M}\frac{\omega_{n,m}}{2}\Big[T_n(\mathbf{x}_{r,m}) - T_n^o(\mathbf{x}_{r,m})\Big]^2. \tag{2}$$

The weight $\omega_{n,m}$ is assigned zero if there is no record of the n-th earthquake at seismic station $\mathbf{x}_{r,m}$ or the data quality is low. Usually, the derivative of the objective function is required to find its minimum.

2.1. Fréchet Derivative

To derive an expression for the Fréchet derivative of the objective function $\chi\Big(s(\mathbf{x})\Big)$ with respect to the slowness, we consider that there is an infinitesimal perturbation $\delta s(\mathbf{x})$ in slowness $s(\mathbf{x})$. All the second-order terms and beyond are ignored in the mathematical derivation below for the first-order Fréchet derivative.

Due to the slowness perturbation $\delta s(\mathbf{x})$, we observe an infinitesimal perturbation $\delta T_n(\mathbf{x})$ in the traveltime field $T_n(\mathbf{x})$; accordingly, the eikonal equation with the boundary condition in the perturbed slowness model $s(\mathbf{x}) + \delta s(\mathbf{x})$ becomes

$$\nabla\Big[T_n(\mathbf{x}) + \delta T_n(\mathbf{x})\Big] \cdot \nabla\Big[T_n(\mathbf{x}) + \delta T_n(\mathbf{x})\Big] = \Big[s(\mathbf{x}) + \delta s(\mathbf{x})\Big]^2, \quad T_n(\mathbf{x}_{s,n}) + \delta T_n(\mathbf{x}_{s,n}) = 0. \tag{3}$$

The traveltime perturbation $\delta T_n(\mathbf{x})$ also causes a change in the objective function of Equation 2, which can be expressed by

$$\begin{aligned}\delta\chi\Big(s(\mathbf{x})\Big) &= \sum_{n=1}^{N}\sum_{m=1}^{M}\omega_{n,m}\Big[T_n(\mathbf{x}_{r,m}) - T_n^o(\mathbf{x}_{r,m})\Big]\delta T_n(\mathbf{x}_{r,m}) \\ &= \sum_{n=1}^{N}\int_\Omega\sum_{m=1}^{M}\omega_{n,m}\Big[T_n(\mathbf{x}) - T_n^o(\mathbf{x}_{r,m})\Big]\delta T_n(\mathbf{x})\delta(\mathbf{x} - \mathbf{x}_{r,m})d\mathbf{x}.\end{aligned} \tag{4}$$

Note that $\delta(\mathbf{x} - \mathbf{x}_{r,m})$ denotes the Dirac delta function, while $\delta\chi\Big(s(\mathbf{x})\Big)$ and $\delta T_n(\mathbf{x})$ are infinitesimal numbers. Subtracting Equation 1 from Equation 3, we obtain

$$\nabla\Big[T_n(\mathbf{x})\Big] \cdot \nabla\Big[\delta T_n(\mathbf{x})\Big] = s(\mathbf{x})\delta s(\mathbf{x}), \quad \delta T_n(\mathbf{x}_{s,n}) = 0. \tag{5}$$

Multiply an arbitrary test function $P_n(\mathbf{x})$ on both sides of Equation 5 and integrate the Earth volume Ω,

$$\int_\Omega P_n(\mathbf{x})\nabla\Big[T_n(\mathbf{x})\Big] \cdot \nabla\Big[\delta T_n(\mathbf{x})\Big]d\mathbf{x} = \int_\Omega P_n(\mathbf{x})s(\mathbf{x})\delta s(\mathbf{x})d\mathbf{x}. \tag{6}$$

The product rule for differentiation implies that

$$P_n(\mathbf{x})\nabla\Big[T_n(\mathbf{x})\Big] \cdot \nabla\Big[\delta T_n(\mathbf{x})\Big] = \nabla \cdot \Big[\delta T_n(\mathbf{x})P_n(\mathbf{x})\nabla T_n(\mathbf{x})\Big] - \delta T_n(\mathbf{x})\nabla \cdot \Big[P_n(\mathbf{x})\nabla T_n(\mathbf{x})\Big]. \tag{7}$$

Thus, by using the divergence theorem Equation 6 can be rewritten as

$$\int_{\partial\Omega}\mathbf{n} \cdot \Big[\delta T_n(\mathbf{x})P_n(\mathbf{x})\nabla T_n(\mathbf{x})\Big]d\mathbf{x} + \int_\Omega \delta T_n(\mathbf{x})\nabla \cdot \Big[P_n(\mathbf{x})(-\nabla T_n(\mathbf{x}))\Big]d\mathbf{x} = \int_\Omega P_n(\mathbf{x})s(\mathbf{x})\delta s(\mathbf{x})d\mathbf{x}. \tag{8}$$

To relate $\delta\chi(s(\mathbf{x}))$ in Equation 4 with $\delta s(\mathbf{x})$ in Equation 8, we make an assumption about the test function $P_n(\mathbf{x})$ that

$$\nabla\cdot\Big[P_n(\mathbf{x})\nabla\big(-T_n(\mathbf{x})\big)\Big]=\sum_{m=1}^{M}\omega_{n,m}\Big[T_n(\mathbf{x})-T_n^o(\mathbf{x}_{r,m})\Big]\delta(\mathbf{x}-\mathbf{x}_{r,m}). \tag{9}$$

If the homogeneous boundary condition is further assumed

$$P_n(\mathbf{x})=0,\quad \mathbf{x}\in\partial\Omega, \tag{10}$$

then the first integral on the left hand side of Equation 8 vanishes. After that, Equation 8 can be expressed by

$$\int_\Omega \delta T_n(\mathbf{x})\sum_{m=1}^{M}\omega_{n,m}\Big[T_n(\mathbf{x})-T_n^o(\mathbf{x}_{r,m})\Big]\delta(\mathbf{x}-\mathbf{x}_{r,m})d\mathbf{x}=\int_\Omega P_n(\mathbf{x})s(\mathbf{x})\delta s(\mathbf{x})d\mathbf{x}. \tag{11}$$

Comparing Equations 11 with 4, we can obtain an expression for the Fréchet derivative of the objective function in terms of the variation in slowness $s(\mathbf{x})$,

$$\delta\chi\big(s(\mathbf{x})\big)=\sum_{n=1}^{N}\int_\Omega P_n(\mathbf{x})s(\mathbf{x})\delta s(\mathbf{x})d\mathbf{x}. \tag{12}$$

It is a common practice to invert for the relative slowness (velocity) perturbation $\delta s(\mathbf{x})/s(\mathbf{x})$ in seismic tomographic inversions. According to Equation 12, the perturbation of the objective function is related to the relative slowness perturbation as

$$\delta\chi\big(s(\mathbf{x})\big)=\int_\Omega K_s(\mathbf{x})\frac{\delta s(\mathbf{x})}{s(\mathbf{x})}d\mathbf{x}, \tag{13}$$

where

$$K_s(\mathbf{x})=\sum_{n=1}^{N}K_{s,n}(\mathbf{x})=\sum_{n=1}^{N}P_n(\mathbf{x})s^2(\mathbf{x}). \tag{14}$$

$K_s(\mathbf{x})$ is defined as the misfit kernel, representing the Fréchet derivative of the objective function $\chi(\mathbf{x})$ with respect to the slowness $s(\mathbf{x})$ (Tromp et al., 2005); whereas $K_{s,n}(\mathbf{x})=P_n(\mathbf{x})s^2(\mathbf{x})$ is the event kernel associated with the n-th earthquake (Tape et al., 2007). Specifically, if only a pair of the n-th earthquake and m-th seismic station is considered, the corresponding sensitivity kernel $K_{s,n,m}(\mathbf{x})=P_n(\mathbf{x})s^2(\mathbf{x})$ is named individual kernel.

We now give a physical interpretation of Equation 9. The traveltime residual $T_n(\mathbf{x}_{r,m})-T_n^o(\mathbf{x}_{r,m})$ measured at every receiver $\mathbf{x}_{r,m}$ is caused by the discrepancy between the true velocity model and the velocity model where $T_n(\mathbf{x}_{r,m})$ is computed. It is gradually accumulated as the wave propagates outward from the source to receiver. The right hand side of Equation 9 is the sum of the traveltime residuals at all the receivers, acting as a source term for the equation itself. $\nabla\cdot\Big[P_n(\mathbf{x})\nabla\big(-T_n(\mathbf{x})\big)\Big]$ on the left hand side describes the transport of a substance or quantity. Here is the transport of the traveltime residuals. Sourcing at the receivers, the magnitude and direction of the transport follow the vector $\mathbf{v}=\nabla\big(-T_n(\mathbf{x})\big)$. The density of the transported traveltime residuals is modeled by the scalar field $P_n(\mathbf{x})$, while the flux $P_n(\mathbf{x})\nabla\big(-T_n(\mathbf{x})\big)$ denotes the backward "loss" of the accumulated traveltime residuals through a unit area that is normal to the direction $\mathbf{v}=\nabla\big(-T_n(\mathbf{x})\big)$. Based on seismic ray theory, $\nabla\big(-T_n(\mathbf{x})\big)$ is the direction opposite to seismic wave propagation. Thus, Equation 9 describes the transport of traveltime residuals that starts at the receivers and ends at the source location $\mathbf{x}_{s,n}$. Solving Equation 9 yields the density function $P_n(\mathbf{x})$, which can also be interpreted as the contribution of every single point to the accumulated traveltime residuals at the receivers. The product of the density function $P_n(\mathbf{x})$ and slowness squared $s^2(\mathbf{x})$ is the sensitivity kernel $K_s(\mathbf{x})$, which measures the single point contribution of strength factor for model perturbation. In analogy to adjoint wave equation which back-projects observational misfits at receivers to the earthquake location (Fichtner et al., 2006;

Tromp et al., 2005), Equation 9 is called adjoint eikonal equation, and $P_n(\mathbf{x})$ is an adjoint traveltime field sourced by the traveltime residuals at the receivers.

It is worth mentioning that the mathematical derivation of the sensitivity kernel $K_{s,n}(\mathbf{x})$ in Equation 14 provides an alternative to the common Lagrangian formulation or penalization technique used by previous similar studies (e.g., Sei & Symes, 1994; Taillandier et al., 2009; Tavakoli et al., 2017). Different from Leung and Qian (2006), Taillandier et al. (2009) and some others, who obtained a homogeneous adjoint eikonal equation but assumed that all the receivers were exactly on the boundary, we derived an inhomogeneous adjoint eikonal equation allowing the receivers to be anywhere in the study area. The strong requirement for the receivers to be located on the boundary limits the wide applications of the corresponding seismic tomography method. On the contrary, the inhomogeneous adjoint eikonal equation can be applied to any configurations of seismic stations, indicating the possible wide applications in the future.

We show examples of individual kernel, event kernel, and misfit kernel in Figure 1. To compute the kernels, we use the fast sweeping method to solve the eikonal Equation 1 first to get the traveltime field $T_n(\mathbf{x})$ and the adjoint eikonal Equation 9 then to obtain the adjoint traveltime field $P_n(\mathbf{x})$ (Zhao, 2004; Leung & Qian, 2006). All the kernels are computed in a simple vertically increasing velocity model $v(\mathbf{x})$ (Figure 1a). This is achieved after computing the observed traveltimes in a velocity model that has alternating positive and negative perturbations (Figure 1b) from the simple velocity model. It is observable that every individual kernel is perpendicular to the wavefronts all the way from the source to receiver (Figures 1c and 1d). The sign of each individual kernel depends on the positive or negative traveltime residual measured at the seismic station. Theoretically, individual kernel is infinitely thin. Due to the numerical approximation on a discretized spatial grid, the individual kernel is typically a few times as wide as the grid interval. The event kernel is simply the sum of all the individual kernels corresponding to the same event (Figure 1e). Further, the sum of all the event kernels is the misfit kernel (Figure 1f), which suggests the proper direction to update the slowness model. A comparison between Figures 1b and 1f clearly illustrates the relationship between the misfit kernel and the velocity perturbation.

2.2. Model Parameterization

Model parameterization is one of the key steps in seismic tomographic inversion (Rawlinson et al., 2010; Tong et al., 2019). Among various options including the grid approach, block approach, spectral parameterization and irregular parameterization (Thurber & Eberhart-Philips, 1999; Tong et al., 2019), we use the trilinear interpolation on a regular grid to represent the relative slowness perturbation $\delta s(\mathbf{x})/s(\mathbf{x})$, considering its easy implementation and wide applications in previous studies (e.g., Thurber, 1983; Zhao et al., 1992; Tong et al., 2019). The regular grid is called the inversion grid, whose nodes are denoted by (X_i, Y_j, Z_k) $(1 \le i \le N_x, 1 \le j \le N_y, 1 \le k \le N_z)$ (Figure 2). Due to the limited resolving ability of seismic data, the inversion grid is usually set to be coarser than the forward grid on which both the eikonal equation and adjoint eikonal equation are solved. Every grid node (X_i, Y_j, Z_k) is associated with one basis function $B_l(\mathbf{x}) = B_l(x, y, z)$ $(l = (k-1)N_xN_y + (j-1)N_x + i)$ and vice versa (Figure 2a). In detail, the basis function $B_l(\mathbf{x}) = B_l(x, y, z)$ is defined as

$$B_l(x, y, z) = w_i(x)w_j(y)w_k(z), \tag{15}$$

where

$$w_i(x) = \begin{cases} \dfrac{x - X_{i-1}}{X_i - X_{i-1}}, & \text{if } X_1 \le X_{i-1} \le x \le X_i, \\ \dfrac{X_{i+1} - x}{X_{i+1} - X_i}, & \text{if } X_i \le x \le X_{i+1} \le X_{N_x}, \\ 0, & \text{otherwise,} \end{cases} \tag{16}$$

and similarly for $w_j(y)$ and $w_k(z)$.

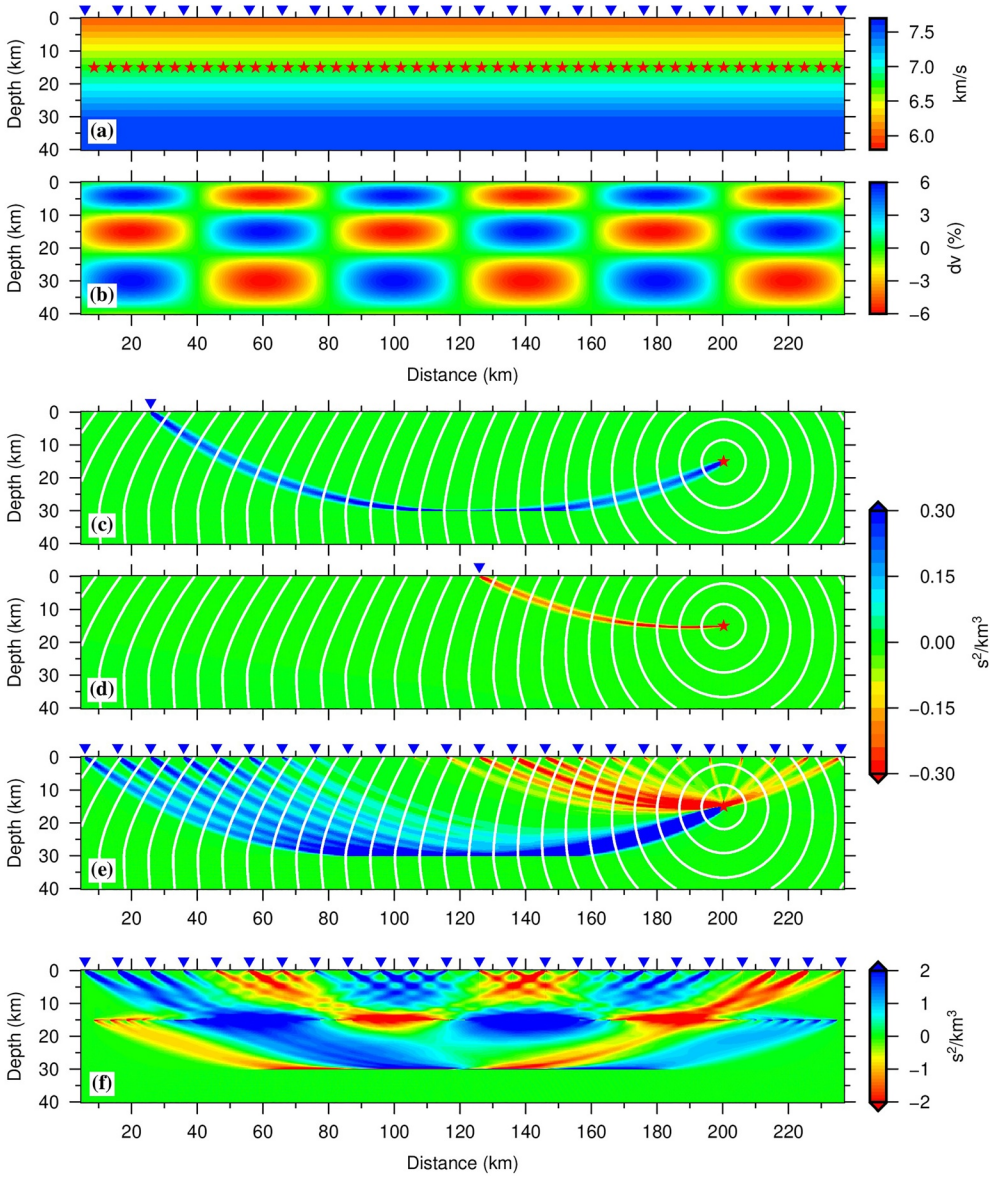


Figure 1. (a) The vertically increasing velocity model $v(\mathbf{x})$. Mathematically it can be expressed by $v(\mathbf{x}) = v(x, y, z) = \min(6.0 + 0.05z, 7.5)$. The red stars and blue inverse triangles represent 47 earthquakes and 24 seismic stations, respectively. (b) The relative velocity perturbation $\Delta v(\mathbf{x})/v(\mathbf{x})$. The observed traveltimes are computed in the perturbed velocity model $v(\mathbf{x}) + \Delta v(\mathbf{x})$. (c) and (d) Two individual kernels corresponding to the same earthquake but two different seismic stations. (e) An example of event kernel. (f) The misfit kernel related to all the 47 earthquakes and 24 seismic stations. The white curves in (c), (d) and (e) are traveltime isochrones at an interval of 1 s. All the sensitivity kernels with respect to slowness in (c)–(f) are computed in $v(\mathbf{x})$. The color scales for velocity, relative velocity perturbation and sensitivity kernels are on the right hand side of the respective plots. (c)–(e) share the same color scale.

The relative slowness perturbation $\delta s(\mathbf{x})/s(\mathbf{x})$ can be written as a linear combination of the basis functions $B_l(\mathbf{x})$ $(l = 1, ..., L; L = N_x N_y N_z)$

$$\frac{\delta s(\mathbf{x})}{s(\mathbf{x})} = \sum_{l=1}^{L} \delta C_l B_l(\mathbf{x}), \tag{17}$$

where the coefficients δC_l are to be determined. Equations 15–17 indicate that the relative slowness perturbation at any point $\mathbf{x} = (x, y, z)$ is a weighted average of the values at the eight grid points surrounding $\mathbf{x}$ (Figure 2b).

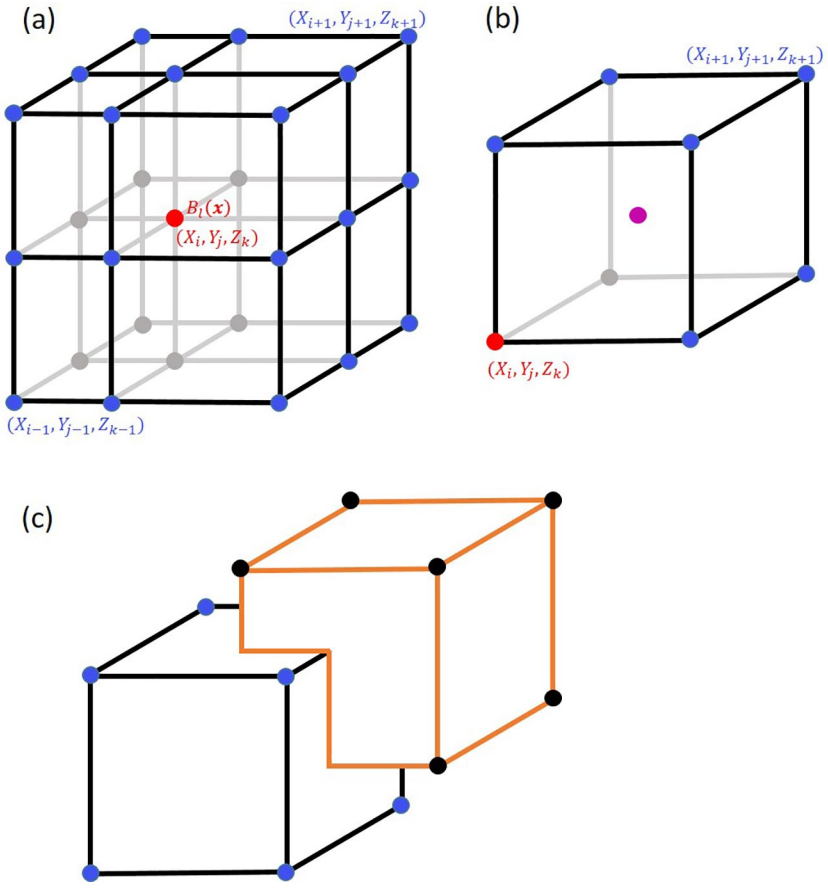


Figure 2. The grid approach for model parameterization. (a) A single regular grid. The red grid point (X_i, Y_j, Z_k) is associated with the basis function $B_l(\mathbf{x})$. (b) An arbitrary point $\mathbf{x}$ (in purple) and the eight surrounding points. (c) An example of multiple-grid. Two regular grids are placed in a staggered way to form the multiple-grid.

The selection of the inversion grid involves subjectivity, which in turn affects the final results. In comparison with the commonly used grid approach, the multiple-grid model parameterization has been proven to be able to produce better or at least comparable results (Tong et al., 2019). Hence, rather than using a single regular inversion grid as most previous studies did, we approximate the relative slowness perturbation $\delta s(\mathbf{x})/s(\mathbf{x})$ on multiple sets of regular grids placed in a staggered way (Figure 2c). Considering that the relative slowness perturbation $\delta s(\mathbf{x})/s(\mathbf{x})$ can be interpolated on every regular grid without preference, we simply use the average of all the interpolated functions. That is,

$$\frac{\delta s(\mathbf{x})}{s(\mathbf{x})} = \frac{1}{H}\sum_{h=1}^{H}\sum_{l=1}^{L_h}\delta C_{l,h}B_{l,h}(\mathbf{x}), \tag{18}$$

where $B_{l,h}(\mathbf{x})$ is the l-th basis function ($l = 1, \ldots, L_h$) defined on the h-th regular grid ($h = 1, \ldots, H$). Plugging Equation 18 in Equation 13, we obtain

$$\frac{\partial \chi\left(s(\mathbf{x})\right)}{\partial C_{l,h}} = \sum_{n=1}^{N}\frac{1}{H}\int_{\Omega} P_n(\mathbf{x})s^2(\mathbf{x})B_{l,h}(\mathbf{x})d\mathbf{x}. \tag{19}$$

So far, the optimization problem 2 in an infinite-dimensional space has been reduced to a minimization problem in a finite-dimensional space. Given the gradient function (19), a set of optimal coefficients ($C_{1,1}$, $C_{2,1}$,…, $C_{L_1,1}$, $C_{1,2}$,…, $C_{L_H,H}$) can be found by using the method of gradient descent, conjugate gradient method or some other variants. A slightly modified gradient descent method which prefers a

magnitude-controlled step size is used in this study (Appendix A). As a whole, the eikonal equation-based inversion method described in this and the previous subsection is named adjoint-state traveltime tomography.

3. Eikonal Equation-Based Earthquake Location

Inaccuracy in earthquake locations may distort dramatically seismic velocity model (Thurber, 1992). To provide accurate earthquake locations for adjoint-state traveltime tomography, an eikonal equation-based earthquake location method is developed in this section.

3.1. Method

We treat seismic stations as virtual sources and earthquakes as virtual receivers. Every virtual source occurs at the same reference time τ_0 (the time on an accurate clock) and a respective known location $\mathbf{x}_{r,m}$ ($m = 1, 2, \dots, M$). The locations of the virtual receivers $\mathbf{x}_{s,n}$ ($n = 1, 2, \dots, N$) need to be determined. We assume that there is a clock at every virtual receiver $\mathbf{x}_{s,n}$, which reads τ_n at the reference time τ_0. But τ_n is unknown. When the seismic wave originating from the virtual source $\mathbf{x}_{r,m}$ arrives at the virtual receiver $\mathbf{x}_{s,n}$, the clock reads $\tau^o_{m,n}$. Given the validity of the reciprocity principle (Aki & Richards, 2002), the above scenario is equivalent to that the first arrival from an unknown hypocenter ($\mathbf{x}_{s,n}$, τ_n) arrives at the seismic station $\mathbf{x}_{r,m}$ at the time $\tau^o_{m,n}$. To relocate the n-th virtual receiver, we attempt to minimize the objective function

$$\chi\left(\mathbf{x}_{s,n}, \tau_n\right) = \sum_{m=1}^{M} \frac{\omega_{m,n}}{2}\left[T_m(\mathbf{x}_{s,n}) - \left(\tau^o_{m,n} - \tau_n\right)\right]^2, \tag{20}$$

where $T_m(\mathbf{x})$ is the solution of the eikonal equation

$$\nabla T_m(\mathbf{x}) \cdot \nabla T_m(\mathbf{x}) = s^2(\mathbf{x}), \quad T_m(\mathbf{x}_{r,m}) = 0. \tag{21}$$

If there is an infinitesimal perturbation ($\delta\mathbf{x}_{s,n}$, $\delta\tau_n$) in ($\mathbf{x}_{s,n}$, τ_n), then the corresponding perturbation in the objective function is

$$\delta\chi\left(\mathbf{x}_{s,n}, \tau_n\right) = \sum_{m=1}^{M} \omega_{m,n}\left[T_m(\mathbf{x}_{s,n}) - \left(\tau^o_{m,n} - \tau_n\right)\right]\left[\delta\mathbf{x}_{s,n} \cdot \nabla T_m(\mathbf{x}_{s,n}) + \delta\tau_n\right]. \tag{22}$$

At the minimizer $(\hat{\mathbf{x}}_{s,n}, \hat{\tau}_n)$ the partial derivative of the objective function with respect to the origin time is zero, that is,

$$0 = \frac{\partial\chi\left(\hat{\mathbf{x}}_{s,n}, \hat{\tau}_n\right)}{\partial\tau_n} = \sum_{m=1}^{M} \omega_{m,n}\left[T_m(\hat{\mathbf{x}}_{s,n}) - \left(\tau^o_{m,n} - \hat{\tau}_n\right)\right]. \tag{23}$$

Accordingly, at an estimated location $\mathbf{x}_{s,n}$ the optimal origin time can be approximated by

$$\hat{\tau}_n = \frac{\sum_{m=1}^{M} \omega_{m,n}\left[\tau^o_{m,n} - T_m(\mathbf{x}_{s,n})\right]}{\sum_{m=1}^{M} \omega_{m,n}}. \tag{24}$$

Substituting Equation 24 into Equation 22, we further have

$$\delta\chi\left(\mathbf{x}_{s,n}, \tau_n\right) = \sum_{m=1}^{M} \omega_{m,n}\left[T_m(\mathbf{x}_{s,n}) - \left(\tau^o_{m,n} - \hat{\tau}_n\right)\right]\delta\mathbf{x}_{s,n} \cdot \nabla T_m(\mathbf{x}_{s,n}). \tag{25}$$

Equations 24 and 25 define the eikonal equation-based earthquake location method, which uses the global optimization to find the origin time τ_n and resorts to the local search for the spatial location $\mathbf{x}_{s,n}$. The optimization algorithm presented in Appendix A is adopted to find $\mathbf{x}_{s,n}$ in this study.

Theoretically, the eikonal equation-based earthquake location method has no difference from the commonly used Geiger's earthquake location method (e.g., Geiger, 1910; Tong et al., 2016). The latter is based on

the following relationship between the traveltime residual δt at a particular receiver $\mathbf{x}_{r,m}$ and the causing perturbation $\delta\mathbf{x}_{s,n}$ in earthquake location $\mathbf{x}_{s,n}$,

$$\delta t(\mathbf{x}_{r,m}) = \delta\mathbf{x}_{s,n} \cdot \frac{\partial t}{\partial \mathbf{x}_{s,n}}. \tag{26}$$

In Equation 25, if only one seismic station is considered, then

$$\begin{aligned} \delta\chi\left(\mathbf{x}_{s,n},\tau_n\right) &= \omega_{m,n}\left[T_m(\mathbf{x}_{s,n}) - \left(T^o_{m,n} - \hat{\tau}_n\right)\right]\delta T_m(\mathbf{x}_{s,n}) \\ &= \omega_{m,n}\left[T_m(\mathbf{x}_{s,n}) - \left(T^o_{m,n} - \hat{\tau}_n\right)\right]\delta\mathbf{x}_{s,n} \cdot \nabla T_m(\mathbf{x}_{s,n}), \end{aligned} \tag{27}$$

or

$$\delta T_m(\mathbf{x}_{s,n}) = \delta\mathbf{x}_{s,n} \cdot \nabla T_m(\mathbf{x}_{s,n}). \tag{28}$$

A direct comparison between Equations 26 and 28 shows that the eikonal equation-based earthquake location method and Geiger's method are rooted in the same theory. In order to compute $\partial t/\partial\mathbf{x}_{s,n}$ or $\nabla T_m(\mathbf{x}_{s,n})$, Geiger's method needs to determine the raypath connecting the earthquake $\mathbf{x}_{s,n}$ and seismic station $\mathbf{x}_{r,m}$ beforehand. On the contrary, no ray paths are required for the eikonal equation-based earthquake location method.

3.2. Synthetic Earthquake Location Test

We relocate 47 earthquakes simultaneously in a simple vertically increasing velocity model (Figure 3a) to demonstrate the performances of the eikonal equation-based earthquake location method and Geiger's method. The location of each earthquake is determined by its 24 seismic arrivals recorded by 24 seismic stations on the surface. If we know exactly the velocity model, all the 47 earthquakes can be precisely located by either earthquake location method (Figures 3a and 3b). However, the real situation is not so ideal. Model inaccuracy has to be taken into account in earthquake location. We assume that all the earthquakes occur in a model (Figure 3c) which has alternative velocity perturbations with a maximum amplitude of 6% from the vertically increasing velocity model (Figure 3a). But we determine the locations of the 47 earthquakes still in the vertically increasing velocity model. The final locations of all the earthquakes have errors in horizontal position, focal depth and origin time, especially the ones around the boundaries of the computational domain (Figures 3c and 3d). The eikonal equation-based earthquake location method and Geiger's method also have very similar and satisfactory performances in the second example with an inaccurate velocity model.

4. Application: Tomography of the 2020 M_w 4.9 Anza Earthquake Area

We propose an integrated workflow for the development of seismic velocity models in real practice (Figure 4). In detail, the velocity model is obtained in an iterative fashion. The iterative process consists of two steps: (a) eikonal equation-based earthquake location and (b) adjoint-state traveltime tomography. In this study the objective function of the eikonal equation-based earthquake location method measures the same quantity as that of the adjoint-state traveltime tomography. The iteration stops when both objective functions become stable.

We choose the April 2020 M_w 4.9 Anza earthquake area in the San Jacinto Fault zone (SJFZ) as the test field to assess the real-world performance of the integrated workflow (Figure 5). Both V_P and V_P/V_S models are developed. On the one hand, we can verify the reliability of the V_P model to be produced by comparing it with velocity models of previous studies such as the SCEC Community Velocity Model (CVM-H 15.1.0) (Shaw et al., 2015). Meanwhile, most earthquakes in southern California have been precisely located by advanced techniques like the double-difference earthquake location method (Waldhauser & Ellsworth, 2000). This allows us to construct an accurate velocity model by using traveltime data of those accurately located earthquakes. On the other hand, the methods developed in the present study provide the necessary tools to

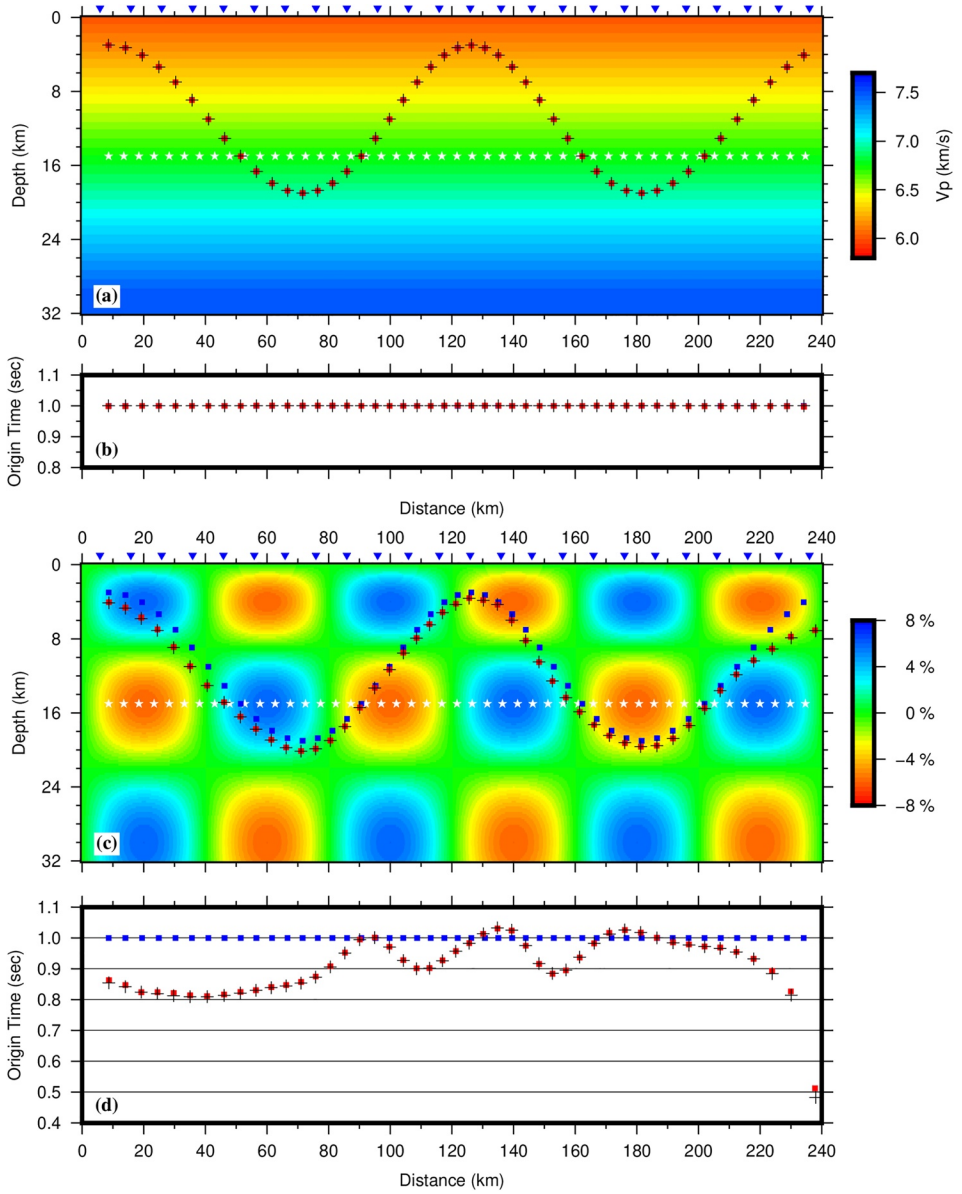


Figure 3. Earthquake locations determined by the eikonal equation-based earthquake location method and Geiger's method. (a) The velocity model for locating earthquakes. The "observed" traveltimes are also computed in this model. The color bar on the right is in km/s. The 24 blue inverse triangles are seismic stations. The initial and true locations of the 47 earthquakes are denoted by white stars and blue squares, respectively. Red squares are the locations determined by the method of this study, while black crosses are the locations given by Geiger's method. (b) The red squares determined by the eikonal equation-based method and black crosses given by Geiger's method are the final origin times of the 47 earthquakes. The true origin times are denoted by blue squares. The initial origin times of all the earthquakes are 0.0 s. In (a)–(b) all the blue squares are overlain by red squares and black crosses. (c) The relative velocity perturbation of the true velocity model from the velocity model in (a). The color bar for the relative velocity perturbation is included on the right. The "observed" traveltimes are computed in the true model. Earthquake locations are determined in the velocity model of (a). All the others in (c) and (d) are the same as those displayed in (a) and (b). For demonstration purpose, both (a) and (c) have a vertical exaggeration of 3:1.

investigate the relationship between velocity heterogeneities and seismic activities. The right-lateral strike-slip SJFZ is one of the most seismologically active areas in the world. In particular, more than 10% of seismicity in southern California since 2000 took place in the trifurcation area of the SJFZ (Ross et al., 2017), including the April 2020 M_w 4.9 Anza earthquake (Figures 5a and 5b). There is an interesting observation that at least one earthquake with magnitude >4.0 occurs in the trifurcation area every 3–4 years during the

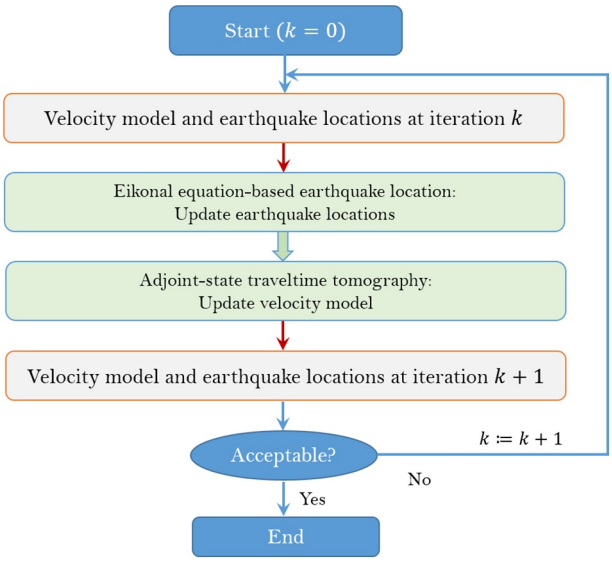


Figure 4. An integrated workflow performing eikonal equation-based earthquake location and adjoint-state traveltime tomography alternatively.

past two decades (Figure 5b). Thus, the source area of the April 2020 M_w 4.9 Anza earthquake as well as the hosting SJFZ is an ideal place to test the performances of the newly developed earthquake location and traveltime tomography methods. In addition to that, the tomographic results can provide new insight into the relationship between earthquake occurrence and structural heterogeneities.

4.1. Data and Initial Models

All the first *P*- and *S*-wave arrival times used in the present study are from Southern California Earthquake Data Center (SCEDC, 2013). But the hypocenters of almost all earthquakes are further updated according to the high-quality 1981–2019 waveform relocated earthquake catalog for southern California (Hauksson et al., 2012). The very few earthquakes that are not included in the waveform relocated catalog are removed.

A total of 1,101,212 first *P*-wave arrivals during the period from January 2000 to December 2019 with the corresponding earthquakes and seismic stations located in the study area (Figure 5a) are inverted to construct an initial model for seismic tomographic inversions. Starting at the standard southern California velocity model which is separated by discontinuities at 5.5, 16.0 and 32.0 km depth (Wald et al., 1995), the velocities of the 3 crustal layers are inverted to be 5.616, 6.263, and 6.702 km/s, respectively. The *P*-wave velocity in the mantle is fixed to 7.8 km/s. The flat discontinuity at 32.0 km depth is then replaced by a curved surface embedded in the CVM-H 15.1.0 model to honor the Moho topography (Shaw et al., 2015). The sharp velocity contrasts across the 3 discontinuities are smoothed by convolving the velocity model with a one-dimensional Gaussian function in the depth direction (Tong et al., 2019). The smoothed velocity model is the initial model for seismic *P*-wave tomography, denoted by $\mathbf{m}_0$.

The *P*-wave arrivals mentioned above are filtered for the subsequent *P*-wave tomography. First, to avoid event clustering only one earthquake is selected from every $0.4 \times 0.4 \times 0.4$ km block. The earthquake is required to have at least 6 and the maximum number of picks among all the earthquakes in that containing block. Second, the study area has a landscape size of 190×190 km, but we only choose the earthquakes occurring in a smaller domain $(5, 185 \text{ km}) \times (5, 185 \text{ km})$, so that seismic stations in the study area are more likely to have a full azimuthal coverage over the selected earthquakes (Figure 5c). In the end, we have selected 629,281 first *P*-wave arrivals of 31,815 earthquakes recorded by 315 seismic stations. Following the same criteria except that every earthquake must have at least 12 picks, 629, 381 first *S*-wave arrivals of 22, 409 earthquakes recorded by 308 seismic stations are selected for *S*-wave tomography (Fig. S1). The large and almost the same amounts of first *P*- and *S*-wave arrivals have fully covered the study area; that is necessary for the reliability of the V_P/V_S model to be produced. The initial model for *S*-wave tomography is the final V_P model divided by a factor of 1.73.

4.2. Multiscale Checkerboard Resolution Test

In real practice of seismic tomographic inversions, we have to deal with imprecisely located earthquakes and noisy observations. A challenging checkerboard resolution test is conducted to verify the reliability of the inversion workflow and the resolving ability of the chosen *P*-wave data. Firstly, none of the 31,815 earthquakes are placed at their true locations. For the *i*-th earthquake in chronological order ($i = 1, \ldots, 31{,}815$), its initial errors in spatial location $\mathbf{x}_{s,i}$ and origin time τ_i are

$$e_i^{\mathbf{x}_s} = \mathbf{x}_{s,i}^{\text{true}} - \mathbf{x}_{s,i} = \left(2\cos\left(2\pi\frac{i}{101}\right),\ 2\sin\left(2\pi\frac{i}{101}\right),\ 4\cos\left(2\pi\frac{i}{103}\right)\right) \tag{29}$$

and

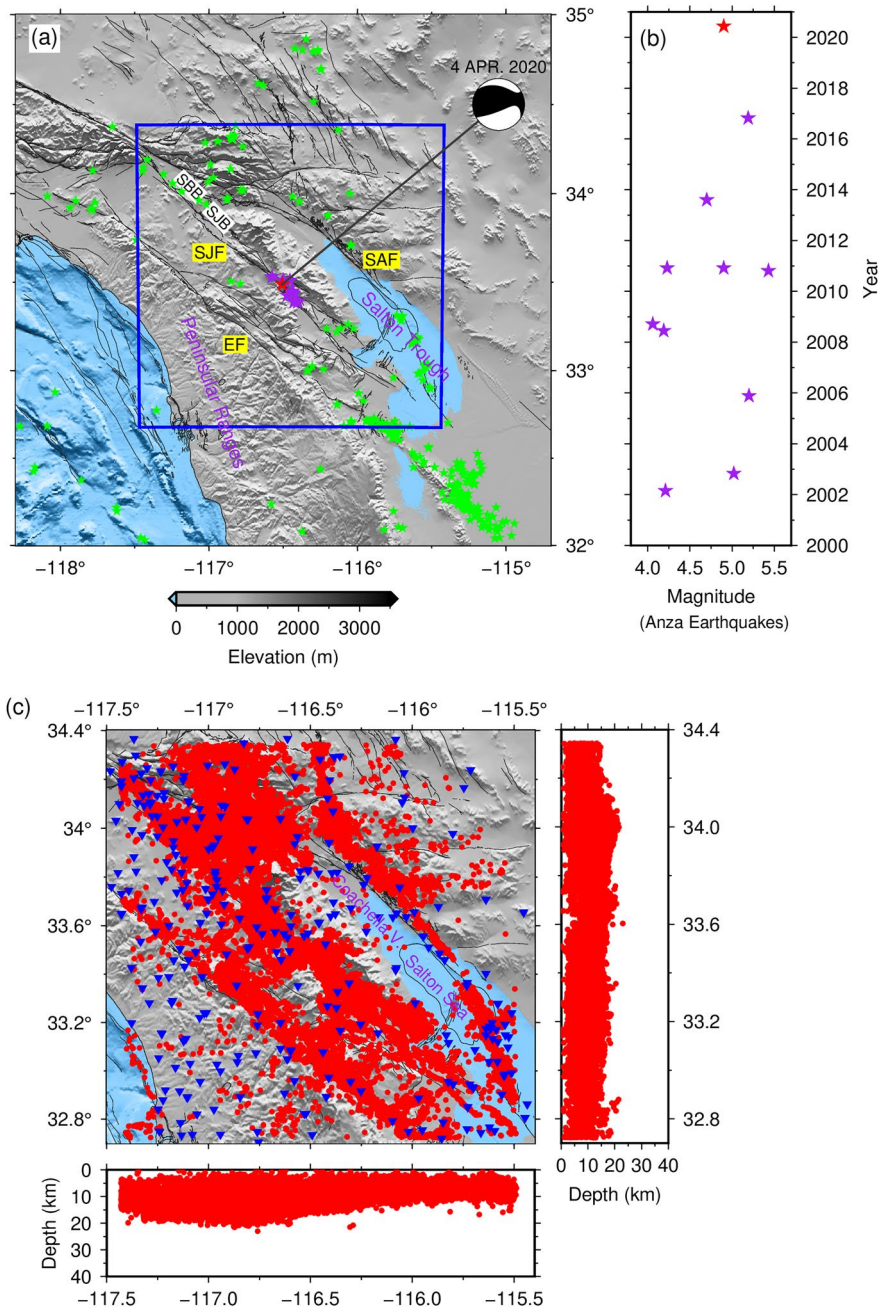


Figure 5. (a) The tectonic setting and surface topography in southern California. The gray curves denote the surface traces of active faults including the San Andreas Fault (SAF), San Jacinto Fault (SJF) and Elsinore Fault (EF) (Jennings & Bryant, 2010). The stars except the red one are $M_w \geq 4.0$ earthquakes during the period from 2000 to 2019. The red and purple stars are earthquakes occurring in the trifurcation area of the San Jacinto Fault Zone. The focal mechanism of the 2020 M_w 4.9 Anza earthquake (red star) is displayed by the beach ball. SBB is short for San Bernardino Basin. SJB is for San Jacinto Basin. The blue box outlines the present study area. (b) The 11 $M_w \geq 4.0$ Anza earthquakes occurring in the trifurcation area in the San Jacinto Fault Zone since 2000. (c) The spatial distributions of the 31,815 earthquakes (red dots) and 315 seismic stations (blue inverse triangles) selected for *P*-wave tomography. The Coachella Valley and Salton Sea are labeled.

$$e_i^{\tau} = \tau_i^{\text{true}} - \tau_i = 0.4\cos\left(2\pi\frac{i}{107}\right). \tag{30}$$

Secondly, the relative velocity perturbation of the checkerboard velocity model with respect to the initial model $\mathbf{m}_0$ has a multiscale feature, which can be expressed by

$$\frac{v_{\text{ckbd}} - v_0}{v_0} = \left[3\%\sin\left(\frac{\pi x}{95}\right)\sin\left(\frac{\pi y}{95}\right) - 5\%\sin\left(\frac{2\pi x}{95}\right)\sin\left(\frac{2\pi y}{95}\right)\right]\sin\left(2\pi\frac{\sqrt{25+8z}-5}{8}\right). \tag{31}$$

Thirdly, the synthetic traveltime data calculated in the checkerboard velocity model are deliberately polluted by Gaussian noise to mimic the picking errors in real data. The mean and standard deviation of the noise are 0.0 and 0.05 s, respectively. Lastly, one multiple-grid formed by 5 sets of regular grids is used to discretize the relative slowness perturbation. The lateral spacing of each regular grid is 15.0 km, and the vertical spacing varies from 3.0 km near the surface to 7.5 km around 25 km depth (Figure S2). Each of the 5 regular grids adequately samples the velocity anomalies. But none of the grid nodes coincide with any peaks or troughs of the checkerboard velocity model, making it difficult to fully recover the anomalous structures in amplitude.

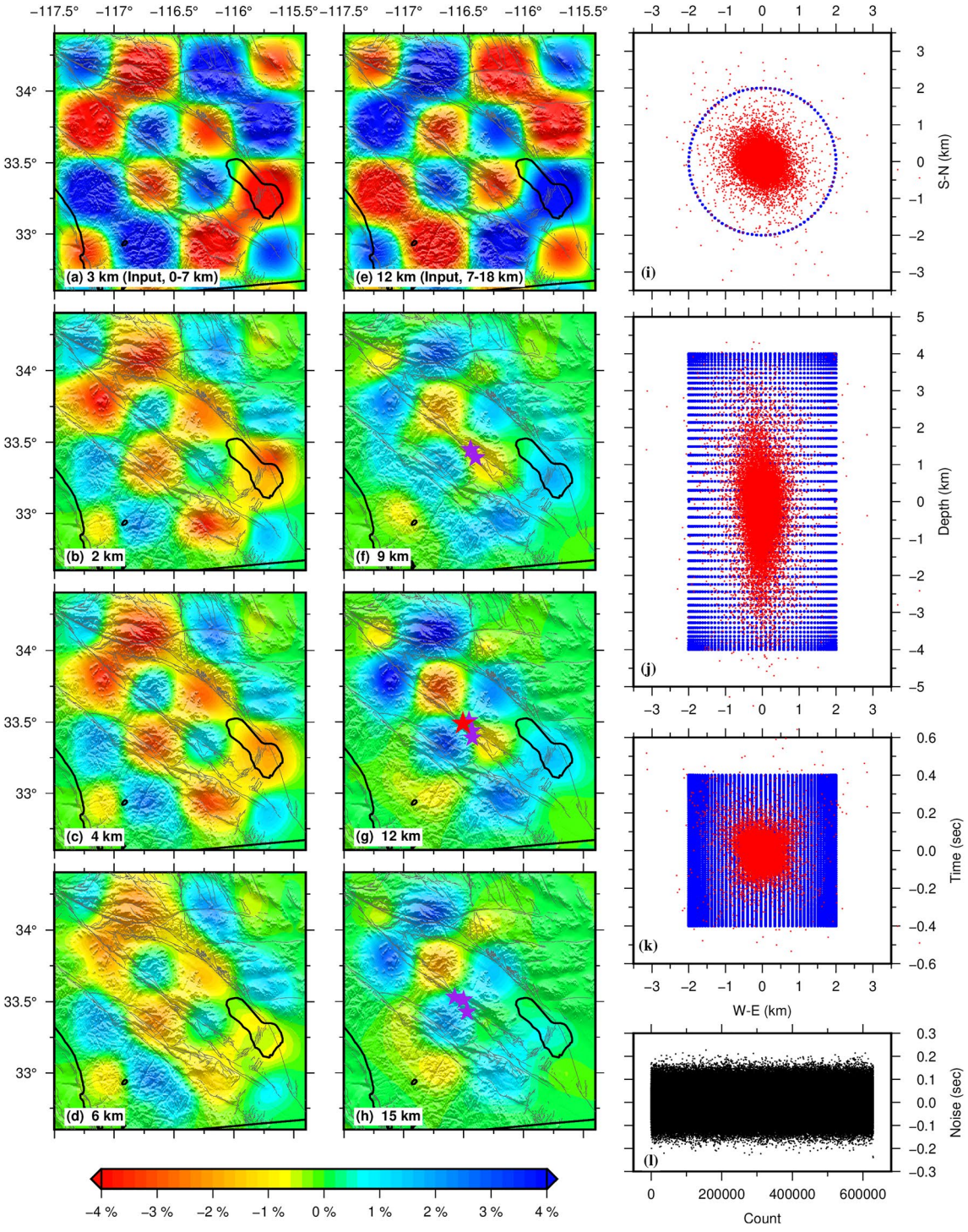
We follow the steps of the integrated workflow described in Figure 4 to recover the multiscale checkerboard velocity anomalies. During the iterative process of earthquake location, an upper bound of 1.0% is set for the non-dimensionalized variables ($\Delta x/D_x$, $\Delta y/D_y$, and $\Delta z/D_z$) by choosing $D_x = D_y = D_z = 20$ km. Similarly, the relative slowness perturbation $\Delta s(\mathbf{x})/s(\mathbf{x})$ is bounded by 1.5% at every update. Earthquake location and adjoint-state traveltime tomography are both iteratively performed four times during each iteration. The workflow process stops after six iterations when the reduction in the objective function is less than 5% of the reduction amount at the first iteration.

The obtained velocity model at six representative depths, the initial hypocenter errors, the final hypocenter errors and the Gaussian noise added to the synthetic data are shown in Figure 6. From the surface to less than 7 km depth (Figures 6a–6d), almost all the fast and slow velocity anomalies of different sizes are recovered. Below 7 km depth (Figures 6e–6h), the velocity structure between the Elsinore Fault and San Andreas Fault is properly revealed, including the hypocentral zone of the 2020 M_w 4.9 Anza earthquake. The initial maximum hypocenter errors in horizontal place, depth and origin time are 2 km, 4 km, and 0.4 s, respectively; whereas the majority of the corresponding final hypocenter errors are less than 1 km, 2 km, and 0.2 s (Figures 6i–6k). The accuracy of earthquake locations has been greatly improved. However, the amplitudes of velocity anomalies are slightly underestimated. This is partly due to the adding noise and still existent errors in earthquake locations. Undoubtedly, the pattens and spatial distributions of fast and slow velocity anomalies are accurately determined, allowing us to interpret velocity heterogeneities revealed by the real data. A similar multiscale checkerboard resolution test is performed for *S*-wave tomography. The results included in Figure S3 show that the *S*-wave data have the same resolving ability as the *P*-wave data. Since both *P*- and *S*-wave data can well resolve the structure between the Elsinore Fault and San Andreas Fault, we claim that the V_P/V_S structure of this specific region can be reliably resolved.

4.3. Tomographic Results

By only replacing the noisy synthetic data with the real data and inheriting all other parameters from the *P*-wave and *S*-wave checkerboard resolution tests, we obtain the optimal V_P and V_S models. The V_P/V_S model is subsequently computed by taking the ratio of the final V_P to the final V_S. Both V_P and V_P/V_S models reveal strong crustal heterogeneities in the study area (Figure 7).

In the upper crust, low-V_P high-V_P/V_S anomalies are related to basins and faults (Figures 7a–7c, 7g–7i). The low-V_P high-V_P/V_S anomaly beneath the Salton Sea and the elongated Coachella Valley in the north (both are parts of the Salton Trough) correlates with the sedimentary basin structure of the Salton Trough. This anomaly extends from the surface to about 6 km depth; its southeastward shortening indicates the asymmetry of the Coachella Valley basin (Ajala et al., 2019). Low-V_P high-V_P/V_S anomalies are also observable beneath the San Bernardino Basin and the San Jacinto Basin. There exists another low-V_P zone west of the


(a) 3 km (Input, 0-7 km)
(b) 2 km
(c) 4 km
(d) 6 km
(e) 12 km (Input, 7-18 km)
(f) 9 km
(g) 12 km
(h) 15 km
(i)
(j)
(k)
(l)
S-N (km)
Depth (km)
Time (sec)
W-E (km)
Noise (sec)
Count
−4 %
−3 %
−2 %
−1 %
0 %
1 %
2 %
3 %
4 %

trifurcation area of the San Jacinto Fault. The about-9-km-thick low-velocity zone has been reported by Allam and Ben-Zion (2012) as well. Low-to-normal V_P and high-V_P/V_S anomalies are revealed beneath the surface traces of the Elsinore Fault, San Jacinto Fault and San Andreas Fault, suggesting the presence of high pore fluid pressure (e.g., Moretti et al., 2009; Wang et al., 2012). Attenuation models also suggest that the southern California crust is generally partially fluid-saturated and areas around the San Jacinto Fault are almost completely fluid-saturated (Hauksson & Shearer, 2006).

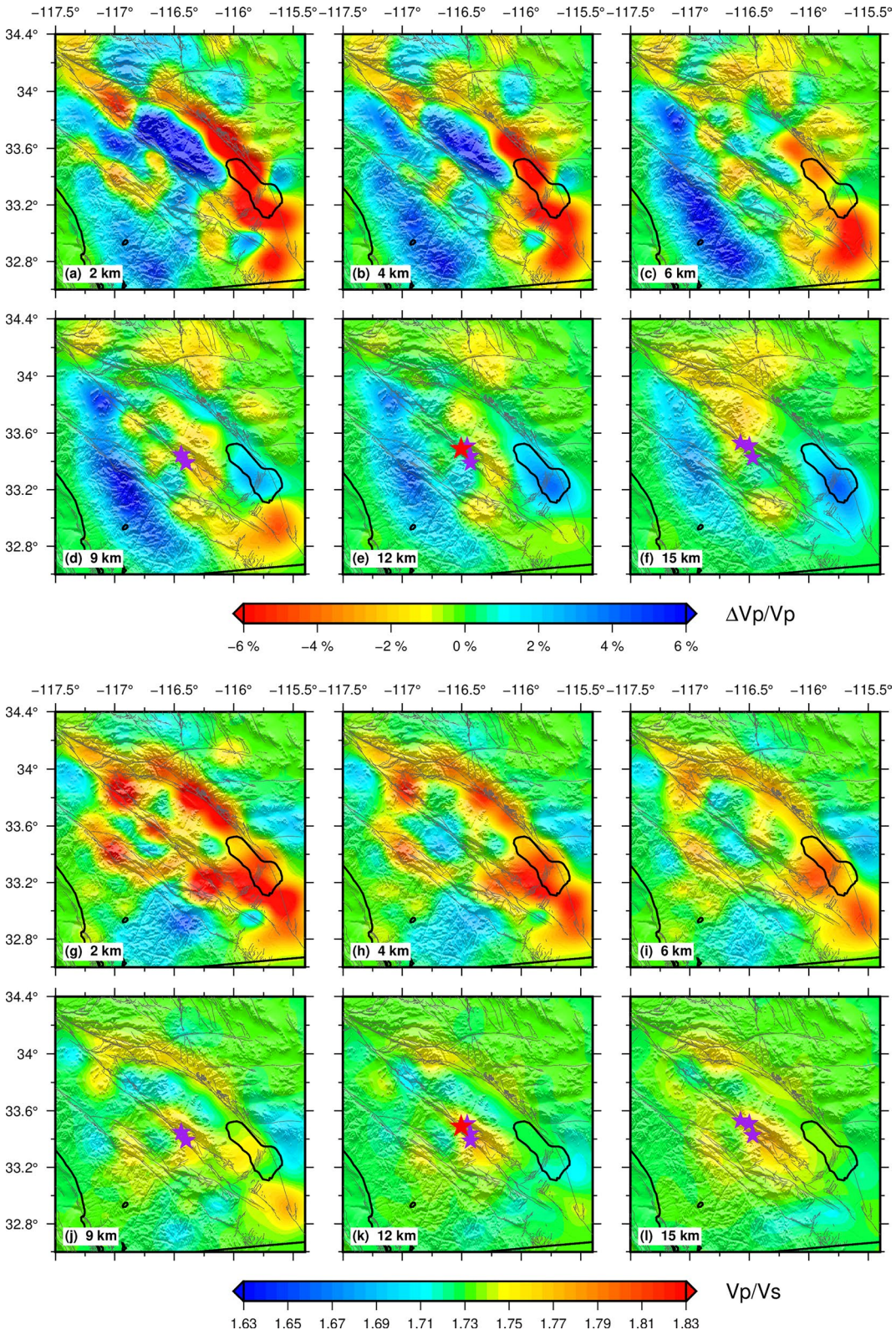
The vast high-velocity structure west of the San Jacinto Fault is related to the mafic part of the Western Peninsular Ranges batholith that is dense, magnetic, and characterized with relatively sparse seismicity (Figures 7a–7f) (Langenheim et al., 2014). The high-velocity rocks east of the San Jacinto Fault at shallow depths reflect the Eastern Peninsular Ranges batholith (Figures 7a–7c), which provide evidence for displacement of the Peninsular Ranges batholith across the San Jacinto Fault (Ajala et al., 2019). The tomographic images also show strong velocity contrasts and reversal in contrast polarity across the San Jacinto Fault. As depth increases, the near-surface low-velocity anomaly beneath Salton Trough Basin is replaced by a pronounced high-velocity structure (Figures 7d–7f). In contrast, a low-velocity body roughly bounded by the Elsinore Fault in the south and the San Andreas Fault in the north emerges in the middle crust at 9–15 km depths. We can also observe that the trifurcation area of the San Jacinto Fault is occupied by a high V_P/V_S structure in the middle crust (Figures 7j–7l).

Figure 7 displays that the 2020 M_w 4.9 Anza earthquake and other moderate-sized Anza earthquakes took place near or at the boundaries of low-velocity anomalies, while the cross-sectional views in Figure 8 show that those earthquakes occurred in normal-V_P rocks with nearby low-velocity anomalies or at the boundaries of low-V_P anomalies. The boundary areas of low-velocity anomalies or normal-V_P areas with low-velocity anomalies surrounded are related to brittle and strong parts of the fault zone that are suitable for the genesis of earthquakes (Zhao & Kanamori, 1995). The same observations were made in source areas of some other large crustal earthquakes. For instance, the 1994 M_w 6.7 Northridge earthquake in southern California, the 2008 M_w 7.9 Wenchuan earthquake in Sichuan, and the 2014 M_w 6.0 South Napa earthquake in northern California all occurred near the boundaries of low-velocity anomalies (Li et al., 2020; Tong et al., 2017; Zhao & Kanamori, 1995). In addition, the hypocenters of the 11 moderate-sized Anza earthquakes after 2000 are located on the top side of a broad high-V_P/V_S body (Figures 7j–7l and Figures 8a2–8d2). Regions with high-V_P/V_S ratios are usually assumed to have a high density of fluid-saturated cracks (e.g., Doi et al., 2013; Moretti et al., 2009). Hence, we infer that the high-V_P/V_S rocks are composed of water-rich materials, which may provide fluids to reduce fault-zone friction and consequently lead to the occurrence of the moderate-sized Anza earthquakes (e.g., Tong et al., 2021; Zhao et al., 2015).

All the findings regarding seismic *P*-wave velocity heterogeneities reported above are consistent with the features of previous models derived from different types of data, including first-arrival traveltimes (e.g., Ajala et al., 2019; Allam & Ben-Zion, 2012; Fang et al., 2016; Share et al., 2019; Tong et al., 2014b), traveltimes of reflected waves (Huang et al., 2016) and ambient noise (e.g., Qiu et al., 2019; Wang et al., 2018; Zigone et al., 2015). The consistence supports the validity of the adjoint-state traveltime tomography method and the integrated workflow.

To compare the performances of different seismic tomography methods, we continue to invert the first *P*-wave arrivals by separately using the adjoint-state traveltime tomography method of the present study and the eikonal ray-based traveltime tomography method (Tong et al., 2017, 2019). The two new V_P models are compared against the velocity model generated by the integrated workflow and the CVM-H 15.1.0 model (Figure 9). All the 4 models have the following features in common: (1) The high-velocity Western

Figure 6. The results of the multiscale checkerboard resolution test for *P*-wave tomography. (a) The relative velocity perturbation of the checkerboard velocity model with respect to the initial model $\mathbf{m}_0$ at a depth of 3 km. As described by Equation 31, the pattern of the relative velocity perturbation from the surface to 7 km depth is the same but the amplitude increases from 0 (0 km) to 6.6% (3 km) and then decreases back to 0 (7 km). (b)–(d) The recovered relative velocity perturbations at 2, 4 and 6 km, respectively. (e) The relative velocity perturbation of the checkerboard velocity model at 12 km depth. The pattern of the relative velocity perturbation from 7 km depth to 18 km depth is the same but the amplitude increases from 0 (7 km) to 6.6% (12 km) and then returns to 0 (18 km). (f)–(h) The recovered relative velocity perturbations at 9, 12 and 15 km, respectively. In (b)–(d) and (f)–(h) stars denote $M_w \geq 4.0$ Anza earthquakes occurring after 2000 and every earthquake is projected to its nearest horizontal slice. The color scale for the relative velocity perturbations shown in (a)–(h) is included at the bottom. (i)–(k) The blue and red dots display the hypocenter errors of the 31,815 earthquakes before and after the multiscale checkerboard resolution test, respectively. (i) Shows the horizontal location errors. The errors in depth are illustrated in (j). The errors of origin time are included in (k). (l) Random errors added to the synthetic data. The errors have a mean value of 0.0 s and a standard deviation of 0.05 s.


−117.5° −117° −116.5° −116° −115.5°
34.4°
34°
33.6°
33.2°
32.8°
(a) 2 km
(b) 4 km
(c) 6 km
(d) 9 km
(e) 12 km
(f) 15 km
−6 % −4 % −2 % 0 % 2 % 4 % 6 %
ΔVp/Vp
(g) 2 km
(h) 4 km
(i) 6 km
(j) 9 km
(k) 12 km
(l) 15 km
1.63 1.65 1.67 1.69 1.71 1.73 1.75 1.77 1.79 1.81 1.83
Vp/Vs

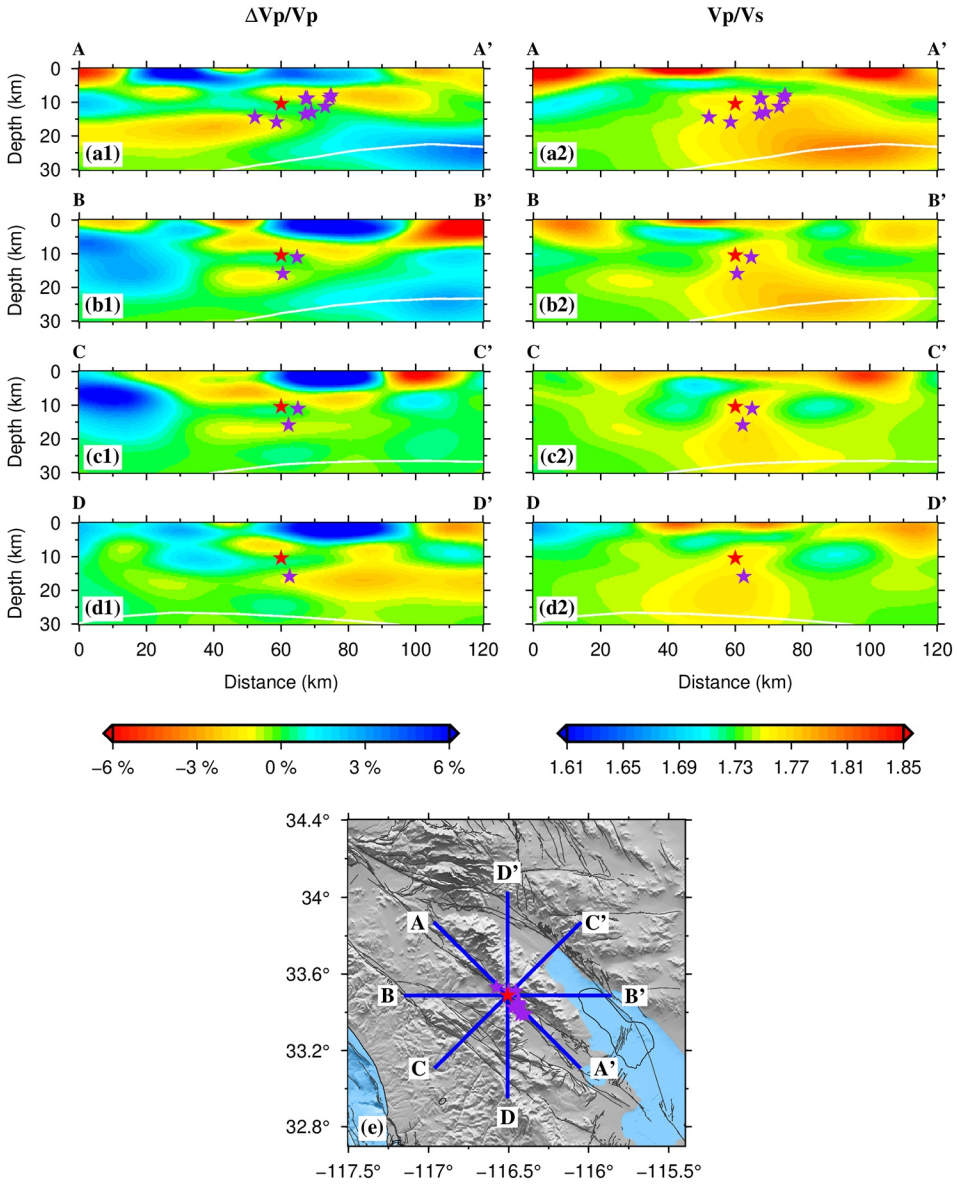


Figure 8. Cross-section views of the V_P (a1–d1) and V_P/V_S (a2–d2) models along 4 profiles. The locations of the 4 vertical profiles are displayed on the inset map (e). Red color denotes low-V_P or high-V_P/V_S structures, and blue color represents high-V_P or low-V_P/V_S anomalies. The respective color scales for V_P and V_P/V_S are shown below (d1) and (d2). The white curves are the Moho discontinuity (Shaw et al., 2015). All the others are the same as Figure 7.

Peninsular Ranges; (2) Existence of low-velocity anomalies beneath the San Bernardino Basin and the San Jacinto Basin and along the San Andreas Fault; (3) Strong velocity contrasts and reversal in contrast polarity across the San Jacinto Fault; (4) The near-surface high-velocity body between the San Jacinto Fault and the San Andreas Fault; (6) Significant low-velocity anomalies around the trifurcation area of the San Jacinto Fault from the near-surface to middle crust; (6) A low-velocity structure atop of high-velocity rocks beneath

Figure 7. (a)–(f) The obtained *P*-wave velocity model after iteratively executing the integrated workflow (Figure 4) six times. Horizontal slices at 6 representative depths are displayed. Red and blue colors denote low- and high-velocity structures, respectively. The color scale for the relative velocity perturbation is below (e). (g)–(l) The V_P/V_S model at 6 representative depths. It is generated by computing the ratio of the final V_P to the final V_S. The final V_P divided by 1.73 is the starting model for *S*-wave tomography. The final V_S is obtained after executing the integrated workflow (Figure 4) six times. The color scale for V_P/V_S is included at the bottom. In (a)–(l) gray curves are active faults (Jennings & Bryant, 2010), the stars represent the moderate-sized Anza earthquakes ($M_w \geq 4.0$) after 2000, and in particular the red star denotes the 2020 April M_w 4.9 Anza earthquake.

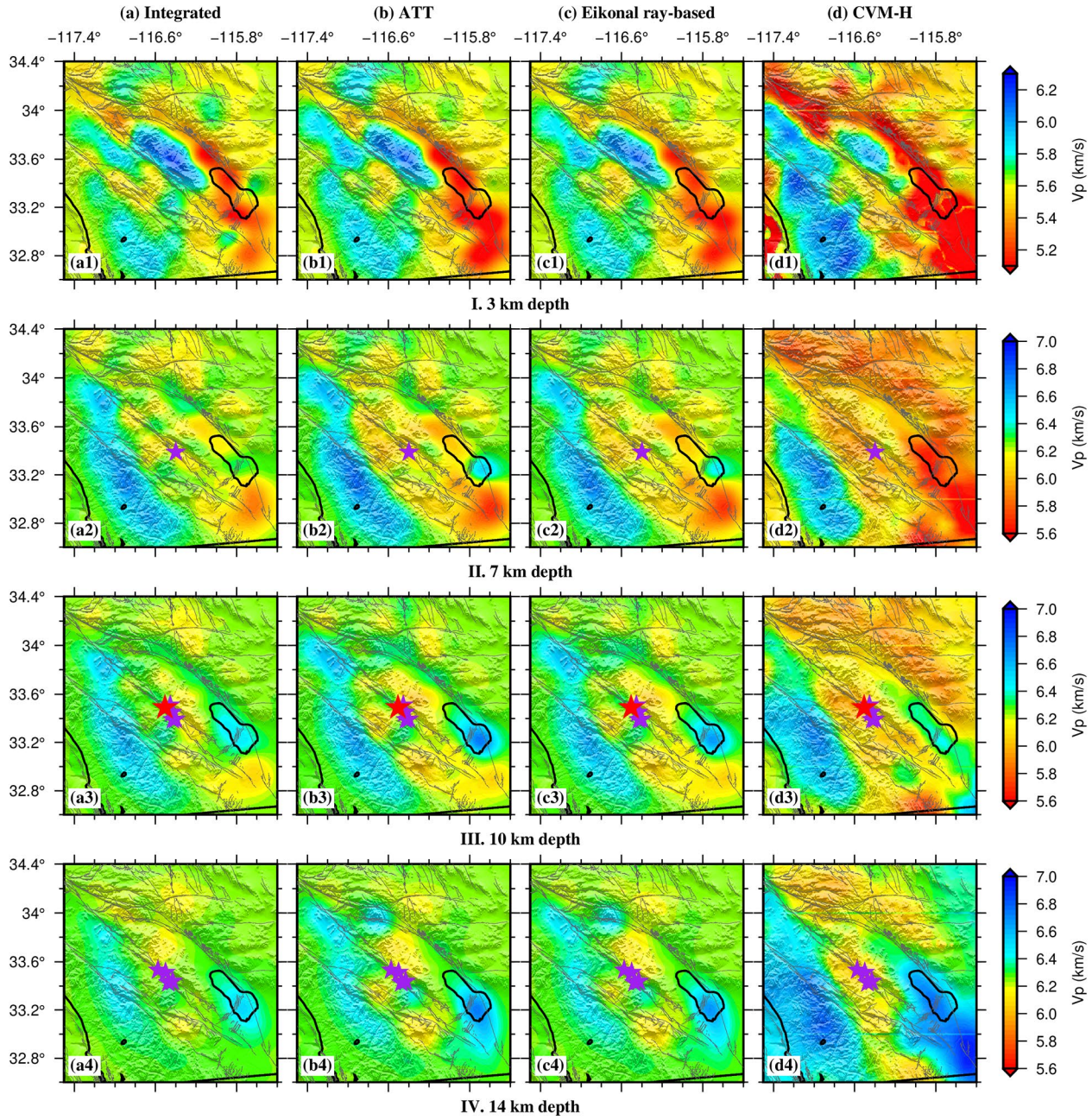


Figure 9. Comparisons of the *P*-wave velocity models generated by different methods. The velocity models at 4 representative depths (3, 7, 10, and 14 km) are displayed. Each color scale inset on the right is for the absolute velocity at the same depth. The horizontal slices of (a1)–(a4) are generated by the integrated workflow. (b1)–(b4) are the results of the adjoint-state traveltime tomography (ATT) method. Tomographic images in (c1)–(c4) are produced by the eikonal ray-based traveltime tomography method (Tong et al., 2017, 2019). (d1)–(d4) are extracted from the SCEC Community Velocity Model (CVM-H 15.1.0) (Shaw et al., 2015). All the others are the same as Figure 7.

Salton Trough. The differences between the V_P models generated by adjoint-state traveltime tomography and eikonal ray-based traveltime tomography are minor and can be ignored (Figures 9b and 9c). There are two main reasons for the nearly identical results: (1) The same data set and same model parameterization are used by the two tomography methods; (2) The using eikonal ray-based traveltime tomography method is able to accurately determine seismic ray paths in complex media by solving the eikonal equation with a highly accurate numerical solver first and then tracing along the negative traveltime gradient direction (Tong et al., 2017). The community velocity model CVM-H 15.1.0 has larger variations or stronger

heterogeneities than the other three models. The results of the integrated workflow, adjoint-state traveltime tomography and eikonal ray-based traveltime tomography all suggest that the moderate-sized Anza earthquakes after 2000 took place near or at the boundaries of low-velocity rocks. However, the CVM-H 15.1.0 model reveals that the hypocenters of those Anza earthquakes are within a wide low-velocity structure. We speculate that the minor discrepancy is mainly caused by the relatively low frequency data that were used to develop the community velocity model. As described by Shaw et al. (2015), waveform adjoint tomography was the final stage for the development of the community velocity model CVM-H 15.1.0, and three-component seismograms filtered between 2 and 30 s were used to refine the crustal model. Considering that only velocity anomalies at or above the wavelength scale can be resolved by waveform adjoint tomography, the community velocity model may not contain the small-scale details of the Anza earthquake source zone as revealed by the other three tomographic models.

5. Discussion and Conclusions

Seismic traveltime tomography is a data-fitting procedure that minimizes some form of the difference measurement between observed and theoretical traveltimes (the objective function) to obtain an optimal subsurface velocity model. Since the pioneer work of Aki and Lee (1976) to present, it is usually necessary to trace seismic wave paths in a certain velocity model so that the theoretical traveltimes and the derivative of the objective function can be computed. In contrast, the adjoint-state traveltime tomography method of this study avoids the cumbersome step of ray tracing but the theoretical traveltimes and the derivative of the objective function still can be computed with high accuracy. If all the involving seismic ray paths can be accurately traced, traditional ray-based seismic tomography methods can produce similar results as the adjoint-state traveltime tomography. Given the fact that common ray tracing schemes such as shooting, bending and pseudo bending may fail even in mildly heterogeneous models (Rawlinson et al., 2008), while the eikonal equation precisely describes the traveltime field in a medium with any level of complexity, it is a technical advance to develop the ray-free eikonal equation-based adjoint-state traveltime tomography method. One can obtain stable and reliable tomographic results by using the new adjoint-state traveltime tomography method.

In this study, first-arrival traveltime is assumed to be governed by the eikonal equation $|\nabla T(\mathbf{x})| = s(\mathbf{x})$. It is usually mentioned that the eikonal equation is derived from the wave equation under the high-frequency approximation (e.g., Aki & Richards, 2002; Rawlinson et al., 2008). Intuitively, it can be viewed as the differentiation expression of the relation: Traveltime/Distance = 1/Velocity. The eikonal equation actually states a precise mathematical physical model for traveltime without considering the finite-frequency effects of wave propagation (e.g., Dahlen et al., 2000; Liu & Tromp, 2006). Accordingly, we reformulate seismic traveltime tomography as an eikonal equation-constrained optimization problem and solve it by using the effective adjoint-state method. An auxiliary adjoint-state equation is introduced to facilitate the computation of the gradient of the objective function or the sensitivity kernel. It is common to add damping and smoothing regularization to the objective function, which is an efficient way to control the amplitude and wavelength of the retrieved model (e.g., Rawlinson et al., 2010; Tape et al., 2007; Tong et al., 2014a). But we do not add regularization terms to the objective function in all the examples of the present study. The two main reasons are: (1) The step-size-controlled gradient descent method can efficiently constrain the amplitude of the model; (2) The coarse inversion grid has a smoothing effect on the inverted model.

The computational cost of the adjoint-state traveltime tomography method is moderate and proportional to the number of earthquakes. If there are more earthquakes than seismic stations, just like the *P*-wave tomography example which involves 31,815 earthquakes and 315 seismic stations, we consider all the receivers as virtual sources and all the earthquakes as virtual receivers. This strategy can save a lot of computational time. In comparison with wave equation-based adjoint tomography, eikonal equation-based adjoint-state traveltime tomography is computationally inexpensive and can be carried out on a moderately equipped desktop workstation of nowadays. But compared with seismic tomography methods using pseudo-bending approach for ray tracing (e.g., Zhao et al., 1992; Zhang & Thurber, 2003), the adjoint-state traveltime tomography method of this study is computationally costly.

An eikonal equation-based earthquake location method using first-arrival traveltimes has been developed to locate earthquakes when needed. Using this method, the number of eikonal equations to be solved equals the number of seismic stations. This method is quite efficient if there are more earthquakes than seismic stations. Although the eikonal equation-based earthquake location method updates earthquake locations iteratively, one can prefer to solve the eikonal equations only once if there is enough space to store the traveltime fields. Theoretically, the new earthquake location method is identical to Geiger's method but no ray tracing is required.

The efficacy of the adjoint-state traveltime tomography and eikonal equation-based earthquake location methods has been demonstrated through a series of tests. An integrated workflow that performs eikonal equation-based earthquake location and adjoint-state traveltime tomography alternatively has been designed and applied to the 2020 M_w 4.9 Anza earthquake source area. High-resolution V_P and V_P/V_S models are generated. The main features of the P-wave velocity model are consistent with those discovered in previous studies. The obtained P-wave velocity model shows that the 2020 M_w 4.9 Anza earthquake and other moderate-sized Anza earthquakes after 2000 occurred near or at the boundaries of low-velocity rocks, regions that are considered to be prone to crustal earthquakes. Moreover, the source zone of the Anza earthquakes is characterized by high-V_P/V_S, suggesting the existence of water-rich materials and hence indicating the critical role of fluids in the genesis of the frequent moderate-sized Anza earthquakes.

In conclusion, the developed adjoint-state traveltime tomography method and eikonal equation-based earthquake location method provide alternatives to ray-based traveltime tomography and ray-based earthquake location methods. They have the potential to become routine tools in future seismic tomographic studies.

Appendix A: Optimization Algorithm

The gradient descent method with a magnitude-controlled step size is designed to iteratively find the minimum of the objective function $\chi(s(\mathbf{x})) = \chi(C_{1,1},\ C_{2,1},\ \ldots\ ,C_{L_H,H})$ in a finite-dimensional vector space. Let $\mathbf{X} = (C_{1,1},C_{2,1},\ldots,C_{L_H,H})$ and set an upper bound γ for the L_∞ norm of its perturbation $\|\Delta\mathbf{X}\|_\infty$ at every iteration. The step-by-step process of the algorithm is given below, which starts from $k = 0$ and $\mathbf{X}^0 = \mathbf{0}$.

1. Compute the objective function $\chi(\mathbf{X}^k)$ and gradient vector $\mathbf{g}^k = \partial\chi/\partial\mathbf{X}^k$. When $k > 0$, divide the upper bound γ by a factor of $\kappa > 1$ as $\gamma \to \gamma/\kappa$ if $\chi(\mathbf{X}^k) > \chi(\mathbf{X}^{k-1})$
2. Let $f = \chi(\mathbf{X}^k)$ and $g = \mathbf{g}^k \cdot \mathbf{g}^k$, compute the step size $\lambda_k = f/(2g)$
3. Compute $\alpha_k = \gamma/\|\lambda_k\mathbf{g}^k\|_\infty$ if $\|\lambda_k\mathbf{g}^k\|_\infty > \gamma$, otherwise $\alpha_k = 1$
4. Compute $\Delta\mathbf{X} = -\alpha_k\lambda_k\mathbf{g}^k$, and update the model $\mathbf{X}^{k+1} = \mathbf{X}^k + \Delta\mathbf{X}$
5. If $g \le \epsilon$, the tolerance level, or k exceeds the pre-defined maximum iteration number, or the objective function value becomes stable as $|\chi(\mathrm{X}^k) - \chi(\mathrm{X}^{k-1})|/|\chi(\mathrm{X}^1) - \chi(\mathrm{X}^0)| \le \theta$ (θ is relatively small), then X^{k+1} is the optimal solution; otherwise go back to the first step with $k + 1$

For the gradient descent method, there is no such a uniform expression for the optimal step size λ_k that works for all the objective functions. If the objective function is quadratic and has a close-to-zero minimum, then $\lambda_k = f/(2g)$ is a good approximation to the optimal step size when the function value is about twice its minimum. That is to say, one can roughly get the current objective function value halved by setting $\lambda_k = f/(2g)$. However, in real practice the objective function $\chi(\mathbf{X}^k)$ always involves noisy observed data, making it even more difficult to estimate the optimal step size. Thus, we prefer to choose a relatively small step size so that the updated model is close to the previous one. This is achieved by setting an upper bound γ in advance for the L_∞ norm of the model perturbation $\|\Delta\mathbf{X}\|_\infty$. Eventually, the model is updated in the negative gradient direction with a magnitude-controlled step size of $\alpha_k\lambda_k$. When there is no reduction or even an increase in the objective function, it is likely that the solution is around the local minimum. A smaller upper bound γ/κ ($\kappa > 1$) is required to further constrain the step size. This conservative strategy can be considered as an implicit implementation of the damping regularization. Liu et al. (2019) reported that an upper bound of about 1% for the relative velocity perturbation has the same effect as the optimal damping parameter.

Of note, some other more sophisticated optimization algorithms such as the nonlinear conjugate gradient method (e.g., Fletcher & Reeves, 1964; Tromp et al., 2005; Tong et al., 2014a) and the Gauss-Newton method (e.g., Virieux & Operto, 2009; Tavakoli et al., 2017) can be utilized to find a minimizer for the objective

function as well. These algorithms attempt to find the optimal model update at each iteration so that fewer iterations are required to reach the minimum, while the gradient descent method with a magnitude-controlled step size usually takes more iterations to converge. But a smaller step size is more likely to get a stable solution.

Data Availability Statement

The seismic data are downloaded from Southern California Earthquake Data Center (SCEDC, 2013). Some figures are made with the Generic Mapping Tool (GMT) (Wessel & Smith, 1991).

Acknowledgments
The author is grateful to Y. Ben-Zion, an associate editor and two anonymous reviewers for their constructive comments. The author thanks Jean Virieux and Shingyu Leung for insightful discussions. This work is partly funded by Singapore MOE AcRF Tier-2 Grant (04MNP002073C230) and Singapore MOE AcRF Tier-1 Grant (04MNP000559C230). This work comprises Earth Observatory of Singapore contribution no. 276. This research is also partly supported by the National Research Foundation Singapore and the Singapore Ministry of Education under the Research Centers of Excellence Initiative (Project Code Number: 04MNS001953A620).

References

Ajala, R., Persaud, P., Stock, J. M., Fuis, G. S., Hole, J. A., Goldman, M., & Scheirer, D. (2019). Three-dimensional basin and fault structure from a detailed seismic velocity model of Coachella valley, southern California. *Journal of Geophysical Research: Solid Earth*, *124*, 4728–4750. https://doi.org/10.1029/2018jb016260

Aki, K., & Lee, W. H. K. (1976). Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model. *Journal of Geophysical Research*, *81*, 4381–4399. https://doi.org/10.1029/jb081i023p04381

Aki, K., & Richards, P. G. (2002). *Quantitative seismology: Theory and methods* (2nd ed.), University Science Books.

Allam, A. A., & Ben-Zion, Y. (2012). Seismic velocity structures in the southern California plate-boundary environment from double-difference tomography. *Geophysical Journal International*, *190*, 1181–1196. https://doi.org/10.1111/j.1365-246x.2012.05544.x

Ammon, C. J., & Vidale, J. E. (1993). Tomography without rays. *Bulletin of the Seismological Society of America*, *83*(2), 509–528.

Chen, P., Zhao, L., & Jordan, T. H. (2007). Full 3D tomography for the crustal structure of the Los Angeles region. *Bulletin of the Seismological Society of America*, *97*(4), 1094–1120. https://doi.org/10.1785/0120060222

Dahlen, F. A., Hung, S.-H., & Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes-I. Theory. *Geophysical Journal International*, *141*, 157–174. https://doi.org/10.1046/j.1365-246x.2000.00070.x

Doi, I., Noda, S., Iio, Y., Horiuchi, S., & Sekiguchi, S. (2013). Relationship between hypocentral distributions and Vp/Vs ratio structures inferred from dense seismic array data: A case study of the 1984 western Nagano Prefecture earthquake, central Japan. *Geophysical Journal International*, *195*, 1323–1336. https://doi.org/10.1093/gji/ggt312

Fang, H., Zhang, H., Yao, H., Allam, A., Zigone, D., Ben-Zion, Y., et al. (2016). A new algorithm for three-dimensional joint inversion of body wave and surface wave data and its application to the Southern California plate boundary region. *Journal of Geophysical Research: Solid Earth*, *121*, 3557–3569. https://doi.org/10.1002/2015jb012702

Fichtner, A., Bunge, H.-P., & Igel, H. (2006). The adjoint method in seismology. *Physics of the Earth and Planetary Interiors*, *157*, 86–104. https://doi.org/10.1016/j.pepi.2006.03.016

Fletcher, R., & Reeves, C. (1964). Function minimization by conjugate gradients. *The Computer Journal*, *7*, 149–154. https://doi.org/10.1093/comjnl/7.2.149

Geiger, L. (1910). Herdbestimmung bei erdbeden ans den ankunftzeiten. *Koniglichen Gesellschaft der Wissenschaften Zu Gottingen*, *4*, 331.

Gou, T., Zhao, D., Huang, Z., & Wang, L. (2018). Anisotropic 3-D ray tracing and its application to Japan subduction zone. *Journal of Geophysical Research: Solid Earth*, *123*(5), 4088–4108. https://doi.org/10.1029/2017jb015321

Hauksson, E., & Shearer, P. M. (2006). Attenuation models (QP and QS) in three dimensions of the southern California crust: Inferred fluid saturation at seismogenic depths. *Journal of Geophysical Research*, *111*(B05), 302. https://doi.org/10.1029/2005jb003947

Hauksson, E., Yang, W., & Shearer, P. M. (2012). Waveform relocated earthquake catalog for southern California (1981 to June 2011). *Bulletin of the Seismological Society of America*, *102*(5), 2239–2244. https://doi.org/10.1785/0120120010

Huang, X., Yang, D., Tong, P., Badal, J., & Liu, Q. (2016). Wave equation-based reflection tomography of the 1992 Landers earthquake area. *Geophysical Research Letters*, *43*, 1884–1892. https://doi.org/10.1002/2016gl067717

Jennings, C. W., & Bryant, W. A. (2010). Fault activity map of California. California Geological Survery Geologic Data Map, 2, Scale1:750000.

Koketsu, K., & Sekine, S. (1998). Pseudo-bending method for three-dimensional seismic ray tracing in a spherical earth with discontinuities. *Geophysical Journal International*, *132*, 339–346. https://doi.org/10.1046/j.1365-246x.1998.00427.x

Langenheim, V., Jachens, R., & Aiken, C. (2014). Geophysical framework of the Peninsular Ranges batholith?Implications for tectonic evolution and neotectonics. In *Peninsular Ranges batholith, Baja California and southern California*. Geological Society of America. https://doi.org/10.1130/2014.1211(0110.1130/2014.1211(01)

Lelièvre, P. G., Farquharson, C. G., & Hurich, C. A. (2011). Inversion of first-arrival seismic traveltimes without rays, implemented on unstructured grids. *Geophysical Journal International*, *185*, 749–763. https://doi.org/10.1111/j.1365-246x.2011.04964.x

Leung, S., & Qian, J. (2006). An adjoint state method for three-dimensional transmission traveltime tomography using first-arrivals. *Communications in Mathematical Sciences*, *4*, 249–266. https://doi.org/10.4310/cms.2006.v4.n1.a10

Li, J., Li, H., Chen, H., Su, J., & Tong, Y. P. (2020). Eikonal equation-based seismic tomography of the source areas of the 2008 Mw 7.9 Wenchuan earthquake and the 2013 Mw 6.6 Lushan earthquake. *Bulletin of the Seismological Society of America*, *110*(2), 886–897. https://doi.org/10.1785/0120190134

Liu, Q., & Tromp, J. (2006). Finite-frequency kernels based on adjoint methods. *Bulletin of the Seismological Society of America*, *96*, 2283–2297. https://doi.org/10.1785/0120060041

Liu, S., Suardi, I., Yang, D., Wei, S., & Tong, P. (2018). Teleseismic traveltime tomography of northern Sumatra. *Geophysical Research Letters*, *45*, GL078. https://doi.org/10.1029/2018gl078610

Liu, S., Suardi, I., Zheng, M., Yang, D., Huang, X., & Tong, P. (2019). Slab morphology beneath northern Sumatra revealed by regional and teleseismic traveltime tomography. *Journal of Geophysical Research: Solid Earth*, *124*(10), 544–564. https://doi.org/10.1029/2019jb017625

Moretti, M., De Gori, P., & Chiarabba, C. (2009). Earthquake relocation and three-dimensional Vp and Vp/Vs models along the low angle Alto Tiberina Fault (Central Italy): Evidence for fluid overpressure. *Geophysical Journal International*, *176*, 833–846. https://doi.org/10.1111/j.1365-246x.2008.03984.x

Qiu, H., Lin, F. C., & Ben-Zion, Y. (2019). Eikonal tomography of the southern California plate boundary region. *Journal of Geophysical Research: Solid Earth*, *124*, 9755–9779. https://doi.org/10.1029/2019jb017806

Rawlinson, N., Hauser, J., & Sambridge, M. (2008). Seismic ray tracing and wavefront tracking in laterally heterogeneous media. *Advances in Geophysics*, *49*, 203–273. https://doi.org/10.1016/s0065-2687(07)49003-3

Rawlinson, N., Pozgay, S., & Fishwick, S. (2010). Seismic tomography: A window into deep Earth. *Physics of the Earth and Planetary Interiors*, *178*, 101–135. https://doi.org/10.1016/j.pepi.2009.10.002

Rawlinson, N., & Sambridge, M. (2004). Wave front evolution in strongly heterogeneous layered media using the fast marching method. *Geophysical Journal International*, *156*, 631–647. https://doi.org/10.1111/j.1365-246x.2004.02153.x

Ross, Z. E., Hauksson, E., and Ben-Zion, Y. (2017). Abundant off-fault seismicity and orthogonal structures in the San Jacinto fault zone, *Science Advance*, *3*, e1601946. https://doi.org/10.1126/sciadv.1601946

SCEDC (2013). *Southern California earthquake center*. Caltech. Dataset. https://doi.org/10.7909/C3WD3xH1

Sei, A., & Symes, W. W. (1994). Gradient calculation of the traveltime cost function without ray tracing. *SEG Expanded Abstracts*, *13*, 1351–1354.

Share, P.-E., Guo, H., Thurber, C. H., Zhang, H., & Ben-Zion, Y. (2019). Seismic imaging of the Southern California plate boundary around the south-central transverse ranges using double-difference tomography. *Pure and Applied Geophysics*, *176*, 1117–1143. https://doi.org/10.1007/s00024-018-2042-3

Shaw, J. H., Plesch, A., Tape, C., Suess, M. P., Jordan, T. H., Ely, G., et al. (2015). Unified structural representation of the southern California crust and upper mantle. *Earth and Planetary Science Letters*, *415*(4), 1–15. https://doi.org/10.1016/j.epsl.2015.01.016

Taillandier, C., Noble, M., Chauris, H., & Calandra, H. (2009). First-arrival traveltime tomography based on the adjoint-state method. *Geophysics*, *74*, WCB57–WCB66. https://doi.org/10.1190/1.3250266

Tape, C., Liu, Q., & Tromp, J. (2007). Finite-frequency tomography using adjoint methods-methodology and examples using membrane surface waves. *Geophysical Journal International*, *168*, 1105–1129. https://doi.org/10.1111/j.1365-246x.2006.03191.x

Tavakoli, F., Operto, S., Ribodetti, A., & Virieux, J. (2017). Slope tomography based on eikonal solvers and the adjoint-state method. *Geophysical Journal International*, *209*, 1629–1647.

Thurber, C., & Eberhart-Phillips, D. (1999). Local earthquake tomography with flexible gridding. *Computers & Geosciences*, *25*, 809–818. https://doi.org/10.1016/s0098-3004(99)00007-2

Thurber, C. H. (1983). Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California. *Journal of Geophysical Research*, *88*, 8226–8236. https://doi.org/10.1029/jb088ib10p08226

Thurber, C. H. (1992). Hypocenter-velocity structure coupling in local earthquake tomography. *Physics of the Earth and Planetary Interiors*, *75*, 55–62. https://doi.org/10.1016/0031-9201(92)90117-e

Tong, P., D. Yang, Q. Liu, X. Yang, and J. Harris, (2016). Acoustic wave-equation-based earthquake location, *Geophysical Journal International*, *205*, 464–478. https://doi.org/10.1093/gji/ggw026

Tong, P., Yang, D., & Huang, X. (2019). Multiple-grid model parametrization for seismic tomography with application to the San Jacinto fault zone. *Journal of Intelligence*, *218*, 200–223. https://doi.org/10.1093/gji/ggz151

Tong, P., Yang, D., Li, D., & Liu, Q. (2017). Time-evolving seismic tomography: The method and its application to the 1989 Loma Prieta and 2014 South Napa earthquake area, California. *Geophysical Research Letters*, *44*, 3165–3175. https://doi.org/10.1002/2017gl072785

Tong, P., Yao, J., Liu, Q., Li, T., Wang, K., Liu, S., et al. (2021). Crustal rotation and fluids: Factors for the 2019 Ridgecrest earthquake sequence?. *Geophysical Research Letters*, *48*, e2020GL090853. https://doi.org/10.1029/2020GL090853

Tong, P., Zhao, D., Yang, D., Yang, X., Chen, J., & Liu, Q. (2014a). Wave-equation-based travel-time seismic tomography—Part 1: Method. *Solid Earth*, *5*, 1151–1168. https://doi.org/10.5194/se-5-1151-2014

Tong, P., Zhao, D., Yang, D., Yang, X., Chen, J., & Liu, Q. (2014b). Wave-equation-based travel-time seismic tomography—Part 2: Application to the 1992 Landers earthquake (Mw 7.3) area. *Solid Earth*, *5*, 1169–1188. https://doi.org/10.5194/se-5-1169-2014

Tromp, J., Tape, C., & Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, *160*, 195–216.

Um, J., & Thurber, C. (1987). A fast algorithm for two-point seismic ray tracing. *Bulletin of the Seismological Society of America*, *77*(3), 972–986.

Virieux, J., & Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, *74*, WCC1–WCC26. https://doi.org/10.1190/1.3238367

Waldhauser, F., & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. *Bulletin of the Seismological Society of America*, *90*(6), 1353–1368. https://doi.org/10.1785/0120000006

Wald, L. A., Hutton, L. K., & Given, D. D. (1995). The southern California network bulletin: 1990–1993 summary. *Seismological Research Letters*, *66*, 9–19. https://doi.org/10.1785/gssrl.66.1.9

Wang, K., Yang, Y., Basini, P., Tong, P., Tape, C., & Liu, Q. (2018). Refined crustal and uppermost mantle structure of southern California by ambient noise adjoint tomography. *Geophysical Journal International*, *215*(2), 844–863. https://doi.org/10.1093/gji/ggy312

Wang, X. Q., Schubnel, A., Fortin, J., David, E. C., Gueguen, Y., & Ge, H. K. (2012). High Vp/Vs ratio: Saturated cracks or anisotropy effects?. *Geophysical Research Letters*, *39*(L11), 307. https://doi.org/10.1029/2012gl051742

Wessel, P., & Smith, W. H. F. (1991). Free software helps map and display data, *Eos, Transactions American Geophysical Union*, *72*(1), 441–448. https://doi.org/10.1029/90eo00319

Zenonos, A., De Siena, L., Widiyantoro, S., & Rawlinson, N. (2020). Direct inversion of S-P differential arrival times for ratio in SE Asia. *Journal of Geophysical Research: Solid Earth*, *125*, e2019JB019. https://doi.org/10.1029/2019JB019152

Zhang, H., & Thurber, C. H. (2003). Double-difference tomography: The method and its application to the Hayward fault, California. *Bulletin of the Seismological Society of America*, *93*(5), 1875–1889. https://doi.org/10.1785/0120020190

Zhao, D., Hasegawa, A., & Horiuchi, S. (1992). Tomographic imaging of P and S wave velocity structure beneath northeastern Japan. *Journal of Geophysical Research*, *97*, 909–928. https://doi.org/10.1029/92jb00603

Zhao, D., & Kanamori, H. (1995). The 1994 Northridge Earthquake: 3-D crustal structure in the rupture zone and its relation to the aftershock locations and mechanisms. *Geophysical Research Letters*, *22*, 763–766. https://doi.org/10.1029/94gl03222

Zhao, D., Kitagawa, H., & Toyokuni, G. (2015). A water wall in the Tohoku forearc causing large crustal earthquakes. *Geophysical Journal International*, *200*, 149–172. https://doi.org/10.1093/gji/ggu381

Zhao, H. (2004). A fast sweeping method for Eikonal equations. *Mathematics of Computation*, *74*(250), 603–628. https://doi.org/10.1090/s0025-5718-04-01678-3

Zigone, D., Ben-Zion, Y., Campillo, M., & Roux, P. (2015). Seismic tomography of the southern California plate boundary region from noise-based Rayleigh and Love waves. *Pure and Applied Geophysics*, *172*, 1007–1032. https://doi.org/10.1007/s00024-014-0872-1