



TomoATT: An open-source package for Eikonal equation-based adjoint-state traveltime tomography for seismic velocity and azimuthal anisotropy

Jing Chen^a, Masaru Nagaso^a, Mijian Xu^{a,1}, Ping Tong^{a,b,*}

^a Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, Singapore

^b Earth Observatory of Singapore, Nanyang Technological University, Singapore, Singapore

ARTICLE INFO

Keywords:

Regional and teleseismic tomography
Anisotropic Eikonal equation
Adjoint-state traveltime tomography
Multi-level parallelization

ABSTRACT

TomoATT is an open-source software package, aiming at determining seismic velocity and azimuthal anisotropy based on adjoint-state traveltime tomography methods. Key features of TomoATT include Eikonal equation modeling, adjoint-state method, sensitivity kernel regularization, and multi-level parallelization. Through several toy experiments, we demonstrate TomoATT's capability in accurate forward modeling, handling multipathing phenomenon, delivering reliable tomographic results, and achieving high-performance parallelization. Additionally, TomoATT is benchmarked with a synthetic experiment and two real-data applications in central California near Parkfield and Thailand. The successful recovery of the synthetic model, along with the imaging results that are consistent with previous studies and regional tectonics, verifies the effectiveness of TomoATT. Each inversion starts with only three simple input files (about model, data, and parameters) and completes within 2 h using 64 processors. Overall, TomoATT offers an efficient and user-friendly tool for regional and teleseismic traveltime tomography, empowering researchers to image subsurface structures and deepen our understanding of the Earth's interior.

1. Introduction

We developed an open-source software package, TomoATT, to invert traveltimes and differential arrival times from regional and teleseismic earthquakes for velocity heterogeneity and azimuthal anisotropy. This package aims to address the growing need for accurate subsurface seismic imaging, as well as to publicize the innovative adjoint-state traveltime tomography (ATT) methods (J. Chen, G. Chen et al., 2023; J. Chen, S. Wu et al., 2023; Tong, 2021a).

Compared with ray-based and wave equation-based tomography methods, the salient features of the ATT methods include the Eikonal equation-based forward modeling and the adjoint-state method for kernel calculation. Solving the Eikonal equation balances accuracy and efficiency: it is accurate and robust for calculating synthetic traveltime, avoiding the potential inaccuracy of bending and shooting methods for ray tracing (Rawlinson et al., 2008; Vidale, 1988). Although it is typically more computationally expensive than ray tracing, its

computational complexity is an order of magnitude lower than that of solving a wave equation. Since the ATT methods use reliable traveltime data, the inversion is less affected by strong nonlinearity compared to waveform inversion (Mora, 1987; Alkhalifah, 2015), albeit at the cost of losing waveform information associated with multiple scattering. The adjoint-state method is applied to derive sensitivity kernels, which has proven effective in partially addressing the multipathing phenomenon (Tong et al., 2023), thus yielding accurate sensitivity kernels. During the inversion procedure, the step size-controlled gradient descent method is employed to update model parameters, with the multiple-grid parameterization and kernel density regularization incorporated to enhance the inversion reliability and accelerate convergence. These attributes make the ATT methods an efficient and robust tomographic approach for imaging subsurface velocity heterogeneity and seismic anisotropy (J. Chen, G. Chen et al., 2023; J. Chen, S. Wu et al., 2023; Tong, 2021a, 2021b; Wu et al., 2022).

The TomoATT package is developed in C++ based on the ATT

* Corresponding author. Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, Singapore.

E-mail address: tongping@ntu.edu.sg (P. Tong).

¹ Mijian Xu is now at the Department of Physics, University of Toronto, Canada.

<https://doi.org/10.1016/j.cageo.2025.105995>

Received 20 November 2024; Received in revised form 19 June 2025; Accepted 21 June 2025

Available online 22 June 2025

0098-3004/© 2025 Elsevier Ltd. All rights reserved, including those for text and data mining, AI training, and similar technologies.

methods. It utilizes the Message Passing Interface (MPI) and shared memory techniques for efficient parallelization. Popular file formats are supported for input and output operations, including YAML for parameter file, HDF5 for model file, and TEXT for data file. To improve user-friendliness, we also develop PyTomoATT, a companion Python module to streamline the processing of input and output files. All resources are accessible through the link provided in the Code Availability Section.

This paper provides a comprehensive introduction to TomoATT. In Section 2, we briefly review the ATT methods and present an overview of the package. Section 3 demonstrates the key features of TomoATT through numerical experiments. Furthermore, we evaluate the performance of TomoATT using both synthetic and real data in Section 4, showcasing its capability to resolve complex subsurface structures. By sharing our experiences and insights, we hope to encourage the geophysical community to adopt and further develop this open-source tool, advancing the field of seismic imaging and improving our understanding of subsurface dynamics.

2. Methodology and overview of TomoATT

2.1. Brief review of the ATT methods

Here we briefly review the ATT methods (Tong, 2021a; J. Chen, G. Chen et al., 2023; J. Chen, S. Wu et al., 2023; Tong et al., 2023). These methods aim at the determination of seismic slowness $s(\mathbf{x})$ (the reciprocal of velocity) and azimuthal anisotropy parameters $\xi(\mathbf{x})$ and $\eta(\mathbf{x})$ by minimizing the discrepancy between synthetic and observational traveltimes (e.g. Thurber, 1983; Kissling et al., 1994), common-source differential arrival times (e.g. Lomax, 2005; Zhang et al., 2010; Eberhart-Phillips and Reyners, 2012; de Vos et al., 2013; Yuan et al., 2016; Terhemba et al., 2022; Chen et al., 2023), and common-receiver differential arrival times (e.g. Waldhauser and Ellsworth, 2000; Zhang

$$\chi_{cr} = \sum_{n=1}^{N_s} \sum_{j=1}^{N_s} \sum_{m=1}^{N_r} \frac{w_{n,j|m}}{2} \left(\Delta T_{n,j|m} - \Delta T_{n,j|m}^{\text{obs}} \right)^2, \quad (4)$$

Here $T_n(\mathbf{x}_{r,m})$ and $T_n^{\text{obs}}(\mathbf{x}_{r,m})$ represent the synthetic and observational traveltimes of a particular seismic phase emanating from the n -th source at $\mathbf{x}_{s,n}$ and recorded by the m -th receiver at $\mathbf{x}_{r,m}$, respectively. The synthetic and observational common source differential arrival times associated with the common source $\mathbf{x}_{s,n}$ and two separate receivers $\mathbf{x}_{r,m}$ and $\mathbf{x}_{r,i}$ are given by

$$\Delta T_{n|m,i} = T_n(\mathbf{x}_{r,m}) - T_n(\mathbf{x}_{r,i}), \Delta T_{n|m,i}^{\text{obs}} = T_n^{\text{obs}}(\mathbf{x}_{r,m}) - T_n^{\text{obs}}(\mathbf{x}_{r,i}), \quad (5)$$

and the synthetic and observational common receiver differential arrival times associated with the common receiver $\mathbf{x}_{r,m}$ and two separate sources $\mathbf{x}_{s,n}$ and $\mathbf{x}_{s,j}$ are

$$\Delta T_{n,j|m} = T_n(\mathbf{x}_{r,m}) - T_j(\mathbf{x}_{r,m}), \Delta T_{n,j|m}^{\text{obs}} = T_n^{\text{obs}}(\mathbf{x}_{r,m}) - T_j^{\text{obs}}(\mathbf{x}_{r,m}). \quad (6)$$

The weight coefficients $w_{n|m}$, $w_{n|m,i}$, $w_{n,j|m}$ are determined by the quality of the observational data $T_n^{\text{obs}}(\mathbf{x}_{r,m})$, $\Delta T_{n|m,i}^{\text{obs}}$, and $\Delta T_{n,j|m}^{\text{obs}}$, respectively. N_s and N_r are the number of sources and receivers, respectively.

The ATT method derives the sensitivity kernels of the objective function with respect to model parameters based on the adjoint-state method (Leung and Qian, 2006) and solves the optimization problem iteratively (Fig. 1). Each iteration consists of four key steps:

1. Calculate the synthetic traveltime field $T_n(\mathbf{x})$ for each source $\mathbf{x}_{s,n}$ by solving the anisotropic Eikonal equation in spherical coordinates (Tsai et al., 2003; J. Chen, G. Chen et al., 2023)

$$[\nabla T_n(\mathbf{x})]^t M(\mathbf{x}; \xi, \eta) \nabla T_n(\mathbf{x}) \triangleq \begin{pmatrix} \partial_r T_n & \frac{\partial_\theta T_n}{r} & \frac{\partial_\phi T_n}{r \cos \theta} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - 2\xi & 2\eta \\ 0 & 2\eta & 1 + 2\xi \end{pmatrix} \begin{pmatrix} \frac{\partial_r T_n}{r} \\ \frac{\partial_\theta T_n}{r} \\ \frac{\partial_\phi T_n}{r \cos \theta} \end{pmatrix} = s^2(\mathbf{x}). \quad (7)$$

and Thurber, 2003; Pesicek et al., 2010; Guo and Zhang, 2017). The related problem is formulated as the following optimization problem

$$\min_{s, \xi, \eta} \chi = \alpha \chi_t + \beta \chi_{cs} + \gamma \chi_{cr}. \quad (1)$$

The objective function χ is a weighted sum of the misfit functions χ_t , χ_{cs} , χ_{cr} corresponding to traveltimes, common-source differential arrival times, and common-receiver differential arrival times, respectively. α , β , γ are weights for these three types of data (see Appendix A for a brief discussion of weights). The definitions are given by

$$\chi_t = \sum_{n=1}^{N_s} \sum_{m=1}^{N_r} \frac{w_{n|m}}{2} \left(T_n(\mathbf{x}_{r,m}) - T_n^{\text{obs}}(\mathbf{x}_{r,m}) \right)^2, \quad (2)$$

$$\chi_{cs} = \sum_{n=1}^{N_s} \sum_{m=1}^{N_r} \sum_{i=1}^{N_r} \frac{w_{n|m,i}}{2} \left(\Delta T_{n|m,i} - \Delta T_{n|m,i}^{\text{obs}} \right)^2, \quad (3)$$

The boundary condition is set as $T_n(\mathbf{x}_{s,n}) = 0$ for earthquakes within the study region Ω or, for teleseismic earthquakes, is set as the travel time from the source to the study region boundary $\partial\Omega$, by solving 2D Eikonal equations (J. Chen, S. Wu et al., 2023). Here ∇ is the gradient operator and the superscript t is the transpose operator. $\mathbf{x} = (r, \theta, \phi)$ is the spherical coordinate, in which r denotes the distance to the Earth's center, θ is latitude, and ϕ is longitude.

2. Calculate the adjoint field $P_n(\mathbf{x})$ based on each traveltime field $T_n(\mathbf{x})$ by solving the adjoint equation (Leung and Qian, 2006)

$$\nabla \cdot (P_n(\mathbf{x}) ([-\nabla T_n(\mathbf{x})]^t M(\mathbf{x}; \xi, \eta))) = \sum_{m=1}^{N_r} R_{n,m} \delta(\mathbf{x} - \mathbf{x}_{r,m}). \quad (8)$$

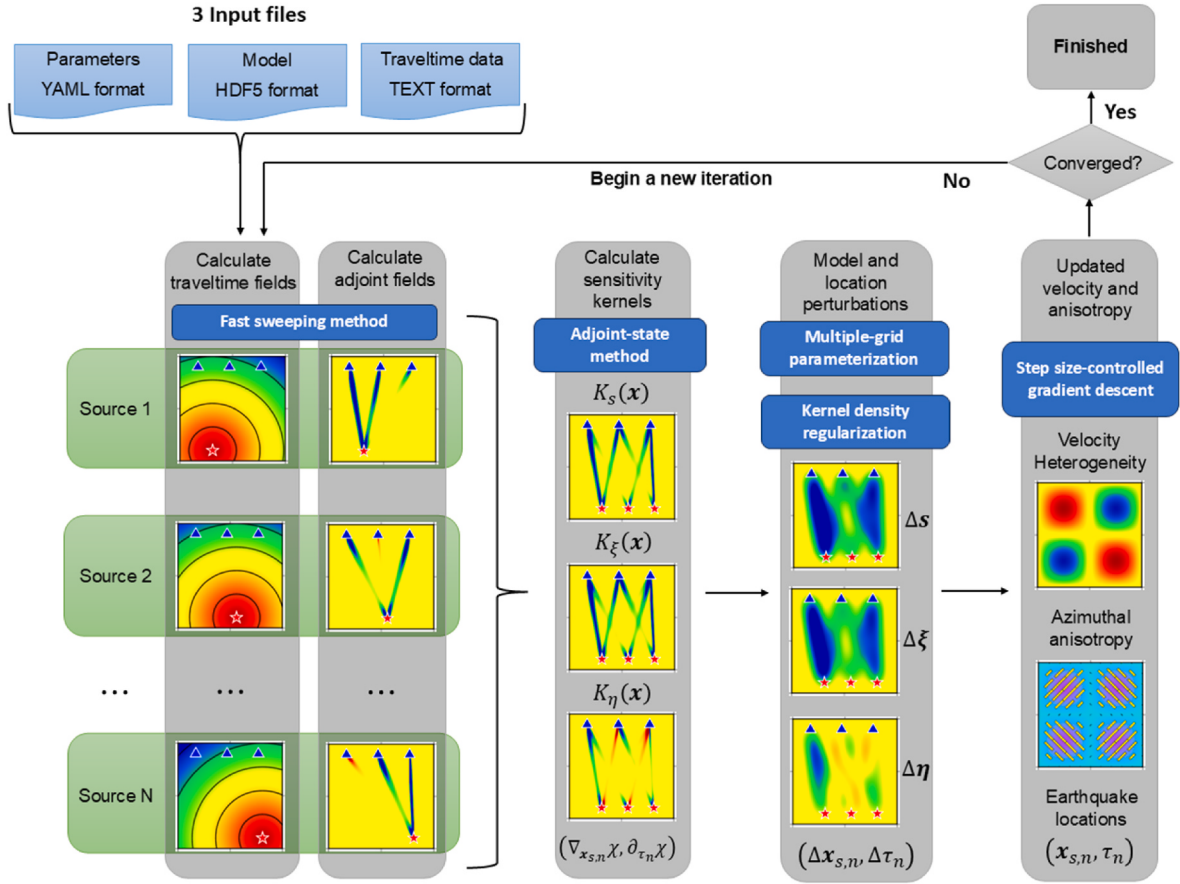


Fig. 1. The workflow of TomoATT to implement adjoint-state traveltime tomography, including anisotropic velocity model determination and earthquake location.

Here $\delta(\cdot)$ is the Dirac delta function. The adjoint source takes the form of $R_{n,m} = \frac{\partial \chi}{\partial T_n(\mathbf{x}_{r,m})}$, specified to be

$$R_{n,m} = w_{n|m}(T_n(\mathbf{x}_{r,m}) - T_n^{obs}(\mathbf{x}_{r,m})) + 2 \sum_{i=1}^{N_r} w_{n|m,i} (\Delta T_{n|m,i} - \Delta T_{n|m,i}^{obs}) + 2 \sum_{j=1}^{N_s} w_{n,j|m} (\Delta T_{n,j|m} - \Delta T_{n,j|m}^{obs}). \quad (9)$$

The derivation of $R_{n,m}$ is elaborated in [Appendix B](#).

3. Calculate the sensitivity kernels with respect to slowness $s(\mathbf{x})$ and anisotropic parameters $\xi(\mathbf{x}), \eta(\mathbf{x})$ based on the adjoint-state method ([J. Chen, G. Chen et al., 2023](#))

$$K_s(\mathbf{x}) = \sum_{n=1}^{N_s} P_n(\mathbf{x}) s^2(\mathbf{x}), \quad (10)$$

$$K_\xi(\mathbf{x}) = \sum_{n=1}^{N_s} P_n(\mathbf{x}) \left(\left(\frac{\partial_\theta T_n(\mathbf{x})}{r} \right)^2 - \left(\frac{\partial_\phi T_n(\mathbf{x})}{r \cos \theta} \right)^2 \right), \quad (11)$$

$$K_\eta(\mathbf{x}) = \sum_{n=1}^{N_s} -2P_n(\mathbf{x}) \frac{\partial_\theta T_n(\mathbf{x})}{r} \frac{\partial_\phi T_n(\mathbf{x})}{r \cos \theta}. \quad (12)$$

An approximate linear relationship between the perturbations of model parameters $\delta s(\mathbf{x}), \delta \xi(\mathbf{x}), \delta \eta(\mathbf{x})$ and objective function $\delta \chi$ is accordingly constructed as follows

$$\delta \chi = \int_{\Omega} K_s(\mathbf{x}) \frac{\delta s(\mathbf{x})}{s(\mathbf{x})} d\mathbf{x} + \int_{\Omega} K_\xi(\mathbf{x}) \delta \xi(\mathbf{x}) d\mathbf{x} + \int_{\Omega} K_\eta(\mathbf{x}) \delta \eta(\mathbf{x}) d\mathbf{x}. \quad (13)$$

4. Update $s(\mathbf{x}), \xi(\mathbf{x}), \eta(\mathbf{x})$ using the step size-controlled gradient descent method ([J. Chen, G. Chen et al., 2023](#)). The multiple-grid parameterization ([Tong et al., 2019](#)) and kernel density regularization techniques are applied to enhance reliability.

The influence of source uncertainty on imaging results can be mitigated by updating hypocenter $\mathbf{x}_{s,n}$ and origin time τ_n during the inversion using the gradient descent method. The sensitivity kernels of the objective function with respect to $\mathbf{x}_{s,n}$ and τ_n take the general forms of

$$K_{\mathbf{x}_{s,n}} = \sum_{m=1}^{N_r} \frac{\partial \chi}{\partial \Gamma_m(\mathbf{x}_{s,n})} \nabla \Gamma_m(\mathbf{x}_{s,n}) \text{ and } K_{\tau_n} = \frac{\partial \chi}{\partial \tau_n}, \quad (14)$$

which are specified in [Appendix C](#).

2.2. Overview of the TomoATT package

One of the core features of the package is its robust utilization of the MPI, which allows for effective multi-level parallelization. This design empowers users to leverage high-performance computing (HPC) platforms, maximizing computational resource efficiency, which is particularly beneficial for large-scale seismic imaging applications. Additionally, we employ shared memory techniques to optimize memory usage and enhance parallel processing efficiency.

To improve accessibility, TomoATT supports popular file formats for input and output operations. The package accommodates the YAML format for parameter files, which provides an intuitive way for users to configure their applications. For storing model parameters, the parallel version of HDF5 is utilized, renowned for its ability to handle large dataset efficiently. This format allows multiple processes to read from

and write to data files simultaneously, enhancing data access speed and overall computational efficiency. Additionally, a TEXT format is supported for managing input traveltime data, often characterized by a non-regular structure. To address the challenges posed by this structure, we developed PyTomoATT, a companion Python module designed to process and handle non-regular data effectively. PyTomoATT includes functionalities for data filtering, weight assignment, coordinate rotation, and other preprocessing tasks. Moreover, it provides tools for model generation and reading, model slicing, checkerboard generation, and the creation and modification of parameter files. These capabilities work together to streamline the preparation and management of input and output files, enabling researchers to set up the inversion process efficiently without needing extensive programming knowledge.

All resources related to TomoATT, including the core package, user documentation, and PyTomoATT, are licensed under the GNU General Public License v3.0. These tools are compatible with multiple platforms, including MacOS, Linux, and Windows through Windows Subsystem for Linux (WSL). TomoATT can be compiled with various compilers, including GNU, Intel, and Clang, ensuring broad accessibility for users across different operating systems and development environments. The tools and documentation are available through the link provided in the Code availability section.

3. Key features of TomoATT

In this section, we design several toy experiments to demonstrate the key features of TomoATT, including forward modeling, adjoint-state method, multiple-grid parameterization, kernel density regularization, and multi-level parallelization.

3.1. Forward modeling

TomoATT solves the anisotropic Eikonal equation for the traveltime field of wavefront propagation from the source to any positions within the study region. We select the fast sweeping method (FSM) (Zhao, 2005) as the Eikonal solver. This grid-based method has proven unconditionally convergent to the solution (Zhao, 2005), and achieves the optimal computational complexity of $O(N)$, where N is the total number of grid nodes. Additionally, we use the multiplicative factorization technique (Luo and Qian, 2012) to eliminate source singularity and solve the equation in spherical coordinates to account for Earth's

curvature (J. Chen, G. Chen et al., 2023), further improving accuracy. Here we design two toy models to evaluate the accuracy of FSM for calculating traveltime and teleseismic differential arrival time, respectively.

First, we consider a linear velocity model within the domain ranging from $[0, 20^\circ] \times [0, 400 \text{ km}]$, where the velocity approximately increases from 4.5 km/s at the surface to 9 km/s at the bottom (Fig. 2a). The earthquake is located at the center $\mathbf{x}_s = (10^\circ, 200 \text{ km})$. The Eikonal equation is solved on four meshes with grid spacings of 20, 10, 5, and 2.5 km, respectively. The L_1 norm error between the calculated and analytical solutions $\|T_{\text{cal}}(\mathbf{x}) - T_{\text{true}}(\mathbf{x})\|_{L_1}$ decreases as grid spacing decreases, from 0.0517 s to 0.0096 s (Fig. 2c), demonstrating the expected first-order accuracy (Luo and Qian, 2012). We also measure the traveltime errors at stations on the surface, which accumulate with epicenter distance (Fig. 2b). The results show that the maximum error in this model remains below 0.1 s for epicenter distances less than 800 km, provided grid spacing is less than 5 km.

Second, we consider the global AK135 model (Kennett et al., 1995) within the region $[11^\circ\text{N}, 21^\circ\text{N}] \times [96^\circ\text{E}, 106^\circ\text{E}] \times [0, 500 \text{ km}]$. A teleseismic earthquake is located at $(5^\circ\text{S}, 142^\circ\text{E}, 20 \text{ km})$, with an epicenter distance of 45.68° from the center of the study region (Fig. 3a). 81 stations are evenly distributed on the surface (Fig. 3b). We solve the Eikonal equation on a mesh with grid spacing of $10 \text{ km} \times 10 \text{ km} \times 5 \text{ km}$, and calculate common-source differential traveltimes between stations less than 300 km apart. Since the analytical solution is unavailable, we compare results obtained from the Eikonal solver with those from the TauP software (Crotwell et al., 1999). The standard deviation of misfits $(\Delta T_{\text{cal}} - \Delta T_{\text{TauP}})$ is 0.03 s (Fig. 3c), verifying the accuracy of our Eikonal solver.

3.2. Sensitivity kernels based on the adjoint-state method

TomoATT computes sensitivity kernels using the adjoint-state method. It solves for adjoint field in the study region, which describes the transportation of adjoint sources from receivers to the source along the opposite direction of wavefront propagation (e.g., along the negative direction of traveltime gradient in isotropic media). The main advantage is that this approach measures the sensitivity field throughout the study region, rather than restricting it to a single ray path. Therefore, it effectively addresses the multipathing phenomenon, where the wavefront reaches one receiver along multiple paths with similar traveltimes

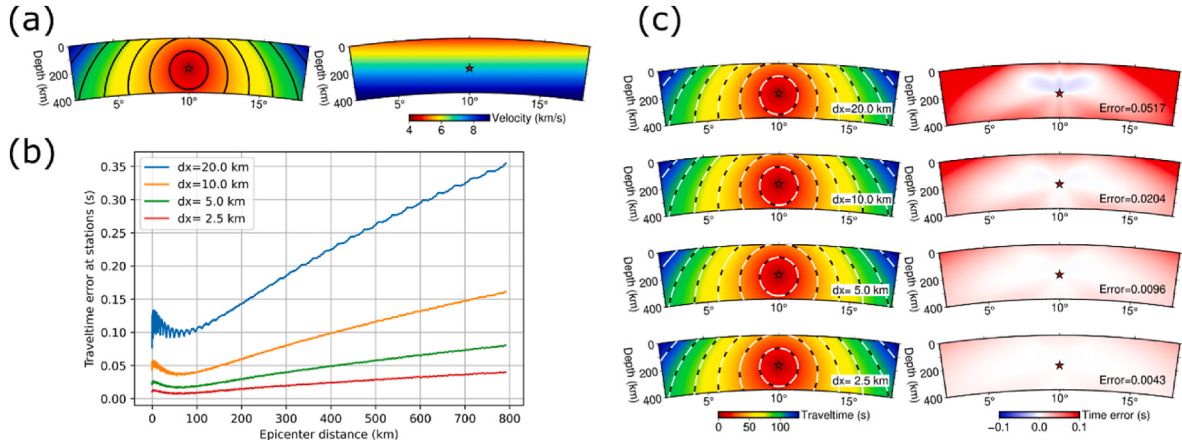


Fig. 2. Validation of the Eikonal solver for calculating traveltime. (a) The true traveltime field $T_{\text{true}}(\mathbf{x})$ calculated in the velocity model. The red star indicates the earthquake. The black curves represent traveltime isochrones at an interval of 20 s. (b) Traveltime errors at surface stations (depth = 0 km) relative to epicenter distance. The blue, yellow, green, and red curves show numerical errors with grid spacings of 20, 10, 5, and 2.5 km, respectively. Oscillations in the blue curve may arise from interpolation errors on the sparse grid. (c) The left column shows the calculated traveltime fields $T_{\text{cal}}(\mathbf{x})$ with grid spacings of 20, 10, 5, and 2.5 km. True and calculated traveltime isochrones are denoted by black and white dashed curves, respectively. The right column displays the numerical error fields $T_{\text{cal}}(\mathbf{x}) - T_{\text{true}}(\mathbf{x})$ for each grid spacing, with the L_1 norm errors quantified. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

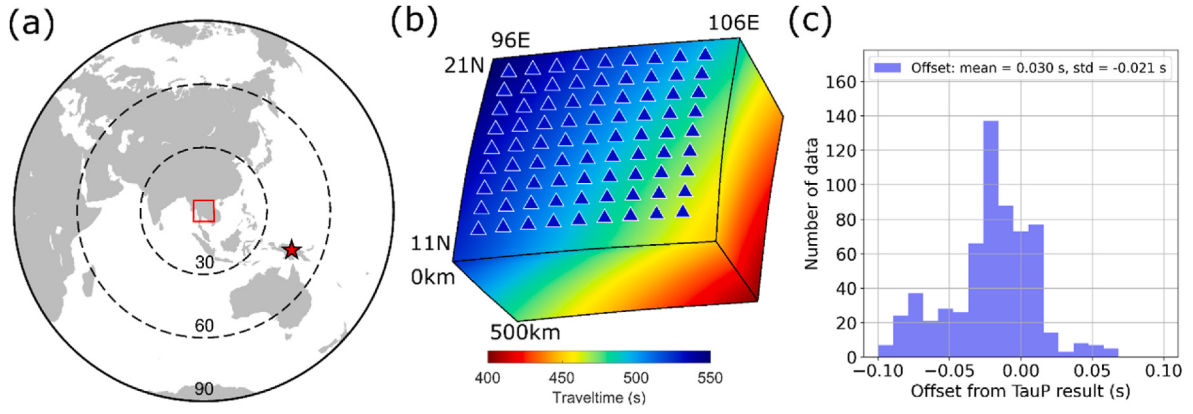


Fig. 3. Validation of the Eikonal solver for teleseismic differential arrival time. (a) The teleseismic earthquake (red star) and the study region (red box). Black dashed circles denote epicenter distances of 30° and 60°. (b) Calculated traveltime field within the study region. The blue triangles denote stations deployed on the surface. (c) Histogram of differential arrival misfits $\Delta T_{cal} - \Delta T_{TauP}$. The mean value and standard deviation are -0.021 s and 0.03 s, respectively, verifying the accuracy of the Eikonal solver. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

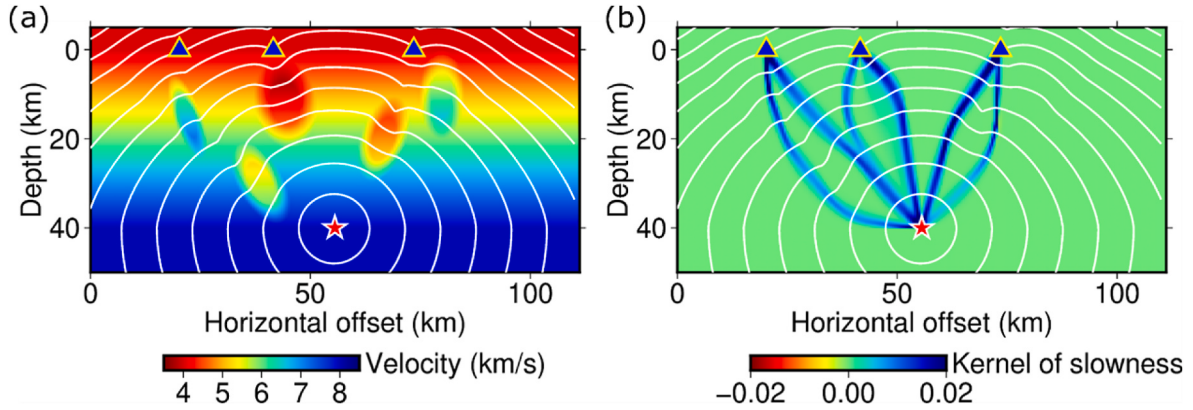


Fig. 4. A toy experiment illustrating the multipathing phenomenon. (a) P-wave velocity model. The blue triangles denote stations, and the red star marks the earthquake. White curves represent traveltime isochrones at an interval of 1 s. (b) Sensitivity kernel computed with artificial adjoint sources $R_{n,m} = 1$. The adjoint-state method captures the sensitivity along multiple paths, demonstrating its capability to address the multipathing phenomenon. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(Tong et al., 2023). The multipathing phenomenon is not pervasive and occurs only where the traveltime gradient is discontinuous, possibly having limited impact on the inversion. Nevertheless, it is still preferable to address this potential issue using the adjoint-state method rather than ignoring it.

Here we present an example to illustrate this issue. We consider a linear velocity model with several high- and low-velocity anomalies embedded (Fig. 4a). An earthquake is located at 40 km depth, and three stations are deployed on the surface. Two paths exist from the source to each station with similar minimum traveltimes. Consequently, the sensitivity kernel should be supported on both paths, as the velocity perturbation along either path would affect the traveltime. The adjoint-state method successfully provides correct sensitivity kernels by measuring the sensitivity throughout the study region (Fig. 4b). In contrast, without special treatment, ray-based methods identify only a single ray path, resulting in an incomplete sensitivity kernel.

3.3. Multiple-grid parameterization and kernel density normalization

After obtaining the sensitivity kernels on discretized grid nodes, TomoATT applies two regularization techniques to enhance inversion reliability: multiple-grid parameterization and kernel density normalization. The grid nodes for forward modeling are typically dense to ensure the accuracy. Directly updating model parameters on the forward grid nodes is unsuitable, as the number of variables may exceed the

resolving ability of the limited traveltime data. The multiple-grid parameterization (Tong et al., 2019) has been proposed to address this problem.

The multiple-grid parameterization method designs H sets of coarse inversion grids, denoting the nodes of the h -th grid as $(r_i^h, \theta_j^h, \phi_k^h)$. A series of basis functions are defined based on these inversion grids, given by

$$B_l^h(r, \theta, \phi) = \frac{1}{H} u_i^h(r) v_j^h(\theta) w_k^h(\phi), l = i + N_i^h(j-1) + N_i^h N_j^h(k-1), \quad (15)$$

in which

$$u_i^h(r) = \begin{cases} (r - r_{i-1}^h) / (r_i^h - r_{i-1}^h), & r_{i-1}^h \leq r \leq r_i^h, \\ (r_{i+1}^h - r) / (r_{i+1}^h - r_i^h), & r_i^h \leq r \leq r_{i+1}^h, \\ 0, & \text{otherwise.} \end{cases} \quad (16)$$

$$v_j^h(\theta) = \begin{cases} (\theta - \theta_{j-1}^h) / (\theta_j^h - \theta_{j-1}^h), & \theta_{j-1}^h \leq \theta \leq \theta_j^h, \\ (\theta_{j+1}^h - \theta) / (\theta_{j+1}^h - \theta_j^h), & \theta_j^h \leq \theta \leq \theta_{j+1}^h, \\ 0, & \text{otherwise.} \end{cases} \quad (17)$$

$$w_i^h(\phi) = \begin{cases} (\phi - \phi_{k-1}^h) / (\phi_k^h - \phi_{k-1}^h), & \phi_1^h \leq \phi_{k-1}^h \leq \phi \leq \phi_k^h, \\ (\phi_{k+1}^h - \phi) / (\phi_{k+1}^h - \phi_k^h), & \phi_k^h \leq \phi \leq \phi_{k+1}^h \leq \phi_{N_K}^h, \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

A model perturbation field is then assumed to be represented by a linear combination of these basis functions

$$\delta m(\mathbf{x}) = \sum_{l,h} \delta C_l^h B_l^h(\mathbf{x}). \quad (19)$$

Consequently, the approximate linear relationship (13) between the perturbations in model parameters and the objective functions is modified to

$$\delta \chi = \sum_{l,h} \frac{\partial \chi}{\partial C_{l,s}^h} \delta C_{l,s}^h + \sum_{l,h} \frac{\partial \chi}{\partial C_{l,\xi}^h} \delta C_{l,\xi}^h + \sum_{l,h} \frac{\partial \chi}{\partial C_{l,\eta}^h} \delta C_{l,\eta}^h, \quad (20)$$

where the new sensitivity kernel with respect to auxiliary parameters $(C_{l,s}^h, C_{l,\xi}^h, C_{l,\eta}^h)$ are

$$\frac{\partial \chi}{\partial C_{l,s}^h} = \int_{\Omega} K_s(\mathbf{x}) B_l^h(\mathbf{x}) d\mathbf{x}, \quad \frac{\partial \chi}{\partial C_{l,\xi}^h} = \int_{\Omega} K_{\xi}(\mathbf{x}) B_l^h(\mathbf{x}) d\mathbf{x}, \quad \frac{\partial \chi}{\partial C_{l,\eta}^h} = \int_{\Omega} K_{\eta}(\mathbf{x}) B_l^h(\mathbf{x}) d\mathbf{x}. \quad (21)$$

This method projects the model perturbation $\delta s(\mathbf{x})$, $\delta \xi(\mathbf{x})$, $\delta \eta(\mathbf{x})$ from an infinite-dimensional function space into a finite-dimensional vector space represented by $(C_{l,s}^h, C_{l,\xi}^h, C_{l,\eta}^h)$. Thus, the number of variables is decreased, and the sensitivity kernel is smoothed by convoluting it with the basis function $B_l^h(\mathbf{x})$. Additionally, averaging contributions across multiple inversion grids mitigates potential error from subjective selection of a single inversion grid, thereby, enhancing inversion stability.

Real-data inversions often suffer from uneven data distribution, leading to slower model updates in regions with sparse data coverage. Various methods have been proposed to address this issue, such as event declustering (Tong, 2021b) and assigning weights (Ruan et al., 2019). Here, we consider the characteristics of the adjoint-state method and propose an innovative approach. Taking the slowness sensitivity kernel $K_s(\mathbf{x})$ as an example, the sensitivity of the data associated with the n -th event, referred to as the event kernel, is given by

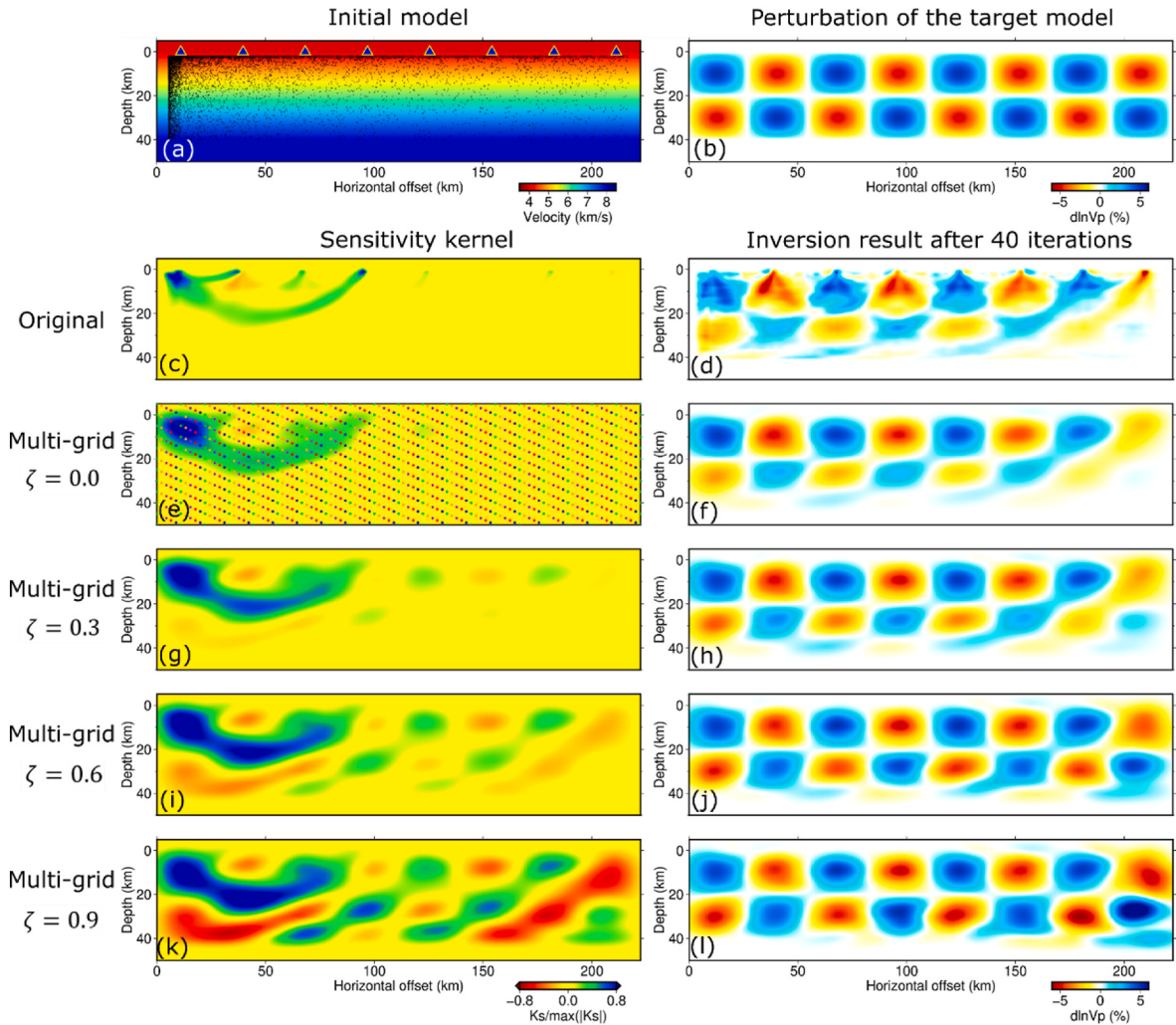


Fig. 5. A toy experiment demonstrating multiple-grid parameterization and kernel density normalization in tomography. (a) Initial velocity model with 8 stations on the surface (blue triangles) and 10,000 unevenly distributed earthquakes. (b) Velocity perturbation of the target model relative to the initial model. (c) Original sensitivity kernel at the first iteration, rescaled to $[-1, 1]$. (d) Imaging result after 40 iterations using the original sensitivity kernel. (e)–(f) Sensitivity kernel and imaging result after applying multiple-grid parameterization. Dots of the same color represents nodes in the same inversion grid. (g)–(l) Sensitivity kernels and imaging results after applying both multiple-grid parameterization and kernel density regularization with different coefficients $\zeta = 0.3, 0.6, 0.9$. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

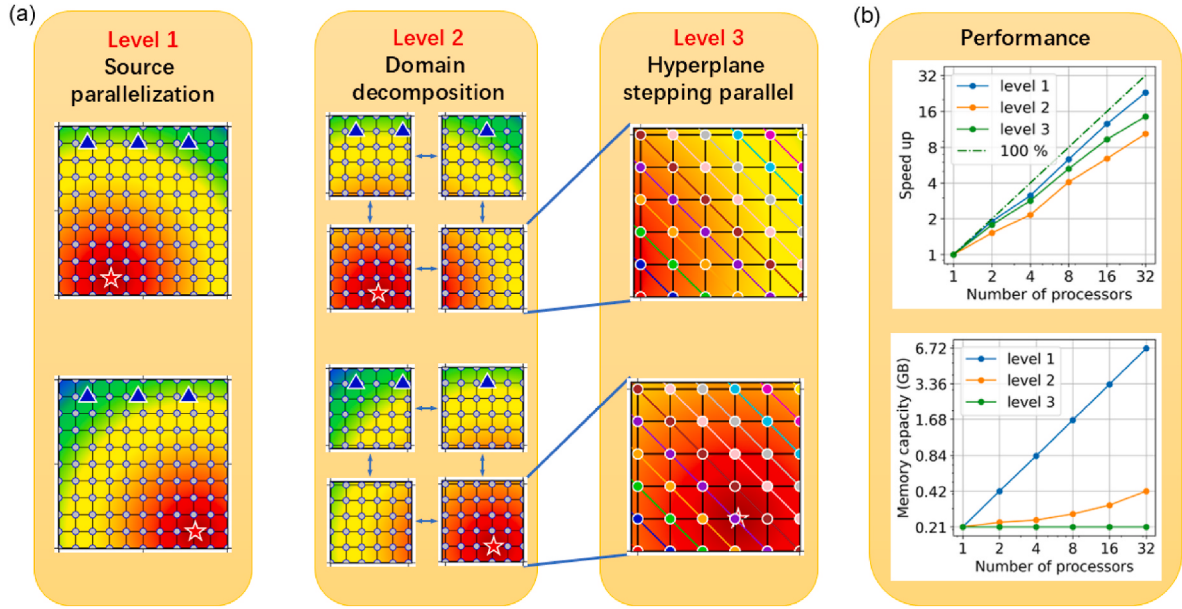


Fig. 6. Illustration of multi-level parallelization in TomoATT. (a) Level 1: Source parallelization demonstrates numerically solving Eikonal equations in parallel for two earthquakes. The blue triangles denote stations, and the red star marks the earthquake. Gray circles denote grid nodes discretizing the computational domain. Level 2: A 2-by-2 domain decomposition. Arrows indicate communications between processors managing adjacent subdomains. Level 3: Hyperplane stepping parallelization. All grid nodes within a subdomain are categorized into multiple hyperplanes (oblique lines in different colors). The grid nodes on the same hyperplane (circles of the same color) can be processed in parallel. (b) Log-log plots showing the speed-up and memory usage relative to the number of processors for each parallelization method. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

$$K_{s,n}(\mathbf{x}) = P_n(\mathbf{x})s^2(\mathbf{x}). \quad (22)$$

The key feature is the adjoint field $P_n(\mathbf{x})$, which satisfies the adjoint equation (8). This equation describes the backward transportation of data residuals, consisting of $T_n(\mathbf{x}_{r,m}) - T_n^{\text{obs}}(\mathbf{x}_{r,m})$, $\Delta T_{n|m,i} - \Delta T_{n|m,i}^{\text{obs}}$, and $\Delta T_{n|j,m} - \Delta T_{n|j,m}^{\text{obs}}$, from the receivers to the source along the negative direction of wavefront propagation. In contrast, if we do not use the data residuals but instead consider the number of data, we can obtain the so-called event kernel density. This can be easily implemented by introducing a specific adjoint source

$$\hat{R}_{n,m} = w_{n|m} \cdot 1 + 2 \sum_{i=1}^{N_r} (w_{n|m,i} \cdot 1) + 2 \sum_{j=1}^{N_s} (w_{n|j,m} \cdot 1). \quad (23)$$

Compared with the original adjoint source $R_{n,m}$ in Equation (9), the only difference lies in replacing the data residuals with the number of data. Then, the specific adjoint field $\hat{P}_n(\mathbf{x})$, satisfying a similar adjoint equation

$$\nabla \cdot (\hat{P}_n(\mathbf{x}) [-\nabla T_n(\mathbf{x})]^t M(\mathbf{x}; \xi, \eta)) = \sum_{m=1}^{N_r} \hat{R}_{n,m} \delta(\mathbf{x} - \mathbf{x}_{r,m}), \quad (24)$$

can be used to characterize the event kernel density. Summing up $\hat{P}_n(\mathbf{x})$ across all sources yields the sensitivity kernel density

$$K_d(\mathbf{x}) = \sum_{n=1}^{N_s} \hat{P}_n(\mathbf{x}). \quad (25)$$

The kernel density $K_d(\mathbf{x})$ is subsequently used to normalize the sensitivity kernel

$$\begin{aligned} K_s(\mathbf{x}) &\leftarrow \frac{K_s(\mathbf{x})}{(K_d(\mathbf{x}) + \epsilon)^\zeta}, \\ K_\xi(\mathbf{x}) &\leftarrow \frac{K_\xi(\mathbf{x})}{(K_d(\mathbf{x}) + \epsilon)^\zeta}, \\ K_\eta(\mathbf{x}) &\leftarrow \frac{K_\eta(\mathbf{x})}{(K_d(\mathbf{x}) + \epsilon)^\zeta}. \end{aligned} \quad (26)$$

The small value ϵ avoids division by zero, and the coefficient ζ controls the degree of normalization. The normalization emphasizes the contribution of data in sparsely constrained regions, which accelerates the model update in these areas. However, it also magnifies the influence of noise on imaging results. Thus, an appropriate coefficient ζ should be carefully determined based on data distribution and noise level.

Here we present a toy experiment to illustrate the multiple-grid parameterization and kernel density regularization. The study region is defined as $[0 \text{ km}, 220 \text{ km}] \times [0 \text{ km}, 50 \text{ km}]$, with 8 stations evenly deployed on the surface. To mimic an uneven earthquake distribution, we randomly place 10,000 earthquakes, with a higher probability in the upper left (Fig. 5a). The initial model is an isotropic model where velocity increases from 4 km/s at the surface to 8 km/s at 40 km depth. The target model is designed as a checkerboard model by adding staggered velocity perturbations to the initial model (Fig. 5b). The observational data is generated by calculating synthetic traveltimes in the target model and adding random Gaussian noise with a standard derivation of 0.1 s. We start with the initial model and invert the observational data to examine the recovery of checkerboard anomalies.

First, we perform the inversion without applying multiple-grid parameterization and kernel density normalization. The sensitivity

kernel at the first iteration and the final inversion result are displayed in Fig. 5c and d. Numerous small artificial anomalies are observable at 1 km, and the geometry of anomalies above 20 km depth is distorted. These likely arise from the excessive number of variables leading to unreliable inversion results. Second, we apply the multiple-grid parameterization, which eliminates the artifacts and produces a more reasonable inversion result (Fig. 5e and f). However, due to uneven earthquake distribution, the anomalies at the right bottom are less well recovered. Furthermore, we apply the kernel density normalization with different coefficients $\zeta = 0.3, 0.6, 0.9$ (Fig. 5g–i). The normalization indeed emphasizes the sensitivity kernel at the right bottom, leading to better recovery of anomalies and reducing smearing caused by uneven earthquake distribution. A higher value of ζ leads to a higher degree of anomaly recovery; however, it also magnifies the artificial anomalies below 40 km depth at the right bottom due to insufficient data constraints.

3.4. Multi-level parallelization

In TomoATT, solving anisotropic Eikonal equations and adjoint equations at each iteration constitutes the primary computation cost. To fully utilize computational resources on high-performance-computing platforms, TomoATT implements three levels of parallelization (Fig. 6a): Source parallelization, domain decomposition (Zhao, 2007), and hyperplane stepping parallelization (Detrixhe and Gibou, 2016; Detrixhe et al., 2013).

Source parallelization (Fig. 6a, Level 1) is the most straightforward approach, as solving Eikonal and adjoint equations corresponding to multiple sources are completely independent. Consequently, source parallelization can achieve high parallel efficiency. The primary drawback is the significant memory required for simultaneously solving three-dimensional (3-D) equations, which may impose a considerable memory burden.

Domain decomposition is a widely applicable parallelization approach, which divides the 3D study region into $I \times J \times K$ subdomains and assigns one processor to solve equations within each subdomain (Fig. 6a, Level 2). As a result, each processor only needs to process $N_{i,j,k} = N_{x,i} \times N_{y,j} \times N_{z,k}$ grid nodes within one subdomain, rather than the total grid nodes $N = \left(\sum_{i=1}^I N_{x,i}\right) \times \left(\sum_{j=1}^J N_{y,j}\right) \times \left(\sum_{k=1}^K N_{z,k}\right)$ within the whole study region. Its main advantage is that it requests almost no additional memory. However, a drawback is the increased computational cost due to boundary communications between processors managing adjacent subdomains. More importantly, increasing the number of subdomains may result in more iterations for the FSM to converge, decreasing parallel efficiency.

After applying source parallelization (Level 1) and domain decomposition (Level 2), $N_{i,j,k} = N_{x,i} \times N_{y,j} \times N_{z,k}$ grid nodes within one subdomain are typically handled by one processor. Hyperplane stepping parallelization enables multiple processors to solve equations within a single subdomain simultaneously (Fig. 6a, Level 3). In this method, $N_{x,i} \times N_{y,j} \times N_{z,k}$ grid nodes within one subdomain are categorized into $N_{x,i} + N_{y,j} + N_{z,k} - 2$ hyperplanes. Since the grid nodes on the same hyperplane are independently handled in the fast sweeping method, they can be processed in parallel by multiple processors. This approach, like domain decomposition, requires no additional memory. By leveraging a shared

memory technique in MPI, the multiple processors can read and write values on $N_{x,i} \times N_{y,j} \times N_{z,k}$ grid nodes within one subdomain. As a result, communication between these processors is avoided, leading to improved parallel efficiency.

We evaluate the performance of three parallelization methods at the National Supercomputing Centre (NSCC) Singapore, using up to 32 processors to solve 64 Eikonal equations on a mesh of $101 \times 101 \times 101$ grid nodes. The speed-up and memory usage with respect to the number of processors are displayed in Fig. 6b and Table 1. Source parallelization achieves the highest parallel efficiency but requires substantial memory. In comparison, domain decomposition and hyperplane stepping parallelization almost require no additional memory, though they yield slightly lower parallel efficiency due to inter-process communication and synchronization overhead. In summary, if memory capacity is sufficient, prioritizing processors for source parallelization is optimal, followed by hyperplane stepping parallelization and domain decomposition. It is worth noting that configuring the parallelization scheme in TomoATT is easy and straightforward. Only three settings in the YAML parameter file need to be modified, while the TEXT data file and HDF5 model file remain unchanged.

4. Case studies

TomoATT provides a powerful and user-friendly tool for imaging subsurface velocity heterogeneity and azimuthal anisotropy. In practical applications, dense and evenly distributed seismicity that offers sufficient data constraints is often a key prerequisite for achieving robust and high-resolution imaging results. TomoATT and ambient noise surface wave tomography methods (e.g., Yang et al., 2007; Lin et al., 2013; Liu et al., 2019; Xu et al., 2025) are complementary, as the latter does not rely on regional seismicity but typically demands a dense station network and a relatively uniform distribution of ambient noise. In this section, we first design a synthetic experiment to validate TomoATT. Subsequently, we benchmark the package with two real-data inversion cases: regional tomography in California near Parkfield, and teleseismic tomography in Thailand and adjacent areas. The tomography for these two real cases has been detailed in our previous studies (J. Chen, G. Chen et al., 2023; J. Chen, S. Wu et al., 2023).

4.1. Earthquake location and tomography of a synthetic model

We design such a synthetic experiment, merely aiming at verifying earthquake location and tomography functions in TomoATT. The study region is $[0, 2^\circ] \times [0, 2^\circ] \times [0, 40 \text{ km}]$, with 25 stations uniformly deployed on the surface. The true model is built by embedding staggered velocity and anisotropic perturbations to an isotropic background model, whose velocity linearly increases from 5.0 km/s at 0 km depth to 8.0 km/s at 40 km depth. 867 earthquakes are regularly located at depths of 10, 20, and 30 km (Fig. 7a). Observational traveltime data are generated using the true model and true earthquake locations. Correspondingly, the initial model is the background velocity model without any velocity or anisotropic perturbations. The initial earthquake locations are randomly deviated from the true locations, following uniform distributions with standard deviations of 0.1° in latitude and longitude, 10 km in depth, and 0.5 s in origin time (Fig. 7b).

Table 1

The computational time and memory usage for each parallelization method.

Number of processors	Source parallelization	Domain decomposition	Hyperplane stepping parallelization
1	662.62 s/0.21 GB	–	–
2	348.03 s/0.42 GB	435.17 s/0.23 GB	372.72 s/0.21 GB
4	211.41 s/0.83 GB	307.20 s/0.24 GB	233.11 s/0.21 GB
8	104.62 s/1.66 GB	162.59 s/0.27 GB	125.53 s/0.21 GB
16	52.64 s/3.32 GB	102.99 s/0.32 GB	71.14 s/0.21 GB
32	28.85 s/6.65 GB	63.99 s/0.42 GB	45.73 s/0.21 GB

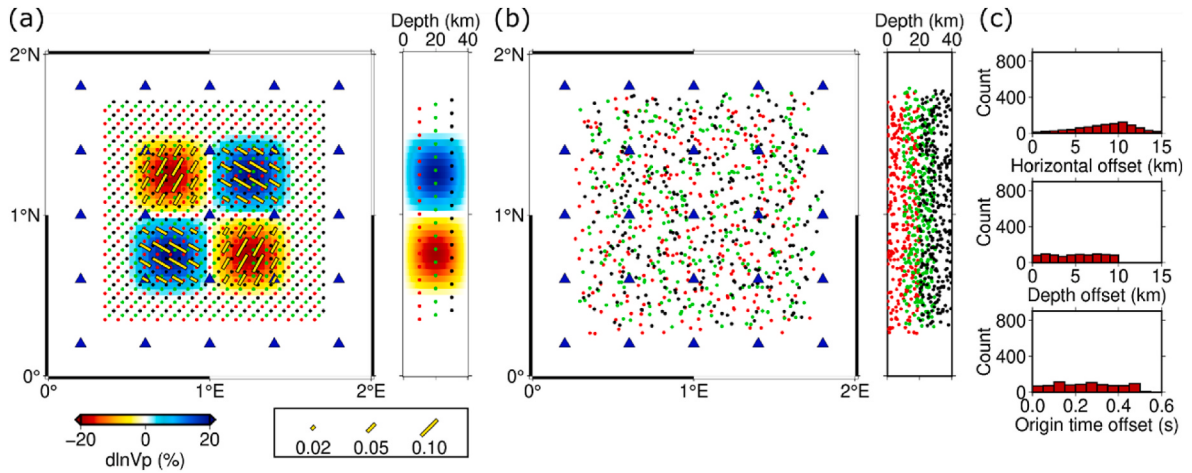


Fig. 7. Model setting of a synthetic experiment. (a) Velocity perturbation of the true model relative to the initial model. Yellow bars represent azimuthal anisotropy, aligned with fast velocity directions. Blue triangles are stations on the surface. True earthquake hypocenters are evenly distributed at depths of 10 km (red dots), 20 km (green dots), and 30 km (black dots). (b) Initial earthquake hypocenters. (c) Histograms of horizontal offsets of hypocenters, vertical offsets of hypocenters, and origin time offsets. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

We perform the following 5 tests:

- Test 1: Locate earthquakes in the true model. The inversion starts with the initial earthquake locations and the true model, performing 200 iterations to update earthquake locations while keeping model parameters fixed.
- Test 2: Locate earthquakes in the initial model. The procedure is the same as in Test 1 but uses the initial model instead.
- Test 3: Update model parameters using the true earthquake locations. The inversion starts with the true earthquake location and the initial model, performing 40 iterations to update velocity and anisotropic parameters while keeping earthquake locations fixed.
- Test 4: Update model parameters using the initial earthquake locations. The procedure is the same as in Test 2 but uses the initial earthquake locations instead.
- Test 5: Simultaneously locate earthquakes and update model parameters. The inversion is conducted in three stages to reduce variable coupling. First, we perform a preliminary earthquake location in the initial model for 50 iterations. Second, we simultaneously update model parameters and earthquake locations for 40 iterations. Finally, we relocate earthquakes in the improved velocity model for 100 iterations to achieve more accurate earthquake locations.

In Test 1, earthquakes are accurately located in the true model (Fig. 8a), verifying the relocation function in TomoATT. In contrast, Test 2 shows that final earthquake locations are slightly offset when using inaccurate velocity and anisotropic model parameters (Fig. 8b), highlighting the importance of an accurate model for precise earthquake location. Similarly, Test 3 indicates that the staggered velocity and anisotropic perturbations are accurately imaged using the true earthquake locations (Fig. 8c), verifying the imaging function in TomoATT. However, in Test 4, inaccurate earthquake locations result in significant distortions and artifacts in the imaged velocity and anisotropy (Fig. 8d), which underlines the critical role of accurate earthquake location in tomography. In practical applications, both initial earthquake locations and model parameters are typically biased. Therefore, it is suggested to update both earthquake locations and model parameters by simultaneous inversion, as illustrated in Test 5 (Fig. 8e). This test result indicates that both model parameters and earthquake locations can be effectively constrained using the simultaneous inversion function in TomoATT.

We also evaluate the influence of data noise in these synthetic experiments. The above five tests are repeated by adding a Gaussian noise

with a standard deviation of 0.1 s to the data to mimic random picking errors. The results are illustrated in Fig. S1. From the comparison, we observe that the event locations are slightly deviated when noise is added, but remain relatively high accuracy (Fig. S1a versus 8a, S1e versus 8e). Meanwhile, the recovered velocities and anisotropies are similar in Figures R1c and R2c, possibly due to the large quantity of data that suppresses the effect of random data noise. Overall, the five synthetic tests yield similar inversion results, indicating that the random data noise has limited influence on the inversion parameters in these cases.

4.2. Regional tomography in central California near Parkfield

We benchmark TomoATT with the regional tomography in central California near Parkfield, covering a study region of 160 km $\times$ 440 km horizontally and 50 km vertically (Fig. 9a). After applying strict data selection criteria (see more details in J. Chen, G. Chen et al. (2023)) to the traveltime data from Northern California Earthquake Data Center (NCEDC, 2014) and Southern California Earthquake Data Center (SCEDC, 2013), we collected 1,218,044 first P-arrival times from 32,721 earthquakes and recorded by 607 stations for inversion. As the earthquake locations has been robustly determined by the data centers, we only inverted for velocity heterogeneity and azimuthal anisotropy, keeping earthquake locations fixed.

The initial model is built as an isotropic layered model used in Tong (2021b). Four constant velocities of 4.17, 6.03, 6.59, and 7.94 km/s are assigned to the sedimentary layer, upper crust, lower crust, and uppermost mantle, respectively. These four layers are separated by two flat interfaces at 4 km and 16 km depths, and an undulating Moho discontinuity extracted from the Crust1.0 model (Laske et al., 2013). Furthermore, a Gaussian filter with a standard deviation of 1 km is applied near the three velocity discontinuities to smooth the initial model.

The study region is discretized into a mesh of 149,787 grid nodes, with an approximate grid spacing of 5 km $\times$ 5 km $\times$ 1 km. This grid spacing is sufficient to ensure the accuracy of our Eikonal solver, as validated in Section 3.1. By applying the reciprocal principle, we can regard the 607 stations as sources and the 32,721 earthquakes as receivers, which allows for numerically solving only 607 anisotropic Eikonal equations and adjoint equations per iteration, rather than 32,721. We perform 80 iterations on the NSCC platform using 64 processors powered by AMD EPYC 7713 64-Core Processor of 2.0 GHz, completing the inversion in 57 min.

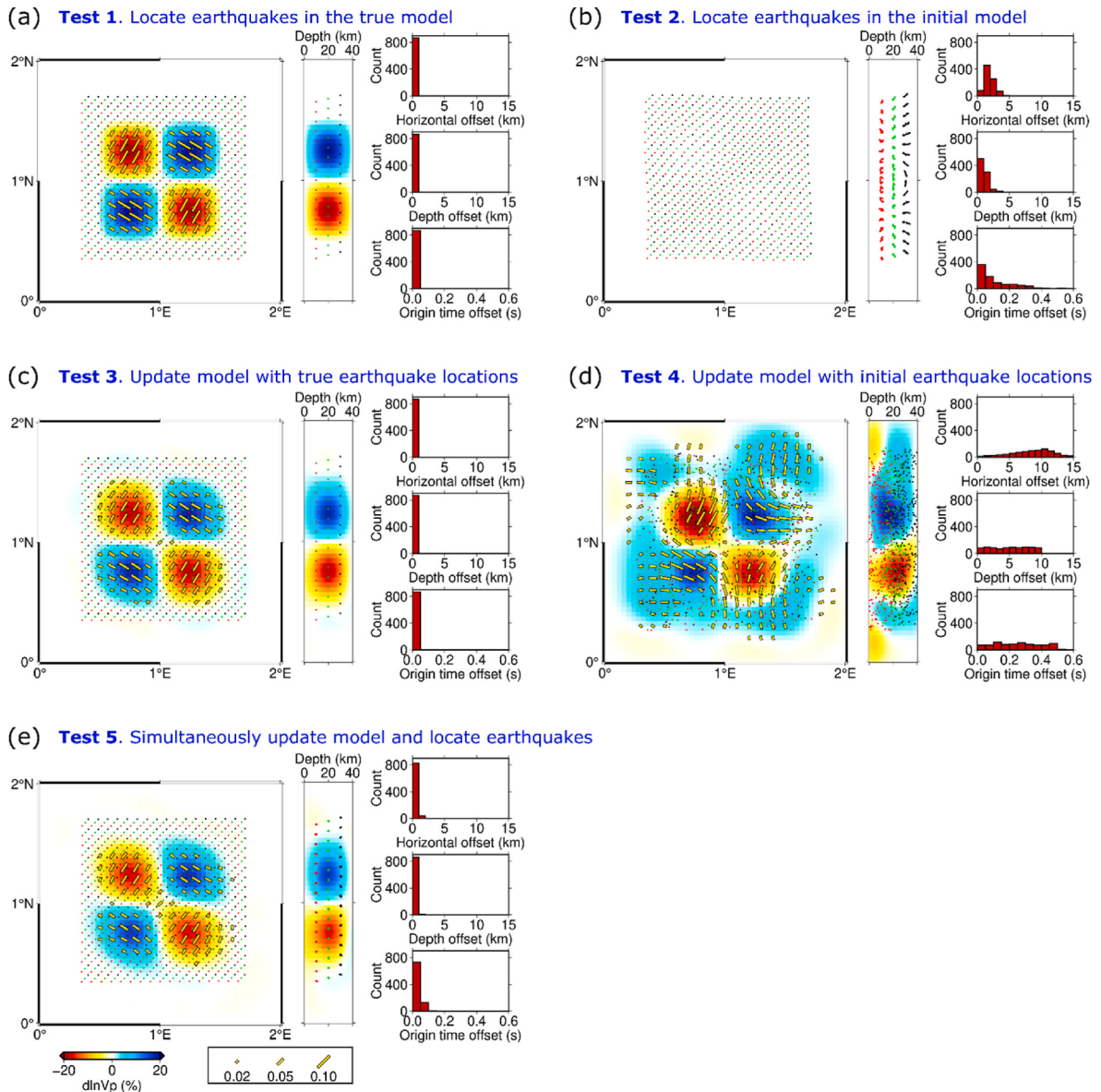


Fig. 8. Test results in the synthetic experiment. Notations follow Fig. 7. (a) Test 1: Earthquake locations determined in the true model, verifying the relocation function. (b) Test 2: Earthquake locations determined in the initial model, showing slight deviation due to inaccurate velocity and anisotropy. (c) Test 3: Velocity and anisotropic parameters updated with the true earthquake locations, verifying the tomography function. (d) Test 4: Velocity and anisotropic parameters updated with the initial earthquake location, resulting in noticeable distortions and artifacts due to source uncertainty. (e) Test 5: Simultaneous inversion of model parameters and earthquake locations.

Checkerboard resolution tests are designed to investigate the data resolving ability, and to assess the influence of random data noise. The size of the checker is approximately $80 \text{ km} \times 70 \text{ km} \times 10 \text{ km}$ in x-y-z directions for velocity and $80 \text{ km} \times 110 \text{ km} \times 10 \text{ km}$ for anisotropy (Fig. S2). Without data noise, velocity perturbation and anisotropy can be well recovered above 17 km depth (Fig. S3). However, the anomalies in the north margin and near the southern coastline are poorly recovered due to less data coverage (Fig. S3e–g). After that, the checkerboard resolution test is repeated with noisy data. The noise follows a Gaussian distribution with a standard deviation of 0.1 s to mimic random picking error. The inversion result (Fig. S4) is similar to that without noise. It suggests that the imaging result is insensitive to random data noise, possibly due to the large volume of data used in the inversion.

Pronounced velocity perturbations are revealed in the real data imaging results, illustrated in horizontal sections in Fig. 9b. At 4 km depth, the San Andreas Fault (SAF) clearly separates the low-velocity zone beneath the Franciscan terrane (FT) in the east from the high-velocity

anomaly beneath the Salinian terrane (ST) in the west. This characteristic structure aligns well with previous imaging results (Eberhart-Phillips and Michael, 1993; Lippoldt et al., 2017; Thurber et al., 2006). The velocity contrast remains at 8 km depth, though the high-velocity anomaly weakens. By 16 km depth, it is replaced by a broad low-velocity anomaly beneath the SAF. This anomaly is suggested by ambient noise tomography (Lippoldt et al., 2017) to expand to the lower crust and uppermost mantle. It is also associated with a high-conductivity zone in the crust (Becken et al., 2011), possibly indicating the presence of crustal fluid (Tong, 2021a). Our result also images a low-velocity anomaly beneath the Santa Maria Basin (SMB), which extends from the surface to 8 km and is connected to the broad low-velocity anomaly at 16 km.

The crustal azimuthal anisotropy beneath the SAF and vicinity shows distinct patterns (Fig. 9c). Specifically, strong anisotropy is observed beneath the creeping and transition segments of the SAF at 8 km depth, with the fast velocity direction parallel to the fault. This fault-parallel

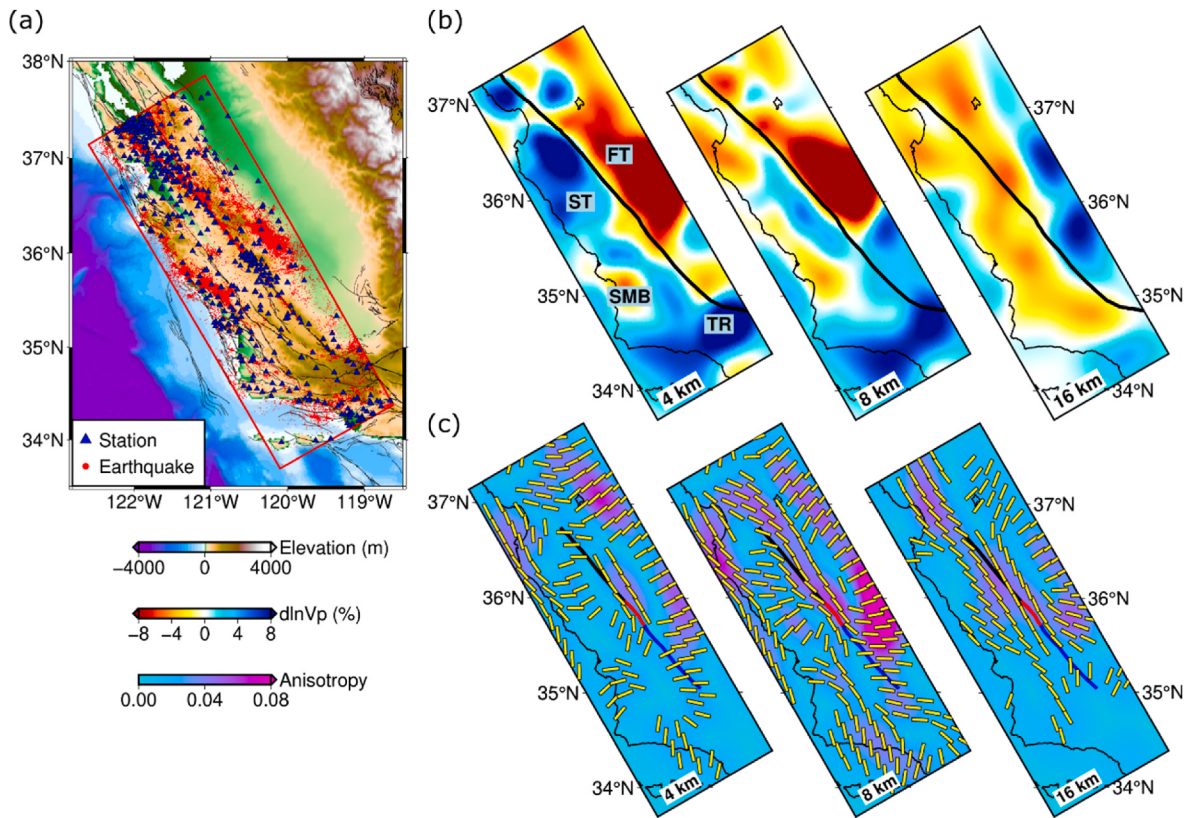


Fig. 9. Tectonic setting and imaging results in central California near Parkfield. (a) Topography map. Red box outlines the study region. Earthquakes and stations are denoted by red dots and blue triangles, respectively. Black lines mark active faults. (b) Horizontal sections of P-wave velocity perturbation relative to the horizontal average. Coastline and the San Andreas Fault (SAF) are indicated by thinner and thicker black lines, respectively. Key tectonic features are labeled: FT (Franciscan Terrane), ST (Salinian Terrane), SMB (Santa Maria Basin), and TR (Transverse Ranges). (c) Horizontal sections of azimuthal anisotropy. Yellow bars align with the fast velocity directions. Black, red, and blue lines denote the creeping, transitional, and locked segments of the SAF. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

fast velocity direction is usually observed near large strike-slip faults (e.g., Zhang and Schwartz, 1994; Peng and Ben-Zion, 2004; Rasendra et al., 2014). Due to the fault parallel crack alignment, P-wave tends to propagate faster along the fault strike than in the orthogonal direction, leading to fault-parallel fast velocity directions (Zhang and Schwartz, 1994). This suggests that anisotropy near the SAF is dominated by the fault structure (Boness and Zoback, 2006). This observation is consistent with the results of teleseismic receiver function analysis (Audet, 2015; Ozacar and Zandt, 2009). In contrast, farther from the SAF, the fast velocity direction shifts, forming a large angle or becoming nearly perpendicular to the fault. This pattern coincides with the maximum horizontal compression (S_{Hmax}) direction observed from borehole data (Townend and Zoback, 2004), indicating that the anisotropy is stress-induced.

In summary, variations in seismic velocity and azimuthal anisotropy in central California near Parkfield are successfully imaged by inverting first P-arrival traveltimes. The consistency of the results with local tectonics and previous studies validates the effectiveness of TomoATT in the regional tomography.

4.3. Teleseismic tomography in Thailand and adjacent areas

We benchmark TomoATT with the teleseismic tomography in Thailand and adjacent areas. A total of 51 stations, consisting of 40 temporary stations from the Thai Seismic Array (Tanaka et al., 2019) and 11 additional permanent stations, are deployed within the study region spanning from 97°E to 106°E and from 12°N to 21°N (Fig. 10a). We download waveforms from 190 earthquakes with magnitudes greater than 5.0 and epicentral distances between 30° and 90° (Fig. 10b)

and use the multi-channel cross correlation technique (VanDecar and Crosson, 1990) to extract common-source differential arrival times. After applying strict data selection criteria (J. Chen, G. Chen et al., 2023), we collect 15,205 common-source differential arrival times to invert for upper mantle seismic velocity. The azimuthal anisotropy is fixed as 0 during the inversion, as the nearly vertical ray paths of teleseismic events provide limited constraints on azimuthal anisotropy.

The initial model is built by adopting the Crust1.0 model (Laske et al., 2013) for the crustal P-wave velocity and the global reference AK135 model (Kennett et al., 1995) for the mantle P-wave velocity. The computational region of $12^\circ \times 12^\circ \times 600$ km is discretized into a mesh of 1,771,561 grid nodes, with an approximate grid spacing of $10 \text{ km} \times 10 \text{ km} \times 5 \text{ km}$. This grid spacing is sufficient to ensure the accuracy of our Eikonal solver, as validated in Section 3.2. We perform 80 iterations on the NSCC platform using 64 processors, completing the inversion in 124 min.

Checkerboard resolution tests are performed to investigate data resolving ability and to assess the influence of data noise (Fig. S5). The size of the checkers is approximately $160 \text{ km} \times 160 \text{ km} \times 150 \text{ km}$. Without data noise, the velocity perturbations are well recovered within the station coverage (Fig. S6). However, oblique smearing is observed outside the network, possibly due to the propagation direction of teleseismic waves. Meanwhile, the amplitude of the recovered velocity perturbation is underestimated. This may be attributed to the fact that teleseismic waves travel from depth to the surface, traversing alternating high- and low-velocity anomalies. The opposing effect of the staggered velocity anomalies on traveltimes tends to cancel out, thereby mitigating the sensitivity of traveltime data to the anomaly amplitude, ultimately leading to underestimation. We also repeat the checkerboard

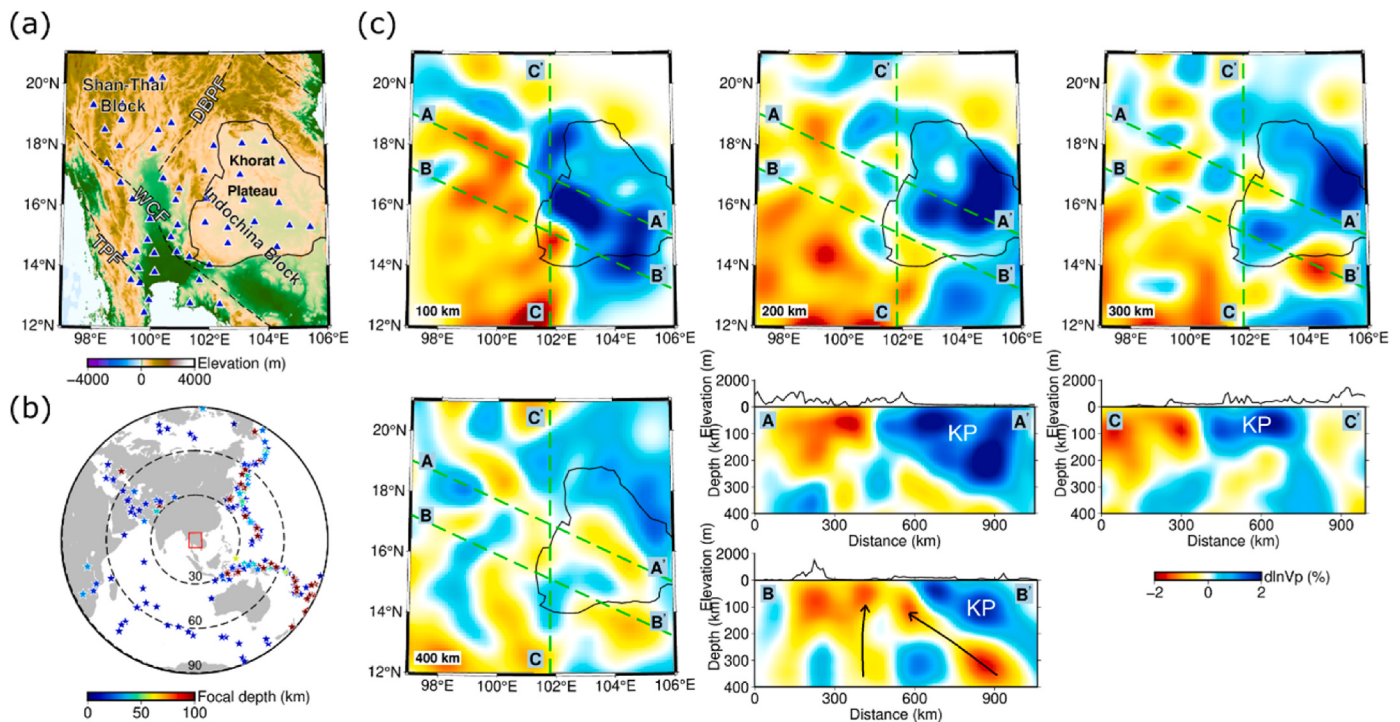


Fig. 10. Tectonic setting and imaging results in Thailand and adjacent regions. (a) Topography map. Black dashed lines denote major faults: DBPF (Dien Bien Phu Fault), WCF (Wang-Chao Fault), and TPF (Three Pagodas Fault). The solid black line indicates the Khorat Plateau. Stations are denoted by blue triangles. (b) Distribution of teleseismic earthquakes used for the inversion. Dashed circles represent epicenter distances of 30° and 60°. (c) Horizontal and vertical sections of velocity perturbation relative to the horizontal average. Vertical section locations are plotted as green dashed lines. KP represents the Khorat Plateau. In BB' profile, two black arrows indicate possible pathways of mantle upwelling, shown as low-velocity perturbations extending from the surface down to the upper mantle. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

tests with noisy data. The data noise follows a Gaussian distribution with a standard deviation of 0.1 s. The comparison between the results with and without noise (Figs. S6 and S7) shows a high level of similarity, indicating that the inversion results are relatively insensitive to random data noise.

Fig. 10c illustrates the horizontal sections of velocity perturbations related to the horizontal average in real data imaging. The most notable feature is the high-velocity perturbation beneath the Khorat Plateau, contrasting with a low-velocity perturbation to the west. This structure shows consistency with previous tomographic images (Li et al., 2006; Yang et al., 2015), and corresponds to a cold and thick lithosphere beneath the Khorat Plateau, as inferred by the thermal analysis based on S-velocity (C. Yu et al., 2017). Furthermore, our result reveals a low-velocity perturbation beneath the southwestern corner of the Khorat Plateau, different from the high-velocity perturbation in the central area. This distinct feature may suggest that the southwestern margin of the Khorat Plateau is partially modified. The vertical section BB' crossing this anomaly reveals two low-velocity channels extending from the surface down to the upper mantle. These low-velocity perturbations may indicate the pathways for mantle material upwelling, potentially driven by the mantle convection associated with surrounding subduction systems of the Indo-Australian, Pacific, and Philippine Sea Plates (Lin et al., 2019). Notably, the western channel appears connected to a possible slab window of the Indian Plate (Pesicek et al., 2008; Y. Yu et al., 2017). The mantle upwelling through the slab window may also be a contributing factor to the mantle upwelling (Arboit et al., 2016).

5. Conclusions

TomoATT provides an effective solution for regional and teleseismic traveltime tomography, based on the Eikonal equation-based adjoint-

state traveltime tomography methods. Synthetic experiments showcase its advantages in accurate forward modeling, handling multipathing phenomenon, sensitivity kernel regularization for reliable inversion, and multi-level parallelization for high efficiency. Users can initiate an inversion easily with only three simple files: an HDF5 model file, a YAML parameter file, and a TEXT data file. Two real-data applications—regional tomography in central California near Parkfield and teleseismic tomography in Thailand and adjacent regions—were completed within 2 h using 64 processors. The imaging results reveal pronounced velocity perturbations and azimuthal anisotropies that align well with previous studies and local tectonics. These benchmarks demonstrate the effectiveness and efficiency of TomoATT for regional and teleseismic tomography.

CRedit authorship contribution statement

Jing Chen: Writing – review & editing, Writing – original draft, Software, Methodology. **Masaru Nagaso:** Writing – review & editing, Software. **Mijian Xu:** Writing – review & editing, Software. **Ping Tong:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition.

Code availability section

Name of the code/library: TomoATT, PyTomoATT.
License: GNU General Public License v3.0.
Contact: tongping@ntu.edu.sg, jing.chen@ntu.edu.sg.
Hardware requirements: PC or HPC.
Program language: C++, Python.
Software required: C++ compiler, CMake, OpenMPI/MPICH, HDF5, YAML, Python 3.9 or later.
Program size: 9.53 MB.

The source codes and document are available at <https://tomoatt.com/>

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Minister of Education, Singapore, under its MOE AcRF Tier-2 Grant (MOE-T2EP20122-0008). The

computational work for this article was partially performed on resources of the National Supercomputing Centre, Singapore (<https://www.nsc.sg>). The authors acknowledge the seismic data for benchmark, available from the Northern California Earthquake Data Center (NCEDC 2014), the Southern California Earthquake Data Center (SCEDC 2013), TSAR (Tanaka et al., 2019) (<http://ohpdm.eri.u-tokyo.ac.jp/breq-fast-tsar/index.html>), and the Incorporated Research Institutions for Seismology (IRIS) Data Management Center (<http://ds.iris.edu/mda/TM/PBKT/?starttime=2008-01-01&endtime=2599-12-31>). The authors thank the editors, Chuanming Liu, and an anonymous reviewer for their constructive suggestions and remarkable comments, which are helpful to improve this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cageo.2025.105995>.

Appendix

A. Chosen of weight coefficients

In general, determining the optimal weight coefficients for different types of data is somewhat subjective and should be evaluated on a case-by-case basis. Nevertheless, we provide some general principles that may guide the selection of reasonable weight coefficients.

A straightforward approach is to balance the contributions from different types of data by setting the weight coefficients as follows

$$\alpha = \frac{1}{\sum_{n=1}^{N_s} \sum_{m=1}^{N_r} w_{n|m}}, \beta = \frac{1}{\sum_{n=1}^{N_s} \sum_{m=1}^{N_r} \sum_{i=1}^{N_r} w_{n|m,i}}, \gamma = \frac{1}{\sum_{n=1}^{N_s} \sum_{j=1}^{N_s} \sum_{m=1}^{N_r} w_{n|jm}}. \quad (\text{A1})$$

This strategy ensures that the contributions from absolute traveltimes, common-source differential arrival times, and common-receiver differential arrival times are comparable in the overall misfit function. It avoids the predominance of any single data type, thereby enhancing the stability of the inversion.

Furthermore, if the standard deviations of the data misfits for the three types of data can be estimated as σ_t , σ_{cs} , and σ_{cr} , they can be further incorporated into the weight coefficient. As suggested by Tarantola (2005) (see Equation 1.60 on Page 22), the weight coefficients can be set to be

$$\alpha = \frac{1}{\sum_{n=1}^{N_s} \sum_{m=1}^{N_r} w_{n|m} \sigma_t^2}, \beta = \frac{1}{\sum_{n=1}^{N_s} \sum_{m=1}^{N_r} \sum_{i=1}^{N_r} w_{n|m,i} \sigma_{cs}^2}, \gamma = \frac{1}{\sum_{n=1}^{N_s} \sum_{j=1}^{N_s} \sum_{m=1}^{N_r} w_{n|jm} \sigma_{cr}^2}. \quad (\text{A2})$$

This formulation helps eliminate the impact of differences in misfit magnitude among the various data types, thus leading to improved stability. If a consistent standard deviation is assumed for the three types of data, the weight coefficients in Equation (A2) reduce to those in Equation (A1).

In addition, the weight coefficients can be subjectively adjusted to emphasize specific types of data. For example, common-source differential arrival time is less sensitive to source uncertainty and better constrains receiver-side structures (Yuan et al., 2016; Guo and Zhang, 2017). Therefore, increasing the weight coefficient β associated with common-source differential arrival time may improve the reliability of the inversion, especially when using earthquake catalogs with relatively large source uncertainties. In contrast, common-receiver differential arrival times are proven less sensitive to near-receiver velocity heterogeneity (Waldhauser and Ellsworth, 2000; Zhang and Thurber, 2003). Accordingly, increasing the weight coefficient γ may enhance the robustness of earthquake location results. In summary, the weight coefficients can be subjectively adjusted based on prior knowledge and the specific goal of the study.

B. Derivation of adjoint sources

The core feature of the adjoint-state method is to solve the adjoint equation (8), in which the adjoint source takes the general form of

$$R_{n,m} = \left(\frac{\partial}{\partial T_n(\mathbf{x}_{r,m})} \chi \right) = \frac{\partial \chi_t}{\partial T_n(\mathbf{x}_{r,m})} + \frac{\partial \chi_{cs}}{\partial T_n(\mathbf{x}_{r,m})} + \frac{\partial \chi_{cr}}{\partial T_n(\mathbf{x}_{r,m})}. \quad (\text{B1})$$

The three derivatives are detailed in Tong et al. (2023), which are also specified as follows

$$\frac{\partial \chi_t}{\partial T_n(\mathbf{x}_{r,m})} = w_{n|m} (T_n(\mathbf{x}_{r,m}) - T_n^{\text{obs}}(\mathbf{x}_{r,m})), \quad (\text{B2})$$

$$\begin{aligned} \frac{\partial \chi_{cs}}{\partial T_n(\mathbf{x}_{r,m})} &= \sum_{i=1}^{N_r} w_{n|mi} ((T_n(\mathbf{x}_{r,m}) - T_n(\mathbf{x}_{r,i})) - (T_n^{\text{obs}}(\mathbf{x}_{r,m}) - T_n^{\text{obs}}(\mathbf{x}_{r,i}))) - \sum_{j=1}^{N_r} w_{n|jm} ((T_n(\mathbf{x}_{r,j}) - T_n(\mathbf{x}_{r,m})) - (T_n^{\text{obs}}(\mathbf{x}_{r,j}) - T_n^{\text{obs}}(\mathbf{x}_{r,m}))) \\ &= 2 \sum_{i=1}^{N_r} w_{n|mi} ((T_n(\mathbf{x}_{r,m}) - T_n(\mathbf{x}_{r,i})) - (T_n^{\text{obs}}(\mathbf{x}_{r,m}) - T_n^{\text{obs}}(\mathbf{x}_{r,i}))), \end{aligned} \quad (\text{B3})$$

$$\begin{aligned} \frac{\partial \chi_{cr}}{\partial T_n(\mathbf{x}_{r,m})} &= \sum_{j=1}^{N_s} w_{nj|m} \left((T_n(\mathbf{x}_{r,m}) - T_j(\mathbf{x}_{r,m})) - (T_n^{obs}(\mathbf{x}_{r,m}) - T_j^{obs}(\mathbf{x}_{r,m})) \right) - \sum_{i=1}^{N_s} w_{in|m} \left((T_i(\mathbf{x}_{r,m}) - T_n(\mathbf{x}_{r,m})) - (T_i^{obs}(\mathbf{x}_{r,m}) - T_n^{obs}(\mathbf{x}_{r,m})) \right) \\ &= 2 \sum_{j=1}^{N_s} w_{nj|m} \left((T_n(\mathbf{x}_{r,m}) - T_j(\mathbf{x}_{r,m})) - (T_n^{obs}(\mathbf{x}_{r,m}) - T_j^{obs}(\mathbf{x}_{r,m})) \right), \end{aligned} \quad (B4)$$

C. The sensitivity kernel with respect to earthquake location

By using the reciprocal principle, the synthetic traveltime $T_n(\mathbf{x}_{r,m})$ from the n -th source $\mathbf{x}_{s,n}$ and recorded by the m -th receiver $\mathbf{x}_{r,m}$ equals the synthetic traveltime $\Gamma_m(\mathbf{x}_{s,n})$ from the receiver $\mathbf{x}_{r,m}$ to the source $\mathbf{x}_{s,n}$ (refer to [Tong et al. \(2023\)](#) for detailed discussion). Here the function $\Gamma_m(\mathbf{x})$ describes the wavefront traveltime from the m -th receiver $\mathbf{x}_{r,m}$ to any position $\mathbf{x}$. Consequently, the objective functions (2)–(4) can be reformulated as

$$\chi_t = \sum_{n=1}^{N_s} \sum_{m=1}^{N_r} \frac{w_{n|m}}{2} (\Gamma_m(\mathbf{x}_{s,n}) + \tau_n - T_n^{obs}(\mathbf{x}_{r,m}))^2, \quad (C1)$$

$$\chi_{cs} = \sum_{n=1}^{N_s} \sum_{m=1}^{N_r} \sum_{i=1}^{N_r} \frac{w_{n|mi}}{2} ((\Gamma_m(\mathbf{x}_{s,n}) - \Gamma_i(\mathbf{x}_{s,n})) - (T_n^{obs}(\mathbf{x}_{r,m}) - T_n^{obs}(\mathbf{x}_{r,i})))^2, \quad (C2)$$

$$\chi_{cr} = \sum_{n=1}^{N_s} \sum_{j=1}^{N_s} \sum_{m=1}^{N_r} \frac{w_{nj|m}}{2} ((\Gamma_m(\mathbf{x}_{s,n}) - \Gamma_m(\mathbf{x}_{s,j})) + \tau_n - \tau_j - (T_n^{obs}(\mathbf{x}_{r,m}) - T_j^{obs}(\mathbf{x}_{r,m})))^2. \quad (C3)$$

The sensitivity kernels of the objective function with respect to $\mathbf{x}_{s,n}$ are derived as follows

$$\nabla_{\mathbf{x}_{s,n}} \chi = \nabla_{\mathbf{x}_{s,n}} \chi_t + \nabla_{\mathbf{x}_{s,n}} \chi_{cs} + \nabla_{\mathbf{x}_{s,n}} \chi_{cr}, \quad (C4)$$

in which

$$\nabla_{\mathbf{x}_{s,n}} \chi_t = \sum_{m=1}^{N_r} w_{n|m} (\Gamma_m(\mathbf{x}_{s,n}) + \tau_n - T_n^{obs}(\mathbf{x}_{r,m})) \nabla \Gamma_m(\mathbf{x}_{s,n}), \quad (C5)$$

$$\begin{aligned} \nabla_{\mathbf{x}_{s,n}} \chi_{cs} &= \sum_{m=1}^{N_r} \sum_{i=1}^{N_r} w_{n|mi} ((\Gamma_m(\mathbf{x}_{s,n}) - \Gamma_i(\mathbf{x}_{s,n})) - (T_n^{obs}(\mathbf{x}_{r,m}) - T_n^{obs}(\mathbf{x}_{r,i}))) (\nabla \Gamma_m(\mathbf{x}_{s,n}) \\ &\quad - \nabla \Gamma_i(\mathbf{x}_{s,n})) = 2 \sum_{m=1}^{N_r} \sum_{i=1}^{N_r} w_{n|mi} ((\Gamma_m(\mathbf{x}_{s,n}) - \Gamma_i(\mathbf{x}_{s,n})) - (T_n^{obs}(\mathbf{x}_{r,m}) - T_n^{obs}(\mathbf{x}_{r,i}))) \nabla \Gamma_m(\mathbf{x}_{s,n}), \end{aligned} \quad (C6)$$

$$\begin{aligned} \nabla_{\mathbf{x}_{s,n}} \chi_{cr} &= \sum_{j=1}^{N_s} \sum_{m=1}^{N_r} w_{nj|m} ((\Gamma_m(\mathbf{x}_{s,n}) - \Gamma_m(\mathbf{x}_{s,j})) + \tau_n - \tau_j - (T_n^{obs}(\mathbf{x}_{r,m}) - T_j^{obs}(\mathbf{x}_{r,m}))) \nabla \Gamma_m(\mathbf{x}_{s,n}) - \sum_{i=1}^{N_s} \sum_{m=1}^{N_r} w_{in|m} ((\Gamma_m(\mathbf{x}_{s,i}) - \Gamma_m(\mathbf{x}_{s,n})) + \tau_i - \tau_n \\ &\quad - (T_i^{obs}(\mathbf{x}_{r,m}) - T_n^{obs}(\mathbf{x}_{r,m}))) \nabla \Gamma_m(\mathbf{x}_{s,n}) = 2 \sum_{j=1}^{N_s} \sum_{m=1}^{N_r} w_{nj|m} ((\Gamma_m(\mathbf{x}_{s,n}) - \Gamma_m(\mathbf{x}_{s,j})) + \tau_n - \tau_j - (T_n^{obs}(\mathbf{x}_{r,m}) - T_j^{obs}(\mathbf{x}_{r,m}))) \nabla \Gamma_m(\mathbf{x}_{s,n}). \end{aligned} \quad (C7)$$

Similarly, the sensitivity kernels of the objective function with respect to τ_n are derived as follows

$$\partial_{\tau_n} \chi = \partial_{\tau_n} \chi_t + \partial_{\tau_n} \chi_{cs} + \partial_{\tau_n} \chi_{cr}, \quad (C8)$$

in which

$$\partial_{\tau_n} \chi_t = \sum_{m=1}^{N_r} w_{n|m} (\Gamma_m(\mathbf{x}_{s,n}) + \tau_n - T_n^{obs}(\mathbf{x}_{r,m})), \quad (C9)$$

$$\partial_{\tau_n} \chi_{cs} = 0 \quad (C10)$$

$$\begin{aligned} \partial_{\tau_n} \chi_{cr} &= \sum_{j=1}^{N_s} \sum_{m=1}^{N_r} w_{nj|m} ((\Gamma_m(\mathbf{x}_{s,n}) - \Gamma_m(\mathbf{x}_{s,j})) + \tau_n - \tau_j - (T_n^{obs}(\mathbf{x}_{r,m}) - T_j^{obs}(\mathbf{x}_{r,m}))) - \sum_{i=1}^{N_s} \sum_{m=1}^{N_r} w_{in|m} ((\Gamma_m(\mathbf{x}_{s,i}) - \Gamma_m(\mathbf{x}_{s,n})) + \tau_i - \tau_n - (T_i^{obs}(\mathbf{x}_{r,m}) \\ &\quad - T_n^{obs}(\mathbf{x}_{r,m}))) = 2 \sum_{j=1}^{N_s} \sum_{m=1}^{N_r} w_{nj|m} ((\Gamma_m(\mathbf{x}_{s,n}) - \Gamma_m(\mathbf{x}_{s,j})) + \tau_n - \tau_j - (T_n^{obs}(\mathbf{x}_{r,m}) - T_j^{obs}(\mathbf{x}_{r,m}))). \end{aligned} \quad (C11)$$

Data availability

The code and data are shared via the link in Code Availability Section.

References

- Alkhalifah, T., 2015. Scattering-angle based filtering of the waveform inversion gradients. *Geophys. J. Int.* 200, 363–373.
- Arboit, F., Collins, A.S., Morley, C.K., Jourdan, F., King, R., Foden, J., Amrouch, K., 2016. Geochronological and geochemical studies of mafic and intermediate dykes from the Khao Khwang Fold–Thrust Belt: implications for petrogenesis and tectonic evolution. *Gondwana Res.* 36, 124–141.
- Audet, P., 2015. Layered crustal anisotropy around the San Andreas Fault near Parkfield, California. *J. Geophys. Res. Solid Earth* 120, 3527–3543.
- Becken, M., Ritter, O., Bedrosian, P.A., Weckmann, U., 2011. Correlation between deep fluids, tremor and creep along the central San Andreas fault. *Nature* 480, 87–90.
- Boness, N.L., Zoback, M.D., 2006. Mapping stress and structurally controlled crustal shear velocity anisotropy in California. *Geology* 34, 825–828.
- Chen, G., Chen, J., Tape, C., Wu, H., Tong, P., 2023. Double-difference adjoint tomography of the crust and uppermost mantle beneath Alaska. *J. Geophys. Res. Solid Earth* 128, e2022JB025168.
- Chen, J., Chen, G., Nagaso, M., Tong, P., 2023. Adjoint-state traveltime tomography for azimuthally anisotropic media in spherical coordinates. *Geophys. J. Int.* 234, 712–736.
- Chen, J., Wu, S., Xu, M., Nagaso, M., Yao, J., Wang, K., Li, T., Bai, Y., Tong, P., 2023. Adjoint-state teleseismic traveltime tomography: method and application to Thailand in Indochina Peninsula. *J. Geophys. Res. Solid Earth* 128, e2023JB027348.
- Crotwell, H.P., Owens, T.J., Ritsema, J., 1999. The TauP Toolkit: flexible seismic travel-time and ray-path utilities. *Seismol. Res. Lett.* 70, 154–160.
- Detrixhe, M., Gibou, F., Min, C., 2013. A parallel fast sweeping method for the Eikonal equation. *J. Comput. Phys.* 237, 46–55.
- Detrixhe, M., Gibou, F., 2016. Hybrid massively parallel fast sweeping method for static Hamilton–Jacobi equations. *J. Comput. Phys.* 322, 199–223.
- de Vos, D., Paulssen, H., Fichtner, A., 2013. Finite-frequency sensitivity kernels for two-station surface wave measurements. *Geophys. J. Int.* 194, 1042–1049.
- Eberhart-Phillips, D., Michael, A.J., 1993. Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield region, central California. *J. Geophys. Res. Solid Earth* 98, 15737–15758.
- Eberhart-Phillips, D., Reyners, M., 2012. Imaging the Hikurangi plate interface region, with improved local-earthquake tomography. *Geophys. J. Int.* 190, 1221–1242.
- Guo, H., Zhang, H., 2017. Development of double-pair double difference earthquake location algorithm for improving earthquake locations. *Geophys. J. Int.* 208, 333–348.
- Kennett, B.L.N., Engdahl, E.R., Buland, R., 1995. Constraints on seismic velocities in the Earth from traveltimes. *Geophys. J. Int.* 122, 108–124.
- Kissling, E., Ellsworth, W.L., Eberhart-Phillips, D., Kradolfer, U., 1994. Initial reference models in local earthquake tomography. *J. Geophys. Res. Solid Earth* 99, 19635–19646.
- Laske, G., Masters, G., Ma, Z., Pasyanos, M., 2013. Update on CRUST1.0—a 1-degree global model of Earth's crust. EGU General Assembly Geophysical Research Abstracts 15, EGU2013-EGU2658.
- Leung, S., Qian, J., 2006. An adjoint state method for three-dimensional transmission traveltime tomography using first-arrivals. *Commun. Math. Sci.* 4, 249–266.
- Li, C., van der Hilst, R.D., Toksöz, M.N., 2006. Constraining P-wave velocity variations in the upper mantle beneath Southeast Asia. *Phys. Earth Planet. Inter.* 154, 180–195.
- Lin, F.C., Li, D., Clayton, R.W., Hollis, D., 2013. High-resolution 3D shallow crustal structure in Long Beach, California: application of ambient noise tomography on a dense seismic array. *Geophysics* 78, Q45–Q56.
- Lin, J., Xu, Y., Sun, Z., Zhou, Z., 2019. Mantle upwelling beneath the South China Sea and links to surrounding subduction systems. *Natl. Sci. Rev.* 6, 877–881.
- Lippoldt, R., Porritt, R.W., Sammis, C.G., 2017. Relating seismicity to the velocity structure of the San Andreas fault near Parkfield, CA. *Geophys. J. Int.* 209, 1740–1745.
- Liu, C., Yao, H., Yang, H.Y., Shen, W., Fang, H., Hu, S., Qiao, L., 2019. Direct inversion for three-dimensional shear wave speed azimuthal anisotropy based on surface wave ray tracing: methodology and application to Yunnan, southwest China. *J. Geophys. Res. Solid Earth* 124, 11394–11413.
- Lomax, A., 2005. A reanalysis of the hypocentral location and related observations for the great 1906 California earthquake. *Bull. Seismol. Soc. Am.* 95, 861–877.
- Luo, S., Qian, J., 2012. Fast sweeping methods for factored anisotropic eikonal equations: multiplicative and additive factors. *J. Sci. Comput.* 52, 360–382.
- Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics* 52, 1211–1228.
- NCEDC, 2014. Northern California earthquake data center. 10.7932/NCEDC.
- Ozacar, A.A., Zandt, G., 2009. Crustal structure and seismic anisotropy near the San Andreas Fault at Parkfield, California. *Geophys. J. Int.* 178, 1098–1104.
- Peng, Z., Ben-Zion, Y., 2004. Systematic analysis of crustal anisotropy along the Karadere-Düzce branch of the North Anatolian fault. *Geophys. J. Int.* 159, 253–27.
- Pesicek, J., Thurber, C.H., Widiyantoro, S., Engdahl, E., DeShon, H., 2008. Complex slab subduction beneath northern Sumatra. *Geophys. Res. Lett.* 35, L20303.
- Pesicek, J.D., Thurber, C.H., Zhang, H., DeShon, H.R., Engdahl, E.R., Widiyantoro, S., 2010. Teleseismic double-difference relocation of earthquakes along the Sumatra–Andaman subduction zone using a 3-D model. *J. Geophys. Res. Solid Earth* 115, B10303.
- Rasendra, N., Bonnin, M., Mazzotti, S., Tiberi, C., 2014. Crustal and upper mantle anisotropy related to fossilized transpression fabric along the Denali fault, northern Canadian Cordillera. *Bull. Seismol. Soc. Am.* 104, 1964–1975.
- Rawlinson, N., Hauser, J., Sambridge, M., 2008. Seismic ray tracing and waveform tracking in laterally heterogeneous media. *Adv. Geophys.* 49, 203–273.
- Ruan, Y., Lei, W., Modrak, R., Orscurran, R., Bozdog, E., Tromp, J., 2019. Balancing unevenly distributed data in seismic tomography: a global adjoint tomography example. *Geophys. J. Int.* 219, 1225–1236.
- SCEDC, 2013. Southern California earthquake center. Caltech. 10.7909/C7903WD7903xH7901.
- Tanaka, S., Siripunvaraporn, W., Boonchaisuk, S., Noisagool, S., Kim, T., Kawai, K., et al., 2019. Thai Seismic Array (TSAR) Project. 東京大学地震研究所彙報= Bulletin of the Earthquake Research Institute, University of Tokyo 94, pp. 1–11.
- Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. Society for industrial and applied mathematics.
- Terhemba, B.S., Yao, H., Luo, S., Gao, L., Zhang, H., Li, J., 2022. P-wave velocity structure in the crust and the uppermost mantle of Chao Lake region of the Tan-Lu fault inferred from teleseismic arrival time tomography. *Earthq. Sci.* 35, 427–447.
- Thurber, C.H., 1983. Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California. *J. Geophys. Res. Solid Earth* 88, 8226–8236.
- Thurber, C.H., Zhang, H., Waldhauser, F., Hardebeck, J., Michael, A., Eberhart-Phillips, D., 2006. Three-dimensional compressional wavespeed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region. *Bull. Seismol. Soc. Am.* 96, S38–S49.
- Tong, P., 2021a. Adjoint-state traveltime tomography for azimuthally anisotropic media and insight into the crustal structure of central California near Parkfield. *J. Geophys. Res. Solid Earth* 126, e2021JB022365.
- Tong, P., 2021b. Adjoint-state traveltime tomography: eikonal equation-based methods and application to the Anza area in Southern California. *J. Geophys. Res. Solid Earth* 126, e2021JB021818.
- Tong, P., Li, T., Chen, J., Nagaso, M., 2023. Adjoint-state differential arrival time tomography. *Geophys. J. Int.*, gga416.
- Tong, P., Yang, D., Huang, X., 2019. Multiple-grid model parametrization for seismic tomography with application to the San Jacinto fault zone. *Geophys. J. Int.* 218, 200–223.
- Townend, J., Zoback, M.D., 2004. Regional tectonic stress near the San Andreas fault in central and southern California. *Geophys. Res. Lett.* 31, L15S11.
- Tsai, Y.H.R., Cheng, L.T., Osher, S., Zhao, H.K., 2003. Fast sweeping algorithms for a class of Hamilton–Jacobi equations. *SIAM J. Numer. Anal.* 41, 673–694.
- VanDecar, J., Crosson, R., 1990. Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares. *Bull. Seismol. Soc. Am.* 80, 150–169.
- Vidale, J., 1988. Finite-difference calculation of travel times. *Bull. Seismol. Soc. Am.* 78, 2062–2076.
- Waldhauser, F., Ellsworth, W.L., 2000. A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California. *Bull. Seismol. Soc. Am.* 90, 1353–1368.
- Wu, S., Jiang, C., Schulte-Pelkum, V., Tong, P., 2022. Complex patterns of past and ongoing crustal deformations in Southern California revealed by seismic azimuthal anisotropy. *Geophys. Res. Lett.* 49, e2022GL100233.
- Xu, M., Hao, S., Chen, J., Zhang, B., Tong, P., 2025. SurfATT: High-performance package for adjoint-state surface-wave travel-time tomography. *Seismol. Res. Lett.* <https://doi.org/10.1785/0220240206>.
- Yang, T., Liu, F., Harmon, N., Le, K.P., Gu, S., Xue, M., 2015. Lithospheric structure beneath Indochina block from Rayleigh wave phase velocity tomography. *Geophys. J. Int.* 200, 1582–1595.
- Yang, Y., Ritzwoller, M.H., Levshin, A.L., Shapiro, N.M., 2007. Ambient noise Rayleigh wave tomography across Europe. *Geophys. J. Int.* 168, 259–274.
- Yu, C., Shi, X., Yang, X., Zhao, J., Chen, M., Tang, Q., 2017. Deep thermal structure of Southeast Asia constrained by S-velocity data. *Mar. Geophys. Res.* 38, 341–355.
- Yu, Y., Gao, S.S., Liu, K.H., Yang, T., Xue, M., Le, K.P., 2017. Mantle transition zone discontinuities beneath the Indochina Peninsula: implications for slab subduction and mantle upwelling. *Geophys. Res. Lett.* 44, 7159–7167.
- Yuan, Y.O., Simons, F.J., Tromp, J., 2016. Double-difference adjoint seismic tomography. *Geophys. J. Int.* 206, 1599–1618.
- Zhang, H., Nadeau, R.M., Toksoz, M.N., 2010. Locating nonvolcanic tremors beneath the San Andreas fault using a station-pair double-difference location method. *Geophys. Res. Lett.* 37 (13).
- Zhang, H., Thurber, C.H., 2003. Double-difference tomography: the method and its application to the Hayward fault, California. *Bull. Seismol. Soc. Am.* 93 (5), 1875–1889.
- Zhang, Z., Schwartz, S.Y., 1994. Seismic anisotropy in the shallow crust of the Loma Prieta segment of the San Andreas fault system. *J. Geophys. Res. Solid Earth* 99, 9651–9661.
- Zhao, H., 2005. A fast sweeping method for eikonal equations. *Math. Comput.* 74, 603–627.
- Zhao, H., 2007. Parallel implementations of the fast sweeping method. *J. Comput. Math.* 25, 421–429.