Petroleum Science 23 (2026) 3180–3212

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

Multiparameter Bayesian full-waveform inversion with uncertainty quantification based on regularized inverse scattering theory for elastic transversely isotropic media

Wen-Rui Ye, Xing-Guo Huang *

College of Instrumentation and Electrical Engineering, Jilin University, Changchun, 130012, Jilin, China

ARTICLE INFO

Article history:
Received 2 April 2025
Received in revised form
20 December 2025
Accepted 4 February 2026
Available online 9 February 2026

Edited by Meng-Jiao Zhou

Keywords:
Inverse scattering theory
Full waveform inversion
Anisotropy
Multi-parameter inversion
Uncertainty quantification
Regularization term
Randomized singular value decomposition

ABSTRACT

Complex subsurface structures exhibit significant anisotropic characteristics, making multi-parameter imaging techniques important for achieving a more comprehensive geological interpretation. Full-waveform inversion (FWI) as a state-of-the-art method for reconstructing subsurface properties based on seismic wavefield modeling and data misfit minimization has been widely applied to isotropic media in both synthetic and field datasets. However, challenges such as crosstalk correlation and inaccuracy of the initial model indicate that further advancements are required to enhance resolution and computational efficiency. We propose an elastic FWI in the frequency domain for two-dimensional (2D) TI media to characterize their physical properties appropriately, as they are common in sedimentary basin environments. Different from traditional inversion schemes, our approach is formulated based on Bayesian inference, which automatically facilitates uncertainty analysis of the inversion results. Seismic data are acquired via the integral equation (IE) method grounded in scattering theory, where the sensitivity kernel is explicitly constructed using Green's functions, hence facilitating the calculation of gradient and Hessian. A Krylov subspace iterative method provides the approximated solution of the Lippmann-Schwinger (L-S) equation without sacrificing the accuracy. Furthermore, we incorporate the minimum support (MS) stabilizing functional as a model misfit term to regularize the objective function. A randomized singular value decomposition (SVD) approach is used to approximate and decompose the prior preconditioned Hessian. Both the model and covariance are updated through the iterative extended Kalman filter (IEKF) that implemented in the form of the Levenberg–Marquardt (LM) algorithm, thereby enabling practical uncertainty quantification. Numerical tests are conducted on two synthetic TI models with vertical and tilted symmetry axes, respectively, illustrating the precision and robustness of our method.

1. Introduction

As an attractive method for reconstructing the physical properties of complex geological models, FWI remains state-of-art in geophysical exploration. It is conceptualized with a data-fitting process, which provides more satisfying images than travel-time-based methods (Tarantola, 1984; Pratt, 1999). By utilizing both kinematic and dynamic information from multi-component pre-stack surveys, FWI aims to recover small-scale structures and lithological details, thereby quantifying high-resolution subsurface models. Several scenarios have reported the successful applications of FWI, such as regional studies, medical imaging, and seismic exploration (Bachmann and Tromp, 2020; Warner and Guasch, 2016; Aghazade et al., 2024), further demonstrating its substantial industrial potential. However, FWI approaches based on acoustic approximations only address P-wave velocity, which is insufficient to describe the actual seismic traces due to the anisotropic nature of the subsurface media, particularly in complex environment such as the sedimentary basin. For some petroleum-related datasets, introducing anisotropic characteristic can help to achieve precise kinematic matching, particularly when

* Corresponding author.
E-mail address: xingguohuang20@gmail.com (X.-G. Huang).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2026.02.005

the field data include widely ranged offsets, azimuths, and wide-angle refracted arrivals (Prieux et al., 2011; Warner et al., 2013).

Several studies have proposed strategies for incorporating anisotropy into the acoustic wave propagation operator, thus achieving acoustic approximation for anisotropic FWI with limited computational cost (Alkhalifah and Plessix, 2014; Aghazade et al., 2024; Alkhalifah et al., 2018). Nevertheless, wavefield propagation modeling methods based on acoustic approximation for TI media may take the undesired S-waves as artifacts (Operto et al., 2013). To better characterize the wave propagation within anisotropic media, FWI based on the elastic wave equation is necessary, as it governs the estimation of multiple parameters, such as attenuation, density, velocity, and so on. Updating any of them is a challenging task, as parameters with different physical characteristics may cause coupled influence on the seismic responses, this generally referred to as cross-correlations or trade-offs (Operto et al., 2013). In addition, parameters of different categories may vary in their orders of magnitude and influence on the wavefield, leading to poorly conditioned inversion results without appropriately scaled. Complex wavefield components in anisotropic FWI encounters more nonlinearity and ill-posedness when compared to isotropic cases. Barnes et al. (2008) applied FWI to cross-well seismic data for TI media, reconstructing only velocities successfully. Chang and McMechan (2009) addressed the limitations when dealing with three-component anisotropic data from a three-layered model, while highlighting the necessity of a high-resolution initial model. Prieux et al. (2013) applied elastic FWI to OBC and hydrophone data from the Valhall field data, but the update only includes the vertical P- and S-wave velocities. Kamath and Tsvankin (2016) performed elastic FWI on multicomponent data for TI media with a vertical symmetry axis (VTI) and proposed a detailed sensitivity analysis. Jakobsen et al. (2020) proposed a volume IE solution for elastic VTI media, achieving the successful reconstruction of five dependent elastic tensors. Shekhar et al. (2024) implemented the volume integral method for 3D microseismic wavefield modeling with different moment tensor sources.

Transforming seismic response into subsurface physical properties is a classic ill-posed and nonlinear geophysical inverse problem. Several theories have been proposed to address this issue, which can generally be categorized into two groups: (1) deterministic methods based on gradient optimization techniques, which are concerned with effective model optimization while addressing challenges such as convergence rate and cycle skipping (Pratt, 1999; Tarantola, 1984, 2005; Warner and Guasch, 2016); and (2) probabilistic methods, most notably Bayesian inference (Gouveia and Scales, 1998; Kaufl et al., 2013; Ye et al., 2023), which aim to identify a group of models that maximize the posterior probability density function (PDF). Markov Chain Monte Carlo (MCMC) methods are extensively applied in Bayesian inference, where they generate a sequence of samples from the full posterior distribution by performing a structured random walk through parameter space (Mosegaard and Tarantola, 1995; Brooks et al., 2011). MCMC approaches are inherently unbiased and, by construction, converge to the true posterior given sufficient time. The Metropolis–Hastings algorithm is one such method (Metropolis and Ulam, 1949; Hastings, 1970) that can be used to tackle the FWI problem (Ray et al., 2017; Guo et al., 2020). However, their practical application is often hindered by the curse of dimensionality, as they converge slowly and require enormous computational cost even for small-scale datasets (Tarantola, 2005; Sambridge and Mosegaard, 2002). Recent developments have focused on advanced MCMC algorithms such as Hamiltonian Monte Carlo (HMC), which sample the full posterior without dimensionality reduction or assumptions about its characteristics (Duane et al., 1987; Gebraad et al., 2020; Kotsi et al., 2020; Fichtner et al., 2019). Other advanced methods, such as Langevin Monte Carlo (Sambridge, 2014) and stochastic Newton MCMC (Martin et al., 2012; Zhao and Sen, 2019), also suffer from similar challenges. Another alternative methodology that makes Bayesian inference computationally efficient is variational inference (VI). It achieves uncertainty quantification through optimization of approximate solutions to posterior probability distributions (Blei et al., 2017). The method can be interpreted as a deterministic optimization approach utilizing gradient-based updates, thus offering a more favorable balance between computational complexity and estimation accuracy in contrast to the stochastic sampling nature of MCMC methods (Nawaz and Curtis, 2018; Zhang and Curtis, 2020). VI is scalable to larger datasets by partitioning the data into smaller minibatches and employing stochastic and distributed optimization techniques (Robbins and Monro, 1951; Kubrusly and Gravier, 1973). Moreover, it is inherently parallelizable at the level of individual samples, enabling efficient real-time computation by leveraging modern high-performance computing resources. VI has been applied to various geophysical research methods, including FWI, traveltime tomography, petrophysical inversion, etc. (Nawaz and Curtis, 2018; Zhang and Curtis, 2020). As a result, variational inference can be more computationally efficient than MCMC methods and provide better scaling to higher dimensionality (Bishop and Nasrabadi, 2006; Blei et al., 2017; Zhang et al., 2018).

For Bayesian FWI, the posterior covariance matrix (PCM) is one of the numerical tools for uncertainty or resolution analysis, while it is hard to be calculated exactly (Tarantola, 2005). Instead, it is generally obtained by manipulating the Hessian information (Tarantola, 2005; Fernandes et al., 2024; Bui-Thanh et al., 2013; Zhu et al., 2016; Chen et al., 2024). Mora (1987) neglected the estimation of the covariance matrix, focusing only on the maximum a posteriori (MAP) estimate, providing a single model as the solution without considering uncertainty issues. By combining prior knowledge with a likelihood function that relates model parameters to measurements, Bayesian inverse frameworks provide the solution in terms of a PDF to better quantify the uncertainty of the posterior candidates, instead of only focusing on a single solution (Tarantola, 2005; Bosch et al., 2010; Sebacher and Hanea, 2024). Bui-Thanh et al. (2013) proposed a tomography method with infinite-dimensional Bayesian inference, obtaining the posterior covariance around the MAP solution. Fichtner and van Leeuwen (2015) performed autocorrelation on the Hessian from random samples to investigate the directional resolution lengths of tomography. Huang et al. (2020) performed the Bayesian FWI on the elastic anisotropic model and provided standard deviations to evaluate the inversion results. In this work, we will utilize the generalization of the Kalman filter, the iterative extended Kalman filter (IEKF), to obtain the uncertainty estimates within the Bayesian framework, where the data are a series of measurements observed over time (Kalman, 1960; Bell and Cathey, 1993; Skoglund et al., 2015). Eikrem et al. (2018) validated the efficiency of the IEFK to deal with the time-lapse data. Thurin et al. (2019) performed a joint FWI based on the ensemble transform Kalman filter to achieve uncertainty estimation in tomography issues. Explicit computation of the Hessian matrix is prohibitively expensive as the scale of it is usually $(N_m \times N)^2$, where N_m is the number of parameters and N denotes the size of the individual parameter. To address large-scale problems, matrix probing theory has gained attention. Bui-Thanh et al. (2013) used eigen decomposition to construct a low-rank approximated data misfit Hessian. Zhu et al. (2016) provided an efficient Hessian manipulation by introducing a randomized SVD and point-spread function. Liu and Peter (2019) combined the randomized SVD with the SRVM method, enhancing both the flexibility and efficiency for FWI.

When subsurface models contain large-contrast structures, overly smooth updates may fail to capture the edge of the boundary during the early iterations. Kazei et al. (2017). Stabilizing functionals can address this issue, especially in electromagnetic and gravity inversion. Thus, regularization terms are also worth to be given further application in seismic inversion, where salt bodies are existing as strong scattering structures in some cases, increasing the nonlinearity of the FWI. The traditional regularization terms, such as L_2-norm and Tikhonov, can capture the high wavenumber components (Candansayar, 2008). The minimum support (MS) stabilizing functional emphasizes the differences between the given prior model and the updated model (Fiandaca et al., 2015). It aims to minimize the discontinuous and steep area, the inversion with the MS regularization term is also regarded as the focusing inversion. Zhdanov (2002) demonstrated that the MS stabilizing functional can realize high wavenumber updates and promoted inverted results. Thus, we would like to introduce the MS stabilizing functional as a regularization term aligned with Bayesian inference to build the inversion scheme for elastic anisotropic models.

The wavefield propagation in anisotropic media is implemented based on scattering theory in this study, namely the IE method. Several distinct advantages make the IE method distinguished from conventional numerical modeling approaches, such as finite-difference (FD) methods: (1) it necessitates the discretization of only the anomalous region, whereas FD methods demand the discretization of the entire model (Taylor, 1972; Kouri et al., 2003; Shekhar et al., 2024). It is also beneficial for developing the target-oriented inversion (Kouri et al., 2003; Huang et al., 2024); (2) it removes the necessity for setting artificial absorbing boundaries, as the Sommerfeld radiation condition is automatically satisfied (Taylor, 1972; Kouri et al., 2003); (3) the Green's functions are derived to characterize the wavefield, and the sensitivity kernel can be naturally and explicitly constructed, facilitating the gradient calculation and sensitivity analysis during the inversion (Taylor, 1972; Kouri et al., 2003; Djebbi et al., 2017); (4) time-stepping is unnecessary within IE methods, thereby avoiding differentiation-related numerical artifacts (Taylor, 1972). The IE methods can provide reliable solutions through the Born and Rytov approximations (Kouri et al., 2003; Osnabrugge et al., 2016; Huang et al., 2019). Besides, matrix-based derivatives within the IE methods are more transparent and explicit but computationally more demanding. To address these problems, iterative methods based on Krylov subspace can be taken into consideration.

Based on the previous work in Ye et al. (2023), this study proposes an effective FWI that focus on the following enhancements: First of all, we incorporate the MS stabilizing functional as a model misfit term to regularize the objective function. Subsequently, a randomized SVD approach is used to approximate and decompose the prior preconditioned Hessian. Then, a more comprehensive uncertainty analysis is presented. The outline of this work begins with an overview of the IE-based forward modeling method and its Krylov subspace iterative solver in elastic anisotropic media, then a glance at elastic moduli construction for VTI and TTI (the TI model with tilted symmetry axis) media. Next, a multiparameter sensitivity kernel is constructed through the third- and fourth-rank tensors of the Green's functions. Later on, a Bayesian inversion with MS stabilization functional is reviewed and followed by the Levenberg-Marquardt recursive Algorithm. After that, we propose a numerical solution by employing the randomized SVD for the prior-preconditioned Hessian to incorporate the uncertainty analysis. Finally, we evaluate our method on both VTI and TTI synthetic experiments and discuss the relationship between the inversion behavior of each parameter and its physical nature. In addition to the satisfactory reconstruction of all independent elastic moduli, the uncertainty is obtained through the estimation of the PCM, validating the effectiveness of the proposed algorithm.

2. Methodology

2.1. Elastic scattering theory for anisotropic media

Wavefields generated in heterogeneous anisotropic elastic media may undergo reflection, refraction, and diffraction. All of these events are considered by the elastodynamic wave equation, enabling a more precise approximation of wavefield propagation. For anisotropic media characterized by the mass density $\boldsymbol{\rho}(\mathbf{x})$ and stiffness tensor component $\mathbf{C}(\mathbf{x})$ at one specific angular frequency ω, the components of the displacement field $\boldsymbol{\psi}$ for the particle $\mathbf{x}$ in response to a single-force source satisfies (Cervený, 2001):

$$\nabla \cdot \mathbf{C}(\mathbf{x}) : \nabla_s \boldsymbol{\psi}(\mathbf{x}, \omega) + \boldsymbol{\rho}(\mathbf{x})\omega^2 \boldsymbol{\psi}(\mathbf{x}, \omega) = -\mathbf{S}(\mathbf{x}, \omega), \tag{1}$$

where $\nabla(\cdot)$ denotes the divergence operator and : represents the summation over the two indices. ∇_s is the gradient of the displacement field that can be expressed as the strain field tensor $\boldsymbol{\varepsilon}(\mathbf{x}, \omega)$, with the relationship characterized by:

$$\nabla_s \boldsymbol{\psi}(\mathbf{x}, \omega) = \boldsymbol{\varepsilon}(\mathbf{x}, \omega) = 1 \Big/ 2\left\{\nabla \boldsymbol{\psi}(\mathbf{x}, \omega) + [\nabla \boldsymbol{\psi}(\mathbf{x}, \omega)]^{\mathrm{T}}\right\} \tag{2}$$

The displacement field generated by the source $\mathbf{S}$ is indicated as the complex-valued data with three components. Under the assumption of constant density, we consider only the elastic moduli $\mathbf{C}(\mathbf{x})$ as the model parameter to be estimated. Based on scattering theory, the subsurface model can be considered homogeneous as the background and heterogeneous as the perturbation, which can be denoted as $\mathbf{C}^{(0)}$ and $\Delta\mathbf{C}$, respectively (Cervený, 2001; Jakobsen et al., 2020; Shekhar et al., 2024). The combination of these two parts can characterize the subsurface medium, namely:

$$\mathbf{C}(\mathbf{x}) = \mathbf{C}^{(0)}(\mathbf{x}) + \Delta\mathbf{C}(\mathbf{x}). \tag{3}$$

By integrating Eq. (3) into Eq. (1), we can obtain:

$$\nabla \cdot \mathbf{C}^{(0)}(\mathbf{x}) : \boldsymbol{\varepsilon}(\mathbf{x}, \omega) + \boldsymbol{\rho}(\mathbf{x})\omega^2 \boldsymbol{\psi}(\mathbf{x}, \omega) = -\mathbf{S}(\mathbf{x}) - \nabla \cdot \Delta\mathbf{C}(\mathbf{x}) : \boldsymbol{\varepsilon}(\mathbf{x}, \omega), \tag{4}$$

then $\boldsymbol{\psi}(\mathbf{x}, \omega)$ can be calculated through the volume integral equation as (Jakobsen et al., 2020; Shekhar et al., 2024):

$$\boldsymbol{\psi}(\mathbf{x}, \omega) = \int \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega)\mathbf{S}(\mathbf{x}', \omega)\mathrm{d}\mathbf{x}', \tag{5}$$

where $\mathbf{x}'$ is a point near $\mathbf{x}$. The Green's function $\mathbf{G}(\mathbf{x}, \mathbf{x}', \omega)$ in scattering theory describes the fingerprint of a particle propagation due to an impulse function (Cervený, 2001), which satisfies:

$$\nabla \cdot \mathbf{C}(\mathbf{x}) : \nabla \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega) + \boldsymbol{\rho}(\mathbf{x})\omega^2 \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega) = -\delta(\mathbf{x} - \mathbf{x}'), \tag{6}$$

where the Dirac function $\delta(\mathbf{x} - \mathbf{x}')$ denotes the unit impulse. The right-hand side in Eq. (4) can be interpreted as the so-called contrast source. Referring to the volume integral equation as Eq.

(5), by treating this term similar to a normal source, after employing Green's function representation, the practical displacement can be reformulated exactly as (Jakobsen et al., 2003, 2020):

$$\boldsymbol{\psi}(\mathbf{x},\omega)=\boldsymbol{\psi}^{(0)}(\mathbf{x},\omega)+\int \mathbf{G}_{qp}^{(0)}(\mathbf{x},\mathbf{x}',\omega)\cdot\nabla\cdot\Delta\mathbf{C}(\mathbf{x}'):\boldsymbol{\varepsilon}(\mathbf{x}',\omega)\mathrm{d}\mathbf{x}', \tag{7}$$

which illustrates the linear superposition of the scattered and background fields. $\mathbf{G}_{qp}^{(0)}(\mathbf{x},\mathbf{x}',\omega)$ is the particle displacement at point $\mathbf{x}$ for the q-th component caused by a point force applied in the p-th component at point $\mathbf{x}'$ (Jakobsen et al., 2020; Shekhar et al., 2024), which is the solution to:

$$\nabla\cdot\mathbf{C}^{(0)}(\mathbf{x}):\nabla\mathbf{G}_{qp}^{(0)}(\mathbf{x},\mathbf{x}',\omega)+\boldsymbol{\rho}(\mathbf{x})\omega^{2}\mathbf{G}_{qp}^{(0)}(\mathbf{x},\mathbf{x}',\omega)=-\delta(\mathbf{x}-\mathbf{x}'). \tag{8}$$

The elastodynamic background Green's functions can be calculated analytically based on its underlying physics interpretation (Cervený, 2001). By conducting a partial integration, we may remove the gradient operator of the elastic stiffness perturbation based on the symmetric of $\boldsymbol{\varepsilon}(\mathbf{x}',\ \omega)$. Therefore, the displacement field can be conveyed as (Ye et al., 2023; Jakobsen et al., 2003):

$$\boldsymbol{\psi}(\mathbf{x},\omega)=\boldsymbol{\psi}^{(0)}(\mathbf{x},\omega)+\int \mathbf{G}_{qps}^{(0)}(\mathbf{x},\mathbf{x}',\omega):\Delta\mathbf{C}(\mathbf{x}'):\boldsymbol{\varepsilon}(\mathbf{x}',\omega)\mathrm{d}\mathbf{x}', \tag{9}$$

where $\mathbf{G}_{pqs}^{(0)}(\mathbf{x},\mathbf{x}',\omega)$ represents the first-order spatial derivative of the background Green's tensor (Jakobsen et al., 2003, 2020). It is constructed as symmetric with respect to an interchange of the last two indices and can be understood as the p-th component of the particle displacement generated by the qs-th component of the stress tensor, namely:

$$\mathbf{G}_{pqs}^{(0)}(\mathbf{x},\mathbf{x}',\omega)=1\Big/2\left[\mathbf{G}_{pq}^{(0)}(\mathbf{x},\mathbf{x}',\omega)\Big/\partial\mathbf{x}_{s}+\mathbf{G}_{ps}^{(0)}(\mathbf{x},\mathbf{x}',\omega)\Big/\partial\mathbf{x}_{q}\right]. \tag{10}$$

The particle displacement field $\boldsymbol{\psi}^{(0)}(\mathbf{x},\omega)$ of the background model can be articulated as (Jakobsen et al., 2003, 2020; Huang et al., 2020):

$$\boldsymbol{\psi}^{(0)}(\mathbf{x},\omega)=\int \mathbf{G}^{(0)}(\mathbf{x},\mathbf{x}',\omega)\mathbf{S}(\mathbf{x}',\omega)\mathrm{d}\mathbf{x}'. \tag{11}$$

The strain field is "scattered" from $\mathbf{x}$ to $\mathbf{x}'$ by the stiffness perturbation $\Delta\mathbf{C}(\mathbf{x})$ via the "propagator" $\mathbf{G}^{(0)}(\mathbf{x},\ \mathbf{x}',\ \omega)$. Thus, the strain field can be determined by taking spatial derivatives to Eq. (9), which can be reformulated exactly as:

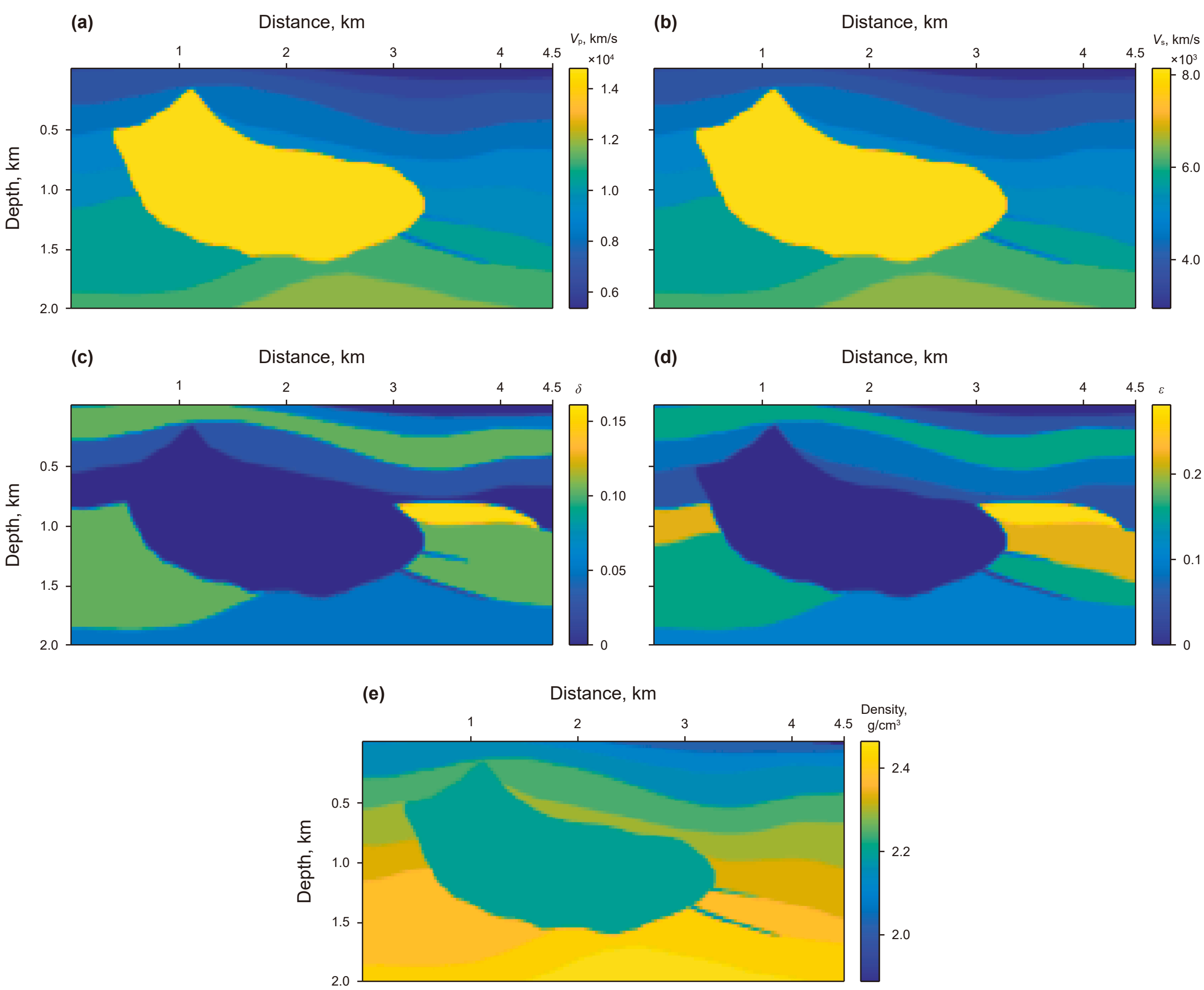

Fig. 1. Hess VTI model described by anisotropy parameters of **(a)** P-wave velocity, **(b)** S-wave velocity, Thomsen parameters **(c)** δ, **(d)** ε, and **(e)** density.

$$\boldsymbol{\varepsilon}(\mathbf{x},\omega)=\boldsymbol{\varepsilon}^{(0)}(\mathbf{x},\omega)+\int\mathbf{G}^{(0)}_{pqst}(\mathbf{x},\mathbf{x}',\omega):\Delta\mathbf{C}(\mathbf{x}'):\boldsymbol{\varepsilon}(\mathbf{x}',\omega)\mathrm{d}\mathbf{x}',\tag{12}$$

where $\boldsymbol{\varepsilon}^{(0)}(\mathbf{x})$ is the strain tensor that is calculated on the background stiffness tensor $\mathbf{C}^{(0)}(\mathbf{x})$. $\mathbf{G}^{(0)}_{pqst}$ represents the fourth-rank tensor for a translation-invariant system, with the individual components calculated by (Jakobsen et al., 2020; Huang et al., 2020):

convenience in derivation when calculating the elastic wavefield for heterogeneous anisotropic media.

Consequently, the strain field in Eq. (12) and the displacement field in Eq. (9) can be represented as coupled Lippmann-Swinger (or Dyson) type integral equations in the operator form:

$$\boldsymbol{\varepsilon}=\boldsymbol{\varepsilon}^{(0)}+\mathbf{G}^{(0)}_{4}\Delta\mathbf{C}\,\boldsymbol{\varepsilon}.\tag{14}$$

$$\boldsymbol{\psi}=\boldsymbol{\psi}^{(0)}+\mathbf{G}^{(0)}_{3}\Delta\mathbf{C}\,\boldsymbol{\varepsilon}.\tag{15}$$

$$\mathbf{G}^{(0)}_{pqst}(\mathbf{x},\mathbf{x}',\omega)=1\Big/4\Big[\partial^2\mathbf{G}^{(0)}_{pq}(\mathbf{x},\mathbf{x}',\omega)\Big/\partial\mathbf{x}_s\partial\mathbf{x}_t+\partial^2\mathbf{G}^{(0)}_{qs}(\mathbf{x},\mathbf{x}',\omega)\Big/\partial\mathbf{x}_p\partial\mathbf{x}_t+\partial^2\mathbf{G}^{(0)}_{pt}(\mathbf{x},\mathbf{x}',\omega)\Big/\partial\mathbf{x}_s\partial\mathbf{x}_q+\partial^2\mathbf{G}^{(0)}_{st}(\mathbf{x},\mathbf{x}',\omega)\Big/\partial\mathbf{x}_p\partial\mathbf{x}_q\Big].\tag{13}$$

Similar to the third-rank tensor in Eq. (10), the physical interpretation of $\mathbf{G}^{(0)}_{pqst}$ can be explained as the pq-th component of the strain field at $\mathbf{x}$ due to the st-th component of the stress field at $\mathbf{x}'$. Following Jakobsen et al. (2020), the third- and fourth-rank tensors $\mathbf{G}^{(0)}_{pqs}$ and $\mathbf{G}^{(0)}_{pqst}$ may be considered as modified Green's functions, and in this paper they are simplified to $\mathbf{G}^{(0)}_{3}$ and $\mathbf{G}^{(0)}_{4}$ for

The structure of Eqs. (14) and (15) support the highly developed iterative solvers to realize a valid calculation. Apparently, the particle displacement field can be naturally derived from the strain field once it is obtained. Given that the strain field in Eq. (15) can be formulated as a linear system, there are various methods available for its solution. Direct solvers tend to be computationally impractical for large-scale or even moderately sized problems.

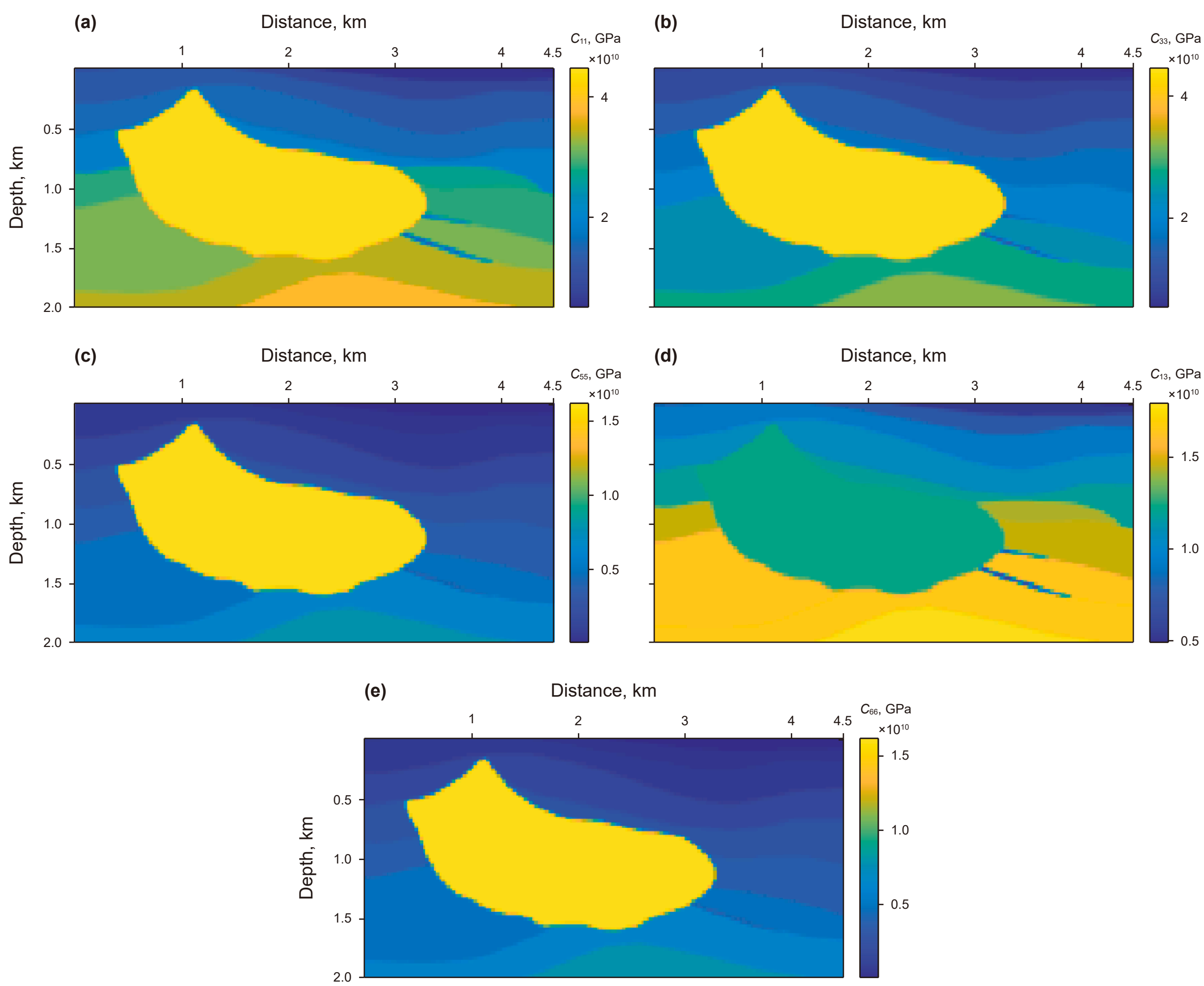

Fig. 2. Hess VTI model described by five independent Voigt elastic moduli.

Therefore, iterative solvers or parallel direct solvers can be applied to avoid this challenge. The linear system of Eq. (15) can be formulated after splitting $\boldsymbol{\varepsilon}^{(0)}$ into n column vectors:

$$\left(\mathbf{I}-\mathbf{G}_4^{(0)}\Delta\mathbf{C}\right)\boldsymbol{\varepsilon}=\left(\boldsymbol{\varepsilon}_1^{(0)},\boldsymbol{\varepsilon}_2^{(0)},\ldots,\boldsymbol{\varepsilon}_n^{(0)}\right)^{\mathrm{T}}, \tag{16}$$

where $\mathbf{I}$ is the identity matrix. We strive to find an approximation that best reduces the misfit between $\boldsymbol{\varepsilon}^{(0)}$ and $\left(\mathbf{I}-\mathbf{G}_4^{(0)}\Delta\mathbf{C}\right)\boldsymbol{\varepsilon}$. Since the matrix $\left(\mathbf{I}-\mathbf{G}_4^{(0)}\Delta\mathbf{C}\right)$ is well-conditioned, nonsymmetric and nonsingular, one of the most effective techniques to tackle this issue is the Krylov subspace iterative method, namely the Generalized Minimal Residual (GMRES) solver. Based on Krylov subspaces, it first projects the problem onto a smaller subspace, computes an approximate solution at a lower computational cost, and then maps the solution back to the full space (Saad and Schultz, 1986). To further improve the efficiency of the GMRES solver, we choose its right preconditioned version with a band-based discrete wavelet transform (DWT) preconditioner Υ in the Frobenius norm, in terms of the factorized sparse approximate preconditioner (Chen, 2005; Ford, 2003; Bertaccini and Durastante, 2018). The right preconditioned linear system of Eq. (16) can be expressed as:

$$\left(\mathbf{I}-\mathbf{G}_4^{(0)}\Delta\mathbf{C}\right)\Upsilon^{-1}y=\left(\boldsymbol{\varepsilon}_1^{(0)},\boldsymbol{\varepsilon}_2^{(0)},\ldots,\boldsymbol{\varepsilon}_n^{(0)}\right)^{\mathrm{T}}, \tag{17}$$

where $\boldsymbol{\varepsilon}=\Upsilon^{-1}y$. Algorithm 1 provides a step-by-step description of the strain field calculation procedure for better understanding.

Algorithm 1. Right preconditioned GMRES (modified according to Algorithm 2.8 of Bertaccini and Durastante (2018))

Input: background strain field $\boldsymbol{\varepsilon}^{(0)}$, initial Guess $\tilde{x}_0$, relative residual $\xi>0$, maximum number of iterations m, the DWT preconditioner Υ
Output: candidate approximation $\boldsymbol{\varepsilon}$
1. Set $A=\left(\mathbf{I}-\mathbf{G}_4^{(0)}\Delta\mathbf{C}\right)$, $\tilde{x}_0=0$
2. Calculate the initial residual $r_0=\boldsymbol{\varepsilon}^{(0)}-A\tilde{x}_0$
3. $\beta=\|r_0\|_2$, $v_1=r_0/\beta$, $\xi=\beta e_1$
4. **for** $j=1,\ldots,m$ do
5. $z_j=\Upsilon^{-1}v_j$, $w_j=Az_j$
6. **for** $i=1,\ldots,j$ **do**
7. $h_{i,j}=\left(w_j,v_i\right)$, $w_j=w_j-h_{i,j}v_i$
8. **end for**
9. $h_{j+1,j}=\|w_j\|_2$, $v_{j+1}=w_j/h_{j+1,j}$
10. Build the Krylov subspace $V_m=[v_1,\ldots,v_m]$
11. $y_m=\mathrm{argmin}\|\beta e_1-H_m y\|_2$, $\tilde{x}_m=\tilde{x}_0+\Upsilon^{-1}V_m y_m$
12.**end for**
12. **if** $\|r_j\|_2/\beta>\xi$ **then**
13. $\tilde{x}_0=\tilde{x}_m$ **return to 2**
14. **else**
15. **build the solution** $\boldsymbol{\varepsilon}=\tilde{x}_m$

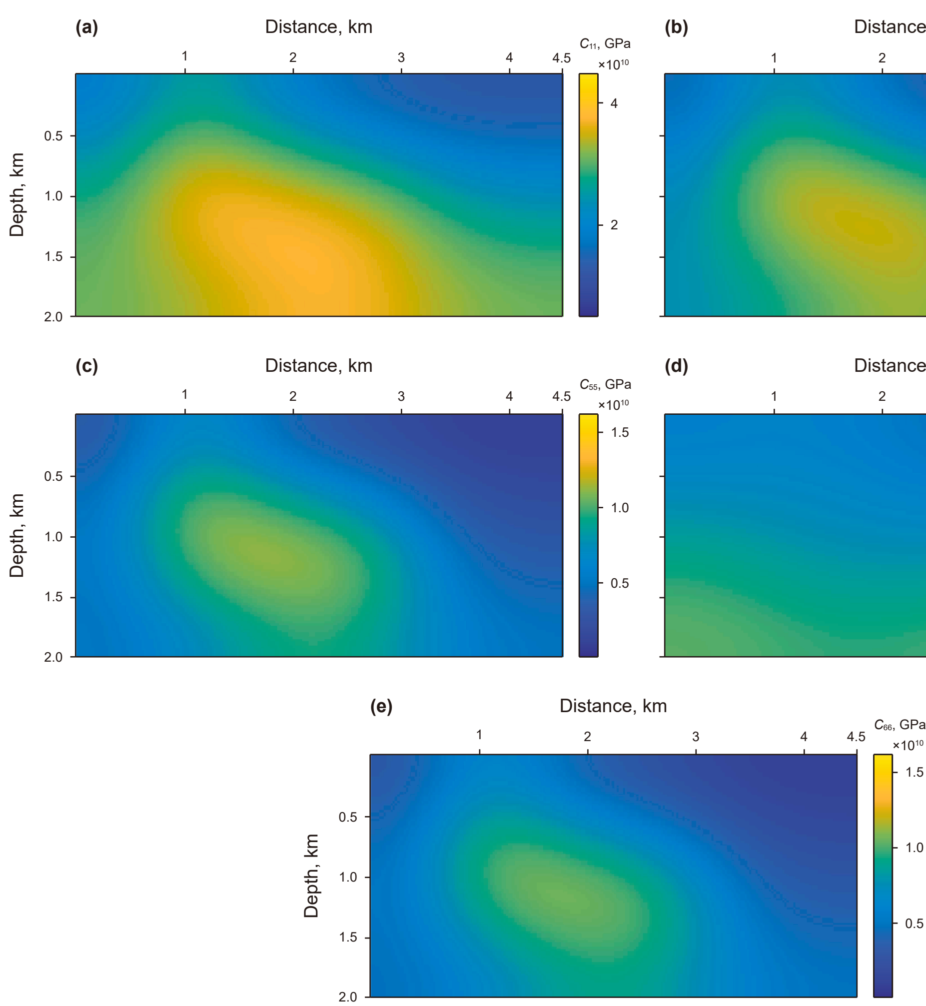

Fig. 3. Initial elastic moduli of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the Hess VTI subset model experiment, respectively.

2.2. Elastic moduli for transversely isotropic medium

The full elastic tensor provides a comprehensive representation to characterize the elastic properties of the subsurface media. It consists of a total of 81 parameters that correspond to different stress and strain responses in different directions. However, their symmetric nature reduces the number of independent elastic constants to merely 21 (Thomsen, 1986; Ting, 1996). Axisymmetric anisotropy is typically observed in sedimentary basin, thus 5 independent coefficients are sufficient to make the description. In this case, one direction is unique, whereas the other two directions are equivalent (Ting, 1996). The full elastic stiffness tensor of the 2D VTI subsurface model can be described by the following expression (Thomsen, 1986):

$$\mathbf{C}_{mn}^{\mathrm{VTI}} = \begin{bmatrix} C_{11} & C_{11}-2C_{66} & C_{13} & 0 & 0 & 0 \\ C_{11}-2C_{66} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{bmatrix}, \tag{18}$$

where the subscriptions m and n stand for the values that range from 1 to 6, within the 6×6 stiffness matrix. Define V_{p} and V_{s} as the P- and S-wave velocities, respectively; ε, γ and δ are the Thomsen anisotropy parameters (Thomsen, 1986; Ting, 1996). The independent elements in Eq. (18) are then given by:

$$\begin{bmatrix} C_{11} \\ C_{33} \\ C_{13} \\ C_{55} \\ C_{66} \end{bmatrix} = \begin{bmatrix} (1+2\varepsilon)C_{33} \\ \rho \cdot V_{\mathrm{p}}^2 \\ \left[2\delta C_{33}(C_{33}-C_{55}) + (C_{33}-C_{55})^2\right]^{1/2} - C_{55} \\ \rho \cdot V_{\mathrm{s}}^2 \\ (1+2\gamma)C_{55} \end{bmatrix}. \tag{19}$$

Note that C_{33} and C_{55} only account for velocities, whereas C_{11}, C_{13} and C_{66} additionally related to different anisotropic parameters. The TTI media usually leads to an internal angle, which may appear in some overthrust regions or parallel inclined fractures (Oh and Alkhalifah, 2016; Tsvankin, 2011). By convention, the azimuth angle is omitted, thus the parameter γ is set to be zero, and C_{66} will not be calculated for 2D TTI media, but it is possible to be determined using VTI coefficients along with the rotation angle. The elastic moduli of a 2D TTI medium can be obtained by rotating a VTI medium through Bond transformation (Bond, 1943; Zhu and Dorman, 2000):

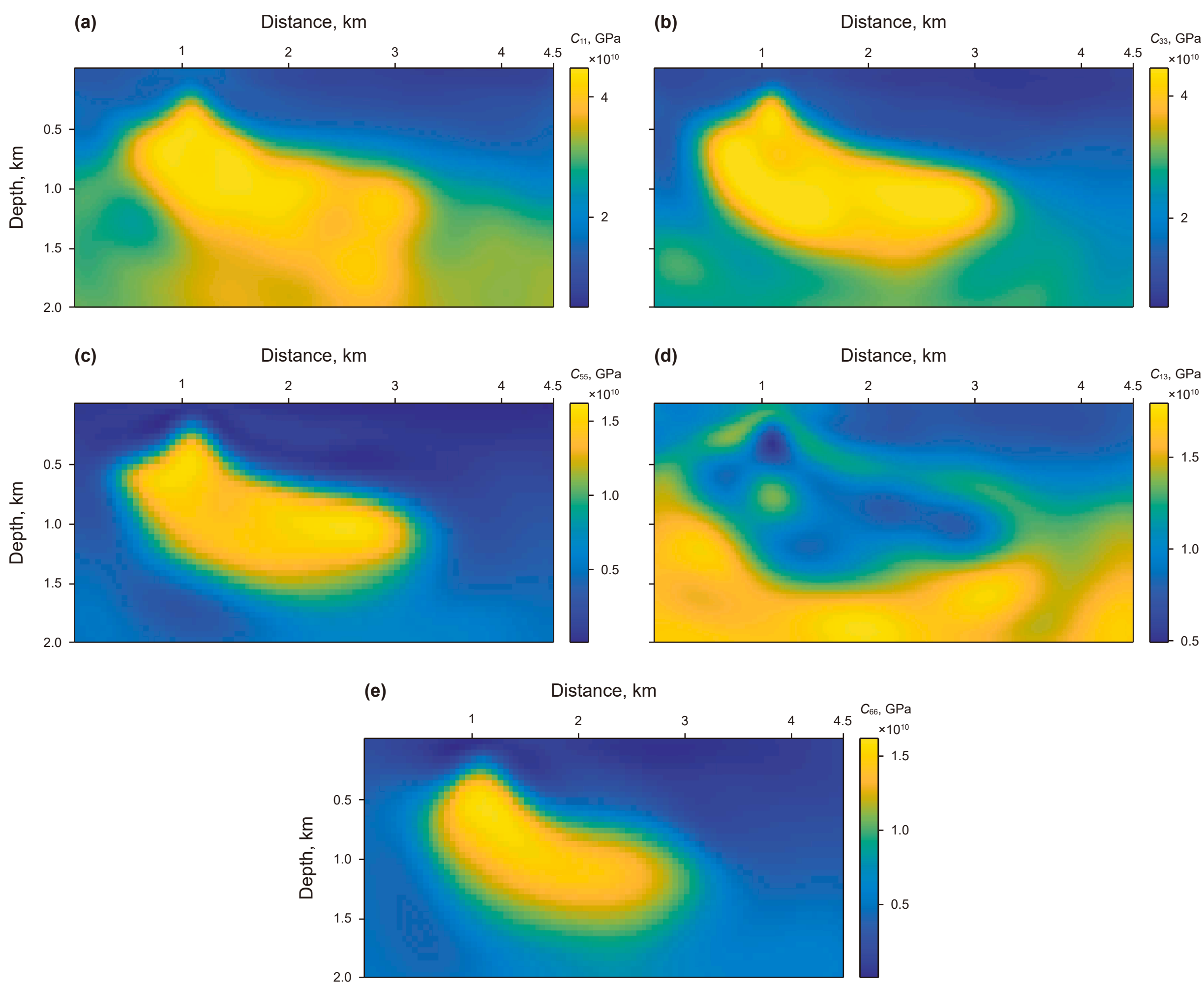

Fig. 4. The inverted results of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ at the first frequency (4 Hz) for the Hess VTI subset model experiment, respectively.

$$\mathbf{C}_{mn}^{\mathrm{TTI}} = \mathbf{M}_{\mathbf{a}} \mathbf{C}_{mn}^{\mathrm{VTI}} \mathbf{M}_{\mathbf{a}}^{\mathrm{T}}. \tag{20}$$

By denoting the tilt angle as θ, and the corresponding rotation matrix is then given by:

$$\mathbf{a} = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix}. \tag{21}$$

And the Bond matrix $\mathbf{M}$ is (Zhu and Dorman, 2000):

$$\mathbf{M} = \begin{bmatrix} \cos^2\theta & 0 & \sin^2\theta & 0 & -\sin 2\theta & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \sin^2\theta & 0 & \cos^2\theta & 0 & \sin 2\theta & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & \sin\theta \\ \frac{1}{2}\sin 2\theta & 0 & -\frac{1}{2}\sin 2\theta & 0 & \cos 2\theta & 0 \\ 0 & 0 & 0 & -\sin\theta & 0 & \cos\theta \end{bmatrix}. \tag{22}$$

Hence, the independent constants for 2D elastic TTI media in Voigt notation are clearly specified as (Macbeth, 2002; Oh and Alkhalifah, 2016):

$$\begin{bmatrix} C_{11}^{\mathrm{TTI}} \\ C_{13}^{\mathrm{TTI}} \\ C_{15}^{\mathrm{TTI}} \\ C_{33}^{\mathrm{TTI}} \\ C_{35}^{\mathrm{TTI}} \\ C_{55}^{\mathrm{TTI}} \end{bmatrix} = \frac{1}{8} \begin{bmatrix} 3C_{11} + 2C_{13} + 3C_{33} + 4C_{55} + 4\lambda_2 \cos 2\theta + \lambda_1 \cos 4\theta \\ C_{11} + 6C_{13} + C_{33} - 4C_{55} - \lambda_1 \cos 4\theta \\ 2(\lambda_2 \sin 2\theta + \lambda_1 \sin 4\theta) \\ 3C_{11} + 2C_{13} + 3C_{33} + 4C_{55} - 4\lambda_2 \cos 2\theta + \lambda_1 \cos 4\theta \\ 2(\lambda_2 \sin 2\theta - \lambda_1 \sin 4\theta) \\ C_{11} - 2C_{13} + C_{33} + 4C_{55} - \lambda_1 \cos 4\theta \end{bmatrix}. \tag{23}$$

In this work, we execute normalization of the elastic stiffness perturbation with their corresponding properties of the isotropic homogeneous model. This treatment aligns with the concept of the scattering potential as described in scattering theory, and we defined the normalized stiffness perturbations by the following expression as the model parameters (Eikrem et al., 2018; Jakobsen et al., 2020; Shekhar et al., 2024):

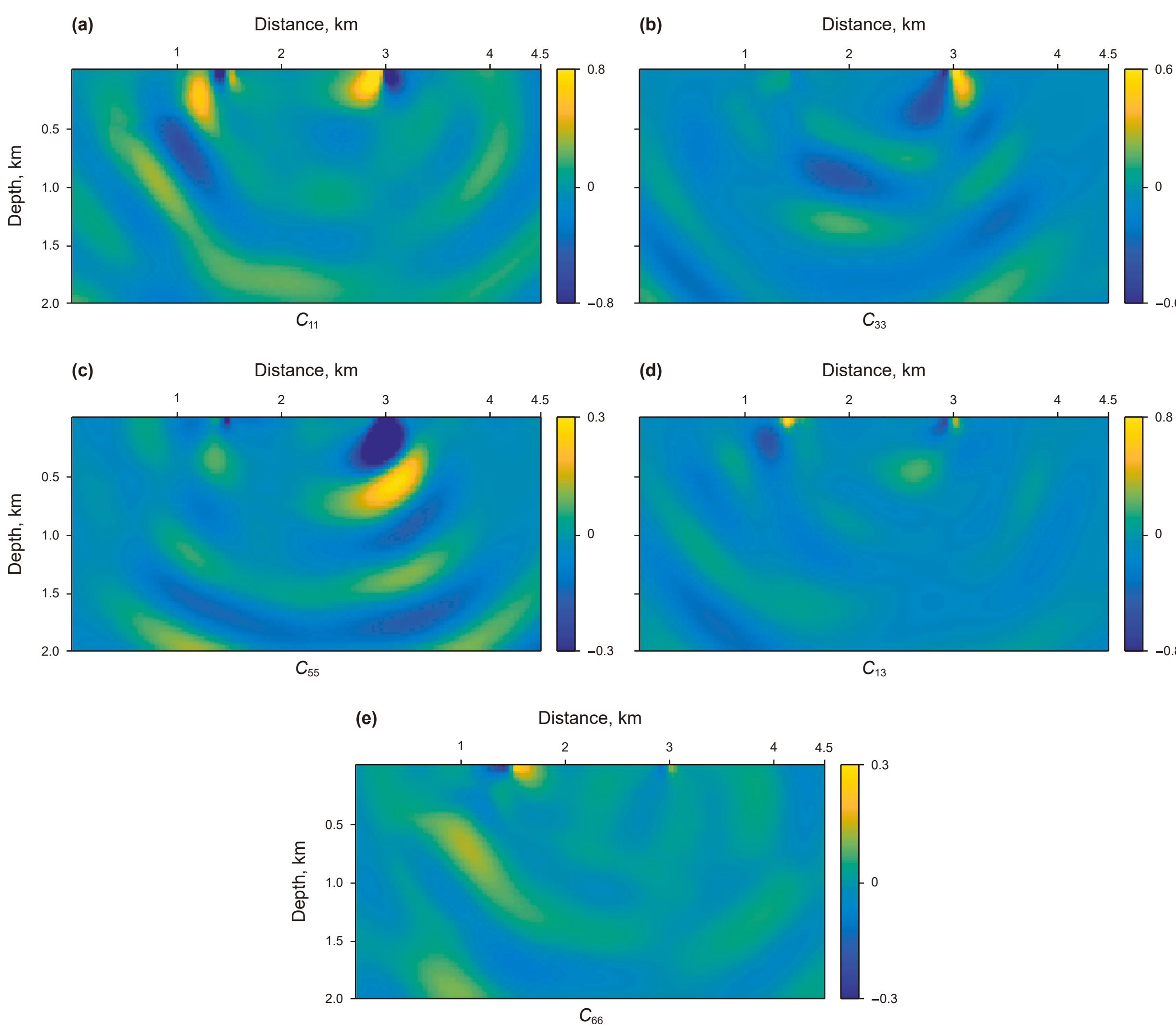

Fig. 5. For a source-receiver configuration with the source at $x = 1.5$ km and the receiver at $x = 3.0$ km, the Fréchet derivatives of **(a)–(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ at the first frequency (4 Hz) for the Hess VTI subset model experiment, respectively.

$$\mathbf{m}^{(p)} = \frac{\left[\mathbf{C}_{mn}(\mathrm{x}) - \mathbf{C}^{(0)}_{mn}(\mathrm{x})\right]}{\mathbf{C}^{(0)}_{mn}(\mathrm{x})}, \tag{24}$$

where the index p ranges from 1 to 5, for example, $\mathbf{m}^{(p)} = [\mathbf{m}^{(1)}, \mathbf{m}^{(2)}, \mathbf{m}^{(3)}, \mathbf{m}^{(4)}, \mathbf{m}^{(5)}]$, where $\mathbf{m}^{(1)} = \Delta C_{11}(\mathrm{x})/C^{(0)}_{11}(\mathrm{x})$. Normalizing the perturbations of model parameters guarantees a uniform magnitude of each parameter, which is beneficial for multi-parameter FWI.

2.3. Scattering integral implementation of multiparameter Fréchet derivative

In anisotropic elastic media, the nonlinear inverse scattering problems are addressed by estimating the elastic moduli via the measurement of the displacement-strain field at the receiver surface. As per Eq. (9), the residual for the particle displacement vector in the real-space coordinates is (Taylor, 1972; Gonis and Butler, 2000):

$$\delta\boldsymbol{\psi}^{(i)}_k(\mathbf{r}) = \boldsymbol{\psi}_k(\mathbf{r}) - \boldsymbol{\psi}^{(i)}_k(\mathbf{r}) = \int \mathbf{G}^{(i)}_3(\mathbf{r}, \mathbf{x}) \cdot \delta\mathbf{C}^{(i+1)}(\mathbf{x}) \cdot \boldsymbol{\varepsilon}^{(i)}(\mathbf{x}) \mathrm{d}\mathbf{x}, \tag{25}$$

where $\delta\mathbf{C}^{(i+1)}(\mathbf{x}) = \Delta\mathbf{C}^{(i+1)}(\mathbf{x}) - \Delta\mathbf{C}^{(i)}(\mathbf{x})$ is the variation in the elastic moduli between two successive iterations. $\boldsymbol{\varepsilon}^{(i)}$ and $\mathbf{G}^{(i)}_3$ represent the strain field and the Green's function at the i-th iteration, respectively. The index k denotes the data components and $\mathbf{r}$ represents the receiver locations. Eq. (25) describes the data residual caused by the model perturbation between two consecutive iterations (Jakobsen et al., 2020).

Referring to Eq. (24), each perturbated elastic tensors can be calculated as the model perturbation processed through the corresponding structure matrix $\mathbf{B}^{(p)}_{mn}(\mathbf{x})$ (Jakobsen et al., 2020):

$$\delta\mathbf{C}(\mathbf{x}) = \sum_{p=1}^{21} \mathbf{B}^{(p)}_{mn}(\mathbf{x})\, \delta\mathbf{m}^{(p)}(\mathbf{x}). \tag{26}$$

Inserting Eq. (26) into Eq. (25), we obtain:

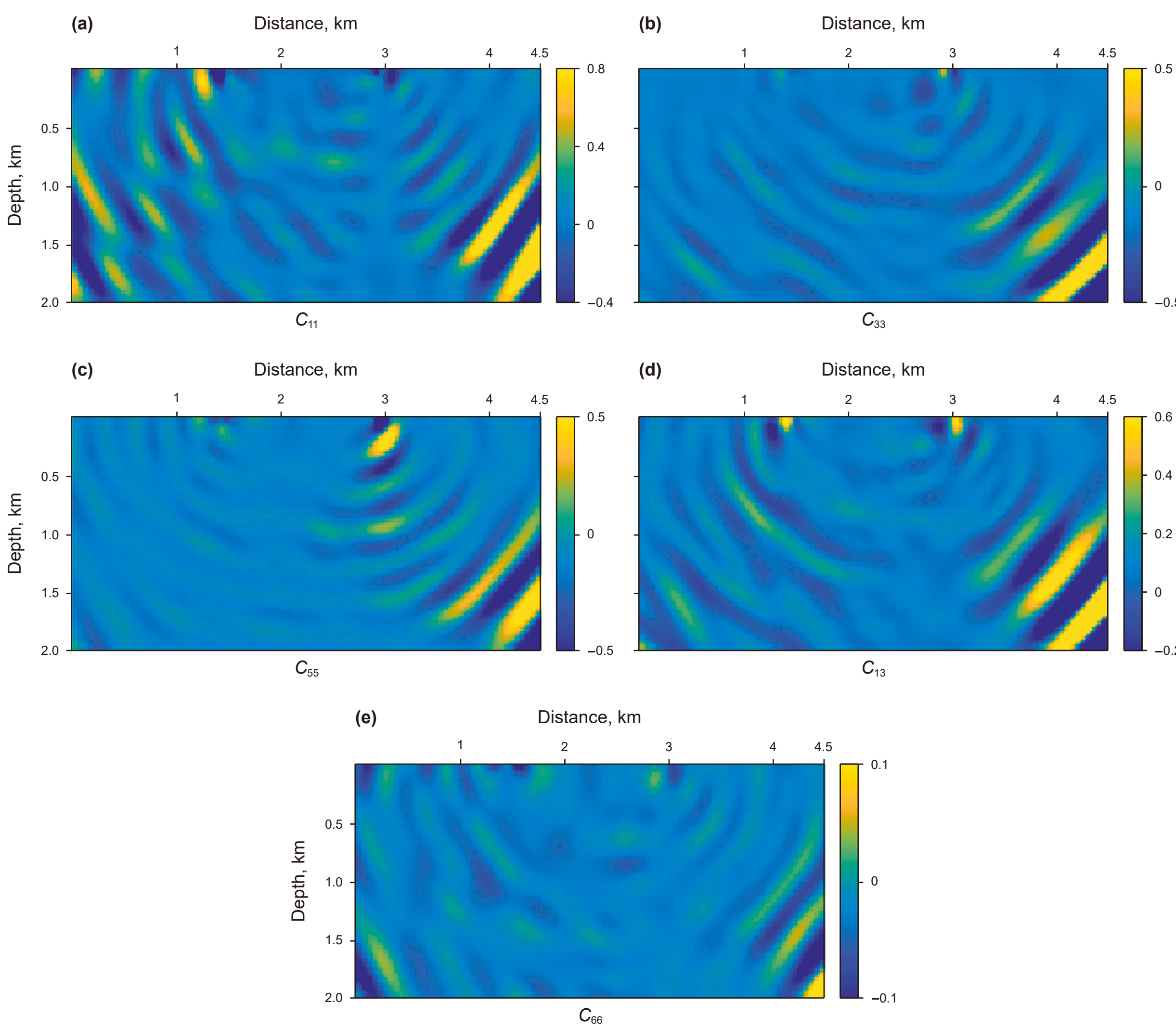

Fig. 6. For a source-receiver configuration with the source at $x = 1.5$ km and the receiver at $x = 3.0$ km, the Fréchet derivatives of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ at 20 Hz for the Hess VTI subset model experiment, respectively.

$$\delta\boldsymbol{\psi}_k^{(i)}(\mathbf{r}) = \int \mathbf{G}_3^{(i)}(\mathbf{r},\mathbf{x})\mathbf{B}^{(p)}(\mathbf{x})\ \boldsymbol{\varepsilon}^{(i)}(\mathbf{x})\delta\mathbf{m}^{(i+1,p)}(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{27}$$

The Fréchet derivative, also known as the sensitivity kernel, is commonly understood to measure the effect of small variations about model parameters on data perturbations as the model perturbation tends to infinity. It provides the mapping between the model and data spaces with its definition of $\mathscr{F} = \delta\boldsymbol{\psi}/\ \delta\mathbf{m}$ (Huang et al., 2020; Tarantola, 2005). It follows from Eqs. (23)–(27) that:

$$\delta\boldsymbol{\psi}_k^{(i)}(\mathbf{r}) = \int \mathscr{F}_k^{(i,p)}\delta\mathbf{m}^{(i+1,p)}(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{28}$$

Hence, we can observe that for each source-receiver pair, the multiparameter nonlinear Fréchet derivative derived from the Lippmann-Schwinger (L-S) scattering integral is:

$$\mathscr{F}^{(i,p)}(\mathbf{x}) = \mathbf{G}_3^{(i)}(\mathbf{r},\mathbf{x})\mathbf{B}^{(p)}(\mathbf{x})\ \boldsymbol{\varepsilon}^{(i)}(\mathbf{x}), \tag{29}$$

where we construct the sensitivity kernel explicitly via an efficient direct approach that uses two actual Green's functions rather than the background. This naturally incorporates the multiple scattering behavior of anisotropic seismic waves, therefore facilitating the calculation of the gradient and Hessian, and providing a notable improvement over traditional adjoint-state methods and conjugate gradient techniques.

2.4. Bayesian inference for inverse scattering problem

FWI is fundamentally an ill-posed, non-unique, and nonlinear inverse problem that is often solved through local optimization techniques. It aims to estimate subsurface model parameters by minimizing the discrepancy between the observed and the numerical wavefields, which is typically framed as a linearized least-squares problem. The solution exhibits high sensitivity to random noise and measurement errors, while the observation configuration is limited and subject to inherent uncertainties. Fortunately, incorporating prior statistical information as additional data weighting or model penalty terms can substantially mitigate these challenges, improving both the accuracy and stability by better constraining the prior information, which is typically introduced via a regularization term and then combined with the objective function (Pratt, 1999).

Bayesian inference, rooted in statistics theory, updates the prior probability as further information is acquired (Tarantola, 2005). When applied to the FWI workflow, it allows for the integration of prior physical information about the subsurface model into

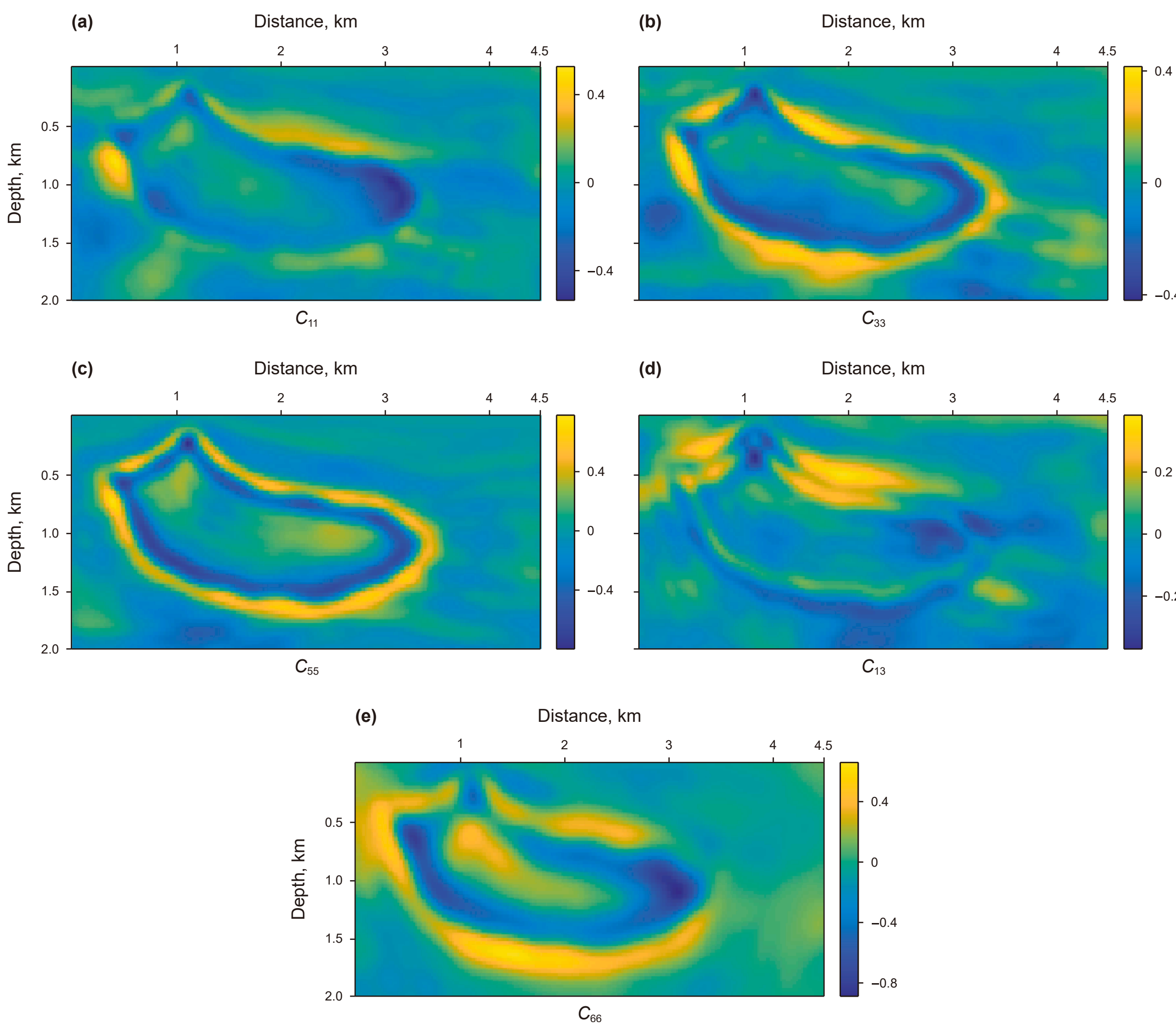

Fig. 7. Frequency-domain normalized gradient of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the Hess VTI subset model experiment, respectively.

seismic traces, thereby enabling the quantification of posterior uncertainties in the inverted results. We believe this characteristic is more vital when two different data sets, such as seismic and production data are accessible. This is precisely where optimization problems within a Bayesian framework come into play, as they simultaneously incorporate both the model parameters and their associated uncertainties, thereby facilitating a more comprehensive evaluation of the solution's reliability.

Bayesian inference defines an a priori probability density over the model space aligned with the data misfit to form a MAP probability density, which reflects both the fit to the data and model reasonableness. In seismic inversion, Bayesian inference describes the correlation between a hypothesis and given evidence, that can be represented by the model parameters $\mathbf{m}$ and the observed data $\mathbf{d}_{\text{obs}}$. The prior distribution $p(\mathbf{m})$ is defined as (Tarantola, 2005):

$$p(\mathbf{m})\propto\exp\left[-\frac{1}{2}\left(\mathbf{m}-\mathbf{m}_{\text{pri}}\right)^{\mathrm{T}}\Gamma_{\mathrm{M}}^{-1}\left(\mathbf{m}-\mathbf{m}_{\text{pri}}\right)\right], \tag{30}$$

where Γ_{M} stands for the model covariance, the superscript T indicates the transpose operator, and $\mathbf{m}_{\text{pri}}$ refers to the prior model, assumed to be a Gaussian distribution (Zhu et al., 2016; Ye et al., 2023). $p(\mathbf{m})$ is independent of the seismic data, and it is typically derived from independent measurements. The likelihood function $p(\mathbf{d}|\mathbf{m})$ characterizes the discrepancy between the given observed $\mathbf{d}_{\text{obs}}$ and predictions, namely (Tarantola, 2005):

$$p(\mathbf{d}|\mathbf{m})\propto\exp\left[-\frac{1}{2}(L(\mathbf{m})-\mathbf{d}_{\text{obs}})^{\mathrm{T}}\boldsymbol{\Gamma}_{\mathrm{D}}^{-1}(L(\mathbf{m})-\mathbf{d}_{\text{obs}})\right], \tag{31}$$

where $\boldsymbol{\Gamma}_{\mathrm{D}}$ represents the data covariance, and $L(\cdot)$ denotes the nonlinear forward operator that is formulated via the L-S forward modeling operator as detailed in Eq. (14) and Eq. (15).

Bayesian frameworks for inverse scattering theory aim to infer the unknown physical parameters of the subsurface model from the observed data (Tarantola, 2005; Gouveia and Scales, 1998). The solution is described as a PDF over the model space, namely:

$$p(\mathbf{m}|\mathbf{d})=\frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{\int p(\mathbf{d}|\mathbf{m})p(\mathbf{m})\mathrm{d}\mathbf{m}}, \tag{32}$$

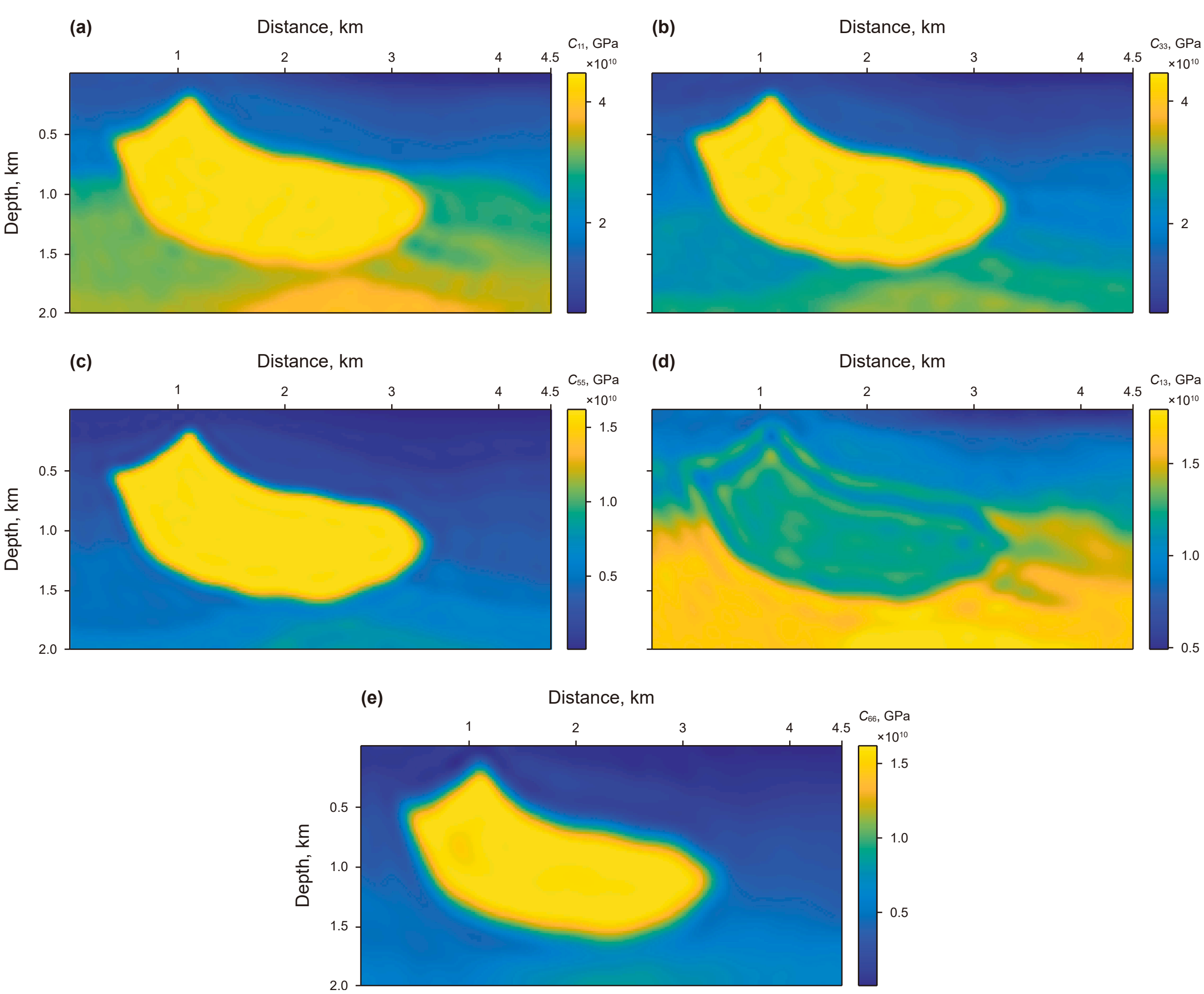

Fig. 8. The final MAP solution of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the Hess VTI subset model experiment, respectively.

where the integral term in the denominator represents the evidence (Ray et al., 2017). It can be expended as:

$$p(\mathbf{d}|\mathbf{m}) \propto \exp\left[-\frac{1}{2}\left(\left(\mathbf{m}-\mathbf{m}_{\text{pri}}\right)^{\text{T}}\boldsymbol{\Gamma}_{\text{M}}^{-1}\left(\mathbf{m}-\mathbf{m}_{\text{pri}}\right)+(L(\mathbf{m})-\mathbf{d}_{\text{obs}})^{\text{T}}\boldsymbol{\Gamma}_{\text{D}}^{-1}(L(\mathbf{m})-\mathbf{d}_{\text{obs}})\right)\right]. \tag{33}$$

By employing local, nonlinear optimization approaches, the parameter-to-observable map in Eq. (33) is typically linearized around the MAP point, leading to the PDF that is conveyed as a multidimensional Gaussian distribution that is outlined as (Bui-Thanh et al., 2013):

$$p(\mathbf{m}|\mathbf{d}) \propto \exp\left[-\frac{1}{2}(\mathbf{m}-\mathbf{m}_{\text{MAP}})^{\text{T}}\boldsymbol{\Gamma}_{\text{post}}^{-1}(\mathbf{m}-\mathbf{m}_{\text{MAP}})\right], \tag{34}$$

where the MAP solution $\mathbf{m}_{\text{MAP}}$ is the candidate model with the maximum posterior probability and is determined across all grid points for each parameter. This approximation is reasonable when the objective function is nearly linear, or in asymptotic regimes such as low-noise or large-data limits, where the posterior tends to be normality. Even when the Gaussian approximation is imperfect, the linearization remains valuable for constructing efficient sampling strategies (Bui-Thanh et al., 2013). $\boldsymbol{\Gamma}_{\text{post}}$ is calculated as $\boldsymbol{\Gamma}_{\text{post}} = \left[\mathbf{H}+\boldsymbol{\Gamma}_{\text{M}}^{-1}\right]^{-1}$, which is crucial for quantifying the uncertainties of the MAP solution and will be discussed later.

2.5. Solution to nonlinear Bayesian inverse problem via modified Levenberg–Marquardt solver

The Levenberg–Marquardt (LM) optimization method that works as an IEKF solver can be employed to implement the nonlinearity of the proposed Bayesian FWI strategy. It yields effective history-matching results by reducing the influence of the data mismatch during the early iterations (Arenas et al., 2001; Zhang and Reynolds, 2002; Skoglund et al., 2015; Eikrem et al., 2018). Let $\mathbf{d}_k$ denotes the first k data batches, a new recursive solution to $p(\mathbf{m} \mid \mathbf{d}_k)$ is available after $p(\mathbf{m} \mid \mathbf{d}_{k-1})$ is calculated by using the previous $k-1$ data batches, namely:

$$p(\mathbf{m} \mid \mathbf{d}_k) \propto p(\mathbf{d}_k \mid \mathbf{m})p(\mathbf{m} \mid \mathbf{d}_{k-1}). \tag{35}$$

Building upon Eq. (33), the MAP solution corresponding to the first k data batches can be given by solving a regularized optimization problem that minimizes the sum of the data misfit and a model penalty term, effectively balancing data fidelity with prior knowledge, as depicted in the problem of minimizing the formulation of:

$$\mathbf{m}_k = \operatorname{argmin}_{\mathbf{m}_k}\left\{\underbrace{\|(\mathbf{d}_k-\mathbf{d}_{\text{cal}})\|^2_{\Gamma_{\text{D},k}^{-1}}}_{\Phi_{\text{D}}(\mathbf{m})} + \alpha\underbrace{\left\|\mathscr{A}\left(\mathbf{m}_k-\mathbf{m}_{\text{pri}}\right)\right\|^2_{\Gamma_{\text{M},k-1}^{-1}}}_{\Phi_{\text{M}}(\mathbf{m})}\right\}, \tag{36}$$

where the symbols $\|\cdot\|^2_{\Gamma_{\text{D}}^{-1}}$ and $\|\cdot\|^2_{\Gamma_{\text{M}}^{-1}}$ represent the weighted L_2-norms scaled by different covariances. $\mathscr{A}$ is the stabilizing functional operator that is introduced to improve the direction of the model update. The regularization weight α governs the balance between data fitting and model smoothness, which effectively mediates the trade-off between achieving the best fit to the observed data and the most reasonable solution. If α is too large, the solution may lose sharpness by excluding large gradients; if too small, the inversion process becomes unstable. In this work, α is determined by the adaptive attenuation mechanism, $\alpha = \alpha_0 \cdot \xi^k$, where $\xi \in (0,1)$ is the user-specified attenuation factor and the initial regularization term $\alpha_0 = \nabla^2\Phi_{\text{D}}(\mathbf{m}_k)/\nabla^2\Phi_{\text{M}}(\mathbf{m}_k)$.

The target of finding the model that best aligns with the observed data that appeared in Eq. (36) is achieved by minimizing the data misfit function $\Phi_{\text{D}}(\mathbf{m})$, while simultaneously ensuring that the model remains as consistent as possible with prior knowledge by minimizing the model misfit term $\Phi_{\text{M}}(\mathbf{m})$. In contrast to traditional objective functions that only consider the data-fitting process, the Bayesian framework takes prior information into account and allows for the construction of an extra model misfit functional. There are various model misfit terms, including smooth, non-smooth, multi-parameter, convex, etc. Common regularization approaches, such as total variation (TV), are dedicated to producing smoother models and ensuring continuity in wavenumber convergence during inversion. In this work, we use minimum support (MS) stabilizing functional (Zhdanov, 2002; Gündogdu and Candansayar, 2018), which penalizes large variations and discontinuities within the model update, making it suitable for sharpening large-contrast structures. The definition for the MS regularization term is given by (Zhdanov, 2002; Portniaguine and Zhdanov, 1999; Gündogdu and Candansayar, 2018):

$$\Phi_{\text{M}}(\mathbf{m}_k) = \left\|\left(\mathbf{m}_k-\mathbf{m}_{\text{pri}}\right)\Big/\sqrt{\left(\mathbf{m}_k-\mathbf{m}_{\text{pri}}\right)^2+\beta^2}\right\|^2. \tag{37}$$

One can observe that the regions where model perturbation approaches zero do not contribute to model updates (Zhdanov, 2002; Portniaguine and Zhdanov, 1999). Therefore, the inversion strategies based on the MS regularization term can also be considered as focusing inversion (Shang et al., 2019). It is nonlinear with the characteristic to establish a balance between low and high wavenumbers at an early stage, thereby enhancing the resolution of anomalous edges. The gradient of the MS regularization term can be calculated as (Portniaguine and Zhdanov, 1999; Gündogdu and Candansayar, 2018):

$$\nabla\Phi_{\text{M}}(\mathbf{m}_k) = \frac{2\beta^2\left(\mathbf{m}_k-\mathbf{m}_{\text{pri}}\right)}{\left[\left(\mathbf{m}_k-\mathbf{m}_{\text{pri}}\right)^2+\beta^2\right]^2}. \tag{38}$$

The data misfit gradient is scaled by $\boldsymbol{\Gamma}_{\text{D}}$, which can be estimated as (Li et al., 2003):

$$\nabla\Phi_{\text{D}}(\mathbf{m}) = \mathscr{F}_l^{\text{T}}\boldsymbol{\Gamma}_{\text{D},k}^{-1}(\mathbf{d}_k-\mathbf{d}_{\text{cal}}), \tag{39}$$

where $\mathscr{F}$ is the nonlinear Fréchet derivative that can be calculated through Eq. (29). The Hessian represents the second derivative of the objective function, which is related to the negative gradient on the steepest descent direction. Based on the objective function defined in Eq. (36), the Hessian of the model misfit and the data misfit are derived as:

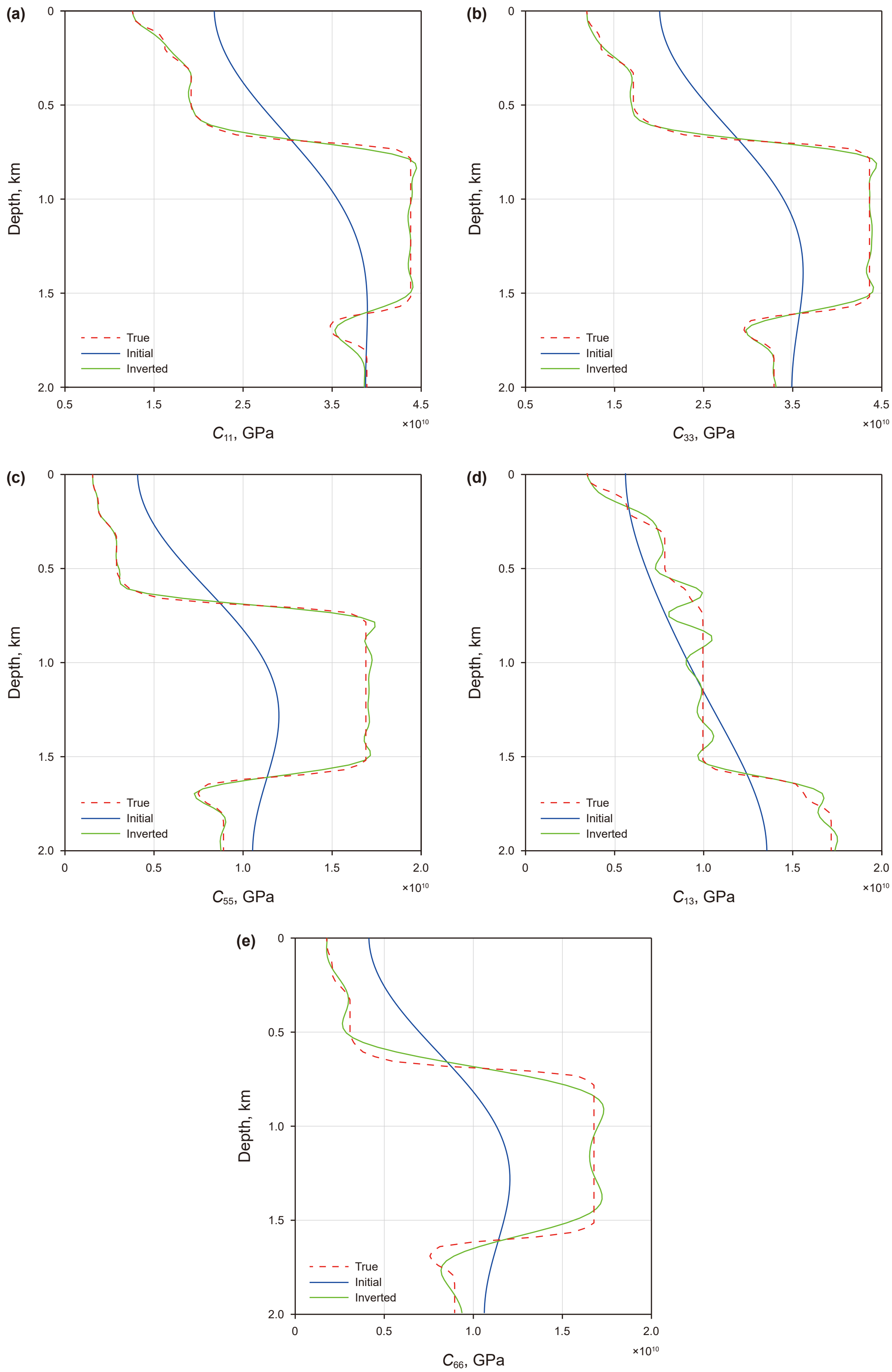

Fig. 9. Comparisons of depth profiles at $x = 2.25$ km of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ between the true model in Fig. 2 (red dashed curves), the initial model in Fig. 3 (blue curves) and the final MAP solution in Fig. 5 (green curves), respectively.

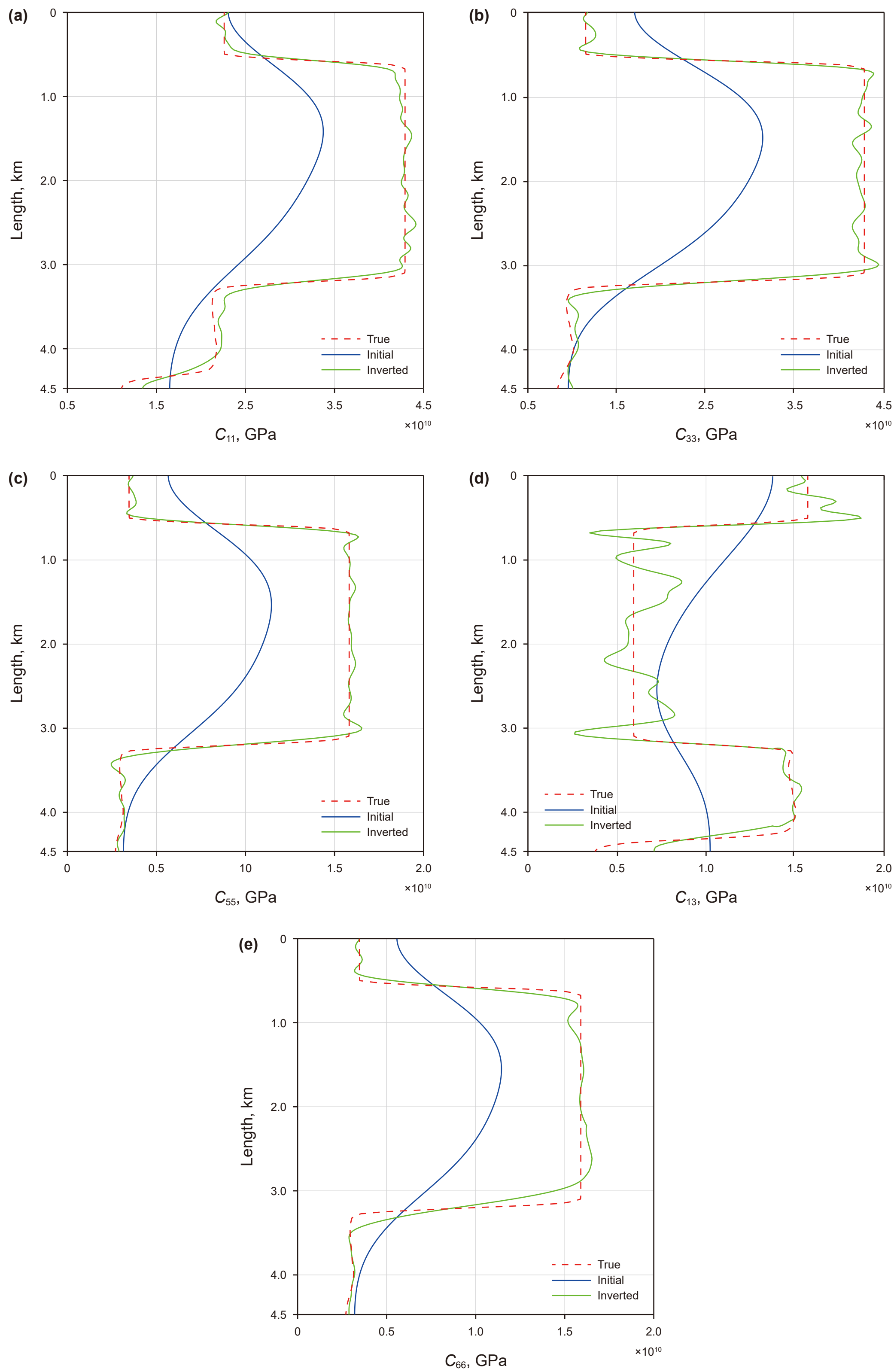

Fig. 10. Comparisons of length profiles at $z = 1$ km of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ between the true model in Fig. 2 (red dashed curves), the initial model in Fig. 3 (blue curves) and the final MAP solution in Fig. 5 (green curves), respectively.

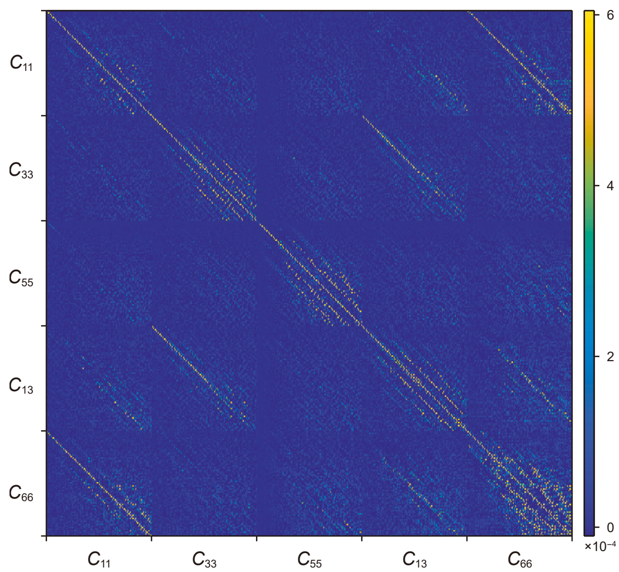

Fig. 11. The full posterior covariance matrix. From upper left to lower right, the blocks correspond to $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the Hess VTI subset model experiment, respectively.

$$\nabla^2 \Phi_\mathrm{M}(\mathbf{m}_k) = \frac{2\beta^2\left[\beta^2 - 3\left(\mathbf{m}_k - \mathbf{m}_\mathrm{pri}\right)^2\right]}{\left[\left(\mathbf{m}_k - \mathbf{m}_\mathrm{pri}\right)^2 + \beta^2\right]^3}, \tag{40}$$

where β is a small positive value that acts as a focusing factor and will be kept as constant during the iterative procedure. Although the inversion is nonlinear, the background model is assumed to be close to the true solution according to scattering theory, then we can set $\beta = 10^{-3}$ in the current work to facilitate the stability of the inversion and prevent over-sharpening of the model (Zhao et al., 2016; Nguyen et al., 2016):

$$\nabla^2\Phi_\mathrm{D}(\mathbf{m}) = \mathscr{F}_l^\mathrm{T}\,\boldsymbol{\Gamma}_{\mathrm{D},k}^{-1}\mathscr{F}_l + \frac{\partial\mathscr{F}^\mathrm{T}}{\partial\mathbf{m}_k}\boldsymbol{\Gamma}_{\mathrm{D},k}^{-1}[L(\mathbf{m}_k) - \mathbf{d}_\mathrm{obs}]. \tag{41}$$

Referring to Eqs. (38) and (39), the regularized gradient is:

$$\begin{aligned}\nabla\Phi(\mathbf{m}_k) &= \nabla\Phi_\mathrm{D}(\mathbf{m}_k) + \nabla\Phi_\mathrm{M}(\mathbf{m}_k)\\ &= \mathscr{F}^\mathrm{T}\boldsymbol{\Gamma}_\mathrm{D}^{-1}[L(\mathbf{m}_k) - \mathbf{d}_\mathrm{obs}] + \alpha\,\boldsymbol{\Gamma}_\mathrm{M}^{-1}\nabla\Phi_\mathrm{M}(\mathbf{m}_k).\end{aligned} \tag{42}$$

For computational efficiency, we simplify the Hessian by ignoring its second-order effects and using its approximation, the

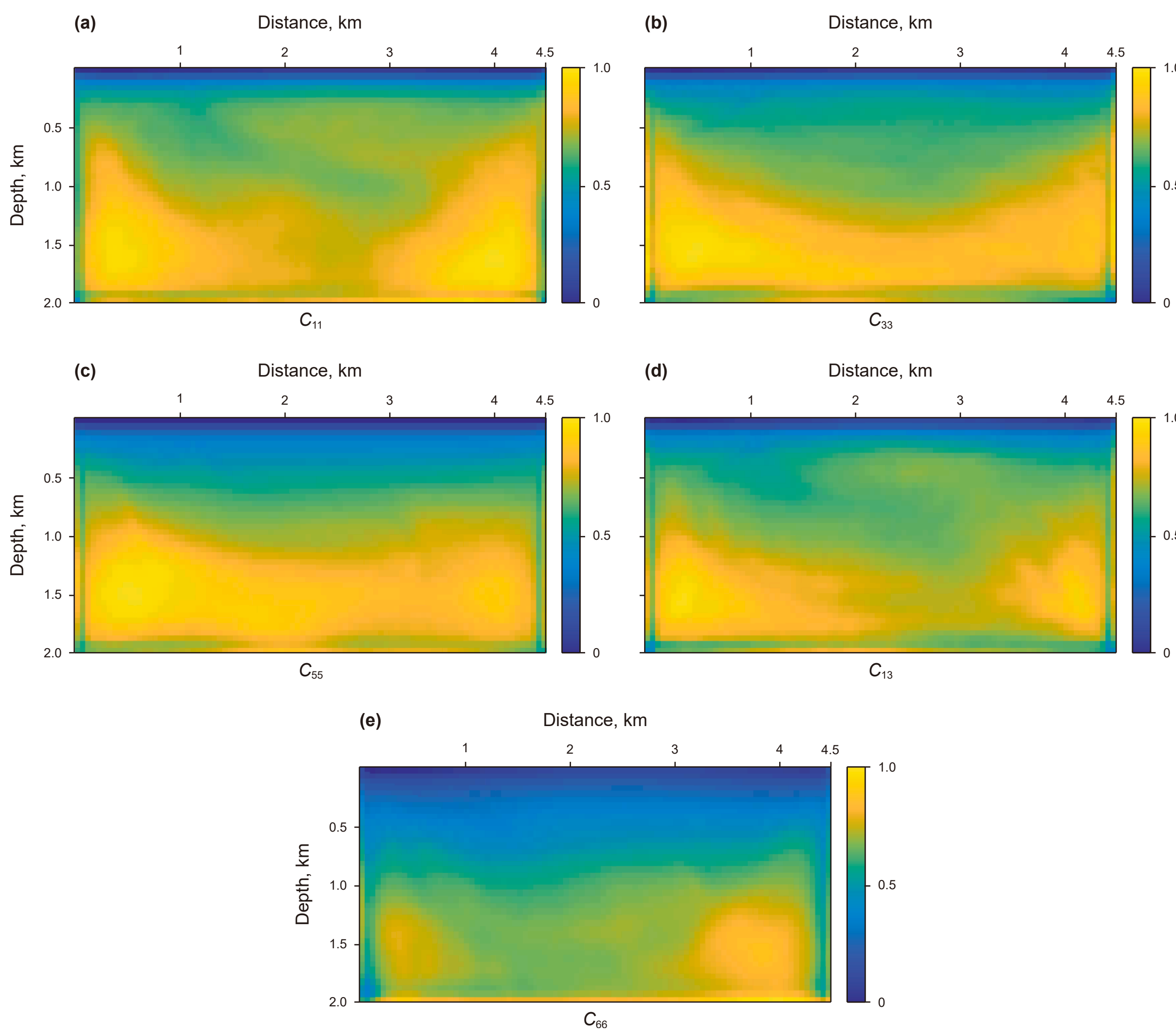

Fig. 12. The posterior standard deviations of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the Hess VTI subset model experiment, respectively.

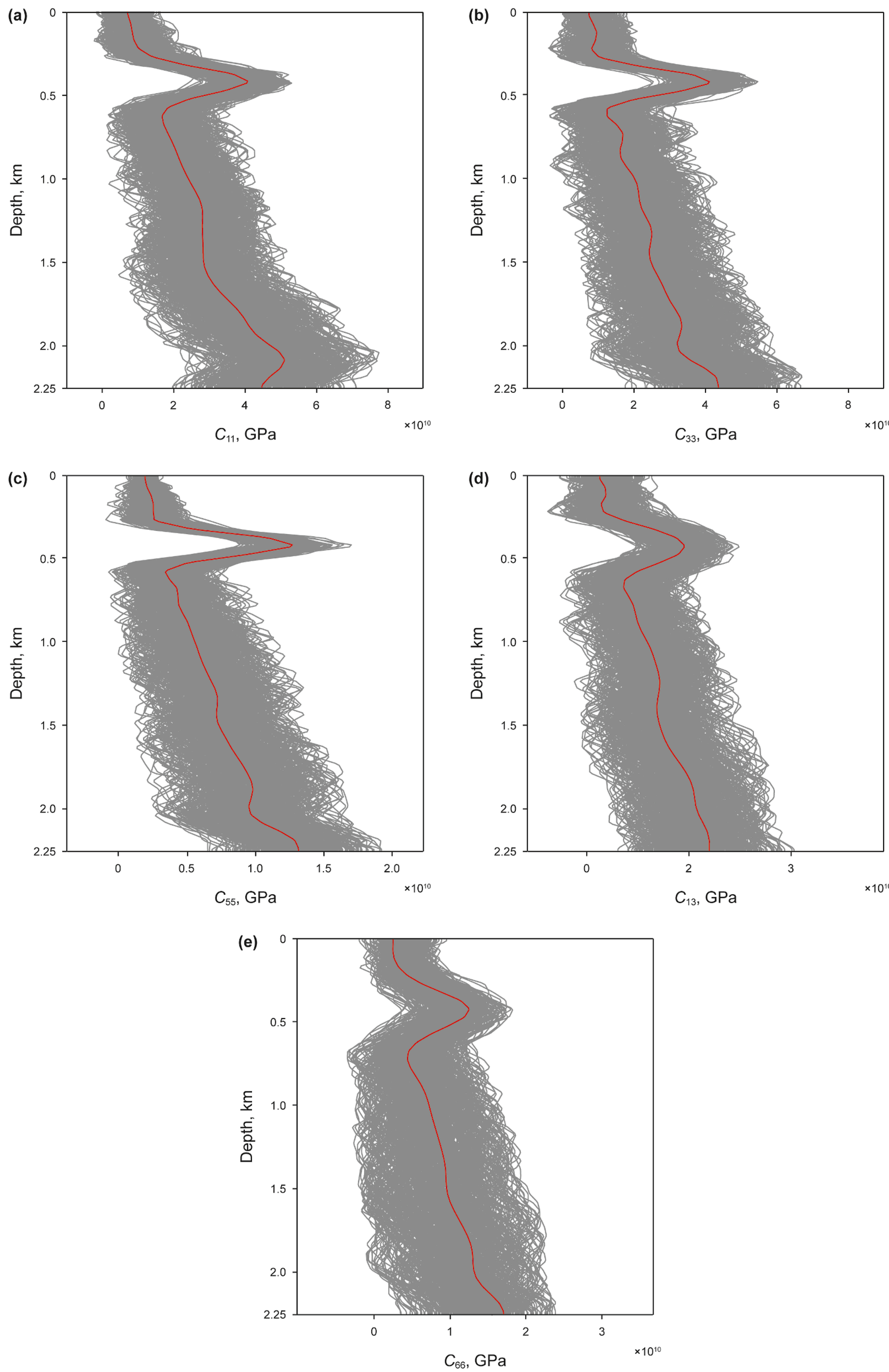

Fig. 13. Comparisons of depth profiles between the posterior distributions at $x = 2.25$ km of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the Hess VTI subset model experiment, respectively. Red curves are MAP solution, and gray profiles are random samples drawn from posterior distributions with 95% confidence level.

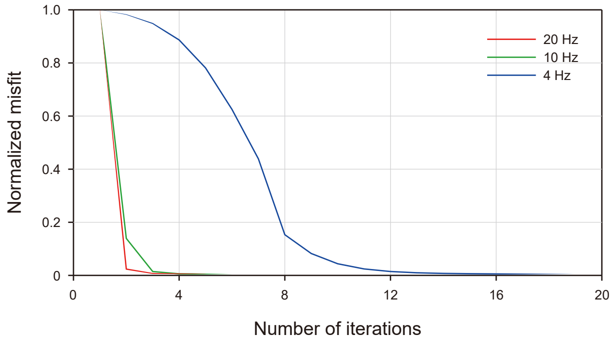

Fig. 14. Convergence rate at three different frequencies for the Hess VTI subset model experiment.

Gauss-Newton version, namely (Djebbi et al., 2017; Abubakar and Habashy, 2013):

$$\begin{aligned}\mathbf{H}(\mathbf{m}_k) &= \nabla^2\Phi_\mathrm{D}(\mathbf{m}_k) + \nabla^2\Phi_\mathrm{M}(\mathbf{m}_k)\\ &\approx \mathscr{F}^\mathrm{T}\mathbf{\Gamma}_{\mathrm{D},k}^{-1}\mathscr{F} + \alpha\,\mathbf{\Gamma}_{\mathrm{M},k}^{-1}\nabla^2\Phi_\mathrm{M}(\mathbf{m}_k)\,.\end{aligned} \tag{43}$$

Bayesian approaches for deterministic frequency-domain FWI substantially resemble traditional workflows by using a similar sequential strategy for frequency data batches (Skoglund et al., 2015). For fixed k batch data, assuming the l-th MAP solution through IEKF is represented by $\mathbf{m}_l$, the well-established expression for determining the updated model is given by (Eikrem et al., 2018):

$$\mathbf{m}_{l+1} = \mathbf{m}_l + (\mathbf{H}(\mathbf{m}_l) + \lambda\mathbf{I})^{-1} \times [\nabla\Phi_\mathrm{D}(\mathbf{m}_l) + \nabla\Phi_\mathrm{M}(\mathbf{m}_l)], \tag{44}$$

where λ represents the damping factor that needs to be adjusted at each iteration. To select an optimal value for λ, various strategies are available, such as the widely used L-curve method (Ostebee, 1998), but come with computational overhead. We adopt a self-adaptive adaptive method to adjust λ (Jakobsen et al., 2020; Zhdanov, 2002). To guarantee the stability of the inversion, a big value is initially selected, producing an update similar to the steepest descent method. Following that, a cooling scheme is implemented to gradually decrease λ in a manner akin to the Gauss–Newton method: Within the first five iterations, the value is reduced by 0.9, and subsequently by 0.1. The initial value is set to be the mean of the diagonal Hessian square root, ensuring that the model inversion begins with a sharp, well-defined solution. This strategy facilitates optimal sharpness of the inverted model early on and the gradual minimization of the data misfit as the process continues, thereby facilitating the gradual update of the detailed features of the model. Once convergence is achieved, the update of the PCM is:

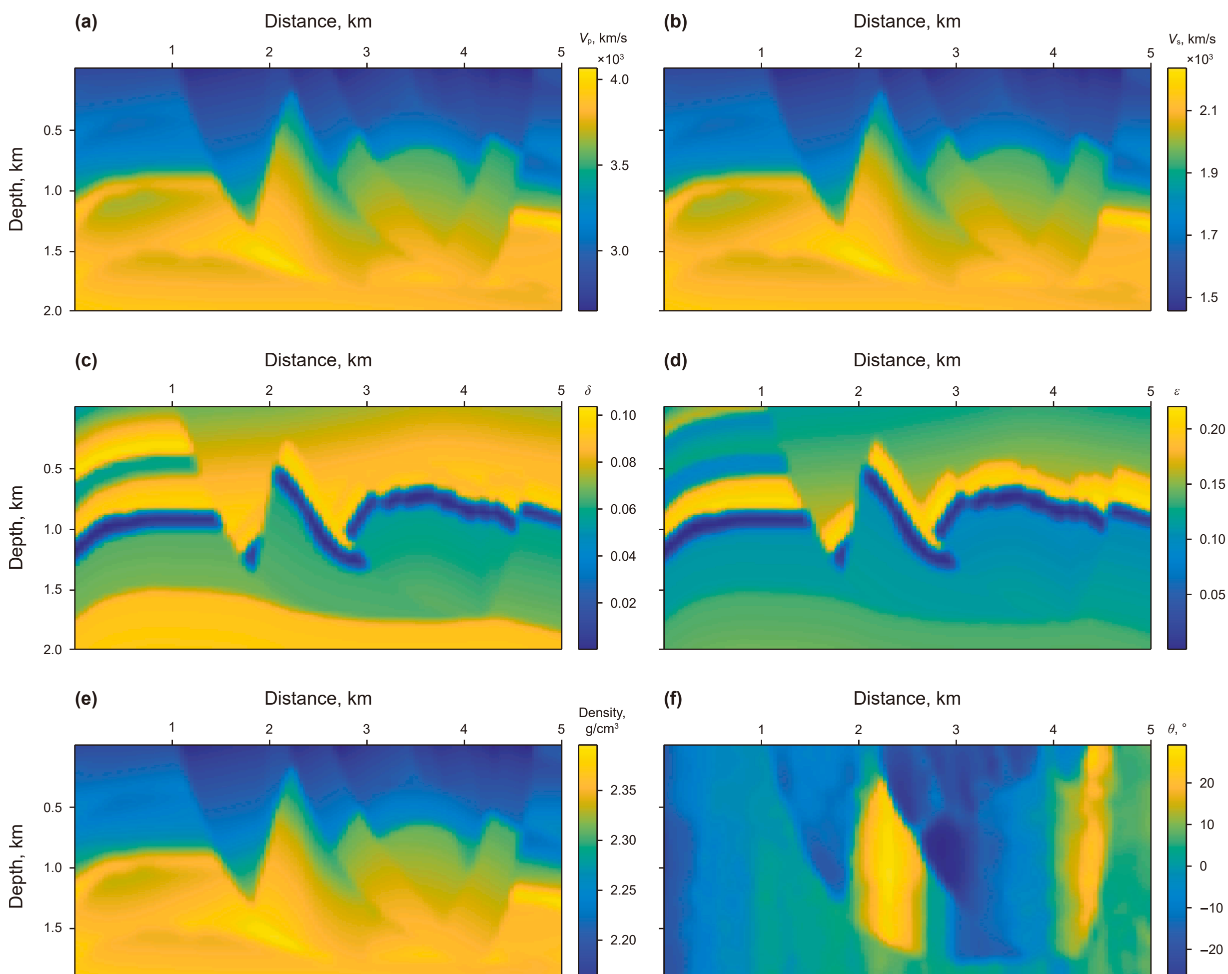

Fig. 15. BP TTI subset model described by anisotropy parameters of **(a)** P-wave velocity, **(b)** S-wave velocity, Thomsen parameters **(c)** δ, **(d)** ε, **(e)** density and **(f)** tilted angle.

$$\boldsymbol{\Gamma}_{\mathrm{M},k}^{-1} = \mathbf{H} + \boldsymbol{\Gamma}_{\mathrm{M},k-1}^{-1} = \mathscr{F}_l^{\mathrm{T}} \boldsymbol{\Gamma}_{\mathrm{D},k}^{-1} \mathscr{F}_l + \boldsymbol{\Gamma}_{\mathrm{M},k-1}^{-1}. \tag{45}$$

Eqs. (44) and (45) provide the Levenberg–Marquardt optimization solution according to the IEKF (Gu and Oliver, 2007; Skoglund et al., 2015). So far, the Bayesian FWI method based on the scattering IE method for multiparameter anisotropic media is complete. Starting from lower frequencies, we utilize the DWT preconditioned GMRES method to solve the scattering IEs (see Eqs. (14) and (15)) to obtain the particle displacement-strain fields. The multiparameter Fréchet derivative has been explicitly formed in Eq. (29) to linearize the relationship between the model and the data. Through the objective function with the MS regularization term, we derive the regularized Hessian and gradient, then update the elastic moduli and the PCM simulations via Eqs. (44) and (45).

2.6. Low rank approximation for uncertainty quantification

The data covariance influences the inversion process in conjunction with the Hessian, while the model covariance serves as a measurement of the uncertainty of each iterated model. By denoting the covariance at the $(k-1)$ th and k th iterations in Eq. (45) as the prior covariance $\boldsymbol{\Gamma}_{\mathrm{prior}}$ and posterior covariance $\boldsymbol{\Gamma}_{\mathrm{post}}$, respectively, we can conclude that in the absence of any information gain from the data, the PCM is infinitely close to the model covariance. Therefore, when the inversion converges, Eq. (45) can be expressed in a practical way as (Tarantola, 2005; Liu and Peter, 2019):

$$\boldsymbol{\Gamma}_{\mathrm{post}} = \boldsymbol{\Gamma}_{\mathrm{prior}}^{1/2} \left(\boldsymbol{\Gamma}_{\mathrm{prior}}^{1/2} \mathbf{H} \boldsymbol{\Gamma}_{\mathrm{prior}}^{1/2} + \mathbf{I} \right)^{-1} \boldsymbol{\Gamma}_{\mathrm{prior}}^{1/2}. \tag{46}$$

The posterior covariance $\boldsymbol{\Gamma}_{\mathrm{post}}$ is typically of immense size, with dimensions on the order of $(N_m \times N)^2$, where N_m is the number of parameters and N denotes the size of the individual parameter, further enhancements are necessary when dealing with large-scale geophysical applications. The randomized SVD method can approximate and perform spectral analysis to factorize large matrices by employing a set of random vectors to investigate the matrix and extract its eigenvalues and eigenvectors (Liberty et al., 2007; Bui-Thanh et al., 2013). It is particularly valid for low-rank matrices. The prior-preconditioned Hessian $\tilde{\mathbf{H}}$ can be approximated with controllable accuracy in a practical way:

$$\tilde{\mathbf{H}}^{-1} = \left(\boldsymbol{\Gamma}_{\mathrm{prior}}^{1/2} \mathbf{H} \boldsymbol{\Gamma}_{\mathrm{prior}}^{1/2} \right)^{-1} \approx \left(\mathbf{V}_r \boldsymbol{\Lambda}_r \mathbf{V}_r^{\mathrm{T}} + \mathbf{I} \right)^{-1}, \tag{47}$$

where $\boldsymbol{\Lambda}_r$ denotes an eigenvector matrix with the diagonal elements corresponding to the r-th singular values of $\tilde{\mathbf{H}}$, while $\mathbf{V}_r$ refers to the matrix of the corresponding eigenvalues. Single-pass randomized SVD allows for efficient eigen decomposition of $\tilde{\mathbf{H}}$, the details of it are summarized in Algorithm 2. The square root of the

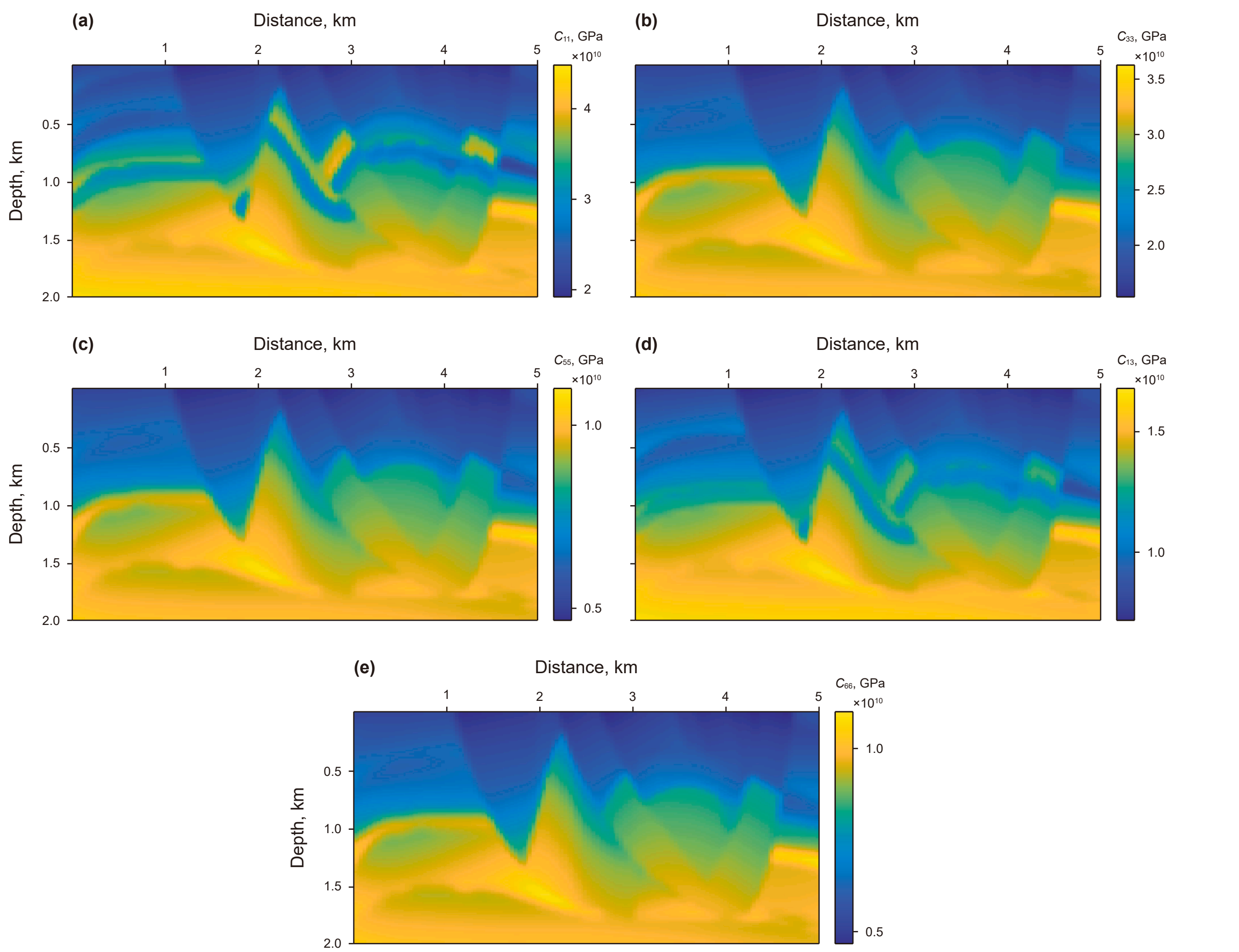

Fig. 16. BP TTI subset model described by five independent Voigt elastic moduli of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$, respectively.

PCM $\boldsymbol{\Gamma}_{\text{post}}^{1/2}$, which can be calculated as (Fichtner and van Leeuwen, 2015; Halko et al., 2011):

$$\boldsymbol{\Gamma}_{\text{post}}^{1/2} = \mathbf{V}_r \mathbf{P}_r^{1/2} \mathbf{V}_r^{\text{T}} \boldsymbol{\Gamma}_{\text{prior}}^{1/2}, \tag{48}$$

where $\mathbf{P}_r = \text{diag}\left(1/\sqrt{\lambda_1 + 1} - 1, \ldots, 1/\sqrt{\lambda_r + 1} - 1\right)$ and λ_r is the singular values of $\tilde{\mathbf{H}}$. Thus, by visually comparing the prior and posterior random samplings, the uncertainty can be evaluated on both the prior and posterior models. After performing the eigen decomposition of $\tilde{\mathbf{H}}$, a Gaussian random sampling can be drawn on the prior and posterior distributions respectively, as (Fichtner and van Leeuwen, 2015; Zhu et al., 2016; Bui-Thanh et al., 2013):

$$\mathbf{m}_{\text{prior}} = \mathbf{m}_0 + \boldsymbol{\Gamma}_{\text{prior}}^{1/2} \mathbf{R}, \tag{49}$$

$$\mathbf{m}_{\text{post}} = \mathbf{m}_{\text{MAP}} + \boldsymbol{\Gamma}_{\text{post}}^{1/2} \mathbf{R}, \tag{50}$$

where $\mathbf{m}_0$ is the initial model, $\mathbf{R}$ in both Eqs. (48) and (49) denote a random sampling vector from a zero-mean normal distribution, with the covariance set to identity.

Algorithm 2. Single-pass randomized singular value decomposition of the prior-preconditioned Hessian (modified according to Algorithm 4.4 of Halko et al., 2011)

Input: A random matrix $\mathbf{X}$ of dimension $N \times r$
Output: the preconditioned Hessian $\tilde{\mathbf{H}}$

1. Sampling $\mathbf{H}$ with $\mathbf{X}$: $\mathbf{Y} = \mathbf{HX}$
2. Perform QR decomposition on $\mathbf{Y}$: $\mathbf{Y} = \mathbf{QR}$
3. Calculate $\mathbf{A}$ via a linear system: $\mathbf{A}\left(\mathbf{Q}^{\text{T}}\mathbf{X}\right) = \mathbf{Q}^{\text{T}}\mathbf{Y}$
4. Perform SVD on $\mathbf{A}$: $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\text{T}}$
5. Calculate the singular vector $\mathbf{V}$: $\mathbf{V} = \mathbf{QU}$
6. Output: $\tilde{\mathbf{H}} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\text{T}}$

3. Numerical examples

Two classic synthetic tests are implemented to show the efficacy of our method. The applications are developed on the MATLAB platform using an in-house computational facility with 1 node, which has 32 cores with 128 GB memory. After adding the noise into the original data, the signal-to-noise ratio S_N is defined as (Tromp et al., 2004):

$$S_N = E\left(\|e\|^2\right)\left(\|\mathbf{d}\| \cdot e / (\mathbf{d}_{\text{noise}} - \mathbf{d})\|e\|\right)^2, \tag{51}$$

where E is the expectation, $\mathbf{d}$ is the data and $\mathbf{d}_{\text{noise}}$ is the data with the noise. The stopping criteria are defined by:

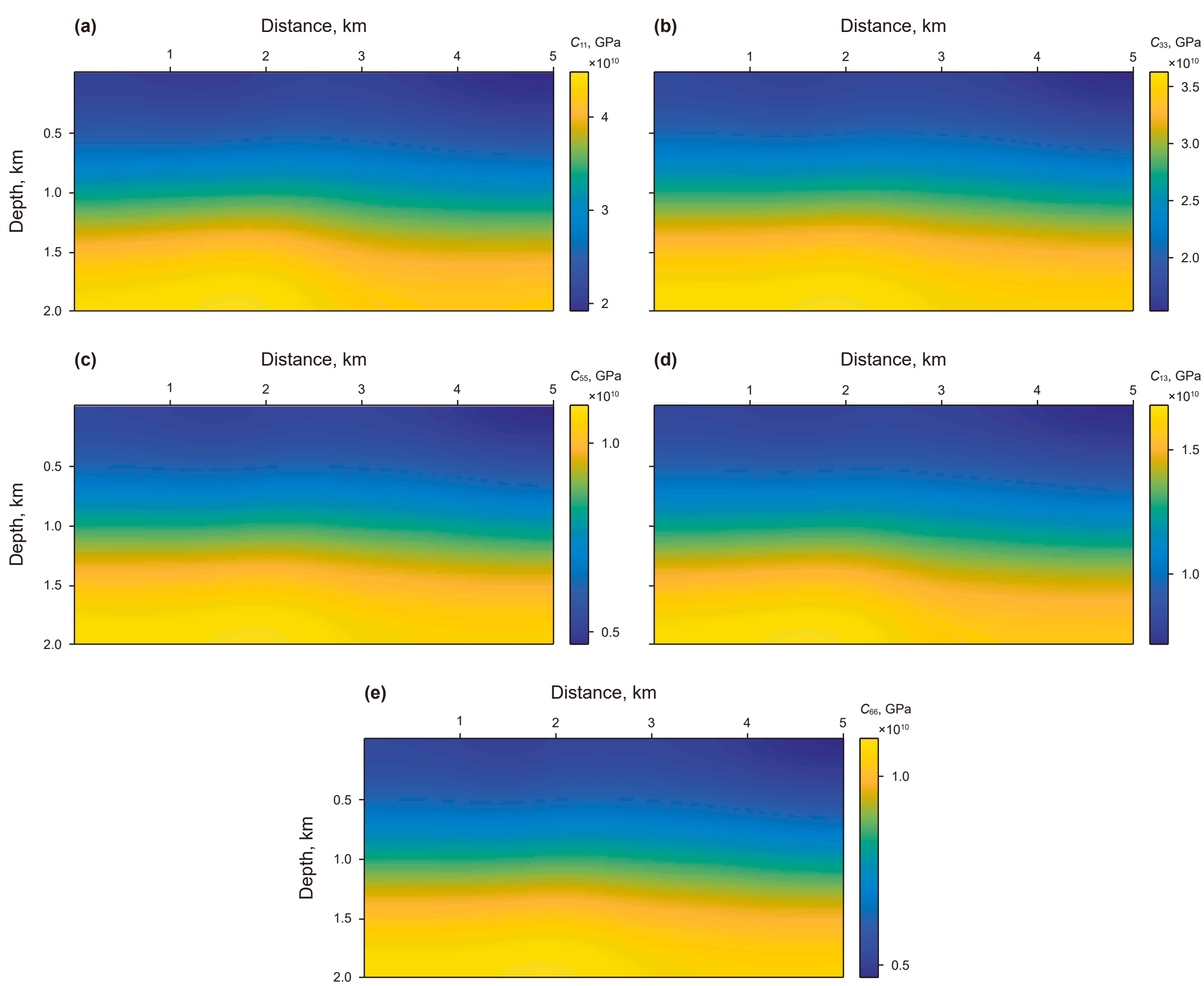

Fig. 17. Initial elastic moduli of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the BP TTI subset model experiment, respectively.

$$rms_{\mathrm{obj}}(\mathbf{m}_k) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\Phi(\mathbf{m}_k) - \Phi\left(\mathbf{m}_{\mathrm{pri}}\right)\right)^2}. \tag{52}$$

3.1. Hess VTI model

We implement the first synthetic test of a classic heterogeneous geophysical model, the SEG/Hess VTI model. The entire model consists of two parts: the irregularly shaped salt body and the surrounding sedimentary layers. The salt located in the middle of the model can be treated as a large-contrast geological anomaly. Since most of the energy is reflected by the near-surface salt domes, only a small portion can reach deeper layers. These factors greatly contribute to the heterogeneity and anisotropy of the model, resulting in complicated wavefield propagation and attenuation, thereby imposing high demands on the inversion algorithm. The original model is modified to test the proposed elastic anisotropic Bayesian FWI algorithm, with P-wave and S-wave velocities, two Thomsen anisotropy parameters and density as depicted in Fig. 1.

We formulate the inverse problem in terms of Voigt parameters. It allows us to construct a stiffness matrix including all five elastic moduli while updating the model and the wavefield. Each modified elastic modulus is modeled with 4.5 km laterally and 2 km vertically, with a grid spacing of 25 m. We take the perturbated elastic moduli that are being normalized to the corresponding properties of a homogeneous isotropic reference medium as the model parameters for the inversion, which are displayed in Fig. 2.

The setup for the acquisition configuration follows a fixed distribution survey, comprising 45 shots and 90 receivers uniformly distributed along the surface. A Ricker wavelet with a dominant frequency of 7.5 Hz is utilized to generate the wavefield, producing three-component datasets recorded in complex values. White noise is artificially introduced to interfere with the data conditions, with a signal-to-noise ratio (SNR) of 4. The temporal sampling interval is set at 0.002 s with a total recording duration of 4 s. By applying a Gaussian smoother with a broad threshold to the true model, we mitigate distorted travel times and remove the discontinuous, yielding the elastic moduli as the initial model (see Fig. 3). The inversion strategy is conducted using 17 discrete frequencies, evenly spread over a range of 4–20 Hz, with the inversion result of the prior frequency acting as the initial model for the subsequent frequency loop. The inversion result at the first frequency (4 Hz) is illustrated in Fig. 4.

Fig. 5 depicts the Fréchet derivative, or sensitivity kernel, for one source-receiver geometry at 4 Hz during the inversion. Fig. 6 is the counterpart at a higher frequency (20 Hz). One can observe that the region shaped as a 'banana-donut' contains the majority of the energy. This area is most affected by the wavefield, also referred to as the first Fresnel zone (Woodward, 1992), with the

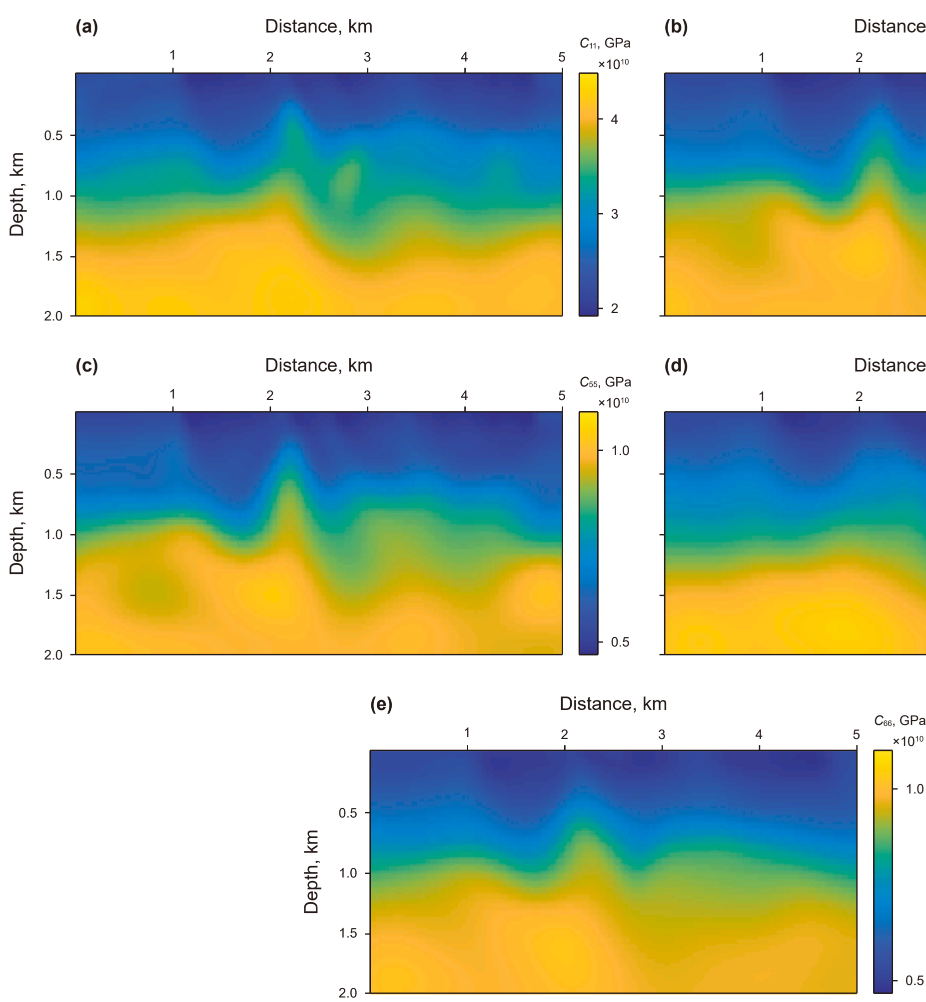

Fig. 18. The inverted results of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ at the first frequency (4 Hz) for the BP TTI subset model experiment, respectively.

main contribution from the direct waves. As the frequency increases, the update shifts progressively towards the long-wavelength components, and the sensitivity kernel becomes more complex. The energy reaching the deeper region gradually weakens, indicating that capturing the detailed structure within the deep parts of the model may be more challenging. Nevertheless, the incorporation of the Hessian information can be introduced to address this issue. By constructing the prior-preconditioned Hessian and refining it through the IEKF method iteratively, noticeable updates in deep layers are observable in the gradient, as shown in Fig. 7. The final MAP solution is illustrated in Fig. 8, and we can observe that it is significantly updated compared to the initial model, where the salt dome and the surrounding sediments are distinctly outlined.

Figs. 9 and 10 show the vertical and horizontal profiles of the five elastic moduli in the initial model, the true model, and the MAP estimation that is located in the middle, providing a more evident comparison of the inverted results as shown in Fig. 8. It is clear that even with a sufficiently smooth initial model and low signal-to-noise ratio, our approach still shows satisfactory accuracy in retrieving all five elastic parameters, especially C_{11}, C_{33}, and C_{55}. Although the depth/length tendencies of reconstructed C_{13} and C_{66} stay aligned with the true model, the resolution and accuracy in the deeper parts still need further enhancement.

Fig. 11 depicts the full PCM, which conveys the information about cross-talk correlations and the accuracy of the inverted results. The PCM can be regarded as a set of 5 × 5 block matrices, with elements arranged from top to bottom and left to right representing C_{11},C_{33},C_{55},C_{13},C_{66}, respectively. The matrices along the main diagonal denote the autocorrelation characteristics of the five independent elastic constants, derived from the second derivative of the objective function, reflecting the uncertainty or variance of each parameter. The non-zero elements in the off-diagonal blocks indicate the cross-correlation between different parameters, with their magnitude reflecting the degree of coupling among the parameters. The square root of the diagonal elements of the PCM signifies the uncertainty associated with each individual elastic constant, as presented in Fig. 12, namely the posterior standard deviation. It quantifies the precision of the parameter estimates, and its value gradually rises as the depth of the model increases. Five hundred random samples from the posterior distribution are attempted to determine the range of possible solutions that fit the data. Then we plotted the uncertainty map of the predicted depth profile at 95% confidence level for the five

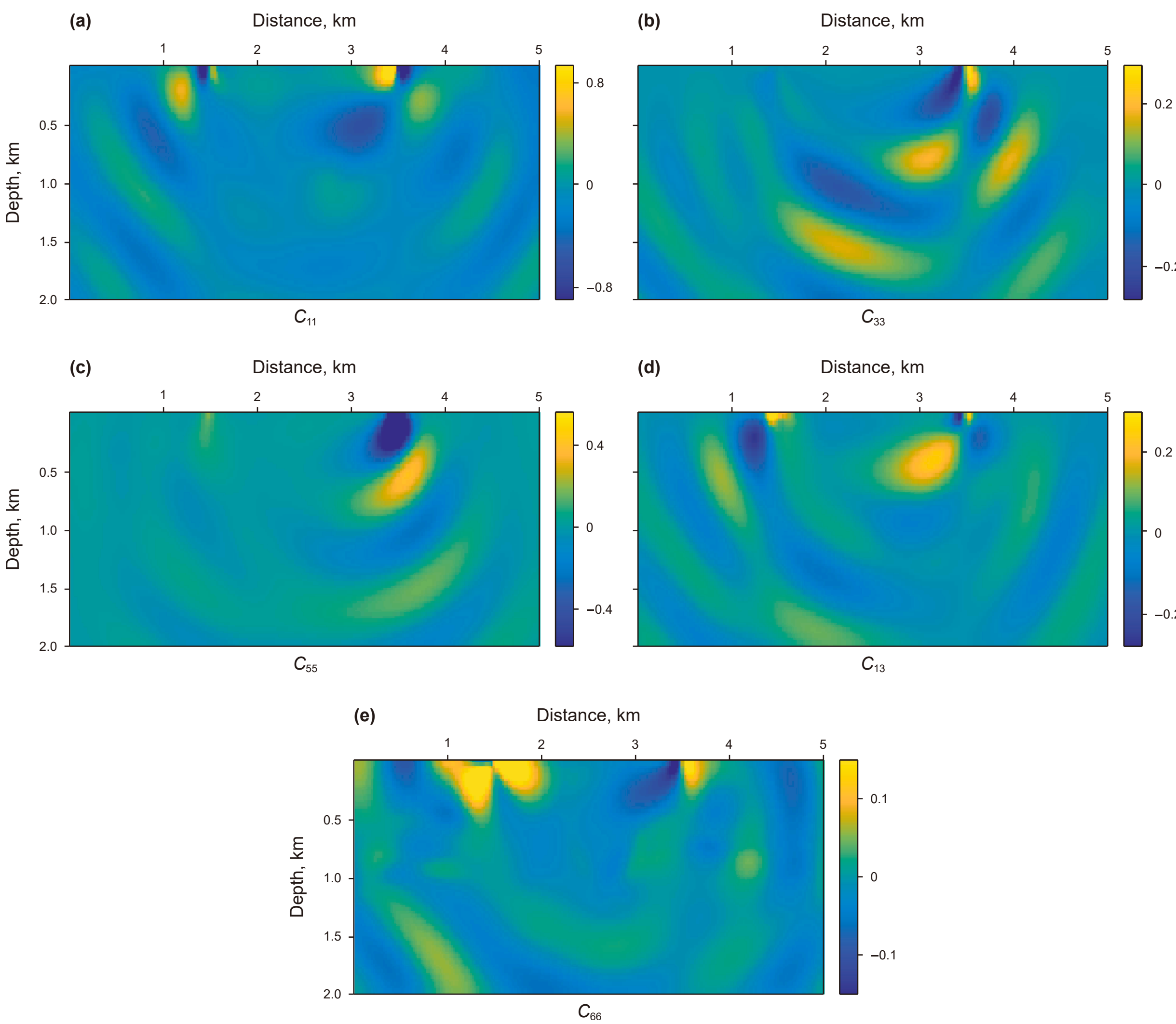

Fig. 19. For a source-receiver configuration with the source at $x = 1.5$ km and the receiver at $x = 3.5$ km, the Fréchet derivatives of **(a)**–**(e)** C_{11}, C_{33}, C_{55}, C_{13}, C_{66} at the first frequency (4 Hz) for the BP TTI subset model experiment, respectively.

parameters as shown in Fig. 13 and compared them with different estimate of the elastic moduli obtained by solving the nonlinear inverse problem via the proposed method. Fig. 14 reports the convergence rate at different frequencies. One can observe that, within a finite number of iterations, our proposed method indicates an almost exponential decline. Furthermore, the higher the frequency, the fewer times the objective function needs to converge to a specific stopping criteria.

3.2. A subset of BP TTI model

Another synthetic example conducted in this section is developed from the BP TTI model. Due to its large scale, the model underwent resampling and rescaling, focusing only on a limited subset of the model. The modified model parameters are shown in Fig. 15, which has a complicated overlying stratum, forming classic thrust structures. The faults and folds within the model cause strong lateral variations, further increasing the heterogeneity and anisotropy. Notably, unlike the previous one, this experiment includes a new parameter, that is, the tilt angle, which complicates the reconstruction of the elastic moduli compared to VTI media. Besides, the model contains diverse rock types with significant physical property variations. Consequently, seismic waves experience multiple reflections and refractions, thereby increasing the complexity of wavefield modeling and leading to more complex seismic data. All of the factors above increase the nonlinearity of the inverse scattering problem and the risk of getting trapped in a local minimum, making it more challenging to reconstruct the model.

Referring to the previous example, we also consider the transformed elastic modulus as the representation of model parameters required to perform the inversion, as shown in Fig. 16. All independent elastic moduli extend 5 km horizontally and 2 km vertically, with a grid spacing of 25 m. The acquisition configuration and the wavelet are identical to the previous test, with 100 receivers and 50 sources employed to generate three-component complex-valued data with 2001 sampling points over a total sampling time of 4 s. The initial elastic moduli for the inversion in Fig. 17 are derived from the true model by utilizing the same smoother. Artificial white noise is introduced to the data, maintaining the signal-to-noise ratio at 4. We slightly broaden the frequency band to 4–24 Hz and execute inversion at a single even

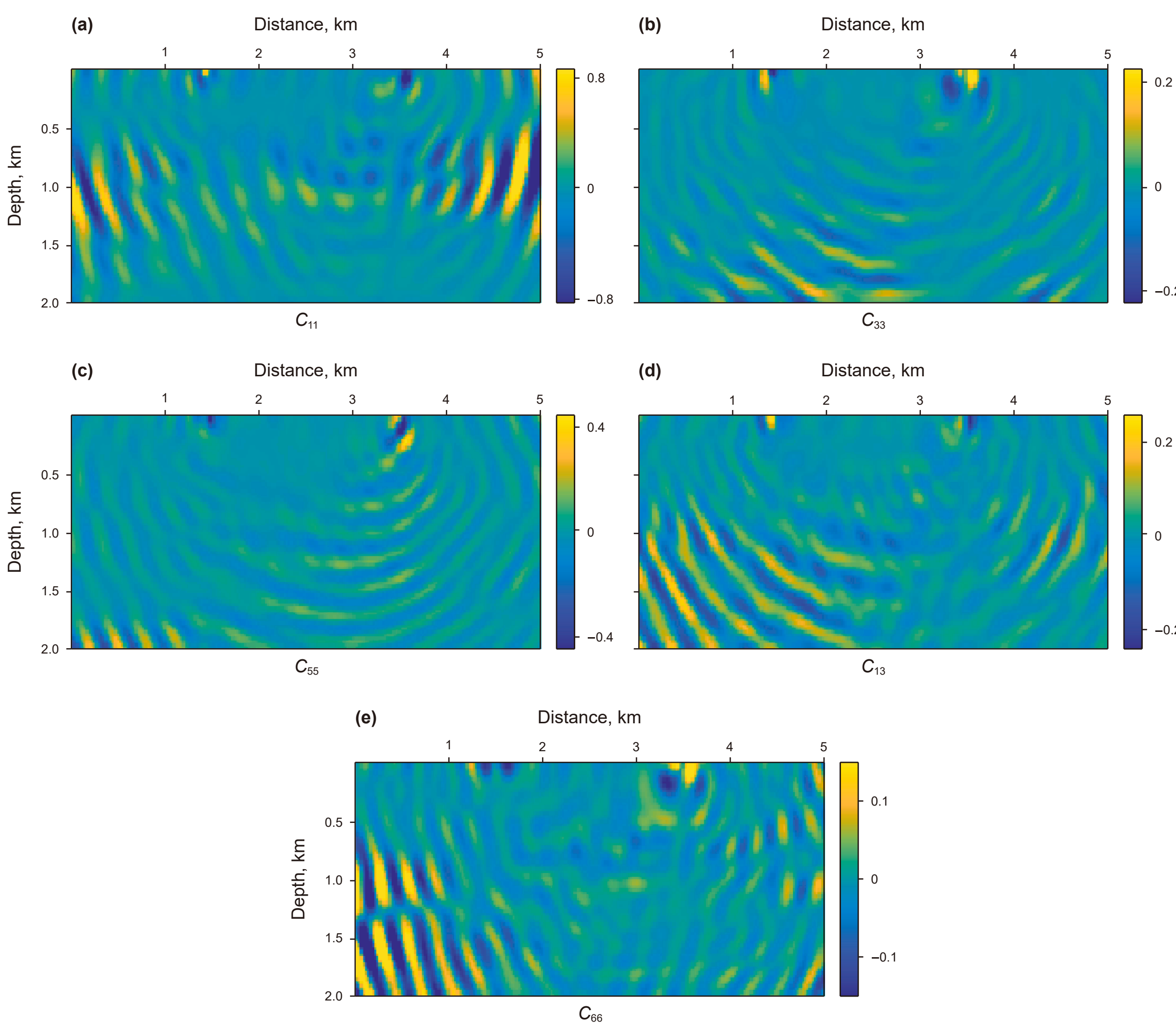

Fig. 20. For a source-receiver configuration with the source at $x = 1.5$ km and the receiver at $x = 3.5$ km, the Fréchet derivatives of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ at 20 Hz for the BP TTI subset model experiment, respectively.

frequency per iteration successively. The inverted result at the first frequency (4 Hz) is shown in Fig. 18, where the large-scale features of the model have manifested.

Referring to the previous example, we also consider the transformed elastic modulus as the representation of model parameters to perform the inversion, as shown in Fig. 16. All independent elastic moduli extend 5 km horizontally and 2 km vertically, with a grid spacing of 25 m. The acquisition configuration and the wavelet are identical to the previous test, with 100 receivers and 50 sources employed to generate three-component complex-valued data with 2001 sampling points over a total sampling time of 4 s. The initial elastic moduli for the inversion in Fig. 17 are derived from the true model by utilizing the same smoother. Artificial white noise is introduced to the data, maintaining the signal-to-noise ratio at 4. We slightly broaden the frequency band to 4–24 Hz and execute inversion at a single even frequency per iteration successively. The inverted result at the first frequency (4 Hz) is shown in Fig. 18, where the large-scale features of the model have manifested.

Figs. 19 and 20 show the Fréchet derivatives in the frequency domain at 4 Hz and 24 Hz, respectively, for a source-receiver pair with the source located at $x = 1.5$ km and the receiver at $x = 3.5$ km. Most of the energy is concentrated in the banana-donut area, with a steady decrease in energy as depth increases. To address this issue, the prior preconditioned Hessian matrix can be applied to compensate for the illumination, yielding the gradients depicted in Fig. 21. The improved resolution in the deeper parts of the model illustrates the effect of long-wavelength components. A comparative experiment is essential to make the experiments more convincing and illustrative.

Using a traditional Newton-type optimization approach that focuses solely on the least-squares data misfit objective function, we obtain the results shown in Fig. 22. To address the function of the MS stabilizing functional, we conduct another experiment using the TV (Total Variation) regularization. By applying the same dataset, with identical settings for the maximum number of iterations and residual thresholds, the inverted results for different elastic moduli obtained by TV regularization are shown in Fig. 23.

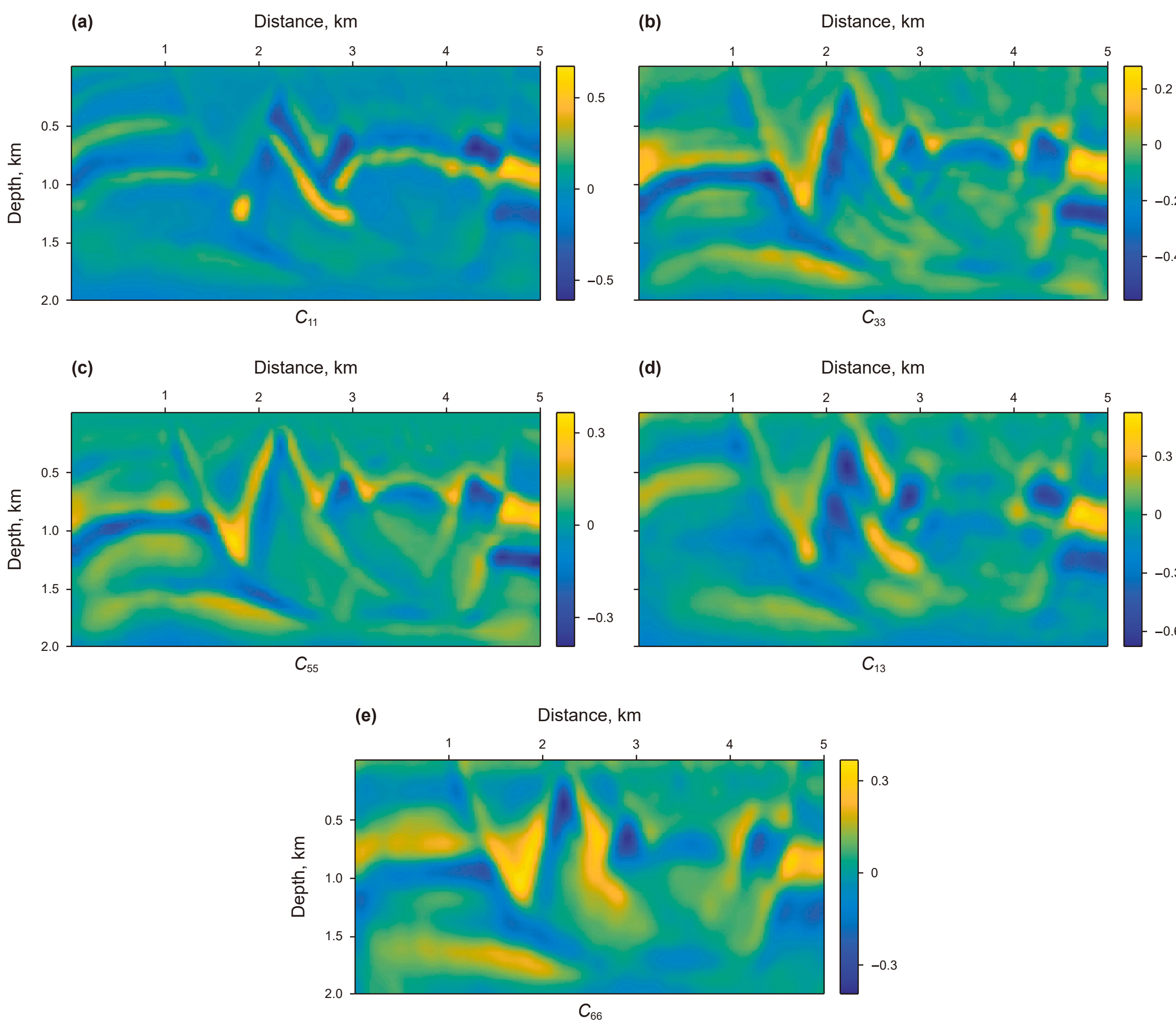

Fig. 21. Frequency-domain normalized gradient at 20 Hz of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the BP TTI subset model experiment, respectively.

With the incorporation of higher frequency datasets, small-scale structures gradually revealed, and the final MAP solutions are displayed in Fig. 24. We can observe that the regularization term is helpful to acquire the high wavenumber components that are necessary for the inversion at an earlier stage, thereby improving the accuracy of the inversion results, clear improvements can be observed in the region highlighted by the red arrows in Fig. 24. Figs. 25 and 26 show the profiles at $x = 2.5$ km and $z = 1$ km from different models for the five elastic moduli, respectively, providing a more intuitive comparison. It is evident that the proposed method can capture both the trend and magnitude of parameter variations, closely matching the true answer no matter how deep the analysis goes. Conventional approaches, on the other hand, can only retrieve the overall geometry of the model, proving the ability of our algorithm in addressing both high wavenumber components and large-scale stratigraphies.

Fig. 27 shows the full posterior covariance, where the narrow energy concentration along the main diagonal implies a low variance of the inversion results, suggesting high accuracy of the solutions. As shown in Fig. 28, the posterior standard deviation, along with the 95% confidence posterior probability distribution with 500 random samplings in Fig. 29, indicates that the uncertainty is lower in the shallow while increasing slightly in the deeper parts. The effectiveness of the inversion methodology reflects not only on the image quality, but also on the computational ability. The convergence curves of the objective function at three different frequencies displayed in Fig. 30 demonstrate that the inclusion of extra parameters does not impose a computational burden. Specifically, during the inversion process, a total number of 88 iterations were conducted across 11 discrete frequencies, compared to 206 iterations required by the conventional method. Moreover, the comparative analysis with conventional approaches further conveys that, despite its probabilistic nature, the Bayesian approach allows us to iterate fewer times and ensures better robustness and accuracy. Iterative strategy and low-rank approximated technique make it possible for practical problems. The average computation time per iteration for a single frequency via the proposed method is approximately 35 min. By employing the

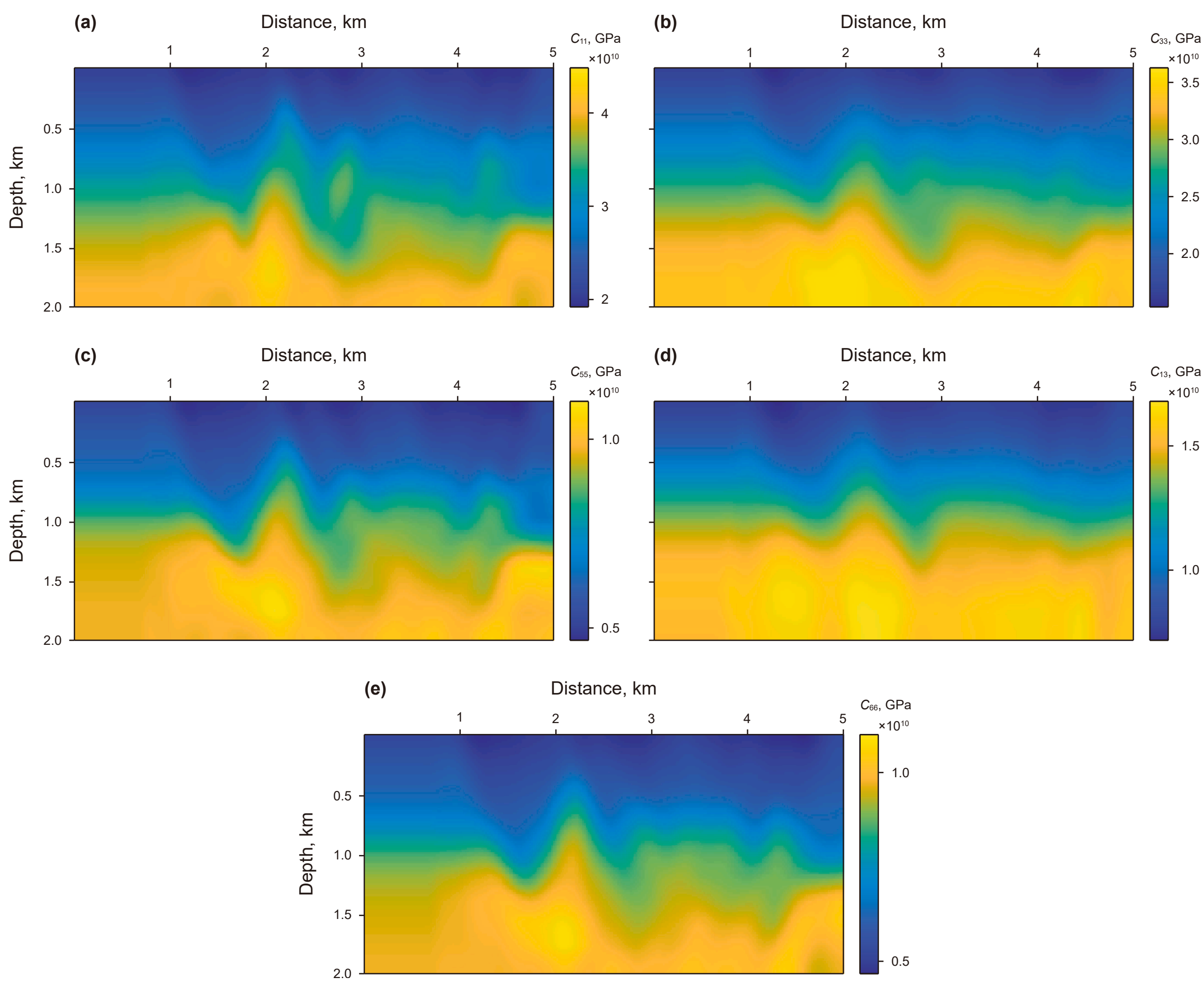

Fig. 22. Inverted results of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the BP TTI subset model experiment obtained by the traditional approach.

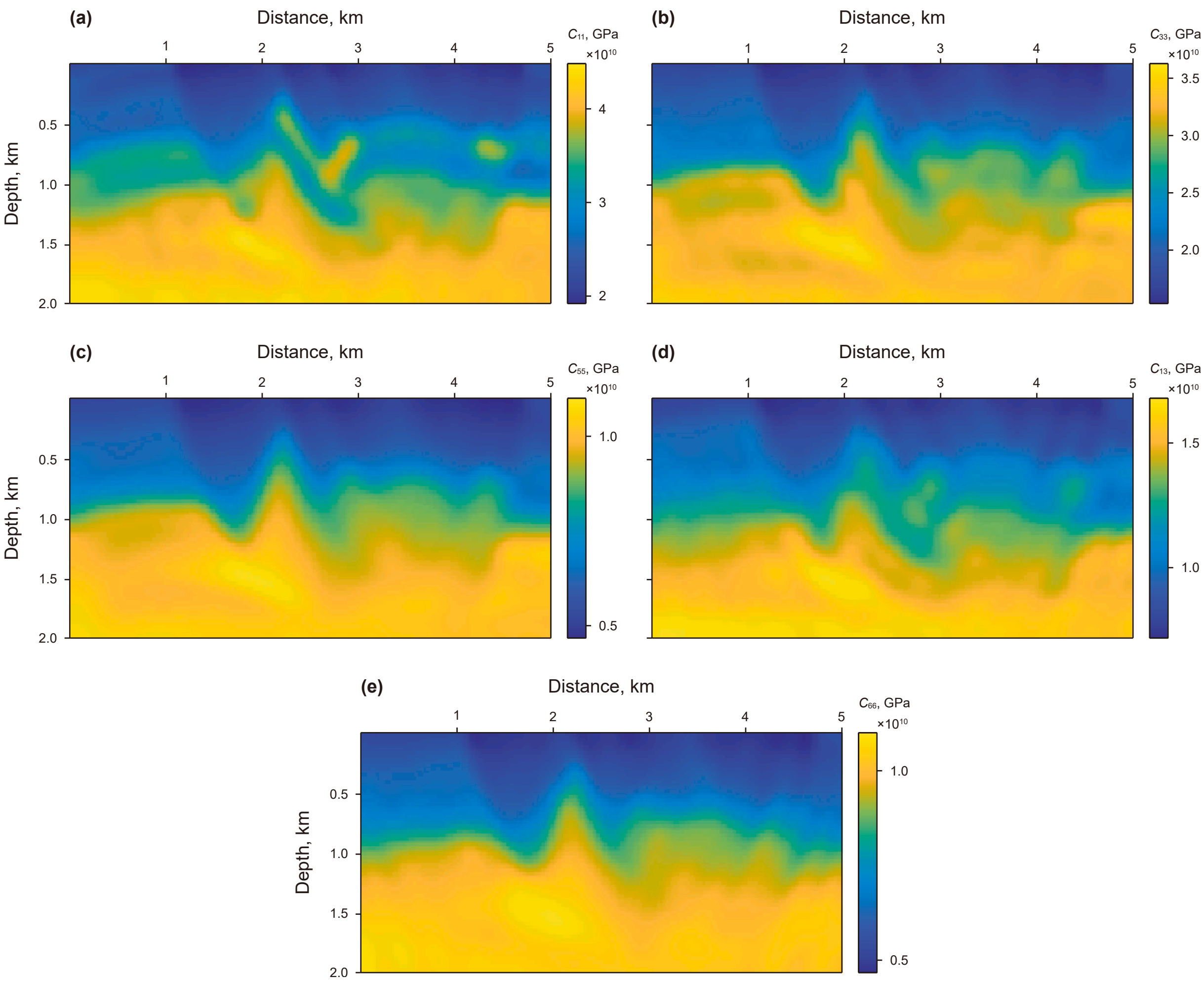

Fig. 23. Inverted results of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the BP TTI subset model experiment obtained by introducing TV regularization term.

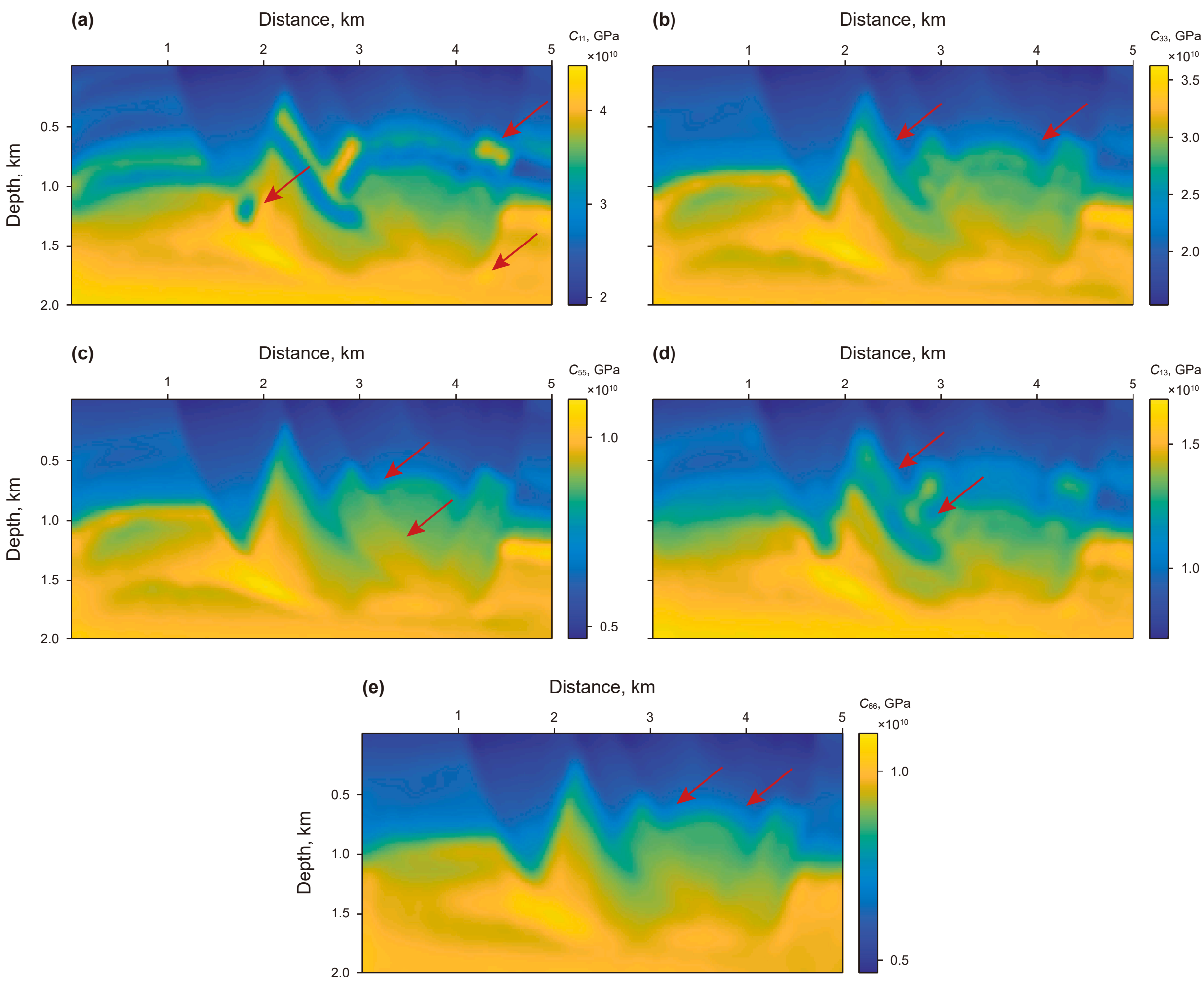

Fig. 24. The final MAP solution of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the BP TTI subset model experiment via the proposed method, respectively.

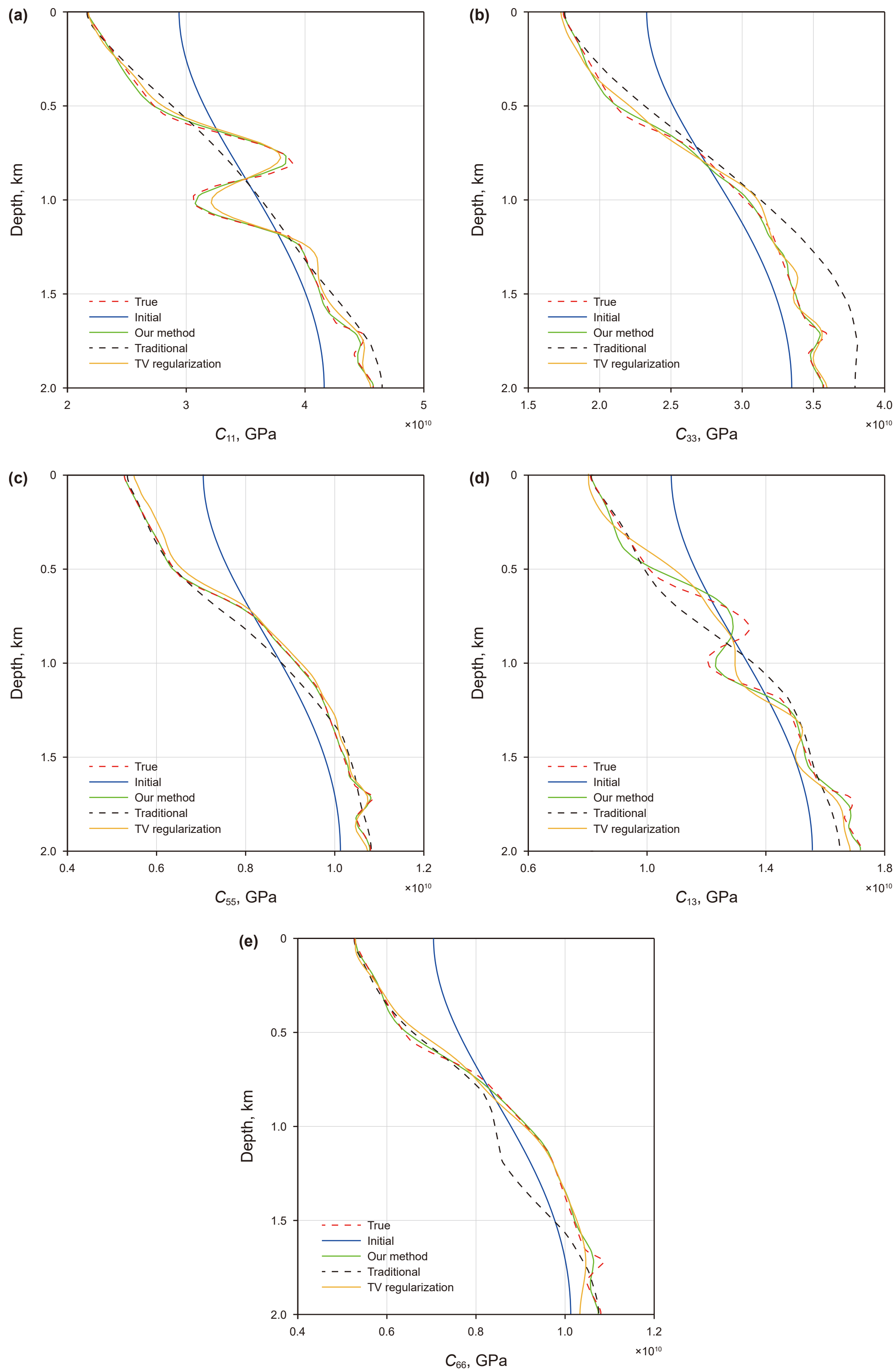

Fig. 25. Comparisons of depth profiles at $x = 2.5$ km of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ between the true model in Fig. 16 (red dashed curves), the initial model in Fig. 17 (blue curves), the inverted results via the traditional approach in Fig. 22 (black dashed curves), the inverted results via the TV regularization in Fig. 23 (yellow curves) and the final MAP solution via the proposed method in Fig. 24 (green curves), respectively.

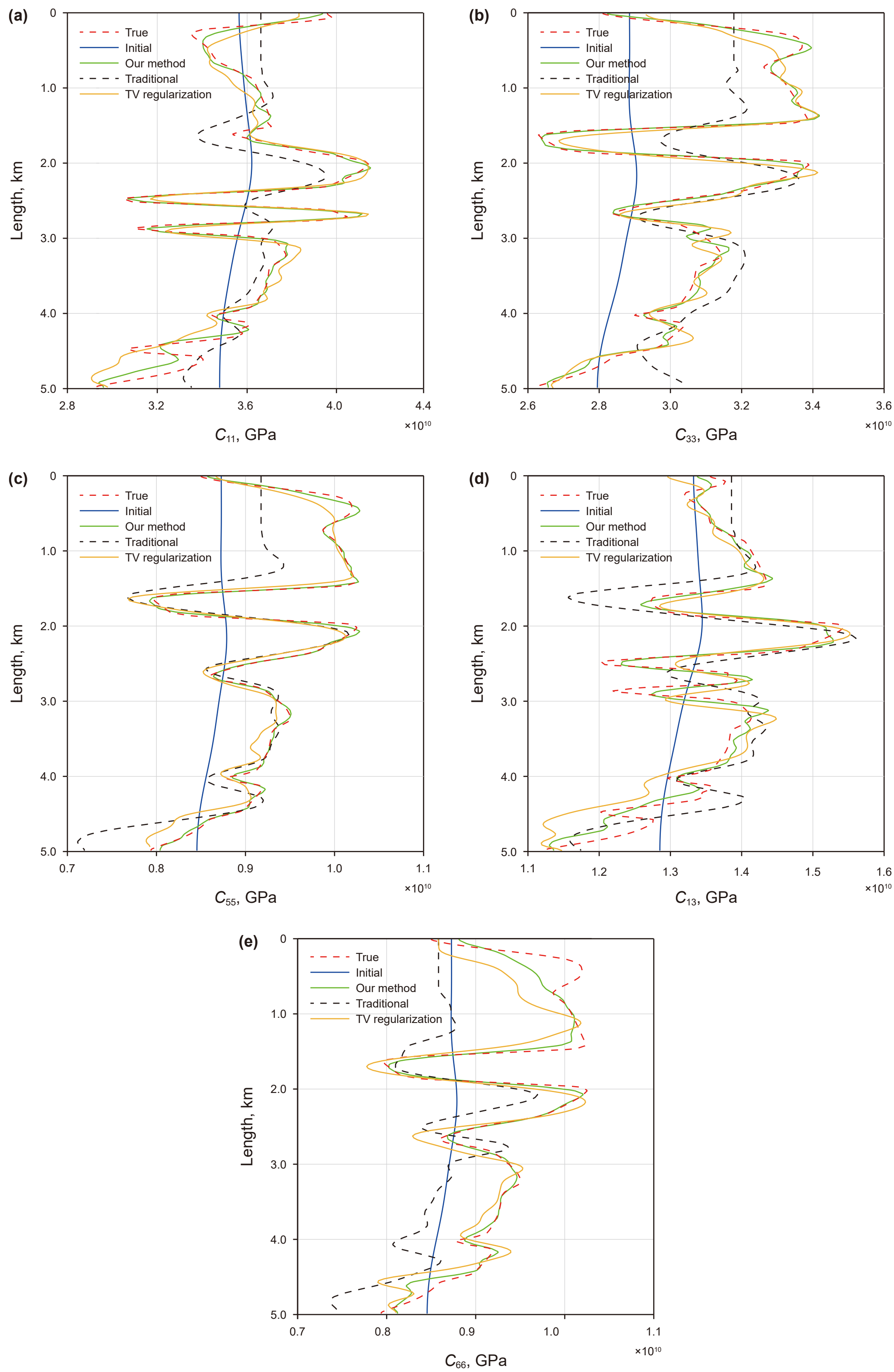

Fig. 26. Comparisons of length profiles at $z = 1$ km of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ between the true model in Fig. 16 (red dashed curves), the initial model in Fig. 17 (blue curves), the inverted results via the traditional approach in Fig. 22 (black dashed curves), the inverted results via the TV regularization in Fig. 23 (yellow curves) and the final MAP solution via the proposed method in Fig. 24 (green curves), respectively.

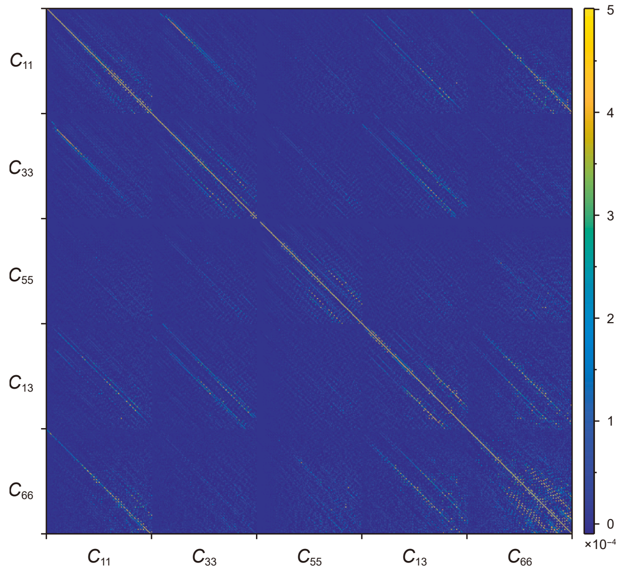

Fig. 27. The full posterior covariance matrix. From upper left to lower right, the blocks correspond to $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$, respectively.

randomized SVD algorithm to approximate and store the posterior covariance matrix, the required memory has been reduced from 3.3185 GB to 322.88 MB, saving about 10.28 times the memory. In addition, by using the GMRES algorithm to solve the Lippmann–Schwinger equation for the wavefield, the memory requirement has been reduced from 11.7 GB to 7.92 GB, saving approximately 32% in computational overhead. Above all, we can conclude that the proposed approach appears to have several good characteristics, no matter the accuracy of the estimated models or the required iterations to reach an acceptable accuracy. Although for complicated large-scale cases, the computational cost may still be a tricky problem.

4. Conclusion

By analyzing the inverted elastic moduli profiles and the posterior covariance matrices from two numerical tests, along with the physical nature of anisotropic media, we can evaluate the difficulties for reconstructing different parameters and the cross-correlations among them. Within the Bayesian framework in the multi-parameter cases, the symmetric full posterior covariance matrix is associated with a given MAP model. The non-zero elements in the off-diagonal blocks indicate the cross-correlations

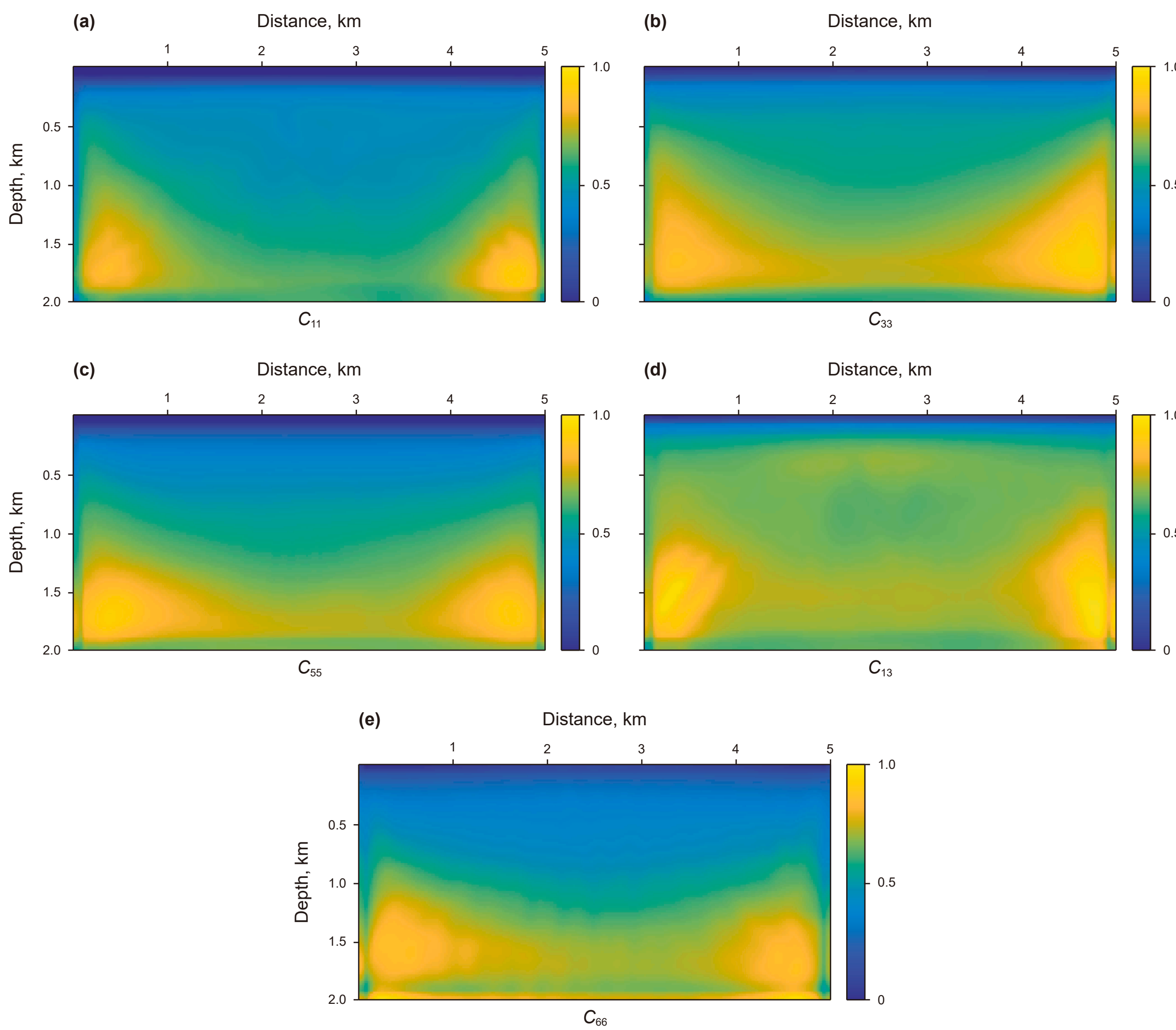

Fig. 28. The posterior standard deviations of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the BP TTI subset model experiment, respectively.

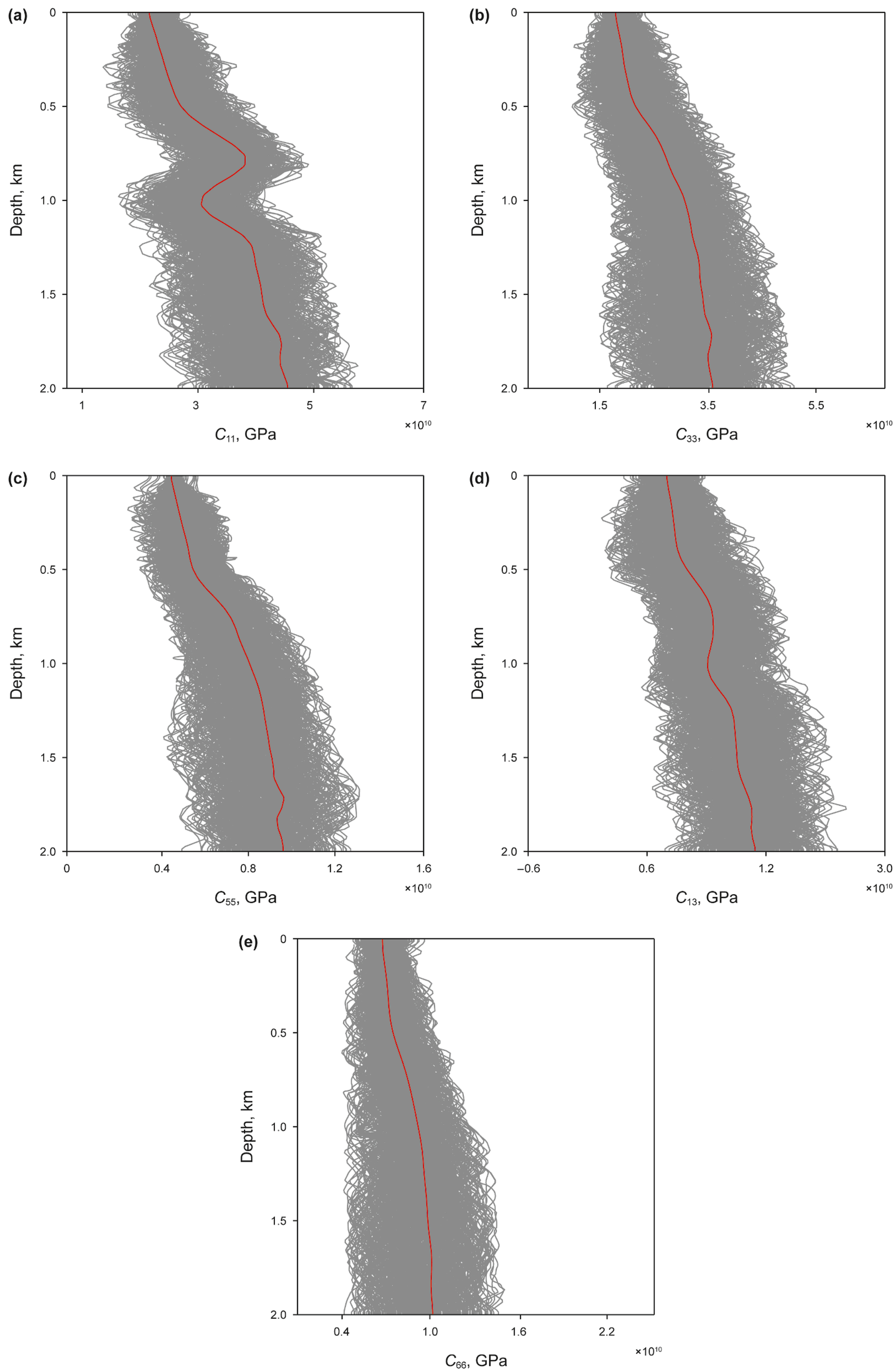

Fig. 29. Comparisons of depth profiles between the posterior distributions at $x = 2.5$ km of **(a)**–**(e)** $C_{11}, C_{33}, C_{55}, C_{13}, C_{66}$ for the BP TTI subset model experiment, respectively. Red curves are MAP solution, and gray profiles are random samples drawn from posterior distributions with 95% confidence level.

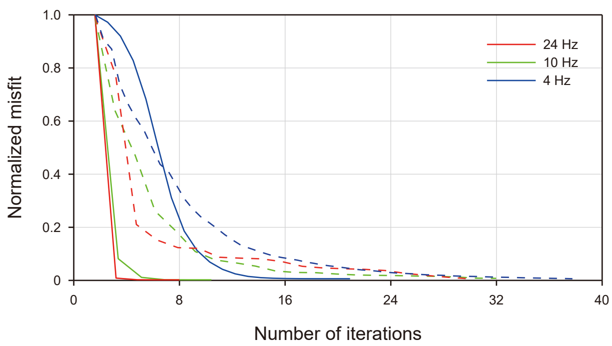

Fig. 30. Convergence rate at three different frequencies for the BP TTI subset model experiment via the traditional approach (dashed curves) and the proposed method (solid curves), respectively.

between different model parameters, with their magnitude reflecting the degree of coupling. These couplings result from the way that different parameters interact with each other in the wavefield. That is, when two parameters jointly influence the data misfit, they often exhibit significant covariance in the posterior distribution. The elements along the main diagonal signify the variances of the fluctuations that characterize the uncertainties of the five independent elastic constants. Model uncertainty also reflects the reliability in parameter reconstruction arising from data uncertainty. On the other hand, we can observe that the standard deviation increases with depth, particularly in the Hess model experiment. This can be attributed to geometry spreading effects, wherein Hessian energy decreases with depth due to geometric attenuation (Pratt, 1999). Additionally, standard deviations are significantly larger near boundaries and perturbed interfaces, as wave paths bend in regions with high perturbation.

Combining the standard deviation and the inversion results of each parameter, it is obvious that C_{33} and C_{55} demonstrate the highest accuracy in recovery due to their relative independence. Specifically, C_{33} is directly associated with the P-wave velocity, making it highly sensitive to seismic measurements. And C_{55} governs vertical S-wave velocity, benefiting from its shorter wavelength, leading to a superior resolution. The Thomsen parameter ε denotes the fractional difference between the vertical and horizontal components of the P-wave velocity, also known as the P-wave anisotropy, which is jointly characterized by C_{11} and C_{33}. Thus, the inversion of C_{11} is slightly challenging because its estimation is constrained by C_{33}. Recovery of C_{66} and C_{13} is less accurate than other coefficients, yet C_{66} is slightly easier due to the relatively independent nature of SH-wave energy regardless of its weakness. Retrieving C_{13} is the most challenging because it is coupled with C_{33} and C_{55}. The three clastic constants are correlated together with δ, which represents the second-order derivative of the P-wave phase velocity at vertical incidence.

We can basically conclude that the proposed method is robust to establish the elastic moduli of the anisotropic medium based on the Bayesian framework and enable simultaneous estimation of multiparameter and their uncertainty quantification. However, further improvements, including the second-order scattering effect of the Hessian matrix, are still necessary to further improve the resolution and computational efficiency.

CRediT authorship contribution statement

Wen-Rui Ye: Writing – original draft, Methodology, Conceptualization. **Xing-Guo Huang:** Funding acquisition, Conceptualization.

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.

Acknowledgement

This work was supported by the Deep Earth National Science and Technology Major Project of China under Grant 2024ZD1002907, and the National Natural Science Foundation of China under Grant 42374149.

References

Abubakar, A., Habashy, T.M., 2013. Three-dimensional visco-acoustic modeling using a renormalized integral equation iterative solver. J. Comput. Phys. 249, 1–12. https://doi.org/10.1016/j.jcp.2013.04.008.

Aghazade, K., Gholami, A., Aghamiry, H.S., et al., 2024. Robust elastic full-waveform inversion using an alternating direction method of multipliers with reconstructed wavefields. Geophysics 89 (3), R287–R302. https://doi.org/10.1190/geo2023-0411.1.

Alkhalifah, T., Plessix, R.-E., 2014. A recipe for practical full-waveform inversion in anisotropic media: an analytical parameter resolution study. Geophysics 79 (3), R91–R101. https://doi.org/10.1190/geo2013-0366.1.

Alkhalifah, T., Sun, B.B., Wu, Z., 2018. Full model wavenumber inversion: identifying sources of information for the elusive middle model wavenumbers. Geophysics 83 (6), R597–R610. https://doi.org/10.1190/geo2017-0775.1.

Arenas, E., van Kruijsdijk, C., Oldenziel, T., 2001. Semi-automatic history matching using the pilot point method including time-lapse seismic data. In: SPE Annual Technical Conference and Exhibition. https://doi.org/10.2118/71634-ms. SPE-71634-MS.

Bachmann, E., Tromp, J., 2020. Source encoding for viscoacoustic ultrasound computed tomography. J. Acoust. Soc. Am. 147 (5), 3221–3235. https://doi.org/10.1121/10.0001191.

Barnes, C., Charara, M., Tsuchiya, T., 2008. Feasibility study for an anisotropic full waveform inversion of cross-well seismic data. Geophys. Prospect. 56 (6), 897–906. https://doi.org/10.1111/j.1365-2478.2008.00702.x.

Bell, B.M., Cathey, F.W., 1993. The iterated Kalman filter update as a gauss-newton method. IEEE Trans. Automat. Control 38 (2), 294–297. https://doi.org/10.1109/9.250476.

Bertaccini, D., Durastante, F., 2018. Iterative Methods and Preconditioning for Large and Sparse Linear Systems with Applications. Chapman and Hall/CRC. https://doi.org/10.1201/9781315153575.

Bishop, C.M., Nasrabadi, N.M., 2006. Pattern Recognition and Machine Learning. Springer, New York.

Blei, D.M., Kucukelbir, A., McAuliffe, J.D., 2017. Variational inference: A review for statisticians. J. Am. Stat. Assoc. 112 (518), 859–877. https://doi.org/10.1080/01621459.2017.1285773.

Bond, W.L., 1943. The mathematics of the physical properties of crystals. Bell Syst. Tech. J. 22 (1), 1–72. https://doi.org/10.1002/j.1538-7305.1943.tb01304.x.

Bosch, M., Mukerji, T., Gonzalez, E.F., 2010. Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review. Geophysics 75 (5), 75A165–75A176. https://doi.org/10.1190/1.3478209.

Brooks, S., Gelman, A., Jones, G., et al., 2011. Handbook of Markov Chain Monte Carlo. CRC Press, Boca Raton.

Bui-Thanh, T., Ghattas, O., Martin, J., et al., 2013. A computational framework for infinite-dimensional Bayesian inverse problems part I: The linearized case, with application to global seismic inversion. SIAM J. Sci. Comput. 35 (6), A2494–A2523. https://doi.org/10.1137/12089586x.

Candansayar, M.E., 2008. Two-dimensional inversion of magnetotelluric data with consecutive use of conjugate gradient and least-squares solution with singular value decomposition algorithms. Geophys. Prospect. 56 (1), 141–157. https://doi.org/10.1111/j.1365-2478.2007.00668.x.

Cervený, V., 2001. Seismic Ray Theory. Cambridge university press, Cambridge.

Chang, H., McMechan, G., 2009. 3D 3-C full-wavefield elastic inversion for estimating anisotropic parameters: A feasibility study with synthetic data. Geophysics 74 (6), WCC159–WCC175. https://doi.org/10.1190/1.3204766.

Chen, K., 2005. Matrix Preconditioning Techniques and Applications (No. 19). Cambridge University Press. https://doi.org/10.1017/cbo9780511543258.024.

Chen, K., Liu, L., Xu, L., et al., 2024. Linearized waveform inversion for vertical transversely isotropic elastic media: Methodology and multi-parameter crosstalk analysis. Pet. Sci. 21 (1), 252–271 doi:10.1016-/j.petsci.2024.01.002.

Djebbi, R., Plessix, R.-É., Alkhalifah, T., 2017. Analysis of the traveltime sensitivity kernels for an acoustic transversely isotropic medium with a vertical axis of symmetry. Geophys. Prospect. 65 (1), 22–34. https://doi.org/10.1111/1365-2478.12361.

Duane, S., Kennedy, A.D., Pendleton, B.J., et al., 1987. Hybrid Monte Carlo. Phys. Lett. B 195 (2), 216–222. https://doi.org/10.1016/0370-2693(87)91197-X.

Eikrem, K.S., Nævdal, G., Jakobsen, M., 2018. Iterated extended Kalman filter method for time-lapse seismic fullwaveform inversion. Geophys. Prospect. 66 (2), 379–394. https://doi.org/10.1111/1365-2478.12730.
Fernandes, F.J.D., Teixeira, L., Freire, A.F.M., et al., 2024. Stochastic seismic inversion and Bayesian facies classification applied to porosity modeling and igneous rock identification. Pet. Sci. 21 (2), 918–935. https://doi.org/10.1016/j.petsci.2023.11.020.
Fiandaca, G., Doetsch, J., Vignoli, G., et al., 2015. Generalized focusing of time-lapse changes with applications to direct current and time-domain induced polarization inversions. Geophys. J. Int. 203 (2), 1101–1112. https://doi.org/10.1093/gji/ggv350.
Fichtner, A., van Leeuwen, T., 2015. Resolution analysis by random probing. J. Geophys. Res. Solid Earth 120 (8), 5549–5573. https://doi.org/10.1002/2015jb012106.
Fichtner, A., Zunino, A., Gebraad, L., 2019. Hamiltonian Monte Carlo solution of tomographic inverse problems. Geophys. J. Int. 216 (2), 1344–1363. https://doi.org/10.1093/gji/ggy496.
Ford, J.M., 2003. An improved discrete wavelet transform preconditioner for dense matrix problems. SIAM J. Matrix Anal. Appl. 25 (3), 642–661. https://doi.org/10.1137/s0895479802416526.
Gebraad, L., Boehm, C., Fichtner, A., 2020. Bayesian elastic full-waveform inversion using hamiltonian Monte Carlo. J. Geophys. Res. Solid Earth 125 (3), e2019JB018428. https://doi.org/10.1029/2019JB018428.
Gonis, A., Butler, W.H., 2000. Multiple Scattering in Solids. Graduate Texts in Contemporary Physics. Springer, New York. https://doi.org/10.1007/978-1-4612-1290-4.
Gouveia, W.P., Scales, J.A., 1998. Bayesian seismic waveform inversion: parameter estimation and uncertainty analysis. J. Geophys. Res. Solid Earth 103 (B2), 2759–2779. https://doi.org/10.1029/97JB02933.
Gu, Y., Oliver, D.S., 2007. An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. SPE J. 12 (4), 438–446. https://doi.org/10.2118/108438-pa.
Gündoğdu, N.Y., Candansayar, M.E., 2018. Three-dimensional regularized inversion of DC resistivity data with different stabilizing functionals. Geophysics 83 (6), E399–E407. https://doi.org/10.1190/geo2017-0558.1.
Guo, P., Visser, G., Saygin, E., 2020. Bayesian Trans-dimensional full waveform inversion: Synthetic and field data application. Geophys. J. Int. 222 (1), 610–627. https://doi.org/10.1093/gji/ggaa201.
Halko, N., Martinsson, P.G., Tropp, J.A., 2011. Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. SIAM Rev. 53 (2), 217–288. https://doi.org/10.1137/090771806.
Hastings, W.K., 1970. Monte Carlo sampling methods using Markov Chains and their applications. Biometrika 57 (1), 97–109. https://doi.org/10.1093/biomet/57.1.97.
Huang, X., Jakobsen, M., Wu, R., 2019. Taming the divergent terms in the scattering series of born by renormalization. In: SEG Technical Program Expanded Abstracts, vol 2019, pp. 5065–5069. https://doi.org/10.1190/segam2019-3216450.1.
Huang, X., Eikrem, K.S., Jakobsen, M., et al., 2020. Bayesian full waveform inversion in anisotropic elastic media using the iterated extended Kalman filter. Geophysics 85 (3), R267–R286. https://doi.org/10.1190/geo2019-0644.1.
Huang, X., Sun, T., Greenhalgh, S., 2024. Modeling of seismic wave scattering in poroelastic media. Geophysics 89 (6), T303–T317. https://doi.org/10.1190/geo2023-0616.1.
Jakobsen, M., Hudson, J.A., Johansen, T.A., 2003. T-matrix approach to shale acoustics. Geophys. J. Int. 154 (2), 533–558. https://doi.org/10.1046/j.1365-246x.2003.01977.x.
Jakobsen, M., Psencik, I., Iversen, E., et al., 2020. Transition operator approach to seismic full-waveform inversion in arbitrary anisotropic elastic media. Commun. Comput. Phys. 25 (1), 297–327. https://doi.org/10.4208/cicp.oa-2018-0197.
Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Trans. ASME J. Basic Eng. 82, 34–45. https://doi.org/10.1115/1.3662552.
Kamath, N., Tsvankin, I., 2016. Elastic full-waveform inversion for VTI media: Methodology and sensitivity analysis. Geophysics 81 (2), C53–C68. https://doi.org/10.1190/geo2014-0586.1.
Kaufl, P., Fichtner, A., Igel, H., 2013. Probabilistic full waveform inversion based on tectonic regionalization—Development and application to the Australian upper mantle. Geophys. J. Int. 193 (1), 437–451. https://doi.org/10.1093/gji/ggs131.
Kazei, V.V., Kalita, M., Alkhalifah, T., 2017. Salt-body inversion with minimum gradient support and Sobolev space norm regularizations. In: 79th EAGE Conference and Exhibition. https://doi.org/10.3997/2214-4609.201700600.
Kotsi, M., Malcolm, A., Ely, G., 2020. Time-lapse full-waveform inversion using Hamiltonian Monte Carlo: A Proof of concept. SEG Technical Program Expanded Abstracts 2020, pp. 845–849. https://doi.org/10.1190/segam2020-3422774.1.
Kouri, D.J., Vijay, A., Hoffman, D.K., 2003. Inverse scattering theory: Renormalization of the Lippmann−Schwinger equation for quantum elastic scattering with spherical symmetry. J. Phys. Chem. 107 (37), 7230–7235. https://doi.org/10.1021/jp030273r.
Kubrusly, C.S., Gravier, J., 1973. Stochastic approximation algorithms and applications. In: 1973 IEEE Conference on Decision and Control Including the 12th Symposium on Adaptive Processes, pp. 763–766. https://doi.org/10.1109/CDC.1973.269114.
Li, R., Reynolds, A.C., Oliver, D.S., 2003. History matching of three-phase flow production data. In: Proceedings of SPE Reservoir Simulation Symposium, RSS. https://doi.org/10.2523/66351-ms.
Liberty, E., Woolfe, F., Martinsson, P.-G., et al., 2007. Randomized algorithms for the low-rank approximation of matrices. Proc. Natl. Acad. Sci. 104 (51), 20167–20172. https://doi.org/10.1073/pnas.0709640104.
Liu, Q., Peter, D., 2019. Square-root variable metric based elastic full-waveform inversion—Part 2: Uncertainty estimation. Geophys. J. Int. 218 (2), 1100–1120. https://doi.org/10.1093/gji/ggz137.
Macbeth, C., 2002. Multi-Component VSP Analysis for Applied Seismic Anisotropy. Pergamon, New York.
Martin, J., Wilcox, L.C., Burstedde, C., et al., 2012. A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion. SIAM J. Sci. Comput. 34 (3), A1460–A1487. https://doi.org/10.1137/110845598.
Metropolis, N., Ulam, S., 1949. The Monte Carlo method. J. Am. Stat. Assoc. 44 (247), 335–341. https://doi.org/10.1080/01621459.1949.10483310.
Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data. Geophysics 52 (9), 1211–1228. https://doi.org/10.1190/1.1442384.
Mosegaard, K., Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems. J. Geophys. Res. Solid Earth 100 (B7), 12431–12447. https://doi.org/10.1029/94JB03097.
Nawaz, M.A., Curtis, A., 2018. Variational Bayesian inversion (VBI) of quasi-localized seismic attributes for the spatial distribution of geological facies. Geophys. J. Int. 214 (2), 845–875. https://doi.org/10.1093/gji/ggy163.
Nguyen, F., Kemna, A., Robert, T., et al., 2016. Data-driven selection of the minimum-gradient support parameter in time-lapse focused electric imaging. Geophysics 81 (1), A1–A5. https://doi.org/10.1190/geo2015-0226.1.
Oh, J., Alkhalifah, T., 2016. Elastic orthorhombic anisotropic parameter inversion: An analysis of parameterization. Geophysics 81 (6), C279–C293. https://doi.org/10.1190/geo2015-0656.1.
Operto, S., Gholami, Y., Prieux, V., et al., 2013. A guided tour of multi-parameter full-waveform inversion with multicomponent data: from theory to practice. Lead. Edge 32 (9), 1040–1054. https://doi.org/10.1190/tle32091040.1.
Osnabrugge, G., Leedumrongwatthanakun, S., Vellekoop, I.M., 2016. A convergent Born series for solving the inhomogeneous Helmholtz equation in arbitrarily large media. J. Comput. Phys. 322, 113–124. https://doi.org/10.1016/j.jcp.2016.06.034.
Ostebee, A., 1998. Rank-Deficient and Discrete ill-posed Problems: Numerical Aspects of Linear Inversion. Society for Industrial and Applied Mathematics.
Portniaguine, O., Zhdanov, M.S., 1999. Focusing geophysical inversion images. Geophysics 64 (3), 874–887. https://doi.org/10.1190/1.1444596.
Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain: Part 1, Theory and verification in a physical scale model. Geophysics 64 (3), 888–901. https://doi.org/10.1190/1.1444597.
Prieux, V., Brossier, R., Gholami, Y., et al., 2011. On the footprint of anisotropy on isotropic full waveform inversion: The Valhall case study. Geophys. J. Int. 187 (3), 1495–1515. https://doi.org/10.1111/j.1365-246X.2011.05209.x.
Prieux, V., Brossier, R., Operto, S., et al., 2013. Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 1: Imaging compressional wave speed, density and attenuation. Geophys. J. Int. 194 (3), 1640–1664. https://doi.org/10.1093/gji/ggt177.
Ray, A., Kaplan, S., Washbourne, J., et al., 2017. Low frequency full waveform seismic inversion within a tree-based Bayesian framework. Geophys. J. Int. 212 (1), 522–542. https://doi.org/10.1093/gji/ggx428.
Robbins, H., Monro, S., 1951. A stochastic approximation method. Ann. Math. Stat. 22 (3), 400–407.
Saad, Y., Schultz, M.H., 1986. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Comput. 7 (3), 856–869. https://doi.org/10.1137/0907058.
Sambridge, M., 2014. A parallel tempering algorithm for probabilistic sampling and multimodal optimization. Geophys. J. Int. 196 (1), 357–374. https://doi.org/10.1093/gji/ggt342.
Sambridge, M., Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems. Rev. Geophys. 40 (3), 3-1–3–29. https://doi.org/10.1029/2000RG000089.
Sebacher, B., Hanea, R., 2024. Bridging element-free Galerkin and Pluri-Gaussian simulation for geological uncertainty estimation in an ensemble smoother data assimilation framework. Pet. Sci. 21 (3), 1683–1698. https://doi.org/10.1016/j.petsci.2023.12.020.
Shang, W., Xue, W., Xu, Y., et al., 2019. Undersea target reconstruction based on coupled Laplacian-of-Gaussian and minimum gradient support regularizations. IEEE Access 7, 171633–171647. https://doi.org/10.1109/access.2019.2954293.
Shekhar, U., Jakobsen, M., Iversen, E., et al., 2024. Microseismic wavefield modelling in anisotropic elastic media using integral equation method. Geophys. Prospect. 72 (2), 403–423. https://doi.org/10.1111/1365-2478.13416.
Skoglund, M.A., Hendeby, G., Axehill, D., 2015. Extended Kalman filter modifications based on an optimization viewpoint. In: 18th International Conference on Information Fusion, pp. 1851–1861.
Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation. Geophysics 49 (8), 1259–1266. https://doi.org/10.1190/1.1441754.
Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM. https://doi.org/10.1137/1.9780898717921.

Taylor, J., 1972. Scattering Theory. John Wiley & Sons. https://doi.org/10.1887/0750304960/b493c11.

Thomsen, L., 1986. Weak elastic anisotropy. Geophysics 51 (10), 1954–1966. https://doi.org/10.1190/1.1442051.

Thurin, J., Brossier, R., Métivier, L., 2019. Ensemble-based uncertainty estimation in full waveform inversion. Geophys. J. Int. 219 (3), 1613–1635. https://doi.org/10.1093/gji/ggz384.

Ting, T.C., 1996. Anisotropic Elasticity: Theory and Applications. Oxford University Press, New York. https://doi.org/10.1093/oso/9780195074475.001.0001.

Tromp, J., Tape, C., Liu, Q., 2004. Seismic tomography, adjoint methods, time reversal, and banana-doughnut kernels. Geophys. J. Int. 160 (1), 195–216. https://doi.org/10.1111/j.1365-246x.2004.02453.x.

Tsvankin, I., 2011. In: Seismic Signatures and Analysis of Reflection Data in Anisotropic Media, third ed. Soc. Explor. Geophys., Tulsa https://doi.org/10.1190/1.9781560803003.

Warner, M., Guasch, L., 2016. Adaptive waveform inversion: Theory. Geophysics 81 (6), R429–R445. https://doi.org/10.1190/geo2015-0387.1.

Warner, M., Ratcliffe, A., Nangoo, T., et al., 2013. Anisotropic 3D full-waveform inversion. Geophysics 78 (2), R59–R80. https://doi.org/10.1190/geo2012-0338.1.

Woodward, M.J., 1992. Wave-equation tomography. Geophysics 57 (1), 15–26. https://doi.org/10.1190/1.1443179.

Ye, W., Huang, X., Han, L., et al., 2023. Bayesian inverse scattering theory for multiparameter full-waveform inversion in transversely isotropic media. Geophysics 88 (3), C91–C109. https://doi.org/10.1190/geo2022-0140.1.

Zhang, X., Curtis, A., 2020. Seismic tomography using variational inference methods. J. Geophys. Res. Solid Earth 125 (4). https://doi.org/10.1029/2019JB018589 e2019JB018589.

Zhang, F., Reynolds, A.C., 2002. Optimization algorithms for automatic history matching of production data. In: ECMOR VIII–8th European Conference on the Mathematics of Oil Recovery. https://doi.org/10.3997/2214-4609.201405958.

Zhang, C., Bütepage, J., Kjellström, H., et al., 2018. Advances in variational inference. IEEE Trans. Pattern Anal. Mach. Intell. 41 (8), 2008–2026. https://doi.org/10.1109/TPAMI.2018.2889774.

Zhao, Z., Sen, M.K., 2019. A gradient based MCMC method for FWI and uncertainty analysis. In: SEG Technical Program Expanded Abstracts 2019, pp. 1465–1469. https://doi.org/10.1190/segam2019-3216560.1.

Zhao, C., Yu, P., Zhang, L., 2016. A new stabilizing functional to enhance the sharp boundary in potential field regularized inversion. J. Appl. Geophys. 135, 356–366. https://doi.org/10.1016/j.jappgeo.2016.10.033.

Zhdanov, M.S., 2002. Geophysical Inverse Theory and Regularization Problems. Elsevier, The Netherlands. https://doi.org/10.1016/s0076-6895(02)80041-5.

Zhu, J., Dorman, J., 2000. Two-dimensional three-component wave propagation in a transversely isotropic medium with arbitrary orientation finite-element modeling. Geophysics 65 (3), 934–942. https://doi.org/10.1190/1.1444789.

Zhu, H., Li, S., Fomel, S., et al., 2016. A Bayesian approach to estimate uncertainty for full-waveform inversion using a priori information from depth migration. Geophysics 81 (5), R307–R323. https://doi.org/10.1190/geo2015-0641.1.