



Near-Surface Rayleigh Wave Dispersion Curve Inversion Algorithms: A Comprehensive Comparison

Xiao-Hui Yang¹ · Yuanyuan Zhou¹ · Peng Han^{2,3,4} · Xuping Feng^{2,3,4} ·
Xiaofei Chen^{2,3,4}

Received: 9 September 2023 / Accepted: 6 February 2024
© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Abstract

Rayleigh wave exploration is a powerful method for estimating near-surface shear-wave (S-wave) velocities, providing valuable insights into the stiffness properties of subsurface materials inside the Earth. The dispersion curve inversion of Rayleigh wave corresponds to the optimization process of searching for the optimal solutions of earth model parameters based on the measured dispersion curves. At present, diversified inversion algorithms have been introduced into the process of Rayleigh wave inversion. However, limited studies have been conducted to uncover the variations in inversion performance among commonly used inversion algorithms. To obtain a comprehensive understanding of the optimization performance of these inversion algorithms, we systematically investigate and quantitatively assess the inversion performance of two bionic algorithms, two probabilistic algorithms, a gradient-based algorithm, and two neural network algorithms. The evaluation indices include the computational cost, accuracy, stability, generalization ability, noise effects, and field data processing capability. It is found that the Bound-constrained limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS-B) algorithm and the broad learning (BL) network have the lowest computational cost among candidate algorithms. Furthermore, the transitional Markov Chain Monte Carlo algorithm, deep learning (DL) network, and BL network outperform the other four algorithms regarding accuracy, stability, resistance to noise effects, and capability to process field data. The DL and BL networks demonstrate the highest level of generalization compared to the other algorithms. The comparison results reveal the variations in candidate algorithms for the inversion task, causing a clear understanding of the inversion performance of candidate algorithms. This study can promote the S-wave velocity estimation by Rayleigh wave inversion.

Keywords Rayleigh wave inversion · Near-surface S-wave velocity · Optimization algorithm · Inversion performance comparison

Extended author information available on the last page of the article

Article Highlights

- A comprehensive comparison of Rayleigh wave inversion performance is conducted among two bionic algorithms, two probabilistic algorithms, a gradient-based algorithm, and two neural network algorithms
- Evaluation factors include the computational cost, accuracy, stability, generalization ability, effects of noise, and field data processing capability
- Comparison results reveal the variations in inversion performance among candidate algorithms

1 Introduction

Near-surface Rayleigh wave method is dependable, economical, and efficient for estimating shear-wave (S-wave) velocities which serve the effective indicators of subsurface material stiffness. (Park et al. 2007; Xia et al. 2007; Renalier et al. 2010; Gao et al. 2016; Pan et al. 2016; Socco et al. 2017; Cheng et al. 2018a; Mi et al. 2018, 2023; Gribler et al. 2020; Zuo et al. 2022). The procedure of the near-surface Rayleigh wave method involves four main steps, i.e., the acquisition of seismic Rayleigh wave data, the imaging of dispersion energy, the extraction of dispersion curves, and the inversion of dispersion curves (Sahadewa et al. 2011; Fang et al. 2015; Cheng et al. 2018b; Gao et al. 2018; Yamanaka and Chimoto 2018; Feng and Chen 2022). The first step measures the vibrations in the ground induced by active and/or passive sources. Active sources refer to the generators of seismic energy used to simulate impulse signals, including sledgehammers, air guns, explosives, etc. In particular, the multichannel analysis of surface waves (MASW) method (Park et al. 1999, 2007) is the preferred approach for acquiring Rayleigh waves using active sources. Passive sources are responsible for the production of ambient noises, which encompass both human activities and natural phenomena, such as traffic, construction work, tidal currents, and volcanic eruptions (Park et al. 2007; Eker et al. 2012; Cheng et al. 2019). The measured waves by the former exhibit rich in the higher frequency signals, and the latter present rich in the lower frequency signals, indicating active and passive sources are better in the shallower and deeper layers, respectively. As the second step, dispersion energy imaging converts the seismic signals from the time-offset domain to the frequency–velocity one. This step upon the active-source seismic signals depends on the integral transformation algorithms, e.g., the phase shift (Park et al. 1998), the frequency decomposition and slant stacking (Xia et al. 2007), the high-resolution linear Radon transform (Luo et al. 2008), etc. Unlike the active-source way, dispersion energy imaging algorithms based on the passive-source waves consider the approximation of empirical Green's function by the cross-correlation of ambient noise between two stations. The spatial autocorrelation (Aki 1957), extended spatial autocorrelation (Ohori et al. 2002), and frequency-Bessel transform (abbreviation: F-J, Wang et al. 2019) were developed based on such theory. Notably, the F-J method has attracted a lot of attention due to its satisfying capability for revealing higher-mode dispersion information and is also applicable to the MASW survey (Wu et al. 2020; Hu et al. 2020; Xi et al. 2021; Fu et al. 2022a, b). Furthermore, the third step extracts multi-modal dispersion curves along the local energy extremes from the obtained dispersion energy. The technique for extracting dispersion curves has undergone a development process, transitioning from manual to semi-manual and eventually to fully automatic (Dai et al. 2021; Dong et al. 2021; Song et al. 2021a, b). Manual and semi-manual methods adapt

to experienced researchers, while the fully automatic manner is more suitable for processing large amounts of images. The last step in the Rayleigh wave prospecting, dispersion curve inversion, employs the extracted dispersion curves to predict the optimal 1-D S-wave velocity profile under the optimization criteria of the best fitting between measured and inverted dispersion curves (Song et al. 2008; Hobiger et al. 2013; Lei et al. 2019; Tremblay and Karray 2019; Poormirzaee and Fister 2021). The inverted 1-D S-wave velocity profile is regarded as the recorded values of the midpoint of the arrays, which reflects the average velocities of the subsurface. If necessary, a pseudo-2D or 3D structure of S-wave velocity can be established by interpolating several 1D profiles (Luo et al. 2008; Zeng et al. 2011; Shakir et al. 2013; Pan et al. 2018; Mi et al. 2020; Zhang et al. 2021). The accuracy and stability of the predicted S-wave velocity profile are influenced by the inversion performance of the applied optimization algorithm, provided with reliably measured dispersion curves (Dal Moro and Pipan 2007; Yao 2015; Poormirzaee 2016; Shen et al. 2018; Li et al. 2022; Lin et al. 2023). Enhancing the inversion performance has always been a persistent focus in the research field of Rayleigh wave inversion (Dal Moro 2008; Pan et al. 2013; Boiero and Socco 2014; Lu et al. 2016; Haney and Tsai 2017; Cercato 2018; Liu et al. 2019; Bergamo et al. 2023).

Rayleigh wave dispersion curve inversion is a continuous optimization problem characterized by multiple local extrema, high nonlinearity, and strong non-uniqueness of solutions. In practice, researchers typically set certain preconditions to reduce the optimization difficulties. These preconditions mainly come in two aspects, i.e., the setting of search space and assumptions of some parameters. Regarding the first aspect, researchers generally define a search space before inversion based on prior geological information and observations of the measured dispersion curves. It is undoubtedly that the performance of constrained optimization benefits from a narrower search space in case the defined search space covers the actual values of unknown parameters. However, pursuing a narrower search space can easily lead to an undesired situation where some parameter values to be inverted lie outside the defined search space. Once this situation occurs, the accuracy and stability of the estimated S-wave velocities cannot be guaranteed by the inversion process. Therefore, a search space with a wide range is typically predetermined before inversion by researchers' experience-based judgment. Under such circumstances, several studies have been recently conducted to address the search space issue for more accurate and stable inversion results. Yang et al. (2022) discussed the inaccurate S-wave velocity estimation determined by a wrong search space and proposed a data-driven broad learning framework for search space design to achieve a suitable search space to improve the inversion performance. Yang et al. (2023) proposed a two-stage inversion strategy to conduct the optimization two times inside the inversion process. This inversion strategy is capable of adjusting the search space. The first inversion stage performs inversion based on a search space prescribed by researchers' experience. In the second inversion stage, a new search space with a narrower range is designed by the two-stage inversion strategy to implement inversion for better estimation of S-wave velocities. In terms of the other aspect of addressing optimization challenges, assuming specific parameters to be known is a widely accepted routine to reduce the optimization difficulties of inversion (Yuen and Yang 2020). For the inversion process, it adopts an assumption of the horizontally layered earth model, constrained by the computational model of theoretical dispersion curves. Based on such assumption, four groups of parameters control the dispersion curves of Rayleigh wave, namely the S-wave velocities, thicknesses, P-wave velocities, and densities of different layers. The P-wave velocities and densities can be assumed known before the inversion process because their contribution to dispersion curves is relatively minor than the first two groups of parameters.

Meanwhile, the number of layers represents an implicit parameter that determines the total number of parameters involved in the inversion process. There are two inversion manners to deal with the number of layers for optimization difficulty reduction: One considers assuming a small number of layers to invert both the S-wave velocity and thickness of each layer. The other adopts the inversion manner of multi-thin-layer while only the S-wave velocities of different layers are inverted. In contrast with each other, these two inversion manners differ in the application background: the former is more suitable for near-surface prospecting; on the contrary, the latter is more commonly applied for deeper surveys.

Given a designed search space and adopted assumptions mentioned above, the inversion performance is influenced by employed optimization algorithm inside inversion process. The inversion task is based on forward modeling for the generation of theoretical dispersion curves (Chen 1993; Wathelet 2005). The inversion algorithm is in charge of performing massive calculations of theoretical dispersion curves for achieving the best fitting between measured and dispersion curves for recovering the S-wave velocities of the subsurface (Boxberger et al. 2011; Li et al. 2012; Gouveia et al. 2016; Bonadio et al. 2018; Liu et al. 2023). Theoretically, grid search may be the best algorithm for Rayleigh wave inversion concerning the evaluation criterion is the searching situation of the global optimal parameter values (Benoit et al. 2006; Caldwell et al. 2009). However, grid search is not suitable for high-dimensional optimization with a wide range of parameter values due to its prohibitively expensive computational cost, which is exactly the case for near-surface Rayleigh wave inversion. Therefore, the optimization algorithms have been widely utilized to accelerate the searching process of global optimal S-wave velocities. During the past nearly 30 years, the introduction of efficient inversion algorithms with powerful global optimization capability has always been an important concern for inversion task. Currently, the widely applied optimization algorithms in dispersion curve inversion can be mainly categorized into four types, containing bionic algorithms (Lu and Zhang 2006; Zhou et al. 2014), probabilistic algorithms (Lu et al. 2014; Yang and Yuen 2021), gradient-based algorithms (Zhang and Alkhalifah 2019; Zhan et al. 2020), neural network algorithms (Luo et al. 2022; Yang et al. 2023), etc.

Bionic algorithms approximate the global optimal values by simulating natural features in the environment, e.g., genetic algorithm, particle swarm optimization algorithm. During the 1990s, the bionic algorithms were introduced into the process of dispersion curve inversion. Lomax and Snieder (1995) utilized the genetic algorithm to invert the S-wave velocity information of the mantle under the East European Platform, western and central Europe. Yamanaka and Ishida (1996) estimated the S-wave profile with the genetic algorithm by dispersion data inversion measured at a site in the Kanto Plain with the S-wave velocities and thicknesses of different layers considered unknown. Dal Moro et al. (2007) applied the genetic algorithm for the estimation of the near-surface S-wave velocities and layer thicknesses and then employed the marginal posterior probability density to provide the uncertainties of these parameters for Rayleigh wave dispersion curve inversion. Since the twenty-first century, other bionic algorithms have been increasingly introduced for inverting S-wave velocities, such as particle swarm optimization algorithm (Song et al. 2012; Wilken and Rabbel 2012), Ant colony algorithm (Xu and Song 2012), bee colony algorithm (Zarean et al. 2013). At present, bionic algorithms are still the popular choice to undertake the optimization task within the dispersion curve inversion process due to their global optimization capability and fast convergency property (Qin et al. 2020; Song et al. 2021a, b; Fu et al. 2022a, b; Vashisth et al. 2022; Zhang et al. 2022; Wang et al. 2023). However, bionic algorithms are sensitive to the initial population, which indicates that the

convergence performance is easily affected by the setting of the initial population, leading to unsatisfactory optimization stability.

Probabilistic algorithms treat the unknown parameters as random variables and search for the globally optimal values by maximizing the probability density of these variables. The most representative probabilistic algorithms include the simulated annealing algorithm and the Markov Chain Monte Carlo algorithm. Introducing these two probabilistic algorithms into dispersion curve inversion for S-wave velocity estimation has begun in the early twentieth century. Beaty et al. (2002) jointly inverted the fundamental and higher-mode Rayleigh wave dispersion curves where the simulated annealing algorithm performed the search of the near-surface optimal S-wave velocities. Pei et al. (2007) utilized the simulated annealing algorithm to determine two S-wave velocity profiles based on field data that fitted the suspension logging data well. Bodin et al. (2012) introduced the reversible jump Markov chain Monte Carlo algorithm to estimate both the S-wave velocities and the number of layers under the joint utilization of receiver functions and surface wave dispersion data. Based on the ambient noises of 315 regional stations, Berg et al. (2018) used the ambient noise data of 315 regional stations to estimate the 3-D shear-wave velocity structures beneath Southern California via the Markov chain Monte Carlo algorithm. Yang and Yuen (2021) introduced a multi-stage Metropolis–Hastings algorithm (transitional Markov chain Monte Carlo, Ching and Chen 2007) to conduct the near-surface Rayleigh wave inversion while the P-wave velocity and density of each layer were also treated as unknown to be inverted. Some of the probabilistic algorithms, e.g., the Markov Chain Monte Carlo algorithm, can quantify the uncertainty of the optimal parameters. However, the Markov Chain Monte Carlo algorithm requires expensively computational cost for producing samples to form a set of ones representing the posterior probability density function (PDF). The substantial disparity between the prior and posterior PDFs can exacerbate such a situation, which is the case for Rayleigh wave inversion in Yang and Yuen (2021).

The gradient-based algorithms take advantage of the gradient information to determine the search direction within the iteration procedure for finding the minimum values of the objective function. The fast iteration property of the gradient-based algorithms attracted the attention of researchers seeking to improve the computational efficiency of dispersion curve inversion for S-wave velocity estimation (Li et al. 2017). There is a history spanning more than 20 years of using gradient-based algorithms for performing dispersion curve inversion. Pasyanos et al. (2001), Yanovskaya and Kozhevnikov (2003), Campman and Dwi Riyanti (2007), Kiełczyński and Szalewski (2011), Yin et al. (2020), etc. utilized the conjugate gradient algorithm to address the optimization problem in dispersion curve inversion for estimating S-wave velocities of the subsurface. Duputel et al. (2010), Gofer et al. (2017), Pan et al. (2019), Wu et al. (2020), Fu et al. (2021), Zhang et al. (2021), etc. conducted the dispersion curve inversion in order to invert the stiffness of materials of the subsurface based on the quasi-Newton algorithms. The disadvantage of the gradient-based algorithms comes at being easily falling into the local minimum values and the optimization performance can be affected by the initial model for the inversion process. Currently, the gradient-based algorithms are still widely applied inside the Rayleigh wave dispersion curve inversion process due to their certain optimization capability and fast convergency advantage.

Neural network algorithms can be capable of simulating the functional mapping between the input–output variables via the connections of neurons which can perform linear and nonlinear transforms. As a subset of machine learning, neural network algorithms can learn the underlying relationships within existing data and automatically establish a data-driven analytical model, enabling optimal decision-making with the minimum of

human being intervention. The deep learning network has been considered as the most representative one of neural networks due to its strong functional modeling capability (Mousavi and Beroza 2019, 2023; Huynh et al. 2023; Zhou et al. 2023). In recent years, the deep learning network has been introduced into the process of dispersion curve inversion. For example, Chen et al. (2022) used the deep learning network to determine the 1-D S-wave velocity profiles by inverting the fundamental Rayleigh wave dispersion curve; Wang et al. (2022a) utilized a deep learning network to achieve a satisfactory 3-D inversion performance based on Rayleigh-wave phase velocities ranging from 8 to 70 s and group velocities ranging from 10 to 70 s. Although the deep learning network possesses powerful modeling capability, it suffers from the disadvantage of time-consuming network training (Chen and Liu 2017; Kuok and Yuen 2020; Wang et al. 2022b; Zhou et al. 2022). To overcome this drawback of the deep learning network, Chen and Liu (2017) proposed a broad learning network with a single-layer architecture. The broad learning network differs from the traditional single-layered one, reflecting that the former comes equipped with much more neurons in width dimension than the latter. The complexity in width dimension compares favorably with the one in depth dimension. As a result, the broad learning network holds comparable functional mapping capability in contrast with the deep learning network without worrying about the time-consuming problem. In practice, the broad learning network has been tried inside the dispersion curve inversion, as demonstrated in the studies by Yang et al. (2022) and Yang et al. (2023). These studies demonstrated that broad learning can obtain equivalent inversion performance as deep learning under much less network training cost with respect to Rayleigh wave inversion.

The optimization algorithm plays a crucial character for accurate and stable estimation of S-wave velocities by Rayleigh wave inversion. As mentioned above, there are various types of algorithms for inversion, seemingly somewhat overwhelming and confusing when selecting the appropriate one for practical applications. At present, few studies comprehensively compare these types of inversion algorithms. In this regard, this study makes a comprehensive comparison among the popular near-surface Rayleigh wave dispersion curve inversion algorithms with the purpose of clarifying the inversion performance of these algorithms. Furthermore, researchers can easily choose an appropriate inversion algorithm by referring to the comparison results in this study. The candidate algorithms include two bionic algorithms, two probabilistic algorithms, a gradient-based algorithm, and two neural network algorithms. The details about the specific candidate algorithms and the inversion performance indicators will be elaborated later.

2 Problem Description

Rayleigh wave dispersion curve inversion is carried out based on the hypothesis of a horizontally layered earth model for modeling theoretical dispersion curves. Such a hypothesis has gained wide acceptance and practical application in the real world. Figure 1 displays an earth model of multi-layered half-space, presenting a horizontally n -layered earth model with four groups of parameters: layer thicknesses, S-wave velocities, P-wave velocities, and densities. It is known that the dispersion curves of Rayleigh wave are dominated by these four groups of parameters under the horizontally layered earth model. Concerning the reduction in the non-uniqueness and mapping difficulties of Rayleigh wave inversion, the P-wave velocities and densities of different layers are typically pre-estimated before

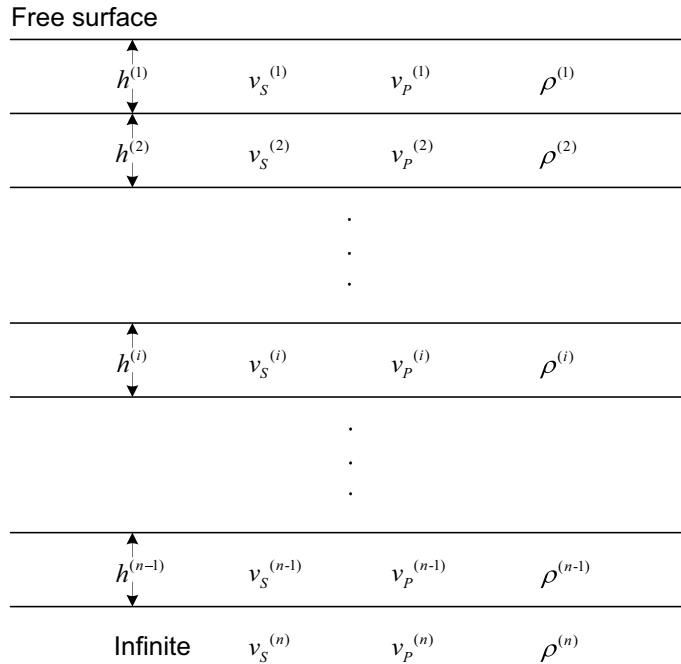


Fig. 1 An earth model of multi-layered half-space (Yang et al. 2023)

inversion in real applications due to their comparably less sensitivity to dispersion curves (Xia et al. 1999; Renalier et al. 2010; Yuen and Yang 2020). The candidate algorithms to be compared also treat these two groups of parameters as known before inversion.

Based on the above assumptions, the dispersion curve inversion is equivalent to determining the layer thicknesses and S-wave velocities in earth model parameter vector $\theta = [\mathbf{h}^T, \mathbf{v}_S^T]^T \in \mathbb{R}^{N_\theta}$ ($N_\theta=2n-1$), including $n-1$ parameters in layer thickness vector $\mathbf{h} = [h^{(1)}, \dots, h^{(i)}, \dots, h^{(n-1)}]^T \in \mathbb{R}^{n-1}$ and n parameters in S-wave velocity vector $\mathbf{v}_S = [v_S^{(1)}, \dots, v_S^{(i)}, \dots, v_S^{(n)}]^T \in \mathbb{R}^n$, where $h^{(i)}$ and $v_S^{(i)}$ are the layer thickness and S-wave velocity of i th layer of subsurface. The functional mapping to be modeled for the inversion task can be expressed as follows:

$$\theta = F(\mathbf{v}_R) \quad (1)$$

where $\mathbf{v}_R \in \mathbb{R}^{N_v}$ is the measured dispersion curve vector, N_v is the number of phase velocities of measured dispersion curves, and $F(\cdot)$ is the unknown functional mapping to be modeled by the candidate algorithms between \mathbf{v}_R and θ . In particular, an objective function is required inside the optimization process in order to evaluate the misfit between measured and theoretical dispersion curves. This study adopts the following objective function for the comparison of candidate algorithms:

$$F_{\text{Obj}} = \frac{1}{N_v} \|\mathbf{v}_R - \hat{\mathbf{v}}_R\|_2^2 \quad (2)$$

where F_{Obj} is the objective function and $\hat{\mathbf{v}}_R$ is the optimal theoretical dispersion curve vector:

$$\hat{\mathbf{v}}_R = T(\hat{\theta}) \quad (3)$$

where $T(\cdot)$ is the modeling function of theoretical dispersion curves and $\hat{\theta}$ is the optimal earth model parameter vector to be inverted.

3 Candidate Inversion Algorithms

The concern of this study is to comprehensively evaluate the inversion performance of frequently used inversion algorithms. As previously stated, the candidate algorithms for comparison belong to four types (namely, two bionic algorithms, two probabilistic algorithms, a gradient-based algorithm, and two neural network algorithms). Specifically, there are seven candidate algorithms to be evaluated for Rayleigh wave inversion, including the genetic algorithm (GA), particle swarm optimization (PSO) algorithm, simulated annealing (SA) algorithm, transitional Markov Chain Monte Carlo (TMCMC) algorithm, Bound-constrained limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS-B) algorithm, deep learning (DL) network, and broad learning (BL) network. In particular, the computer used in this study for testing is a desktop PC equipped with 3.80GHz Intel(R) Xeon(R) W-2235 CPU, six cores, and 32.0 GB RAM.

3.1 Genetic Algorithm

The GA (Holland 1975; Goldberg 1989) is developed by mimicking the genetic inheritance and survival of the fittest in the biological evolution process of nature. The process of biological evolution involves the preservation of favorable genes and the elimination of inferior ones, allowing species that are more adapted to the natural environment to survive. For Rayleigh wave inversion, the objective function F_{Obj} represents the natural environment. An individual in the GA is typically expressed by a binary string to simulate genes. The GA enables the population to evolve toward optimizing the objective function F_{Obj} , where individuals with favorable genes can come into the next generation. The operations, such as gene selection, crossover, and mutation, are carried out within the genetic algorithm to achieve the optimization of the objective function. The iterations of the GA in this study terminate when the inversion process reaches the preset maximum number of iterations. The number of generations, the number of individuals, generation gap, crossover rate, and mutation rate equal 100, 100, 0.9, 0.7, and 0.0175, respectively. The implementation steps of the GA for optimization are shown in Fig. 2.

3.2 Particle Swarm Optimization Algorithm

The PSO algorithm (Kennedy and Eberhart 1995) is inspired by the collectively cooperative behavior of bird flocks, such as searching for food and avoiding obstacles, etc. It is found that individuals within a bird flock communicate and exchange information so that the population collectively moves to more favorable positions. The PSO algorithm simulates such behaviors of bird flocks to achieve the global optimal point for a

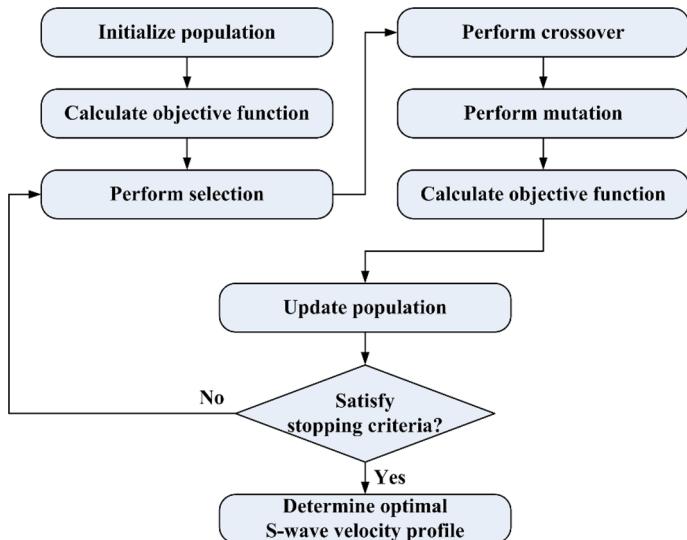


Fig. 2 Implementation steps of the GA for Rayleigh wave inversion

given optimization problem. There are four definitions in the particle swarm algorithm, namely, particle, swarm, personal best, and global best. A particle inside the search space denotes a sample. Its position can be considered a potential solution, and its velocity affects the movement. Each particle corresponds to an objective function value. All particles comprise a swarm, where the personal and global best are the best result for a particle and the swarm, respectively. Searching for the optimal solution can be implemented by modeling the movement of particles of the swarm, like the collectively cooperative behavior of bird flocks. Each particle moves based on the current position and updated velocity, where the latter can be determined by the current velocity, the distance from the current position to the personal best, and the one to the global best. The maximum number of iteration and the number of particles inside the swarm in this study are set as 60 and 100, respectively. The implementation steps of the PSO algorithm for Rayleigh wave inversion are shown in Fig. 3.

3.3 Simulated Annealing Algorithm

The SA algorithm (Kirkpatrick et al. 1983; Černý 1985) drew its inspiration from the annealing process in metallurgy. During the high-temperature annealing period of a material, the atoms automatically move to positions where the internal energy decreases. As a result, the material reaches the minimum internal energy value at the end of cooling. The principle of the SA algorithm is similar to the one of the cooling of a material. It requires a high temperature and a certain number of random samples within search space for initiating the SA algorithm. Throughout the annealing process, each sample moves in accordance with the Metropolis rule. The temperature gradually decreases as the annealing step progresses. The acceptance rate of a new solution relies on the objective function values of current and candidate samples and the current temperature. Upon completion of the cooling scheme, the SA algorithm can obtain globally optimal parameters for an optimization

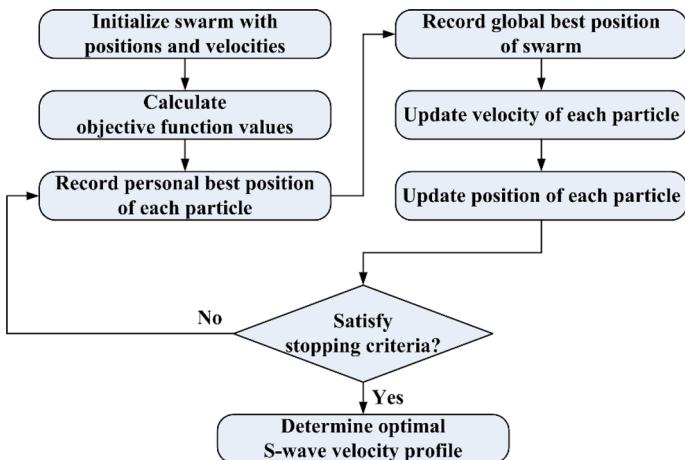


Fig. 3 Implementation steps of the PSO algorithm for Rayleigh wave inversion

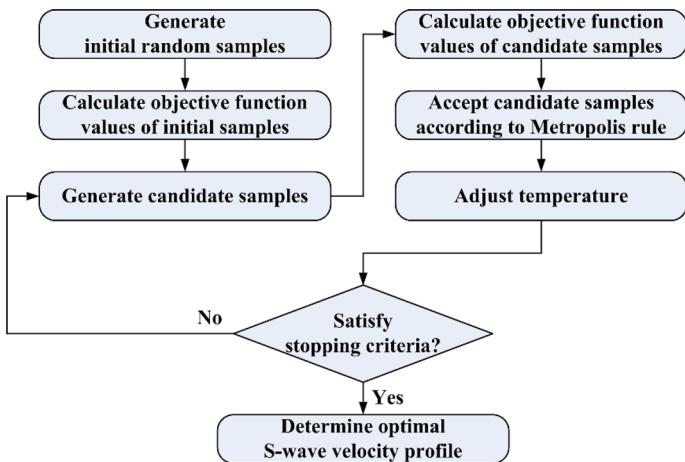


Fig. 4 Implementation steps of the SA algorithm for Rayleigh wave inversion

problem. In this study, the maximum number of iterations and the initial temperature is equal to 3000 and 1000, respectively. The implementation steps of the SA algorithm for Rayleigh wave inversion are shown in Fig. 4.

3.4 Transitional Markov Chain Monte Carlo Algorithm

The TCMC algorithm (Ching and Chen 2007) is a multi-stage Metropolis–Hastings (MH) algorithm (Metropolis et al. 1953; Hastings 1970). As a representative Markov Chain Monte Carlo algorithm, the MH algorithm presents unsatisfactory sampling performance facing optimization problems where there is a substantial difference between the prior probability density function (PDF) and posterior one (Ching and Chen 2007). The core

idea of the TMCMC algorithm is to obtain easy sampling between two adjacent PDFs by inserting some intermediate PDFs between prior and posterior PDFs, resulting in a gradual transition of samples from the prior PDF to the posterior one. As a result, drawing samples from a target posterior PDF that is flat, highly peaked, multimodal, and multidimensional can be achieved efficiently. In contrast with the MH algorithm, the TMCMC algorithm possesses better optimization performance at the cost of heightened computational effort. For Rayleigh wave inversion, the TMCMC algorithm is more appropriate than the MH algorithm (Yang and Yuen 2021). As some intermediate PDFs introduced, there are several sampling stages inside the TMCMC algorithm. In each sampling stage, the number of samples is set as 2000. Meanwhile, the uniform distribution is regarded as the prior PDF that guides the initial sampling process for Rayleigh wave inversion by the TMCMC algorithm. During the execution of the algorithm, re-sampling is performed to select samples transitioning to the next stage, while the MH algorithm is also employed to supplement samples. The TMCMC algorithm terminates when all sampling stages have been completed. The implementation steps of the TMCMC algorithm for Rayleigh wave inversion are shown in Fig. 5.

3.5 Bound-Constrained Limited-Memory Broyden–Fletcher–Goldfarb–Shanno Algorithm

The L-BFGS-B algorithm (Byrd et al. 1995; Zhu et al. 1997) incorporates the constant lower and upper bound constraints on variables to be optimized to extend the L-BFGS (limited-memory BFGS) algorithm which provides an approximation of the BFGS (Broyden–Fletcher–Goldfarb–Shanno) algorithm employing limited-memory resources. It is widely recognized that Newton's method requires the availability of the Hessian matrix to perform the iterative search for solving the optimization problems. It is incapable of execution when the Hessian matrix or its inverse one is unavailable. As an alternative to Newton's method, quasi-Newton methods can approximate the Hessian matrix by the iterations of gradient information. In practice, the BFGS algorithm is a representative and widely applied one of quasi-Newton methods. Compared with the BFGS algorithm, the L-BFGS

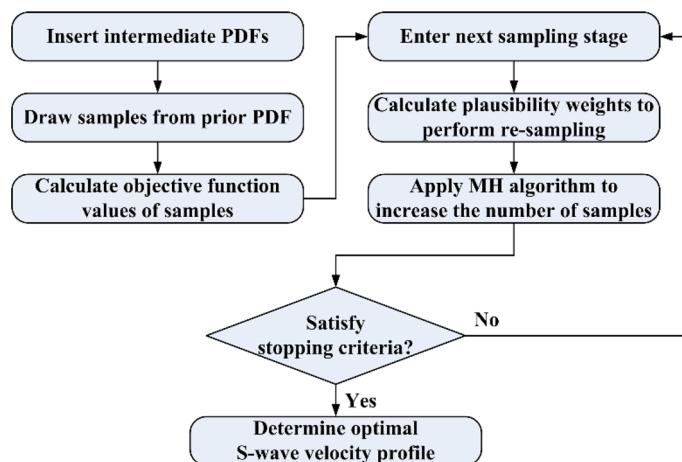


Fig. 5 Implementation steps of the TMCMC algorithm for Rayleigh wave inversion

algorithm significantly saves computational resources by limiting the storage of information about Hessian matrix. As for the L-BFGS-B algorithm, it is specifically designed for constrained optimization problems where the calculation of the Hessian matrix or its inverse one encounters difficulties, showcasing similar performance to the L-BFGS algorithm. The maximum of iterations and iterative convergence tolerance of the L-BFGS-B algorithm for inversion process are defined as 350 and 1e-6, respectively. The implementation steps of the L-BFGS-B algorithm for Rayleigh wave inversion are shown in Fig. 6.

3.6 Deep Learning Network

The DL network is constructed by the configuration of a multi-layered neural network equipped with numerous neurons. In practice, the deep learning network has been considered as the representative of machine learning methods due to its powerful functional mapping capability. Figure 7 depicts the configuration of deep learning network. A neuron can perform linear and nonlinear transformations on input variables. The mapping capability of a deep learning network is associated with the number of neurons. The number of layers (N_{layers}) and the number of neurons in each layer (N_{neurons}) determine the number of all neurons in the deep learning network, manipulating the network mapping complexity. For a Rayleigh wave inversion application case, the mapping function to be modeled is provided with a fixed complexity. The grid search manner is adopted to select a deep learning network with appropriate mapping complexity within the space $[3 : 2 : 15] \times [3 : 2 : 15]$ ($[A : B : C]$ denotes the set of numbers between A and C with interval B). Specifically, both N_{layers} and N_{neurons} range from 3 to 15, with intervals of 2. In the meantime, we employ the Levenberg–Marquardt algorithm (Moré 1978) to train the candidate networks with different mapping complexities. According to Yang et al.

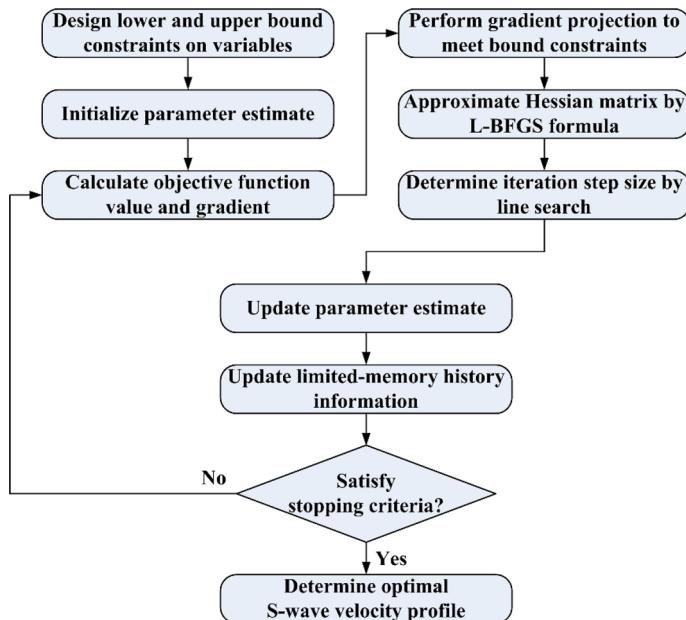


Fig. 6 Implementation Steps of the L-BFGS-B algorithm for Rayleigh Wave Inversion

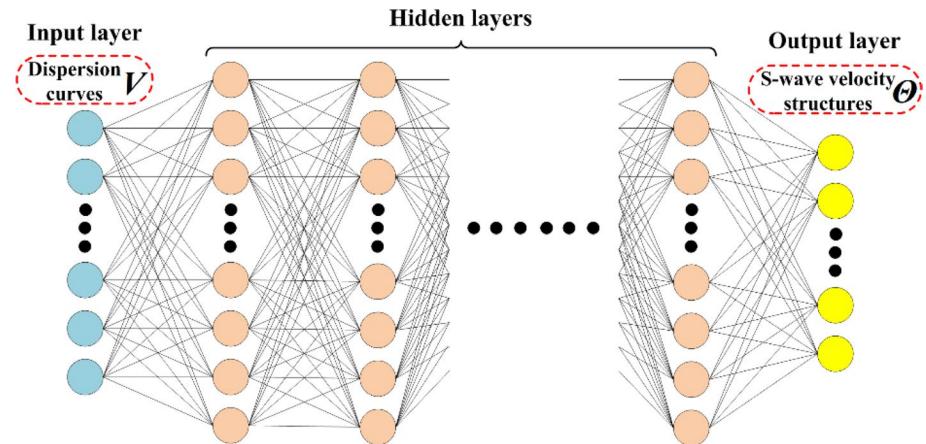


Fig. 7 Configuration of the DL network

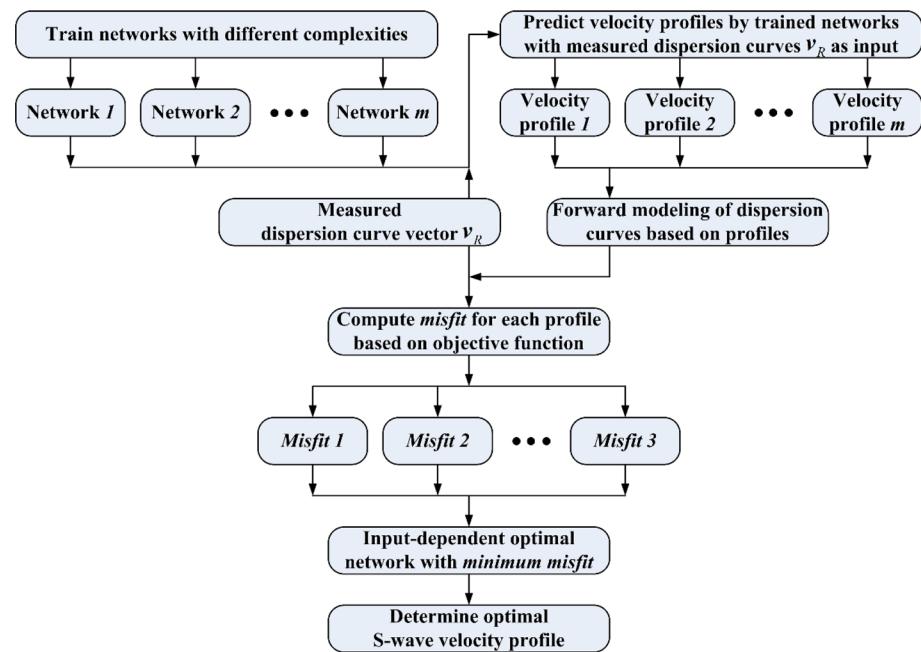


Fig. 8 Implementation steps of the network complexity selection for the DL and BL network for Rayleigh wave inversion

(2023), the utilization of the forward modeling presents better performance than the validation dataset with respect to the determination of suitable network complexity. This study applies the forward modeling way to perform the network complexity selection, which can refer to the flowchart in Fig. 8. For each network, there are 500 input–output sample pairs provided for network training.

3.7 Broad Learning Network

The BL network (Chen and Liu 2017; Chen et al. 2018) establishes a single-layered neuron network with complex architecture along the width direction in the input layer. The idea of proposing the BL network comes to address the expensive computational cost issue of network training of the DL network due to the complex feature in the depth direction. Figure 9 shows the configuration of the BL network. The constitution of mapped features and enhancement nodes enables numerous neurons in the input layer, ensuring the powerful mapping capability of the BL network. The complexity in the width direction as a substitute for the depth direction guarantees the high computational efficiency of network training. As a result, the BL network requires significantly less training time than the DL network while possessing comparable functional mapping capability. For the network complexity selection, the grid search manner is also applied for the BL network. Specifically, the space $[4 : 2 : 10] \times [4 : 2 : 20] \times [4 : 2 : 30]$ is given for determining the most suitable values for parameters N_C (the number of neurons in each group of mapped features), N_M (the number of the groups of mapped features), and N_E (the number of the groups of enhancement nodes) of the BL network. The implementation steps of the network complexity selection can be followed according to Fig. 8. Similar to the network training of a DL network, 500 input–output sample pairs are provided for function modeling.

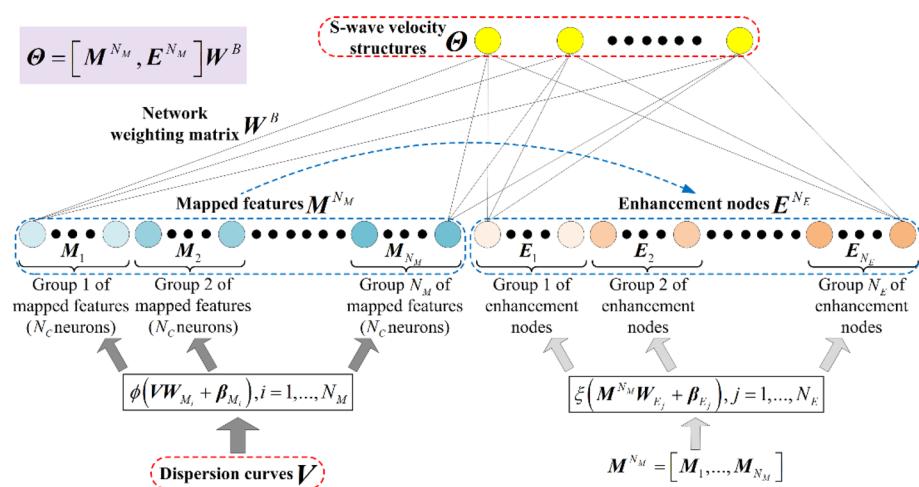
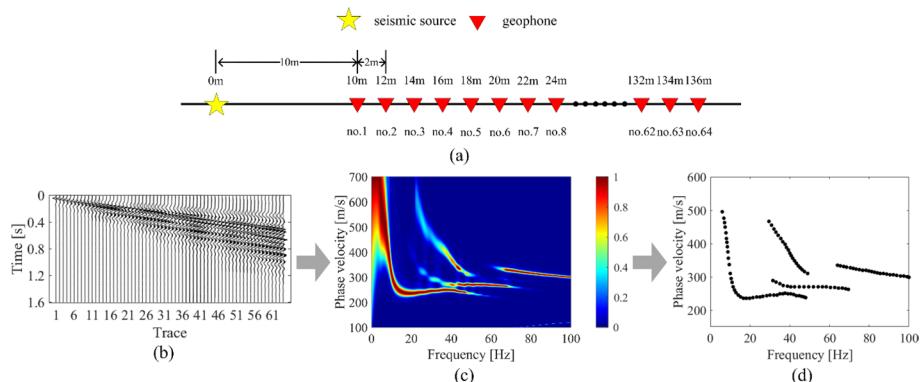


Fig. 9 Configuration of the BL network (Yang et al. 2023)

Table 1 Actual parameter values of the theoretical earth model

Layer	h (m)	v_S (m/s)	v_P (m/s)	ρ (g/cm ³)
1	2.00	400.0	700.0	1.90
2	4.00	200.0	300.0	1.70
3	5.00	300.0	500.0	1.80
4	5.00	500.0	900.0	2.00
5	N/A	650.0	1100.0	2.10

**Fig. 10** Extracted Rayleigh wave dispersion curves based on the numerical data: **a** seismic array of active-source survey, **b** theoretical Rayleigh wave data, **c** dispersion energy image, **d** extracted dispersion curves

4 Numerical Simulation

The comprehensive comparison of candidate algorithms is first performed by a theoretical earth model with five layers. Table 1 lists the actual parameter values of the earth model. Based on the theoretical earth model, the CPS330 (Computer Program in Seismology) program developed by Herrmann (2013) was utilized to produce the simulated Rayleigh wave data for the generation of theoretical multi-modal dispersion curves. Figure 10a, b shows the distribution of the seismometers and their recorded Rayleigh waves, respectively. There were 64 seismometers under the vertical seismic source with 10 m minimum source offset for the simulation of the theoretical data, where the spatial interval was set as 2 m. Furthermore, as shown in Fig. 10c, we applied the F-J method (Wang et al. 2019) to image the dispersion energy by processing the Rayleigh wave records. The dispersion curves (Fig. 10d) were then extracted by manually clicking on the energy image along the local energy extremes. In this section, the inversion performance of candidate algorithms is contrastively evaluated based on the extracted dispersion curve in Fig. 10d. The evaluation indicators involve the inversion computational cost, accuracy, stability, generalization ability, and effects of noise. Specifically, for each candidate algorithm, Sect. 4.1 evaluates the computational cost, accuracy, and stability of inversion; Sect. 4.2 shows the generalization ability regarding the designed search space; Sect. 4.3 examines the generalization ability considering the setting of the number of layers; Sect. 4.4 assesses the effects of noise. The parameter settings of candidate inversion algorithms adhere to the descriptions provided in Sect. 3.

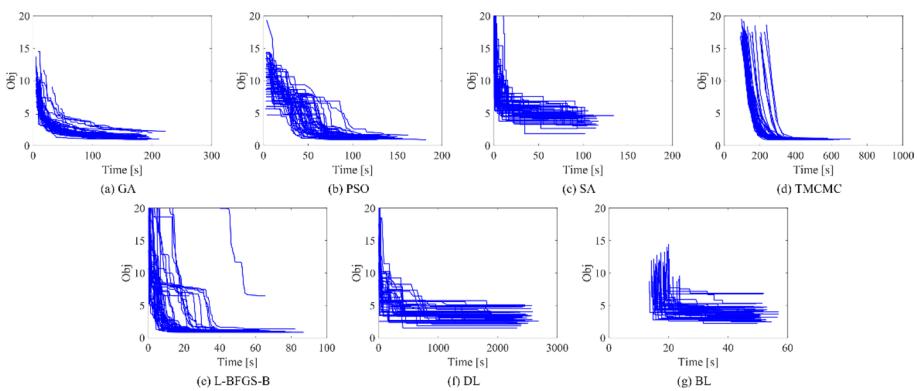


Fig. 11 Objective function curves of 50-time inversion of candidate algorithms based on the theoretical data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

Table 2 Quantification of the convergence ability of candidate inversion algorithms based on the theoretical data

	GA	PSO	SA	TMCMC	L-BFGS-B	DL	BL
Average inversion time (s)	178.0	131.7	102.4	527.0	57.8	2405.1	50.3
Average objective function value	1.42	1.12	4.06	0.94	1.32	3.27	3.78
Std of the objective function values	0.31	0.18	0.80	0.06	1.54	0.87	0.99

The time spent on network training has been factored into the overall time cost of inversion for both DL and BL networks; average inversion time refers to the average time consumption of 50 inversions; average objective function value refers to the average of the final objective function values of 50 inversions; std of the objective function values refers to the standard deviation of the final objective function values of 50 inversions; the value in a row highlighted in bold with underlines indicates the best performance among the seven algorithms

4.1 Comparison of Computational Cost, Accuracy, and Stability

This section evaluates the inversion accuracy, stability, and computational cost of seven candidate optimization algorithms using the theoretical dispersion curves depicted in Fig. 10d. Each algorithm was executed 50 times to provide a clear understanding of the inversion performance of the candidate algorithms. The identical search space was designed by the experience manner for each algorithm to ensure a fair comparison. Figure 11 displays the objective functions curves of 50-time inversion of each candidate algorithm. In terms of inversion time consumption, the GA, PSO algorithm, SA algorithm, L-BFGS-B algorithm, and BL network exhibit better performance than the remaining two algorithms. The BL network presents the best performance among these algorithms, with each single inversion taking less than 60 s. Specifically, the objective function curves for both the DL and BL networks were constructed by considering the misfit values of networks with various complexities, and the corresponding training times were recorded. The distinct contrast of the time consumption between the BL and DL networks verifies that the BL network can overcome the expensive computational cost problem of network training. Regarding the convergence value of the objective function, the 50 curves of the TMCMC algorithm present smaller

convergence values and a higher level of proximity to each other compared to the remaining six algorithms. We also quantified the convergence performance to provide a clearer understanding of the convergence characteristics of each algorithm, as shown in Table 2. The BL network achieves the lowest average computational cost for a single inversion among the candidate algorithms, while the L-BFGS-B algorithm demonstrates a comparable level of time consumption to the BL network. It is not surprising to observe such a phenomenon because these two algorithms are characterized by fast convergence. Correspondingly, the TMCMC algorithm yields the minimum average and the minimum std of the 50 convergence objective function values for inversion. The approximate average and std of the objective function values between the DL and BL networks demonstrate that these two types of neural networks possess equivalent mapping capability. Figure 12 shows the measured and inverted dispersion curves of 50 inversions for each candidate algorithm. The degree of fitting between inverted and measured dispersion curves for the candidate algorithms can match the objective function curves in Fig. 11 and the quantified data in Table 2. A smaller average objective function value indicates a better fitting degree between inverted and measured dispersion curves, and a smaller std of the objective function values suggests clusters closer to each other for 50 inverted dispersion curves. In particular, the approximate of these two values in Table 2 among the SA algorithm, DL network, and BL network produces the similarity of fitting degree between inverted and measured data; the inverted dispersion curves by the TMCMC algorithm provide a best fit to the observed data among candidate algorithms; some inverted dispersion curves of the L-BFGS-B algorithm poorly fit the measured ones, illustrating the disadvantage of being easily trapped in local optima.

In practice, the researchers are more concerned with the inverted S-wave velocity profiles than either the objective function curves or inverted dispersion curves. The strong non-uniqueness of Rayleigh wave inversion may result in a situation where a smaller convergence objective function value or a better data fitting degree cannot always produce a more accurate S-wave velocity profile. The crucial aspect in evaluating the inversion performance of the candidate algorithms is to compare the accuracy and stability of the inverted S-wave velocity profiles. Figure 13 depicts the inverted S-wave velocities of 50 inversions for each candidate algorithm under a unique search space. By observing this figure, the TMCMC algorithm, DL network, and BL network acquire more satisfactory

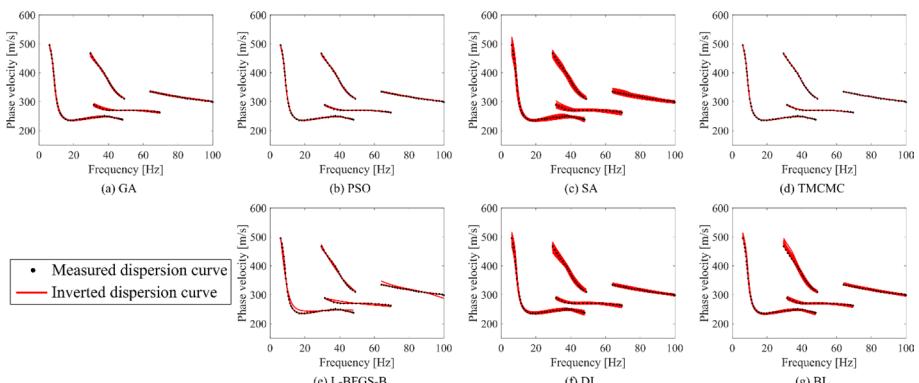


Fig. 12 Inverted dispersion curves of 50-time inversion of candidate algorithms based on the theoretical data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

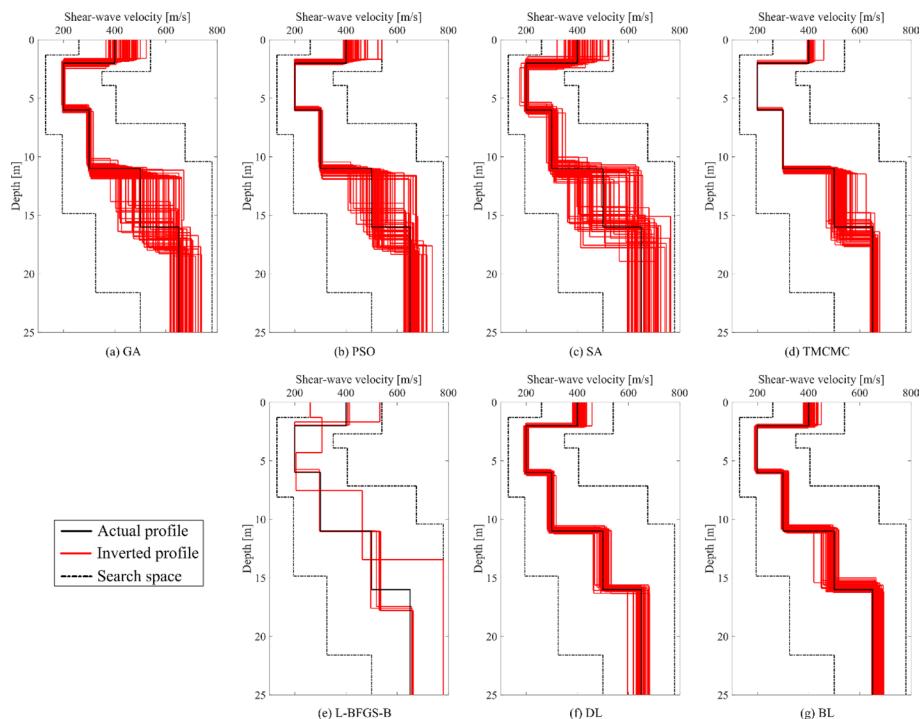


Fig. 13 Inverted S-wave velocity profiles of 50-time inversion of candidate algorithms based on the theoretical data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

inversion results, reflected in that the 50 inverted profiles by these three algorithms are more concentrated in the area around the actual one than the remaining algorithms. Such a phenomenon violates the quantized data in Table 2, in which the DL and BL networks only present a medium level of convergence capability toward the objective function among candidate seven algorithms. Particularly, the GA and PSO algorithms both provide worse estimations of S-wave velocity profiles than the DL network and BL one though formers obtain better average and std values of the objective function. It is worth noting that there are two commonalities among the TMCMC algorithm, DL network, and BL network. The first commonality comes at they are all large-sample-size-based optimization algorithms. In this study, the TMCMC algorithm is equipped with 2000 samples in each sampling stage and the DL and BL networks are both provided 500 input–output sample pairs for network training. On the contrary, the remaining four algorithms only require a few samples to support iteration. The second commonality reflects in that these three algorithms can comprehensively utilize the sample values. Specifically, the TMCMC algorithm treats the average values of the final stage samples as the optimal layer thicknesses and S-wave velocities of different layers; the DL and BL networks utilize all input–output sample pairs to establish mapping networks for the inversion task. The comprehensive utilization of a large number of samples allows the optimization algorithm to focus on exploring the functional relationship within the parameter space, rather than searching for a single optimal point. When the optimization problem exhibits strong non-uniqueness, namely, the case with Rayleigh

Table 3 Quantification of the inversion accuracy of candidate inversion algorithms based on the theoretical data

	$h^{(1)}$	$h^{(2)}$	$h^{(3)}$	$h^{(4)}$	$V_S^{(1)}$	$V_S^{(2)}$	$V_S^{(3)}$	$V_S^{(4)}$	$V_S^{(5)}$	Average
GA	9.9%	2.7%	7.7%	19.1%	10.2%	0.9%	1.5%	14.3%	4.0%	7.8%
PSO	6.0%	0.9%	6.4%	22.7%	7.3%	0.2%	0.8%	11.7%	3.0%	6.6%
SA	10.0%	4.7%	8.7%	12.9%	9.7%	2.1%	3.6%	16.3%	5.8%	8.2%
TMCMC	2.1%	0.5%	2.7%	14.3%	2.2%	0.2%	0.4%	6.1%	1.1%	3.3%
L-BFGS-B	3.9%	2.4%	4.0%	33.5%	5.0%	4.5%	2.9%	6.4%	3.1%	7.3%
DL	2.3%	2.4%	2.3%	2.2%	3.5%	2.0%	2.3%	2.3%	2.0%	2.4%
BL	2.6%	3.1%	3.2%	3.3%	3.3%	2.4%	2.1%	4.2%	2.7%	3.0%

Each cell in columns 2 to 10 of the table represents the MAPE between 50 inverted values and actual value of an earth model parameter for a certain algorithm; the last column of the table represents the average value of columns 2 to 10; a smaller value indicates higher inversion accuracy; the value in a row highlighted in bold with underlines indicates the best performance among the seven algorithms

Table 4 Quantification of the inversion stability of candidate inversion algorithms based on the theoretical data

	$h^{(1)}$	$h^{(2)}$	$h^{(3)}$	$h^{(4)}$	$V_S^{(1)}$	$V_S^{(2)}$	$V_S^{(3)}$	$V_S^{(4)}$	$V_S^{(5)}$	Average
GA	9.7%	3.2%	7.1%	18.3%	8.6%	1.2%	1.8%	15.6%	4.6%	7.8%
PSO	6.4%	1.1%	5.8%	19.7%	8.5%	0.3%	1.0%	12.9%	3.6%	6.6%
SA	12.5%	6.4%	10.0%	15.7%	11.0%	3.0%	4.4%	18.6%	6.4%	9.8%
TMCMC	2.8%	0.3%	2.4%	13.2%	2.5%	0.1%	0.4%	6.3%	1.2%	3.2%
L-BFGS-B	9.9%	7.0%	10.1%	3.5%	11.2%	14.1%	8.8%	3.7%	4.9%	8.1%
DL	3.2%	2.7%	3.1%	3.0%	4.1%	2.4%	2.7%	3.0%	2.7%	3.0%
BL	2.9%	2.3%	3.5%	4.2%	3.9%	1.1%	2.4%	4.1%	1.8%	2.9%

Each cell in columns 2 to 10 of the table represents the std divided by the average of 50 inverted values of an earth model parameter for a certain algorithm; the last column of the table represents the average value of columns 2 to 10; a smaller value indicates stronger inversion stability; the value in a row highlighted in bold with underlines indicates the best performance among the seven algorithms

wave inversion, large-sample-size-based optimization algorithms can effectively mitigate the influence of local optima and converge toward a solution that approximates the globally optimal value. In contrast, the GA, PSO algorithm, SA algorithm, and L-BFGS-B algorithm emphasize the iteration process of finding a single optimal point, which makes them more prone to getting trapped in local optima.

The accuracy and stability of inverted S-wave velocity profiles were also quantified to clarify the inversion performance of candidate algorithms. Table 3 reveals the quantification of the inversion accuracy of each algorithm by calculating the mean absolute percentage error (MAPE) between inverted and actual earth model parameter values. Obviously, the inversion process achieved smaller indicator values of quantifying inversion accuracy by the TMCMC algorithm, DL network, and BL network than other algorithms. The TMCMC algorithm exhibits best in the estimation of six earth model parameters, and the DL network shows best for estimating the remaining earth model parameters. The DL network obtains the best overall performance by comparing the average of the accuracy

indicator values. Notably, the TMCMC algorithm and BL network also determine similar average accuracy indicator values compared with the DL network. Table 4 quantifies the inversion stability of candidate algorithms. It normalizes the std of 50 inverted values for a given earth model parameter to represent the stability indicators. The normalization way refers to dividing the std by the average of 50 inverted values of an earth mode parameter for a candidate algorithm. The TMCMC algorithm and DL network acquire the best performance of inversion stability in seven and two earth model parameters, respectively. The BL network presents the best overall performance of inversion stability in the estimation of all earth model parameters, as it determines the smallest average value in Table 4. In particular, similar to Table 3, the average values of the TMCMC algorithm, DL network, and BL network in Table 4 are very close. The quantized data in Table 3 and 4 indicate that these three algorithms exhibit a comparable level of inversion performance, notably superior to the other four algorithms regarding the accuracy and stability of the inverted S-wave velocity profiles.

4.2 Comparison of Generalization Ability Regarding a Wrong Search Space

The candidate algorithms require an initial search space to perform constrained optimization for Rayleigh wave inversion. The numerical test in Sect. 4.1 is conducted under an implicit assumption, i.e., the design of a suitable search space before inversion, which ensures that the actual model parameter values are encompassed within the search space. Under such an assumption, the GA, PSO algorithm, SA algorithm, TMCMC algorithm, and L-BFGS-B algorithm can perform iterations within the correct search space. Similarly, the DL and BL networks can generate training samples within the correct space to model the unknown functional mapping. As a result, we can observe that the majority of inverted dispersion curves can well fit the measured ones, and most inverted S-wave velocity profiles accurately match the actual ones. In practice, the complex materials of subsurface may lead to a wrong search space designed before inversion. To better reveal the inversion performance of each algorithm, this section investigates their generalization abilities when faced with a slightly incorrect search space for the inversion task. Specifically, the upper boundary of the S-wave velocity in the last layer is purposely set as a smaller value that cannot cover the actual value.

In this section, the objective function curves and inverted dispersion curves are omitted from the main text for the purpose of simplification. These curves, generated by each algorithm, are shown in Appendix 1. Figure 14 shows the inverted S-wave velocity profiles of 50 inversions by all algorithms under a search space with slight deviations. We can clearly observe that the actual S-wave velocity of the last layer is not covered by the search space for each algorithm. The candidate algorithms showcase two distinct responses when faced with the artificial setting where the upper boundary of the last layer's velocity is slightly lower than the actual value. The inverted S-wave velocities in the last layer by the GA, PSO algorithm, SA algorithm, TMCMC algorithm, and L-BFGS-B algorithm are enclosed by the designed search space, determining unsatisfactory S-wave velocity estimations. In stark contrast, the inverted S-wave velocities of the last layer obtained by the DL and BL networks exceed the limitations of the incorrect search space. Surprisingly, the estimated velocities in the last layer by the BL network are all located outside of the search space. As a result, the DL and BL networks produce more accurate S-wave velocity profiles compared with the other five algorithms. The distinct differences in inverting the last layer's velocity demonstrate the

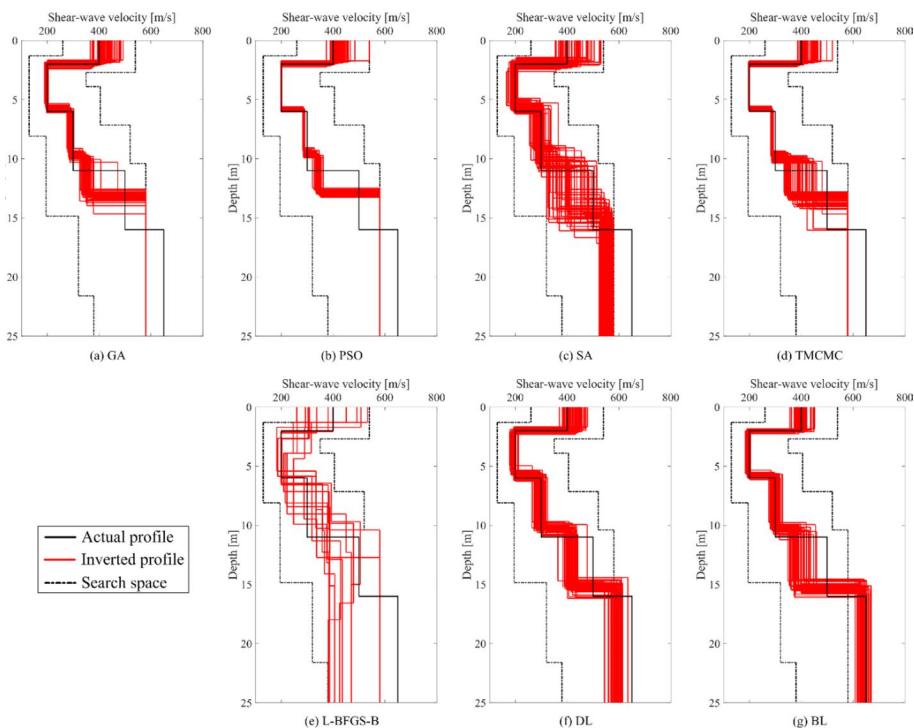


Fig. 14 Inverted S-wave velocity profiles of 50-time inversion of candidate algorithms based on the theoretical data under a wrong search space: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

varied generalization capability of these algorithms when confronted with a wrongly designed search space. The two types of neural networks feature better generalization capability than other algorithms. It is because the search space has different effects on these optimization algorithms. For the DL and BL networks, the search space only controls the numerical ranges of the training samples instead of the ones of the earth model parameters to be inverted, providing possibilities of enabling these two algorithms to be equipped with generalization ability to the search space. The mechanism of modeling by these two neural networks comes in simulating the functional relationship by exploring the one among training samples. When the training samples slightly deviate the parameter values to be predicted, the trained networks still can provide satisfactory estimations of S-wave velocity, as observed in Fig. 14. On the contrary, regarding the other five algorithms, the search space constrains the lower and upper boundaries of unknown parameter values to be searched within the iterations, leading to unsatisfactory generalization capability to the search space. Table 5 attempts to quantify the inversion accuracy to further compare the generalization capability of the search space of candidate algorithms. Most algorithms produce higher accuracy indicator values (namely, less accurate estimations) in four parameters by the wrong search space than by the correct one, including the layer thicknesses in the third and fourth layer and the S-wave velocities in the fourth and fifth layer. Simultaneously, the average values of the accuracy indicators all show an increase under the wrong search space compared to the correct

Table 5 Quantification of the inversion accuracy of candidate inversion algorithms based on the theoretical data under a wrong search space

	$h^{(1)}$	$h^{(2)}$	$h^{(3)}$	$h^{(4)}$	$V_S^{(1)}$	$V_S^{(2)}$	$V_S^{(3)}$	$V_S^{(4)}$	$V_S^{(5)}$	Average
GA	7.9%	4.4%	21.3%	30.1%	8.7%	1.2%	4.1%	29.1%	10.8%	13.1%↑
PSO	5.2%	<u>3.1%</u>	24.6%	33.9%	6.6%	<u>0.4%</u>	4.4%	31.8%	10.8%	13.4%↑
SA	11.4%	10.0%	14.8%	14.8%	10.7%	4.5%	5.2%	18.4%	14.4%	11.6%↑
TMCMC	5.7%	3.8%	19.5%	26.0%	7.6%	0.8%	4.7%	24.2%	10.9%	11.5%↑
L-BFGS-B	10.7%	9.8%	27.7%	32.6%	7.9%	8.8%	10.0%	28.9%	15.9%	16.9%↑
DL	5.9%	8.2%	<u>8.3%</u>	<u>4.7%</u>	6.5%	4.0%	4.3%	<u>17.1%</u>	8.7%	7.5%↑
BL	<u>4.4%</u>	6.3%	8.9%	6.5%	<u>4.2%</u>	2.5%	<u>3.5%</u>	21.2%	<u>2.0%</u>	<u>6.6%</u> ↑

Each cell in columns 2 to 10 of the table represents the MAPE between 50 inverted values and actual value of an earth model parameter for a certain algorithm; the last column of the table represents the average value of columns 2 to 10; a smaller value indicates higher inversion accuracy; the value in a row highlighted in bold with underlines indicates the best performance among the seven algorithms; the upward arrow indicates an increase in the value compared to the one in Table 3; on the contrary, the downward arrow means a decrease

Table 6 Quantification of the inversion stability of candidate inversion algorithms based on the theoretical data under a wrong search space

	$h^{(1)}$	$h^{(2)}$	$h^{(3)}$	$h^{(4)}$	$V_S^{(1)}$	$V_S^{(2)}$	$V_S^{(3)}$	$V_S^{(4)}$	$V_S^{(5)}$	Average
GA	9.3%	4.5%	8.2%	9.4%	7.2%	1.5%	1.8%	6.5%	<u>0.03%</u>	5.4%↓
PSO	5.8%	<u>0.9%</u>	7.0%	<u>2.4%</u>	6.7%	<u>0.3%</u>	<u>0.5%</u>	<u>2.4%</u>	<u>0.03%</u>	<u>2.9%</u> ↓
SA	13.9%	10.5%	15.1%	17.5%	11.1%	5.2%	7.2%	13.6%	2.9%	10.8%↑
TMCMC	6.3%	1.3%	7.4%	15.1%	6.9%	0.6%	0.8%	10.6%	0.07%	5.5%↑
L-BFGS-B	17.1%	13.8%	13.8%	22.3%	15.2%	17.7%	13.9%	10.8%	12.5%	15.2%↑
DL	7.4%	7.6%	6.9%	5.5%	6.2%	3.9%	5.0%	4.5%	3.2%	5.6%↑
BL	<u>4.9%</u>	4.4%	<u>6.3%</u>	7.3%	<u>5.4%</u>	2.0%	4.2%	6.5%	2.2%	4.8%↑

Each cell in columns 2 to 10 of the table represents the std divided by the average of 50 inverted values of an earth model parameter for a certain algorithm; the last column of the table represents the average value of columns 2 to 10; a smaller value indicates stronger inversion stability; the value in a row highlighted in bold with underlines indicates the best performance among the seven algorithms; the upward arrow indicates an increase in the value compared to the one in Table 4; on the contrary, the downward arrow means a decrease

one, confirming the dominant role of the search space in accurately estimating S-wave velocities. Significantly, the BL network demonstrates a remarkable performance in estimating the S-wave velocity in the last layer, with the corresponding accuracy indicator value being as low as 2.0%. The BL and DL networks occupy the top two positions in terms of the average accuracy indicator values among the candidate algorithms. Table 6 lists the inversion stability indicator values of candidate algorithms under the wrong search space. The PSO algorithm and BL network stand out as the top performers in terms of inversion stability among candidate algorithms. In contrast with the stability indicator values in Table 4, the GA and PSO algorithms achieve smaller stability indicator values by a wrong search space, but the other five obtain larger ones. It illustrates that a wrong search space may result in more stable inversion results but with less

accuracy. Considering both the indicator values in Tables 5 and 6, the BL network presents the best generalization ability to the wrong search space, followed by DL network.

4.3 Comparison of Generalization Ability Regarding a Wrong Number of Layers

This study inverts the layer thicknesses and S-wave velocities of different layers assuming the known P-wave velocities and densities for Rayleigh wave inversion. Besides the initial search space discussed in Sect. 4.2, the number of layers is also an implicit parameter as it governs the number of unknowns of the optimization problem inside the inversion process. The comparison of inversions by candidate algorithms in Sect. 4.1 is fulfilled in the presence of the known number of layers. However, the complex materials of the subsurface and the usual lack of borehole data may lead to the number of layers cannot be accurately pre-estimated before inversion. A higher number of layers setting is a common approach to address this problem. In this subsection, such a situation is considered to ensure a comprehensive comparison of candidate inversion algorithms. Specifically, in this case, a setting of eight layers is artificially adopted before inversion to compare the generalization ability of candidate algorithms regarding a wrong number of layers.

Under the setting of eight layers, the initial search space in Table 7 was designed to conduct dispersion curve inversion for each algorithm. For the sake of simplicity, the main text does not include the objective function curves and inverted dispersion curves of candidate algorithms. In particular, we can refer to Appendix 2 for more information on these curves. Figure 15 depicts 50 S-wave velocity profiles inverted by each algorithm based on the theoretical dispersion curves under the setting of eight layers. Due to the overlapping of lower and upper boundaries of layer thicknesses within the 2-dimensional image, this figure does not provide a visualization of the search space. It is clear to see the significant differences among the inverted profiles by different algorithms, reflected in two aspects, i.e., the level of aggregation among the 50 curves and the deviation of these curves from the actual ones. The DL and BL networks exhibit better inversion performance than the other five algorithms in these two aspects. However, there is a shared feature among the profiles inverted by the candidate algorithms when comparing these profiles with the actual one: it appears that the second, third, and fourth layers inside the inverted profiles try to divide a single layer into two layers. The velocities of the two layers split in the second, third, or fourth layer almost coincide for the DL and BL networks, indicating the powerful generalization ability of these two neural networks facing the wrong number of layers.

Table 7 Designed search space for candidate inversion algorithms based on the theoretical data under a wrong number of layers

Layer	$h^{\min}(\text{m})$	$h^{\max}(\text{m})$	$v_S^{\min}(\text{m/s})$	$v_S^{\max}(\text{m/s})$
1	1.3	2.7	260	540
2	1.3	2.7	130	350
3	1.3	2.7	130	350
4	1.3	2.7	195	405
5	2.0	4.1	195	405
6	1.3	2.7	325	675
7	2.0	4.1	325	675
8	N/A	N/A	500	780

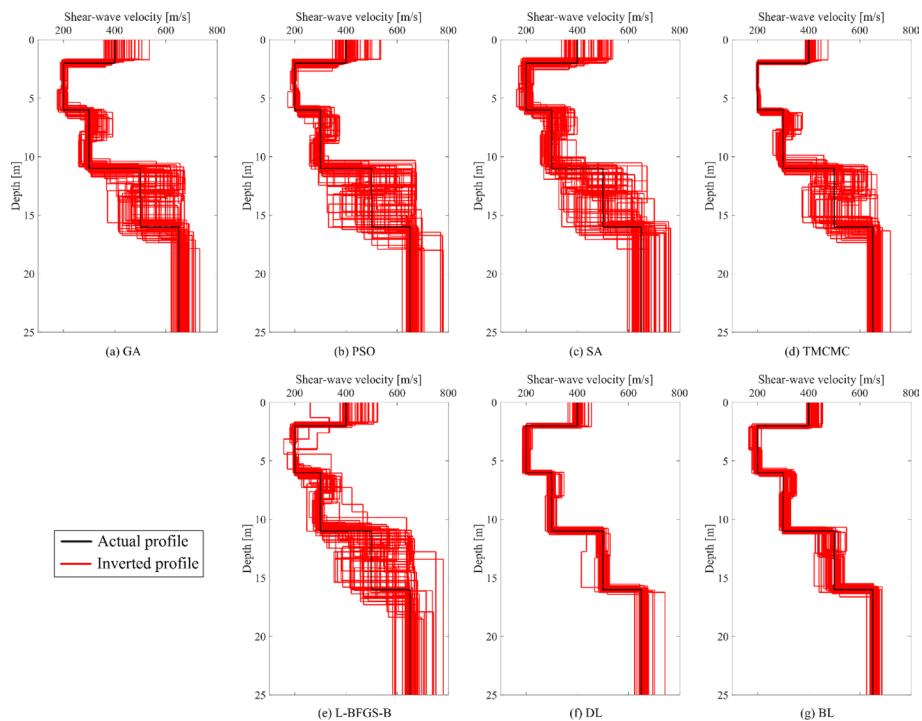


Fig. 15 Inverted S-wave velocity profiles of 50-time inversion of candidate algorithms based on the theoretical data under a wrong number of layers: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

Table 8 Quantification of the inversion accuracy of candidate inversion algorithms based on the theoretical data under a wrong number of layers

	GA	PSO	SA	TMCMC	L-BFGS-B	DL	BL
Average MAPE between inverted and actual profiles	14.6%	14.7%	15.0%	13.0%	15.0%	11.7%	12.3%

For a given algorithm, average MAPE refers to the average among the 50 MAPE values between inverted and actual profiles

To make a better comparison of the generalization ability among candidate algorithms, the inversion accuracy was tried to quantified by calculating the average MAPE between the inverted and actual S-wave velocity profiles, as shown in Table 8. Note that the inversion accuracy quantification way is different from the one in Sect. 4.1 due to the setting of the wrong number of layers. Table 8 shows that the DL and BL networks determine a lower average MAPE between the inverted and actual profiles than the other five algorithms. This indicates that these two neural networks can determine S-wave velocity profiles with greater accuracy compared to other algorithms, which is consistent with the results shown in Fig. 15. Concurrently, the inversion stability was also quantified by adopting the quantification manner utilized in Sect. 4.1. Table 9 lists

Table 9 Quantification of the inversion stability of candidate inversion algorithms based on the theoretical data under a wrong number of layers

	$h^{(1)}$	$h^{(2)}$	$h^{(3)}$	$h^{(4)}$	$h^{(5)}$	$h^{(6)}$	$h^{(7)}$	$V_s^{(1)}$	$V_s^{(2)}$	$V_s^{(3)}$	$V_s^{(4)}$	$V_s^{(5)}$	$V_s^{(6)}$	$V_s^{(7)}$	$V_s^{(8)}$	Average
GA	9.5%	15.3%	13.5%	18.7%	17.7%	19.5%	16.2%	9.7%	3.3%	4.6%	9.3%	4.7%	16.5%	15.0%	3.4%	11.8%↑
PSO	11.0%	21.5%	19.0%	23.4%	23.7%	22.1%	23.7%	10.3%	2.4%	3.9%	8.9%	4.7%	20.1%	19.7%	5.0%	14.6%↑
SA	13.6%	14.0%	18.2%	18.6%	15.9%	15.9%	16.9%	13.8%	6.6%	7.9%	11.1%	9.1%	20.8%	18.3%	6.9%	13.8%↑
TMCMC	4.4%	13.3%	11.9%	14.6%	15.2%	16.1%	14.0%	4.0%	0.9%	1.9%	6.3%	3.5%	13.3%	14.0%	2.7%	9.1%↑
L-BFGS-B	9.9%	28.8%	26.6%	29.0%	22.6%	29.2%	28.4%	12.7%	13.8%	9.6%	12.0%	7.6%	19.7%	17.1%	6.0%	18.2%↑
DL	3.5%	3.4%	2.6%	2.3%	2.8%	2.2%	2.1%	4.3%	4.0%	3.9%	3.1%	2.7%	3.5%	2.9%	3.2%	↑
BL	3.7%	4.8%	4.6%	4.6%	4.9%	5.3%	5.1%	3.7%	4.6%	3.9%	4.3%	3.5%	5.5%	5.0%	1.8%	4.4%↑

Each cell in columns 2 to 10 of the table represents the std divided by the average of 50 inverted values of an earth model parameter for a certain algorithm; the last column of the table represents the average value of columns 2 to 10; a smaller value indicates stronger inversion stability; the value in a row highlighted in bold with underlines indicates the best performance among the seven algorithms; the upward arrow indicates an increase in the value compared to the one in Table 4; on the contrary, the downward arrow means a decrease

the stability indicator values of candidate algorithms under a wrong number of layers. The DL network achieves the minimum stability indicator value among the majority of earth model parameters, followed by the BL network and the TMCMC algorithm. Considering the average stability indicator values of all earth model parameters, the DL and BL networks determine smaller values than other algorithms. Notably, the DL network demonstrates a remarkable level of stability with an average indicator value of only 3.2%, which is the smallest among various algorithms and only 0.2 higher than the value shown in Table 4. The quantification results in Table 9 verifies that the DL and BL networks possess strong generalization ability regarding a wrong number of layers for Rayleigh wave inversion. Meanwhile, the average stability indicator values of the candidate algorithms were compared between the inversions under the incorrect and correct number of layers. Each algorithm provides a higher value of stability indicator under the incorrect number of layers than the correct one, illustrating that the higher number of layers than the actual situation leads to less stable inversion results.

4.4 Comparison of Effects of Noise

The dispersion curves shown in Fig. 10d were artificially perturbed by adding 2.5% root-mean-square noise for evaluating the effects of noise on the candidate algorithms. Each algorithm runs 50 times under the same conditions as Sect. 4.1 except for the presence of

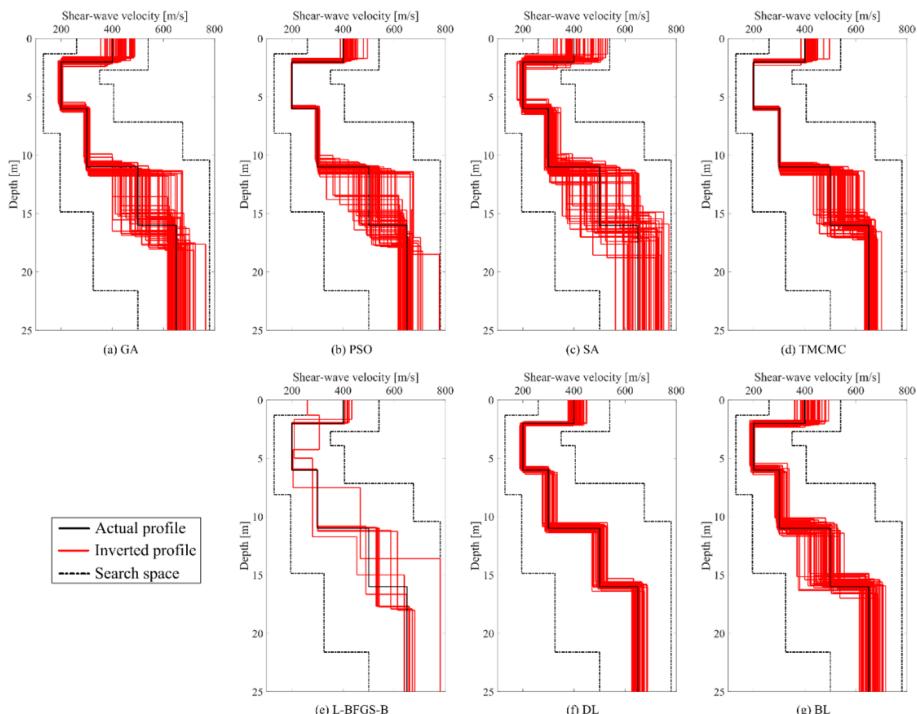


Fig. 16 Inverted S-wave velocity profiles of 50-time inversion of candidate algorithms based on noisy theoretical data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

noise in the observed data. The objective function curves and inverted dispersion curves are presented in Appendix 3, which is not included in the main text for the sake of simplicity. Figure 16 depicts the 50 inverted S-wave velocity profiles by each algorithm under noisy data. It is obvious to observe that the TMCMC algorithm, DL network, and BL network provide better S-wave velocity estimations than the other four algorithms, with the DL network showcasing the best performance. The GA, PSO algorithm, and SA algorithm exhibit similar inversion performance. The L-BFGS-B algorithm presents some individual poor profile results, caused by its susceptibility to converging to local optima. By considering Fig. 13 collectively, the introduction of noise deteriorates the stability of Rayleigh wave inversion for each algorithm. To better analyze the inversion performance of different algorithms, Table 10 lists the MAPE between inverted and actual earth model parameters of different layers to quantify the inversion accuracy. Similar to the noise-free inversion case discussed in Sect. 4.1, the TMCMC algorithm, DL network, and BL network present superior performance with lower accuracy indicator values in contrast with other algorithms. The deep learning network and TMCMC algorithm rank first and second in determining inversion accuracy indicators. In particular, the second one in Sect. 4.1 corresponds to the BL network rather than the TMCMC algorithm, illustrating the worse noise resilience of the BL network than the TMCMC algorithm and DL network. Moreover, Table 11 presents the std divided by the average of 50 inverted values of an earth model parameter to denote the quantification of the inversion stability of candidate inversion algorithms. The top three algorithms with the best inversion stability performance are the DL network, TMCMC algorithm, and BL network, being the same as the top three rankings of inversion accuracy quantification in Table 4. In Table 10 and 11, each average stability indicator value is annotated with an upward/downward arrow or a black dot. In contrast with the inversion test without artificial noise in dispersion curves, for the inversion in this section, an average indicator value marked with an upward arrow indicates an increase in the value, a marked with a downward arrow signifies a decrease, and a black dot means remaining the same for the inversion using 2.5% noisy data in this section. There are only 3 downward arrows out of the total 14 arrows in these two tables. Overall, noise tends to decrease both the

Table 10 Quantification of the inversion accuracy of candidate inversion algorithms based on noisy theoretical data

	$h^{(1)}$	$h^{(2)}$	$h^{(3)}$	$h^{(4)}$	$V_S^{(1)}$	$V_S^{(2)}$	$V_S^{(3)}$	$V_S^{(4)}$	$V_S^{(5)}$	Average
GA	10.0%	3.6%	7.3%	18.1%	9.4%	1.4%	1.7%	12.9%	4.1%	7.6%↓
PSO	5.6%	1.1%	4.8%	19.6%	6.1%	0.5%	0.9%	12.6%	2.8%	6.0%↓
SA	10.2%	5.6%	12.0%	14.6%	12.2%	2.5%	4.2%	15.7%	7.4%	9.4%↑
TMCMC	4.7%	0.7%	4.9%	10.7%	7.1%	0.5%	0.6%	9.6%	2.0%	4.5%↑
L-BFGS-B	3.6%	2.5%	3.3%	33.8%	6.2%	4.0%	2.8%	7.0%	2.5%	7.3%*
DL	2.8%	3.1%	3.0%	1.6%	4.0%	2.3%	2.9%	2.3%	2.0%	2.7%↑
BL	5.7%	4.4%	5.9%	5.5%	7.0%	2.8%	3.0%	8.2%	3.1%	5.1%↑

Each cell in columns 2 to 10 of the table represents the MAPE between 50 inverted values and actual value of an earth model parameter for a certain algorithm; the last column of the table represents the average value of columns 2 to 10; a smaller value indicates higher inversion accuracy; the value in a row highlighted in bold with underlines indicates the best performance among the seven algorithms; the black dot represents remaining the same in the value compared to the one in Table 3; the upward arrow indicates an increase in the value compared to the one in Table 3; on the contrary, the downward arrow means a decrease

Table 11 Quantification of the inversion stability of candidate inversion algorithms based on noisy theoretical data

	$h^{(1)}$	$h^{(2)}$	$h^{(3)}$	$h^{(4)}$	$V_s^{(1)}$	$V_s^{(2)}$	$V_s^{(3)}$	$V_s^{(4)}$	$V_s^{(5)}$	Average
GA	11.1%	4.4%	7.9%	18.5%	7.3%	1.8%	2.0%	15.2%	4.9%	8.1%↑
PSO	6.1%	1.4%	6.7%	21.0%	5.0%	0.5%	1.0%	16.0%	4.3%	6.9%↑
SA	13.4%	7.3%	13.7%	17.2%	13.4%	3.5%	5.0%	18.5%	7.2%	11.0%↑
TMCMC	4.6%	1.0%	4.6%	13.0%	4.0%	0.3%	0.7%	9.5%	2.2%	4.4%↑
L-BFGS-B	8.7%	6.6%	10.1%	8.0%	9.3%	12.7%	7.7%	4.7%	4.5%	8.0%↓
DL	3.8%	3.9%	3.8%	2.1%	4.5%	2.8%	4.0%	3.0%	2.5%	3.4%↑
BL	6.2%	4.3%	7.7%	6.9%	7.4%	2.1%	3.8%	9.5%	3.6%	5.7%↑

Each cell in columns 2 to 10 of the table represents the std divided by the average of 50 inverted values of an earth model parameter for a certain algorithm; the last column of the table represents the average value of columns 2 to 10; a smaller value indicates stronger inversion stability; the value in a row highlighted in bold with underlines indicates the best performance among the seven algorithms; the upward arrow indicates an increase in the value compared to the one in Table 4; on the contrary, the downward arrow means a decrease

accuracy and stability of the inversion. The indicator values of most algorithms only show a limited increase, indicating that most algorithms exhibit a certain level of noise resilience in Rayleigh wave inversion. Notably, the TMCMC algorithm, DL network, and BL network demonstrate remarkable noise resilience.

5 Field Data Application

This section attempts to reveal the inversion performance of candidate algorithms faced with the field dispersion curve data. The InterPACIFIC project (Garofalo et al. 2016a, 2016b) conducted an extensive survey of Rayleigh waves at the Mirandola site in Italy, and the field data are publicly available for further research. Yang et al. (2023) processed the raw data of Rayleigh waves measured by the InterPACIFIC project at the Mirandola site to image the dispersion energy and then successfully extracted multi-modal dispersion curves for the inversion test. This study applies the dispersion curve data in Yang et al. (2023) to evaluate the field data processing capability of the candidate algorithms. Figure 17 depicts the extracted passive and active Rayleigh wave dispersion curves based on the field data at Mirandola site. Figure 17a–c displays the dispersion energy imaging results obtained from three separate active surveys. Each survey utilized 48 geophones and a vertical hammer source. A 15-m minimum source offset with a 1-m spatial interval, a 3-m minimum source offset with a 1-m spatial interval, and a 4-m minimum source offset with a 2-m spatial interval were adopted for the surveys in Fig. 17a–c, respectively. Figure 17d presents the dispersion energy imaging results obtained from the passive survey. In this survey, 15 geophones were deployed azimuthally to form two circles, one with a radius of approximately 26 m and another with a radius of around 78 m. Furthermore, the active and passive dispersion curves were extracted based on the dispersion energy images in Fig. 17a–d, as shown in Fig. 17e.

Each candidate algorithm was applied to perform inversion 50 times based on the multi-modal dispersion curves shown in Fig. 17e. The assumption of 5 layers was adopted for the inversions by each algorithm. Figure 18 illustrates the inverted objective function curves

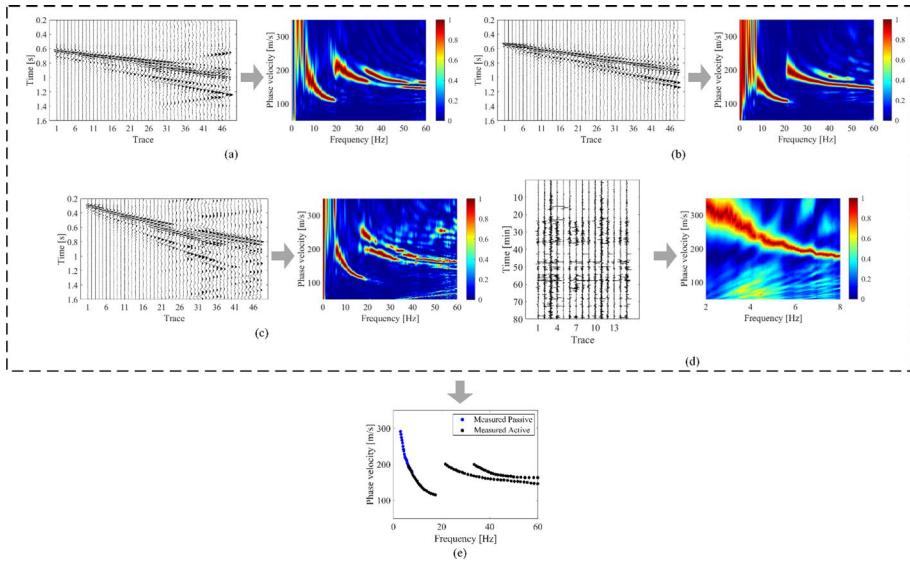


Fig. 17 Extracted passive and active Rayleigh wave dispersion curves based on the field data at Mirandola site (Yang et al. 2023): **a** imaging dispersion curves based on active seismic wave data 1, **b** imaging dispersion curves based on active seismic wave data 2, **c** imaging dispersion curves based on active seismic wave data 1, **d** imaging dispersion curves based on passive seismic wave data, **e** extracted dispersion curves based on dispersion images in **(a)–(d)**

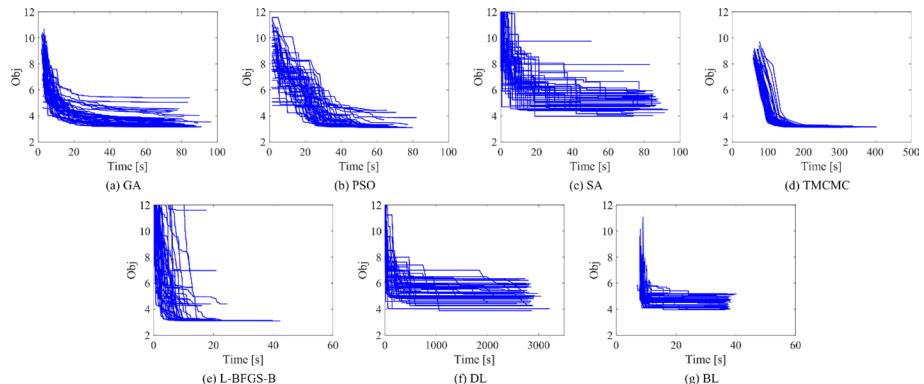


Fig. 18 Objective function curves of 50-time inversion of candidate algorithms based on the field data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

of 50 inversions by the candidate algorithms under the field data. These algorithms present varying convergence capabilities facing the same filed data, reflected in different time consumptions and different positions in the vertical axis of objective function curves. Similar to the case in Sect. 4.1, the objective function curves of the TMCMC algorithm converge to smaller values than the ones of other algorithms. Table 12 quantifies the convergence capability of the candidate algorithm for the inversion process. Regarding the computational

Table 12 Quantification of the convergence ability of candidate inversion algorithms based on the field data

	GA	PSO	SA	TMCMC	L-BFGS-B	DL	BL
Average inversion time (s)	80.7	67.9	80.4	296.1	23.2	2783.4	37.3
Average objective function value	3.53	3.29	5.22	3.17	4.22	5.12	4.58
Std of the objective function values	0.50	0.35	0.99	0.04	2.68	0.61	0.32

The time spent on network training has been factored into the overall time cost of inversion for both DL and BL networks; average inversion time refers to the average time consumption of 50 inversions; average objective function value refers to the average of the final objective function values of 50 inversions; std of the objective function values refers to the standard deviation of the final objective function values of 50 inversions; the value in a row highlighted in bold with underlines indicates the best performance among the seven algorithms

cost, the L-BFGS-B algorithm and the BL network rank in the first and second positions, respectively, followed by three algorithms with approximately equivalent levels of time consumption, namely the PSO algorithm, SA algorithm, and GA. The TMCMC algorithm experiences significantly higher time consumption than these algorithms and the issue of time-consuming stands out during the inversion process using the DL network. Among the candidate algorithms, the TMCMC algorithm demonstrates the smallest values for the average objective function value and the std of the objective function values. The differences among other algorithms in these two performance indicators are not significant. Figure 19 shows the inverted and measured dispersion curves of candidate algorithms. The level of agreement between the inverted and measured dispersion curves for the candidate algorithms can match the objective function curves in Fig. 18 and the quantified data in Table 12. Notably, some inverted dispersion curves generated by the GA, SA, and L-BFGS-B algorithms exhibit inadequate fitting to the measured ones. It should be due to these algorithms getting trapped in local optima during corresponding iterations.

The accuracy and stability of the S-wave velocity profiles are the main concerns when evaluating the inversion performance among the candidate algorithms. Figure 20 presents the inverted S-wave velocity profiles of 50 inversions by candidate algorithms based on the field data. The inverted S-wave velocities locate inside the search space, especially for

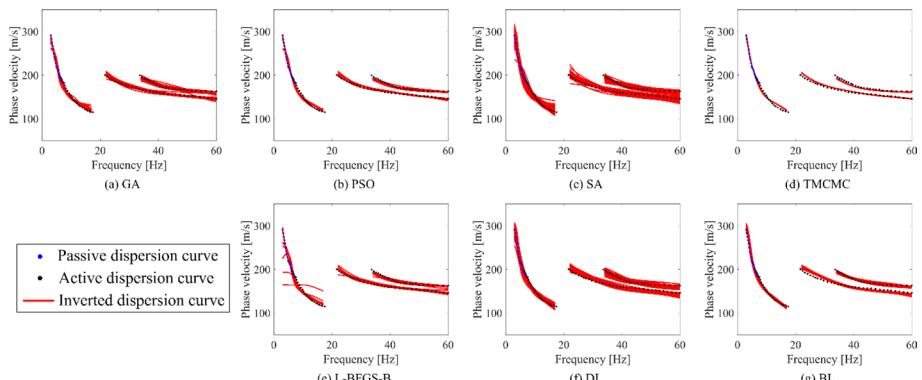


Fig. 19 Inverted dispersion curves of 50-time inversion of candidate algorithms based on the field data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

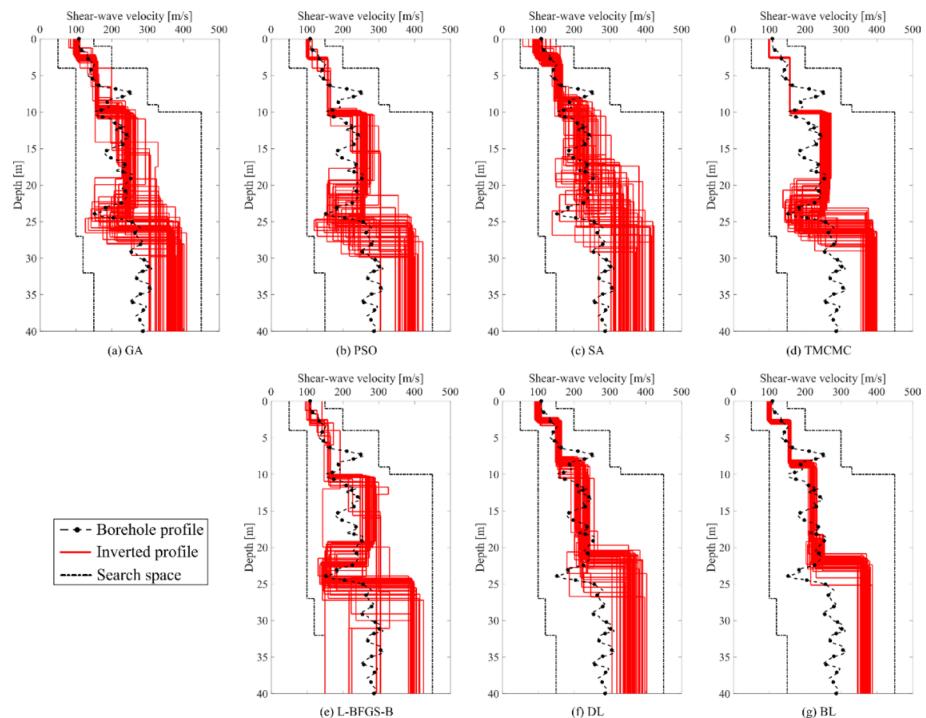


Fig. 20 Inverted S-wave velocity profiles of 50-time inversion of candidate algorithms based on the field data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

the DL and BL networks, demonstrating the designed initial search space is suitable for launching the inversion task. There are significant differences among the algorithms in the S-wave velocity profiles of the third and fourth layers. Using the TMCMC algorithm and BL network as examples, it is observed that the TMCMC algorithm assigns noticeably different velocities to these two layers; on the contrary, the BL network appears to merge these two layers into a single layer. This results in that the velocities inverted by the BL network better match the borehole profile between the depths of 10 m and 20 m but fit the one between the depths of 20 m and 25 m less well compared to the TMCMC algorithm. Meanwhile, the stability levels of inverted S-wave velocity profiles exhibit notable variations among different algorithms. The TMCMC algorithm, DL network, and BL network are superior to other algorithms concerning stability performance. Furthermore, we also tried to quantify the accuracy and stability of inverted S-wave velocity profiles for each algorithm. Table 13 lists the average MAPE between the inverted and borehole S-wave velocities for different algorithms. The GA determines the minimum average MAPE. It is worth noting that there is only a slight difference among the algorithms in terms of this performance indicator. Table 14 displays the quantification results of the inversion stability for candidate inversion algorithms based on the field data. It is clearly seen that the BL network achieves the minimum average indicator value among all algorithms. The TMCMC algorithm, DL network, and BL network obtain significantly smaller values of inversion stability indicator than other algorithms, illustrating the strong inversion stability of these

Table 13 Quantification of the inversion accuracy of candidate inversion algorithms based on the field data

	GA	PSO	SA	TMCMC	L-BFGS-B	DL	BL
Average MAPE between inverted and borehole profiles	24.2%	25.6%	24.9%	26.0%	28.2%	24.8%	26.9%

For a given algorithm, average MAPE refers to the average among the 50 MAPE values between inverted and actual profiles

Table 14 Quantification of the inversion stability of candidate inversion algorithms based on the field data

	$h^{(1)}$	$h^{(2)}$	$h^{(3)}$	$h^{(4)}$	$V_S^{(1)}$	$V_S^{(2)}$	$V_S^{(3)}$	$V_S^{(4)}$	$V_S^{(5)}$	Average
GA	16.7%	17.0%	26.0%	30.6%	5.3%	4.9%	12.2%	19.7%	7.0%	15.5%
PSO	12.5%	17.4%	24.0%	30.4%	26.7%	3.9%	11.4%	21.8%	6.9%	14.6%
SA	24.5%	14.0%	27.8%	36.2%	12.3%	3.9%	11.0%	20.4%	9.9%	17.8%
TMCMC	1.3%	1.7%	11.8%	23.9%	0.7%	0.2%	2.9%	19.2%	2.8%	7.2%
L-BFGS-B	21.8%	31.0%	30.3%	25.1%	6.9%	8.0%	21.2%	30.9%	18.4%	21.5%
DL	9.5%	9.7%	10.7%	8.8%	4.8%	2.6%	6.1%	6.5%	6.0%	7.2%
BL	5.3%	4.1%	6.4%	9.2%	2.6%	1.3%	2.3%	5.4%	2.4%	4.4%

Each cell in columns 2 to 10 of the table represents the std divided by the average of 50 inverted values of an earth model parameter for a certain algorithm; the last column of the table represents the average value of columns 2 to 10; a smaller value indicates stronger inversion stability; the value in a row highlighted in bold with underlines indicates the best performance among the seven algorithms

three algorithms. A clear contrast comes in that the L-BFGS-B algorithm provides the maximum stability indicator value. This is likely due to the disadvantage of being easily trapped in local optima.

6 Discussion

This study aims to investigate the variations in optimization performance among seven candidate algorithms in the context of Rayleigh wave inversion, providing valuable insights for researchers seeking to improve the estimation of S-wave velocities. The numerical and field data were applied to conduct a comprehensive comparison by considering the time-consuming, accuracy and stability, generalization ability, impacts of noise, field data processing capability during the inversion process. This section highlights several key points that warrant discussion.

The first aspect explores the unconformity between the convergence of the objective function curves and the inverted S-wave velocity profiles. Regarding the multi-time inversion results, a satisfactory convergence of the objective function curves does not guarantee a corresponding one of the S-wave velocity profiles, and vice versa. It is satisfactorily observed that the better convergence of the S-wave velocity profiles than the objective function curves for the inversion results produced by the DL and BL networks. In stark contrast, such a phenomenon did not occur in the inversion results obtained by the remaining candidate algorithms. This study believes that the mechanism disparity in convergence should be the main factor contributing to this situation. The DL and BL networks establish a set of mapping functions by learning the internal relationship of training samples. The objective function in Eq. (2) was only utilized to choose the most suitable network mapping complexity to approximate the unknown functional mapping of the input dispersion curves and output earth model. For a trained network with a fixed mapping complexity, the objective function is useless for functional modeling. On the other hand, the remaining five algorithms depend on the objective function to perform the iterations within the inversion process. It is widely recognized that Rayleigh wave inversion is characterized by a strong non-uniqueness of solutions. Consequently, achieving a lower objective function value does not necessarily imply a better approximation of the actual S-wave velocity

profile, which conforms to the observed discrepancy between the convergence of the objective function curves and the inverted S-wave velocity profiles. As a result, only the DL and BL networks among the candidate algorithms exhibit inverted S-wave velocity profiles that are more stable compared to the objective function curves. Undoubtedly, the impressive performance of these two neural networks is attributed to their exceptional mapping and generalization capabilities. In real applications, this finding can contribute to the better estimation of S-wave velocities of subsurface.

The second aspect focuses on the large-sample-size-based optimization algorithms among the candidate algorithms, specifically the TMCMC algorithm, the DL network, and the BL network. These three algorithms utilize the entire samples to estimate the layer thicknesses and S-wave velocities of different layers, instead of searching for a single optimal solution. This study believes that these large-sample-size-based optimization algorithms can effectively overcome the influence of local optima and converge toward a solution that closely approximates the globally optimal value. As a result, these three algorithms determine more accurate and stable S-wave velocity profiles than the remaining four algorithms. Furthermore, these algorithms also possess unique advantages. As a Bayesian algorithm, the TMCMC algorithm can provide estimates of the posterior probability density, which is essential for quantifying the uncertainty associated with the unknown parameters to be estimated. The DL and BL networks demonstrate superior generalization capability even facing a slightly incorrect search space or an inaccurate number of layers. This adaptability makes them well suited for real-world applications with limited geological information. It is crucial to highlight that the inversion performance of the DL and BL networks is significantly influenced by the training samples. In this study, random training samples are employed for both types of networks to ensure a fair comparison. In practical applications, optimizing the representativeness of the training samples for the inversion problem has the potential to enhance the inversion performance of both the DL and BL networks.

The third aspect pertains to the average profile obtained from multi-time inversion. The numerical and field data examples reveal variations among the multiple inversion profiles obtained from a specific candidate algorithm. It means that the accuracy of individual inversion profile may be compromised, despite a reasonable fit between the inverted

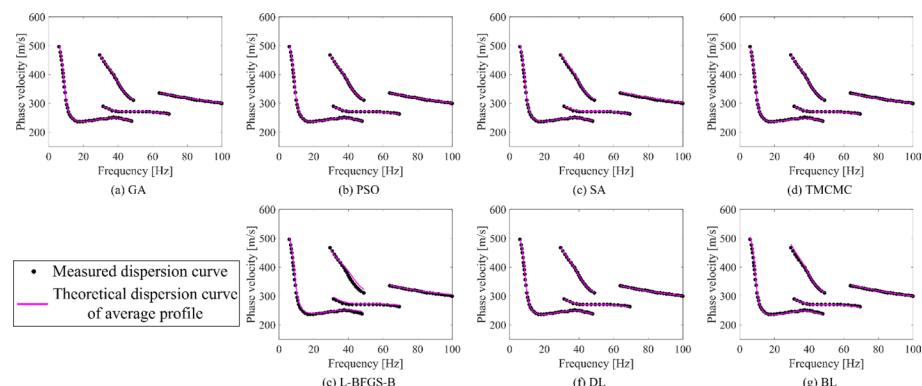


Fig. 21 Theoretical dispersion curves of average S-wave velocity profile obtained based on the theoretical data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

dispersion curve and the observed one. It is worth noting that calculating the average profile based on the multiple inversion profiles can address this problem. By utilizing the multiple inversion profiles in Sect. 4.1, the average profile and its corresponding theoretical dispersion curves were calculated to substantiate this perspective. For each candidate algorithm, Fig. 21 displays the fitting between the theoretical dispersion curve of the average S-wave velocity profile and measured dispersion curves; Fig. 22 illustrates the average S-wave velocity profile. In contrast with the inversion results in Sect. 4.1, there is no doubt that the average S-wave velocity profile can overcome the disadvantage of optimization instability of each algorithm. In practice, researchers can employ multi-time inversion to compute the average profile in order to promote accurate and stable inversion.

The last aspect delves into recommendations regarding the appropriateness of candidate algorithms in dealing with varied geological structures. Firstly, it considers the geological structures characterized by a gradual increase in velocity in the depth dimension. These regular velocity structures facilitate the acquisition of suitable prior information for the initial search space. This study recommends the TMCMC algorithm or BL network as the inversion algorithm for these geological structures based on their comprehensive performance in the conducted inversion tests. With the appropriate design of the search space, the primary distinction in inversion performance between these two methods is related to computational cost, with the latter method requiring significantly less time than the former. However, the advantage of the TMCMC algorithm is its capability

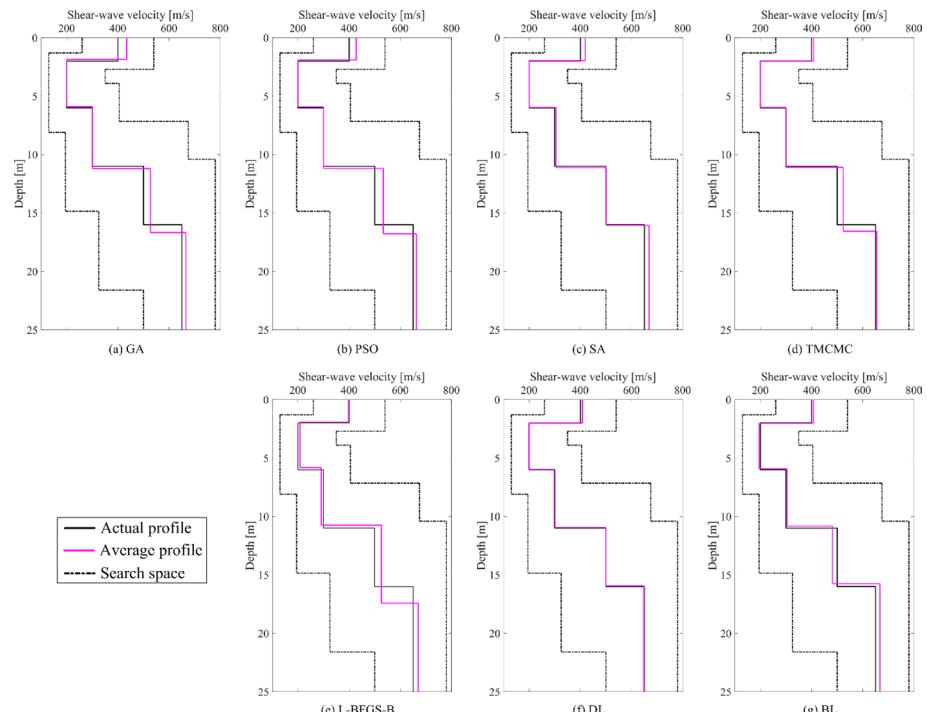


Fig. 22 Average S-wave velocity profiles of 50 inversion profiles of candidate algorithms based on the theoretical data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

to provide the credibility interval for inverted S-wave velocities under a single-time inversion scenario. Correspondingly, the advantage of the BL network is rooted in its time-saving capability, facilitating the effortless execution of multi-time inversion for real-time inversion and stability analysis. Afterward, it addresses the geological structure encompassing velocity anomaly layers. The presence of low and/or high velocity layers, especially with noticeable velocity contrasts, can easily lead to an incorrect initial search space. By observing the inversion tests in Sect. 4.2, only the DL and BL networks demonstrated satisfactory inversion performance when employing an incorrect search space. Notably, the BL network demands significantly less training time than the DL network, which is crucial for multi-time inversion. Therefore, opting for the BL network as the inversion algorithm is deemed beneficial for Rayleigh wave inversion, given these geological structures. After that, it explores the geological structures featuring lateral velocity variation. Lateral heterogeneity can lead to variations in the number of layers and search space for S-wave velocity profiles from different field records, making it difficult to set up a universal inversion model. Under such circumstances, this study recommends inversion algorithms of generalization capability to overcome the complex situation. Obviously, the DL and BL networks demonstrated superior generalization capability compared to other candidate algorithms. Users can choose either of these two inversion methods when dealing with lateral velocity variation. Based on the above discussion, the BL network might exhibit commendable inversion performance facing different geological structures.

7 Conclusions

This study comprehensively compares four categories of algorithms for Rayleigh wave inversion, i.e., two bionic algorithms, two probabilistic algorithms, a gradient-based algorithm, and two neural network algorithms. Specifically, the candidate algorithm contains the genetic algorithm, the particle swarm optimization algorithm, the simulated annealing algorithm, the transitional Markov Chain Monte Carlo algorithm, the Bound-constrained limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm, the deep learning algorithm, and the broad learning algorithm. The evaluation factors include the computational cost of inversion, the accuracy and stability of inversion, the generalization ability when facing an incorrect search space, the generalization ability related to the setting of a wrong number of layers, the impact of noise, and the capability to process field data. The comparison of candidate algorithms for the inversion process adheres to a fair principle, refraining from misunderstanding the inversion performance of each algorithm.

The comparison results reveal varied inversion performance among the candidate algorithms about the evaluation factors. For computational cost, the L-BFGS-B algorithm and BL network demonstrate superior optimization efficiency compared to other algorithms. The corresponding performance of the GA, PSO algorithm, and SA algorithm presents a moderate level among all algorithms, while the remaining two algorithms, particularly the DL network, exhibit the lowest efficiency performance. Regarding the aspects of inversion accuracy and stability, the TMCMC algorithm, DL network, and BL network outperform the remaining four algorithms. The results illustrate that optimization algorithms based on large sample sizes may contribute to enhancing both the accuracy and stability of the inversion process. In terms of the designed cases with a

slightly wrong search space or a higher number of layers, the DL and BL networks present better inversion generalization capability compared with other algorithms. Even in such scenarios, these two neural networks successfully generate inverted S-wave velocity profiles comparable to those obtained under normal circumstances. When faced with numerical dispersion curves containing noise, the TMCMC algorithm, DL network, and BL network exceed the remaining algorithms considering the inverted S-wave velocities. The DL network determines the most satisfactory inverted results under noisy data. Concerning the field Rayleigh wave data, the TMCMC algorithm, DL network, as well as the BL network outperform the other four algorithms. Given a specific algorithm, increasing the number of inversions can enhance the accuracy of S-wave velocity estimation through Rayleigh wave inversion.

In this study, seven candidate algorithms are employed for the Rayleigh wave dispersion curve inversion process. With identical measured dispersion curves, the candidate algorithms demonstrate diverse inversion performance. In real applications, for Rayleigh wave dispersion curve inversion, the researchers can choose an appropriate algorithm inside the inversion process for adapting to the practical necessities according to the comprehensive comparison of candidate algorithms. This study can promote the S-wave velocity estimation by Rayleigh wave inversion.

Appendix 1

Appendix 1 describes the objective function curves and inverted dispersion curves of 50 inversions by candidate algorithms based on theoretical data under a slightly incorrect search space in Sect. 4.2. Figure 23 displays the convergence curves of objective function produced by different algorithms. In Fig. 24, it illustrates the data fitting between the inverted and measured dispersion curves for each algorithm.

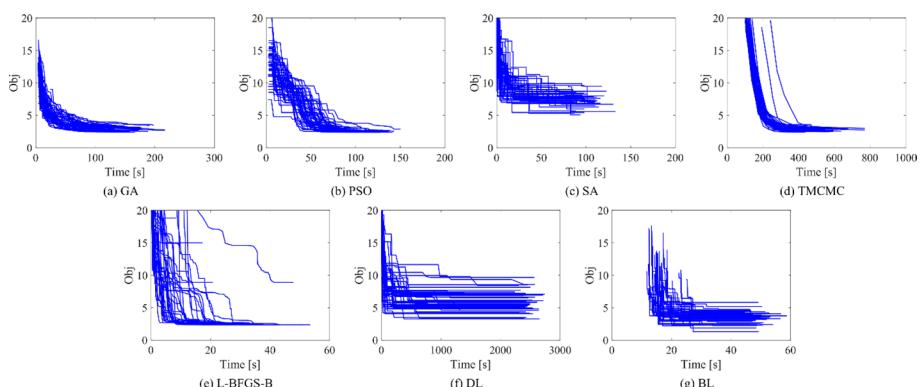


Fig. 23 Objective function curves of 50-time inversion of candidate algorithms based on theoretical data under a wrong search space: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

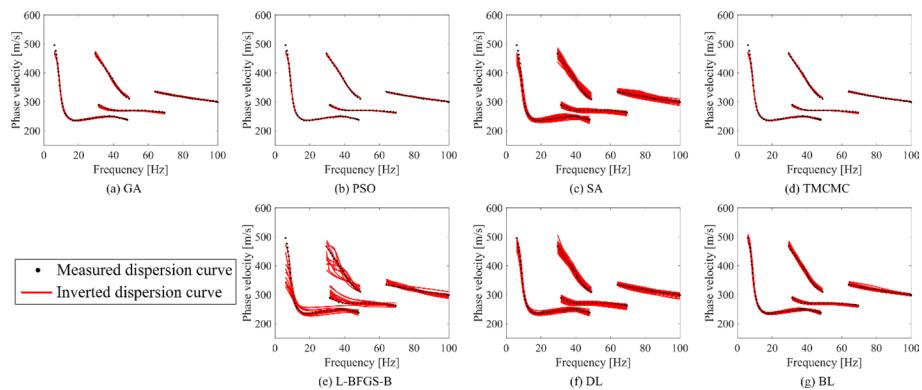


Fig. 24 Inverted dispersion curves of 50-time inversion of candidate algorithms based on theoretical data under a wrong search space: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

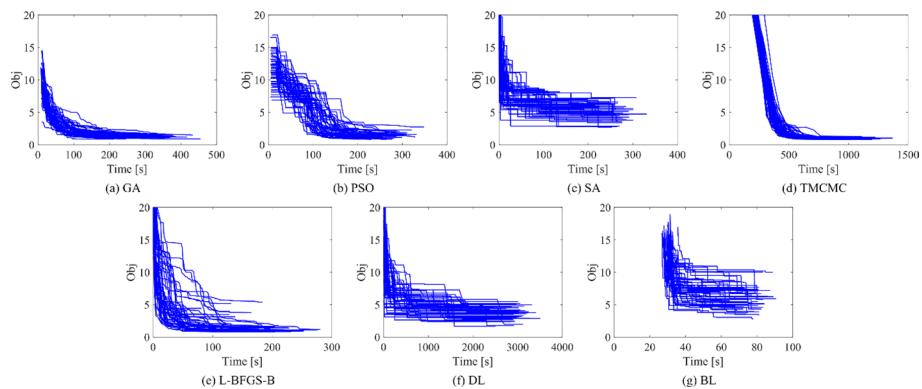


Fig. 25 Objective function curves of 50-time inversion of candidate algorithms based on theoretical data under a wrong number of layers: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

Appendix 2

Appendix 2 aims to provide the objective function curves and inverted dispersion curves of the inversion performed in Sect. 4.3, where an incorrect number of layers is applied to simulate the corresponding scenario in real applications. The wrong number of layers equals eight, while the actual number of layers equals five. Figure 25 displays the 50 objective function curves generated by each algorithm using theoretical data, assuming the presence of eight layers. Figure 26 shows the corresponding data fitting between inverted and measured dispersion curves of 50 inversions for the candidate algorithms.

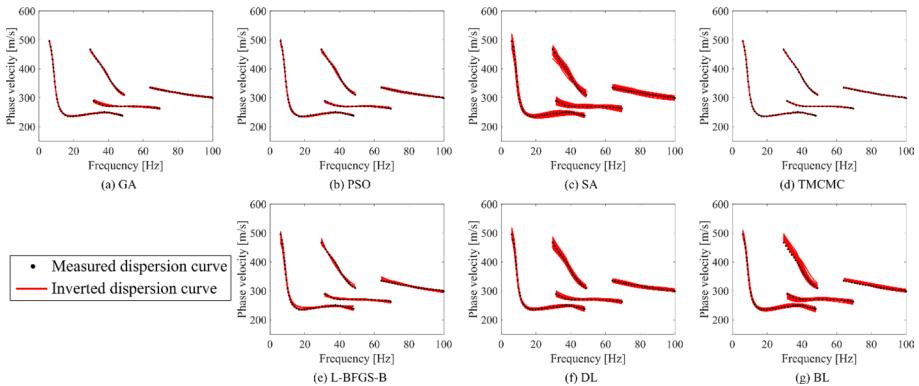


Fig. 26 Inverted dispersion curves of 50-time inversion of candidate algorithms based on theoretical data under a wrong number of layers: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

Appendix 3

Appendix 3 presents the objective function curves and inverted dispersion curves of the candidate algorithms for the numerical simulation in Sect. 4.4. Figure 27 illustrates the convergence objective function curves of 50 inversions by the candidate algorithms under 2.5% noisy dispersion curves. Figure 28 shows the data fitting between inverted and measured dispersion curves.

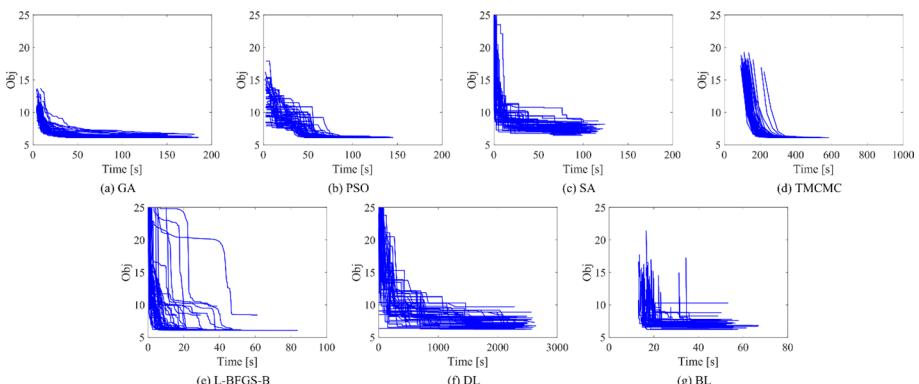


Fig. 27 Objective function curves of 50-time inversion of candidate algorithms based on noisy theoretical data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

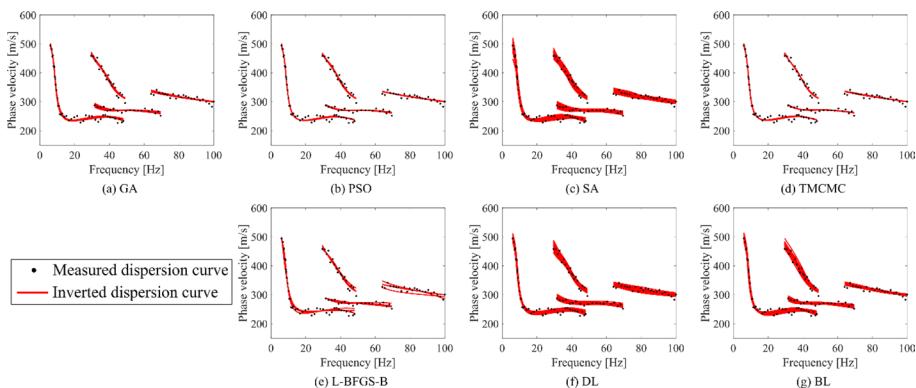


Fig. 28 Inverted dispersion curves of 50-time inversion of candidate algorithms based on noisy theoretical data: **a** GA, **b** PSO algorithm, **c** SA algorithm, **d** TMCMC algorithm, **e** L-BFGS-B algorithm, **f** DL network, **g** BL network

Acknowledgements The forward modeling program of theoretical Rayleigh wave dispersion curve comes from the website (<https://github.com/eespr/MuLTI>) provided by Killingbeck et al. (2018). According to Killingbeck et al. (2018), this program is created by Richard Rigby (at University of Leeds) by cross-platform programming based on the gpdc function of the Geopsy software (Wathelet 2005). The Rayleigh wave dispersion curves applied in Section 5 comes from Yang et al. (2023), where the field Rayleigh wave data were obtained by the InterPACIFIC project (Garofalo et al. 2016a; b). This research is partly supported by the National Key Research & Development Program of China (Grant No. 2023YFE0101800), Talent Launch Project of Chengdu University of Information Technology (Grant No. KYTZ2023035 and KYTZ202220), Guangdong Provincial Key Laboratory of Geophysical High-resolution Imaging Technology (Grant No. 2022B1212010002), and Research Project on Disciplinary Development Strategy, Academic Divisions of the Chinese Academy of Sciences (Grant No. XK2018DXA001 and XK2018DXC003).

Open Research The codes for the inversion algorithms evaluated in this study for Rayleigh wave inversion are available for access on the website (<https://github.com/X-H-Yang/Comprehensive-comparison-Rayleigh-wave-inversion-algorithms>).

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

References

- Aki K (1957) Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bull Earthq Res Inst* 35:415–456
- Beaty KS, Schmitt DR, Sacchi M (2002) Simulated annealing inversion of multimode Rayleigh wave dispersion curves for geological structure. *Geophys J Int* 151(2):622–631
- Benoit MH, Nyblade AA, Pasanos ME (2006) Crustal thinning between the Ethiopian and East African plateaus from modeling Rayleigh wave dispersion. *Geophys Res Lett* 33(13)
- Berg EM, Lin FC, Allam A, Qiu H, Shen W, Ben-Zion Y (2018) Tomography of Southern California via Bayesian joint inversion of Rayleigh wave ellipticity and phase velocity from ambient noise cross-correlations. *J Geophys Res Solid Earth* 123(11):9933–9949
- Bergamo P, Maranò S, Fäh D (2023) Joint estimation of s-wave velocity and damping ratio of the near-surface from active Rayleigh wave surveys processed with a Wavefield decomposition approach. *Geophys J Int*
- Bodin T, Sambridge M, Tkalcic H, Arroucau P, Gallagher K, Rawlinson N (2012) Transdimensional inversion of receiver functions and surface wave dispersion. *J Geophys Res Solid Earth* 117(B2):429

- Boiero D, Socco LV (2014) Joint inversion of Rayleigh-wave dispersion and P-wave refraction data for laterally varying layered models. *Geophysics* 79(4):EN49–EN59
- Bonadio R, Geissler WH, Lebedev S, Fullea J, Ravenna M, Celli NL et al (2018) Hot upper mantle beneath the Tristan da Cunha hotspot from probabilistic Rayleigh-wave inversion and petrological modeling. *Geochem Geophys Geosyst* 19(5):1412–1428
- Boxberger T, Picozzi M, Parolai S (2011) Shallow geology characterization using Rayleigh and Love wave dispersion curves derived from seismic noise array measurements. *J Appl Geophys* 75(2):345–354
- Byrd RH, Lu P, Nocedal J, Zhu C (1995) A limited memory algorithm for bound constrained optimization. *SIAM J Sci Comput* 16(5):1190–1208
- Caldwell WB, Klemperer SL, Rai SS, Lawrence JF (2009) Partial melt in the upper-middle crust of the northwest Himalaya revealed by Rayleigh wave dispersion. *Tectonophysics* 477(1–2):58–65
- Campman X, Dwi Riyanti C (2007) Non-linear inversion of scattered seismic surface waves. *Geophys J Int* 171(3):1118–1125
- Cercato M (2018) Sensitivity of Rayleigh wave ellipticity and implications for surface wave inversion. *Geophys J Int* 213(1):489–510
- Černý V (1985) Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *J Optim Theory Appl* 45:41–51
- Chen X (1993) A systematic and efficient method of computing normal modes for multilayered half-space. *Geophys J Int* 115(2):391–409
- Chen CP, Liu Z (2017) Broad learning system: an effective and efficient incremental learning system without the need for deep architecture. *IEEE Trans Neural Netw Learn Syst* 29(1):10–24
- Chen CP, Liu Z, Feng S (2018) Universal approximation capability of broad learning system and its structural variations. *IEEE Trans Neural Netw Learn Syst* 30(4):1191–1204
- Chen X, Xia J, Pang J, Zhou C, Mi B (2022) Deep learning inversion of Rayleigh-wave dispersion curves with geological constraints for near-surface investigations. *Geophys J Int* 231(1):1–14
- Cheng F, Xia J, Shen C, Hu Y, Xu Z, Mi B (2018a) Imposing active sources during high-frequency passive surface-wave measurement. *Engineering* 4(5):685–693
- Cheng F, Xia J, Xu Z, Hu Y, Mi B (2018b) Frequency–wavenumber (FK)-based data selection in high-frequency passive surface wave survey. *Surv Geophys* 39:661–682
- Cheng F, Xia J, Behm M, Hu Y, Pang J (2019) Automated data selection in the Tau–p domain: Application to passive surface wave imaging. *Surv Geophys* 40:1211–1228
- Ching J, Chen YC (2007) Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging. *J Eng Mech* 133(7):816–832
- Dai T, Xia J, Ning L, Xi C, Liu Y, Xing H (2021) Deep learning for extracting dispersion curves. *Surv Geophys* 42:69–95
- Dal Moro G (2008) VS and VP vertical profiling via joint inversion of Rayleigh waves and refraction travel times by means of bi-objective evolutionary algorithm. *J Appl Geophys* 66(1–2):15–24
- Dal Moro G, Pipan M (2007) Joint inversion of surface wave dispersion curves and reflection travel times via multi-objective evolutionary algorithms. *J Appl Geophys* 61(1):56–81
- Dal Moro G, Pipan M, Gabrielli P (2007) Rayleigh wave dispersion curve inversion via genetic algorithms and marginal posterior probability density estimation. *J Appl Geophys* 61(1):39–55
- Dong S, Li Z, Chen X, Fu L (2021) DisperNet: An effective method of extracting and classifying the dispersion curves in the frequency–bessel dispersion spectrum. *Bull Seismol Soc Am* 111(6):3420–3431
- Duputel Z, Cara M, Rivera L, Herquel G (2010) Improving the analysis and inversion of multimode Rayleigh-wave dispersion by using group-delay time information observed on arrays of high-frequency sensors. *Geophysics* 75(2):R13–R20
- Eker AM, Akgün H, Koçkar MK (2012) Local site characterization and seismic zonation study by utilizing active and passive surface wave methods: a case study for the northern side of Ankara, Turkey. *Eng Geol* 151:64–81
- Fang H, Yao H, Zhang H, Huang YC, van der Hilst RD (2015) Direct inversion of surface wave dispersion for three-dimensional shallow crustal structure based on ray tracing: methodology and application. *Geophys J Int* 201(3):1251–1263
- Feng X, Chen X (2022) Rayleigh-wave dispersion curves from energetic hurricanes in the Southeastern United States. *Bull Seismol Soc Am* 112(2):622–633
- Fu L, Pan L, Ma Q, Dong S, Chen X (2021) Retrieving S-wave velocity from surface wave multimode dispersion curves with DispINet. *J Appl Geophys* 193:104430
- Fu L, Pan L, Li Z, Dong S, Ma Q, Chen X (2022a) Improved high-resolution 3D Vs model of Long Beach, CA: inversion of multimodal dispersion curves from ambient noise of a dense array. *Geophys Res Lett* 49(4):21

- Fu Y, Yang A, Yao Z, Liu Y, Li H, Chen H, Wang X (2022) Inversion of Rayleigh wave dispersion curves via long short-term memory combined with particle swarm optimization. *Comput Intell Neurosci* 2022
- Gao L, Xia J, Pan Y, Xu Y (2016) Reason and condition for mode kissing in MASW method. *Pure Appl Geophys* 173:1627–1638
- Gao L, Pan Y, Tian G, Xia J (2018) Estimating Q factor from multi-mode shallow-seismic surface waves. *Pure Appl Geophys* 175:2609–2622
- Garofalo F, Foti S, Hollender F, Bard PY, Cornou C, Cox BR et al (2016a) InterPACIFIC project: comparison of invasive and non-invasive methods for seismic site characterization. Part I: Intra-comparison of surface wave methods. *Soil Dyn Earthq Eng* 82:222–240
- Garofalo F, Foti S, Hollender F, Bard PY, Cornou C, Cox BR et al (2016b) InterPACIFIC project: Comparison of invasive and non-invasive methods for seismic site characterization. Part II: Inter-comparison between surface-wave and borehole methods. *Soil Dyn Earthq Eng* 82:241–254
- Gofer E, Bachrach R, Marco S (2017) Anisotropic surface-wave characterization of granular mediaAnisotropic Rayleigh wave inversion. *Geophysics* 82(6):MR191–MR200
- Golberg DE (1989) Genetic algorithms in search, optimization, and machine learning. Addison Wesley 1989(102):36
- Gouveia F, Lopes I, Gomes RC (2016) Deeper VS profile from joint analysis of Rayleigh wave data. *Eng Geol* 202:85–98
- Gribler G, Liberty LM, Mikesell TD (2020) High-velocity surface layer effects on Rayleigh waves: Recommendations for improved shear-wave velocity modeling. *Bull Seismol Soc Am* 110(1):279–287
- Haney MM, Tsai VC (2017) Perturbational and nonperturbational inversion of Rayleigh-wave velocities—Inversion of Rayleigh-wave velocities. *Geophysics* 82(3):F15–F28
- Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biométrika* 57(1):97–109
- Herrmann RB (2013) Computer programs in seismology: An evolving tool for instruction and research. *Seismol Res Lett* 84(6):1081–1088
- Hobiger M, Cornou C, Wathelet M, Giulio GD, Knapmeyer-Endrun B, Renalier F et al (2013) Ground structure imaging by inversions of Rayleigh wave ellipticity: sensitivity analysis and application to European strong-motion sites. *Geophys J Int* 192(1):207–229
- Holland JH (1975) Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor (2nd edn, MIT Press, 1992)
- Hu S, Luo S, Yao H (2020) The frequency-Bessel spectrograms of multicomponent cross-correlation functions from seismic ambient noise. *J Geophys Res Solid Earth* 125(8):e2020JB019630
- Huynh NNT, Martin R, Oberlin T, Plazolles B (2023) Near-surface seismic arrival time picking with transfer and semi-supervised learning. *Surv Geophys* 1–25
- Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of ICNN'95-international conference on neural networks, vol 4. IEEE, pp 1942–1948
- Kiełczyński P, Szalewski M (2011) An inverse method for determining the elastic properties of thin layers using Love surface waves. *Inverse Probl Sci Eng* 19(1):31–43
- Killingbeck SF, Livermore PW, Booth AD, West LJ (2018) Multimodal layered transdimensional inversion of seismic dispersion curves with depth constraints. *Geochem Geophys Geosyst* 19(12):4957–4971
- Kirkpatrick S, Gelatt CD Jr, Vecchi MP (1983) Optimization by simulated annealing. *Science* 220(4598):671–680
- Kuok SC, Yuen KV (2020) Broad learning for nonparametric spatial modeling with application to seismic attenuation. *Comput Aided Civil Infrastruct Eng* 35(3):203–218
- Lei Y, Shen H, Li X, Wang X, Li Q (2019) Inversion of Rayleigh wave dispersion curves via adaptive GA and nested DLS. *Geophys J Int* 218(1):547–559
- Li Y, Wu Q, Pan J, Sun L (2012) S-wave velocity structure of northeastern China from joint inversion of Rayleigh wave phase and group velocities. *Geophys J Int* 190(1):105–115
- Li J, Feng Z, Schuster G (2017) Wave-equation dispersion inversion. *Geophys J Int* 208(3):1567–1578
- Li X, Li Q, Shen H, Yang F, Zhang X (2022) Research on oil exploration seismic Rayleigh wave imaging based on multi-channel analysis of surface waves and genetic-damped least squares joint inversion. *J Appl Geophys* 202:104670
- Lin S, Ashlock JC, Zhao G, Lai Q, Xu L, Zhai C (2023) Genetic-simulated annealing optimization for surface wave inversion of shear-wave velocity profiles of geotechnical sites. *Comput Geotech* 160:105525
- Liu Z, Li J, Hanafy SM, Schuster G (2019) 3D wave-equation dispersion inversion of Rayleigh waves. *Geophysics* 84(5):R673–R691

- Liu Q, Lu L, Wang K, Chang L, Zhu Y (2023) Rayleigh wave phase velocity maps at regional scale inferring from SPAC of ambient noise at a dense array: a case study in Northeastern Tibetan Plateau. *Pure Appl Geophys* 1–16
- Lomax A, Snieder R (1995) The contrast in upper mantle shear-wave velocity between the East European Platform and tectonic Europe obtained with genetic algorithm inversion of Rayleigh-wave group dispersion. *Geophys J Int* 123(1):169–182
- Lu L, Zhang B (2006) Inversion of Rayleigh waves using a genetic algorithm in the presence of a low-velocity layer. *Acoust Phys* 52(6):701–712
- Lu J, Li S, Li W, Tang L (2014) A hybrid inversion method of damped least squares with simulated annealing used for Rayleigh wave dispersion curve inversion. *Earthq Eng Eng Vib* 13(1):13–21
- Lu Y, Peng S, Du W, Zhang X, Ma Z, Lin P (2016) Rayleigh wave inversion using heat-bath simulated annealing algorithm. *J Appl Geophys* 134:267–280
- Luo Y, Xia J, Liu J, Xu Y, Liu Q (2008) Generation of a pseudo-2D shear-wave velocity section by inversion of a series of 1D dispersion curves. *J Appl Geophys* 64(3–4):115–124
- Luo Y, Huang Y, Yang Y, Zhao K, Yang X, Xu H (2022) Constructing shear velocity models from surface wave dispersion curves using deep learning. *J Appl Geophys* 196:104524
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. *J Chem Phys* 21(6):1087–1092
- Mi B, Xia J, Shen C, Wang L (2018) Dispersion energy analysis of Rayleigh and Love waves in the presence of low-velocity layers in near-surface seismic surveys. *Surv Geophys* 39:271–288
- Mi B, Xia J, Bradford JH, Shen C (2020) Estimating near-surface shear-wave-velocity structures via multi-channel analysis of Rayleigh and love waves: an experiment at the Boise hydrogeophysical research site. *Surv Geophys* 41:323–341
- Mi B, Xia J, Xu Y, You B, Chen Y (2023) Retrieval of surface waves from high-speed-train-induced vibrations using seismic interferometry. *Geophysics* 88(5):1–85
- Moré JJ (1978) The Levenberg-Marquardt algorithm: implementation and theory. In: Numerical analysis. Springer, Berlin, pp 105–116
- Mousavi SM, Beroza GC (2019) Bayesian-deep-learning estimation of earthquake location from single-station observations. *arXiv preprint arXiv:1912.01144*.
- Mousavi SM, Beroza GC (2023) Machine learning in earthquake seismology. *Annu Rev Earth Planet Sci* 51:105–129
- Ohori M, Nobata A, Wakamatsu K (2002) A comparison of ESAC and FK methods of estimating phase velocity using arbitrarily shaped microtremor arrays. *Bull Seismol Soc Am* 92(6):2323–2332
- Pan Y, Xia J, Gao L, Shen C, Zeng C (2013) Calculation of Rayleigh-wave phase velocities due to models with a high-velocity surface layer. *J Appl Geophys* 96:1–6
- Pan Y, Xia J, Xu Y, Xu Z, Cheng F, Xu H, Gao L (2016) Delineating shallow S-wave velocity structure using multiple ambient-noise surface-wave methods: an example from Western Junggar, China. *Bull Seismol Soc Am* 106(2):327–336
- Pan Y, Schaneng S, Steinweg T, Bohlen T (2018) Estimating S-wave velocities from 3D 9-component shallow seismic data using local Rayleigh-wave dispersion curves—a field study. *J Appl Geophys* 159:532–539
- Pan L, Chen X, Wang J, Yang Z, Zhang D (2019) Sensitivity analysis of dispersion curves of Rayleigh waves with fundamental and higher modes. *Geophys J Int* 216(2):1276–1303
- Park CB, Miller RD, Xia J (1999) Multichannel analysis of surface waves. *Geophysics* 64(3):800–808
- Park CB, Miller RD, Xia J, Ivanov J (2007) Multichannel analysis of surface waves (MASW)—active and passive methods. *Lead Edge* 26(1):60–64
- Park CB, Miller RD, Xia J (1998) Imaging dispersion curves of surface waves on multi-channel record. In: SEG technical program expanded abstracts 1998. Society of Exploration Geophysicists, pp 1377–1380
- Pasyanos ME, Walter WR, Hazler SE (2001) A surface wave dispersion study of the Middle East and North Africa for monitoring the Comprehensive Nuclear-Test-Ban Treaty. In: Monitoring the comprehensive nuclear-test-ban treaty: surface waves, pp 1445–1474
- Pei D, Louie JN, Pullammanappallil SK (2007) Application of simulated annealing inversion on high-frequency fundamental-mode Rayleigh wave dispersion curves. *Geophysics* 72(5):R77–R85
- Poormirzaee R (2016) S-wave velocity profiling from refraction microtremor Rayleigh wave dispersion curves via PSO inversion algorithm. *Arab J Geosci* 9(16):673
- Poormirzaee R, Fister I Jr (2021) Model-based inversion of Rayleigh wave dispersion curves via linear and nonlinear methods. *Pure Appl Geophys* 178:341–358
- Qin T, Zhao Y, Hu S, An C, Bi W, Ge S et al (2020) An interactive integrated interpretation of GPR and Rayleigh wave data based on the genetic algorithm. *Surv Geophys* 41:549–574

- Renalier F, Jongmans D, Savvaidis A, Wathélet M, Endrun B, Cornou C (2010) Influence of parameterization on inversion of surface wave dispersion curves and definition of an inversion strategy for sites with a strong vs contrast. *Geophysics* 75(6):B197–B209
- Sahadewa, A., Zekkos, D., Lobbestael, A., & Woods, R. D. (2011). Shear wave velocity measurements at municipal solid waste landfills in Michigan. In: Proceedings of 14th Pan-American conference on soil mechanics and geotechnical engineering. Toronto, Canada
- Shakir AM, Foti S, Garofalo F, Hijab BR, Laftah AA (2013) Laterally constrained inversion of surface wave data at Najaf city (Iraq). *Soil Dyn Earthq Eng* 45:89–95
- Shen W, Wiens DA, Anandakrishnan S, Aster RC, Gerstoft P, Bromirski PD et al (2018) The crust and upper mantle structure of central and West Antarctica from Bayesian inversion of Rayleigh wave and receiver functions. *J Geophys Res Solid Earth* 123(9):7824–7849
- Socco LV, Comina C, Anjom FK (2017) Time-average velocity estimation through surface-wave analysis: Part 1—S-wave velocity. *Geophysics* 82(3):U49–U59
- Song X, Gu H, Zhang X, Liu J (2008) Pattern search algorithms for nonlinear inversion of high-frequency Rayleigh-wave dispersion curves. *Comput Geosci* 34(6):611–624
- Song X, Tang L, Lv X, Fang H, Gu H (2012) Application of particle swarm optimization to interpret Rayleigh wave dispersion curves. *J Appl Geophys* 84:1–13
- Song W, Feng X, Wu G, Zhang G, Liu Y, Chen X (2021a) Convolutional neural network, Res-Unet++, based dispersion curve picking from noise cross-correlations. *J Geophys Res Solid Earth* 126(11):e2021JB022027
- Song Z, Zeng X, Thurber CH (2021b) Surface-wave dispersion spectrum inversion method applied to Love and Rayleigh waves recorded by distributed acoustic sensingDAS surface-wave inversion. *Geophysics* 86(1):EN1–EN12
- Tremblay SP, Karray M (2019) Practical considerations for array-based surface-wave testing methods with respect to near-field effects and shear-wave velocity profiles. *J Appl Geophys* 171:103871
- Vashisth D, Shekar B, Srivastava S (2022) Joint inversion of Rayleigh wave fundamental and higher order mode phase velocity dispersion curves using multi-objective grey wolf optimization. *Geophys Prospect* 70(3):479–501
- Wang J, Wu G, Chen X (2019) Frequency-Bessel transform method for effective imaging of higher-mode Rayleigh dispersion curves from ambient seismic noise data. *J Geophys Res Solid Earth* 124(4):3708–3723
- Wang F, Song X, Li M (2022a) A deep-learning-based approach for seismic surface-wave dispersion inversion (SfNet) with application to the Chinese mainland. *Earthq Sci* 36(2):147–168
- Wang S, Yuen KV, Yang X, Zhang B (2022b) A nonparametric tropical cyclone wind speed estimation model based on dual-polarization SAR observations. *IEEE Trans Geosci Remote Sens* 60:1–13
- Wang Y, Song X, Zhang X, Yuan S, Zhang K, Wang L et al (2023) Multi-objective particle swarm optimization for multimode surface wave analysis. *Comput Geosci* 176:105343
- Wathélet M (2005) Array recordings of ambient vibrations: surface-wave inversion. Ph.D. Diss., Liège University, 161
- Wilken D, Rabbel W (2012) On the application of particle swarm optimization strategies on Scholte-wave inversion. *Geophys J Int* 190(1):580–594
- Wu GX, Pan L, Wang JN, Chen X (2020) Shear velocity inversion using multimodal dispersion curves from ambient seismic noise data of USArray transportable array. *J Geophys Res Solid Earth* 125(1):e2019JB018213
- Xi C, Xia J, Mi B, Dai T, Liu Y, Ning L (2021) Modified frequency–Bessel transform method for dispersion imaging of Rayleigh waves from ambient seismic noise. *Geophys J Int* 225(2):1271–1280
- Xia J, Miller RD, Park CB (1999) Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *Geophysics* 64(3):691–700
- Xia J, Xu Y, Miller RD (2007) Generating an image of dispersive energy by frequency decomposition and slant stacking. *Pure Appl Geophys* 164:941–956
- Xu J, Song X (2012) Ant colony optimization for nonlinear inversion of Rayleigh waves. In: Bio-Inspired computing and applications: 7th international conference on intelligent computing, ICIC 2011, Zhengzhou, China, August 11–14. 2011, Revised Selected Papers 7. Springer, Berlin, pp 370–377
- Yamanaka H, Chimoto K (2018) Variability of shallow soil amplification from surface-wave inversion using the Markov-chain Monte Carlo method. *Soil Dyn Earthq Eng* 107:141–151
- Yamanaka H, Ishida H (1996) Application of genetic algorithms to an inversion of surface-wave dispersion data. *Bull Seismol Soc Am* 86(2):436–444
- Yang XH, Yuen KV (2021) All-parameters Rayleigh wave inversion. *Earthq Eng Eng Vib* 20:517–534
- Yang XH, Han P, Yang Z, Miao M, Sun YC, Chen X (2022) Broad learning framework for search space design in rayleigh wave inversion. *IEEE Trans Geosci Remote Sens* 60:1–17

- Yang XH, Han P, Yang Z, Chen X (2023) Two-stage broad learning inversion framework for shear-wave velocity estimation. *Geophysics* 88(1):WA219–WA237
- Yanovskaya TB, Kozhevnikov VM (2003) 3D S-wave velocity pattern in the upper mantle beneath the continent of Asia from Rayleigh wave data. *Phys Earth Planet Inter* 138(3–4):263–278
- Yao H (2015) A method for inversion of layered shear wavespeed azimuthal anisotropy from Rayleigh wave dispersion using the Neighborhood Algorithm. *Earthq Sci* 28:59–69
- Yin X, Xu H, Mi B, Hao X, Wang P, Zhang K (2020) Joint inversion of Rayleigh and Love wave dispersion curves for improving the accuracy of near-surface S-wave velocities. *J Appl Geophys* 176:103939
- Yuen KV, Yang XH (2020) Bayesian Rayleigh wave inversion with an unknown number of layers. *Earthq Eng Eng Vib* 19:875–886
- Zarean A, Mirzaei N, Shabani E (2013) Introducing artificial bee colony optimization to invert surface wave dispersion curve. In: Near surface geoscience 2013–19th EAGE European meeting of environmental and engineering geophysics. EAGE Publications BV, pp. cp-354
- Zeng C, Xia J, Miller RD, Tsoflias GP (2011) Feasibility of waveform inversion of Rayleigh waves for shallow shear-wave velocity using a genetic algorithm. *J Appl Geophys* 75(4):648–655
- Zhan W, Pan L, Chen X (2020) A widespread mid-crustal low-velocity layer beneath Northeast China revealed by the multimodal inversion of Rayleigh waves from ambient seismic noise. *J Asian Earth Sci* 196:104372
- Zhang ZD, Alkhalifah T (2019) Wave-equation Rayleigh-wave dispersion inversion using fundamental and higher modesSurface wave inversion. *Geophysics* 84(4):EN57–EN65
- Zhang D, Yang B, Yang Z, Zhang M, Xiong Z, Zhu D, Zhang X (2022) Multimodal inversion of Rayleigh wave dispersion curves based on a generalized misfit function. *J Appl Geophys* 207:104849
- Zhang ZD, Saygin E, He L, Alkhalifah T (2021) Rayleigh wave dispersion spectrum inversion across scales. *Surv Geophys* 1–23
- Zhou TF, Peng GX, Hu TY, Duan WS, Yao FC, Liu YM (2014) Rayleigh wave nonlinear inversion based on the Firefly algorithm. *Appl Geophys* 11(2):167–178
- Zhou Y, Li X, Tang Q, Kuok SC, Fei K, Gao L (2022) An assimilating model using broad learning system for incorporating multi-source precipitation data with environmental factors over Southeast China. *Earth Space Sci* 9(4):e2021EA002043
- Zhou Y, Ghosh A, Fang L, Yue H, Zhou S (2023) Construction of Long-term Seismic Catalog with Deep Learning and Characterization of Preseismic Fault Behavior in the Ridgecrest-Coso Region (2008–2019)
- Zhu C, Byrd RH, Lu P, Nocedal J (1997) Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans Math Softw (TOMS)* 23(4):550–560
- Zuo J, Niu F, Liu L, Da S, Zhang H, Yang J et al (2022) 3D anisotropic P-and S-mode wavefields separation in 3D elastic reverse-time migration. *Surv Geophys* 43(3):673–701

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Xiao-Hui Yang¹  · Yuanyuan Zhou¹  · Peng Han^{2,3,4}  · Xuping Feng^{2,3,4} ·
Xiaofei Chen^{2,3,4}

 Peng Han
hanp@sustech.edu.cn

Xiao-Hui Yang
yangxh@cuit.edu.cn

¹ Plateau Atmosphere and Environment Key Laboratory of Sichuan Province, School of Atmospheric Sciences, Chengdu University of Information Technology, Chengdu 610225, China

² Shenzhen Key Laboratory of Deep Offshore Oil and Gas Exploration Technology, Southern University of Science and Technology, Shenzhen 518055, China

³ Guangdong Provincial Key Laboratory of Geophysical High-Resolution Imaging Technology, Southern University of Science and Technology, Shenzhen 518055, China

⁴ Department of Earth and Space Sciences, Southern University of Science and Technology, Shenzhen 518055, China