A Deep Learning-Based Forward Surrogate Model for Accelerating Surface-Wave Dispersion Inversion

Qianyun Dong[1,2], Hongrui Qiu[*1,2], Jing Hu[*3], Yong Zheng[1,2,4], Sanxi Ai[1], Siyuan He[5], and Lei Qin[1]

Abstract

We present a deep learning-based surrogate forward operator to accelerate 1D shear-wave velocity (V_S) inversion from Rayleigh-wave dispersion. A gated recurrent unit-based surrogate forward model (FGRU) is trained to predict phase-velocity dispersion curves from the V_S profiles. Training data are generated using a cubic B-spline crust–mantle parameterization with sufficiently broad variability to ensure accurate learning of the mapping. Benchmarks on a consumer-grade workstation show that FGRU reproduces the reference dispersion curves computed by a physics-based surface-wave dispersion solver accelerated using numba (DISBA) with high fidelity. In contrast, graphic processing unit (GPU)-accelerated batch inference enables forward evaluations that are orders of magnitude faster. When embedded in particle swarm optimization (FGRU–PSO), synthetic tests demonstrate that the surrogate forward operator accelerates the inversion by reducing the computational cost of large forward-model ensembles, while producing results comparable to those obtained with the conventional solver (DISBA–PSO). Application to South China phase-velocity data allows up to one million model evaluations per grid point at a manageable cost. Compared with previous tomographic models, the resulting 3D V_S model not only recovers the major tectonic features but also reveals clearer shallow basin-related low-velocity anomalies, owing to a more thorough exploration of model space. These results demonstrate that the GPU-parallelized surrogate forward operator provides an efficient and practical route for accelerating regional surface-wave inversion.

Cite this article as Dong, Q., H. Qiu, J. Hu, Y. Zheng, S. Ai, S. He, and L. Qin (2026). A Deep Learning-Based Forward Surrogate Model for Accelerating Surface-Wave Dispersion Inversion, *Seismol. Res. Lett.* **XX**, 1–12, doi: 10.1785/0220260035.

Supplemental Material

Introduction

Surface-wave tomography has long been used to probe Earth's interior across diverse spatial scales, incorporating isotropic and anisotropic imaging of crustal and mantle structures (e.g., Moschetti *et al.*, 2010; Lin *et al.*, 2011; Xie *et al.*, 2013; Shaw *et al.*, 2015; Lynner *et al.*, 2018; Ai *et al.*, 2020; Wang *et al.*, 2020; Zhang *et al.*, 2020; Zheng, 2023; Cheng *et al.*, 2025). Compared to body waves, surface waves propagate primarily along Earth's surface in a quasi-2D manner (Aki and Richards, 2002) and can be efficiently extracted from ambient-noise cross correlations (e.g., Bensen *et al.*, 2007; Nakata *et al.*, 2019; Yang *et al.*, 2026). This has led to the development and widespread application of ambient-noise tomography, which significantly alleviates limitations associated with the uneven spatial and temporal distribution of natural earthquakes (e.g., Shapiro and Campillo, 2004; Lin *et al.*, 2008). Although full 3D inversion methods for surface-wave data (e.g., Fang *et al.*, 2015; Sager *et al.*, 2018; Zhang *et al.*, 2018; Wang *et al.*, 2019) offer higher accuracy and have been successfully applied in various regions, the conventional approach—constructing 2D surface-wave velocity maps (e.g., Barmin *et al.*, 2001; Lin *et al.*, 2009), followed by inversion for piecewise 1D velocity structures and assembling them into a pseudo-3D model—remains widely used (e.g., Zigone *et al.*, 2015; Shen *et al.*, 2016; Mordret *et al.*, 2019; Qiu *et al.*, 2019).

---

1. Hubei Subsurface Multi-Scale Imaging Key Laboratory, School of Geophysics and Geomatics, China University of Geosciences, Wuhan, China, https://orcid.org/0000-0002-4219-0039 (HQ); https://orcid.org/0000-0001-6157-8584 (YZ); https://orcid.org/0000-0001-9349-5820 (SA); https://orcid.org/0000-0001-5933-0029 (LQ); 2. State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan, China; 3. School of Geological Engineering and Geomatics, Chang'an University, Xi'an, China; 4. National Precise Gravity Measurement Facility, Huazhong University of Science and Technology, Wuhan, China; 5. Hunan Earthquake Agency, Changsha, China, https://orcid.org/0000-0001-8992-9550 (SH)

*Corresponding authors: qiuhongrui@cug.edu.cn; qiuhonrui@gmail.com; hujing@chd.edu.cn

This is primarily due to its simplicity and its ability to provide a first-order approximation of the true structure in most cases.

The inversion of piecewise 1D velocity structures in surface-wave tomography is based on locally measured surface-wave observations, including phase and group velocities. These observations are inherently dispersive, meaning their sensitivity to subsurface velocity structure varies with frequency (Aki and Richards, 2002). As a result, different frequencies sample different depths, enabling the recovery of depth-dependent velocity structures from dispersion measurements (e.g., Wathelet *et al.*, 2008; Herrmann, 2013). Given the predominant sensitivity of surface-wave dispersion to shear-wave velocity (V_S) over compressional-wave velocity (V_P), we hereinafter refer to this process as 1D V_S inversion (e.g., Zhou *et al.*, 2004). The essence of 1D V_S inversion is to parameterize the local subsurface structure as a layered V_S model, whereas the corresponding V_P and density are constrained via prescribed empirical relationships (e.g., a constant V_P/V_S ratio and density) to calculate theoretical dispersion curves. The inversion then seeks a set of 1D V_S models whose theoretical surface-wave dispersion curves match the observed dispersion measurements within their estimated uncertainties (e.g., Shen *et al.*, 2013; Berg *et al.*, 2018).

Several established approaches can effectively address the inversion problem with good computational efficiency and satisfactory data fit in most scenarios. These range from classic linearized inversion methods (e.g., Herrmann, 2013) to forward-modeling-based global optimization methods such as Markov chain Monte Carlo (MCMC; e.g., Shen *et al.*, 2016; Berg *et al.*, 2018) and particle swarm optimization (PSO; e.g., Song *et al.*, 2012), alongside emerging machine learning-based (ML-based) methods (e.g., Hu *et al.*, 2020; Luo *et al.*, 2022; Huang *et al.*, 2023). While efficient, linearized inversions are often ill-conditioned and susceptible to initialization bias, potentially converging to local minima. This has prompted a shift toward global optimization and ML-based methods; the latter, in particular, show significant promise in mitigating dependence on the starting model. Nevertheless, the reliability and generalizability of ML-based approaches across diverse datasets still require further rigorous validation (Hu *et al.*, 2020; Cai *et al.*, 2022).

Forward-modeling-based inversion approaches directly integrate a forward solver with optimization algorithms to efficiently search the model space for solutions that provide an acceptable fit to the dispersion observations. Because exhaustive sampling is computationally infeasible, these methods rely on prior constraints and nonlinear optimization algorithms (e.g., Zhang and Curtis, 2024) to guide the search. Nevertheless, achieving robust convergence and a reliable characterization of non-uniqueness typically requires an extremely large number of forward evaluations. When conventional physics-based forward modeling (e.g., Herrmann, 2013) is used, the computational cost of repeated forward calculations becomes prohibitive, even with Numba-accelerated implementations such as the Python dispersion solver DISBA (Luu, 2021). Consequently, the efficiency of forward modeling constitutes the primary computational bottleneck, motivating the development of even more efficient forward operators.

Leveraging advances in ML and graphic processing unit (GPU) acceleration, we propose a fast deep learning-based surrogate forward operator—an efficient neural-network inference scheme—that learns a reliable mapping from 1D velocity profiles to theoretical surface-wave phase-velocity dispersion curves. Such surrogate forward operators have been successfully applied to a range of geophysical inversion problems (e.g., Shahriari *et al.*, 2020; Wu *et al.*, 2023; Chang *et al.*, 2024; Liu *et al.*, 2025) in recent years, demonstrating their ability to substantially reduce computational cost while maintaining high forward-modeling accuracy. Compared with end-to-end, fully data-driven deep learning inversion approaches (e.g., Hu *et al.*, 2020; Cai *et al.*, 2022; Huang *et al.*, 2023), surrogate forward operators offer three key advantages: (1) they preserve the physical structure of the traditional inversion workflow by explicitly computing forward responses and minimizing data misfit, thereby ensuring physical interpretability; (2) they are used solely to accelerate forward modeling, avoiding the potential "black-box" behavior inherent in direct data-to-model mappings and improving robustness and controllability; and (3) they can be seamlessly integrated with existing nonlinear optimization algorithms (e.g., PSO and MCMC), significantly improving computational efficiency without altering the inversion strategy, enabling more thorough exploration of the model space and a reliable characterization of the inherent non-uniqueness of the inversion problem. The trade-off is that surrogate forward models still require iterative optimization rather than producing one-shot inverse predictions, but they provide a more transparent and controllable pathway for incorporating deep learning into geophysical inversion.

In this study, we develop a gated recurrent unit (GRU)-based surrogate forward model (FGRU) to improve the efficiency of the surface-wave dispersion inversion, thereby enabling substantially denser exploration of the parameter space than is feasible with conventional physics-based forward solvers. The architecture and training procedure of the FGRU forward operator are described in the GRU-based forward operator section, and its computational performance is evaluated against the conventional DISBA forward solver using a consumer-grade workstation. To illustrate its robustness and effectiveness in accelerating 1D V_S inversion, the FGRU forward operator is integrated into a particle swarm optimization framework (FGRU–PSO; Fig. 1) and evaluated using synthetic tests (see the FGRU-based *Vs* inversion section) and a field dataset from South China (see the Application to South China section). The field-data application suggests that the high computational efficiency of the FGRU–PSO framework enables a more thorough exploration of the velocity model space, as reflected by clearer V_S structural imaging in regions with thick sedimentary cover compared with results from traditional inversion methods.

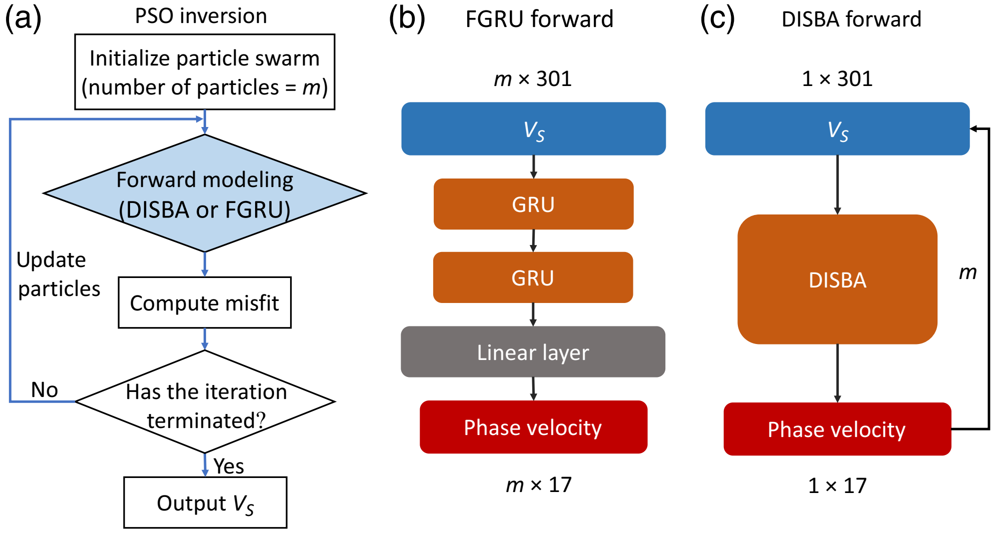

Figure 1. Particle swarm optimization (PSO)-based surface-wave inversion and forward-modeling strategies. (a) In PSO, a swarm of m particles (candidate V_S models) is iteratively updated based on the misfit between predicted and observed dispersion curves. (b) FGRU forward modeling evaluates the m particles in parallel: each particle is a 301-layer V_S profile (0.5 km spacing, 0–150 km), and batch inference through stacked GRU layers and a linear output layer yields m phase-velocity curves sampled at 17 periods. (c) DISBA forward modeling evaluates particles sequentially. The GRU cell structure is shown in Figure S1. DISBA, a physics-based forward solver; and FGRU, a gated recurrent unit (GRU)-based surrogate forward model. The color version of this figure is available only in the electronic edition.

GRU-Based Forward Operator

The surface-wave dispersion forward problem involves computing theoretical phase or group velocities for a given layered Earth model. In this work, we use DISBA as the reference ("ground truth") solver because of its robustness and long-standing use in computing Rayleigh-wave dispersion curves (Aquino *et al.*, 2022; Korostelev *et al.*, 2022). Although DISBA provides highly accurate forward modeling of surface-wave dispersion with accelerated implementations, each evaluation requires the explicit numerical solution of the propagator-matrix equations, leading to an approximately fixed computational cost per model. When complex velocity structures are considered or when a large number of forward evaluations is required, as in nonlinear global-search inversion algorithms, the cumulative computational cost becomes substantial. The resulting overhead from repeated forward calculations, therefore, constitutes a major efficiency bottleneck in practical inversion workflows.

A variety of deep learning architectures have been explored to accelerate geophysical applications (Cui *et al.*, 2023; Wu *et al.*, 2023; Chang *et al.*, 2024). Building on these advances, we propose a fast forward surrogate model based on the GRU neural network (Cho *et al.*, 2014; Chung *et al.*, 2014), hereinafter referred to as FGRU, which approximates the nonlinear mapping from the V_S models to their corresponding theoretical dispersion curves after training. The GRU architecture is a gated variant of recurrent neural networks designed to efficiently capture sequential dependencies while alleviating vanishing-gradient issues and maintaining a compact model structure. These properties make GRU particularly well-suited for modeling the depth-ordered structure of 1D velocity profiles and their associated period-dependent surface-wave dispersion, while retaining low computational and training complexity. Although GRU is a suitable choice for accelerating forward dispersion calculations in this study, it is not necessarily the optimal architecture among all possible deep learning models. A systematic evaluation of alternative network architectures is beyond the scope of this work; however, exploring and optimizing network design for improved performance represents a promising direction for future research.

The FGRU architecture (Fig. 1b), which is similar to the network structure proposed by Wang *et al.* (2025), consists of two stacked bidirectional GRU layers with 32 and 128 hidden units, respectively, designed to extract sequential features from the depth-dependent V_S profile (an $N \times 1$ vector, in which N is the number of sublayers). The GRU layers are concatenated and mapped through a fully connected (linear) layer to produce an $M \times 1$ (M is the number of sampling periods) vector of theoretical phase velocities at the target periods, in which M is the number of sampled periods. The FGRU model was trained on an NVIDIA RTX 3060 GPU. Using the hyperparameters summarized in Table S1, available in the supplemental material to this article, training took approximately 6 hr and exhibited stable convergence on both the training and validation datasets (Fig. S2). Replacing numerical forward modeling with fast surrogate inference substantially reduces computational cost by effectively learning a mapping between plausible velocity models and their theoretical phase-velocity dispersion curves. Once trained, the surrogate enables the inversion to exploit rapid neural-network inference rather than repeatedly solving the forward problem numerically across iterations and grid points.

The generalization capability of the surrogate forward operator is primarily governed by the diversity and representativeness of its training dataset. To ensure broad coverage of the

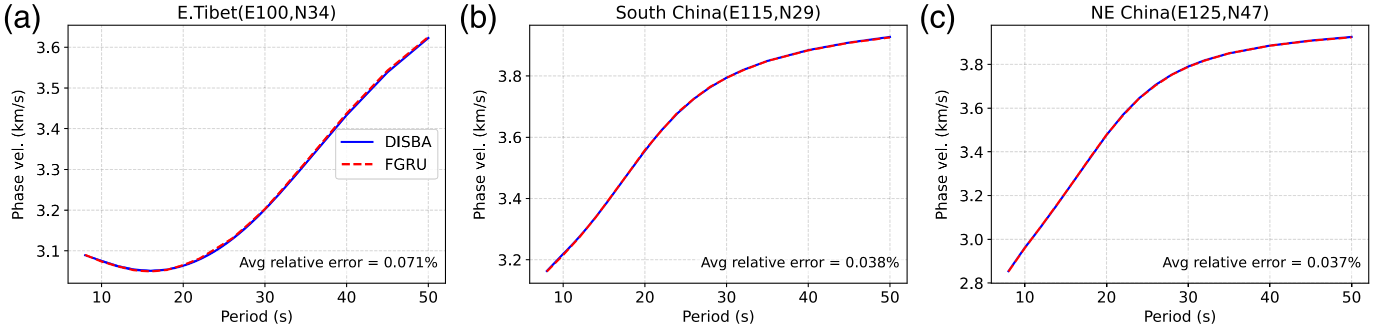

Figure 2. Accuracy assessment of Rayleigh-wave phase velocity dispersion predicted by FGRU (red), benchmarked against the DISBA solver (blue), at three representative locations: (a) Eastern Tibet, (b) South China, and (c) Northeast China. Relative errors are shown in the lower-right corner of each panel. The spatial distribution of relative errors across all grid points in South China is shown in Figure S5. The color version of this figure is available only in the electronic edition.

model space, the training dataset is generated by constructing 1D V_S models using a basis-function (cubic B-spline) parameterization commonly employed in large-scale seismic velocity inversions (e.g., Shen *et al.*, 2016). This parameterization captures continuous and physically plausible velocity variations using a small number of parameters, thereby avoiding the high-dimensional redundancy introduced by layered models with thin fixed layers. As a result, it enables more efficient and diverse sampling of the model space. The same parameterization is also used during the inversion and for the real-observational data in later sections, ensuring structural consistency and robustness between the training and application of the surrogate forward operator.

Following Shen *et al.* (2016), we interpolate the V_S profile at each grid point using 13 cubic B-spline parameters, including sediment thickness, V_S values at the upper- and lower-sediment boundaries, and Moho depth. The crustal and upper-mantle V_S structures are represented by four and five different cubic B-spline basis functions, respectively (Fig. S3). The training dataset is constructed through the following steps. First, 1D V_S profiles across China are extracted from the Crust1.0 model (Laske *et al.*, 2013). Second, each profile is represented using 13 parameters based on cubic B-spline functions. Finally, these parameters at each grid point are randomly perturbed within prescribed ranges (Table S2) to generate an ensemble of V_S profiles, yielding a total of 500,000 models in the training dataset. The distribution of this dataset is shown in Figure S4. This ensures broad coverage of the parameter space while keeping the computational cost manageable. Each parameter set is then converted into a V_S model consisting of 301 layers with 0.5-km spacing over the 0–150 km depth range, using cubic B-spline interpolation. V_P values are derived from V_S using the Brocher (2005) empirical relationship above 120 km depth and a fixed V_P/V_S ratio of 1.79 between 120 and 150 km (Kennett *et al.*, 1995). Density is calculated from V_S following Brocher (2005). Each velocity model is paired with its theoretical Rayleigh-wave phase-velocity dispersion curve sampled at 17 periods—matching the period range of the subsequent field data application (8–50 s)—computed using the DISBA solver. Collectively, these V_S profiles and their corresponding dispersion curves constitute the training dataset.

We evaluated the accuracy and efficiency of the trained FGRU forward operator using 1D velocity models from Shen *et al.* (2016). For a fair comparison, all dispersion forward evaluations are conducted on the same consumer-grade workstation equipped with an NVIDIA RTX 3060 GPU and an Intel i7-12700 central processing unit (CPU). In terms of accuracy, FGRU forward predictions closely match the DISBA results (Fig. 2 and Fig. S5), with average relative errors predominantly below 0.05%, well below the typical 1%–5% measurement uncertainty of observed dispersion curves. In terms of efficiency, the largest speedup arises from combining the learned surrogate with GPU-based batched inference: the per-model runtime of FGRU decreases rapidly with increasing number of forward evaluations and stabilizes at $\sim 10^{-5}$ s per model on the RTX 3060 GPU once the total number of evaluations exceeds 10^5 (purple circles in Fig. 3). In contrast, the per-model runtime of the DISBA solver decreases by only about one order of magnitude—from $\sim 10^{-2}$ to $\sim 10^{-3}$ s per model—as the number of forward evaluations increases from fewer than 100 to more than 1000, even with 16-core parallelization on the i7-12700 CPU (purple triangles in Fig. 3). As a result, for ultra-large numbers of forward evaluations ($>10^5$), FGRU achieves nearly a 50-fold improvement in computational efficiency over DISBA (Table S3).

Overall, the FGRU forward operator provides a highly accurate and computationally efficient alternative. This is critical for nonlinear inversion methods (e.g., PSO and MCMC) that rely on massive numbers of forward calculations, as it enables more thorough exploration of the model space, reduces the risk of convergence to local minima, and allows a more reliable characterization of the inherent nonuniqueness of the inversion problem.

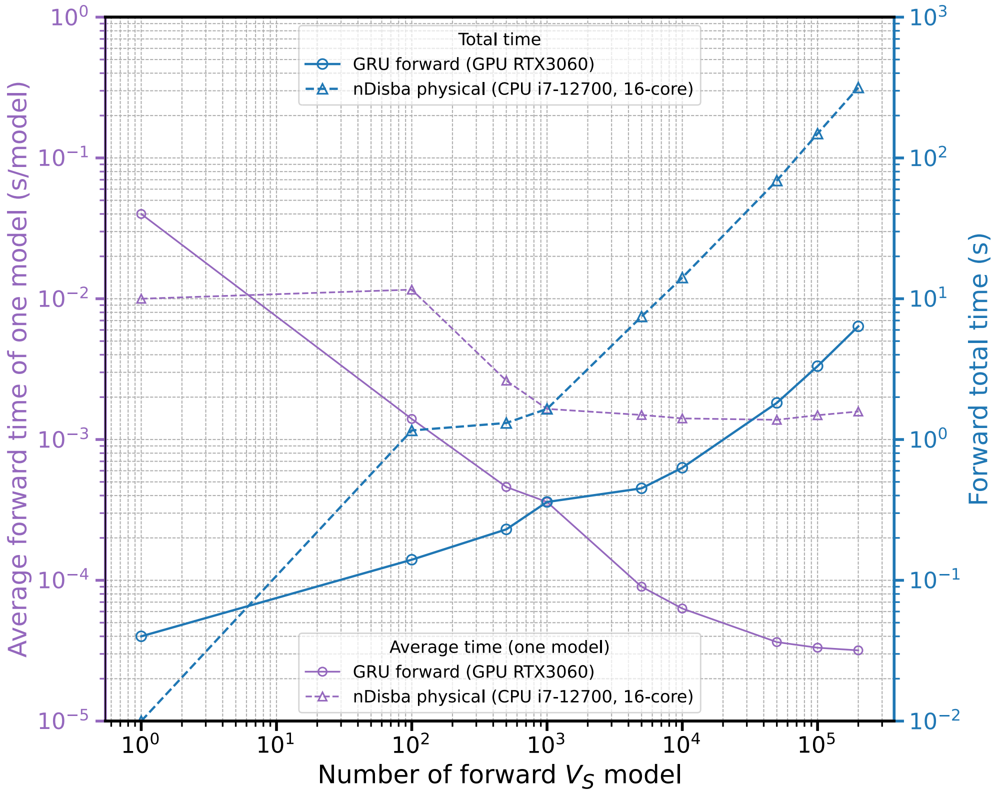

Figure 3. Computational efficiency (purple) and total runtime (blue) of FGRU (circles) versus the DISBA solver (triangles) across different forward-modeling scales (detailed values are listed in Tables S3 and S4). CPU, central processing unit; GPU, graphics processing unit. The color version of this figure is available only in the electronic edition.

FGRU-Based V_S Inversion

In 1D V_S inversion, surface-wave velocity structures are commonly parameterized in three ways: (1) a fixed number of layers with prescribed thicknesses and variable V_S values in each layer (e.g., Herrmann, 2013); (2) a fixed number of layers with both variable layer thicknesses and V_S values (e.g., Marc *et al.*, 2020); and (3) a smooth representation using basis functions, such as a sum of cubic B-spline curves (e.g., Shen *et al.*, 2016). Each parameterization offers distinct advantages and limitations in terms of model flexibility, stability, and computational efficiency. To enable a fair comparison between our field-data results presented in a later section and the model inverted by Shen *et al.* (2016), we adopt the same basis-function parameterization for the 1D V_S model as described in the previous section.

The next step is to select an optimization strategy that can efficiently explore a prescribed parameter space associated with the chosen model parameterization. Conventional MCMC methods (Shen *et al.*, 2016; Berg *et al.*, 2018) evaluate model misfit sequentially along Markov chains during the random-walk process and are therefore not well suited to exploiting the batch-parallel inference capability of FGRU. In contrast, PSO evaluates a large number of particles—each representing a candidate V_S model—simultaneously at each iteration and updates their positions based on both individual and global optima. Therefore, we adopt PSO rather than MCMC to demonstrate the potential of replacing the DISBA solver with FGRU for accelerating 1D V_S inversion, fully leveraging the computational efficiency of the GRU framework. The PSO-based inversion aims to minimize an objective function defined as the chi-square misfit (χ) between the predicted (c_i^{pred}) and observed (c_i^{obs}) dispersion curves:

$$\chi = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{c_i^{\mathrm{pred}} - c_i^{\mathrm{obs}}}{\sigma_i^{\mathrm{obs}}}\right)^2}, \tag{1}$$

in which σ_i^{obs} represents the uncertainty of the observed dispersion curve at the ith period. N is the total number of data points on the dispersion curve. It should be noted that the inversion scheme presented here is not fully optimized, and future studies may explore improved combinations of parameterization strategies and optimization methods.

To evaluate the robustness and efficiency of the FGRU-based V_S inversion, we extract three representative 1D velocity models from the model of Shen *et al.* (2016) at grid points (100° E, 34° N) in Tibet, (115° E, 29° N) in South China, and (125° E, 47° N) in North China, and compute their corresponding theoretical phase-velocity dispersion curves using the DISBA solver. Gaussian random noise is added to mimic real observational data. In the PSO framework, the parameter search ranges are shown in Figure S6. The inertia weight w and the cognitive ($c1$) and social acceleration ($c2$) coefficients are all set to 0.5. All inversion tests employ 100 particles per iteration over 100 iterations, resulting in a total of 10,000 forward evaluations. On the same consumer-grade workstation, the FGRU–PSO inversion completes in approximately 2 s using the RTX 3060 GPU, whereas the DISBA–PSO implementation requires substantially longer runtimes (Table 1), taking about 16 s even with 16-core parallelization on an Intel i7-12700 CPU. The good data fits (Fig. 4a, Figs. S7a and S8a) and the consistency between the inverted and true V_S models (Fig. 4b, Figs. S7b and S8b) at depths where the data have sufficient

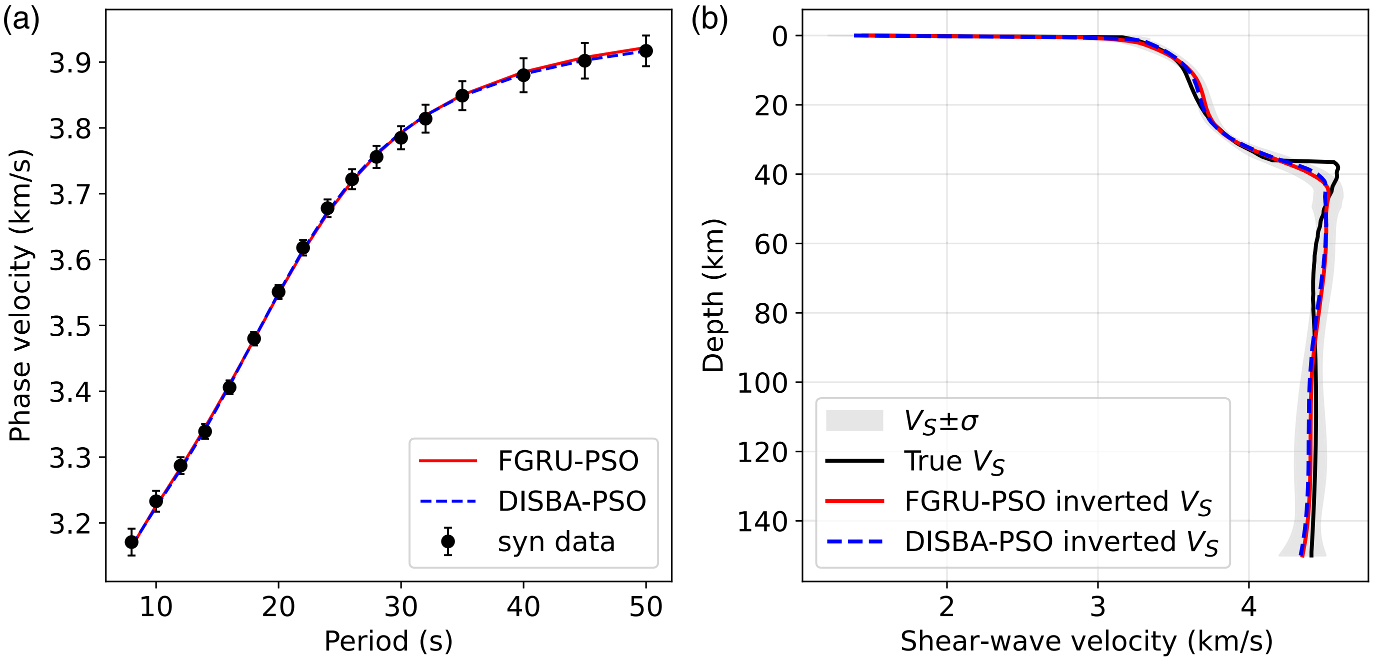

Figure 4. Synthetic inversion results for the South China grid point (115° E, 29° N). (a) Comparison between the input (synthetic) and predicted phase velocity curves; error bars denote the prescribed observational uncertainty. (b) Comparison between the inverted and true V_S models. The gray shaded region indicates the uncertainty of the FGRU–PSO inverted V_S model, estimated from 100 independent FGRU–PSO inversions. The color version of this figure is available only in the electronic edition.

sensitivity (i.e., above the Moho) demonstrate the robustness of both methods. These results indicate that FGRU–PSO effectively exploits batch parallelism, enabling efficient execution on both CPUs and GPUs, and substantially reducing computational cost while preserving inversion consistency under the tested settings. We note that the reported GPU runtime of approximately 2 s is not a strict lower bound, and further reductions may be achievable with additional implementation refinements and optimization.

Application to South China

We further applied the FGRU–PSO method to Rayleigh-wave phase-velocity dispersion measurements in South China. The observed phase velocities were taken from Shen *et al.* (2016); details of the dataset are provided in Text S1 and Figure S9. Because the efficiency and robustness of the FGRU–PSO framework are validated using synthetic datasets, we leverage its exceptional computational efficiency not merely to speed up the inversion, but to substantially enhance exploration of the model space. Specifically, we increase the number of evaluated velocity models to 1,000,000 at each grid point while keeping the computational cost manageable (~165 s), a scale of sampling that would be infeasible with the DISBA solver. This dramatic increase in sampling density allows for a much more thorough exploration of the model space, thereby increasing the likelihood of convergence to the global minimum. To assess the reliability of field-data application, we perform 100 independent PSO inversions at each grid point and use the spread of the resulting models as a measure of solution

TABLE 1
Comparison of Inversion Runtime for Different Forward-Inversion Frameworks

Inversion Framework	Model Count (Particles × Iterations)	Hardware Platform	Runtime	Speedup (Relative to DISBA–PSO CPU Baseline)
FGRU–PSO	100 × 100	GPU (RTX 3060)	2 s	~45.5×
FGRU–PSO	100 × 100	CPU (Intel i7-12700)	5 s	~18.2×
DISBA–PSO	100 × 100	CPU (Intel i7-12700)	91 s	1.0× (baseline)
nDISBA–PSO (parallel)	100 × 100	CPU (16-Cores Parallel)	16 s	~5.7×

CPU, central processing unit; DISBA–PSO, a particle swarm optimization framework incorporating the physics-based DISBA forward solver; FGRU–PSO, a particle swarm optimization framework incorporating the FGRU forward operator; and GPU, graphics processing units.

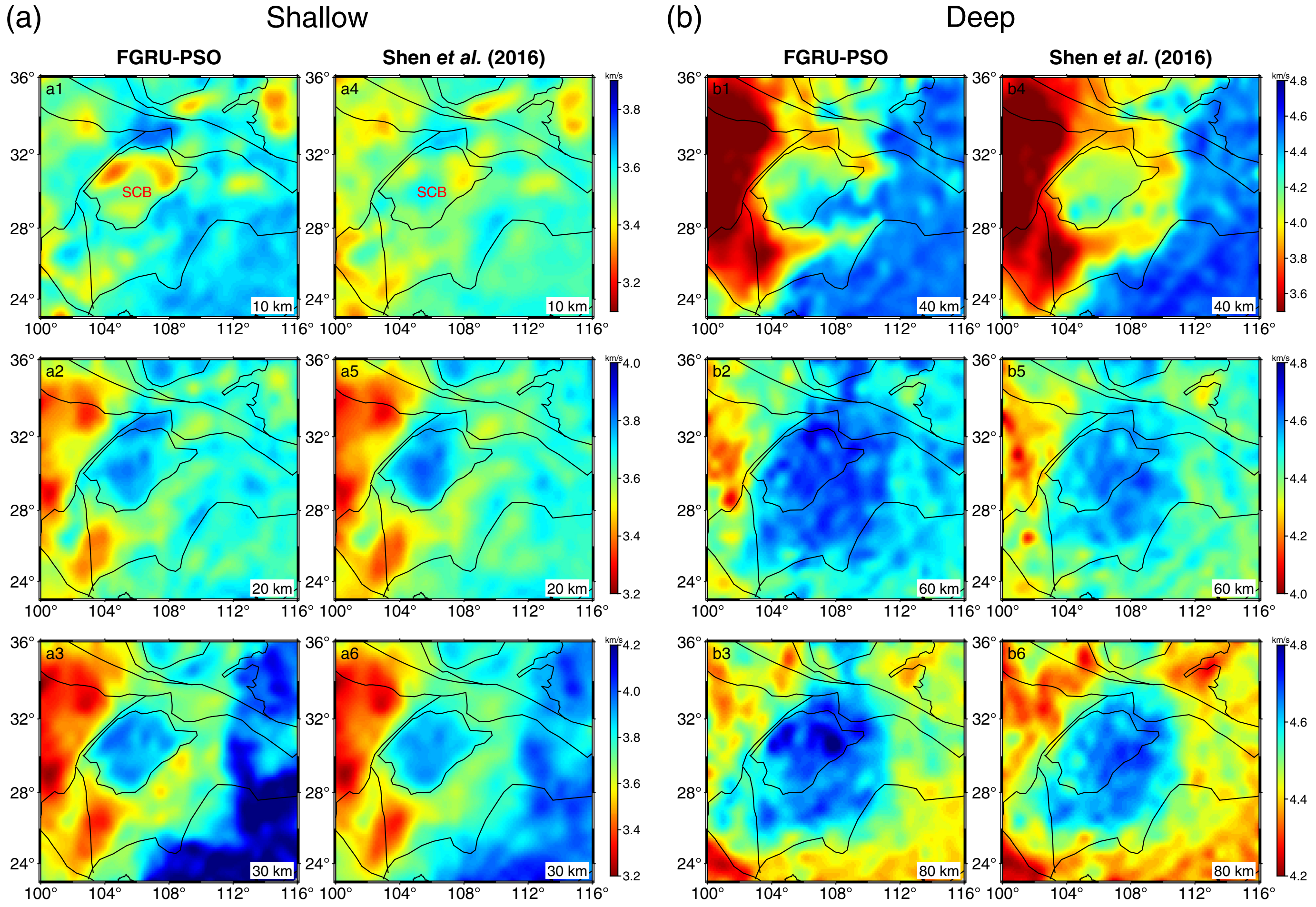

Figure 5. Depth slices of inverted V_S models from the FGRU–PSO method and Shen *et al.* (2016). The depth value of each slice is indicated in the bottom-right corner, with shallower slices shown in (a) and deeper slices shown in (b). SCB, Sichuan basin. The color version of this figure is available only in the electronic edition.

stability under the adopted parameterization and inversion settings (Fig. S10).

Comparison shown in Figure 5 and Figure S11 indicates good overall consistency between our V_S model and that of Shen *et al.* (2016). At a shallow depth of 10 km, pronounced low-velocity anomalies are observed beneath the Sichuan basin (SCB), consistent with its thick sedimentary cover. In the depth range of 20–40 km, corresponding to mid-to-lower crust, prominent low-velocity zones are observed beneath the Songpan-Ganzi, Sichuan-Yunnan, and West Qinling blocks, in good agreement with previous tomographic and geodynamic studies (Zheng *et al.*, 2010, 2013; Wang *et al.*, 2016; Guo *et al.*, 2022). These large-scale low-velocity features likely reflect crustal eastward flow driven by sustained compression from the Tibetan Plateau. In the uppermost mantle (60–80 km), the stable lithospheric root of the Yangtze craton is characterized by a high-velocity anomaly in both models, a hallmark of cratonic lithosphere. In contrast, the lithosphere beneath Eastern China exhibits lower velocities, supporting the hypothesis of craton destruction or lithospheric thinning in this region.

The overall agreement between our V_S model and that of Shen *et al.* (2016), along with the successful recovery of well-known structural features from the shallow crust to the upper mantle, supports the robustness and reliability of the FGRU–PSO inversion scheme. However, a closer, zoom-in comparison reveals several notable differences at smaller spatial scales. The most striking difference lies in the imaging of the SCB. The V_S model obtained from the FGRU–PSO inversion yields two shallow low-velocity anomalies (red ellipses in Fig. 6a) in the northern SCB, extending continuously from the surface to approximately 10 km depth and forming a crescent-shaped low-velocity zone. In contrast, the model of Shen *et al.* (2016) does not clearly capture these anomalies and shows limited lateral continuity (Fig. 6c), a feature consistent with assessments of V_S models in South China using full-wave simulations (Miao *et al.*, 2025). Our FGRU–PSO results are also in close agreement with velocity model derived from oil wells and seismic reflection profiles (Wang *et al.*, 2016), as well as with joint body- and surface-wave inversion results (Gao *et al.*, 2022; Han *et al.*, 2022). The continuous low-velocity anomalies in the northern SCB are well explained by the thick sedimentary cover in this region.

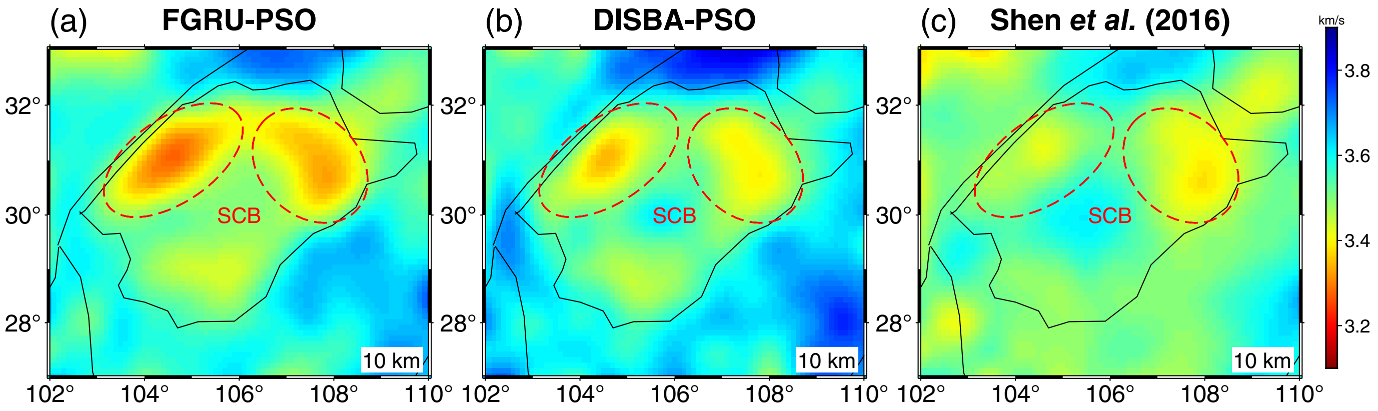

Figure 6. Zoomed-in comparison of V_S maps near the SCB at 10 km depth. (a) V_S model derived from the FGRU–PSO inversion; (b) V_S model derived from the benchmark DISBA–PSO inversion, conducted under the same total runtime as panel (a) but with fewer evaluated models per grid point; and (c) V_S model from Shen *et al.* (2016). Red dashed ellipses outline major low-velocity anomalies, which are more pronounced in the FGRU–PSO model. The color version of this figure is available only in the electronic edition.

These observations indicate that the FGRU–PSO method can better resolve V_S structures in sedimentary basins than the approach of Shen *et al.* (2016).

Because both inversions use identical dispersion data and model parameterization, the enhanced shallow low-velocity features recovered by FGRU–PSO in sedimentary basins are most likely a result of the denser sampling of the parameter space enabled by the accelerated forward modeling of FGRU. To more rigorously isolate this effect, we further compare FGRU–PSO with DISBA–PSO using the same inversion algorithm and identical total computational time (Fig. 6a,b). Under this controlled setting, DISBA–PSO performs 100,000 model evaluations at each grid point (the parameter settings described in Text S1), whereas FGRU–PSO can evaluate up to 1,000,000 models within the same total runtime due to its faster forward predictions, thereby increasing the sampling density of the parameter space. This improved sampling increases the likelihood of converging to the global minimum rather than local minima, particularly for grid points within the SCB. Misfit maps (Figs. S12 and S13) support this interpretation, showing consistently lower misfit for FGRU–PSO (average misfit ≈ 0.6) compared with Shen *et al.* (2016). These results underscore the practical value of deep learning-based surrogate forward models for improving regional seismic imaging.

Discussion and Conclusions

We demonstrate that 1D V_S inversion can be substantially accelerated by replacing conventional physics-based surface-wave forward modeling with a deep learning-based surrogate forward operator. Using the state-of-the-art DISBA solver as a reference, we show that, after training on a dataset generated from the low-resolution global model Crust1.0, FGRU can map layered V_S models to Rayleigh-wave phase-velocity dispersion curves with high fidelity (average relative error $< 0.05\%$). Moreover, FGRU provides exceptional computational acceleration—over an order of magnitude faster than the DISBA solver on a consumer-grade workstation when a large number of forward models are evaluated simultaneously. This gain arises from GPU-based batched inference enabled by the learned surrogate model.

Specifically, FGRU achieves $\sim 10^{-5}$ s per model under batch inference using GPU (RTX 3060), compared with $\sim 10^{-3}$ s per model for the DISBA solver with 16-core CPU (Intel i7-12700) parallelization (Fig. 3). This pronounced performance difference primarily reflects the efficiency of GPU-parallelized batch inference. We note that the comparison is not strictly hardware equivalent, as the DISBA solver is not currently GPU compatible, whereas FGRU is naturally suited for GPU acceleration. To provide a more direct comparison, we also benchmarked both forward operators using the same CPU (Table S3). Although the performance gap is reduced under this configuration, FGRU still outperforms the DISBA solver by several times. The initial GPU-based comparison is not intended to diminish the efficiency of the DISBA solver, but rather to highlight the readily available and easily implemented GPU acceleration enabled by the FGRU framework, in contrast to the absence of a GPU implementation for the DISBA solver at present.

Although FGRU may not represent the optimal architecture for accelerating surface-wave dispersion calculations, it is sufficiently effective to illustrate the central idea of this study—that improving the efficiency of forward modeling alone can lead to enhanced performance in 1D V_S inversion. This acceleration enables much denser sampling of the parameter space in forward-modeling-based inversion frameworks, such as PSO and MCMC, as well as other highly parallelizable stochastic search and optimization algorithms, under fixed computational budgets. To demonstrate this, we adopt PSO and integrate it with both FGRU and the DISBA solver to assess the acceleration performance of FGRU relative to DISBA in the context of 1D V_S inversion. We do not include a comparison based on the same inversion scheme (e.g., combining FGRU with MCMC), because conventional MCMC approaches cannot efficiently exploit the GPU-based batched inference that underlies the computational advantage of the proposed method.

In a synthetic test using the same consumer-grade workstation, the FGRU–PSO framework completes 10,000 forward evaluations in ~2 s on the GPU, whereas the DISBA–PSO workflow requires substantially longer runtimes (~16 s) on the CPU with 16-core parallelization (Table 1). We note that the current workflow is not yet fully optimized; for example, particle updates in PSO are performed on the CPU while forward evaluations are executed on the GPU, leading to additional overhead from CPU–GPU data transfers. Even without fully optimized parameterization and optimization strategies, these results highlight the potential of FGRU as an effective replacement for DISBA in large-scale, computationally intensive 1D V_S inversion. Although FGRU predictions may introduce small numerical errors for each forward evaluation, these do not accumulate across iterations because each evaluation in PSO is independent. As long as the surrogate errors remain much smaller than the observational uncertainties, their impact on the inversion results is negligible.

Application of the FGRU–PSO framework to Rayleigh-wave phase-velocity dispersion data from Shen *et al.* (2016) in South China further demonstrates the practical value of this approach. Leveraging the high efficiency of FGRU, we evaluate up to one million velocity models at each grid point while maintaining a manageable computational cost (~165 s), a level of sampling that would be prohibitive with the DISBA solver. The enhanced shallow low-velocity features recovered by FGRU–PSO in sedimentary basins—consistent with previous local studies—indicate that the denser parameter-space sampling enabled by accelerated forward modeling increases the likelihood of convergence to the global minimum rather than local minima, thereby improving imaging quality.

As with all ML models trained on finite datasets, the accuracy of FGRU depends on the completeness and representativeness of the training dataset—that is, whether the key features of real 1D V_S profiles (e.g., absolute velocities and depth gradients) in the target region are adequately sampled. In this study, we construct a robust training dataset using a widely adopted large-scale parameterization (Shen *et al.*, 2016), ensuring broad coverage of physically reasonable crust–mantle velocity variations. However, in cases where the training dataset is inevitably incomplete—for example, in imaging shallow structures characterized by strong heterogeneity and limited prior constraints—the accuracy of FGRU may be reduced. A potential solution to this limitation is to adopt an iterative, incremental training strategy. Specifically, the accuracy of FGRU can be evaluated using the inverted models, and those that show poor agreement with the DISBA solver can be identified. Additional training samples are then generated in the vicinity of these misfit-prone models to expand the training dataset, and the network is retrained. This process can be repeated iteratively until FGRU achieves consistently high accuracy across all inverted models.

The present network is trained and applied using a widely adopted cubic B-spline parameterization, which provides sufficient variability in the training dataset to ensure robust learning of the FGRU model. However, we note that the FGRU framework is not restricted to this specific parameterization and can, in principle, be extended to alternative representations of 1D V_S structure (e.g., models with a fixed number of layers but variable layer thicknesses and layer-top velocities, with velocity assumed to vary linearly with depth within each layer). A systematic evaluation of how different parameterizations influence model performance is beyond the scope of this study, but the transferability of the trained network across parameterization schemes remains an important topic for future investigation. In addition, because FGRU is fully differentiable, future work could exploit this property to compute gradients of dispersion curves with respect to model parameters via backpropagation, and to compare these neural-network-derived Jacobians with conventional sensitivity kernels (Fig. S14). Such developments may provide a pathway toward more efficient and physically informed gradient-based inversion frameworks.

In conclusion, we develop a deep learning-based surrogate forward operator, FGRU that substantially improves computational efficiency relative to the physics-based DISBA solver while maintaining high fidelity, after training on datasets generated from a low-resolution global model. We demonstrate that this acceleration enables much denser exploration of the parameter space in 1D V_S inversion when combined with optimization algorithms such as PSO, as shown through both synthetic tests and field-data applications. Unlike direct deep learning inversion schemes, the key contribution of this work is the use of deep learning to accelerate the forward solver, while the inversion itself is still carried out through explicit data-misfit evaluation and optimization. This preserves a more transparent and controllable framework for regional surface-wave tomography. This approach is particularly promising for imaging shallow structures, where strong heterogeneity and limited prior constraints make dense parameter-space sampling essential to avoid local minima and better characterize the inherent non-uniqueness of the inversion.

Data and Resources

The training dataset, trained gated recurrent unit-based surrogate forward model (FGRU), Rayleigh-wave phase-velocity dispersion data measured by Shen *et al.* (2016) for South China, and the code package to reproduce this work is available at https://github.com/spaopaos/dqy. The DISBA solver is available at https://github.com/keurfonluu/disba. Both the websites were last accessed in May 2026. The supplemental material for this article includes Text S1, Tables S1–S4, and Figures S1–S14, which provide details of the FGRU hyperparameters, model perturbation ranges, forward-modeling runtime benchmarks, training behavior, model parameterization, synthetic and field-data inversion results, misfit analysis, and sensitivity-kernel comparisons.

Declaration of Competing Interests

The authors acknowledge that there are no conflicts of interest recorded.

Acknowledgments

The authors thank the two anonymous reviewers for their constructive and insightful comments, as well as the editors for their time and effort devoted to this article. This study is supported by the National Key Research and Development Program of China (2024YFF0807300). Jing Hu is supported by the National Natural Science Foundation of China (NSFC; Grant Number 42104051) and the Deep Earth Probe and Mineral Resources Exploration - National Science and Technology Major Project of China (2024ZD1000408). Hongrui Qiu and Yong Zheng are also supported by the NSFC grants (Grant Numbers 42574126, 42274082, and 42030108), the Open Fund of SINOPEC Key Laboratory of Geophysics (36750000-24-FW0399-0003) and the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan).

References

Ai, S., Y. Zheng, and S. Wang (2020). Crustal deformations of the central North China Craton constrained by radial anisotropy, *J. Geophys. Res.: Solid Earth* **125,** e2019JB018374, doi: 10.1029/2019JB018374.

Aki, K., and P. G. Richards (2002). *Quantitative Seismology*, Second Ed., University Science Books, Sausalito, California.

Aquino, M., G. Marquis, and J. Vergne (2022). Joint one-dimensional inversion of magnetotelluric data and surface-wave dispersion curves using correspondence maps, *Geophys. Prospect.* **70,** 1455–1470, doi: 10.1111/1365-2478.13239.

Barmin, M. P., M. H. Ritzwoller, and A. L. Levshin (2001). A fast and reliable method for surface wave tomography, *Pure Appl. Geophys.* **158,** 1351–1375, doi: 10.1007/PL00001225.

Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F.-C. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2007). Processing seismic ambient noise data to obtain reliable broadband surface wave dispersion measurements, *Geophys. J. Int.* **169,** no. 3, 1239–1260, doi: 10.1111/j.1365-246X.2007.03374.x.

Berg, E. M., F.-C. Lin, A. Allam, H. Qiu, W. Shen, and Y. Ben-Zion (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,** 9933–9949, doi: 10.1029/2018JB016269.

Brocher, T. M. (2005). Empirical relations between elastic wavespeeds and density in the earth's crust, *Bull. Seismol. Soc. Am.* 95**,** no. 6, 2081–2092, doi: 10.1785/0120050077.

Cai, A., H. Qiu, and F. Niu (2022). Semi-supervised surface wave tomography with Wasserstein cycle-consistent GAN: Method and application to the Southern California plate boundary region, *J. Geophys. Res.: Solid Earth* **127,** e2021JB023598, doi: 10.1029/2021JB023598.

Chang, Y., D. Yang, and X. He (2024). A deep learning operator-based numerical scheme method for solving 1D wave equations, *J. Geophys. Eng.* **21,** no. 4, 1194–1215, doi: 10.1093/jge/gxae062.

Cheng, F., J. Xia, Z. Li, C. Shi, and X. Chen (2025). Guided P-wave extracting and imaging with underground urban seismic noise, *Geophys. J. Int.* **243,** no. 1, ggaf311, doi: 10.1093/gji/ggaf311.

Cho, K., B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation, *Proc. of the EMNLP 2014*, 1724–1734, doi: 10.3115/v1/D14-1179.

Chung, J., C. Gulcehre, K. Cho, and Y. Bengio (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling, *NIPS 2014 Workshop on Deep Learning*, available at http://arxiv.org/abs/1412.3555 (last accessed May 2026).

Cui, D., L. Shi, and K. Gao (2023). Rapid construction of Rayleigh wave dispersion curve based on deep learning, *Front. Earth Sci.* **10,** 1084414, doi: 10.3389/feart.2022.1084414.

Fang, H., H. Yao, H. Zhang, Y.-C. Huang, and R. D. van der Hilst (2015). Direct inversion of surface wave dispersion for three-dimensional shallow crustal structure based on ray tracing: methodology and application, *Geophys. J. Int.* **201,** 1251–1263, doi: 10.1093/gji/ggv080.

Gao, L., H. Zhang, L. Gao, C. He, H. Xin, and W. Shen (2022). High-resolution Vs tomography of South China by joint inversion of body wave and surface wave data, *Tectonophysics* **824,** 229228, doi: 10.1016/j.tecto.2022.229228.

Guo, Z., S. Li, Y. Yu, Y. J. Chen, Y. Yang, B. Xu, and X. Liang (2022). Eastward asthenospheric flow from NE Tibet inferred by joint inversion of teleseismic body and surface waves: Insight into widespread continental deformation in Eastern China, *J. Geophys. Res.: Solid Earth* **127,** e2022JB024410, doi: 10.1029/2022JB024410.

Han, S., H. Zhang, L. Gao, Y. Liu, C. Chai, and M. Maceira (2022). Joint inversion of body wave arrival times, surface wave dispersion data and receiver functions: Method and application to South China, *J. Geophys. Res.: Solid Earth* **127,** e2022JB024083, doi: 10.1029/2022JB024083.

Herrmann, R. B. (2013). Computer programs in seismology: An evolving tool for instruction and research, *Seismol. Res. Lett.* **84,** no. 6, 1081–1088, doi: 10.1785/0220110096.

Hu, J., H. Qiu, H. Zhang, and Y. Ben-Zion (2020). Using deep learning to derive shear-wave velocity models from surface-wave dispersion data, *Seismol. Res. Lett.* **91,** 1738–1751, doi: 10.1785/0220190222.

Huang, X., Z. Yu, W. Wang, and F. Wang (2023). JointNet: A multimodal deep learning-based approach for joint inversion of Rayleigh-wave dispersion and ellipticity, *Bull. Seismol. Soc. Am.* **114,** 627–641, doi: 10.1785/0120230199.

Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995). Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.* **122,** 108–124, doi: 10.1111/j.1365-246X.1995.tb03540.x.

Korostelev, F., Y. Lu, F. Magrini, L. Boschi, S. Leroy, and W. Vétel (2022). Images of the East African Rift System by global adaptive-resolution surface-wave tomography, *J. Geophys. Res.: Solid Earth* **127,** no. 6, e2021JB023570, doi: 10.1029/2021JB023570.

Laske, G., G. Masters, Z. Ma, and M. E. Pasyanos (2013). Update on CRUST1.0—A 1-degree global model of Earth's crust, *Geophys. Res. Abstr. 15, EGU2013-2658.*

Lin, F.-C., M. P. Moschetti, and M. H. Ritzwoller (2008). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.* **173,** 281–298, doi: 10.1111/j.1365-246X.2008.03720.x.

Lin, F.-C., M. H. Ritzwoller, and R. Snieder (2009). Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array, *Geophys. J. Int.* **177,** 1091–1110, doi: 10.1111/j.1365-246X.2009.04105.x.

Lin, F.-C., M. H. Ritzwoller, Y. Yang, M. P. Moschetti, and M. J. Fouch (2011). Complex and variable crustal and uppermost mantle seismic anisotropy in the western United States, *Nature Geosci.* **4,** 55–61, doi: 10.1038/ngeo1036.

Liu, W. Y., C. C. Yin, Y. Su, B. Zhang, S. Xue, Y. Liu, and X. Ren (2025). Trans-dimensional Bayesian inversion of airborne electromagnetic data based on artificial neural network, *Chin. J. Geophys.* **68,** no. 8, 2928–2940, doi: 10.6038/cjg2025S0572 (in Chinese).

Luo, Y., Y. Huang, Y. Yang, K. Zhao, X. Yang, and H. Xu (2022). Constructing shear velocity models from surface wave dispersion curves using deep learning, *J. Appl. Geophys.* **196,** 104524, doi: 10.1016/j.jappgeo.2021.104524.

Luu, K. (2021). DISBA: Numba-accelerated computation of surface wave dispersion (Version 0.6.0) [Computer software], *Zenodo*, doi: 10.5281/zenodo.5637204.

Lynner, C., S. L. Beck, G. Zandt, R. W. Porritt, F.-C. Lin, and Z. C. Eilon (2018). Midcrustal deformation in the Central Andes constrained by radial anisotropy, *J. Geophys. Res.: Solid Earth* **123,** 4798–4813, doi: 10.1029/2017JB014936.

Marc, W., J.-L. Chatelain, C. Cornou, G. Di Giulio, B. Guillier, M. Ohrnberger, and A. Savvaidis (2020). Geopsy: A user-friendly open-source tool set for ambient vibration processing, *Seismol. Res. Lett.* **91,** no. 3, 1878–1889, doi: 10.1785/0220190360.

Miao, W. P., G. L. Li, F. L. Niu, K. Tao, and Y. H. Li (2025). Assessment of shear wave velocity models in the southeast Qinghai-Xizang plateau with full-wave simulation, *Earthq. Sci.* **38,** doi: 10.1016/j.eqs.2025.01.007.

Mordret, A., P. Roux, P. Boué, and Y. Ben-Zion (2019). Shallow three-dimensional structure of the San Jacinto fault zone revealed from ambient noise imaging with a dense seismic array, *Geophys. J. Int.* **216,** 896–905, doi: 10.1093/gji/ggy464.

Moschetti, M. P., M. H. Ritzwoller, F.-C. Lin, and Y. Yang (2010). Seismic evidence for widespread western-US deep-crustal deformation caused by extension, *Nature* **464,** 885–889, doi: 10.1038/nature08951.

Nakata, N., L. Gualtieri, and A. Fichtner (Editors) (2019). *Seismic Ambient Noise*, Cambridge University Press, Cambridge, United Kingdom, doi: 10.1017/9781108264808.

Qiu, H., F.-C. Lin, and Y. Ben-Zion (2019). Eikonal tomography of the Southern California plate boundary region, *J. Geophys. Res.: Solid Earth* **124,** 9755–9779, doi: 10.1029/2019JB017806.

Sager, K., L. Ermert, A. Boehm, J. van Driel, B. Krischer, D. S. Peter, and A. Fichtner (2018). Towards full waveform ambient noise inversion, *Geophys. J. Int.* **212,** 566–590, doi: 10.1093/gji/ggx429.

Shahriari, M., D. Pardo, B. Moser, and F. Sobieczky (2020). A deep neural network as surrogate model for forward simulation of borehole resistivity measurements, *Procedia Manuf.* **42,** 235–238, doi: 10.1016/j.promfg.2020.02.075.

Shapiro, N. M., and M. Campillo (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.* **31,** L07614, doi: 10.1029/2004GL019491.

Shaw, J. H., A. Plesch, C. Tape, M. P. Suess, T. H. Jordan, G. Ely, E. Hauksson, J. Tromp, T. Tanimoto, R. Graves, *et al.* (2015). Unified structural representation of the Southern California crust and upper mantle, *Earth Planet. Sci. Lett.* **415,** 1–15, doi: 10.1016/j.epsl.2015.01.016.

Shen, W., M. H. Ritzwoller, D. Kang, Y. Kim, F.-C. Lin, J. Ning, W. Wang, Y. Zheng, and L. Zhou (2016). A seismic reference model for the crust and uppermost mantle beneath China from surface wave dispersion, *Geophys. J. Int.* **206,** no. 2, 954–979, doi: 10.1093/gji/ggw175.

Shen, W., M. H. Ritzwoller, V. Schulte-Pelkum, and F.-C. Lin (2013). Joint inversion of surface wave dispersion and receiver functions: A Bayesian Monte-Carlo approach, *Geophys. J. Int.* **192,** 807–836, doi: 10.1093/gji/ggs050.

Song, X., L. Tang, X. Lv, H. Fang, and H. Gu (2012). Application of particle swarm optimization to interpret Rayleigh wave dispersion curves, *J. Appl. Geophys.* **84,** 1–13, doi: 10.1016/j.jappgeo.2012.05.011.

Wang, F., X. Zhang, and X. Song (2025). Surface-wave tomography with transfer learning and Moho constraints: Method and application to China Mainland, *Seismol. Res. Lett.* doi: 10.1785/0220240437.

Wang, K., C. Jiang, Y. Yang, V. Schulte-Pelkum, and Q. Liu (2020). Crustal deformation in Southern California constrained by radial anisotropy from ambient noise adjoint tomography, *Geophys. Res. Lett.* **47,** e2020GL088580, doi: 10.1029/2020GL088580.

Wang, K., Q. Liu, and Y. Yang (2019). Three-dimensional sensitivity kernels for multicomponent empirical Green's functions from ambient noise: Methodology and application to adjoint tomography, *J. Geophys. Res.: Solid Earth* **124,** 5794–5810, doi: 10.1029/2018JB017020.

Wang, M., J. Hubbard, A. Plesch, J. H. Shaw, and L. Wang (2016). Three-dimensional seismic velocity structure in the Sichuan basin, China, *J. Geophys. Res.: Solid Earth* **121,** 1007–1022, doi: 10.1002/2015JB012644.

Wathelet, M., D. Jongmans, M. Ohrnberger, and S. Bonnefoy-Claudet (2008). Array performances for ambient vibrations on a shallow structure and consequences over Vs inversion, *J. Seismol.* **12,** 1–19.

Wu, S. H., Q. H. Huang, and L. Zhao (2023). Fast Bayesian inversion of airborne electromagnetic data based on the invertible neural network, *IEEE Trans. Geosci. Remote Sens.* **61,** 5907211, doi: 10.1109/TGRS.2023.3264777.

Xie, J., M. H. Ritzwoller, W. Shen, Y. Yang, Y. Zheng, and L. Zhou (2013). Crustal radial anisotropy across eastern Tibet and the western Yangtze craton, *J. Geophys. Res.: Solid Earth* **118,** 4226–4252, doi: 10.1002/jgrb.50296.

Yang, Y., Z. Li, T. Xie, J. Xie, X. Chen, L. Zhou, and G. Li (2026). A review of recent advances in methods for extracting broadband surface waves from ambient noise data, *Sci. China Earth Sci.* **69,** 839–861, doi: 10.1007/s11430-025-1803-x.

Zhang, X., and A. Curtis (2024). Bayesian variational time-lapse full waveform inversion, *Geophys. J. Int.* **237,** no. 3, 1624–1638, doi: 10.1093/gji/ggae129.

Zhang, X., A. Curtis, E. Galetti, and S. de Ridder (2018). 3-D Monte Carlo surface wave tomography, *Geophys. J. Int.* **215,** 1644–1658, doi: 10.1093/gji/ggy362.

Zhang, X., F. Hansteen, A. Curtis, and S. de Ridder (2020). 1D, 2D and 3D Monte Carlo ambient noise tomography using a dense passive seismic array installed on the North Sea seabed, *J. Geophys. Res.: Solid Earth* **125,** e2019JB018552, doi: 10.1029/2019JB018552.

Zheng, Y. (2023). Southwest China community velocity model V2.0: The key to decipher active tectonics and seismic hazards in the

Sichuan–Yunnan region, *Sci. China Earth Sci.* **66,** 2395–2398, doi: 10.1007/s11430-023-1179-4.

Zheng, Y., C. Ge, Z. J. Xie, Y. J. Yang, X. Xiong, and H. Hsu (2013). Crustal and upper mantle structure and the deep seismogenic environment in the source regions of the Lushan earthquake and the Wenchuan earthquake, *Sci. China Earth Sci.* **56,** no. 7, 1158–1168, doi: 10.1007/s11430-013-4641-2.

Zheng, Y., Y. Yang, and M. H. Ritzwoller (2010). Crustal structure of the northeastern Tibetan plateau, the Ordos block and the Sichuan basin from ambient noise tomography, *Earthq. Sci.* **23,** 465–476.

Zhou, Y., F. A. Dahlen, and G. Nolet (2004). Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.* **158,** 142–168, doi: 10.1111/j.1365-246X.2004.02324.x.

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

Manuscript received 20 January 2026