mathematics

MDPI

Article

Residual Physics-Informed Neural Networks for Seismic Tomography with Multi-Source Prior Constraints

Yanjin Xiang, Ziang Song, Zhiliang Wang and Guojie Song *

School of Sciences, Southwest Petroleum University, Chengdu 610500, China; yanzinxyz@gmail.com (Y.X.); haluze0725@gmail.com (Z.S.); 202199010053@swpu.edu.cn (Z.W.)
* Correspondence: cylsgj@126.com

Abstract

Seismic traveltime tomography is essential for constructing subsurface velocity models that underpin high-resolution imaging and inversion. Traditional ray- and eikonal-based methods are sensitive to the starting model and lack a unified, physically consistent framework to integrate seismic data with high-confidence prior information. PINN-based approaches offer flexible, grid-free inversion but often suffer from training instability and limited use of prior constraints. We propose a Multi-Source Prior-Guided Residual Physics-Informed Neural Network (MSP-ResPINNs) to address these limitations through two key technical advancements. First, MSP-ResPINNs integrates a Residual Network (ResNet) architecture with Sinusoidal Representation (SIREN) activation to replace standard MLPs, ensuring the robust capture of high-frequency velocity gradients. Second, the framework implements a unified loss function that rigorously enforces multi-source constraints, including well-logs and geological horizons. Numerical experiments demonstrate that MSP-ResPINNs accurately reconstructs sharp velocity contrasts and complex geological features compared with conventional PINN-based approaches. Among the tested variants, the multiplicative factorization consistently provides the most stable and physically consistent results, outperforming the additive factorization.

Keywords: seismic tomography; eikonal equation; multi-source fusion; physics-informed neural networks

MSC: 86-10

1. Introduction

Seismic tomography is a foundational technique for investigating subsurface velocity structures. At global and regional scales, seismic tomography images mantle and crustal anomalies by inverting seismic traveltimes and waveforms, providing critical insights into Earth's dynamic processes [1,2]. At the exploration scale, seismic tomography constructs background velocity models, forming the essential basis for high-resolution imaging methods like depth migration and full-waveform inversion. The accuracy of these background models directly controls the reliability of structural imaging and reservoir characterization [3–7].

Traditional traveltime tomography, based on ray theory or the eikonal equation, estimates velocity structures by minimizing misfits between observed and calculated traveltimes. Ray-based methods, reliant on explicit tracing, can become unstable in complex media and computationally expensive for large 3D problems [8–10]. Eikonal-based methods improve stability by solving traveltime fields but typically require strong smoothing

Check for updates

Academic Editor: Ioannis Tsoulos

Received: 15 January 2026
Revised: 10 March 2026
Accepted: 14 March 2026
Published: 21 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

regularization to mitigate ill-posedness, which often blunts sharp velocity contrasts and fine-scale features [11–13]. A common and profound challenge for these traditional approaches lies in their high sensitivity to the starting model and in the lack of a unified and physically consistent framework that can simultaneously assimilate seismic data with other high-confidence prior information, such as well-log velocities or known geological interfaces, within the inversion process.

The advent of scientific machine learning, particularly Physics-Informed Neural Networks (PINNs), offers a promising paradigm shift. PINNs provide a general framework for solving forward and inverse problems governed by partial differential equations (PDEs) by embedding the physical laws directly into the neural network's loss function [14]. PINNs allow the incorporation of diverse prior information, such as point measurements, boundary conditions, and interface constraints, within a unified optimization procedure. In geophysical applications, PINNs have demonstrated considerable potential in eikonal-based traveltime modeling [15–18], time-domain and frequency-domain wave equation modeling [19–22], seismic tomography [15,23–25] and full waveform inversion [26–29].

PINN-based formulations have been shown to reconstruct seismic traveltime fields from sparse observations and infer subsurface velocity structures without explicit grid discretization or ray tracing [15,30]. Subsequent studies have focused on improving the robustness and flexibility of PINN-based traveltime tomography frameworks through the introduction of Bayesian formulations [29,31,32], hard data enforcement strategies [33], limited incorporation of well-log velocity constraints [33,34], and the adoption of effective-slowness-based factorized eikonal equations to enhance convergence behavior and multiparameter inversion capability [25]. Despite these advances, most PINN-based traveltime tomography studies have been validated primarily on simplified or smoothly varying velocity models, and the systematic, physically consistent integration of diverse high-confidence prior information commonly available in exploration-scale applications remains insufficient.

In exploration-scale seismic inversion, subsurface velocity modeling is commonly constrained not only by seismic traveltime observations but also by a wide range of geological and geophysical prior information. Such prior information typically includes well-log velocity measurements, interpreted geological interfaces and stratigraphic horizons, as well as auxiliary constraints derived from integrated geophysical surveys. For example, Chen et al. [35] constructed high-quality initial velocity models for full-waveform inversion by interpolating well-log velocities under geological structure constraints extracted from reverse time migration images. Heir et al. [36] embedded poststack seismic attributes, geological information, and field calibration data such as well tops and interpreted horizons as discrete features into deep learning models to predict porosity and acoustic impedance. Zhou et al. [37] incorporated horizon position encoding together with seismic traces to improve the inversion of gamma-ray logs. Such prior information provides localized or layer-specific constraints that are often more reliable than seismic inversion results alone, particularly in near-surface or structurally complex settings, thereby reducing inversion nonuniqueness, enhancing model resolution, and improving the geological consistency of subsurface models.

Motivated by the demonstrated value of geological and geophysical prior information in seismic inversion, we propose the Multi-Source Prior-Guided Residual Physics-Informed Neural Network (MSP-ResPINNs) framework. Unlike conventional PINN-based tomography, which often relies on simplistic architectures and limited data types, our approach delivers improvements in two key dimensions. (1) We integrate a Residual Neural Network (ResNet [38]) architecture with Sinusoidal Representation Network (SIREN) activation for seismic traveltime tomography. This hybrid design effectively addresses the gradient vanishing issues inherent in deep MLPs and significantly enhances the network's capac-

ity to represent high-frequency velocity variations and sharp geological discontinuities. (2) We develop a unified, physically consistent loss function that assimilates multi-source constraints, including eikonal-based physics residuals, sparse seismic traveltimes, pinpoint well-log measurements, and interpreted geological horizons. This framework systematically reduces the non-uniqueness of the inverse problem and ensures that the reconstructed models honor both physical laws and geological data. To further clarify the positioning of MSP-ResPINNs within the current research landscape, Table 1 provides a comparative analysis of the proposed framework against traditional and existing PINN-based methodologies.

Table 1. Comparative analysis of seismic tomography methodologies.

Method	Methodology	Pros	Cons/Limitations	References
Traditional Tomography	Ray-tracing/Grid-based eikonal	Mature	Sensitive to initial model, smoothing artifacts	[8–13]
Standard PINNs	Fully-connected NN + eikonal loss	Grid-free, unified inversion	Training instability, struggles with sharp interfaces	[15,30]
Advanced PINNs	Factorized eikonal/Bayesian PINN	Handles singularity, robust convergence	Limited integration of diverse geological priors	[25,29,31–34]
MSP-ResPINNs	ResNet + SIREN + PINNs + Multi-source Priors	High structural fidelity, stable training	–	This study

The remainder of this paper is organized as follows. Section 2 reviews the eikonal equation and its multiplicative and additive factorized formulations, followed by an overview of PINN-based seismic traveltime inversion. The proposed MSP-ResPINNs framework and the associated evaluation metrics are then introduced. Section 3 presents numerical experiments to demonstrate the effectiveness of the proposed method. Finally, Section 4 summarizes the main conclusions and outlines potential directions for future research.

2. Methodology

To enhance the interpretability of the mathematical formulations and ensure clarity across the domains of geophysics and machine learning, a comprehensive summary of symbols, indices, and acronyms used in this study is provided in Abbreviations section.

2.1. Eikonal Equation and Its Factorization

In high-frequency seismic wave propagation, the first-arrival traveltime field satisfies the eikonal equation, which constitutes the theoretical foundation of traveltime tomography. Let $T(\mathbf{x})$ denote the seismic traveltime from a source location $\mathbf{x}_s$ to a spatial point $\mathbf{x} \in \Omega \subset \mathbb{R}^d$, where d is the spatial dimension. The eikonal equation is given by:

$$|\nabla T(\mathbf{x})|^2 = \frac{1}{\mathbf{v}^2(\mathbf{x})}, \quad \forall \mathbf{x} \in \Omega, \tag{1}$$

subject to the source condition:

$$T(\mathbf{x_s}) = 0, \tag{2}$$

where $\mathbf{v}(\mathbf{x})$ is the seismic velocity model.

However, the traveltime field generally exhibits a strong singularity at the source location, which poses challenges for numerical solvers and learning-based approaches such as PINNs [15,25,30,33]. Specifically, while $T(\mathbf{x})$ is continuous, its derivatives approach infinity as $\mathbf{x} \to \mathbf{x}_s$, leading to numerical instability and potential gradient explosion dur-

ing automatic differentiation. Moreover, the inherent smoothness of standard activation functions (e.g., Sigmoid or Tanh) results in a representation mismatch with the sharp, cone-shaped geometry of the traveltime field near the source. These issues create an optimization imbalance, where excessively large PDE residuals near the source dominate the total loss, hindering the network's ability to resolve far-field velocity structures.

To alleviate this difficulty, factorization techniques are commonly employed to separate the known singular behavior from the smooth component of the solution [39]. By reformulating the problem to learn a smooth correction factor rather than the raw traveltime, the unbounded gradients are eliminated from the optimization target. In this study, both multiplicative and additive factorizations of the eikonal equation are considered.

2.1.1. Multiplicative Factorization

In the multiplicative factorization approach, the traveltime field is decomposed as:

$$T(\mathbf{x}) = T_0(\mathbf{x})\tau(\mathbf{x}), \tag{3}$$

where $T_0(\mathbf{x})$ is a known background traveltime that captures the singularity at the source, and $\tau(\mathbf{x})$ is a smooth correction term to be learned. A common choice for $T_0(\mathbf{x})$ is the analytical solution in a homogeneous medium with a constant reference velocity $\mathbf{v}(\mathbf{x_s})$, given by:

$$T_0(\mathbf{x}) = \frac{|\mathbf{x} - \mathbf{x}_s|}{\mathrm{v}(\mathbf{x}_s)}, \tag{4}$$

Substituting the multiplicative factorization Equation (3) into the eikonal Equation (1) yields:

$$|\nabla T_0(\mathbf{x}) \cdot \tau(\mathbf{x}) + T_0(\mathbf{x})\nabla\tau(\mathbf{x})|^2 = \frac{1}{\mathrm{v}^2(\mathbf{x})}, \quad \forall \mathbf{x} \in \Omega, \tag{5}$$

The source condition is naturally enforced by:

$$\tau(\mathbf{x}_s) = 1, \tag{6}$$

2.1.2. Additive Factorization

In the additive factorization approach, the traveltime field is decomposed as:

$$T(\mathbf{x}) = T_0(\mathbf{x}) + \tau(\mathbf{x}), \tag{7}$$

where $T_0(\mathbf{x})$ is the background traveltime defined above, and $\tau(\mathbf{x})$ is a dimensionless smooth factor.

Substituting the additive factorization Equation (7) into the eikonal Equation (1) yields:

$$|\nabla T_0 + \nabla\tau|^2 = \frac{1}{\mathrm{v}^2(\mathbf{x})}, \tag{8}$$

The source condition is naturally enforced by:

$$\tau(\mathbf{x}_s) = 0, \tag{9}$$

2.2. Physics-Informed Neural Network for Seismic Tomography

Based on the multiplicative factorization introduced in Section 2.1, a physics-informed neural network framework is employed to jointly approximate the traveltime field and infer the subsurface velocity model. Two neural networks are used to approximate the traveltime factor $\hat{\tau}_\theta(\mathbf{x}, \mathbf{x}_s)$ and the velocity model $\hat{\mathrm{v}}_\phi(\mathbf{x})$ [15]. The network used to predict $\hat{\tau}_\theta(\mathbf{x}, \mathbf{x}_s)$ takes the spatial coordinates and the source location as inputs, whereas the network used to

predict $\hat{v}_\phi(\mathbf{x})$ depends only on the spatial coordinates. Spatial derivatives required by the governing equation are computed via automatic differentiation [40]. The physics constraint is imposed by minimizing the residual of the multiplicatively factorized eikonal equation:

$$\mathcal{R}_{\text{Eik}}(\mathbf{x}) = |\hat{\tau}_\theta(\mathbf{x})\nabla T_0(\mathbf{x}) + T_0(\mathbf{x})\nabla\hat{\tau}_\theta(\mathbf{x})|^2 - \frac{1}{\hat{v}_\phi^2(\mathbf{x})}, \tag{10}$$

which is enforced at a set of collocation points $\{\mathbf{x}_i\}_{i=1}^{N_f}$ within the computational domain, where N_f denotes the total number of collocation points. The corresponding physics loss is defined as:

$$\mathcal{L}_{\text{Eik}} = \frac{1}{N_f}\sum_{i=1}^{N_f}|\mathcal{R}_{\text{Eik}}(\mathbf{x}_i)|^2, \tag{11}$$

The source condition is incorporated through:

$$\mathcal{L}_{\text{src}} = \frac{1}{N_s}\sum_{i=1}^{N_s}|\hat{\tau}_\theta(\mathbf{x}_s) - 1|^2, \tag{12}$$

where N_s is the number of sources.

The observational constraint is enforced by minimizing the misfit between predicted and observed first-arrival traveltimes at receiver locations:

$$\mathcal{L}_{\text{data}} = \frac{1}{N_d}\sum_{j=1}^{N_d}|T_0(\mathbf{x}_r^j)\hat{\tau}_\theta(\mathbf{x}_r^j) - T_j^{\text{obs}}|^2, \tag{13}$$

where N_d denotes the total number of receiver measurements.

The overall training objective is formulated as a weighted sum of the above loss terms:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{Eik}}\mathcal{L}_{\text{Eik}} + \lambda_{\text{src}}\mathcal{L}_{\text{src}} + \lambda_{\text{data}}\mathcal{L}_{\text{data}}, \tag{14}$$

where λ_{Eik}, λ_{src}, and λ_{data} are weighting parameters.

2.3. The Proposed MSP-ResPINNs Framework

The MSP-ResPINNs framework is shown in Figure 1. MSP-ResPINNs improves both the network representation capability and the physical regularization of the inversion by incorporating residual network architectures and multiple sources of prior velocity information, the corresponding pseudocode is presented in Algorithm 1. The key components of MSP-ResPINNs are detailed below.

2.3.1. Residual Network Architecture

Standard fully connected PINN architectures may suffer from limited expressive capability and inefficient gradient propagation when modeling complex velocity structures. To enhance representation power, we employ a residual network architecture [38] with sinusoidal activation functions (SIREN [41]) for both the traveltime factor network and the velocity network.

As illustrated in Figure 2, the residual architecture introduces skip connections that enable efficient gradient flow and facilitate the learning of high-frequency components. Let $\mathbf{h}_l$ denote the hidden feature representation at the l-th residual block. Each block updates the features according to:

$$\mathbf{h}_{l+1} = \mathbf{h}_l + \mathcal{F}(\mathbf{h}_l), \tag{15}$$

where $\mathcal{F}(\cdot)$ denotes a residual mapping implemented by a fully connected subnetwork composed of multiple SIREN layers, each consisting of a linear transformation followed by a sinusoidal activation function.

Algorithm 1: Training Procedure of MSP-ResPINNs

Input: Observed traveltimes T^{obs}; Well-logs $\{\mathbf{x}_w^k, v_w^k\}$; Interfaces $\{\mathbf{x}_\Gamma^i, v_\Gamma^i\}$; Domain Ω; Background factor $T_0(\mathbf{x})$; Learning rate η.

Output: Optimized velocity $\hat{v}_\phi$ and traveltime factor $\hat{\tau}_\theta$.

Initialization:

Initialize θ for $\hat{\tau}_\theta$ and ϕ for $\hat{v}_\phi$ using Xavier initialization;

Data Preparation:

Generate collocation points $\{\mathbf{x}_j\}_{j=1}^{N_f} \in \Omega$ via Latin Hypercube Sampling;

while not converged **do**

`// Step 1: Forward Pass`

Predict $\hat{\tau}_\theta$ and $\hat{v}_\phi$ via Residual Blocks with SIREN activation;

Reconstruct $\hat{T}(\mathbf{x})$ using multiplicative (Equation (3)) or additive (Equation (7)) strategy;

`// Step 2: Physics-based Differentiation`

Compute spatial gradients $\nabla\hat{T}(\mathbf{x})$ via Automatic Differentiation (AD);

`// Step 3: Multi-objective Loss Evaluation`

$\mathcal{L}_{\text{eik}} \leftarrow$ Eikonal residual (Equation (11));

$\mathcal{L}_{\text{src}} \leftarrow$ Source condition (Equation (12));

$\mathcal{L}_{\text{data}}, \mathcal{L}_{\text{well}}, \mathcal{L}_{\text{int}} \leftarrow$ Data-driven constraints (Equations (13)–(17));

$\mathcal{L}_{\text{total}} = \sum \lambda_i \mathcal{L}_i$ (Equation (18));

`// Step 4: Joint Optimization`

Compute gradients: $g_\theta = \nabla_\theta \mathcal{L}_{\text{total}}$, $g_\phi = \nabla_\phi \mathcal{L}_{\text{total}}$;

Update $\{\theta, \phi\} \leftarrow \text{Adam}(g_\theta, g_\phi, \eta)$;

end

return $\hat{v}_\phi, \hat{\tau}_\theta$;

The superiority of this integrated architecture in PINN-based inversion stems from the synergistic effects of residual learning and periodic activation. From the perspective of optimization, the identity mapping within the residual blocks ensures that the gradient of the loss function during backpropagation can be expressed as $\frac{\partial \mathbf{h}_{l+1}}{\partial \mathbf{h}_l} = \mathbf{I} + \frac{\partial \mathcal{F}}{\partial \mathbf{h}_l}$. The presence of the identity term $\mathbf{I}$ allows gradients to bypass complex nonlinear mappings, effectively mitigating the gradient vanishing problem that can be exacerbated by the differential operators inherent in eikonal-based loss functions. Complementing this, the differential consistency of SIREN provides a unique advantage over traditional activations like Tanh or ReLU. Since the derivative of a sine function is a phase-shifted cosine, the derivatives required for the eikonal equation remain well-behaved and preserve rich high-frequency structural information. Consequently, the MSP-ResPINNs framework can learn incremental corrections to the identity mapping while accurately reconstructing sharp velocity gradients and fine-scale geological interfaces, ultimately improving the representation of multiscale components within the subsurface model.

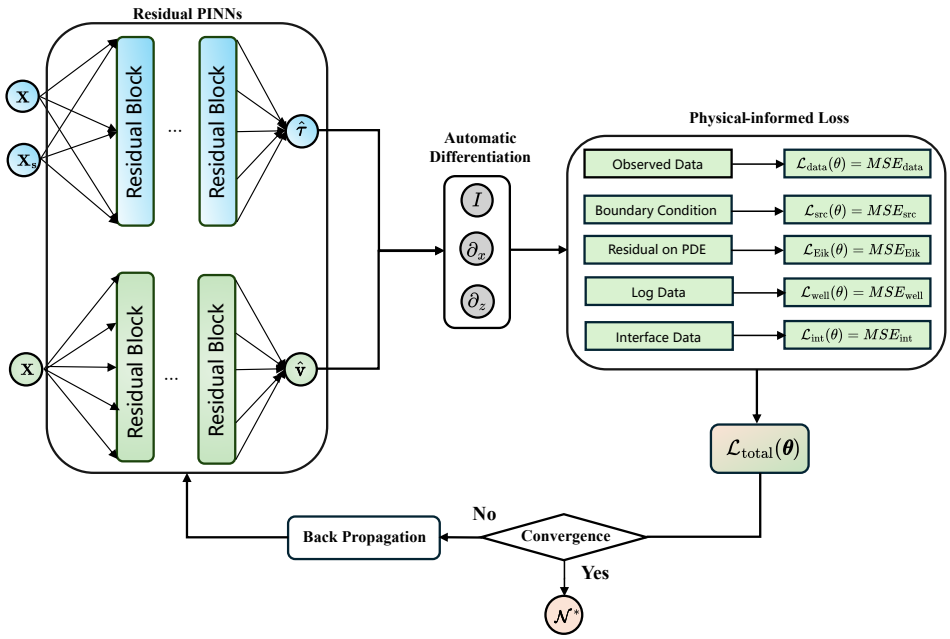


Figure 1. Overview of the proposed MSP-ResPINNs framework. The computational flow geometrically links the mathematical series as follows: the input coordinates $\mathbf{x}_s$, $\mathbf{x}$ drive the approximation of traveltime factors $\hat{\tau}$ (Equation (3) or (7)) and velocity $\hat{v}$ (Equation (11)) via Residual Blocks. Through automatic differentiation, these outputs are mapped to the physics-informed loss $\mathcal{L}_{\text{total}}$ (Equation (18)), which integrates the eikonal residual $\mathcal{L}_{\text{Eik}}$ (Equation (11)), source conditions $\mathcal{L}_{\text{src}}$ (Equation (13)), and multi-source priors including well-logs $\mathcal{L}_{\text{well}}$ (Equation (16)) and interfaces $\mathcal{L}_{\text{int}}$ (Equation (17)).

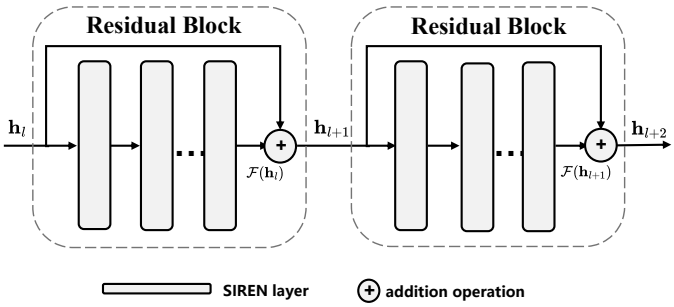


Figure 2. Illustration of the residual block used in MSP-ResPINNs. Each block updates the hidden feature $\mathbf{h}_l$ by adding a residual mapping $\mathcal{F}(\mathbf{h}_l)$, implemented as a fully connected subnetwork with multiple SIREN layers.

2.3.2. Multi-Source Prior Constraints

Seismic traveltime tomography is a highly ill-posed inverse problem, in which different velocity models may produce similar traveltime responses, particularly when only first-arrival traveltime data are used. To reduce non-uniqueness and guide the inversion toward physically meaningful solutions, we incorporate multi-source prior constraints into the proposed MSP-ResPINNs framework, together with the physics-based eikonal residual.

(a) Well-log velocity constraint

Well-log measurements provide accurate pointwise velocity information along boreholes and serve as strong local constraints on the velocity model. Let $\{\mathbf{x}_w^k, v_w^k\}_{k=1}^{N_w}$ denote the locations and measured velocities from well logs, where N_w is the total number of well-log measurements. The corresponding loss term is defined as:

$$\mathcal{L}_{\text{well}} = \frac{1}{N_w} \sum_{k=1}^{N_w} \left| \hat{v}_\phi(\mathbf{x}_w^k) - v_w^k \right|^2, \tag{16}$$

(b) Interface velocity constraint

In practical scenarios, geological interfaces such as layer boundaries or horizons are known from seismic interpretation or other geophysical surveys. These interfaces often correspond to velocity discontinuities or characteristic velocity values. Let Γ denote a known interface, and $v_\Gamma(\mathbf{x}_\Gamma)$ the prescribed velocity along this interface. The interface constraint is enforced through:

$$\mathcal{L}_{\text{int}} = \frac{1}{N_\Gamma}\sum_{i=1}^{N_\Gamma}|\hat{v}_\phi(\mathbf{x}_\Gamma^i) - v_\Gamma(\mathbf{x}_\Gamma^i)|^2, \quad \mathbf{x}_\Gamma^i \in \Gamma, \tag{17}$$

where N_Γ is the total number of sampled points along the interface.

(c) Overall loss function

By combining the physics-based constraints with multiple sources of prior information, the overall training objective of the MSP-ResPINNs framework is formulated as:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{Eik}}\mathcal{L}_{\text{Eik}} + \lambda_{\text{src}}\mathcal{L}_{\text{src}} + \lambda_{\text{data}}\mathcal{L}_{\text{data}} + \lambda_{\text{well}}\mathcal{L}_{\text{well}} + \lambda_{\text{int}}\mathcal{L}_{\text{int}}, \tag{18}$$

where the weighting coefficients λ control the relative influence of each constraint.

2.4. Evaluation Metrics

To quantitatively evaluate the reconstruction accuracy of the predicted velocity models, several commonly used error- and structure-based metrics are employed, including mean absolute error (MAE), relative mean absolute error (RMAE), R^2, peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) [42], and a cross-relation indicator [43]. These metrics jointly evaluate the point-wise error, structural similarity, and spatial consistency between the predicted velocity model $\hat{\mathbf{v}}$ and the reference model $\mathbf{v}$.

2.4.1. MAE and RMAE

The MAE and RMAE are adopted to evaluate the average point-wise discrepancy between the predicted and reference velocity models. They are defined as:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|v_i - \hat{v}_i|, \tag{19}$$

$$\text{RMAE} = \frac{\sum_{i=1}^{N}|v_i - \hat{v}_i|}{\sum_{i=1}^{N}|v_i|}, \tag{20}$$

where v_i and $\hat{v}_i$ represent the velocity values of the reference and predicted models at the i-th grid point in Ω, respectively. A lower value for both metrics indicates higher reconstruction accuracy.

2.4.2. PSRN

The PSNR is adopted to measure the overall reconstruction quality of the predicted velocity model. It is defined as:

$$\text{PSNR} = 10\log_{10}\left(\frac{v_{\max}^2}{\text{MSE}}\right), \tag{21}$$

where $v_{\max}$ denotes the maximum velocity value of the reference model, and MSE represents the mean square error between the predicted model $\hat{\mathbf{v}}$ and reference velocity models $\mathbf{v}$. A higher PSNR indicates a smaller overall discrepancy between the two models.

2.4.3. SSIM

The SSIM is used to evaluate the structural consistency between the predicted model $\hat{\mathbf{v}}$ and reference velocity models $\mathbf{v}$. SSIM considers local variations in luminance, contrast, and structural information, and is defined as:

$$\mathrm{SSIM}(\mathbf{v}, \hat{\mathbf{v}}) = \frac{(2\mu_v\mu_{\hat{v}} + C_1)(2\sigma_{v\hat{v}} + C_2)}{(\mu_v^2 + \mu_{\hat{v}}^2 + C_1)(\sigma_v^2 + \sigma_{\hat{v}}^2 + C_2)}, \tag{22}$$

where μ_v and $\mu_{\hat{v}}$ are the mean velocities, σ_v^2 and $\sigma_{\hat{v}}^2$ are the variances, and $\sigma_{v\hat{v}}$ is the covariance between the predicted and reference velocity models. Constants C_1 and C_2 are introduced to stabilize the calculation. SSIM values range from -1 to 1, with larger values indicating higher structural similarity and better preservation of layer continuity.

2.4.4. Cross-Relation Indicator

To further evaluate the spatial consistency between the predicted model $\hat{\mathbf{v}}$ and the reference velocity model $\mathbf{v}$, a cross-relation indicator I_{corr} is introduced and defined as

$$I_{\mathrm{corr}} = \frac{\langle \mathbf{v}^2, \hat{\mathbf{v}}^2 \rangle}{\|\mathbf{v}^2\|_2 \, \|\hat{\mathbf{v}}^2\|_2}, \tag{23}$$

where the inner product $\langle \mathbf{v}^2, \hat{\mathbf{v}}^2 \rangle$ is defined as $\sum_k v_k^2 \hat{v}_k^2$, and the 2-norm $\|\mathbf{v}\|_2 = \sqrt{\sum_k v_k^2}$ denotes the Euclidean norm of the vectorized velocity model. The values of I_{corr} lie in the range $[0, 1]$, with larger values indicating a stronger correlation between the predicted and reference velocity distributions.

3. Numerical Results

In this section, we conduct velocity inversion experiments on the Curvel-A, Salt, and Marmousi models using PINNtomo [15], MSP-ResPINNs with additive factorization, referred to as MSP-ResPINNs (additive), and MSP-ResPINNs with multiplicative factorization, referred to as MSP-ResPINNs (multiplicative).

3.1. Curvel-A Model

We first evaluated MSP-ResPINNs using the Curvel-A model from OpenFWI [44], as shown in Figure 3a. The model contains three well-defined layer interfaces, with velocity increasing gradually from top to bottom. The model grid consists of 71 × 71 points with a spacing of 0.01 km. Fifteen sources are positioned at x = 0 km with an interval of 0.05 km, and 71 receivers are located at x = 0.71 km. Well velocity profiles are available at x = 0.1 km, x = 0.35 km, and x = 0.6 km, with nine velocity points along each profile.

In MSP-ResPINNs, the network architecture is kept consistent for both multiplicative and additive factorizations. For the network $\mathcal{N}_{\hat{\tau}_\theta}$, each layer contains 64 neurons, with six residual blocks, each comprising two SIREN layers. For the network $\mathcal{N}_{\hat{v}_\phi}$, each layer contains 64 neurons, with six residual blocks, each also containing two SIREN layers. For PINNtomo, the network $\mathcal{N}_{\hat{\tau}_\theta}$ consists of 12 hidden layers with 64 neurons per layer, while the network $\mathcal{N}_{\hat{v}_\phi}$ also has 12 hidden layers with 64 neurons per layer. In this experiment, 3000 sets of data $\{x, z, x_s, z_s\}$ were generated. The input coordinates $\{x, z, x_s, z_s\}$, uniformly and randomly sampled within the model domain, were used as inputs to $\mathcal{N}_{\hat{\tau}_\theta}$, while the corresponding coordinates $\{x, z\}$ were used as inputs to $\mathcal{N}_{\hat{v}_\phi}$. The learning rate was set to 0.002, and all networks were trained for 1500 iterations.

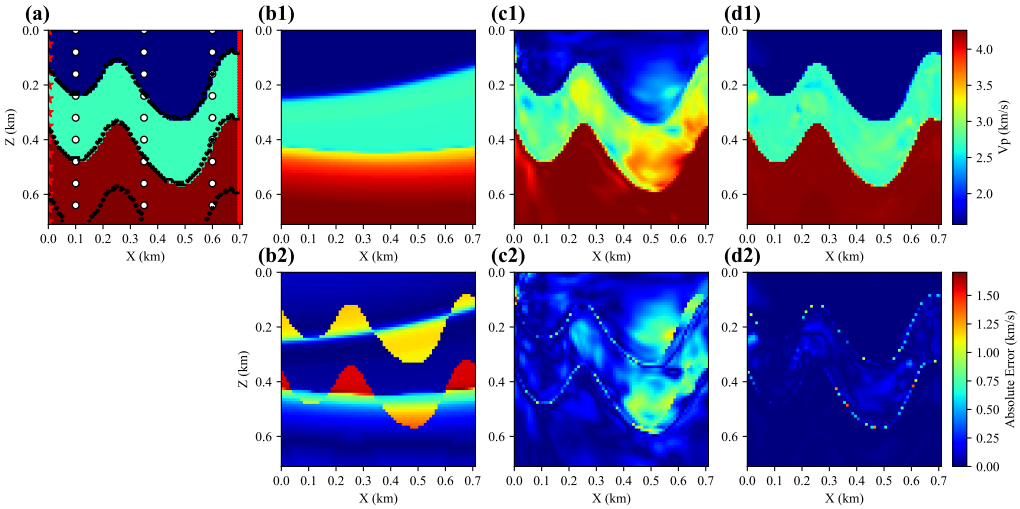


Figure 3. Curvel-A velocity model and inversion results. (**a**) True Curvel-A velocity model. Red pentagrams indicate sources, red downward triangles indicate receivers, white circles represent well velocity points, and black curves denote interface locations. (**b1**) Inverted velocity model by PINNtomo. (**c1**) Inverted velocity model by MSP-ResPINNs (Additive). (**d1**) Inverted velocity model by MSP-ResPINNs (Multiplicative). (**b2**) Absolute error predicted by PINNtomo. (**c2**) Absolute error predicted by MSP-ResPINNs (Additive). (**d2**) Absolute error predicted by MSP-ResPINNs (Multiplicative).

Figure 3(b1–d1) compares the velocity inversion results of PINNtomo and MSP-ResPINNs on the Curvel-A model. The inversion result of PINNtomo (Figure 3(b1)) exhibits strong smoothing, failing to accurately capture the pronounced layer interfaces in the true model. Due to the lack of explicit interface constraints, the velocity gradients near the layer interfaces are significantly weakened, which is reflected in the corresponding absolute error map (Figure 3(b2)) as large, concentrated errors around the interfaces.

With the introduction of interface constraints, the inversion performance of MSP-ResPINNs is significantly improved, though notable differences exist between different factorization forms. MSP-ResPINNs (Additive) (Figure 3(c1)) effectively recovers the overall undulating trend of the interfaces, showing clear improvement in interface continuity compared with PINNtomo. However, some oscillations still appear in localized high-velocity regions, and the corresponding error map (Figure 3(c2)) shows concentrated error bands near the interfaces. In contrast, MSP-ResPINNs (Multiplicative) (Figure 3(d1)) achieves the best performance in interface localization and intra-layer velocity recovery. This method accurately reconstructs the interface geometry of the three-layer medium, with clear velocity contrasts between layers and stable deep structures. The corresponding error map (Figure 3(d2)) exhibits the lowest overall error, with only minor local errors at a few interface locations, demonstrating higher inversion accuracy and numerical stability.

Figure 4 presents a comparison of vertical velocity slices obtained by PINNtomo and MSP-ResPINNs at x = 0.25 km and x = 0.5 km. All methods reproduce the overall trend of the true velocity model. However, MSP-ResPINNs recover velocity amplitudes and interface locations more accurately than PINNtomo, with the multiplicative factorization form performing best and clearly outperforming the additive form.

Furthermore, MSP-ResPINNs demonstrate clear advantages in inversion accuracy and structural consistency, as summarized in Table 2. For the Curvel-A model, the multiplicative factorization achieves substantial relative improvements compared with PINNtomo. Specifically, the MAE decreases from 0.3070 to 0.0302, corresponding to a relative improvement of 90.2%, and the RMAE decreases from 0.1043 to 0.0103, yielding a relative improvement of 90.1%. For structural and correlation metrics, R^2 improves from 0.7551 to 0.9936 (31.6% relative improvement), PSNR increases from 17.89 to 33.73 (88.5% relative improvement), SSIM

increases from 0.8725 to 0.9968 (14.3% relative improvement), and Icorr improves from 0.9610 to 0.9990 (4.0% relative improvement). These results consistently demonstrate that MSP-ResPINNs with multiplicative factorization significantly enhances velocity amplitude recovery and structural fidelity compared with both PINNtomo and the additive variant.

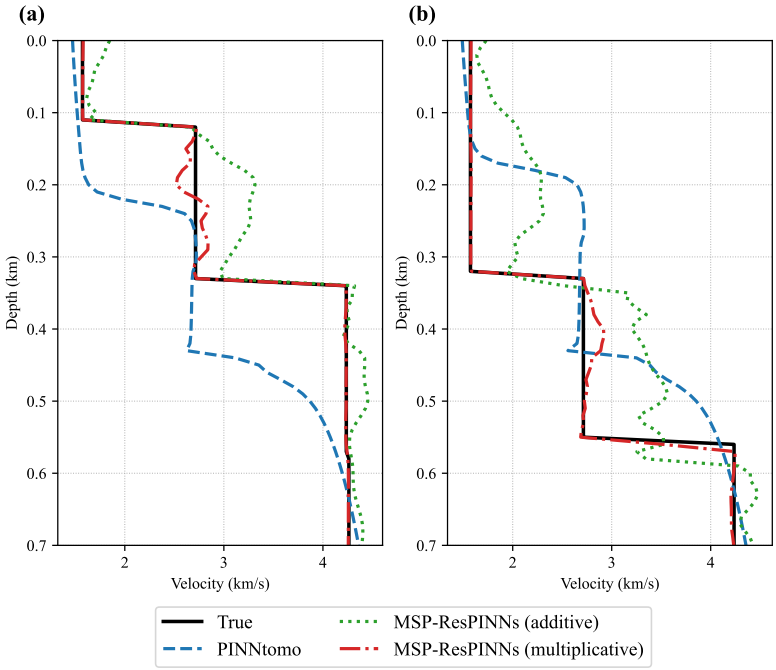


Figure 4. The vertical velocity profiles extracted at x = 0.25 km (**a**) and x = 0.5 km (**b**).

Table 2. Comparison results of the evaluation metrics on CurveA, Salt and Marmousi models.

Model	Method	MAE ↓	RMAE ↓	R^2 ↑	PSNR ↑	SSIM ↑	Icorr ↑
CurveA	PINNtomo	0.3070	0.1043	0.7551	17.89	0.8725	0.9610
	MSP-ResPINNs (Additive)	0.2401	0.0816	0.9045	21.98	0.9585	0.9880
	MSP-ResPINNs (Multiplicative)	**0.0302**	**0.0103**	**0.9936**	**33.73**	**0.9968**	**0.9990**
Salt	PINNtomo	0.1319	0.0414	0.9003	26.53	0.9457	0.9401
	MSP-ResPINNs (Additive)	0.0471	0.0148	0.9720	33.51	0.9800	0.9861
	MSP-ResPINNs (Multiplicative)	**0.0408**	**0.0128**	**0.9892**	**36.20**	**0.9946**	**0.9979**
Marmousi	PINNtomo	0.3198	0.1181	0.7668	21.07	0.8822	0.9097
	MSP-ResPINNs (Additive)	0.2071	0.0765	0.8921	24.42	0.9465	0.9642
	MSP-ResPINNs (Multiplicative)	**0.1803**	**0.0666**	**0.9093**	**25.17**	**0.9523**	**0.9827**

In the training of physics-informed neural networks, differences in the loss terms used by different algorithms make it difficult to objectively evaluate convergence performance solely from the decrease in the total loss function. Therefore, we adopt a unified velocity model MAE to assess the evolution of different methods during the training process, as shown in Figure 5. All methods exhibit a significant error reduction at the early stage of iteration, particularly before approximately 600 iterations. The PINNtomo method demonstrates the fastest convergence, with the MAE eventually stabilizing at around 0.3 km/s.

In contrast, although the proposed MSP-ResPINNs method introduces additional learning complexity due to the incorporation of residual modules, SIREN activation functions, and interface-based physical constraints, resulting in a slightly more complicated convergence process, its final convergence accuracy is substantially superior to that of the comparison methods, with the MAE reduced by approximately one order of magnitude. Furthermore, a comparison of different factorization strategies reveals that MSP-ResPINNs (additive) exhibits frequent and pronounced numerical oscillations in the later stages of

training. This instability is clearly reflected in the error distribution shown in Figure 3(c1,c2). By comparison, MSP-ResPINNs (multiplicative) presents the smoothest and most stable MAE evolution curve.

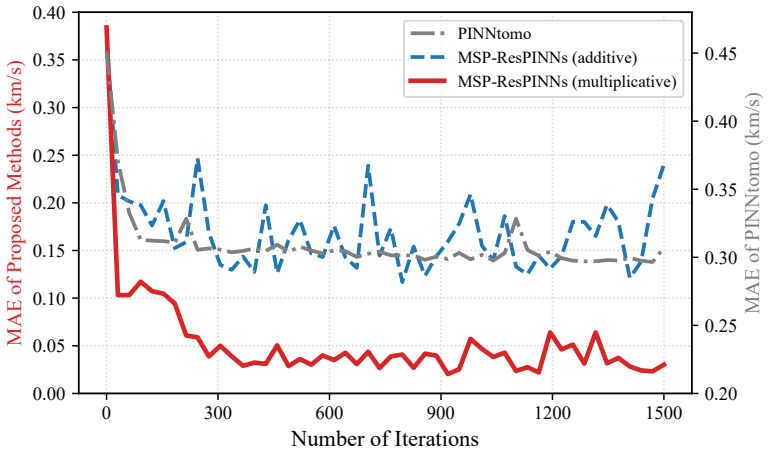


Figure 5. Curvel-A velocity MAE evolution during training for different inversion frameworks.

3.2. Salt Model

We evaluated MSP-ResPINNs using 2D simulated models extracted from a 3D salt model [45], as shown in Figure 6a. The velocity model consists of a 302 × 202 grid with a grid spacing of 0.01 km. Twenty-one sources are positioned at x = 0 with an interval of 0.15 km. One hundred receivers are evenly distributed along z = 0 km and x = 3.02 km, respectively. Well-log velocity profiles are located at x = 0.1 km, x = 0.8 km, x = 1.5 km, x = 2.3 km, and x = 2.9 km, with each profile containing 55 well-log velocity points.

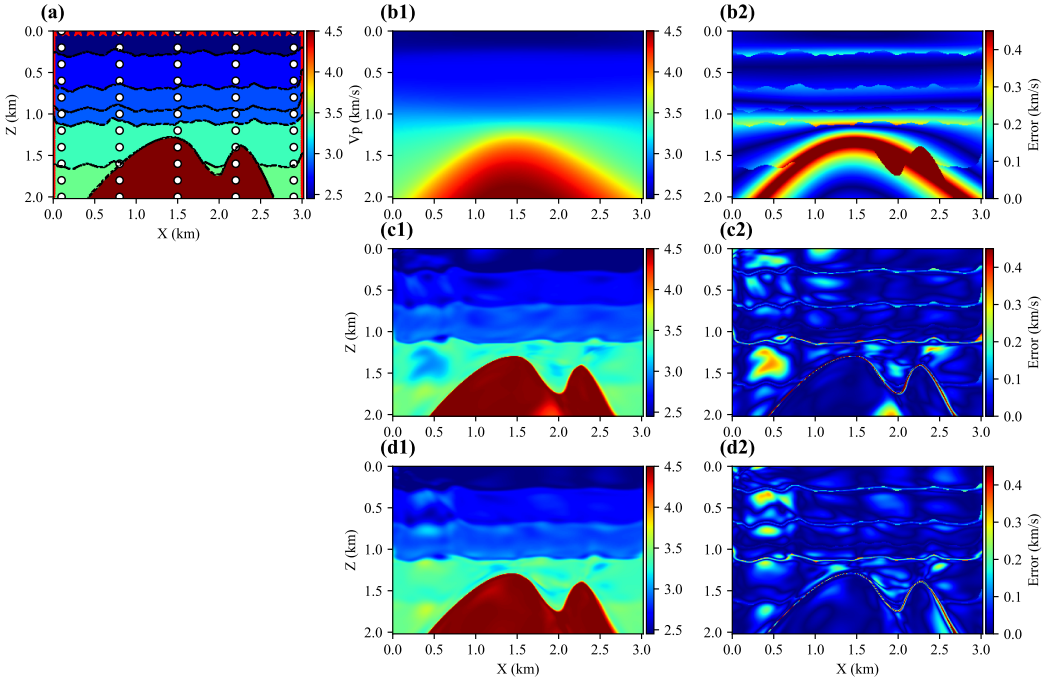


Figure 6. Salt velocity model and inversion results. (**a**) True Salt velocity model. Red pentagrams indicate sources, red downward triangles indicate receivers, white circles indicate well-log velocity points, and black lines represent interface information. (**b1**) Velocity model inverted by PINNtomo. (**c1**) Velocity model inverted by MSP-ResPINNs (additive). (**d1**) Velocity model inverted by MSP-ResPINNs (multiplicative). (**b2**) Absolute error of PINNtomo predictions. (**c2**) Absolute error of MSP-ResPINNs (additive). (**d2**) Absolute error of MSP-ResPINNs (multiplicative).

In MSP-ResPINNs, the network architecture remains the same regardless of whether multiplicative or additive factorization is employed. For the network $\mathcal{N}_{\hat{t}_\theta}$, each layer contains 64 neurons, and the network includes 6 residual blocks, with each residual block consisting of two SIREN layers. For the network $\mathcal{N}_{\hat{v}_\phi}$, each layer contains 128 neurons, and each residual block also comprises two SIREN layers, with the total number of residual blocks equal to 6. In PINNtomo, the network $\mathcal{N}_{\hat{t}_\theta}$ has 64 neurons per layer and 12 hidden layers, while $\mathcal{N}_{\hat{v}_\phi}$ has 128 neurons per layer and 12 hidden layers. In this study, we generated 20,000 data samples $\{x, z, x_s, z_s\}$. The input coordinates $\{x, z, x_s, z_s\}$ were randomly sampled from a uniform distribution over the model domain to serve as the inputs for $\mathcal{N}_{\hat{t}_\theta}$, while the corresponding $\{x, z\}$ were used as inputs for $\mathcal{N}_{\hat{v}_\phi}$. The learning rate was set to 0.002, and the networks were trained for 3000 iterations.

Figure 6(b1–d1) compare the velocity inversion results obtained by PINNtomo and MSP-ResPINNs on the Salt model. The PINNtomo results shown in Figure 6(b1) exhibit pronounced smoothing. Due to the absence of interface constraints, PINNtomo struggles to accurately capture the layered velocity structures in shallow regions and the complex geometry of the salt body. This limitation is reflected in the corresponding absolute error map (Figure 6(b2)), where high errors are concentrated along the salt boundaries and in shallow layers. In contrast, the inclusion of interface information in MSP-ResPINNs significantly improves inversion quality. Both additive and multiplicative factorization strategies effectively reduce errors at interfaces and produce clearer salt body outlines. In particular, MSP-ResPINNs (multiplicative) (Figure 6(d1)) achieves high-fidelity reconstruction of the salt boundaries and demonstrates superior resolution of shallow structures. Its corresponding error map (Figure 6(d2)) indicates the lowest overall error, with only minor residual errors in some deep regions.

Figure 7 compares the vertical velocity profiles inverted by PINNtomo and MSP-ResPINNs at horizontal positions x = 1.0 km and x = 2.0 km. Due to the absence of interface constraints, PINNtomo exhibits pronounced over-smoothing, deviating significantly from the step-like layered structures of the true velocity profiles. As shown in Figure 7a, in the depth range between 1.25 km and 1.5 km, PINNtomo systematically overestimates velocities and fails to accurately capture critical velocity discontinuities. For instance, at 1.5 km depth, the true velocity abruptly increases from approximately 3.2 km/s to 4.5 km/s, whereas PINNtomo predicts only around 3.8 km/s, underrepresenting the sharp transition. In contrast, MSP-ResPINNs explicitly incorporate interface constraints, effectively mitigating the overestimation in deep regions below 1.25 km and enabling high-fidelity reconstruction of velocity discontinuities. At x = 2.0 km (Figure 7b), the predicted velocity at 1.75 km depth closely matches the true value, demonstrating the method's capability to accurately resolve interface structures and non-continuous features.

Table 2 presents the quantitative evaluation results of PINNtomo and MSP-ResPINNs for velocity inversion in the Salt model. Overall, MSP-ResPINNs with multiplicative factorization outperforms the addictive factorization across all metrics. The MAE of MSP-ResPINNs (multiplicative) is 0.0408, representing a 69.1% reduction compared to PINNtomo, which has an MAE of 0.1319. The RMAE is 0.0128, decreased by 69.0% relative to PINNtomo, indicating a substantial reduction in the absolute difference between predicted and true velocities. Meanwhile, the coefficient of determination R^2 increases from 0.9003 for PINNtomo to 0.9892 for MSP-ResPINNs (multiplicative), an improvement of approximately 9.9%, reflecting a significant enhancement in overall fitting accuracy.

In terms of structural and perceptual quality, the PSNR of MSP-ResPINNs (multiplicative) reaches 36.20 dB, corresponding to a relative improvement of 36.5% over PINNtomo, while the SSIM increases to 0.9946, yielding a 5.1% relative improvement. These quantitative gains indicate enhanced structural continuity and boundary preservation in the

reconstructed velocity field. This observation is consistent with the clearer delineation of salt boundaries in Figure 6(d1), as well as the more accurate reconstruction of velocity discontinuities at depths of 1.5 km in Figure 7a and 1.75 km in Figure 7b. Furthermore, the correlation coefficient Icorr reaches 0.9979, demonstrating strong spatial consistency between the predicted and reference velocity models.

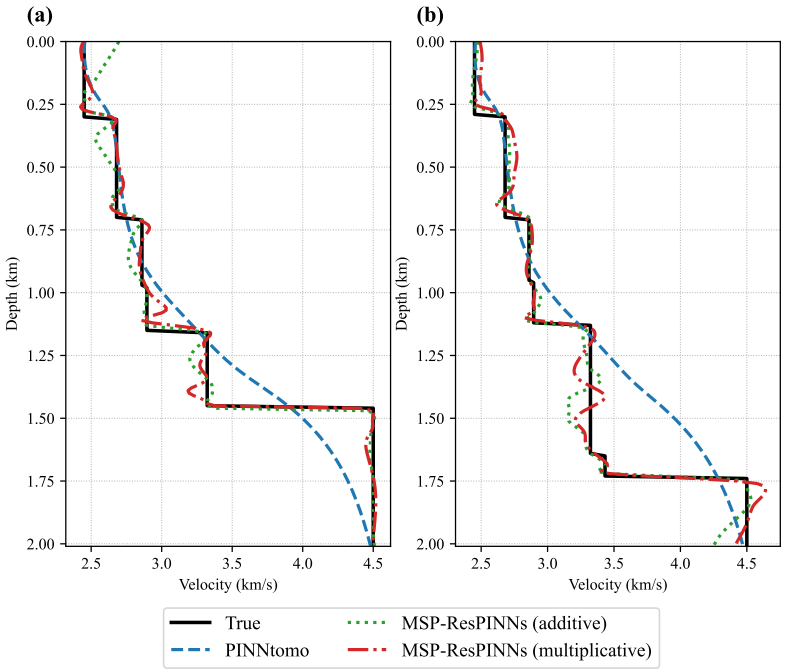


Figure 7. Vertical velocity profiles extracted at horizontal positions x = 1.0 km (**a**) and x = 2.0 km (**b**).

We further conduct ablation experiments to verify the necessity of the residual modules and the SIREN activation function in the proposed framework. Specifically, the MSP-ResPINNs model without residual modules is denoted as MSP-PINNs, while the model in which the SIREN activation function in MSP-ResPINNs is replaced with the tanh activation function is referred to as MSP-ResPINNs (without SIREN). In these ablation settings, the multiplicative factorization strategy is adopted, while the remaining network architecture and training-related parameters are kept identical.

The ablation results in Figure 8 provide insights into the roles of the residual modules and the SIREN activation function. In the absence of residual modules, MSP-PINNs shows increased difficulty in optimizing the deep network. When SIREN is replaced by the tanh activation function, the model tends to show reduced capability in representing the high-frequency interfaces of the salt body. Figure 9 presents the MAE evolution curves of the Salt model under different inversion frameworks. The results indicate that, compared with the conventional PINNtomo, the proposed MSP-ResPINNs (multiplicative) exhibits significant advantages in both convergence accuracy and stability. The ablation experiments further confirm that the SIREN activation function, through its unique periodic mapping capability, mitigates the spectral bias problem commonly encountered by traditional tanh activation functions when representing high-frequency velocity discontinuities such as salt boundaries. This effectively suppresses the frequent and severe numerical oscillations observed in the without SIREN curve. Meanwhile, the ResNet module ensures stable and continuous optimization during the training of deep neural networks.

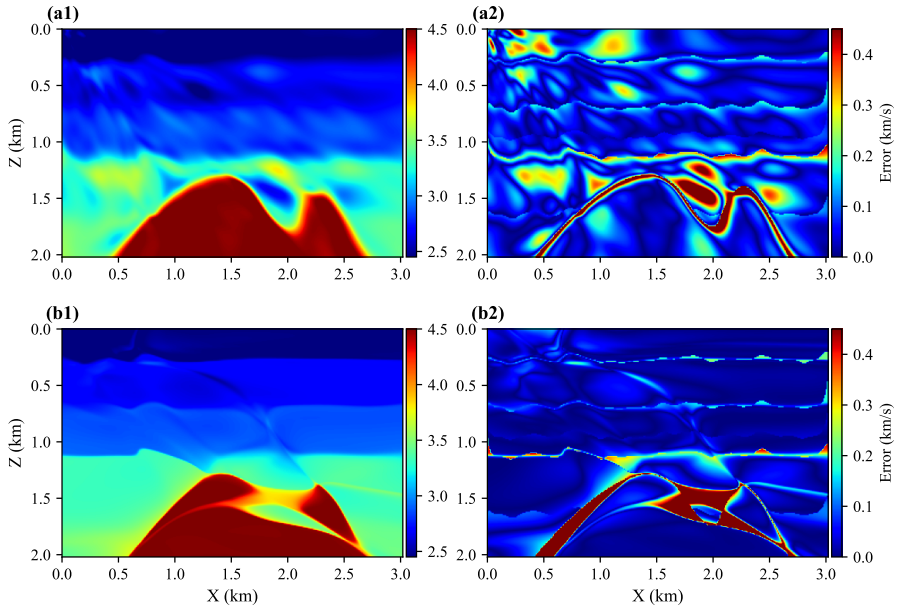


Figure 8. Ablation study results for the Salt velocity model inversion. (**a1**) Inversion result of MSP-PINNs. (**a2**) Absolute error of MSP-PINNs. (**b1**) Inversion result of MSP-ResPINNs (without SIREN). (**b2**) Absolute error of MSP-ResPINNs (without SIREN).

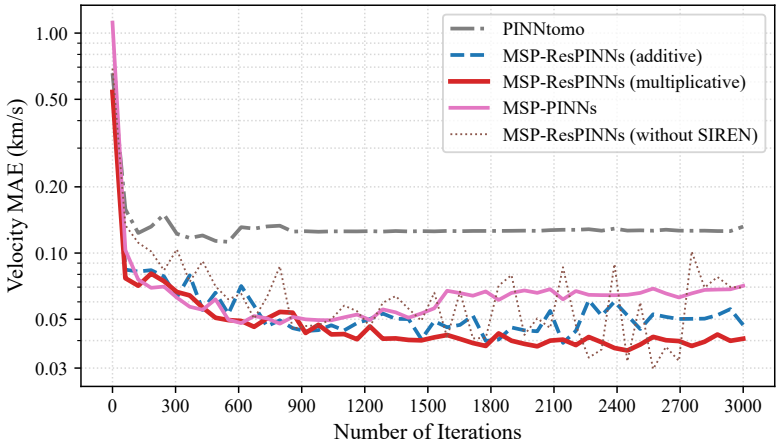


Figure 9. Salt velocity MAE evolution during training for different inversion frameworks.

3.3. Marmousi Model

We applied MSP-ResPINNs to invert the Marmousi model, as shown in Figure 10a. The model grid consists of 426 × 131 points with a spacing of 0.01 km. We positioned twenty-one sources along z = 0 at intervals of 0.21 km and deployed one hundred receivers at x = 0 km and x = 4.26 km. Well-log velocity profiles were placed at x = 0.2 km, 0.84 km, 1.48 km, 2.12 km, 2.76 km, 3.40 km, and 4.05 km, with each profile containing 55 well-log velocity points.

In MSP-ResPINNs, the network architecture remains the same regardless of whether multiplicative or additive factorization is employed. For the network $\mathcal{N}_{\hat{\tau}_\theta}$, each layer contains 128 neurons, and the network includes six residual blocks, with each residual block consisting of two SIREN layers. Similarly, for the network $\mathcal{N}_{\hat{v}_\phi}$, each layer contains 128 neurons, and each residual block also comprises two SIREN layers, with a total of six residual blocks. In contrast, in PINNtomo, the network $\mathcal{N}_{\hat{\tau}_\theta}$ has 128 neurons per layer and twelve hidden layers, while $\mathcal{N}_{\hat{v}_\phi}$ has 128 neurons per layer and twelve hidden layers. In this study, we generated 40,000 data samples $\{x, z, x_s, z_s\}$. The input coordinates $\{x, z, x_s, z_s\}$ were randomly sampled from a uniform distribution over the model domain and used as inputs

for $\mathcal{N}_{\hat{t}_\theta}$, while the corresponding coordinates $\{x, z\}$ were used as inputs for $\mathcal{N}_{\hat{v}_\phi}$. We set the learning rate to 0.002 and trained the networks for 5000 iterations.

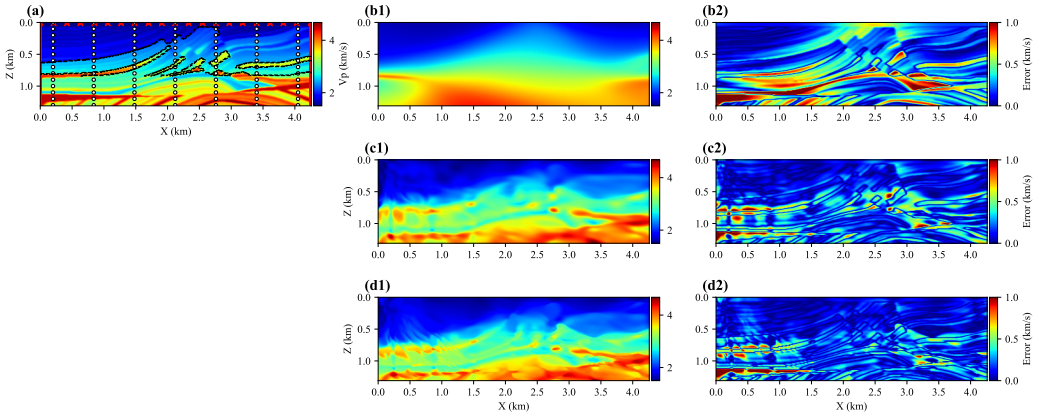


Figure 10. Marmousi velocity model and inversion results. (**a**) True Marmousi velocity model. Red pentagrams indicate sources, red downward triangles indicate receivers, white circles represent well-log velocity points, and black lines denote interface information. (**b1**) Velocity model inverted by PINNtomo. (**c1**) Velocity model inverted by MSP-ResPINNs (additive). (**d1**) Velocity model inverted by MSP-ResPINNs (multiplicative). (**b2**) Absolute error of PINNtomo predictions. (**c2**) Absolute error of MSP-ResPINNs (additive). (**d2**) Absolute error of MSP-ResPINNs (multiplicative).

Figure 10(b1) shows that in the shallow region above 0.5 km, PINNtomo produces over-smoothed results and fails to recover the low-velocity structures present in the true model. In addition, the lateral variations in the low-velocity zone at x = 1–2 km and the high-velocity zone at x = 3 km are weaker in the inversion model than in the true model. The corresponding absolute error map in Figure 10(b2) reveals widespread errors greater than 0.25 km/s in shallow regions and areas with strong lateral variations, indicating that PINNtomo struggles to fit complex structures accurately. Figure 10(c1) illustrates that MSP-ResPINNs (additive) better captures shallow details, such as the low-velocity zone at x = 1–2 km, as well as deep lateral variations, including the high-velocity zone at x = 3 km, showing closer agreement with the true model. Figure 10(d1) presents the results of MSP-ResPINNs (multiplicative), which slightly outperforms the additive model. Both shallow and deep structures are effectively recovered, and the incorporation of well-log and interface constraints enhances local accuracy.

Figure 11 presents comparisons of vertical velocity profiles at different horizontal positions. At x = 0.5 km (Figure 11a), all methods closely match the true model in the shallow region above 0.6 km, with minimal differences. However, at depths greater than 0.8 km, PINNtomo exhibits significant deviations from the true values, indicating limited fitting capability in deep, complex structures. In contrast, MSP-ResPINNs with both additive and multiplicative factorization better reproduce the oscillatory patterns of the true model, particularly capturing the cyclic variations between 1.0 and 1.2 km depth. At x = 1.8 km (Figure 11b), PINNtomo fails to accurately capture the downward trend in velocity, whereas MSP-ResPINNs closely follow this variation and show strong agreement with the true profile. At x = 2.7 km (Figure 11c), PINNtomo shows significant deviations within the oscillatory interval between 0.8 and 1.2 km, while MSP-ResPINNs effectively reproduces the high-frequency oscillations of the true model. At x = 4.0 km (Figure 11d), PINNtomo deviates substantially from the true model in both velocity values and overall trend. By contrast, MSP-ResPINNs closely tracks the true profile, particularly in the 1.0–1.2 km oscillatory section, demonstrating a clear advantage in deep-layer reconstruction.

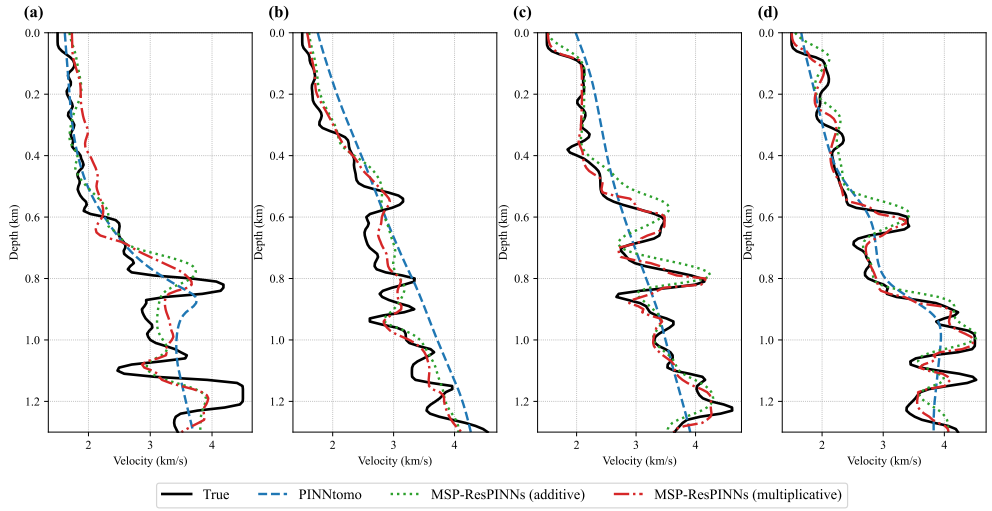


Figure 11. Vertical velocity profiles extracted at horizontal positions x = 0.5 km (**a**), x = 1.8 km (**b**), x = 2.7 km (**c**), and x = 4.0 km (**d**).

Figure 12 presents the MAE evolution curves of the Marmousi model under different inversion frameworks. The results indicate that, compared with the conventional PINNtomo, the proposed MSP-ResPINNs family of methods achieves improved convergence accuracy. Among them, MSP-ResPINNs (multiplicative) (red solid line) exhibits the best overall performance, reaching the lowest MAE after convergence (approximately 0.17 km/s). In addition, its evolution curve appears smoother, suggesting improved stability and effectively mitigating the numerical oscillations observed in the additive strategy during the middle and late stages of training.

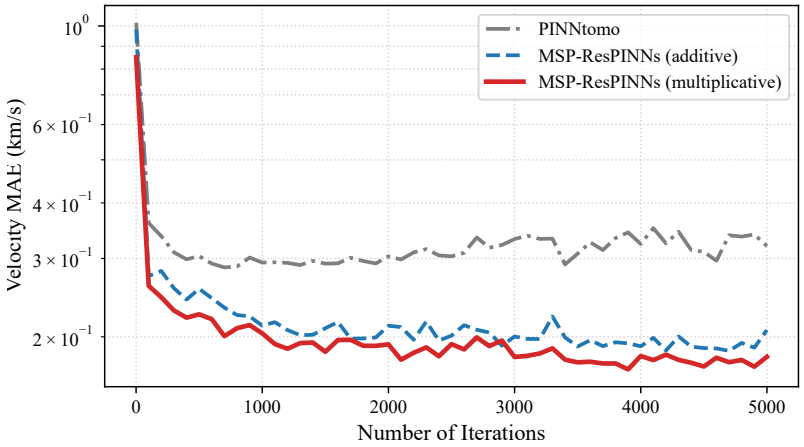


Figure 12. Marmousi velocity MAE evolution during training for different inversion frameworks.

Figure 13(b1) shows that the PINNtomo results are close to the true values, although the edges are slightly blurred, with minor deviations observed on the right side. Figure 13(b2) demonstrates that MSP-ResPINNs (multiplicative) exhibits very strong agreement with the true traveltimes, as the color distribution almost perfectly matches the reference with only negligible differences. Figure 13(b3) indicates that MSP-ResPINNs (additive) produces a clear shift in the predicted traveltimes, especially in the right region where the values are significantly overestimated, suggesting a substantial deviation from the true solution.

Figure 13(c1) illustrates that the τ field predicted by PINNtomo is non-uniform but overall smooth, exhibiting certain structural characteristics. Figure 13(c2) shows that MSP-ResPINNs (multiplicative) yields a τ field with well-defined spatial variations and smooth color transitions, reflecting physically meaningful smoothing corrections. Figure 13(c3)

reveals that the τ field predicted by MSP-ResPINNs (additive) is dominated by low values combined with random speckle patterns, indicating a lack of coherent spatial structure and an inability to capture the physical significance of τ.

Figure 13(d1,e1) show that the gradient fields of τ predicted by PINNtomo are smooth and exhibit clear directional trends. Figure 13(d2,e2) demonstrate that MSP-ResPINNs (multiplicative) produces gradient fields that are complex yet organized, characterized by distinct wave-like or banded structures. Figure 13(d3,e3) indicate that the gradient fields predicted by MSP-ResPINNs (additive) are dominated by random noise and lack meaningful spatial organization.

Although both MSP-ResPINNs with additive and multiplicative factorizations are able to reconstruct velocity structures similar to the true model at the early training stage, Figure 13(b3) shows that the additive approach produces a severe overestimation of traveltimes in the right region, indicating inferior physical consistency compared with the multiplicative strategy. Figure 13(c3) further reveals that the τ field predicted by the additive method lacks coherent spatial structure. Figure 13(d3,e3) show that its corresponding gradient fields are dominated by random noise, demonstrating a failure of the network to learn the physical meaning of τ. As training proceeds beyond approximately 5000 iterations, this inconsistency becomes increasingly pronounced. The velocity model obtained with additive factorization gradually deteriorates into high-frequency noise, whereas the velocity and τ fields produced by the multiplicative factorization remain smooth and well organized, and the predicted traveltimes preserve high accuracy.

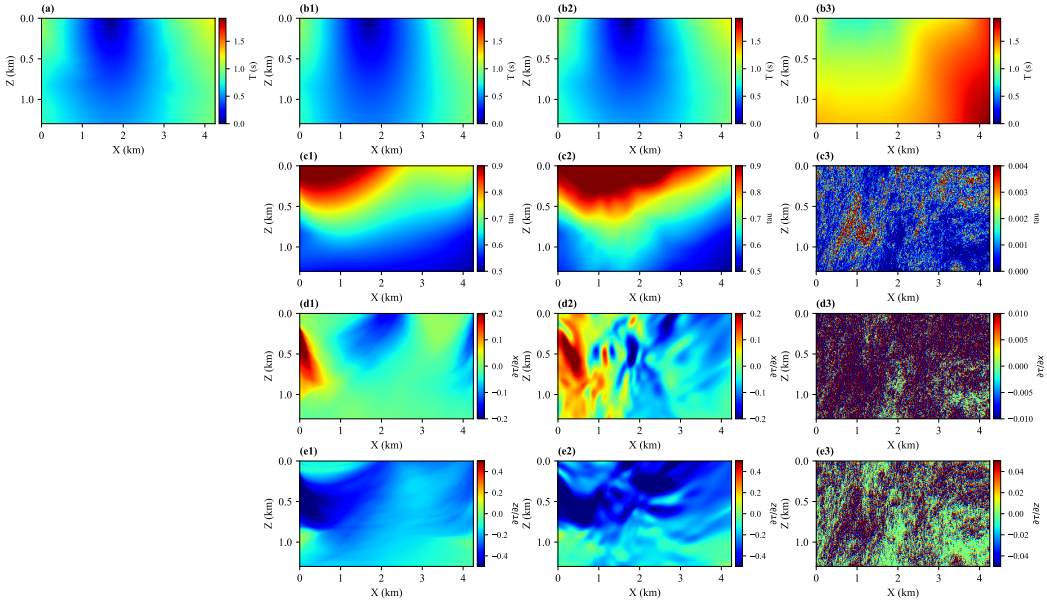


Figure 13. Traveltime T, traveltime smoothing correction τ, and the partial derivatives of τ along each direction. (**a**) True traveltime. (**b1**–**b3**) Predicted traveltimes by PINNtomo, MSP-ResPINNs (multiplicative), and MSP-ResPINNs (additive), respectively, sharing the same colorbar as (**a**). (**c1**–**c3**) Predicted traveltime smoothing corrections τ by PINNtomo, MSP-ResPINNs (multiplicative), and MSP-ResPINNs (additive), respectively; (**c1**,**c2**) use the same colorbar range. (**d1**–**d3**) Partial derivatives of τ along the x direction predicted by PINNtomo, MSP-ResPINNs (multiplicative), and MSP-ResPINNs (additive), respectively; (**d1**,**d2**) share the same colorbar range. (**e1**–**e3**) Partial derivatives of τ along the z direction predicted by PINNtomo, MSP-ResPINNs (multiplicative), and MSP-ResPINNs (additive), respectively; (**e1**,**e2**) use the same colorbar range.

The underlying reason for this behavior lies in formulating the additive factorization of the eikonal equation, which constrains only the magnitude of the gradient while leaving its direction and spatial smoothness unconstrained. As a result, the network can exploit high-frequency gradient noise to compensate for residuals during optimization, without enforcing physical consistency among τ, velocity, and traveltime. With increasing iterations, this excess gradient freedom is progressively amplified, leading to traveltime deviations and ultimately to velocity degradation. In contrast, the multiplicative factorization enforces a proportional coupling between and the background traveltime T_0, such that high-frequency perturbations are penalized through amplified residuals. Consequently, the τ field remains smooth and ordered, and both traveltime predictions and velocity models remain stable and physically consistent.

4. Conclusions

We present a Multi-Source Prior-Guided Residual Physics-Informed Neural Network (MSP-ResPINNs) framework for seismic traveltime tomography. MSP-ResPINNs extends conventional PINN-based tomography by incorporating a residual network to improve training stability and representing complex velocity structures, while enabling the unified integration of observed seismic traveltimes, well-log velocities, and geological interface constraints within a physically consistent eikonal-based inversion framework. Numerical experiments demonstrate that MSP-ResPINNs markedly enhances velocity reconstruction, effectively suppressing excessive smoothing, recovering sharp velocity contrasts, and resolving complex geological structures. A systematic comparison of additive and multiplicative factorized eikonal formulations shows that the multiplicative approach provides superior physical consistency, greater training stability, and improved robustness during long-term optimization, whereas the additive formulation may suffer from gradient noise amplification and velocity degradation. Future work will extend the framework to three-dimensional problems, incorporate additional geophysical constraints, and couple traveltime tomography with waveform-based inversion.

Author Contributions: Y.X. conceived and designed the study, developed the methodology, prepared the software, performed the formal analysis, and wrote the original draft. Z.S. was responsible for the software implementation, performed the investigation and validation, and visualized the results. Z.W. contributed to the methodology, provided data curation, managed the project, and was responsible for funding acquisition and supervision. G.S. contributed to the conceptualization, supervised the study, and participated in manuscript review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This study was jointly funded by the Special Project of the Central Government to Guide Local Science and Technology Development (2021ZYD0003), the Nanchong City University Science and Technology Strategic Cooperation Project (SXHZ045, 19SXHZ0030), the scientific research starting project of SWPU (No. 2021QHZ016), and the Southwest Petroleum University Graduate Research Innovation Fund (2020cxyb043).

Data Availability Statement: Data associated with this research are available and can be obtained by contacting the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Category	Symbol/Acronym	Description
	$T(\mathbf{x})$	Seismic traveltime field at spatial point $\mathbf{x}$
	$v(\mathbf{x})$	Seismic velocity model
	$\tau(\mathbf{x})$	Dimensionless traveltime factor or smooth correction term
Physics and Seismology	$T_0(\mathbf{x})$	Background traveltime
	$\mathbf{x}_s, \mathbf{x}_r$	Spatial coordinates of seismic sources and receivers
	$\Gamma, \mathbf{x}_\Gamma$	Geological interface and coordinates on the interface
	PINNs	Physics-Informed Neural Networks
	SIREN	Sinusoidal Representation Networks (Periodic activation)
Machine Learning	ResNet	Residual Neural Network
	θ, ϕ	Trainable parameters of traveltime and velocity networks
	$\mathcal{F}(\cdot)$	Residual mapping within the neural network blocks
	$\mathcal{L}_{\text{total}}$	Total composite loss function
Optimization	λ	Weighting coefficients for individual loss terms
	$N_f, N_s, N_d, N_w, N_\Gamma$	Number of collocation points, source points, data points, well-logs, and interface points

References

1. Rawlinson, N.; Pozgay, S.; Fishwick, S. Seismic tomography: A window into deep Earth. *Phys. Earth Planet. Inter.* **2010**, *178*, 101–135. [CrossRef]
2. Wu, X.; Guo, Z.; Li, S.; Yu, Y.; Bai, Q.; Chen, Y.J. Seismic Azimuthal Anisotropy of Northeastern Tibetan Plateau from Ambient Noise Double Beamforming Tomography: Implications for Crustal Deformation. *J. Geophys. Res. Solid Earth* **2023**, *128*, e2022JB026109. [CrossRef]
3. Levin, S.A. Principle of reverse-time migration. *Geophysics* **1984**, *49*, 581–583. [CrossRef]
4. Virieux, J.; Operto, S. An overview of full-waveform inversion in exploration geophysics. *Geophysics* **2009**, *74*, WCC1–WCC26. [CrossRef]
5. Berkhout, A.J. *Seismic Migration: Imaging of Acoustic Energy by Wave Field Extrapolation*; Elsevier: Amsterdam, The Netherlands, 2012.
6. Alkhalifah, T.; Choi, Y. From tomography to full-waveform inversion with a single objective function. *Geophysics* **2014**, *79*, R55–R61. [CrossRef]
7. Tavakoli, F.B.; Operto, S.; Ribodetti, A.; Virieux, J. Slope tomography based on eikonal solvers and the adjoint-state method. *Geophys. J. Int.* **2017**, *209*, 1629–1647. [CrossRef]
8. Rawlinson, N.; Hauser, J.; Sambridge, M. Seismic ray tracing and wavefront tracking in laterally heterogeneous media. *Adv. Geophys.* **2008**, *49*, 203–273.
9. Lecomte, I.; Lavadera, P.L.; Anell, I.; Buckley, S.J.; Schmid, D.W.; Heeremans, M. Ray-based seismic modeling of geologic models: Understanding and analyzing seismic images efficiently. *Interpretation* **2015**, *3*, SAC71–SAC89. [CrossRef]
10. Jordi, C.; Schmelzbach, C.; Greenhalgh, S. Frequency-dependent traveltime tomography using fat rays: Application to near-surface seismic imaging. *J. Appl. Geophys.* **2016**, *131*, 202–213. [CrossRef]
11. Leung, S.; Qian, J. An adjoint state method for three-dimensional transmission traveltime tomography using first-arrivals. *Commun. Math. Sci.* **2006**, *4*, 249–266. [CrossRef]
12. Tong, P. Adjoint-state traveltime tomography: Eikonal equation-based methods and application to the Anza area in southern California. *J. Geophys. Res. Solid Earth* **2021**, *126*, e2021JB021818. [CrossRef]

13. Taillandier, C.; Noble, M.; Chauris, H.; Calandra, H. First-arrival traveltime tomography based on the adjoint-state method. *Geophysics* **2009**, *74*, WCB1–WCB10. [CrossRef]
14. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [CrossRef]
15. bin Waheed, U.; Alkhalifah, T.; Haghighat, E.; Song, C.; Virieux, J. PINNtomo: Seismic tomography using physics-informed neural networks. *arXiv* **2021**, arXiv:2104.01588. [CrossRef]
16. Taufik, M.H.; bin Waheed, U.; Alkhalifah, T.A. Upwind, no more: Flexible traveltime solutions using physics-informed neural networks. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5922512. [CrossRef]
17. Wang, C.; Chen, H.; Zhan, K.; Kong, C.; Li, G. Fourier-ResPINN: A new solution for solving the travel time of first arrival. *Acta Geophys.* **2025**, *73*, 2509–2522. [CrossRef]
18. Liu, Z.; Archibong, G.; Waheed, U.B.; Wang, S.; Song, C. Physics-informed Fourier-DeepONet for a generalized eikonal solution. *Comput. Geosci.* **2026**, *206*, 106026. [CrossRef]
19. Song, C.; Alkhalifah, T.; Waheed, U.B. Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks. *Geophys. J. Int.* **2021**, *225*, 846–859. [CrossRef]
20. Chai, X.; Gu, Z.; Long, H.; Liu, S.; Yang, T.; Wang, L.; Zhan, F.; Sun, X.; Cao, W. Modeling multisource multifrequency acoustic wavefields by a multiscale Fourier feature physics-informed neural network with adaptive activation functions. *Geophysics* **2024**, *89*, T79–T94. [CrossRef]
21. Chang, Y.; Yang, D.; He, X. A robust numerical method based on a deep-learning operator for solving the 2D acoustic wave equation. *Geophysics* **2024**, *89*, T195–T205. [CrossRef]
22. Cheng, S.; Alkhalifah, T. Meta learning for improved neural network wavefield solutions. *Surv. Geophys.* **2025**, *46*, 145–167. [CrossRef]
23. Chen, Y.; de Ridder, S.A.; Rost, S.; Guo, Z.; Wu, X.; Chen, Y. Eikonal tomography with physics-informed neural networks: Rayleigh wave phase velocity in the Northeastern Margin of the Tibetan Plateau. *Geophys. Res. Lett.* **2022**, *49*, e2022GL099053. [CrossRef]
24. Chen, Y.; de Ridder, S.A.; Rost, S.; Guo, Z.; Wu, X.; Li, S.; Chen, Y. Physics-Informed Neural Networks for Elliptical-Anisotropy Eikonal Tomography: Application to Data From the Northeastern Tibetan Plateau. *J. Geophys. Res. Solid Earth* **2023**, *128*, e2023JB027378. [CrossRef]
25. Song, C.; Geng, H.; Wang, Y.; Waheed, U.B.; Liu, C. Simultaneous P- and S-wave seismic traveltime tomography using physics-informed neural networks. *Geophys. Prospect.* **2025**, *73*, e70034. [CrossRef]
26. Rasht-Behesht, M.; Huber, C.; Shukla, K.; Karniadakis, G.E. Physics-informed neural networks (PINNs) for wave propagation and full waveform inversions. *J. Geophys. Res. Solid Earth* **2022**, *127*, e2021JB023120. [CrossRef]
27. Song, C.; Wang, Y. Simulating seismic multifrequency wavefields with the Fourier feature physics-informed neural network. *Geophys. J. Int.* **2023**, *232*, 1503–1514. [CrossRef]
28. Lu, C.; Wang, Y.; Zou, X.; Zong, J.; Su, Q. Elastic full-waveform inversion via physics-informed recurrent neural network. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 4510616. [CrossRef]
29. Li, Y.; Pensoneault, A.; Zhang, Y.; Zhu, X.; Gou, R.; Gao, J. Efficient Bayesian physics-informed neural networks for surface tomography via ensemble Kalman inversion. *Geophysics* **2025**, *90*, U1–U14. [CrossRef]
30. Smith, J.D.; Azizzadenesheli, K.; Ross, Z.E. EikoNet: Solving the Eikonal Equation With Deep Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10685–10696. [CrossRef]
31. Gou, R.; Zhang, Y.; Zhu, X.; Gao, J. Bayesian physics-informed neural networks for the subsurface tomography based on the eikonal equation. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4503012. [CrossRef]
32. Agata, R.; Shiraishi, K.; Fujie, G. Bayesian seismic tomography based on velocity-space Stein variational gradient descent for physics-informed neural network. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4506917. [CrossRef]
33. Taufik, M.H.; Alkhalifah, T.; Waheed, U.b. Stable neural network-based traveltime tomography using hard-constrained measurements. *Geophysics* **2024**, *89*, U87–U99. [CrossRef]
34. Zhao, T.; Liu, C.; Song, C.; Waheed, U.B.; Zhang, X. Smoothness: The key factor in well-log information-assisted PINNtomo. *J. Appl. Geophys.* **2024**, *226*, 105417. [CrossRef]
35. Chen, Y.; Chen, H.; Xiang, K.; Chen, X. Geological structure guided well log interpolation for high-fidelity full waveform inversion. *Geophys. Suppl. Mon. Not. R. Astron. Soc.* **2016**, *207*, 1313–1331. [CrossRef]
36. Heir, A.; Aghayev, S.; Tran, C.; Molder, A. Inversion with stratigraphy-guided deep learning. *Geophysics* **2024**, *89*, R377–R386. [CrossRef]
37. Zhou, Y.; Qi, H.; Zhang, W.; Shan, X. Gamma Log Inversion of Seismic Data Based on Transformer with Stratigraphic Position Encoding. *IEEE Geosci. Remote Sens. Lett.* **2025**, *22*, 7503705. [CrossRef]
38. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.

39. Fomel, S.; Luo, S.; Zhao, H. Fast sweeping method for the factored eikonal equation. *J. Comput. Phys.* **2009**, *228*, 6440–6455. [CrossRef]
40. Baydin, A.G.; Pearlmutter, B.A.; Radul, A.; Siskind, J.M. Automatic differentiation in machine learning: A survey. *J. Mach. Learn. Res.* **2015**, *18*, 153:1–153:43.
41. Sitzmann, V.; Martel, J.; Bergman, A.; Lindell, D.; Wetzstein, G. Implicit Neural Representations with Periodic Activation Functions. In *Proceedings of the Advances in Neural Information Processing Systems*; Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., Lin, H., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2020; Volume 33, pp. 7462–7473.
42. Wang, F.; Huang, X.; Alkhalifah, T. Controllable Seismic Velocity Synthesis Using Generative Diffusion Models. *J. Geophys. Res. Mach. Learn. Comput.* **2024**, *1*, e2024JH000153. [CrossRef]
43. Zang, Y.; Koutsourelakis, P.S. DGenNO: A novel physics-aware neural operator for solving forward and inverse PDE problems based on deep, generative probabilistic modeling. *J. Comput. Phys.* **2025**, *538*, 114137. [CrossRef]
44. Deng, C.; Feng, S.; Wang, H.; Zhang, X.; Jin, P.; Feng, Y.; Zeng, Q.; Chen, Y.; Lin, Y. OpenFWI: Large-scale multi-structural benchmark datasets for full waveform inversion. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 6007–6020.
45. Aminzadeh, F.; Burkhard, N.; Long, J.; Kunz, T.; Duclos, P. Three dimensional SEG/EAEG models—An update. *Lead. Edge* **1996**, *15*, 131–134. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.