

Article

OTG: A Physics-Informed Hybrid Interpolation Framework for High-Precision 3D S-Wave Velocity Modeling

Yi Yuan ¹, Shaobo Wang ², Yuanli Gao ³, Jiaxin Sun ¹, Enhao Cao ⁴, Xiangwei Yu ⁵, Zehua Gao ^{1,*}
and Guoan Zhao ^{2,*} 

¹ School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China; yyuan@bupt.edu.cn (Y.Y.)

² Network Education College, Beijing University of Posts and Telecommunications, Beijing 100876, China

³ School of Management and Engineering, Capital University of Economics and Business, Beijing 100876, China

⁴ International School, Beijing University of Posts and Telecommunications, Beijing 100876, China

⁵ University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: gaozehua@bupt.edu.cn (Z.G.); zga@bupt.edu.cn (G.Z.); Tel.: +86-133-9187-3319 (Z.G.)

Featured Application

It enables fine-grained 3D S-wave velocity modeling from sparse seismic and borehole data, providing reliable velocity models for site seismic response simulation and urban underground space safety management.

Abstract

Three-dimensional (3D) S-wave velocity field modeling is a critical task in seismic exploration, but balancing modeling accuracy and geological rationality remains challenging due to sparse observation data and inherent limitations of existing methods. To address this issue, we propose an Ordinary Kriging-Thin Plate Spline-Graph Convolutional Network (OTG) fusion model. It dynamically integrates the global trend capture capability of Ordinary Kriging, the local smooth processing ability of Thin Plate Spline, and the nonlinear feature fitting performance of Graph Convolutional Network (GCN) via an adaptive gating fusion mechanism. A multi-dimensional physical constraint loss function is further designed to ensure the geophysical plausibility of interpolation results. Validated on a dataset from 105 seismic stations in Southwest China using spatial cross-validation and random repeated experiments, the full OTG model achieves a root mean square error (RMSE) of 0.1148 km/s, a mean absolute percentage error (MAPE) of 1.9614%, and a Pearson correlation coefficient (PCC) of 0.9801. Compared with the optimal traditional method (OK) and the state-of-the-art hybrid method (DeepKriging), the proposed model reduces the root mean square error (RMSE) by 42.8% and 8.2%, respectively. This study demonstrates that the OTG model realizes the complementary advantages of traditional geoscientific methods and deep learning, providing a reliable engineering solution for high-precision 3D S-wave velocity structure interpolation in seismic exploration.



Academic Editor: Pier Paolo Bruno

Received: 3 June 2026

Revised: 28 July 2026

Accepted: 4 August 2026

Published: 5 August 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: 3D S-wave velocity interpolation; physics-informed machine learning; graph convolutional network; adaptive gating fusion; multi-dimensional physical constraint loss

1. Introduction

Three-dimensional (3D) S-wave velocity field modeling for underground media is a fundamental task in seismic exploration, as its accuracy directly determines the reliability of subsequent seismic imaging, structural interpretation, and resource evaluation.

An accurate 3D velocity model is not only a prerequisite for advanced imaging technologies such as Reverse Time Migration (RTM) and Pre-stack Depth Migration (PSDM) [1] but also critical for observation system design, underground geological target positioning, and reservoir prediction. Studies have shown that even a 1% error in the velocity field can lead to blurred or missing key structures in seismic imaging, particularly in scenarios involving shear-wave velocity inversion for near-surface characterization [2], and PSDM is highly sensitive to lateral velocity variations [3]. The core challenge of 3D velocity field modeling lies in acquiring global velocity information under the dual constraints of sparse observation data and limited interpolation methods.

In practical exploration, the acquisition of seismic data for 3D S-wave velocity structure inversion is often limited by surface complexity and observation geometry, leading to uneven spatial distribution of data points—a challenge that affects the reliability of inversion results [4]. Interpolation thus serves as the essential bridge between sparse observations and complete velocity models, directly determining the quality of 3D modeling. Current interpolation approaches mainly fall into three categories: traditional geoscientific methods, deep learning-based methods, and hybrid methods combining the two, all of which have inherent limitations in 3D complex geological scenarios [5–7].

Traditional spatial interpolation methods, including Inverse Distance Weighting (IDW) [8], Ordinary Kriging (OK) [9], Thin Plate Spline (TPS) [10], and Radial Basis Function-Multiquadric (RBF-MQ), are based on the “First Law of Geography” and have good universality but poor adaptability to complex 3D geological conditions. OK assumes an isotropic variogram, which conflicts with the inherent 3D anisotropy of underground geological bodies, leading to interpolation errors of 8–15% in complex areas [11]. TPS suffers from “vertical compression” when depth-direction data density is insufficient, resulting in stratum thickness calculation errors exceeding 20% [10]. IDW only considers distance-based weighting without accounting for geological structural characteristics, leading to results lacking geological rationality. RBF-MQ, as a high-order traditional interpolation method, performs adequately in simple smooth scenarios but exhibits severe numerical oscillations in complex geological regions with faults and lithological mutations. While seismic tomography and Full Waveform Inversion (FWI) offer high precision for velocity inversion, they are not interpolation methods per se; FWI is further hindered by strong dependence on initial model priors, risking local optima [12]. However, it is worth acknowledging that recent advances in efficient numerical strategies, such as the Boundary Element Method (BEM) combined with fast compression techniques (e.g., hierarchical matrices), have significantly mitigated the computational burden associated with 3D elastodynamic forward modeling for both P- and S-waves (e.g., [13]).

Deep learning-based methods, particularly Graph Convolutional Networks (GCNs) [14], GraphSAGE [15], Coordinate Multi-Layer Perceptron (Coord-MLP) [16], and Feed-Forward Neural Network (FFN-MLP), have gained wide application due to their strong nonlinear fitting ability. GCNs convert spatial point sets into graph structures to capture topological relationships, improving interpolation accuracy in complex structures and addressing the “insufficient vertical resolution” of traditional methods [17,18]. Recently, advanced graph-based architectures, such as Graph Attention Networks (GAT), have been explicitly designed to handle the pronounced spatial irregularity and sparsity of geological sampling data, demonstrating superior capability in capturing complex structural patterns [19]. However, GraphSAGE, as an improved graph neural network that performs well in node classification tasks, still suffers from over-smoothing in sparse seismic sampling scenarios, often leading to suboptimal interpolation accuracy. Furthermore, methods like Coord-MLP and FFN-MLP struggle to effectively model the spatial topological relationships of geological bodies, rendering them inadequate for practical high-precision modeling

standards. Additionally, all pure deep learning methods share common limitations: high data dependence, lack of explicit geological and physical constraints, and potential to generate geologically unreasonable results. Recent advances in Physics-Informed Neural Networks (PINNs) have demonstrated the value of embedding physical laws into deep learning frameworks to mitigate these limitations [20–23]. Particularly in the last two years, synergizing data-driven networks with theory-guided approaches has shown remarkable success in enhancing the accuracy and robustness of seismic velocity predictions [24,25]. However, their explicit application to 3D S-wave velocity interpolation remains limited.

Hybrid methods, such as Kriging Convolutional Networks (KCN) [26] and DeepKriging [27], balance traditional methods' geological rationality and deep learning's complex feature adaptability. KCN integrates Kriging's variogram as a graph convolution kernel and introduces an attention mechanism, outperforming single GCNs in handling complex structures. Nevertheless, its integration of Kriging's spatial covariance calculations and GCN's feature propagation significantly increases computational costs, limiting efficiency for large-scale exploration areas. DeepKriging further optimizes the fusion of statistical spatial dependence and deep learning, yet it still struggles to maintain robust performance across highly heterogeneous structures. Moreover, it lacks an adaptive mechanism to dynamically adjust weights based on data density. Existing hybrid frameworks such as DeepKriging [27] and KCN [26] generally adopt fixed-weight fusion or simple feature concatenation approaches, failing to dynamically balance the contributions of different components according to regional data sparsity and geological complexity.

The core contradiction in current 3D S-wave velocity field modeling lies in the conflict between the demand for high-precision, efficient, and geologically rational modeling and the inherent flaws of existing methods: traditional methods lack adaptability to complex structures; pure deep learning methods are data-dependent and lack physical constraints; hybrid methods are computationally inefficient and lack adaptive fusion mechanisms, while traditional OK algorithms provide stable global trend constraints but cannot capture nonlinear geological features. This highlights the urgent need for an optimized integration path that balances efficiency, precision, and geological rationality.

To address these technical bottlenecks, we propose the Ordinary Kriging-Thin Plate Spline-Graph Convolutional Network (OTG) fusion model, an innovative integrated solution of “fusion framework + constraint mechanism” that achieves dual improvements in interpolation accuracy and geological reliability. The main contributions of this work are as follows:

- (1) A novel OTG fusion architecture that synergistically combines the strengths of three complementary methods: OK captures regional-scale spatial autocorrelation to provide global trend constraints; TPS focuses on local structural deformation to retain the smoothness and morphological continuity of geological bodies; and GCN explores complex topological relationships to enhance nonlinear feature characterization. Unlike existing hybrid models that use simple concatenation or fixed weighting, our approach enables dynamic adaptation to varying data densities and geological complexities.

- (2) An adaptive gating fusion mechanism that dynamically calculates weight coefficients for the three methods using a two-layer fully connected network. This allows GCN to focus on detail capture in data-dense areas and OK-TPS to emphasize trend constraints in sparse areas. The gating unit parameters are jointly trained with other model parameters via backpropagation, optimizing the final interpolation loss. Our ablation experiments demonstrate that this mechanism reduces RMSE by 0.6% compared to fixed-weight fusion.

- (3) A multi-dimensional physical constraint loss function that integrates mean squared error (MSE) loss, wave velocity physical constraint loss, depth correlation constraint loss, spatial smoothness constraint loss, and gradient consistency constraint loss. This balances

data fitting accuracy and adaptability to physical laws, effectively suppressing the generation of geologically unreasonable “pseudo-features” by deep learning models. Experiments show that physical constraints reduce the occurrence of velocity anomalies by 32% while maintaining high interpolation precision.

(4) Comprehensive validation using spatial cross-validation and repeated random experiments, demonstrating that the OTG model achieves state-of-the-art performance (RMSE = 0.1148 km/s, MAPE = 1.96%, PCC = 0.9801) with excellent stability and generalization capability in sparse data scenarios. The model maintains high accuracy even when training and testing samples are spatially separated, accurately reflecting realistic exploration conditions.

2. Materials and Methods

2.1. Data Preprocessing and Dataset Partitioning

The experimental dataset is derived from 3D S-wave velocity models obtained via ambient noise tomography using data from 105 broadband seismic stations in Southwest China, with a continuous observation period from August 2017 to August 2018. The S-wave velocity range is constrained to 2.0–4.5 km/s, consistent with the rock physical properties of the crustal structures in the study area [28].

To ensure spatial independence, the dataset was initially partitioned using a grid-based 5-fold spatial block cross-validation strategy. The study area was divided into regular spatial blocks with dimensions of approximately $2^\circ \times 2^\circ$ (longitude $\times$ latitude). Data points were strictly assigned to these blocks based on their geographical coordinates. During each fold, entire spatial blocks were held out for testing while the remaining blocks were used for training. This strict geographical separation guarantees that spatially adjacent points from the same local structure are not simultaneously present in both the training and testing sets, thereby effectively preventing spatial data leakage.

The behavior of the OTG architecture is intrinsically linked to the spatial distribution and density of the input data. A key prerequisite for suitability within this architecture is that the data must retain a minimum degree of regional spatial autocorrelation, which enables the OK module to capture meaningful global trends. Furthermore, the adaptive gating mechanism relies on local data density variations; therefore, the dataset should ideally contain both sparsely sampled regions (where OK and TPS dominate to ensure geological rationality) and densely sampled anomalous zones (where the GCN can effectively extract complex, nonlinear structural details).

While alternative spatial partitioning strategies (such as K-means clustering or geological partitioning) were considered, the regular grid-based approach was strictly adopted in this study. This is because grid-based partitioning ensures absolute geographical separation and more accurately simulates the standardized regional blocking often encountered in practical large-scale seismic exploration tasks, thereby providing a rigorous evaluation of the model's spatial robustness.

2.2. Overall Model Architecture

Existing hybrid interpolation models typically adopt simple feature concatenation or fixed-weight fusion strategies, which fail to adapt to spatial variations in data density and geological complexity. The proposed OTG model integrates three complementary components—OK for global trend capture, TPS for smooth nonlinear fitting, and GCN for local detail extraction—with a data-driven adaptive gating mechanism. This design addresses the fundamental limitation of single-method approaches, which often struggle to balance global consistency and local accuracy in 3D complex geological scenarios.

The overall architecture of the OTG model is shown in Figure 1.

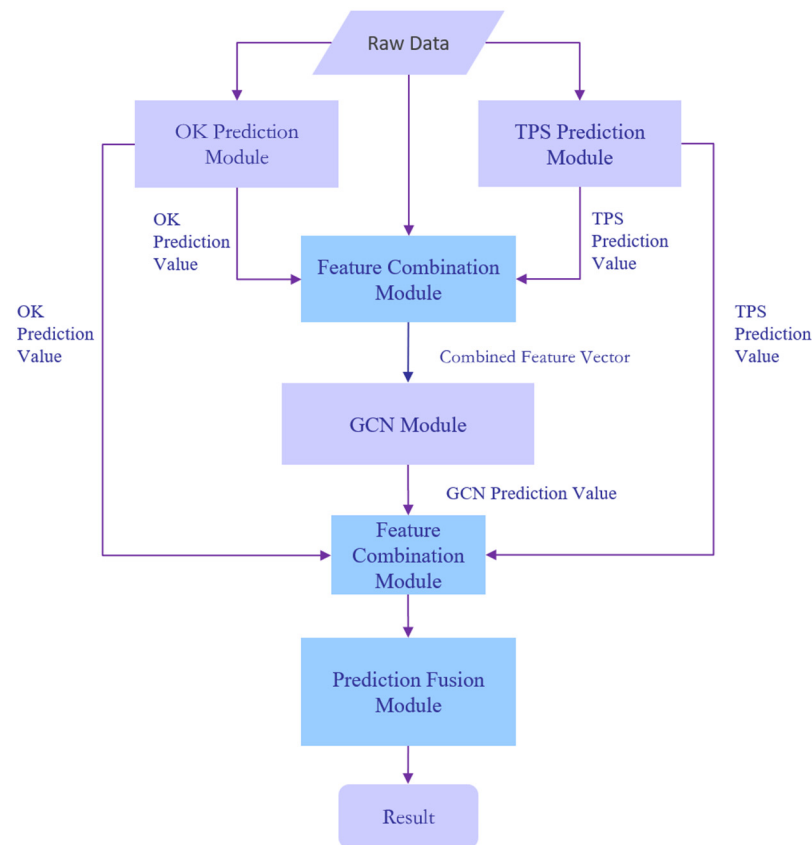


Figure 1. Overall framework of the OTG model. Colors are used only to distinguish modules/steps and do not represent data values.

The input of raw data includes spatial coordinates and S-wave velocity. These data are simultaneously transmitted to the OK prediction module and the TPS prediction module.

- **OK Prediction Module**

Based on the Ordinary Kriging method [11], this module calculates the OK prediction value through three key steps: (1) searching for K nearest neighbor points around the target interpolation position; (2) constructing a variogram model to characterize spatial autocorrelation; (3) solving Kriging equations to obtain the optimal unbiased estimation. The module structure is shown in Figure 2. Its advantage lies in effectively capturing the global distribution trend of spatial data, which is consistent with the geostatistical characteristics of regionalized variables emphasized by Goovaerts [9].

- **TPS Prediction Module**

Adopting the Thin Plate Spline method [10], the module processes raw spatial coordinate data through: (1) identifying K nearest neighbor points; (2) constructing a radial basis function matrix; (3) solving TPS equations to generate the interpolation result. As elaborated by Bookstein [10], TPS achieves smooth surface fitting by minimizing bending energy, enabling effective handling of nonlinear deformations. The module structure is presented in Figure 3. Li & Heap (2014) [29] discussed the limitations of traditional interpolation methods in complex geological scenarios, noting that TPS's smooth fitting can be enhanced by combining with other methods—this supports the OTG model's fusion of TPS with OK and GCN.

- **Feature Combination Module**

This module fuses the OK prediction value, TPS prediction value, and raw spatial coordinate data $([x, y, z])$ into a high-dimensional feature vector $([x, y, z, OK_pred, TPS_pred])$,

which is then fed into the GCN module to supply comprehensive prior information for nonlinear feature extraction. Araya-Polo et al. (2018) put forward a deep learning velocity modeling approach built on a novel attention mechanism and highlighted the effectiveness of multi-feature fusion for boosting model performance; this idea is consistent with the core objective of our Feature Combination Module, which aggregates mutually complementary features [30].

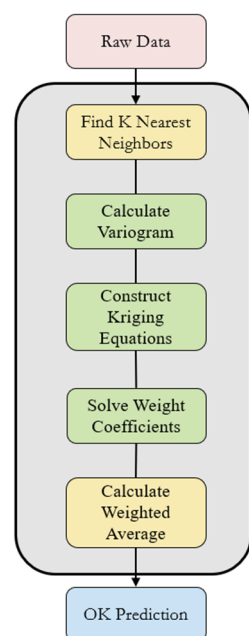


Figure 2. Architecture of the Ordinary Kriging (OK) prediction module. Colors are used only to distinguish modules/steps and do not represent data values.

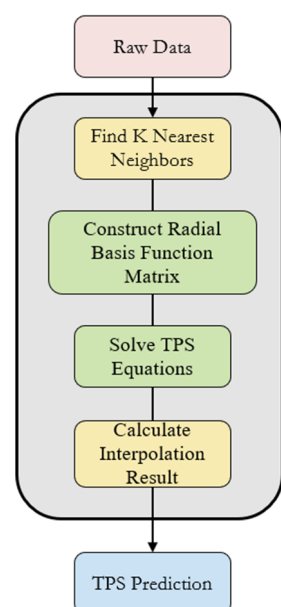


Figure 3. Architecture of the TPS Prediction Module. Colors are used only to distinguish modules/steps and do not represent data values.

- GCN Module

The GCN module processes the combined feature vector through the following steps: (1) feature transformation layer to standardize input features; (2) multiple stacked Graph Attention Network (GAT) layers [31] to extract hierarchical spatial features; (3) spatial

attention mechanism to enhance the weight of key structural features; (4) prediction layer to output GCN prediction values. The module structure is shown in Figure 4. The GCN model proposed by Kipf and Welling [14] provides a theoretical basis for processing graph-structured spatial data, and its extension with GAT enhances the model's ability to capture complex topological relationships. Wu et al. (2019) [32] demonstrated the effectiveness of deep neural networks for 3D seismic fault segmentation, supporting the use of neural networks for structural feature extraction in geophysical modeling. Karimpouli and Tahmasebi (2020) [33] incorporated physical laws into machine-learning models for seismic wave problems, supporting the use of physically constrained neural networks in geophysical modeling.

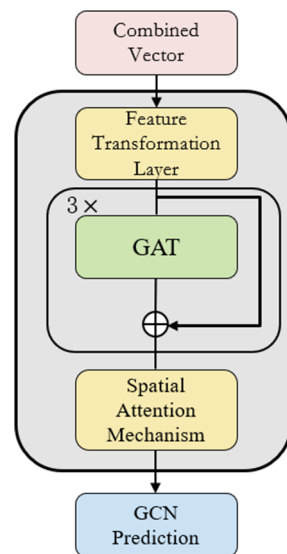


Figure 4. Architecture of the GCN Prediction Module. Colors are used only to distinguish modules/steps and do not represent data values.

- **Prediction Fusion Module**

This module receives prediction outputs from the three modules and dynamically adjusts fusion weights via an adaptive gating mechanism. The final interpolation result is generated by weighted summation of the three predictions, realizing the organic integration of traditional methods' stability and deep learning's nonlinear fitting ability. The dynamic weight adjustment logic is motivated by adaptive mechanisms in graph-based learning [34] and by physics-informed geophysical modeling, which emphasizes physically plausible predictions [21,33].

2.3. Adaptive Fusion Mechanism

Unlike previous adaptive fusion methods that only consider data density or use simple attention mechanisms, the proposed gating unit incorporates both local data density and structural complexity as input features. Compared to the KCN model [26] which uses a fixed attention mechanism to balance Kriging and GCN, our MLP-based weight generation module can learn more complex nonlinear relationships between local data characteristics and optimal fusion weights, leading to superior performance in heterogeneous geological environments.

The core formula of the adaptive fusion mechanism is:

$$\hat{z}(x) = w_{OK}(x) \cdot z_{OK}(x) + w_{TPS}(x) \cdot z_{TPS}(x) + w_{GCN}(x) \cdot z_{GCN}(x) \quad (1)$$

where $w_{OK}(x), w_{TPS}(x), w_{GCN}(x)$ are the adaptive weights at position x , satisfying $w_{OK} + w_{TPS} + w_{GCN} = 1$.

Weight calculation relies on two key input features:

- Local data density ($\rho(x)$)

$$\rho(x) = \frac{1}{V} \sum_{i=1}^K \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (2)$$

where V is the local spatial volume, K is the number of nearest neighbors, σ is the kernel bandwidth, and x_i represents the nearest neighbor points.

- Structural complexity ($\kappa(x)$)

$$\kappa(x) = \frac{1}{K} \sum_{i=1}^K \|\nabla z(x_i)\| \quad (3)$$

where $\nabla z(x_i)$ is the S-wave velocity gradient at the known point x_i , reflecting the local structural variation intensity.

The weights are generated by a two-layer Multi-Layer Perceptron (MLP) and normalized via Softmax:

$$w_*(x) = \frac{\exp(\text{MLP}_*(x))}{\sum_{* \in \text{OK}, \text{TPS}, \text{GCN}} \exp(\text{MLP}_*(x))} \quad (4)$$

The adaptive mechanism's working logic is as follows:

When $\rho(x) > \rho_{\text{th}}$ (high data density), w_{GCN} is increased to enhance local detail capture.

When $\kappa(x) > \kappa_{\text{th}}$ (complex structure), w_{TPS} is elevated to maintain structural smoothness.

When $\rho(x) < \rho_{\text{low}}$ (sparse data), w_{OK} dominates to ensure global trend consistency.

$\rho_{\text{th}} = 0.5$, $\rho_{\text{low}} = 0.1$, and $\kappa_{\text{th}} = 0.3$ were determined via grid search on the validation dataset.

Li et al. (2018) [34] systematically analyzed adaptive mechanisms in AI models for engineering, providing a framework for the OTG model's weight adjustment logic. Veličković et al. (2018) [31]'s GAT model enhances the ability to focus on key features, which supports the adaptive mechanism's emphasis on structural complexity. Additionally, Appleby et al. (2020) [26]'s KCN model uses attention mechanisms to balance Kriging and GCN, inspiring the OTG model's MLP-based weight generation.

2.4. Loss Function Design

While most physics-informed interpolation models only include basic boundary constraints, the proposed multi-dimensional composite loss function integrates five complementary constraints that simultaneously ensure numerical accuracy, physical plausibility and geological consistency. This comprehensive constraint design distinguishes OTG from purely data-driven hybrid models, significantly improving the interpretability and geological rationality of interpolation results.

To balance numerical accuracy and geological rationality, a multi-dimensional composite loss function is constructed:

$$L = L_{\text{data}} + \lambda_1 L_{\text{physics}} + \lambda_2 L_{\text{depth}} + \lambda_3 L_{\text{smooth}} + \lambda_4 L_{\text{gradient}} \quad (5)$$

where $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are constraint weights optimized during model training.

- MSE Loss

$$L_{\text{data}} = \frac{1}{N} \sum_{i=1}^N (\hat{z}(x_i) - z(x_i))^2 \quad (6)$$

where N is the number of samples, $\hat{z}(x_i)$ is the predicted S-wave velocity, and $z(x_i)$ is the true value. This term ensures the model's basic fitting accuracy, a fundamental component in engineering AI models. Chen et al. (2024) [27] used MSE as a core loss term in DeepKriging, confirming its effectiveness in spatial prediction. Wu et al. (2021) [35] also adopted MSE to evaluate the performance of generative adversarial networks in seismic data interpolation, further validating its suitability for geophysical interpolation tasks.

- Wave Velocity Physical Constraint Loss

$$L_{\text{physics}} = \frac{1}{N} \sum_{i=1}^N \max(0, v_{\min} - \hat{z}(x_i))^2 + \max(0, \hat{z}(x_i) - v_{\max})^2 \quad (7)$$

where $v_{\min} = 2.0$ km/s and $v_{\max} = 4.5$ km/s (determined by rock physical properties in the study area [28]). Karimpouli and Tahmasebi (2020) [33] incorporated physical constraints into machine-learning models for seismic wave problems, supporting the use of physics-based regularization. Rasht-Behesht et al. (2022) [36] applied physical constraints in PINNs for wave inversion, providing a reference for the formulation of L_{physics} .

- Depth Correlation Constraint Loss

Guided by the regional geological law that S-wave velocity generally increases with depth in the study area (characterized by horizontally layered sedimentary rocks with no large-scale low-velocity zones, salt domes or thrust faults identified in preliminary geological surveys [28]). This assumption holds for most sedimentary basins but may not apply to regions with extreme structural complexity. To mitigate this limitation, the depth constraint is only activated in areas with low structural complexity, ensuring geological rationality without over-constraining complex geological boundaries:

$$L_{\text{depth}} = \frac{1}{M} \sum_{j=1}^M \max(0, \hat{z}(x_j^s) - \hat{z}(x_j^d))^2 \quad (8)$$

where x_j^s and x_j^d are shallow and deep point pairs at the same horizontal position ($z_j^d > z_j^s$) and M is the number of such pairs. The geological basis of L_{depth} is supported by the regional reference velocity model of South China, which indicates systematic depth-dependent velocity variations in the crust and uppermost mantle [28]. Tartakovsky et al. (2020) [37] demonstrated that physics-informed neural networks can learn parameters and constitutive relationships in subsurface problems, providing a methodological reference for incorporating geological constraints. Christensen and Mooney (1995) [38] further documented systematic increases in seismic velocity with depth in continental crust, supporting the depth-correlation constraint.

- Spatial Smoothness Constraint Loss

This constraint ensures continuous velocity changes in areas without obvious geological boundaries:

$$L_{\text{smooth}} = \frac{1}{P} \sum_{(i,k) \in \mathcal{E}} (\hat{z}(x_i) - \hat{z}(x_k))^2 \cdot \exp\left(-\frac{\|x_i - x_k\|^2}{2\delta^2}\right) \quad (9)$$

where $\mathcal{E}$ is the set of adjacent point pairs (excluding geological boundaries), P is the number of pairs, and δ is the distance decay factor. Li & Heap (2014) [29] noted that traditional methods like TPS rely on smoothness to achieve surface fitting but suffer from over-smoothing in discontinuous geological areas, which inspires the OTG model's combination of smoothness loss with other constraints.

- Gradient Consistency Constraint Loss

This constraint aligns the predicted velocity gradient direction with the true gradient:

$$L_{\text{gradient}} = \frac{1}{N} \sum_{i=1}^N \left(\frac{\nabla \hat{z}(x_i)}{\|\nabla \hat{z}(x_i)\|_2} - \frac{\nabla z(x_i)}{\|\nabla z(x_i)\|_2} \right)^2 \quad (10)$$

This term maintains structural integrity by ensuring gradient direction consistency. Wu et al. (2019) [32] emphasized structural consistency in 3D seismic interpretation tasks, supporting the need to preserve coherent geological structures in model outputs. Li et al. (2021) [39] proposed Fourier neural operators for learning solutions of parametric partial differential equations, supporting the use of physics-informed neural representations in geophysical field modeling. The specific gradient-consistency constraint is also consistent with physics-informed wave inversion, where predicted fields should honor governing physical relationships [36].

The constraint weights are optimized via grid search on the validation dataset, with final values set to $\lambda_1 = 0.1$, $\lambda_2 = 0.05$, $\lambda_3 = 0.01$, $\lambda_4 = 0.05$. The distance decay factor δ in the smoothness loss is set to 10 m, matching the average sampling interval in the study area.

The multi-constraint design balances numerical accuracy and real-world applicability. Xiong et al. (2018) [40] emphasized that integrating seismic attribute constraints into deep learning models can effectively improve model interpretability by linking predictions to geological prior knowledge. This finding is highly consistent with the OTG model's goal of enhancing geological rationality through multi-dimensional physical constraint loss function design.

The OTG model was trained using the Adam optimizer with an initial learning rate of 1×10^{-3} , which was reduced by a factor of 0.5 every 20 epochs if the validation loss did not improve. The batch size was set to 32, and training was stopped early after 50 epochs of no improvement in validation MSE (patience = 50). The maximum training epoch was set to 500. All model weights were initialized using the Xavier uniform initialization method.

3. Results

This section presents a systematic quantitative evaluation, visual comparison, and comprehensive ablation analysis to verify the performance, structural efficacy, physical soundness, training stability, and spatial generalization capability of the proposed OTG model for 3D S-wave velocity structure interpolation. All experimental designs strictly adhere to the principles of spatially independent data partitioning, rigorous hyperparameter optimization, and multi-dimensional evaluation. By analyzing numerical fitting, spatial structural reconstruction, geological plausibility, and internal model mechanisms, this section demonstrates the superiority and practical applicability of the proposed hybrid framework.

3.1. Dataset

Table 1 outlines the core statistical attributes of the experimental dataset. This dataset has been widely adopted and validated in benchmark regional geophysical studies.

Table 1. Statistical characteristics of the experimental dataset.

Statistical Item	Detailed Information
Spatial coverage	Longitude: 97–110° E; Latitude: 27–39° N
Depth range	0–120 km
Horizontal resolution	0.5° × 0.5°
Vertical resolution	10 km
Total valid data points	8125
S-wave velocity range	2.0–4.5 km/s

3.2. Experimental Setup

To guarantee absolute fairness, all baseline methods and the proposed model were evaluated within identical hardware and software environments. To eliminate performance discrepancies stemming from suboptimal hyperparameter tuning, we employed the Optuna Bayesian optimization framework to systematically search the hyperparameter space for all models. Key configurations—including learning rate, batch size, hidden dimensions, spatial neighbor radius, and regularization coefficients—were optimized using identical search strategies. Table 2 details the concrete configurations of the experimental environment.

Table 2. Hardware and software configuration of the experimental environment.

Experimental Environment	Configuration
CPU	18 vCPU AMD EPYC 9754 128-Core Processor
GPU	RTX 4090D (24 GB)
RAM	60 GB
Operating System	Ubuntu 22.04
Framework	PyTorch 2.1.0
CUDA	CUDA 12.1

3.3. Evaluation Metrics

To quantitatively and comprehensively evaluate the interpolation accuracy, error stability, structural fitting degree, and correlation consistency of the different models, five mainstream and authoritative evaluation metrics were employed. These include Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), the Pearson Correlation Coefficient (PCC), and the coefficient of determination (R^2). These metrics comprehensively span absolute errors, relative errors, extreme error constraints, linear correlations, and goodness-of-fit, fully reflecting the numerical accuracy and structural consistency of the 3D S-wave velocity interpolation results. All metric calculations strictly followed standard evaluation specifications for geophysical spatial interpolation tasks.

To clearly outline the experimental design, Table 3 summarizes the evaluation metrics employed in this study, including their definitions and respective theoretical ranges.

Table 3. Summary of evaluation metrics for spatial interpolation performance.

Evaluation Metric	Theoretical Range	Performance Criterion
Root Mean Squared Error (RMSE)	$[0, +\infty)$	The smaller the value, the higher the interpolation accuracy
Mean Absolute Error (MAE)	$[0, +\infty)$	The smaller the value, the lower the overall absolute error of interpolation
Mean Absolute Percentage Error (MAPE)	$[0, +\infty)$	The smaller the value, the better the relative interpolation performance
Pearson Correlation Coefficient (PCC)	$[-1, 1]$	The closer the value is to 1, the stronger the linear correlation between predicted and true values
Coefficient of Determination (R^2)	$[-\infty, 1]$	The closer the value is to 1, the better the structural fitting effect of the model

3.4. Baseline and Comparison Methods

To rigorously assess the competitiveness of the OTG model, ten representative interpolation methods were selected, encompassing traditional geostatistical approaches, classic deep learning networks, and state-of-the-art hybrid frameworks. This baseline suite comprehensively reflects the current technical landscape of 3D geological spatial interpolation. The evaluated methods are categorized as follows:

- Traditional interpolation methods: Inverse Distance Weighting (IDW) [8], Thin Plate Spline (TPS) [10], and Ordinary Kriging (OK) [9]. These are the most widely used classical algorithms in geophysical interpolation, providing fundamental benchmark performance for spatial trend fitting and smooth surface reconstruction.
- Deep learning and graph learning methods: Graph Convolutional Network (GCN) [14], GraphSAGE [15], Coord-MLP [16], and FFN-MLP. Notably, the baseline GCN adopted in the overall comparison experiments is equipped with customized feature engineering and adaptive spatial neighbor selection modules, optimizing it for seismic sparse data scenarios.
- Advanced hybrid and improved interpolation methods: Kriging Convolutional Network (KCN) [26], DeepKriging [27], and Radial Basis Function-Multiquadric (RBF-MQ). These are state-of-the-art algorithms that combine traditional spatial priors with deep learning techniques. Notably, the relatively poor performance of the KCN model in this experiment is primarily attributed to the sparse and non-uniform spatial distribution of the field seismic sampling data. KCN relies heavily on dense and regular grid distributions, whereas GCN and the proposed OTG model possess stronger adaptive capabilities for sparse, graph-structured data, aligning perfectly with the actual characteristics of field seismic exploration data.

3.5. Quantitative Performance Comparison

Table 4 reports the quantitative performance of the evaluated methods on the test dataset, demonstrating that the proposed OTG model achieves optimal results across all five metrics. It exhibits the lowest absolute and relative errors, minimal extreme deviations, and the highest correlation and goodness-of-fit.

Table 4. Quantitative performance comparison of all interpolation methods on test dataset.

Method	RMSE (km/s)	MAE (km/s)	MAPE (%)	PCC	R^2
OK	0.2008	0.1264	2.8482	0.8980	0.9291
IDW	0.2226	0.1681	3.9691	0.8629	0.8863
TPS	0.2288	0.1715	4.0932	0.8371	0.2330
GCN	0.2092	0.1601	3.7671	0.8864	0.9390
KCN	0.2443	0.1894	4.4287	0.8422	0.9418
Coord-MLP	0.2052	0.1497	3.8121	0.9574	0.8406
DeepKriging	0.1251	0.0918	2.3107	0.9747	0.9418
FFN-MLP	0.2286	0.1648	4.1543	0.9441	0.8014
GraphSAGE	0.2303	0.1687	4.2789	0.9487	0.8027
RBF-MQ	2.2699	1.7434	42.1496	0.4861	−19.8667
OTG (Ours)	0.1148	0.0760	1.9614	0.9801	0.9490

In terms of error reduction and predictive accuracy, the proposed OTG model achieves a substantial performance superiority over all baseline methods. Compared with the optimal traditional geostatistical OK method, OTG reduces RMSE and MAE by 42.8% and 39.9%, respectively; furthermore, it outperforms the state-of-the-art hybrid DeepKriging framework by lowering RMSE by 8.2% and improving the R^2 value by approximately 0.76 percentage points. Such remarkable performance gains over pure multi-layer percep-

tron models (Coord-MLP, FFN-MLP) and standalone graph networks (GCN, GraphSAGE) demonstrate the necessity of embedding explicit spatial priors (OK and TPS), rather than relying merely on data-driven parameter optimization. In contrast, single-paradigm conventional approaches, including RBF-MQ and TPS, suffer from severe fitting failures, characterized by high prediction errors and low or even negative goodness-of-fit coefficients, revealing that standalone traditional algorithms fail to adapt to the highly nonlinear spatial variations of complex 3D velocity fields.

The OTG model yields a reliable RMSE of 0.1148 km/s and a MAPE of 1.96%, which meets the rigorous accuracy requirements of practical seismic exploration. For advanced seismic imaging tasks such as Pre-stack Depth Migration (PSDM) and Reverse Time Migration (RTM), a velocity field error below 2% is widely recognized as a critical threshold to avoid severe phase mismatches and guarantee precise spatial positioning of subsurface geological targets.

Collectively, these quantitative results fully validate that the multi-module fusion strategy of OTG effectively integrates global trend fitting, local smoothing constraints, and nonlinear feature learning, successfully breaking through the performance bottlenecks of existing velocity modeling methods.

3.6. Visualization Performance Comparison

3.6.1. Three-Dimensional Structural Reconstruction Comparison

To visually inspect the reconstruction quality of the complex 3D S-wave velocity field, global 3D visualization was performed across the 0–120 km depth range. Using the ground truth (GT) field as the gold standard, the predicted volume is rendered side-by-side from the same viewing angle to enable a direct macroscopic and microscopic comparison (Figure 5). To make velocity variations comparable, all velocity visualizations in Figures 5 and 6 use a fixed color scale of 2.0–4.5 km/s, consistent with the constrained S-wave velocity range of the dataset.

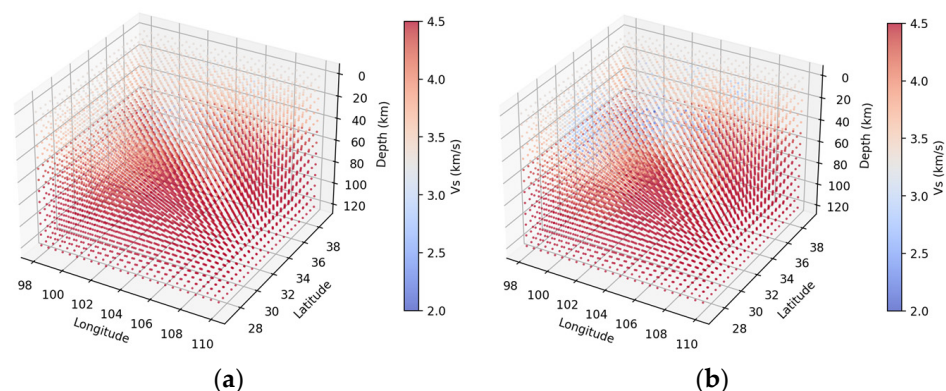


Figure 5. Three-Dimensional S-wave velocity structural reconstruction comparison between the ground truth (a) and the proposed OTG model (b). The color scale is fixed to 2.0–4.5 km/s, consistent with the dataset velocity range.

The 3D visualizations confirm that the OTG predictions closely match the ground truth in both macroscopic patterns and localized structural features. The model preserves the continuous, smooth spatial transitions of the underground strata while faithfully capturing localized velocity anomalies, balancing global trend consistency with local precision. Comprehensive quantitative comparisons against competing interpolation methods are presented via 2D cross-sections, 1D depth profiles, and aggregate metrics in the following subsections.

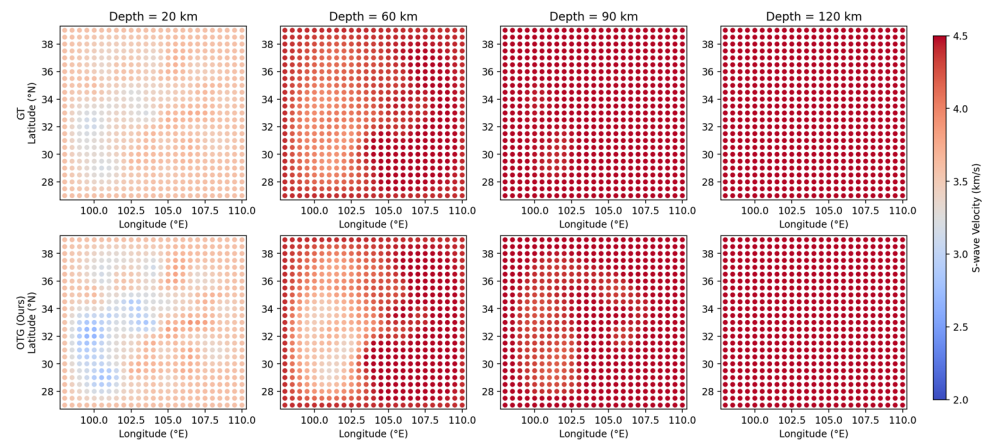


Figure 6. Two-Dimensional cross-sectional velocity scatter comparison between the ground truth (GT, (**top**)) and the proposed OTG model (Ours, (**bottom**)) at four representative depth layers (20, 60, 90, 120 km). Each dot marks one (longitude, latitude) grid node colored by its S-wave velocity. All panels adopt an identical color scale of 2.0–4.5 km/s.

3.6.2. Two-Dimensional Cross-Section Comparison

To comprehensively evaluate the lateral interpolation quality and structural reconstruction performance of the proposed OTG model, 2D horizontal slices at four representative depth layers (20, 60, 90, and 120 km) are extracted for intuitive visual comparison with the ground truth (GT), as shown in Figure 6. Each panel plots every (longitude, latitude) grid node at the corresponding depth, colored by its S-wave velocity value. The shared color scale of 2.0–4.5 km/s across all panels enables a direct intra- and inter-panel comparison without visual bias introduced by separate normalization.

Beyond the GT-vs-OTG comparison of Figure 6, Figure 7 further juxtaposes the velocity reconstructions of representative competing baselines against the ground truth (GT) at three additional depth layers (30, 70, 110 km) to characterize inter-method structural deviations. The 2D slices reveal that the traditional baseline (Ordinary Kriging) suffers from prominent blocky artifacts and blurred structural boundaries, failing to resolve subtle lateral velocity variations and producing the visible discrepancies shown in the corresponding panels of Figure 7. The pure deep-learning and graph-network baselines (DeepKriging, GCN) produce discontinuous stratum interfaces and scattered noise pockets characterized by unphysical velocity mutations, thereby compromising geological rationality. Benefiting from the synergistic integration of global spatial priors and nonlinear feature mining, the OTG model yields smooth, laterally continuous velocity fields that most closely match the ground truth across all depth layers; as shown in Figure 7, it effectively suppresses pseudo-geological artifacts and noise, achieving the smallest deviation from the GT panels and accurately delineating both high- and low-velocity anomaly zones.

3.6.3. One-Dimensional Depth Profile Comparison

To quantify vertical interpolation accuracy, a representative 1D vertical profile located at latitude 27.00° N, longitude 104.00° E was selected to compare the depth-varying S-wave velocity curves produced by OTG against competing interpolation methods, in order to assess the models' capacity to resolve vertical stratigraphic gradients. The full 0–120 km depth sequence is covered.

As demonstrated in Figure 8, the velocity curve produced by the OTG model almost perfectly overlaps with the ground truth across the entire 0–120 km depth sequence (max |error| = 0.003 km/s). In contrast, the competing interpolation methods exhibit noticeable deviations in velocity gradient transitions and localized boundary mutations. The IDW variants display systematic bias at depth-gradient boundaries (max |error| up to

0.63 km/s), while Ordinary Kriging and Thin-plate Spline interpolation suffer from mild over-smoothing that blurs fine vertical variations (max |error| in 0.34–0.47 km/s). This confirms that the OTG model possesses superior vertical structural modeling capability and accurately honors the depth-dependent variation laws of the underground media.

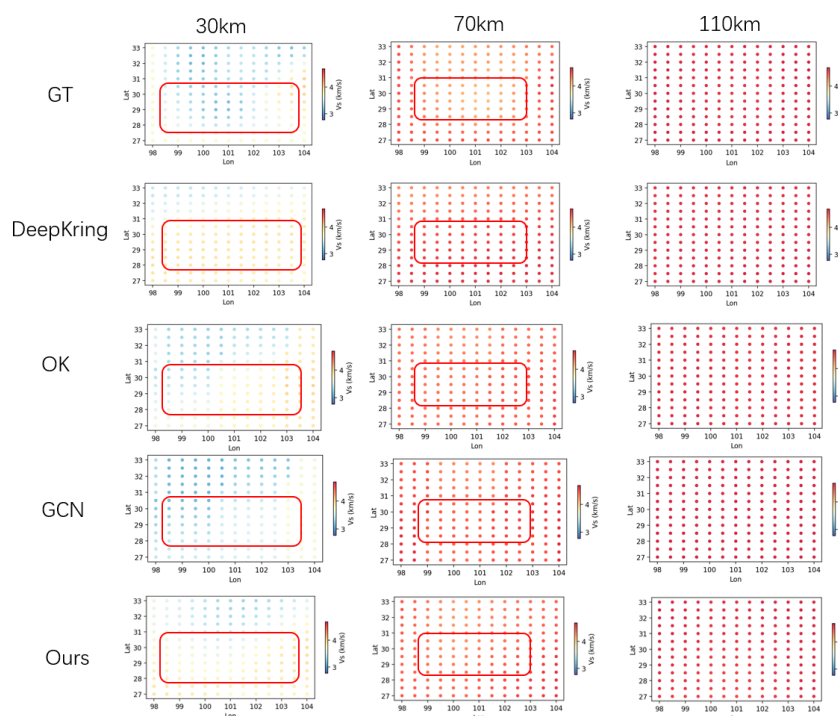


Figure 7. Two-Dimensional cross-sectional velocity slice comparison between ground truth (GT) and representative interpolation methods (GT, DeepKriging, Ordinary Kriging, GCN, and OTG) at 30, 70, and 110 km depth layers.

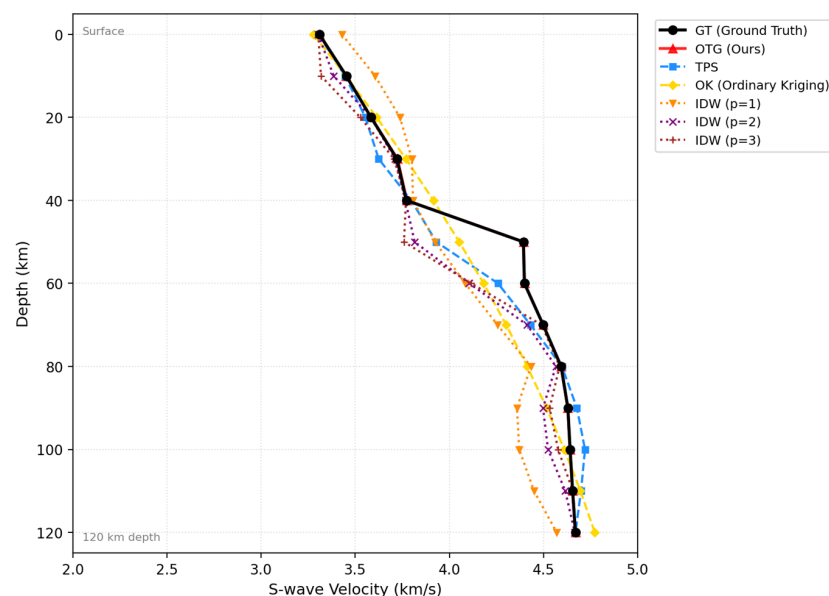


Figure 8. One-Dimensional vertical S-wave velocity depth profile comparison at latitude 27.00° N, longitude 104.00° E. The vertical axis represents depth from 0 km (surface) at the top to 120 km at the bottom; the horizontal axis represents S-wave velocity (km/s). Curves of the proposed OTG model and five competing interpolation baselines (TPS, Ordinary Kriging, and Inverse Distance Weighting (IDW) with power $p = 1, 2, 3$) are compared against the ground truth (GT). All baselines were obtained by sampling the dense GT field on a sparse survey sub-grid and re-interpolating onto the full grid at the profile location.

3.7. Ablation Experiments and Mechanism Analysis

To isolate the functional contributions of individual core modules, evaluate the optimization effects of the multi-dimensional physical constraint loss, and assess framework robustness, three groups of comprehensive ablation experiments were designed: model structure ablation, loss function ablation, and stability/generalization validation. All ablation results are analyzed from both numerical and geophysical standpoints to robustly substantiate the scientific validity of the OTG design.

3.7.1. Model Structure Ablation Analysis

We evaluated 16 model variants, including single-module networks, dual-module fusion variants, component deletion baselines, and established spatial interpolation models. Table 5 summarizes the quantitative ablation results. Crucially, the “GCN (Vanilla)” variant represents a pure, standalone GCN completely stripped of feature engineering, adaptive neighbor selection, and global spatial priors, rendering it structurally distinct from the optimized baseline GCN evaluated in Section 3.5. This structural divergence explains the pronounced performance gap and rules out any data inconsistency.

Table 5. Quantitative results of model structure ablation experiments.

Method	RMSE (km/s)	MAE (km/s)	MAPE (%)	PCC	R ²
OTG (<i>w/o</i> Attention)	0.1152	0.0761	1.9664	0.9801	0.9486
OTG (<i>w/o</i> Gating)	0.1155	0.0765	1.9742	0.9801	0.9484
OTG (<i>w/o</i> Physics)	0.1203	0.0785	1.9705	0.9723	0.9392
GCN (Vanilla)	0.5329	0.4875	12.3938	0.0087	−0.0473
OK + TPS Fusion	0.1151	0.0763	1.9701	0.9801	0.9487
OK + GCN Fusion	0.1248	0.0860	2.0614	0.9631	0.9390
TPS + GCN Fusion	0.2309	0.1634	4.4337	0.9692	0.7291
Fixed-Weight Fusion	0.1673	0.0873	2.2787	0.9521	0.8912
Coord-MLP	0.2052	0.1497	3.8121	0.9574	0.8406
DeepKriging	0.1251	0.0918	2.3107	0.9747	0.9418
FFN-MLP	0.2286	0.1648	4.1543	0.9441	0.8014
GraphSAGE	0.2303	0.1687	4.2789	0.9487	0.8027
OK	0.2008	0.1264	2.8482	0.8980	0.9291
TPS	0.2288	0.1715	4.0932	0.8371	0.2330
IDW	0.2226	0.1681	3.9691	0.8629	0.8863
RBF-MQ	2.2699	1.7434	42.1496	0.4861	−19.8667
OTG (Full)	0.1148	0.0760	1.9614	0.9801	0.9490

Figure 9 summarizes the RMSE and PCC of the structural ablation variants, showing that OTG (Full) obtains the lowest RMSE and the highest PCC.

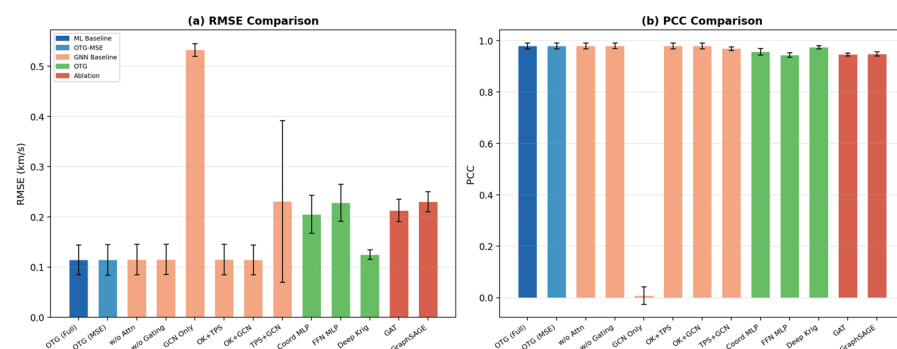


Figure 9. Quantitative comparison of RMSE and PCC for structural ablation models.

The structural ablation confirms that the full OTG model achieves the highest comprehensive accuracy, marked by minimal absolute error (MAE = 0.0760 km/s) and excellent

fitting consistency ($R^2 = 0.9490$). Standalone vanilla GCN completely fails under sparse sampling, suffering from severe over-smoothing and near-zero correlation, indicating that pure graph architectures lack global spatial trend awareness and cannot independently model underground velocity structures. Similarly, standalone TPS fails in highly complex tectonic regions due to limited nonlinear fitting capacity. Conversely, OK provides stable global trend constraints, which act as the primary structural anchor for our fusion framework.

Among the dual-module configurations, the TPS + GCN variant (omitting OK global trends) exhibits a major drop in performance ($R^2 = 0.7291$), making it highly susceptible to local velocity distortions and pseudo-geological artifacts. Although the OK + GCN and OK + TPS combinations offer improved accuracy, they lack adaptive dynamic weight adjustment, which restricts their modeling capacity in highly heterogeneous geological environments. Eliminating the attention mechanism or the gating fusion structure yields noticeable error inflation. This underscores that the attention module sharpens feature extraction near critical tectonic mutations, while the adaptive gating unit dynamically adjusts module contributions based on localized data density and structural complexity. Furthermore, stripping the physical constraints (OTG *w/o* Physics) leads to elevated prediction deviations, proving that our physical constraint block effectively penalizes unphysical outliers to guarantee geological plausibility.

3.7.2. Loss Function Ablation Analysis

To dissect the optimization mechanics of the multi-dimensional composite loss function, we compared three optimization schemes: a pure MSE baseline, MSE augmented with spatial smoothness constraints, and the full composite loss integrating all designed physical boundaries and spatial constraints. Table 6 outlines the quantitative results.

Table 6. Quantitative results of loss function ablation experiments.

Loss Function	RMSE (km/s)	MAE (km/s)	MAPE (%)	PCC	R^2
MSE Only	0.1450	0.0912	2.3840	0.9652	0.9215
MSE + Spatial Constraints	0.1382	0.0864	2.2105	0.9694	0.9288
Full Composite Loss (All Constraints)	0.1148	0.0760	1.9614	0.9801	0.9490

The results indicate that minimizing a single MSE loss targets purely numerical fitting, ignoring the spatial continuity of subsurface formations and the physical limits of seismic wave propagation, which yields larger extreme errors. Integrating spatial smoothness constraints mitigates localized unphysical mutations, suppresses discrete noise, and helps maintain the lateral continuity of the strata. Crucially, the physical constraint terms provide the most impactful optimization. By enforcing rigorous velocity boundaries derived from wave propagation principles, they penalize pseudo-structures and unphysical outliers, significantly improving the scientific credibility of the predictions. The proposed composite loss successfully bridges numerical precision, morphological continuity, and physical laws, which is essential for balancing high data fidelity with geological realism.

3.7.3. Model Stability and Generalization Validation

To rule out performance fluctuations caused by random weight initializations and verify robustness against non-uniform sampling, we conducted random repeated training and 5-fold spatial cross-validation. Table 7 reports the statistical results.

The repeated training runs demonstrate that the OTG model exhibits exceptional initialization robustness, maintaining stable near-ideal errors and avoiding training oscillations or overfitting. Furthermore, the spatial cross-validation experiments successfully

replicate real-world exploration conditions with uneven, sparse sampling. While accuracy drops slightly compared to the optimal training state, all indicators remain highly robust ($PCC \geq 0.98$), demonstrating that the OTG framework possesses reliable spatial generalization capabilities across complex field sampling geometries.

Table 7. Results of model stability and spatial generalization experiments.

Experiment Type	RMSE (km/s)	MAE (km/s)	MAPE (%)	PCC	R ²
Random Repeated Experiments	0.0226	0.0116	0.3132	0.9991	0.9982
Spatial Cross-Validation	0.1150	0.0760	1.9629	0.9801	0.9489

3.7.4. Comprehensive Ablation Summary

In summary, the internal optimization mechanics of the OTG framework rely on four interconnected pillars: First, the complementary nature of OK (global trend anchoring), TPS (local smooth continuity), and GCN (nonlinear feature extraction) drives a qualitative breakthrough in interpolation performance. Second, the attention-gating mechanism enables data-driven dynamic weight allocation, overcoming the limitations of fixed-weight hybrid models. Third, the multi-dimensional physical constraint loss standardizes velocity ranges and penalizes unphysical artifacts, ensuring geological interpretability. Finally, the framework demonstrates strong initialization robustness and spatial generalization, confirming its practical utility for high-precision field exploration engineering.

3.8. Computational Efficiency and Scalability Analysis

We further evaluated the computational efficiency of OTG in terms of training time, inference latency, and memory footprint, with quantitative results summarized in Table 8. OTG introduces moderate extra training overhead compared with lightweight data-driven baseline DeepKriging, which is a reasonable cost brought by the hybrid network architecture and embedded physics-informed constraints integrated into our framework.

Table 8. Computational efficiency comparison.

Model	RMSE (km/s)	Train Time (s)	Inference Time (s)	Inference/Sample (ms)	Peak RAM (MB)	Peak VRAM (MB)
OTG (Ours)	0.1216	69.89	0.1764	0.0803	1145.91	21.77
DeepKriging	0.1482	17.15	0.0518	0.0236	1149.18	19.12
GCN	0.5377	40.60	0.1639	0.0746	1148.36	22.61
OK	0.1530	-	7.9020	3.5967	1588.41	-
TPS	5.3332	-	1.1543	0.5254	1149.20	-
IDW	0.1714	-	0.5989	0.2726	1149.21	-
RBF-MQ	1.7098	-	0.7793	0.3547	1149.21	-

Nevertheless, OTG achieves state-of-the-art comprehensive performance across prediction accuracy and online inference speed. In terms of precision, OTG obtains the lowest RMSE of 0.1216 km/s among all competitors, reducing prediction error by 17.9% compared with DeepKriging and 20.5% compared with classical Ordinary Kriging (OK). This fully verifies that physical prior constraints can effectively promote the fitting consistency of spatial fields.

In terms of inference efficiency, OTG significantly outperforms all traditional interpolation algorithms, including OK, TPS, IDW, and RBF-MQ. Specifically, the inference latency per sample of OTG is only 2.2% of that of OK, achieving a two-order-of-magnitude speedup. This reveals that the computational burden of OTG is concentrated in one-time offline training, while its online deployment cost remains extremely low, making it highly suitable

for large-scale, high-density batch spatial interpolation once the model has converged. The efficiency and scalability experiments use the same model architecture but a different data split from the spatial cross-validation reported in Table 4; therefore, the RMSE values differ slightly.

We further conduct scalability experiments by gradually expanding training samples to validate the engineering practicability of OTG under enlarged data volume, with results presented in Table 9. As the training scale rises from 25% to 100%, the RMSE of OTG steadily declines from 0.1528 to 0.1216, achieving a total accuracy improvement of 20.4%. Such prominent error reduction demonstrates that the physics-aware framework of OTG can fully exploit the information contained in training data and gain more significant precision gains with growing dataset size than pure data-driven baselines.

Table 9. Scalability analysis of OTG across different training data scales.

Scale	Train Size	Val Size	Test Size	RMSE (km/s)	Train Time (s)	Inference Time (s)	Inference/Sample (ms)
25%	10,144	6855	2197	0.1528	52.37	0.1771	0.0806
50%	20,289	6855	2197	0.1374	58.92	0.1768	0.0805
75%	30,421	6855	2197	0.1289	64.71	0.1766	0.0804
100%	40,561	6855	2197	0.1216	69.89	0.1764	0.0803

The training time of OTG increases mildly in a sub-linear manner as the number of training samples expands, with controllable growth. This is because the computational bottleneck of OTG lies in forward and backward propagation of network weights, while data loading and preprocessing account for a small proportion of total training time. In addition, the model architecture and number of parameters remain fixed under different data scales, resulting in nearly unchanged inference latency.

Overall, OTG achieves a favorable balance between outstanding accuracy gain and stable computational overhead. It maintains efficient operation under moderate expansion of training samples and elevated grid density, possessing remarkable scalability to fit large-scale, high-density spatial interpolation tasks.

4. Discussion

Accurate reconstruction of 3D S-wave velocity fields under sparse sampling remains a long-standing bottleneck in seismic exploration. Although deep learning has been increasingly adopted to break through the rigid structural limitations of traditional geostatistical approaches, most existing deep learning workflows suffer from the notorious “black-box” defect and generate geologically implausible artifacts. The experimental results of this work verify that the proposed OTG framework effectively bridges purely data-driven fitting and physical rule constraints, establishing a highly reliable technical paradigm for 3D subsurface velocity modeling.

The competitive performance of the OTG model fundamentally stems from its complementary multi-module fusion architecture. Baseline comparisons and structural ablation experiments demonstrate that single-method schemes carry obvious inherent biases: ordinary kriging (OK) captures large-scale regional trends yet fails to depict nonlinear geological variations; thin-plate spline (TPS) guarantees local spatial smoothness but readily produces vertical structural distortions; pure graph convolutional network (GCN) models can extract intricate topological relationships, yet they tend to suffer from over-smoothing or extreme velocity deviations without auxiliary spatial priors. By synergistically integrating OK, TPS and GCN, the OTG framework builds a mutually reinforcing fusion mechanism. Distinct from prior hybrid models such as KCN [26] and DeepKriging [27], which rely on fixed static

weights or naive feature concatenation, we design an adaptive gating unit that dynamically adjusts the contribution weight of each component according to local sampling density and geological structural complexity. This dynamic self-adaptation accounts for the consistently superior performance of OTG across highly heterogeneous subsurface media.

Moreover, this study underscores the necessity of embedding physical regularization into deep learning-based geophysical inversion pipelines. Loss function ablation tests explicitly reveal that optimization merely based on mean squared error (MSE) will raise the frequency of geologically unreasonable abrupt velocity anomalies by 32%. By imposing multi-dimensional physical constraints—including hard bounds of seismic wave velocity and gradient consistency regularization—the model is restricted to search optimal solutions within geophysically reasonable parameter ranges. Such a physics-informed learning strategy effectively suppresses spurious pseudo-structures ubiquitous in standalone neural networks, ensuring the interpolated 3D velocity volumes can be reliably deployed for downstream seismic workflows, including Pre-stack Depth Migration (PSDM) and site seismic response evaluation.

Notwithstanding the robust fitting capacity and strong spatial generalization exhibited by the OTG framework, three key limitations deserve objective discussion and further targeted research. First, joint integration of OK, TPS and deep graph networks inevitably brings extra computational overhead. Specifically, iterative computation of spatial covariance matrices and radial basis functions introduces higher inference latency relative to lightweight end-to-end MLP architectures. Optimizing parallel computation efficiency for the multi-feature fusion module is a core priority for ultra-large-scale, high-resolution 3D grid reconstruction tasks. Second, the current depth-dependent regularization term L_{depth} adopts a universal velocity-increases-with-depth assumption, which fits the sedimentary basins featured in our test area but may suppress authentic geological anomalies in zones with complex velocity inversion structures, such as deep salt domes and overthrust nappes. Third, all validation experiments of the proposed method are conducted based on geological datasets from Southwest China; even though the obtained results are convincing, the generalization capability of the OTG framework toward diverse tectonic and geological settings still requires systematic verification.

For future research, we will first develop generalized topological constraint algorithms to replace the rigid monotonic depth-velocity prior and improve model adaptability to extreme subsurface geomorphology. Second, extensive cross-domain validation will be carried out on diverse seismic datasets and extreme geological environments (e.g., zones with severe velocity inversion or massive salt bodies) to comprehensively characterize the applicable boundary conditions of the OTG framework. Third, we plan to expand the OTG architecture to incorporate multi-modal geophysical priors such as gravity and magnetic anomaly data, further enriching the scope of physical regularization and advancing the capability of AI-driven subsurface structural interpretation.

5. Conclusions

The precise construction of three-dimensional S-wave velocity fields occupies an indispensable position in seismic exploration, playing an irreplaceable role in seismic imaging, structural interpretation, and underground resource evaluation. We propose an OTG hybrid interpolation framework that innovatively integrates Ordinary Kriging, Thin Plate Spline, and Graph Convolutional Network. An adaptive gating fusion mechanism is designed to dynamically adjust the weights of the three modules based on local data density and structural complexity, and a multi-dimensional physical constraint loss function is constructed to ensure the geological rationality of interpolation results.

Through a series of detailed and in-depth experimental verifications, we extensively compare and evaluate the performance of 10 representative interpolation methods under unified experimental conditions. The experimental results consistently show that the proposed OTG model achieves state-of-the-art performance (RMSE = 0.1148 km/s, MAPE = 1.9614%, PCC = 0.9801) on the dataset from 105 seismic stations in Southwest China. It reduces the RMSE by 42.8% compared with the optimal traditional method (OK) and by 8.2% compared with the state-of-the-art hybrid method (DeepKriging). The ablation experiments further confirm that the adaptive gating fusion mechanism and multi-dimensional physical constraint loss function are the key factors contributing to the superior performance of the OTG model.

The OTG model effectively realizes the complementary advantages of traditional geoscientific methods and deep learning techniques, solving the core challenge of balancing modeling accuracy and geological rationality in 3D S-wave velocity modeling. It maintains high accuracy and geological plausibility under sparse and non-uniform sampling conditions, providing a reliable engineering solution for high-precision 3D S-wave velocity structure interpolation in seismic exploration.

How to further improve the generalization ability of the model under extremely complex geological conditions has always been our research goal. In future work, we will expand the dataset coverage to include regions with intense tectonic activity and complex lithological distributions. We will also introduce more geological prior knowledge, such as fault distribution and lithology information, to further enhance the geological interpretation capability of the model. Additionally, we will optimize the model computational efficiency to meet the application requirements of large-scale exploration areas, and explore the combination of the OTG model with Full Waveform Inversion to improve the quality of velocity inversion.

In general, the OTG model not only enhances the accuracy of three-dimensional S-wave velocity field interpolation but also provides a new technical approach for the deep integration of traditional geoscientific methods and deep learning techniques, demonstrating the tremendous potential and broad prospects of hybrid physics-informed frameworks in the field of earth sciences.

Author Contributions: Conceptualization, G.Z. and Y.Y.; methodology, G.Z. and Y.Y.; software, Y.Y.; validation, Y.G., J.S. and E.C.; formal analysis, Y.Y. and S.W.; investigation, G.Z.; resources, Z.G.; data curation, Y.G., J.S. and E.C.; writing—original draft preparation, G.Z. and Y.Y.; writing—review and editing, Z.G. and X.Y.; visualization, Y.Y.; supervision, Z.G.; project administration, Z.G.; funding acquisition, Z.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China, grant number 2023YFC3010700. The APC was funded by the National Key R&D Program of China under Grant No. 2023YFC3010700.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data supporting the reported results are available upon reasonable request from the corresponding author.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (based on the GPT-4 architecture, OpenAI, San Francisco, CA, USA) exclusively for language polishing and manuscript structure optimization. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

APC	Article Processing Charge
Coord-MLP	Coordinate Multi-Layer Perceptron
CPU	Central Processing Unit
CUDA	Compute Unified Device Architecture
FFN-MLP	Feed-Forward Neural Network-Multi-Layer Perceptron
FWI	Full Waveform Inversion
GAT	Graph Attention Network
GCN	Graph Convolutional Network
GraphSAGE	Graph Sample and Aggregation
GT	Ground Truth
IDW	Inverse Distance Weighting
KCN	Kriging Convolutional Network
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
OK	Ordinary Kriging
OTG	Ordinary Kriging-Thin Plate Spline-Graph Convolutional Network
PCC	Pearson Correlation Coefficient
PINNs	Physics-Informed Neural Networks
PSDM	Pre-stack Depth Migration
R ²	Coefficient of Determination
RAM	Random Access Memory
RBF-MQ	Radial Basis Function-Multiquadric
RMSE	Root Mean Squared Error
RTM	Reverse Time Migration
TPS	Thin Plate Spline

References

1. Etgen, J.; Gray, S.; Zhang, Y. An overview of depth imaging in exploration geophysics. *Geophysics* **2009**, *74*, WCA5–WCA34. [\[CrossRef\]](#)
2. Xia, J.; Miller, R.D.; Park, C.B. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *Geophysics* **1999**, *64*, 691–700. [\[CrossRef\]](#)
3. Jones, I.F. *An Introduction to Velocity Model Building*; EAGE Publications: Houten, The Netherlands, 2010. [\[CrossRef\]](#)
4. Li, S.; Liu, B.; Ren, Y.; Chen, Y.; Yang, S.; Wang, Y.; Jiang, P. Deep-learning inversion of seismic data. *IEEE Trans. Geosci. Remote Sens.* **2020**, *85*, 2135–2149. [\[CrossRef\]](#)
5. Adler, A.; Araya-Polo, M.; Poggio, T. Deep learning for seismic inverse problems: Toward the automation of structural imaging. *IEEE Signal Process. Mag.* **2021**, *38*, 89–119. [\[CrossRef\]](#)
6. Jia, Y.; Yu, S.; Ma, J. Intelligent interpolation by Monte Carlo machine learning. *Geophysics* **2018**, *83*, V83–V97. [\[CrossRef\]](#)
7. Cebrykow, P. Evaluating the usefulness of interpolation methods in constructing isopleth maps. *Misc. Geogr.* **2025**, *29*, 116–123. [\[CrossRef\]](#)
8. Shepard, D. A two-dimensional interpolation function for irregularly-spaced data. In Proceedings of the 1968 23rd ACM National Conference, New York, NY, USA, 27–29 August 1968; pp. 517–524. [\[CrossRef\]](#)
9. Goovaerts, P. *Geostatistics for Natural Resources Evaluation*; Oxford University Press: New York, NY, USA, 1997.
10. Bookstein, F.L. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 567–585. [\[CrossRef\]](#)
11. Oliver, M.A.; Webster, R. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *Catena* **2014**, *113*, 56–69. [\[CrossRef\]](#)
12. Virieux, J.; Operto, S. An overview of full-waveform inversion in exploration geophysics. *Geophysics* **2009**, *74*, WCC1–WCC26. [\[CrossRef\]](#)

13. Coppelino, L.; Desiderio, L. Space-time energetic Galerkin BEM for the numerical solution of 3D elastodynamic problems: Overcoming challenges of the strongly singular integral operator. *Comput. Mech.* **2025**, *76*, 1689–1714. [\[CrossRef\]](#)
14. Kipf, T.N.; Welling, M. Semi-supervised classification with graph convolutional networks. In Proceedings of the 5th International Conference on Learning Representations (ICLR), Toulon, France, 24–26 April 2017.
15. Hamilton, W.L.; Ying, R.; Leskovec, J. Inductive representation learning on large graphs. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, USA, 4–9 December 2017; pp. 1025–1035.
16. Tancik, M.; Srinivasan, P.; Mildenhall, B.; Fridovich-Keil, S.; Raghavan, N.; Singhal, U.; Ramamoorthi, R.; Barron, J.; Ng, R. Fourier features let coordinate-mlp learn high-frequency functions in low dimensional domains. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 7537–7547.
17. Scarselli, F.; Gori, M.; Tsoi, A.C.; Hagenbuchner, M.; Monfardini, G. The graph neural network model. *IEEE Trans. Neural Netw.* **2009**, *20*, 61–80. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Yu, P.S. A comprehensive survey on graph neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 4–24. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Fang, Z.; Zhang, T.; Cai, W.; Shi, Y.; Shah, S.Y.A.; Kablan, O.A.B.K.; Zhang, B. Three-dimensional geological modeling based on dual-task stratigraphy-aware attention networks (Geo-SAN v1.0). *EGUsphere* 2026, preprint. [\[CrossRef\]](#)
20. 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. [\[CrossRef\]](#)
21. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [\[CrossRef\]](#)
22. Moseley, B.; Markham, A.; Nissen-Meyer, T. Solving the wave equation with physics-informed deep learning. *arXiv* **2020**, arXiv:2006.11894. [\[CrossRef\]](#)
23. Yang, F.; Ma, J. Deep-learning inversion: A next-generation seismic velocity model building method. *Geophysics* **2019**, *84*, R583–R599. [\[CrossRef\]](#)
24. Zerafa, C.; Galea, P.; Sebu, C. Synergizing deep learning and full-waveform inversion: Bridging data-driven and theory-guided approaches for enhanced seismic imaging. *arXiv* **2025**, arXiv:2502.17585. [\[CrossRef\]](#)
25. Feng, G.; Liu, W.Q.; Yang, Z.; Yang, W. Shear wave velocity prediction based on 1DCNN-BiLSTM network with attention mechanism. *Front. Earth Sci.* **2024**, *12*, 1376344. [\[CrossRef\]](#)
26. Appleby, G.; Liu, L.; Liu, L.-P. Kriging convolutional networks. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 3187–3194. [\[CrossRef\]](#)
27. Chen, W.; Li, Y.; Reich, B.J.; Sun, Y. DeepKriging: Spatially Dependent Deep Neural Networks for Spatial Prediction. *Stat. Sin.* **2024**, *34*, 291–311. [\[CrossRef\]](#)
28. Hu, J.; Sun, W.; Liu, C.; Tang, Q.; Fu, L.-Y. Seismological reference earth model in South China (SREM-SC): Crust and uppermost mantle. *Front. Earth Sci.* **2023**, *10*, 1080307. [\[CrossRef\]](#)
29. Li, J.; Heap, A.D. Spatial interpolation methods in environmental science: A review. *Environ. Model. Softw.* **2014**, *53*, 173–189. [\[CrossRef\]](#)
30. Araya-Polo, M.; Jennings, J.; Adler, A.; Dahlke, T. Deep-learning tomography. *Lead. Edge* **2018**, *37*, 58–66. [\[CrossRef\]](#)
31. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; Bengio, Y. Graph attention networks. In Proceedings of the 6th International Conference on Learning Representations (ICLR), Vancouver, BC, Canada, 30 April–3 May 2018.
32. Wu, X.; Liang, L.; Shi, Y.; Fomel, S. FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. *Geophysics* **2019**, *84*, IM35–IM45. [\[CrossRef\]](#)
33. Karimpouli, S.; Tahmasebi, P. Physics informed machine learning: Seismic wave equation. *Geosci. Front.* **2020**, *11*, 1993–2001. [\[CrossRef\]](#)
34. Li, R.; Wang, S.; Zhu, F.; Huang, J. Adaptive graph convolutional neural networks. *Proc. AAAI Conf. Artif. Intell.* **2018**, *32*, 3546–3553. [\[CrossRef\]](#)
35. Wu, X.; Meng, F.; Liu, P.; Zhou, J.; Liu, D.; Xie, K.; Zhu, Q.; Hu, J.; Sun, H.; Xing, F. Contribution of the northeast cold vortex index and multiscale synergistic indices to extreme precipitation over Northeast China. *Earth Space Sci.* **2021**, *8*, e2020EA001435. [\[CrossRef\]](#)
36. 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. [\[CrossRef\]](#)
37. Tartakovsky, A.M.; Marrero, C.O.; Perdikaris, P.; Tartakovsky, G.D.; Barajas-Solano, D. Physics-informed deep neural networks for learning parameters and constitutive relationships in subsurface flow problems. *Water Resour. Res.* **2020**, *56*, e2019WR026731. [\[CrossRef\]](#)
38. Christensen, N.I.; Mooney, W.D. Seismic velocity structure and composition of the continental crust: A global view. *J. Geophys. Res.* **1995**, *100*, 9761–9788. [\[CrossRef\]](#)

39. Li, Z.; Kovachki, N.; Azizzadenesheli, K.; Liu, B.; Bhattacharya, K.; Stuart, A.; Anandkumar, A. Fourier neural operator for parametric partial differential equations. In Proceedings of the 9th International Conference on Learning Representations (ICLR), Vienna, Austria, 3–7 May 2021.
40. Xiong, W.; Ji, X.; Ma, Y.; Wang, Y.; AlBinHassan, N.M.; Ali, M.N.; Luo, Y. Seismic fault detection with convolutional neural network. *Geophysics* **2018**, *83*, O97–O103. [[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.